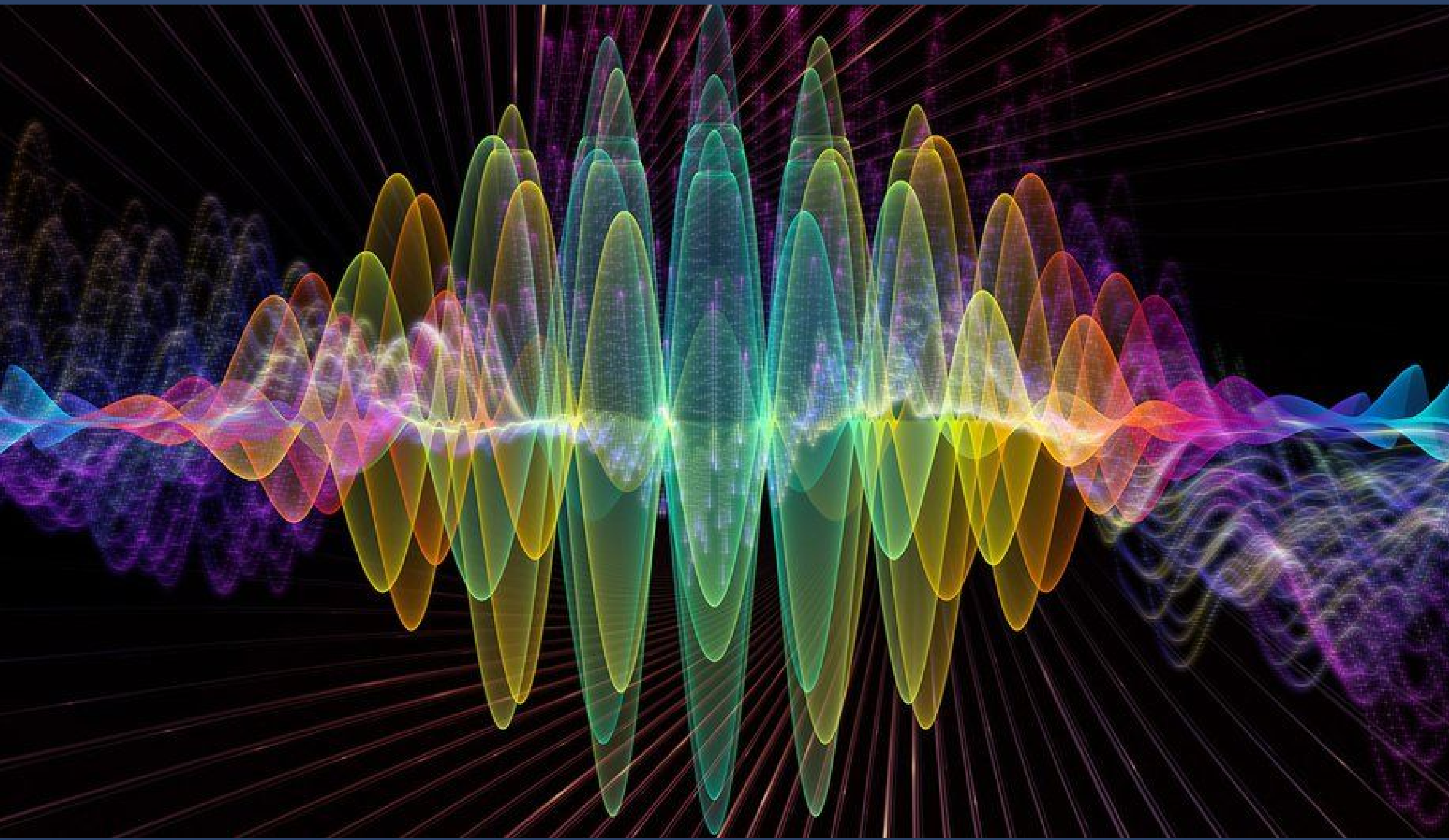
TEXT BOOK FOR

SR.SECONDARY COURSE

PHYSICS



BOARD OF SCHOOL EDUCATION HUBLI, KARNATAKA

ಬೋರ್ಡ್ ಆಫ್ ಸ್ಕೂಲ್ ಎಜುಕೇಶನ್ ಹುಬ್ಬಳ್ಳಿ, ಕರ್ನಾಟಕ

Upper Ground Floor, IT Park, Deshpande Nagar

0836-4061132

www.bseh.ac.in

contact@bseh.ac.in

CONTENT-A

NO. **CHAPTER** **PAGE NO.**

MODULE-1
Motion, Force and Energy

MODULE-2
Mechanics of Solids and Fluids

MODULE-3
Thermal Physics

MODULE-4
Oscillations and Waves

CONTENT-B

312en01

1

UNITS, DIMENSIONS AND VECTORS

In science, particularly in physics, we try to make measurements as precisely as possible. Several times in the history of science, precise measurements have led to new discoveries or important developments. Obviously, every measurement must be expressed in some units. For example, if you measure the length of your room, it is expressed in suitable units. Similarly, if you measure the interval between two events, it is expressed in some other units. The unit of a physical quantity is derived, by expressing it in base units fixed by international agreement. The idea of base units leads us to the concept of **dimensions**, which as we shall see, has important applications in physics.

You will learn that physical quantities can generally be divided in two groups: **scalars** and **vectors**. Scalars have only magnitudes while vectors have both magnitude and direction. The mathematical operations with vectors are somewhat different from those which you have learnt so far and which apply to scalars. The concepts of vectors and scalars help us in understanding physics of different natural phenomena. You will experience it in this course.

OBJECTIVES

After studying this lesson, you should be able to:

- *describe the scope of physics, nature of its laws and applications of the principles of physics in our life;*
- *identify the number of significant figures in measurements and give their importance;*
- *distinguish between the fundamental and derived quantities and give their SI units;*
- *write the dimensions of various physical quantities;*

- *apply dimensional analysis to check the correctness of an equation and determine the dimensional nature of 'unknown' quantities;*
- *differentiate between scalar and vector quantities and give examples of each;*
- *add and subtract two vectors and resolve a vector into its components; and*
- *calculate the product of two vectors.*

1.1 PHYSICAL WORLD AND MEASUREMENTS

1.1.1 Physics: Scope and Excitement

The scope of Physics is very wide. It covers a vast variety of natural phenomena. It includes the study of mechanics; heat and thermodynamics; optics; waves and oscillations; electricity and magnetism; atomic and nuclear physics; electronics and computers etc. Of late, need for solutions of quite a few problems has led to the development of subjects like biophysics, chemical physics, astrophysics, soil physics, geophysics etc., thus widening the scope of physics further. In physics, we study large objects such as stars, planets etc.; and tiny objects like elementary particles; large distances such as 10^{26} m (size of the universe) as well as small distances such as 10^{-14} m (size of the nucleus of an atom); large masses such as 10^{55} kg (mass of universe) as well as tiny masses of 10^{-30} kg (mass of an electron).

Physics is perhaps the most basic of all sciences. All developments in engineering or technology are nothing but the applications of Physics.

The study of Physics has led to many exciting discoveries, inventions and their applications for example:

(i) A falling apple led to the understanding of gravitation.

(ii) Production of electrical energy by hydro, thermal or nuclear power plants (imagine the life and the world without electricity).

(iii) Receiving messages and visuals from anywhere on the globe by telephone and television,

(iv) Landing on the moon and the study of planets like Mars and other astronomical objects with robotic control from the ground,

(v) The study of the outer space with the help of artificial satellites, and satellite mounted telescopes,

(vi) Lasers and its numerous applications

(vii) High speed computers, and many more.

1.1.2 Nature of Physical Laws

Physicists explore the universe. Their investigations based on scientific process range from sub-atomic particles to big stars.

Physical laws are typical conclusions based on repeated scientific experiments and observations over many years and which have been accepted universally within the scientific community. Physical laws are:

- True at least within their regime of validity.
- Universal. They appear to apply everywhere in the universe.
- Simple. They are typically expressed in terms of a single mathematical equation.
- Absolute. Nothing in the universe appears to affect them.
- Stable. Unchanged since discovered (although they may have some approximations and/or exceptions).
- Omnipotent. Everything in the universe apparently must comply with them.

1.1.3 Physics, Technology and Society

Technology is the application of the principles of physics for the manufacture of machines, gadgets etc. and improvements in them, which leads to better quality of our physical life. For example:

(i) Different types of Engines (steam, petrol, diesel etc.) are based on the laws of thermodynamics.

(ii) Means of communication e.g. radio, telephone, television etc. are based on the propagation of electromagnetic waves.

(iii) Generation of electricity is based on the principle of electromagnetic induction.

(iv) Nuclear reactors – are based on the phenomenon of controlled nuclear fission.

(v) Jet aeroplanes and rockets are based on Newton's second and third laws of motion.

(vi) X–rays, ultraviolet rays and infrared rays are used in medical science for diagnostic and healing purposes.

(vii) Mobile phones, calculators and computers are based on the principles of electronics.

(viii) Lasers are based on the phenomenon of population inversion, and so on.

1.1.4 Need of Measurement

Every new discovery brings in revolutionary change in the structure of society and life of its people. Can you illustrate this fact with the help of some examples?

Physics, as we know, is a branch of science which deals with nature and natural phenomena. For complete and proper study of any phenomenon, measurement of quantities involved is essential. For example, to study the motion of a particle, measurement of its displacement, velocity, and acceleration at any instant has to be made accurately. For this, measurement of time and distance has to be done. Similarly, measurement of volume, pressure and temperature is necessary to study the state of a gas fully. Measurement of mass, volume and temperature of a liquid has to be made to study the effect of heat on it. Thus, we find that measurement of quanties, such as, distance, time, temperature, mass, force etc. has to be made to study every natural phenomena. This explains the need for measurement.

1.2 UNIT OF MEASUREMENT

The laws of physics are expressed in terms of physical quantities such as distance, speed, time, force, volume, electric current, etc. For measurement, each physical quantity is assigned a unit. For example, time could be measured in minutes, hours or days. But for the purpose of useful communication among different people, this unit must be compared with a standard unit acceptable to all. As another example, when we say that the distance between Mumbai and Kolkata is nearly 2000 kilometres, we have for comparison a basic unit in mind, called a kilometre. Some other units that you may be familiar with are a kilogram for mass and a second for time. It is essential that all agree on the standard units, so that when we say 100 kilometres, or 10 kilograms, or 10 hours, others understand what we mean by them. In science, international agreement on the basic units is absolutely essential; otherwise scientists in one part of the world would not understand the results of an investigation conducted in another part.

Suppose you undertake an investigation on the solubility of a chemical in water. You weigh the chemical in tolas and measure the volume of water in cupfuls. You communicate the results of your investigation to a scientist friend in Japan. Would your friend understand your results?

It is very unlikely that your friend would understand your results because he/she may not be familiar with tola and the cup used in your measerments, as they are not standard units.

Do you now realize the need for agreed standards and units?

Remember that in science, the results of an investigation are considered established only if they can be reproduced by investigations conducted elsewhere under identical conditions.

Measurements in Indian Traditions

Practices of systematic measurement are very old in India. The following quote from Manusmriti amply illustrates this point :

"The king should examine the weights and balance every six months to ensure true measurements and to mark them with royal stamp." – Manusmriti, 8th Chapter, sloka–403.

In **Harappan Era,** signs of systematic use of measurement are found in abundance : the equally wide roads, bricks having dimensions in the ratio 4 : 2 : 1, Ivory scale in Lothal with smallest division of 1.70 mm, Hexahedral weights of 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200 and 500 units (1 unit = 20 g)

In Mauriyan Period, the following units of length were prevalent

8 Parmanu	= 1 Rajahkan
8 Rajahkan	= 1 Liksha
8 Liksha	= 1 Yookamadhya
8 Yookamadhya	= 1 Yavamadhya
8 Yavamadhya	= 1 Angul
8 Angul	= 1 Dhanurmushthi

In **Mughal Period**, Shershah and Akbar tried to re-establish uniformity of weights and measures. Akbar introduced gaz of 41 digits for measuring length. For measuring area of land, bigha was the unit. 1 bigha was 60 gaz × 60 gaz.

Units of mass and volume were also well obtained in Ayurveda.

1.2.1 The SI Units

With the need of agreed units in mind, the 14th General Conference on Weights and Measures held in 1971, adopted seven **base** or **fundamental units**. These units form the SI system. The name SI is abbreviation for **Système International d'Unités** for the International System of units. The system is popularly known as the metric system. The SI units along with their symbols are given in Table 1.1.

Table 1.1 : Base SI Units

Quantity	Unit	Symbol
Length	metre	m
Mass	kilogram	kg
Time	second	s
Electric Current	ampere	A
Temperature	kelvin	K
Luminous Intensity	candela	cd
Amount of Substance	mole	mol

The mile, yard and foot as units of length are still used for some purposes in India as well in some other countries. **However, in scientific work we always use SI units.**

As may be noted, the SI system is a metric system. It is quite easy to handle because the smaller and larger units of the base units are always submultiples or multiples of ten. These multiples or submultiples are given special names. These are listed in Table 1.2.

Table 1.2 : Prefixes for powers of ten

Power of ten	Prefix	Symbol	Example	
10^{-18}	atto	a	attometre	(am)
10^{-15}	femto	f	femtometre	(fm)
10^{-12}	pico	p	picofarad	(pF)
10^{-9}	nano	n	nanometre	(nm)
10^{-6}	micro	μ	micron	(μm)
10^{-3}	milli	m	milligram	(mg)
10^{-2}	centi	c	centimetre	(cm)
10^{-1}	deci	d	decimetre	(dm)
10^{1}	deca	da	decagram	(dag)
10^{2}	hecto	h	hectometre	(hm)
10^{3}	kilo	k	kilogram	(kg)
10^{6}	mega	M	megawatt	(MW)
10^{9}	giga	G	gigahertz	(GHz)
10^{12}	tera	T	terahertz	(THz)
10^{15}	peta	P	peta kilogram	(Pkg)
10^{18}	exa	E	exa kilogram	(Ekg)

Just to get an idea of the masses and sizes of various objects in the universe, see Table 1.3 and 1.4. Similarly, Table 1.5 gives you an idea of the time scales involved in the universe.

Table 1.3 : Order of magnitude of some masses

Mass	kg
Electron	10^{-30}
Proton	10^{-27}
Amino acid	10^{-25}
Hemoglobin	10^{-22}
Flu virus	10^{-19}
Giant amoeba	10^{-8}
Raindrop	10^{-6}
Ant	10^{-2}
Human being	10^{2}
Saturn 5 rocket	10^{6}
Pyramid	10^{10}
Earth	10^{24}
Sun	10^{30}
Milky Way galaxy	10^{41}
Universe	10^{55}

Table 1.4 : Order of magnitude of some lengths

Length	m
Radius of proton	10^{-15}
Radius of atom	10^{-10}
Radius of virus	10^{-7}
Radius of giant amoeba	10^{-4}
Radius of walnut	10^{-2}
Height of human being	10^{0}
Height of highest mountain	10^{4}
Radius of earth	10^{7}
Radius of sun	10^{9}
Earth-sun distance	10^{11}
Radius of solar system	10^{13}
Distance to nearest star	10^{16}
Radius of Milky Way galaxy	10^{21}
Radius of visible universe	10^{26}

Table 1.5 : Order of magnitude of some time intervals

Interval	s
Time for light to cross nucleus	10^{-23}
Period of visible light	10^{-15}
Period of microwaves	10^{-10}
Half-life of muon	10^{-6}
Period of highest audible sound	10^{-4}
Period of human heartbeat	10^{0}
Half-life of free neutron	10^{3}
Period of the Earth's rotation (day)	10^{5}
Period of revolution of the Earth (year)	10^{7}
Lifetime of human beings	10^{9}
Half-life of plutonium-239	10^{12}
Lifetime of a mountain range	10^{15}
Age of the Earth	10^{17}
Age of the universe	10^{18}

1.2.2 Standards of Mass, Length and Time

Once we have chosen to use the SI system of units, we must decide on the set of standards against which these units will be measured. We define here standards of mass, length and time.

(i) **Mass :** The SI unit of mass is **kilogram**. The standard kilogram was established in 1887. **It is the mass of a particular cylinder made of platinum-iridium alloy**, which is an unusually stable alloy. The standard is kept in the International Bureau of Weights and Measures in Paris, France. The prototype kilograms made of the same alloy have been distributed to all countries the world over. For India, the national prototype is the kilogram no. 57. This is maintained by the National Physical Laboratory, New Delhi (Fig. 1.1).

Fig. 1.1 : Prototype of kilogram

(ii) **Length :** The SI unit of length is metre. It is defined in terms of a natural phenomenon: **One metre is defined as the distance travelled by light in vacuum in a time interval of 1/299792458 second.**

This definition of metre is based on the adoption of the speed of light in vacuum as 299792458 ms^{-1}

(iii) Time : One second is defined as **the time required for a Cesium - 133 (^{133}Cs) atom to undergo 9192631770 vibrations between two hyperfine levels of its ground state.**

This definition of a second has helped in the development of a device called atomic clock (Fig. 1.2). The cesium clock maintained by the National Physical Laboratory (NPL) in India has an uncertainty of $\pm 1 \times 10^{-12}$ s, which corresponds to an accuracy of one picosecond in a time interval of one second.

Cesium Atomic Clock
(S60,000)
Cesium Beam Tube

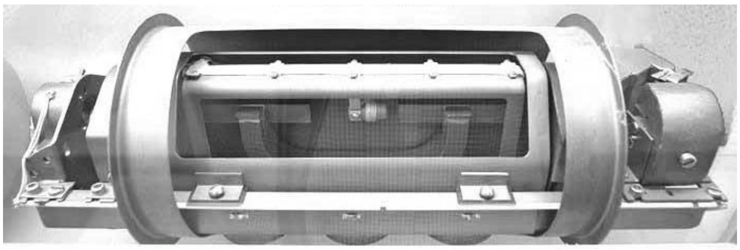


Fig. 1.2 : Atomic Clock

As of now, clock with an uncertainty of 5 parts in 10^{15} have been developed. This means that if this clock runs for 10^{15} seconds, it will gain or lose less than 5 seconds. You can convert 10^{15} s to years and get the astonishing result that this clock could run for 6 million years and lose or gain less than a second. This is not all. Researches are being conducted today to improve upon this accuracy constantly. Ultimately, we expect to have a clock which would run for 10^{18} second before it could gain or lose a second. To give you an idea of this technological achievement, if this clock were started at the time of the birth of the universe, an event called the Big Bang, it would have lost or gained only two seconds till now.

Role of Precise Measurements in New Discoveries

A classical example of the fact that precise measurements may lead to new discoveries are the experiments conducted by Lord Rayleigh to determine density of nitrogen.

In one experiment, he passed the air bubbles through liquid ammonia over red hot copper contained in a tube and measured the density of pure nitrogen so obtained. In another experiment, he passed air directly over red hot copper and measured the density of pure nitrogen. The density of nitrogen obtained in second experiment was found to be 0.1% higher than that obtained in the first case. The experiment suggested that air has some other gas heavier than nitrogen present in it. Later he discovered this gas – Argon, and got Nobel Prize for this discovery.

Another example is the failed experiment of Michelson and Morley. Using Michelson's interferometer, they were expecting a shift of 0.4 fringe width in the interference pattern obtained by the superposition of light waves travelling in the direction of motion of the earth and those travelling in a transverse direction. The instrument was hundred times more sensitive to detect the shift than the expected shift. Thus they were expecting to measure the speed of earth with respect to ether and conclusively prove that ether existed. But when they detected no shift, the world of science entered into long discussions to explain the negative results. This led to the concepts of length contraction, time dilation etc and ultimately to the theory of relativity.

Several discoveries in nuclear physics became possible due to the new technique of spectroscopy which enabled detection, with precision, of the traces of new atoms formed in a reaction.

1.2.3 Significant Figures

When a student measures the length of a line as 6.8 cm, the digit 6 is certain, while 8 is uncertain as a little less or more than 0.8 cm is reported as 0.8 cm. Normally those digits in measurement that are known with certainly plus the first uncertain digit, are called significant figures.

Thus, there are two significant figures in 1.4 cm. The number of significant figures in any quantity depends upon the accuracy of the measuring instrument. More the number of significant number of figures, less is the percentage of error in the measurement of the quantity. If there are lesser number of significant figures (in a measurement) more will be the percentage error in the measurement.

The number of significant figures of a quantity may be found by the following rules:

(i) All non-zero digits are significant. For example, 315.58 has five significant figures.

(ii) All zeros between two non-zero digits are significant. For example, 5300405.003 has ten significant figures.

(iii) All zeros which are to the right of a decimal point and also to the right of a non-zero digit are significant. For example, 50.00 has four significant figures, and so has .04050. It should be noted that in .04050, the first zero to the right of the decimal is not significant but, the last zero is significant.

(iv) All zeros to the right of a decimal point and to the left of a non-zero digit in a decimal fraction are not significant. For example, .00043 has only two significant figures but 2.00023 has 6 significant figures. It is also to be noted that zero conventionally placed to the left of a decimal point is not significant.

(v) All zero to the right of last of non-zero digit are significant, if they come from some measurement. For example, if the distance between two objects is 4050 m (measured to the nearest metre), then 4050 m contains 4 significant figures.

(vi) The number of significant figures does not vary with the change in unit. For example, if the length of an object is 348.6 cm, it has 4 significant figures. If the length is expressed in metre, then it is equal to 3.486 m. It still has 4 significant figures.

(vii) In a whole number all zeros to the right of the last non zero number are not significant, for example 5000 has only one significant figure.

Importance of significant figures in measurement.

As stated earlier, the accuracy of the measurement determines the number of significant figures in the quantity. Suppose the diameter of a coin is 2 cm. If a student measures the diameter with a metre scale which can read up to .1 cm only (i.e. cannot read less than 0.1 cm) the student will report the diameter to be 2.0 cm i.e. upto 2 significant figures only. If the diameter is measured by an instrument which can read upto .01 cm only (or which cannot measure less than .01cm), viz a Vernier Callipers, he will report the diameter as 2.00 cm i.e. upto 3 significant figure. Similarly if the measurement is made by an instrument like a screw gauge which can measure upto .001 cm only (i.e. cannot measure less than .001 cm), the diameter will be recorded as 2.000 cm i.e. upto 4 significant figures. Thus any measurement should be recorded keeping in view the accuracy of the measuring instrument.

Importance of significant figures in expressing the result of calculations

Suppose a student measures the side of a cube with the help of a metre scale which comes to be 3.2 cm. He calculates the volume of this cube mathematically and reports it to be $(3.2 \times 3.2 \times 3.2)$ cubic centimetre or 32.768 cm^3. The reported result is mathematically correct but is not correct in scientific measurement. The correct volume should be recorded as 33 cm^3. This is because there are only two significant figures in the length of the side of the cube, hence the volume should also have two significant figures, whereas there are 5 significant figures in 32.768 which is not correct.

Significant figures in addition, subtraction, multiplication and division

(i) Addition and subtraction – Suppose we have to add three quantities, 2.7 m, 3.68 m and 0.486 m. In these quantities, the first measurement is known upto one decimal place only, hence the sum of these numbers will be definite upto one decimal place only. Therefore, the correct sum of these numbers should not be written as 6.848 m but 6.8 m.

Similarly, to find the sum of quantities like 2.65×10^3 cm and 2.63×10^2 cm, all quantities should be converted to the same power of 10. These quantities will then be, 2.65×10^3 cm and $.263 \times 10^3$ cm. Since, the first number is known upto 2 decimal places, their sum will also be upto 2 decimal places. Hence $2.65 \times 10^3\text{cm} + .263 \times 10^3 \text{ cm} = 2.91 \times 10^3$ cm.

The same is done with subtraction. For example the result of subtracting 2.38 cm from 4.6 cm will be 2.2 cm, not 2.22 cm.

(ii) Multiplication and division – Suppose the length of a plate is measured as 3.003 m and its width as 2.26 m. According to Mathematical Calculation, the area of the plate will be 6.78678 m^2. But, it is not correct in scientific measurement. There are six significant figures in this result. But, the least number of significant figures (in the width) are only 3. Hence, the multiplication should also be writen upto 3 significant figures. Therefore, the correct area would be 6.79 m^2.

The same method is applied for division also. For example, dividing 248.57 by 56.9 gives 4.3685413. But, the result should be recorded upto 3 significant figures only as the least number of significant figures in the given numbers is only 3. Hence, the result will be 4.37.

Similarly, if a body travels a distance of 1452 m in 142 seconds, its speed according to mathematical calculations will be $\frac{1452}{142}$ m per second or 10.225352 m s^{-1}, but in scientific measurements it should be 10.2 m s^{-1}, as there are only 3 significant figures in the number for time.

(iii) Value of constants used in Calculation

If the radius (r) of a circle is 3.35 cm, to calculate its area (πr^2) the value of π should be taken upto two places of decimal (i.e $\pi = 3.14$, not 3.1416). Hence, the area of this circle $\pi r^2 = (3.14 \times 3.35 \times 3.35)\text{ cm}^2 = 35.2 \text{ cm}^2$, not 35.23865 cm^2.

(iv) If a measured quantity is multiplied by a constant, all the digits in the product are significant that are obtained by multiplication. For example, if the mass of a ball is 32.59 g the mass of 10 similar balls will be $32.59 \times 10 = 325.90$ g. Note that there are five significant figures in the number.

1.2.4. Derived Units

We have so far defined three basic units for the measurement of mass, length and time. For many quantities, we need units which we get by combining the basic units. These units are called derived units. For example, combination of the units of length and time gives us the derived unit of speed or velocity, m s^{-1}. Another example is the interaction of the unit of length with itself. We get derived units of area and volume as m^2 and m^3, respectively.

Now are would like you to list all the physical quantities that you are familiar with and the units in which they are expressed.

Some derived units have been given special names. Examples of most common of such units are given in Table 1.6.

Table 1.6 : Examples of derived units with special names

Quantity	Name	Symbol	Unit Symbol
Force	newton	N	kg m s^{-2}
Pressure	pascal	Pa	N m^{-2}
Energy/work	joule	J	N m
Power	watt	W	J s^{-1}

One of the advantages of the SI system of units is that they form a coherent set in the sense that the product or division of the SI units gives a unit which is also the SI unit of some other derived quantity. For example, product of the SI units of force and length gives directly the SI unit of work, namely, newton-metre (Nm) which has been given a special name joule. **Some care should be exercised in the order in which the units are written**. For example, Nm should be written in this order. If by mistake we write it as mN, it becomes millinewton, which is something entirely different.

Remember that in physics, a quantity must be written with correct units. Otherwise, it is meaningless and, therefore, of no significance.

Example 1.1 : Anand, Rina and Kaif were asked by their teacher to measure the volume of water in a beaker.

Anand wrote : 200; Rina wrote : 200 mL; Kaif wrote : 200 Lm

Which one of these answers is correct?

Solution : The first one has no units. Therefore, we do not know what it means. The third is also not correct because there is no unit like Lm. The second one is the only correct answer. It denotes millilitre.

Note that the mass of a book, for example, can be expressed in kg or g. **You should not use gm for gram because the correct symbol is g and not gm.**

Nomenclature and Symbols

(i) Symbols for units should not contain a full stop and should remain the same in the plural. For example, the length of a pencil should be expressed as 7cm and not 7cm. or 7cms.

(ii) Double prefixes should be avoided when single prefixes are available, e.g., for nanosecond, we should write ns and not mμs; for pico farad we write pF and not μμf.

(iii) When a prefix is placed before the symbol of a unit, the combination of prefix and symbol should be considered as one symbol, which can be raised to a positive or a negative power without using brackets, e.g., μs^{-1}, cm^2, mA^2.

$\mu s^{-1} = (10^{-6} s)^{-1}$ (and not $10^{-6} s^{-1}$)

(iv) Do not write cm/s/s for cm s^{-2}. Similarly 1 poise = 1 g $s^{-1} cm^{-1}$ and not 1 g/s/cm.

(v) When writing a unit in full in a sentence, the word should be spelt with the letter in lower case and not capital, e.g., 6 hertz and not 6 Hertz.

(vi) For convenience in reading of large numbers, the digits should be written in groups of three starting from the right but no comma should be used: 1 532; 1 568 320.

Albert Abraham Michelson (1852-1931)

German-American Physicst, inventor and experimenter devised Michelson's interferometer with the help of which, in association with Morley, he tried to detect the motion of earth with respect to ether but failed. However, the failed experiment stirred the scientific world to reconsider all old theories and led to a new world of physics.

He developed a technique for increasing the resolving power of telescopes by adding external mirrors. Through his stellar interferometer along with 100″ Hookes telescope, he made some precise measurements about stars.

Now, it is time to check your progress. Solve the following questions. In case you have any problem, check answers given at the end of the lesson.

INTEXT QUESTIONS 1.1

1. Discuss the nature of laws of physics.
2. How has the application of the laws of physics led to better quality of life?
3. What is meant by significant figures in measurement?
4. Find the number of significant figures in the following quantity, quoting the relevant laws:

   (i) 426.69 (ii) 4200304.002 (iii) 0.3040 (iv) 4050 m (v) 5000
5. The length of an object is 3.486 m, if it is expressed in centimetre (i.e. 348.6 cm) will there be any change in number of significant figures in the two cases.
6. What are the four applications of the principles of dimensions? On what principle are the above based?
7. The mass of the sun is 2×10^{30} kg. The mass of a proton is 2×10^{-27} kg. If the sun was made only of protons, calculate the number of protons in the sun?
8. Earlier the wavelength of light was expressed in angstroms. One angstrom equals 10^{-8} cm. Now the wavelength is expressed in nanometers. How many angstroms make one nanometre?
9. A radio station operates at a frequency of 1370 kHz. Express this frequency in GHz.
10. How many decimetres are there in a decametre? How many MW are there in one GW?

1.3 DIMENSIONS OF PHYSICAL QUANTITIES

Most physical quantities you would come across in this course can be expressed in terms of five basic dimensions : mass (M), length (L), time (T), electrical current (I) and temperature (θ). Since all quantities in mechanics can be expressed in terms of mass, length and time, it is sufficient for our present purpose to deal with only these three dimensions. Following examples show how dimensions of the physical quantities are combinations of the powers of M, L and T :

(i) Volume requires 3 measurements in length. So it has 3 dimensions in length (L^3).

(ii) Density is mass divided by volume. Its dimensional formula is ML^{-3}.

(iii) Speed is distance travelled in unit time or length divided by time. Its dimensional formula is LT^{-1}.

(iv) Acceleration is change in velocity per unit time, i.e., length per unit time per unit time. Its dimensionsal formula is LT^{-2}.

(v) Force is mass multiplied by acceleration. Its dimensions are given by the formula MLT^{-2}.

Similar considerations enable us to write dimensions of other physical quantities.

Note that numbers associated with physical quantities have no significance in dimensional considerations. Thus if dimension of x is L, then dimension of $3x$ will also be L.

Write down the dimensions of momentum, which is product of mass and velocity and work which is product of force and displacement.

Remember that dimensions are not the same as the units. For example, speed can be measured in m s^{-1} or kilometre per hour, but its dimensions are always given by length divided by time, or simply LT^{-1}.

Dimensional analysis is the process of checking the dimensions of a quantity, or a combination of quantities. One of the important principles of dimensional analysis is that **each physical quantity on the two side of an equation must have the same dimensions**. Thus if $x = p + q$, then p and q will have the same dimensions as x. This helps us in checking the accuracy of equations, or getting the dimensions of a quantity using an equation. The following examples illustrate the use of dimensional analysis.

1.3.1 Applications of Dimensions (or dimensional equations)

There are four applications of dimensions (or dimensional equations)

(i) Derivation of a relationship between different physical quantities (or formula);

(ii) Checking up of accuracy of a formula (or relationship between different physical quantities);

(iii) Conversion of one system of units into another; and

(iv) Derivation of units of a physical quantity

The above applications are based on the principle that the dimensions of physical quantities on the two sides of a relation/equation/formula must be the same. This is called 'the Principle of Homogeneity of Dimensions'.

Example 1.2 : You know that the kinetic energy of a particle of mass m is $\frac{1}{2}m\upsilon^2$ while its potential energy is mgh, where υ is the velocity of the particle, h is its height from the ground and g is the acceleration due to gravity. Since the two expressions represent the same physical quantity i.e, energy, their dimensions must be the same. Let us prove this by actually writing the dimensions of the two expressions.

Solution : The dimensions of $\frac{1}{2}m\upsilon^2$ are $M.(LT^{-1})^2$, or ML^2T^{-2}. (Remember that the numerical factors have no dimensions.) The dimensions of mgh are $M.LT^{-2}.L$, or ML^2T^{-2}. Clearly, the two expressions are the same and hence represent the same physical quantity.

Let us take another example to find an expression for a physical quantity in terms of other quantities.

Example 1.3 : Experience tells us that the distance covered by a car, say x, starting from rest and having uniform acceleration depends on time t and acceleration a. Let us use dimensional analysis to find expression for the distance covered.

Solution : Suppose x depends on the mth power of t and nth power of a. Then we may write

$$x \propto t^m . a^n$$

Expressing the two sides now in terms of dimensions, we get

$$L^1 \propto T^m (LT^{-2})^n,$$

or, $$L^1 \propto T^{m-2n} L^n.$$

Comparing the powers of L and T on both sides, you will easily get $n = 1$, and $m = 2$. Hence, we have

$$x \propto t^2 a^1, \text{ or } x \propto at^2.$$

This is as far as we can go with dimensional analysis. It does not help us in getting the numerical factors, since they have no dimensions. To get the numerical factors, we have to get input from experiment or theory. In this particular case, of course, we know that the complete relation is $x = (1/2)at^2$.

Besides numerical factors, other quantities which do not have dimensions are angles and arguments of trigonometric functions (sine, cosine, etc), exponential and logarithmic functions. In sin x, x is said to be the argument of sine function. In e^x, x is said to be the argument of the exponential function.

Now take a pause and attempt the following questions to check your progress.

INTEXT QUESTIONS 1.2

1. Experiments with a simple pendulum show that its time period depends on its length (l) and the acceleration due to gravity (g). Use dimensional analysis to obtain the dependence of the time period on l and g .

2. Consider a particle moving in a circular orbit of radius r with velocity υ and acceleration a towards the centre of the orbit. Using dimensional analysis, show that $a \propto \upsilon^2/r$.

3. You are given an equation: $m\upsilon = Ft$, where m is mass, υ is speed, F is force and t is time. Check the equation for dimensional correctness.

1.4 VECTORS AND SCALARS

1.4.1 Scalar and Vector Quantities

In physics we classify physical quantities in two categories. In one case, we need only to state their magnitude with proper units and that gives their complete description. Take, for example, mass. If we say that the mass of a ball is 50 g, we do not have to add anything to the description of mass. Similarly, the statement that the density of water is 1000 kg m^{-3} is a complete description of density. Such quantities are called scalars. **A scalar quantity has only magnitude; no direction.**

On the other hand, there are quantities which require both magnitude and direction for their complete description. A simple example is velocity. The statement that the velocity of a train is 100 km h^{-1} does not make much sense unless we also tell the direction in which the train is moving. Force is another such quantity. We must specify not only the magnitude of the force but also the direction in which the force is applied. Such quantities are called vectors. **A vector quantity has both magnitude and direction.**

Some examples of vector quantities which you come across in mechanics are: displacement (Fig. 1.3), acceleration, momentum, angular momentum and torque etc.

What is about energy? Is it a scalar or a vector?

To get the answer, think if there is a direction associated with energy. If not, it is a scalar.

1.4.2 Representation of Vectors

A vector is represented by a line with an arrow indicating its direction. Take vector AB in Fig. 1.4. The length of the line represents its magnitude on some scale. The arrow indicates its direction. Vector CD is a vector in the same direction but its magnitude is smaller. Vector EF is a vector whose magnitude is the same as that of vector CD, but its direction is different. In any vector, the initial point, (point A in **AB),** is called the **tail** of the vector and the final point, (point B in **AB)** with the arrow mark is called its **tip** (or **head**).

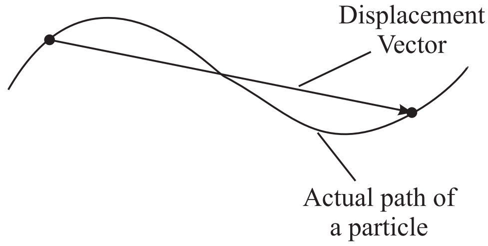


Fig. 1.3 : Displacement vector

A vector is written with an arrow over the letter representing the vector, for example, $\vec{A}$. The magnitude of vector $\vec{A}$ is simply A or $|\vec{A}|$. In print, a vector is indicated by a bold letter as **A**.

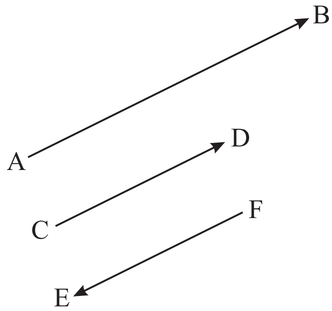


Fig. 1.4 : Directions and magnitudes of vectors

Two vectors are said to be equal if their magnitudes are equal and they point in the same direction. This means that all vectors which are parallel to each other have the same magnitude and point in the same direction are equal. Three vectors **A, B** and **C** shown in Fig. 1.5 are equal. We say **A = B = C**. But **D** is not equal to **A**.

A vector (here **D**) which has the same magnitude as A but has opposite direction, is called **negative** of **A**, or **–A**. Thus, **D** = **–A**.

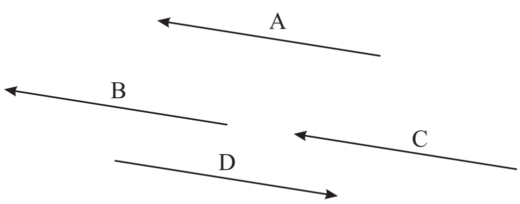


Fig. 1.5 : Three vectors are equal but fourth vector D is not equal.

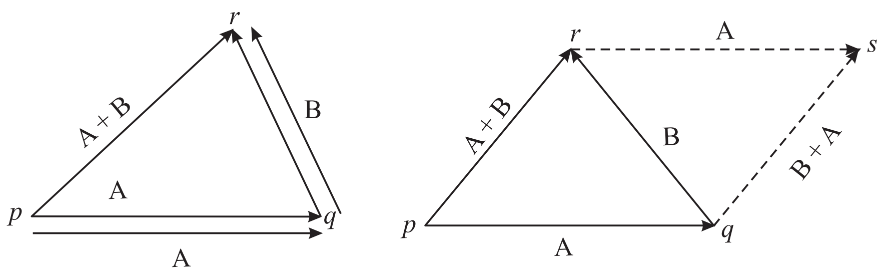
For respresenting a physical vector quantitatively, we have to invariably choose a proportionality scale. For instance, the vector displacement between Delhi and Agra, which is about 300 km, is represented by choosing a scale 100 km = 1 cm, say. Similarly, we can represent a force of 30 N by a vector of length 3cm by choosing a scale 10N = 1cm.

From the above we can say that if we translate a vector parallel to itself, it remains unchanged. This important result is used in addition of vectors. Let us sec how.

1.4.3 Addition of Vectors

You should remember that only **vectors of the same kind can be added**. For example, two forces or two velocities can be added. But a force and a velocity cannot be added.

Suppose we wish to add vectors **A** and **B.** First redraw vector **A** [Fig. 1.6 (a)]. For this draw a line (say *pq*) parallel to vector **A.** The length of the line i.e. *pq* should be equal to the magnitude of the vector. Next draw vector **B** such that its tail coincides with the tip of vector **A**. For this, draw a line *qr* from the tip of **A** (i.e., from the point *q*) parallel to the direction of vector **B.** The sum of two vectors then is the vector from the tail of **A** to the tip of **B,** i.e. the resultant will be represented in magnitude and direction by line *pr.* You can now easily prove that **vector addition is commutative. That is, A + B = B + A,** as shown in Fig. 1.6 (b). In Fig. 1.6(b) we observe that *pqr is* a triangle and its two sides *pq* and *qr* respectively represent the vectors **A** and **B** in magnitude and direction, and the third side *pr*, of the triangle represents the resultant vector with its direction from *p* to *r*. This gives us a rule for finding the resultant of two vectors :



(a) (b)

Fig. 1.6 : Addition of vectors A and B

If two vectors are represented in magnitude and direction by the two sides of a triangle taken in order, the resultant is represented by the third side of the triangle taken in the opposite order. This is called triangle law of vectors.

The sum of two or more vectors is called the **resultant** vector. In Fig. 1.6(b), **pr** is the resultant of **A** and **B**. What will be the resultant of three forces acting along the three sides of a triangle in the same order? If you think that it is zero, you are right.

Notes

Let us now learn to calculate resultant of more than two vectors.

The resultant of more than two vectors, say **A**, **B** and **C**, can be found in the same manner as the sum of two vectors. First we obtain the sum of **A** and **B**, and then add the resultant of the two vectors, (**A + B**), to **C**. Alternatively, you could add **B** and **C**, and then add **A** to (**B + C**) (Fig. 1.7). In both cases you get the same vector. Thus, vector addition is associative. That is, **A + (B + C) = (A + B) + C.**

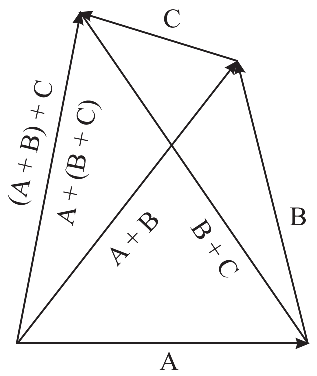


Fig. 1.7 : Addition of three vectors in two different orders

If you add more than three vectors, you will discover that **the resultant vector is the vector from the tail of the first vector to the tip of the last vector.**

Many a time, the point of application of vectors is the same. In such situations, it is more convenient to use parallelogram law of vector addition. Let us now learn about it.

1.4.4 Parallelogram Law of Vector Addition

Let **A** and **B** be the two vectors and let θ be the angle between them as shown in Fig. 1.8. To calculate the vector sum, we complete the parallelogram. Here side PQ represents vector **A**, side PS represents **B** and the diagonal PR represents the resultant vector **R**. Can you recognize that the diagonal PR is the sum vector **A + B**? It is called the **resultant** of vectors **A** and **B**. The resultant makes an angle α with the direction of vector **A**. Remember that vectors **PQ** and **SR** are equal to **A**, and vectors **PS** and **QR** are equal, to **B**. To get the magnitude of the resultant vector **R**, drop a perpendicular RT as shown. Then in terms of magnitudes

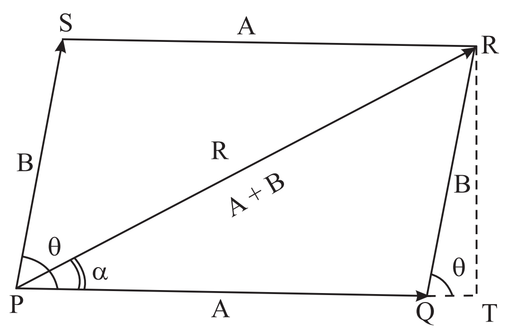


Fig. 1.8: Parallelogram law of addition of vectors

$$
\begin{aligned}
(PR)^2 &= (PT)^2 + (RT)^2 \\
&= (PQ + QT)^2 + (RT)^2 \\
&= (PQ)^2 + (QT)^2 + 2PQ.QT + (RT)^2 \\
&= (PQ)^2 + [(QT)^2 + (RT)^2] + 2PQ.QT \qquad (1.1) \\
&= (PQ)^2 + (QR)^2 + 2PQ.QT \\
&= (PQ)^2 + (QR)^2 + 2PQ.QR\ (QT / QR) \\
R^2 &= A^2 + B^2 + 2AB.\cos\theta
\end{aligned}
$$

Therefore, the magnitude of **R** is

$$|R| = \sqrt{A^2 + B^2 + 2AB.\cos\theta} \qquad (1.2)$$

For the direction of the vector **R**, we observe that

$$\tan\alpha = \frac{RT}{PT} = \frac{RT}{PQ + QT} = \frac{B\sin\theta}{A + B\cos\theta} \qquad (1.3)$$

So, the direction of the resultant can be expressed in terms of the angle it makes with base vector.

Special Cases

Now, let us consider as to what would be the resultant of two vectors when they are parallel?

To find answer to this question, note that the angle between the two parallel vectors is zero and the resultant is equal to the sum of their magnitudes and in the direction of these vectors.

Suppose that two vectors are perpendicular to each other. What would be the magnitude of the resultant? In this case, $\theta = 90^\circ$ and $\cos\theta = 0$.

Suppose further that their magnitudes are equal. What would be the direction of the resultant?

Notice that $\tan\alpha = B/A = 1$. So what is α?

Also note that when $\theta = \pi$, the vectors become anti-parallel. In this case $\alpha = 0$. The resultant vector will be along **A** or **B**, depending upon which of these vectors has larger magnitude.

Example 1.4: A cart is being pulled by Ahmed north-ward with a force of magnitude 70 N. Hamid is pulling the same cart in the south-west direction with a force of magnitude 50 N. Calculate the magnitude and direction of the resulting force on the cart.

Solution :

Here, magnitude of first force, say, A = 70 N.

The magnitude of the second force, say, B = 50 N.

Angle θ between the two forces = 135 degrees.

So, the magnitude of the resultant is given by Eqn. (1.2) :

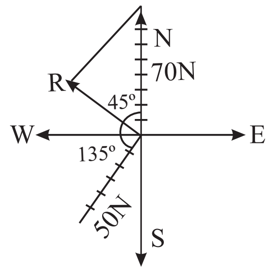


Fig. 1.9: Resultant of forces inclined at an angle

$$R = \sqrt{(70)^2 + (50)^2 + 2 \times 70 \times 50 \times \cos(135)}$$

$$= \sqrt{4900 + 2500 - 7000 \times \sin 45}$$

$$= 49.5 \text{ N}$$

The magnitude of **R** = 49.5 N.

The direction is given by Eqn. (1.3):

$$\tan\alpha = \frac{B\sin\theta}{A + B\cos\theta} = \frac{50 \times \sin(135)}{70 + 50\cos(135)} = \frac{50 \times \cos 45}{70 - 50\sin 45} = 1.00$$

Therefore, α = 45.0° (from the tables). Thus **R** makes an angle of 45° with 70 N force. That is, **R** is in North-west direction as shown in Fig. 1.9.

1.4.5 Subtraction of Vectors

How do we subtract one vector from another? If you recall that the difference of two vectors, **A – B**, is actually equal to **A + (–B)**, then you can adopt the same method as for addition of two vectors. It is explained in Fig. 1.10. Draw vector **–B** from the tip of **A**. Join the tail of **A** with the tip of **–B**. The resulting vector is the difference **(A – B)**.

You may now like to check your progress.

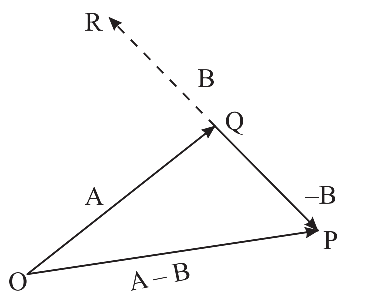
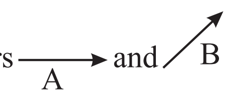


Fig. 1.10 : Subtraction of vector B from vector A

INTEXT QUESTIONS 1.3

Given vectors  and B

1. Make diagrams to show how to find the following vectors:

   (a) **B – A**, (b) **A + 2B**, (c) **A – 2B** and (d) **B – 2A**.

Notes

2. Two vectors **A** and **B** of magnitudes 10 units and 12 units are anti-parallel. Determine **A + B** and **A – B**.

3. Two vectors **A** and **B** of magnitudes A = 30 units and B = 60 units respectively are inclined to each other at angle of 60 degrees. Find the resultant vector.

1.5 MULTIPLICATION OF VECTORS

1.5.1 Multiplication of a Vector by a Scalar

If we multiply a vector **A** by a scalar k, the product is a vector whose magnitude is the absolute value of k times the magnitude of **A**. This means that the magnitude of the resultant vector is $k\,|\mathbf{A}|$. The direction of the new vector remains unchanged if k is positive. If k is negative, the direction of the new vector is opposite to its original direction. For example, vector 3**A** is thrice the magnitude of vector **A**, and it is in the same direction as **A**. But vector –3**A** is in a direction opposite to vector **A**, although its magnitude is thrice that of vector **A**.

1.5.2 Scalar Product of Vectors

The **scalar product** of two vectors **A** and **B** is written as **A.B** and is equal to AB cosθ, where θ is the angle between the vectors. If you look carefully at Fig. 1.11, you would notice that B cosθ is the projection of vector **B** along vector **A**. Therefore, the scalar product of **A** and **B** is the product of magnitude of **A** with the length of the projection of **B** along **A**. Another thing to note is that even if we take the angle between the two vectors as 360 – θ, it does not matter because the cosine of both angles is the same. Since a dot between the two vectors indicates the scalar product, it is also called the **dot product. Remember that the scalar product of two vectors is a scalar quantity.**

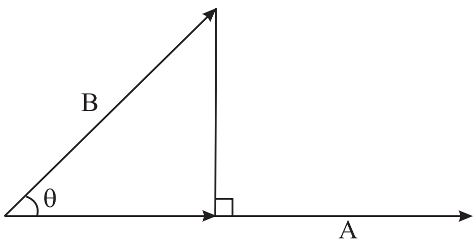


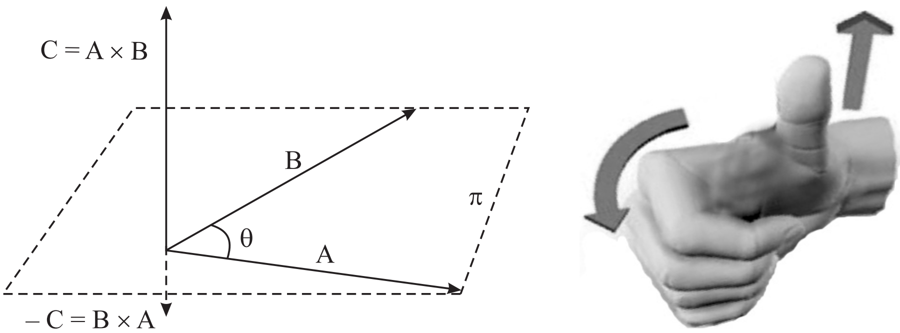
Fig. 1.11: Projection of B on A

A familiar example of the scalar product is the work done when a force **F** acts on a body moving at an angle to the direction of the force. If **d** is the displacement of the body and θ is the angle between **F** and **d**, then the work done by the force is $Fd\cos\theta$.

Since dot product is a scalar, it is commutative: **A.B** = **B.A** = ABcosθ. It is also distributive: **A.(B + C) = A.B + A.C**.

1.5.3 Vector Product of Vectors

Suppose we have two vectors **A** and **B** inclined at an angle θ. We can draw a plane which contains these two vectors. Let that plane be called Ω ((Fig. 1.12 a)



(a) (b)

Fig.1.12 (a) : Vector product of Vectors; (b) Direction of the product vector C =A × B is given by the right hand rule. If the right hand is held so that the curling fingers point from A to B through the smaller angle between the two, then the thumb strectched at right angles to fingers will point in the direction of C.

which is perpendicular to the plane of paper here. Then the vector product of these vectors, written as **A** × **B**, is a vector, say **C**, whose magnitude is AB sinθ and whose direction is perpendicular to the plane Ω. The direction of the vector **C** can be found by **right-hand rule** (Fig. 1.12 b). Imagine the fingers of your right hand curling from **A** to **B** along the smaller angle between them. Then the direction of the thumb gives the direction of the product vector **C**. If you follow this rule, you can easily see that direction of vector **B** × **A** is opposite to that of the vector **A** × **B**. This means that **the vector product is not commutative**. Since a cross is inserted between the two vectors to indicate their vector product, the vector product is also called the cross product.

A familiar example of vector product is the angular momentum possessed by a rotating body.

To check your progress, try the following questions.

INTEXT QUESTIONS 1.4

1. Suppose vector **A** is parallel to vector **B**. What is their vector product? What will be the vector product if **B** is anti-parallel to **A**?

2. Suppose we have a vector **A** and a vector **C** = $\frac{1}{2}$ **B.** How is the direction of vector **A** × **B** related to the direction of vector **A** × **C**.

3. Suppose vectors **A** and **B** are rotated in the plane which contains them. What happens to the direction of vector **C** = **A** × **B**.

4. Suppose you were free to rotate vectors **A** and **B** through arbitrary amounts keeping them confined to the same plane. Can you make vector **C** = **A** × **B** to point in exactly opposite direction?

5. If vector **A** is along the x-axis and vector **B** is along the y-axis, what is the direction of vector $\mathbf{C} = \mathbf{A} \times \mathbf{B}$? What happens to **C** if **A** is along the y-axis and **B** is along the x-axis?
6. **A** and **B** are two mutually perpendicular vectors. Calculate (a) **A . B** and (b) **A × B.**

1.6 RESOLUTION OF VECTORS

Resolution of vectors is converse of addition of vectors. Here we calculate components of a given vector along any set of coordinate axes. Suppose we have vector **A** as shown in Fig. 1.13 and we need to find its components along x and y-axes. Let these components be called A_x and A_y respectively. Simple trigonometry shows that

$$A_x = A \cos\theta \qquad (1.4)$$

and

$$A_y = A \sin\theta, \qquad (1.5)$$

where θ is the angle that **A** makes with the x - axis. What about the components of vector **A** along X and Y-axes (Fig. 1.13)? If the angle between the X-axis and **A** is φ, then

$$A_X = A \cos\phi$$

and

$$A_Y = A \sin\phi.$$

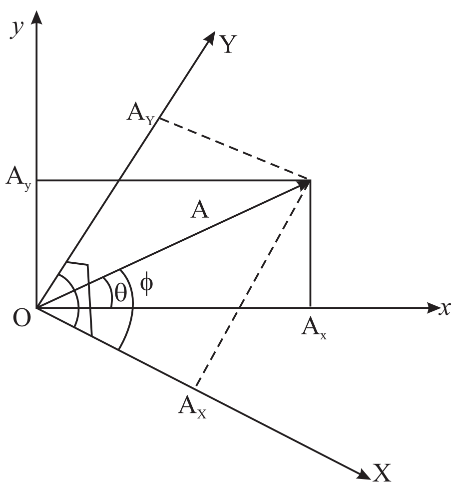


Fig. 1.13 : Resolution of vector A along two sets of coordinates (x, y) and (X, Y)

It must now be clear that the components of a vector are not fixed quantities; they depend on the particular set of axes along which components are required. Note also that the magnitude of vector **A** and its direction in terms of its components are given by

$$A = \sqrt{A_x^2 + A_y^2} = \sqrt{A_X^2 + A_Y^2} \qquad (1.6)$$

and $\tan\theta = A_y / A_x$, $\tan\phi = A_Y / A_X$. (1.7)

So, if we are given the components of a vector, we can combine them as in these equations to get the vector.

1.7 UNIT VECTOR

At this stage we introduce the concept of a **unit vector**. As the name suggests, a unit vector has unitary magnitude and has a specified direction. It has no units and no dimensions. As an example, we can write vector **A** as A $\hat{\mathbf{n}}$ where a cap on **n** (i.e. $\hat{\mathbf{n}}$) denotes a unit vector in the direction of **A**. Notice that a unit vector has been introduced to take care of the direction of the vector; the magnitude has been taken care of by A. Of particular importance are the unit vectors along coordinate axes. Unit vector along *x*-axis is denoted by $\hat{\mathbf{i}}$, along *y*-axis by $\hat{\mathbf{j}}$ and along *z*-axis by $\hat{\mathbf{k}}$. Using this notation, vector **A**, whose components along *x* and *y* axes are respectively A_x and A_y, can be written as

$$\mathbf{A} = A_x\,\hat{\mathbf{i}} + A_y\,\hat{\mathbf{j}}\,. \qquad (1.8)$$

Another vector **B** can similarly be written as

$$\mathbf{B} = B_x\,\hat{\mathbf{i}} + B_y\,\hat{\mathbf{j}}\,. \qquad (1.9)$$

The sum of these two vectors can now be written as

$$\mathbf{A} + \mathbf{B} = (A_x + B_x)\,\hat{\mathbf{i}} + (A_y + B_y)\,\hat{\mathbf{j}} \qquad (1.10)$$

By the rules of scalar product you can show that

$$\hat{\mathbf{i}}.\,\hat{\mathbf{i}}=1,\ \hat{\mathbf{j}}.\hat{\mathbf{j}}=1,\ \hat{\mathbf{k}}.\hat{\mathbf{k}}=1,\ \hat{\mathbf{i}}.\hat{\mathbf{j}}=0,\ \hat{\mathbf{i}}.\hat{\mathbf{k}}=0, \text{ and } \hat{\mathbf{j}}.\hat{\mathbf{k}}=0 \qquad (1.11)$$

The dot product between two vectors **A** and **B** can now be written as

$$\begin{aligned}\mathbf{A}\,.\,\mathbf{B} &= (A_x\,\hat{\mathbf{i}} + A_y\,\hat{\mathbf{j}}).\,(B_x\,\hat{\mathbf{i}} + B_y\,\hat{\mathbf{j}}) \\ &= A_xB_x\,(\hat{\mathbf{i}}\,.\,\hat{\mathbf{i}}) + A_xB_y\,(\hat{\mathbf{i}}\,.\,\hat{\mathbf{j}}) + A_yB_x\,(\hat{\mathbf{j}}\,.\,\hat{\mathbf{i}}) + A_yB_y\,(\hat{\mathbf{j}}\,.\,\hat{\mathbf{j}}) \\ &= A_xB_x + A_yB_y, \qquad (1.12)\end{aligned}$$

Here, we have used the results contained in Eqn. (1.11).

Example 1.4: On a coordinate system (showing all the four quadrants) show the following vectors:

$$\mathbf{A} = 4\hat{\mathbf{i}} + 0\hat{\mathbf{j}},\ \mathbf{B} = 0\hat{\mathbf{i}} + 5\hat{\mathbf{j}},\ \mathbf{C} = 4\hat{\mathbf{i}} + 5\hat{\mathbf{j}},$$

$$\mathbf{D} = 6\hat{\mathbf{i}} - 4\hat{\mathbf{j}}.$$

Find their magnitudes and directions.

Solution : The vectors are given in component form. The factor multiplying $\hat{\mathbf{i}}$ is the x component and the factor multiplying $\hat{\mathbf{j}}$ is the y component. All the vectors are shown on the coordinate grid (Fig. 1.14).

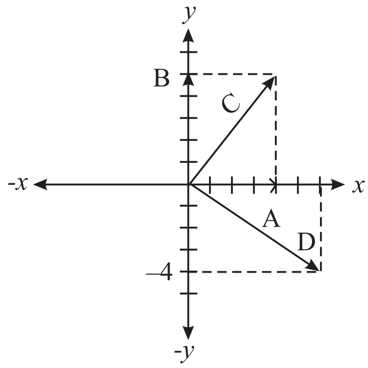


Fig. 1.14

The components of **A** are $\mathbf{A}_x = 4$, $\mathbf{A}_y = 0$. So, the magnitude of **A** = 4. Its direction is $\tan^{-1}\left(\frac{A_y}{A_x}\right)$ in accordance with Eqn. (1.7). This quantity is zero, since $A_y = 0$. This makes it to be along the x-axis, as it is. Vector **B** has x-component = 0, so it lies along the y-axis and its magnitude is 5.

Let us consider vector **C**. Here, $\mathbf{C}_x = 4$ and $\mathbf{C}_y = 5$. Therefore, the magnitude of **C** is $C = \sqrt{4^2+5^2} = \sqrt{41}$. The angle that it makes with the x-axis is $\tan^{-1}(\mathbf{C}_y/\mathbf{C}_x)$ = 51.3 degrees. Similarly, the magnitude of **D** is $D = \sqrt{60}$. Its direction is $\tan^{-1}(\mathbf{D}_y/\mathbf{D}_x) = \tan^{-1}(0.666) = -33.7°$ (in the fourth quadrant).

Example 1.5: Calculate the product **C . D** for the vectors given in Example 1.4.

Solution : The dot product of **C** with **D** can be found using Eqn. (1.12):

$$\mathbf{C} \cdot \mathbf{D} = C_xD_x + C_yD_y = 4\times6 + 5\times(-4) = 24 - 20 = 4.$$

The cross product of two vectors can also be written in terms of the unit vectors. For this we first need the cross product of unit vectors. For this remember that the angle between the unit vectors is a right angle. Consider, for example, $\hat{\mathbf{i}} \times \hat{\mathbf{j}}$. Sine of the angle between them is one. The magnitude of the product vector is also 1. Its direction is perpendicular to the xy - plane containing $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$, which is the z-axis. By the right hand rule, we also find that this must be the positive z-axis. And what is the unit vector in the positive z - direction. The unit vector $\hat{\mathbf{k}}$. Therefore,

$$\hat{\mathbf{i}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}}. \qquad (1.13)$$

Using similar arguments, we can show,

$$\hat{\mathbf{j}} \times \hat{\mathbf{k}} = \hat{\mathbf{i}},\ \hat{\mathbf{k}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}},\ \hat{\mathbf{j}} \times \hat{\mathbf{i}} = -\hat{\mathbf{k}},\ \hat{\mathbf{k}} \times \hat{\mathbf{j}} = -\hat{\mathbf{i}},\ \hat{\mathbf{i}} \times \hat{\mathbf{k}} = -\hat{\mathbf{j}}, \qquad (1.14)$$

and $$\hat{\mathbf{i}} \times \hat{\mathbf{i}} = \hat{\mathbf{j}} \times \hat{\mathbf{j}} = \hat{\mathbf{k}} \times \hat{\mathbf{k}} = 0. \qquad (1.15)$$

Example 1.6: Calculate the cross product of vectors **C** and **D** given in Example (1.4).

Solution : We have

$$\mathbf{C} \times \mathbf{D} = (4\,\hat{\mathbf{i}} + 5\,\hat{\mathbf{j}}) \times (6\,\hat{\mathbf{i}} - 4\,\hat{\mathbf{j}})$$

$$= 24\,(\hat{\mathbf{i}} \times \hat{\mathbf{i}}) - 16\,(\hat{\mathbf{i}} \times \hat{\mathbf{j}}) + 30\,(\hat{\mathbf{j}} \times \hat{\mathbf{i}}) - 20\,(\hat{\mathbf{j}} \times \hat{\mathbf{j}})$$

Notes

Using the results contained in Eqns. (1.13 – 1.15), we can write

$$\mathbf{C} \times \mathbf{D} = -16\,\hat{\mathbf{k}} - 30\,\hat{\mathbf{k}} = -46\,\hat{\mathbf{k}}$$

So, the cross product of **C** and **D** is a vector of magnitude 46 and in the negative z direction. Since **C** and **D** are in the xy-plane, it is obvious that the cross product must be perpendicular to this plane, that is, it must be in the z-direction.

INTEXT QUESTIONS 1.5

1. A vector **A** makes an angle of 60 degrees with the x-axis of the xy-system of coordinates. If its magnitude is 50 units, find its components in x, y directions. If another vector **B** of the same magnitude makes an angle of 30 degrees with the X-axis of the XY- system of coordinates. Find its components now. Are they same as before?

2. Two vectors **A** and **B** are given respectively as $3\,\hat{\mathbf{i}} - 4\hat{\mathbf{j}}$ and $-2\,\hat{\mathbf{i}} + 6\hat{\mathbf{j}}$. Sketch them on the coordinate grid. Find their magnitudes and the angles that they make with the x-axis (see Fig. 1.14).
3. Calculate the dot and cross product of the vectors given in the above question.

You now know that each term in an equation must have the same dimensions. Having learnt vectors, we must now add this: **For an equation to be correct, each term in it must have the same character: either all of them be vectors or all of them be scalars.**

WHAT YOU HAVE LEARNT

- The number of significant figures determines the accuracy of a measurement.
- Every physical quantity must be measured in some unit and also expressed in this unit. The SI system has been accepted and followed universally for scientific reporting.
- Base SI units for mass, length and time are respectively kg, m and s. In addition to base units, there are derived units.
- Every physical quantity has dimensions. Dimensional analysis is a useful tool for checking correctness of equations.
- In physics, we deal generally with two kinds of quantities, scalars and vectors. A scalar has only magnitude. A vector has both direction and magnitude.
- Vectors are added according to the parallelogram rule.
- The scalar product of two vectors is a scalar.

- The vector product of two vectors is a vector perpendicular to the plane containing the two vectors.
- Vectors can be resolved into components along a specified set of coordinates axes.

TERMINAL EXERCISE

1. A unit used for measuring very large distances is called a light year. It is the distance covered by light in one year. Express light year in metres. Take speed of light as 3×10^8 m s^{-1}.

2. Meteors are small pieces of rock which enter the earth's atmosphere occasionally at very high speeds. Because of friction caused by the atmosphere, they become very hot and emit radiations for a very short time before they get completely burnt. The streak of light that is seen as a result is called a 'shooting star'. The speed of a meteor is 51 km s^{-1} In comparison, speed of sound in air at about 20^{0}C is 340 m s^{-1} Find the ratio of magnitudes of the two speeds.

3. The distance covered by a particle in time t while starting with the initial velocity u and moving with a uniform acceleration a is given by $s = ut + (1/2)at^2$. Check the correctness of the expression using dimensional analysis.

4. Newton's law of gravitation states that the magnitude of force between two particles of mass m_1 and m_2 separated by a distance r is given by

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

   where G is the universal constant of gravitation. Find the dimensions of G.

5. Hamida is pushing a table in a certain direction with a force of magnitude 10 N. At the same time her, classmate Lila is pushing the same table with a force of magnitude 8 N in a direction making an angle of 60^o to the direction in which Hamida is pushing. Calculate the magnitude of the resultant force on the table and its direction.

6. A physical quantity is obtained as a dot product of two vector quantities. Is it a scalar or a vector? What is the nature of a physical quantity obtained as cross product of two vectors?

7. John wants to pull a cart applying a force parallel to the ground. His friend Ramu suggests that it would be easier to pull the cart by applying a force at an angle of 30 degrees to the ground. Who is correct and why?

8. Two vectors are given by $5\,\hat{\mathbf{i}} - 3\,\hat{\mathbf{j}}$ and $3\,\hat{\mathbf{i}} - 5\,\hat{\mathbf{j}}$. Calculate their scalar and vector products.

ANSWERS TO INTEXT QUESTIONS

1.1

4. (i) 5 (ii) 10 (iii) 4 (iv) 4 (v) 1

5. No, in both cases, the number of significant figures will be 4.

7. Mass of the sun = 2×10^{30} kg

   Mass of a proton = 2×10^{-27} kg

   (No of protons in the sun = $\dfrac{2\times10^{30}\ \text{kg}}{2\times10^{-27}\ \text{kg}} = 10^{57}$.

8. 1 angstrom = 10^{-8} cm = 10^{-10} m

   1 nanometer (nm) = 10^{-9} m

   $\therefore$ 1 nm/1 angstrom = 10^{-9} m /10^{-10} m = 10 so 1 nm = 10 Å

9. 1370 kHz = 1370×10^3 Hz = $(1370 \times 10^3)/10^9$ GHz = 1.370×10^{-3} GHz

10. 1 decameter (dam) = 10 m

    1 decimeter (dm) = 10^{-1} m

    $\therefore$ 1 dam = 100 dm

    1 MW = 10^6 W

    1 GW = 10^9 W

    $\therefore$ 1 GW = 10^3 MW

1.2

1. Dimension of length = L

   Dimension of time = T

   Dimensions of g = LT^{-2}

   Let time period t be proportional to l^α and g^β

   Then, writing dimensions on both sides $T = L^\alpha (LT^{-2})^\beta = L^{\alpha+\beta} T^{-2\beta}$

   Equating powers of L and T,

   $\alpha + \beta = 0,\ 2\beta = -1 \Rightarrow \beta = -1/2$ and $\alpha = 1/2$

   So, $t \ \alpha \sqrt{\dfrac{l}{g}}$.

2. Dimension of $a = LT^{-2}$

   Dimension of $\upsilon = LT^{-1}$

   Dimension of $r = L$

   Let a be proportional to υ^{α} and r^{β}

   Then dimensionally,

   $$LT^{-2} = (LT^{-1})^{\alpha} L^{\beta} = L^{\alpha+\beta} T^{-\alpha}$$

   Equating powers of L and T,

   $$\alpha + \beta = 1, \alpha = 2, \Rightarrow \alpha = -1$$

   So, $\alpha \propto \upsilon^2/r$

3. Dimensions of $m\upsilon = MLT^{-1}$

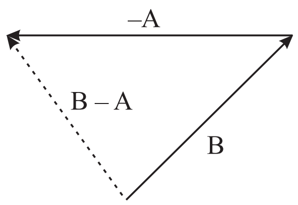
   Dimensions of $Ft = MLT^{-2}\,T^{1} = MLT^{-1}$

   Dimensions of both the sides are the same, therefore, the equation is dimensionally correct.

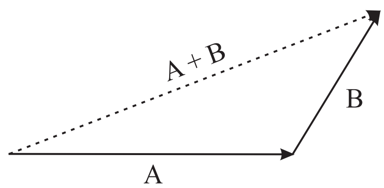
1.3

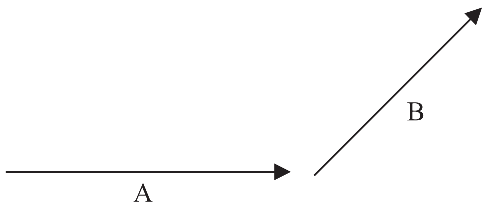
1. Suppose

(a)

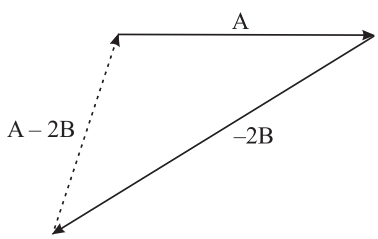


(b)

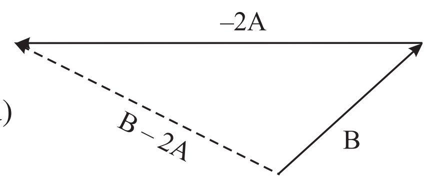




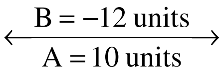
(c)



(d)

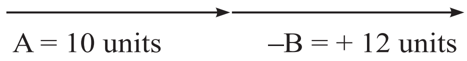


2. 

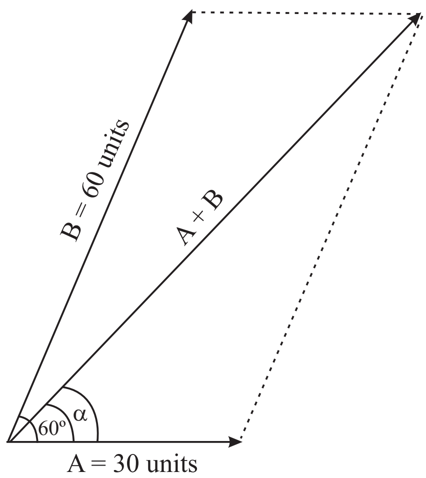


$$A + B = 10 + (-12)$$

$$= -2 \text{ units}$$

also 

$$\mathbf{A - B} = 22 \text{ units}$$

3. 

$$|A + B| = 77 \text{ units}$$

1.4

1. If **A** and **B** are parallel, the angle θ between them is zero. So, their cross product

   $\mathbf{A} \times \mathbf{B} = AB \sin\theta = 0.$

   If they are antiparallel then the angle between them is 180°. Therefore,

   $\mathbf{A} \times \mathbf{B} = AB \sin\theta = 0$, because $\sin 180° = 0$.

2. If magnitude of **B** is halved, but it remains in the same plane as before, then the direction of the vector product $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ remains unchanged. Its magnitude may change.

3. Since vectors **A** and **B** rotate without change in the plane containing them, the direction of $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ will not change.

4. Suppose initially the angle between **A** and **B** is between zero and 180°. Then **C** = **A** × **B** will be directed upward perpendicular to the plane. After rotation through arbitrary amounts, if the angle between them becomes > 180°, then **C** will drop underneath but perpendicular to the plane.

5. If **A** is along x-axis and **B** is along y-axis, then they are both in the xy plane. The vector product **C** = **A** × **B** will be along z-direction. If **A** is along y-axis and **B** is along x-axis, then **C** is along the negative z-axis.

6. (a) $\mathbf{A} \cdot \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \cos\theta = 0$ when $\theta = 90°$

   (b) $\mathbf{A} \times \mathbf{B} = |\mathbf{A}|\,|\mathbf{B}| \sin\theta = |\mathbf{A}|\,|\mathbf{B}|$ as $\sin\theta = 1$ at $\theta = 90°$

1.5

1. When **A** makes an angle of 60° with the x-axis:

   $A_x = A \cos 60 = 50 \cdot ½ = 25$ units

   $A_y = A \sin 60 = 50.\sqrt{3}/2 = 50 \cdot 0.866$

   $= 43.3$ units

   When **A** makes an angle of 30° with the x-axis

   $A_x = 50 \cos 30 = 50 \cdot 0.866 = 43.3$ units

   $A_y = 50 \sin 30 = 50 \cdot ½ = 25$ units

   The components in the two cases are obviously not the same.

2. The position of vectors on the coordinate grid is shown in Fig. 1.14.

   Suppose **A** makes an angle θ with the x-axis, then

   $\tan\theta = -4/3 \Rightarrow \theta = \tan^{-1}(-4/3)$

   $= -53° \, 6' \text{ or } 306° \, 54'$

   after taking account of the quadrant in which the angle lies.

   If B makes an angle ϕ with the x-axis, then

   $\tan\phi = 6/-2 = -3 \Rightarrow \phi = \tan^{-1}(-3)$

   $= 108° \, 24'$

3. The dot product of **A** and **B**:

   $$\mathbf{A} \cdot \mathbf{B} = (3\hat{\mathbf{i}} - 4\hat{\mathbf{j}}).(-2\hat{\mathbf{i}} + 6\hat{\mathbf{j}})$$

   $$= -6(\hat{\mathbf{i}}.\hat{\mathbf{i}}) - 24(\hat{\mathbf{j}}.\hat{\mathbf{j}}) = -30$$

   because $\hat{\mathbf{i}}.\hat{\mathbf{j}} = \hat{\mathbf{j}}.\hat{\mathbf{i}} = 0$, and $\hat{\mathbf{i}}.\hat{\mathbf{i}} = \hat{\mathbf{j}}.\hat{\mathbf{j}} = 1$

   The cross product of **A** and **B**:

$$\mathbf{A} \times \mathbf{B} = (3\hat{\mathbf{i}} - 4\hat{\mathbf{j}}) \times (-2\hat{\mathbf{i}} + 6\hat{\mathbf{j}})$$

$$= 18\,(\hat{\mathbf{i}} \times \hat{\mathbf{j}}) + 8\,(\hat{\mathbf{j}} \times \hat{\mathbf{i}}) = 18\,\hat{\mathbf{k}} - 8\,\hat{\mathbf{k}} = 10\,\hat{\mathbf{k}}$$

on using Eqs.(1.14) and (1.15). So, the cross product is in the direction of z-axis, since **A** and **B** lie in the *xy* plane.

Answers to Terminal Problems

1. 1 ly = 9.4673×10^{15} m.

2. $\dfrac{\text{Speed of meteor}}{\text{Speed of sound in air of } 20^\circ\text{C}} = \dfrac{51}{340} = \dfrac{3}{20}\,\text{S}$

5. 15.84 N and $\alpha = \tan^{-1}\left(\dfrac{1}{2}\right)$

8. $\mathbf{A} \cdot \mathbf{B} = 30$

   $\mathbf{A} \times \mathbf{B} = \left(5\hat{\mathbf{i}} - 3\hat{\mathbf{j}}\right) \times \left(3\hat{\mathbf{i}} - 5\hat{\mathbf{j}}\right)$ is a single vector **C** such that $|\mathbf{C}|$ = 16 units along negative *z*-direction.

312en02

2

MOTION IN A STRAIGHT LINE

We see a number of things moving around us. Humans, animals, vehicles can be seen moving on land. Fish, frogs and other aquatic animals move in water. Birds and aeroplanes move in air. Though we do not feel it, the earth on which we live also revolves around the sun as well as its own axis. It is, therefore, quite apparent that we live in a world that is very much in constant motion. Therefore to understand the physical world around us, the study of motion is essential. Motion can be in a straight line(1D), in a plane(2D) or in space(3D). If the motion of the object is only in one direction, it is said to be the motion in a straight line. For example, motion of a car on a straight road, motion of a train on straight rails, motion of a freely falling body, motion of a lift, and motion of an athlete running on a straight track, etc.

In this lesson you will learn about motion in a straight line. In the following lessons, you will study the laws of motion, motion in plane and other types of motion. You will also learn the concept of Differentiation and Integration.

OBJECTIVES

After studying this lesson, you should be able to,

- *distinguish between distance and displacement, and speed and velocity;*
- *explain the terms instantaneous velocity, relative velocity and average velocity;*
- *define acceleration and instantaneous acceleration;*
- *interpret position - time and velocity - time graphs for uniform as well as non-uniform motion;*
- *derive equations of motion with constant acceleration;*
- *describe motion under gravity;*
- *solve numericals based on equations of motion; and*
- *understand the concept of differentiation and integration.*

2.1 SPEED AND VELOCITY

We know that the total length of the path covered by a body is the **distance** travelled by it. But the difference between the initial and final position vectors of a body is called its **displacement**. Basically, **displacement is the shortest distance between the two positions and has a certain direction**. Thus, the displacement is a vector quantity but distance is a scalar. You might have also learnt that the rate of change of distance with time is called **speed** but the rate of change of displacement is known as **velocity**. Unlike speed, velocity is a vector quantity. For 1-D motion, the directional aspect of the vector is taken care of by putting + and – signs and we do not have to use vector notation for displacement, velocity and acceleration for motion in one dimension.

2.1.1 Average Velocity

When an object travels a certain distance with different velocities, its motion is specified by its average velocity. The **average velocity** of an object is defined as the displacement per unit time. Let x_1 and x_2 be its positions at instants t_1 and t_2, respectively. Then mathematically we can express average velocity as

$$\bar{\upsilon} = \frac{\text{displacement}}{\text{time taken}}$$

$$= \frac{x_2 - x_1}{t_2 - t_1} = \frac{\Delta x}{\Delta t} \tag{2.1}$$

where $x_2 - x_1$ signifies change in position (denoted by Δx) and $t_2 - t_1$ is the corresponding change in time (denoted by Δt). Here the bar over the symbol for velocity ($\bar{\upsilon}$) is standard notation used to indicate an average quantity. Average velocity can be represented as υ_{av} also. The average speed of an object is obtained by dividing the total distance travelled by the total time taken:

$$\text{Average speed} = \frac{\text{total distance travelled}}{\text{total time taken}} \tag{2.2}$$

If the motion is in the same direction along a straight line, the average speed is the same as the magnitude of the average velocity. However, this is always not the case (see example 2.2).

Following examples will help you in understanding the difference between average speed and average velocity.

Example 2.1 : The position of an object moving along the x-axis is defined as $x = 20t^2$, where t is the time measured in seconds and position is expressed in metres. Calculate the average velocity of the object over the time interval from 3s to 4s.

Solution : Given,

$$x = 20t^2$$

Note that x and t are measured in metres and seconds. It means that the constant of proportionality (20) has dimensions ms^{-2}.

We know that the average velocity is given by the relation

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

At $t_1 = 3s$,

$$x_1 = 20 \times (3)^2$$
$$= 20 \times 9 = 180 \text{ m}$$

Similarly, for $t_2 = 4s$

$$x_2 = 20 \times (4)^2$$
$$= 20 \times 16 = 320 \text{ m}$$

$$\therefore \quad \bar{v} = \frac{x_2 - x_1}{t_2 - t_1} = \frac{(320 - 180)\text{ m}}{(4-3)\text{ s}} = \frac{140\text{ m}}{1\text{ s}} = 140 \text{ ms}^{-1}$$

Hence, average velocity = 140 ms^{-1}.

Example 2.2 : A person runs on a 300m circular track and comes back to the starting point in 200s. Calculate the average speed and average velocity.

Solution : Given,

Total length of the track = 300m.

Time taken to cover this length = 200s

Hence,

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

$$= \frac{300}{200} \text{ ms}^{-1} = 1.5 \text{ ms}^{-1}$$

As the person comes back to the same point, the displacement is zero. Therefore, the average velocity is also zero.

Note that in the above example, the average speed is not equal to the magnitude of the average velocity. Do you know the reason?

2.1.2 Relative Velocity

When we say that a bullock cart is moving at 10km h^{-1} due south, it means that the cart travels a distance of 10km in 1h in southward direction from its starting position. Thus it is implied that the referred velocity is with respect to some reference point. In fact, the velocity of a body is always specified with respect to some other body. Since all bodies are in motion, we can say that every velocity is relative in nature.

The relative velocity of an object with respect to another object is the rate at which it changes its position relative to the object / point taken as reference. For example, if υ_A and υ_B are the velocities of the two objects along a straight line, the relative velocity of B with respect to A will be $\upsilon_B - \upsilon_A$.

The rate of change of the relative position of an object with respect to the other object is known as the ***relative velocity*** *of that object with respect to the other.*

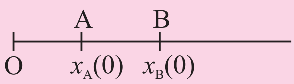
Importance of Relative Velocity

The position and hence velocity of a body is specified in relation with some other body. If the reference body is at rest, the motion of the body can be described easily . You will learn the equations of kinematics in this lesson. But what happens, if the reference body is also moving? Such a motion is seen to be of the two body system by a stationary observer. However, it can be simplified by invoking the concept of relative motion.

Let the initial positions of two bodies A and B be $x_A(0)$ and $x_B(0)$. If body A moves along positive x-direction with velocity υ_A and body B with velocity υ_B, then the positions of points A and B after t seconds will be given by

$$x_A(t) = x_A(0) + \upsilon_A t$$

$$x_B(t) = x_B(0) + \upsilon_B t$$



Therefore, the relative separation of B from A will be

$$x_{BA}(t) = x_B(t) - x_A(t) = x_B(0) - x_A(0) + (\upsilon_B - \upsilon_A)\,t$$

$$= x_{BA}(0) + \upsilon_{BA} t$$

where $\upsilon_{BA} = (\upsilon_B - \upsilon_A)$ is called the relative velocity of B with respect to A. Thus by applying the concept of relative velocity, a two body problem can be reduced to a single body problem.

Example 2.3 : A train A is moving on a straight track (or railway line) from North to South with a speed of 60km h^{-1}. Another train B is moving from South to North with a speed of 70km h^{-1}. What is the velocity of B relative to the train A?

Solution : Considering the direction from South to North as positive, we have

velocity (υ_B) of train B = + 70km h^{-1}

and, velocity (v_A) of train A = – 60km h^{-1}

Hence, the velocity of train B relative to train A

$$= v_B - v_A$$

$$= 70 - (-60) = 130\text{km h}^{-1}.$$

In the above example, you have seen that the relative velocity of one train with respect to the other is equal to the sum of their respective velocities. This is why a train moving in a direction opposite to that of the train in which you are travelling appears to be travelling very fast. But, if the other train were moving in the same direction as your train, it would appear to be very slow.

2.1.3 Acceleration

While travelling in a bus or a car, you might have noticed that sometimes it speeds up and sometimes it slows down. That is, its velocity changes with time. Just as the velocity is defined as the time rate of change of displacement, ***the acceleration is defined as time rate of change of velocity.*** Acceleration is a vector quantity and its SI unit is ms^{-2}. In one dimension, there is no need to use vector notation for acceleration as explained in the case of velocity. The average acceleration of an object is given by,

$$\text{Average acceleration } (\bar{a}) = \frac{\text{Final velocity - Initial velocity}}{\text{Time taken for change in velocity}}$$

$$\bar{a} = \frac{v_2 - v_1}{t_2 - t_1} = \frac{\Delta v}{\Delta t} \qquad (2.3)$$

In one dimensional motion, when the acceleration is in the same direction as the motion or velocity (normally taken to be in the positive direction), the acceleration is positive. But the acceleration may be in the opposite direction of the motion also. Then the acceleration is taken as negative and is often called deceleration or ***retardation***. So we can say that an increase in the rate of change of velocity is ***acceleration***, whereas the decrease in the rate of change of velocity is ***retardation***.

Example 2.4 : The velocity of a car moving towards the East increases from 0 to 12ms^{-1} in 3.0 s. Calculate its average acceleration.

Solution : Given,

$$v_1 = 0 \text{ m s}^{-1}$$

$$v_2 = 12 \text{ m s}^{-1}$$

$$t = 3.0 \text{ s}$$

$$a = \frac{(12.0\text{m s}^{-1})}{3.0\text{s}}$$

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

INTEXT QUESTIONS 2.1

1. Is it possible for a moving body to have non-zero average speed but zero average velocity during any given interval of time? If so, explain.
2. A lady drove to the market at a speed of 8 km h^{-1}. Finding market closed, she came back home at a speed of 10 km h^{-1}. If the market is 2km away from her home, calculate the average velocity and average speed.
3. Can a moving body have zero relative velocity with respect to another body? Give an example.
4. A person strolls inside a train with a velocity of 1.0 m s^{-1} in the direction of motion of the train. If the train is moving with a velocity of 3.0 m s^{-1}, calculate his

   (a) velocity as seen by passengers in the compartment, and (b) velocity with respect to a person sitting on the platform.

2.2 POSITION - TIME GRAPH

If you roll a ball on the ground, you will notice that at different times, the ball is found at different positions. The different positions and corresponding times can be plotted on a graph giving us a certain curve. Such a curve is known as position-time curve. Generally, the time is represented along x-axis whereas the position of the body is represented along y-axis.

Let us plot the position - time graph for a body at rest at a distance of 20m from the origin. What will be its position after 1s, 2s, 3s, 4s and 5s? You will find that the graph is a straight line parallel to the time axis, as shown in Fig. 2.1

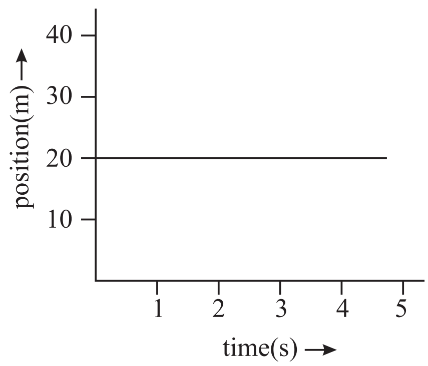


Fig. 2.1 : Position-time graph for a body at rest

2.2.1 Position-Time Graph for Uniform Motion

Now, let us consider a case where an object covers equal distances in equal intervals of time. For example, if the object covers a distance of 10m in each second for 5 seconds, the positions of the object at different times will be as shown in the following table.

Time (t) in s	1	2	3	4	5
Position (x) in m	10	20	30	40	50

In order to plot this data, take time along x-axis assuming 1cm as 1s, and position along y-axis with a scale of 1cm to be equal to 10m. The position-time graph will be as shown in Fig. 2.2

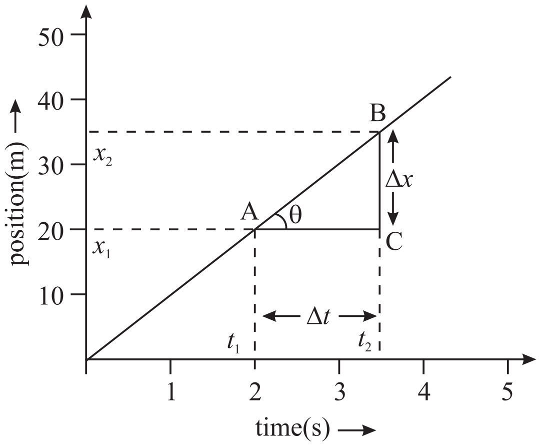


Fig. 2.2 : Position-time graph for uniform motion

The graph is a straight line inclined with the x-axis. ***A motion in which the velocity of the moving object is constant is known as uniform motion.*** Its position-time graph is a straight line inclined to the time axis.

In other words, we can say that when a moving object covers equal distances in equal intervals of time, it is in ***uniform motion.***

2.2.2 Position-Time Graph for Non-Uniform Motion

Let us now take an example of a train which starts from a station, speeds up and moves with uniform velocity for certain duration and then slows down before steaming in the next station. In this case you will find that the ***distances covered in equal intervals of time are not equal. Such a motion is said to be non-uniform motion.*** If the distances covered in successive intervals are increasing, the motion is said to be accelerated motion. The position-time graph for such an object is as shown in Fig.2.3.

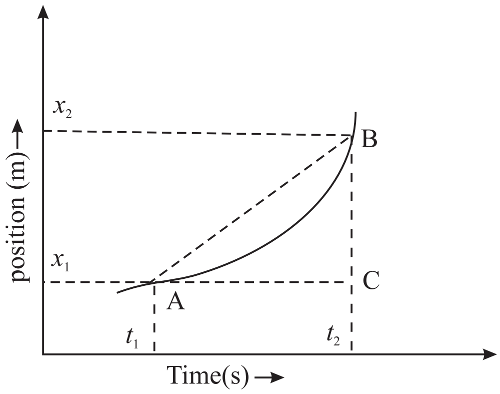


Fig. 2.3 : Position-time graph of accelerated motion as a continuous curve

Note that the position-time graph of accelerated motion is a continuous curve. Hence, the velocity of the body changes continuously. In such a situation, it is more appropriate to define average velocity of the body over an extremely small interval of time or instantaneous velocity. Let us learn to do so now.

2.2.3 Interpretation of Position - Time Graph

As you have seen, the position - time graphs of different moving objects can have different shapes. If it is a straight line parallel to the time axis, you can say that the body is at rest (Fig. 2.1). And the straight line inclined to the time axis shows that the motion is uniform (Fig.2.2). A continuous curve implies continuously changing velocity.

(a) Velocity from position - time graph : The slope of the straight line of position - time graph gives the average velocity of the object in motion. For determining the slope, we choose two widely separated points (say A and B) on the straight line (Fig.2.2) and form a triangle by drawing lines parallel to y-axis and x-axis. Thus, the average velocity of the object

$$\bar{\upsilon} = \frac{x_2 - x_1}{t_2 - t_1} = \frac{\Delta x}{\Delta t} = \frac{\mathrm{BC}}{\mathrm{AC}} \qquad (2.4)$$

Hence, average velocity of object equals the slope of the straight line AB.

It shows that greater the value of the slope ($\Delta x/\Delta t$) of the straight line position - time graph, more will be the average velocity. Notice that the slope is also equal to the tangent of the angle that the straight line makes with a horizontal line, i.e., $\tan\theta = \Delta x/\Delta t$. Any two corresponding Δx and Δt intervals can be used to determine the slope and thus the average velocity during that time internal.

Example 2.5 : The position - time graphs of two bodies A and B are shown in Fig. 2.4. Which of these has greater velocity?

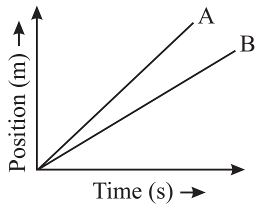


Fig. 2.4 : Position-time graph of bodies A and B

Solution : Body A has greater slope and hence greater velocity.

(b) Instantaneous velocity : As you have learnt, a body having uniform motion along a straight line has the same velocity at every instant. But in the case of non-uniform motion, the position - time graph is a curved line, as shown in Fig.2.5. As a result, the slope or the average velocity varies, depending on the size of the time intervals selected. The velocity of the particle at any instant of time or at some point of its path is called its instantaneous velocity.

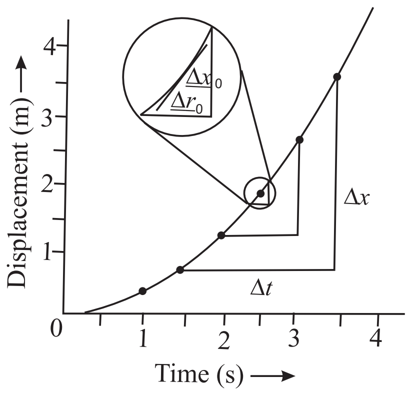


Fig. 2.5 : Displacement-time graph for non- uniform motion

Note that the average velocity over a time interval Δt is given by $\bar{\upsilon} = \frac{\Delta x}{\Delta t}$. As Δt is made smaller and smaller the average velocity approaches instantaneous velocity.

In the limit $\Delta t \to 0$, the slope $(\Delta x/\Delta t)$ of a line tangent to the curve at that point gives the instantaneous velocity. However, for uniform motion, the average and instantaneous velocities are the same.

Example 2.6 : The position - time graph for the motion of an object for 20 seconds is shown in Fig. 2.6. What distances and with what speeds does it travel in time intervals (i) 0 s to 5 s, (ii) 5 s to 10 s, (iii) 10 s to 15 s and (iv) 15 s to 17.5 s? Calculate the average speed for this total journey.

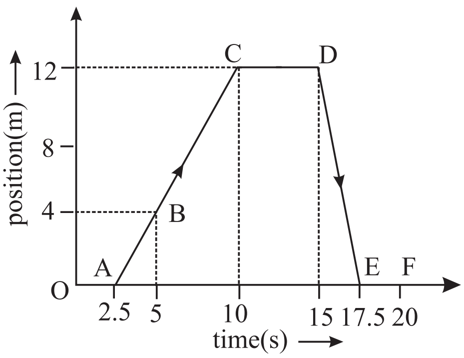


Fig. 2.6: Position-time graph

Solution :

(i) During 0 s to 5 s, distance travelled = 4 m

$$\therefore \text{ speed } = \frac{\text{Distance}}{\text{Time}} = \frac{4\text{ m}}{(5-0)\,s} = \frac{4\text{ m}}{5\,s} = 0.8\text{ m s}^{-1}$$

(ii) During 5 s to 10 s, distance travelled = 12 – 4 = 8 m

$$\therefore \text{ speed} = \frac{(12-4)\text{ m}}{(10-5)\text{ s}} = \frac{8\text{ m}}{5\text{ s}} = 1.6\text{ m s}^{-1}$$

(iii) During 10 s to 15 s, distance travelled = 12 – 12 = 0 m

$$\therefore \text{speed } = \frac{\text{Distance}}{\text{Time}} = \frac{0}{5} = 0$$

(iv) During 15 s to 17.5 s, distance travelled = 12 m

$$\therefore \text{ Speed} = \frac{12\text{ m}}{2.5\text{ s}} = 4.8\text{ m s}^{-1}$$

Now we would like you to pause for a while and solve the following questions to check your progress.

INTEXT QUESTIONS 2.2

1. Draw the position-time graph for a motion with zero acceleration.
2. The following figure shows the displacement - time graph for two students A and B who start from their school and reach their homes. Look at the graphs carefully and answer the following questions.

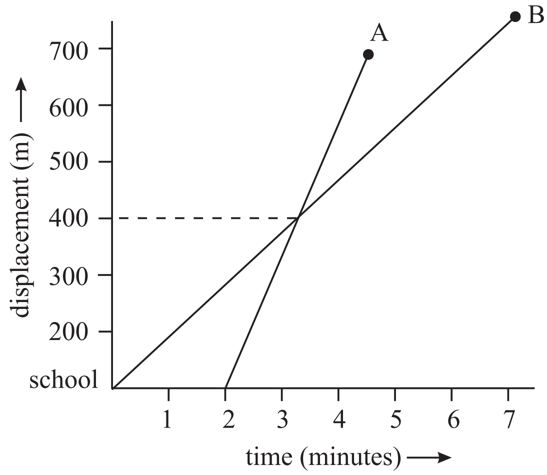
   (i) Do they both leave school at the same time?

   (ii) Who stays farther from the school?

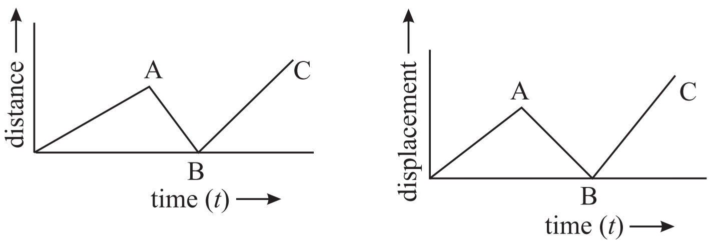
   (iii) Do they both reach their respective houses at the same time?

   (iv) Who moves faster?

   (v) At what distance from the school do they cross each other?



3. Under what conditions is average velocity of a body equal to its instantaneous velocity?
4. Which of the following graphs is not possible? Give reason for your answer?



(a) (b)

2.3 VELOCITY - TIME GRAPH

Just like the position-time graph, we can plot velocity-time graph. While plotting a velocity-time graph, generally the time is taken along the x-axis and the velocity along the y-axis.

2.3.1 Velocity-Time Graph for Uniform Motion

As you know, in uniform motion the velocity of the body remains constant, i.e., there is no change in the velocity with time. The velocity-time graph for such a

uniform motion is a straight line parallel to the time axis, as shown in the Fig. 2.7.

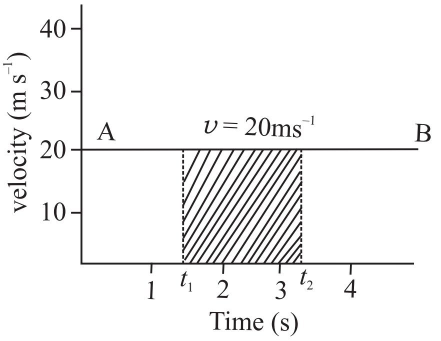


Fig. 2.7 : Velocity-time graph for uniform motion

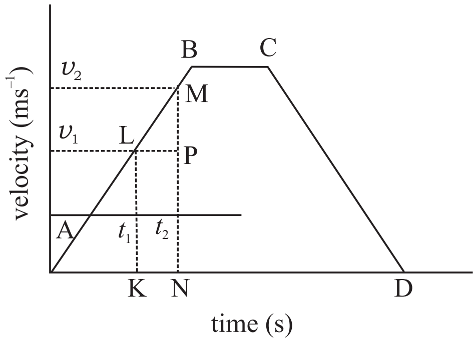


Fig. 2.8 : Velocity-time graph for motion with three different stages of constant acceleration

2.3.2 Velocity-Time Graph for Non-Uniform Motion

If the velocity of a body changes uniformly with time, its acceleration is constant. The velocity-time graph for such a motion is a straight line inclined to the time axis. This is shown in Fig. 2.8 by the straight line AB. It is clear from the graph that the velocity increases by equal amounts in equal intervals of time. The average acceleration of the body is given by

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

$$= \text{slope of the straight line}$$

Since the slope of the straight line is constant, the average acceleration of the body is constant. However, it is also possible that the rate of variation in the velocity is not constant. Such a motion is called non-uniformly accelerated motion. In such a situation, the slope of the velocity-time graph will vary at every instant, as shown in Fig.2.9. It can be seen that θ_A, θ_B and θ_C are different at points A, B and C.

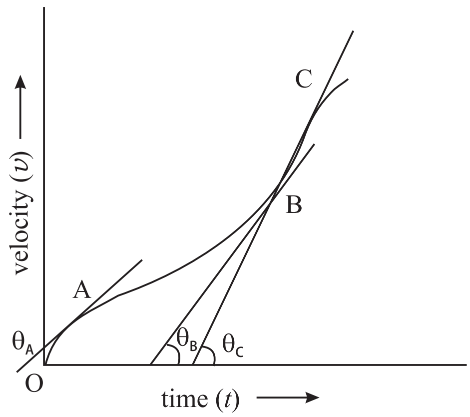


Fig. 2.9 : Velocity-time graph for a motion with varying acceleration

2.3.3 Interpretation of Velocity-Time Graph

Using $v - t$ graph of the motion of a body, we can determine the distance travelled by it and the acceleration of the body at different instants. Let us see how we can do so.

(a) Determination of the distance travelled by the body : Let us again consider the velocity-time graph shown in Fig. 2.10. The portion AB shows the motion with constant acceleration, whereas the portion CD shows the constantly retarded motion. The portion BC represents uniform motion (i.e., motion with zero acceleration).

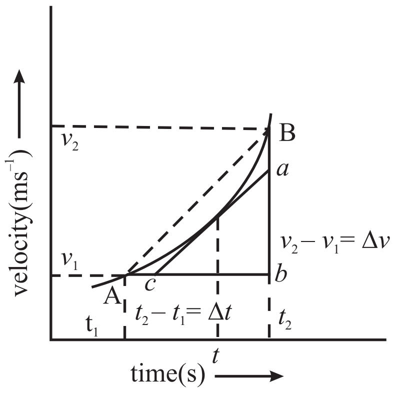


Fig. 2.10 : Velocity-time graph of non-uniformly accelerated motion

For uniform motion, the distance travelled by the body from time t_1 to t_2 is given by $s = \upsilon(t_2 - t_1)$ = the area under the curve between t_1 and t_2 Generalising this result for Fig. 2.10, we find that the distance travelled by the body between time t_1 and t_2

$$s = \text{area of trapezium KLMN}$$
$$= (½) \times (KL + MN) \times KN$$
$$= (½) \times (\upsilon_1 + \upsilon_2) \times (t_2 - t_1)$$

(b) Determination of the acceleration of the body : We know that acceleration of a body is the rate of change of its velocity with time. If you look at the velocity-time graph given in the Fig.2.10, you will note that the average acceleration is represented by the slope of the chord AB, which is given by

$$\text{average acceleration } (\bar{a}) = \frac{\Delta \upsilon}{\Delta t} = \frac{\upsilon_2 - \upsilon_1}{t_2 - t_1}.$$

If the time interval Δt is made smaller and smaller, the average acceleration becomes instantaneous acceleration. Thus, instantaneous acceleration

$$a = \lim_{\Delta t \to 0} \frac{\Delta \upsilon}{\Delta t} = \frac{d\upsilon}{dt} = \text{slope of the tangent at } (t = t) = \frac{ab}{bc}$$

Thus, the slope of the tangent at a point on the velocity-time graph gives the acceleration at that instant.

Example 2.7 : The velocity-time graphs for three different bodies A,B and C are shown in Fig. 2.11.

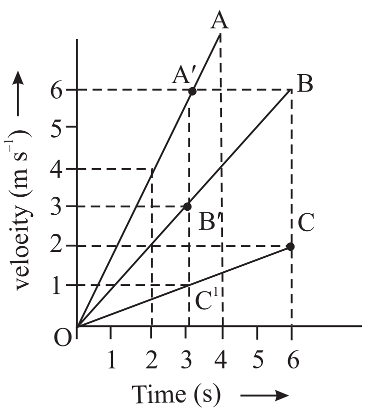


Fig. 2.11 : Velocity-time graph of uniformly accelerated motion of three different bodies

(i) Which body has the maximum acceleration and how much?

(ii) Calculate the distances travelled by these bodies in first 3s.

Notes

(iii) Which of these three bodies covers the maximum distance at the end of their journey?

(iv) What are the velocities at $t = 2$s?

Solution :

(i) As the slope of the υ-t graph for body A is maximum, its acceleration is maximum:

$$a = \frac{\Delta\upsilon}{\Delta t} = \frac{6-0}{3-0} = \frac{6}{3} = 2 \text{ ms}^{-2}.$$

(ii) The distance travelled by a body is equal to the area of the υ-t graph.

∴ In first 3s,

the distance travelled by A = Area OA′L

= (½) × 6 × 3 = 9m.

the distance travelled by B = Area OB′L

= (½) × 3 × 3 = 4.5 m.

the distance travelled by C = (½) × 1 × 3 = 1.5 m.

(iii) At the end of the journey, the maximum distance is travelled by B.

= (½) × 6 × 6 = 18 m.

(iv) Since υ-t graph for each body is a straight line, instantaneous acceleration is equal to average acceleration.

At 2s, the velocity of A = 4 m s^{-1}

the velocity of B = 2 m s^{-1}

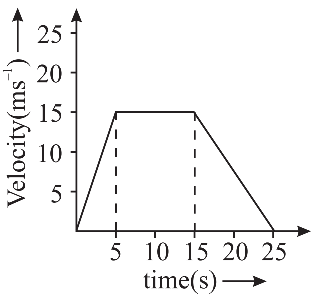
the velocity of C = 0.80 m s^{-1} (approx.)

INTEXT QUESTIONS 2.3

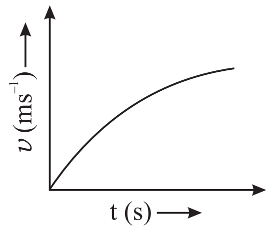
1. The motion of a particle moving in a straight line is depicted in the adjoining $v - t$ graph.

   (i) Describe the motion in terms of velocity, acceleration and distance travelled

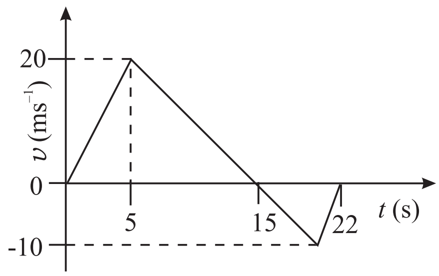
   (ii) Find the average speed.



2. What type of motion does the adjoining graph represent - uniform motion, accelerated motion or decelerated motion? Explain.



3. Using the adjoining υ-t graph, calculate the (i) average velocity, and (ii) average speed of the particle for the time interval 0 – 22 seconds. The particle is moving in a straight line all the time.



2.4 EQUATIONS OF MOTION

As you now know, for describing the motion of an object, we use physical quantities like distance, velocity and acceleration. In the case of constant acceleration, the velocity acquired and the distance travelled in a given time can be calculated by using one or more of three equations. These equations, generally known as equations of motion for constant acceleration or kinematical equations, are easy to use and find many applications.

2.4.1 Equation of Uniform Motion

In order to derive these equations, let us take initial time to be zero i.e. $t_1 = 0$. We can then assume $t_2 = t$ to be the elapsed time. The initial position (x_1) and initial velocity (υ_1) of an object will now be represented by x_0 and υ_0 and at time t they will be called x and υ (rather than x_2 and υ_2). According to Eqn. (2.1), the average velocity during the time t will be

$$\bar{\upsilon} = \frac{x - x_0}{t}. \qquad (2.4)$$

2.4.2 First Equation of Uniformly Accelerated Motion

The first equation of uniformly accelerated motion helps in determining the velocity of an object after a certain time when the acceleration is given. As you know, by definition

$$\text{Acceleration } (a) = \frac{\text{Change in velocity}}{\text{Time taken}} = \frac{\upsilon_2 - \upsilon_1}{t_2 - t_1}$$

If at $t_1 = 0$, $\upsilon_1 = \upsilon_0$ and at $t_2 = t$, $\upsilon_2 = \upsilon$. Then

$$a = \frac{\upsilon - \upsilon_0}{t} \qquad (2.5)$$

$$\Rightarrow \quad \upsilon = \upsilon_0 + at \qquad (2.6)$$

Example 2.8 : A car starting from rest has an acceleration of $10ms^{-2}$. How fast will it be going after 5s?

Solution : Given,

Initial velocity $\upsilon_0 = 0$

Acceleration $a = 10\ ms^{-2}$

Time $t = 5s$

Using first equation of motion

$$\upsilon = \upsilon_0 + at$$

we find that for $t = 5s$, the velocity is given by

$$\upsilon = 0 + (10\ ms^{-2}) \times (5s)$$

$$= 50\ ms^{-1}$$

2.4.3 Second Equation of Uniformly Accelerated Motion

Second equation of motion is used to calculate the position of an object after time t when it is undergoing constant acceleration a.

Suppose that at $t = 0$, $x_1 = x_0$; $\upsilon_1 = \upsilon_0$ and at $t = t$, $x_2 = x$; $\upsilon_2 = \upsilon$.

The distance travelled = area under $\upsilon - t$ graph

= Area of trapezium OABC

$$= \frac{1}{2}(CB + OA) \times OC$$

$$x - x_0 = ½\,(\upsilon + \upsilon_0)\,t$$

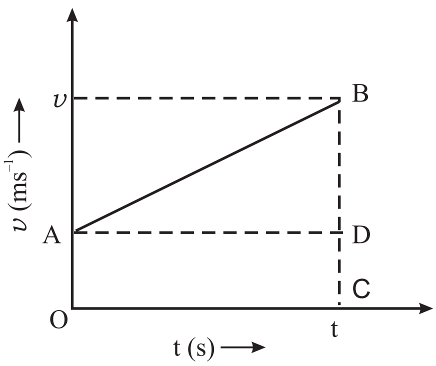


Fig. 2.12 : v–t graph for uniformly accelerated motion

Since $\upsilon = \upsilon_0 + at$, we can write

$$x - x_0 = ½\,(\upsilon_0 + at + \upsilon_0)t$$

$$= \upsilon_0 t + ½\,at^2$$

or $$x = x_0 + \upsilon_0 t + ½\,at^2 \qquad (2.7)$$

Example 2.9 : A car A is travelling on a straight road with a uniform speed of 60 km h^{-1}. Car B is following it with uniform velocity of 70 km h^{-1}. When the distance between them is 2.5 km, the car B is given a decceleration of 20 km h^{-1}. At what distance and time will the car B catch up with car A?

Solution : Suppose that car B catches up with car A at a distance x after time t.

For car A, the distance travelled in t time, $x = 60 \times t$.

For car B, the distance travelled in t time is given by

$$x' = x_0 + v_0 t + ½\, at^2$$

$$= 0 + 70 \times t + ½\, (–20) \times t^2$$

$$x' = 70\, t – 10\, t^2$$

But the distance between two cars is

$$x' – x = 2.5$$

$\therefore$ $$(70\, t – 10\, t^2) – (60\, t) = 2.5$$

or $$10\, t^2 – 10\, t + 2.5 = 0$$

It gives $t = ½$ hour

$\therefore$ $$x = 70t – 10t^2$$

$$= 70 \times ½ – 10 \times (½)^2$$

$$= 35 – 2.5 = 32.5 \text{ km.}$$

2.4.4 Third Equation of Uniformly Accelerated Motion

The third equation is used in a situation when the acceleration, position and initial velocity are known, and the final velocity is desired but the time t is not known. From Eqn. (2.7.), we can write

$$x – x_0 = ½\, (v + v_0)\, t.$$

Also from Eqn. (2.6), we recall that

$$t = \frac{v - v_0}{a}$$

Substituting this in above expression we get

$$x – x_0 = ½\, (v + v_0) \left(\frac{v - v_0}{a}\right)$$

$\Rightarrow$ $$2a\, (x – x_0) = v^2 – v^0$$

$\Rightarrow$ $$v^2 = v_0^2 + 2a\, (x – x_0) \qquad (2.8)$$

Thus, the three equations for constant acceleration are

$$v = v_0 + at$$

$$x = x_0 + v_0 t + ½\, at^2$$

and $$v^2 = v_0^2 + 2a\, (x – x_0)$$

Example 2.10 : A motorcyclist moves along a straight road with a constant acceleration of 4m s^{-2}. If initially she was at a position of 5m and had a velocity of 3m s^{-1}, calculate

(i) the position and velocity at time t = 2s, and

(ii) the position of the motorcyclist when its velocity is 5ms^{-1}.

Solution : We are given

$$x_0 = 5\text{m},\ \upsilon_0 = 3\text{m s}^{-1},\ a = 4\ \text{ms}^{-2}.$$

(i) Using Eqn. (2.7)

$$x = x_0 + \upsilon_0 t + \tfrac{1}{2} at^2$$

$$= 5 + 3 \times 2 + \tfrac{1}{2} \times 4 \times (2)^2 = 19\text{m}$$

From Eqn. (2.6)

$$\upsilon = \upsilon_0 + at$$

$$= 3 + 4 \times 2 = 11\text{ms}^{-1}$$

Velocity, $\upsilon = 11\text{ms}^{-1}$.

(ii) Using equation

$$\upsilon^2 = \upsilon_0^2 + 2a\,(x - x_0)$$

$$(5)^2 = (3)^2 + 2 \times 4 \times (x - 5)$$

$$\Rightarrow \quad x = 7\text{m}$$

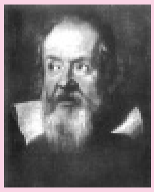
Hence position of the motor cyclist (x) = 7m.

2.5 MOTION UNDER GRAVITY

You must have noted that when we throw a body in the upward direction or drop a stone from a certain height, they come down to the earth. Do you know why they come to the earth and what type of path they follow? It happens because of the gravitational force of the earth on them. The gravitational force acts in the vertical direction. Therefore, motion under gravity is along a straight line. It is a one dimensional motion. ***The free fall of a body towards the earth is one of the most common examples of motion with constant acceleration.*** In the absence of air resistance, it is found that all bodies, irrespective of their size or weight, fall with the same acceleration. Though the acceleration due to gravity varies with altitude, for small distances compared to the earth's radius, it may be taken constant throughout the fall. For our practical use, the effect of air resistance is neglected.

The acceleration of a freely falling body due to gravity is denoted by g. At or near the earth's surface, its magnitude is approximately 9.8 ms^{-2}. More precise values, and its variation with height and latitude will be discussed in detail in lesson 5 of this book.

Galileo Galilei (1564 – 1642)



He was born at Pisa in Italy in 1564. He enunciated the laws of falling bodies. He devised a telescope and used it for astronomical observations. His major works are : Dialogues about the Two great Systems of the World and Conversations concerning Two New Sciences. He supported the idea that the earth revolves around the sun.

Example 2.11 : A stone is dropped from a height of 50m and it falls freely. Calculate the (i) distance travelled in 2 s, (ii) velocity of the stone when it reaches the ground, and (iii) velocity at 3 s i.e., 3 s after the start.

Solution : Given

Height $h = 50$ m and Initial velocity $\upsilon_0 = 0$

Consider, initial position (y_0) to be zero and the origin at the starting point. Thus, the y-axis (vertical axis) below it will be negative. Since acceleration is downward in the negative y-direction, the value of $a = -g = -9.8$ ms^{-2}.

(i) From Eqn. (2.7), we recall that

$$y = y_0 + \upsilon_0 t + ½\, at^2$$

For the given data, we get

$$y = 0 + 0 - ½\, gt^2 = -½ \times 9.8 \times (2)^2$$

$$= -19.6\text{m}.$$

The negative sign shows that the distance is below the starting point in downward direction.

(ii) At the ground $y = -50$m,

Using equation (2.8),

$$\upsilon^2 = \upsilon_0^2 + 2a\,(y - y_0)$$

$$= 0 + 2\,(-9.8)\,(-50 - 0)$$

$$\upsilon = 9.9 \text{ ms}^{-1}$$

(iii) Using $\upsilon = \upsilon_0 + at,$ at $t = 3$s, we get

$$\therefore \quad \upsilon = 0 + (-9.8) \times 3$$

$$\upsilon = -29.4 \text{ ms}^{-1}$$

This shows that the velocity of the stone at $t = 3$ s is 29.4 m s^{-1} and it is in downward direction.

Note : *It is important to mention here that in kinematic equations, we use certain sign convention according to which quantities directed upwards and rightwards are taken as positive and those downwards and leftward are taken as negative.*

2.6 CONCEPT OF DIFFERENTIATION AND INTEGRATION

All branches of Mathematics have been very useful tools in understanding and explaining the laws of Physics and finding the relations between different Physical quantities. You are already familiar with the use of Algebra and Trigonometry in this connection. In the further study of Physics, you will come across the use of Differentiation (or Differential Calculus) and Integration (or Integral Calculus). A brief and simple description of the concept of Differentiation and Integration is, therefore, being given below. You may consult books on Mathematics for more information on these topics.

We will often come across the following terms in this topic. Let us define these terms:

Constant: It is a quantity whose value does not change during mathematical operations, e.g. integers like 1, 2, 3, …., fractions, π, e, etc.

Variable: It is a quantity which can take different values during mathematical operations. A variable is generally denoted by x, y, z etc.

Function: 'y' is said to be a function of 'x', if for every value of 'x' there is definite value of 'y'.

Mathematically, it is represented by

$$y = f(x)$$

i.e. 'y' is a function of 'x'

Differential Coefficient: Of any variable 'y' with respect to any other variable 'x', is the instantaneous rate of change of 'y' with respect to 'x'.

Let 'y' be a function of 'x' i.e. $y = f(x)$. Suppose 'x' is increased by a very small amount δx or say there is a very small increament 'δx' in 'x'. Let there be a corresponding increament 'δy' in 'y'. Then, $y + \delta y$ is a function of $(x + \delta x)$

or $$y + \delta y = f(x + \delta x)$$

or $$\delta y = f(x + \delta x) - y$$

or $$\frac{\delta y}{\delta x} = \frac{f(x + \delta x) - f(x)}{\delta x}$$

The quantity $\frac{\delta y}{\delta x}$ is called increment ratio and represents the average rate of change of '*y*' with respect to '*x*' in the range between the time interval *x* and $(x + \delta x)$.

To find the instantaneous rate of change of '*y*' with respect to '*x*', we will have to calculate the limit of $\frac{\delta y}{\delta x}$ as δx tends to zero $(\delta x \to 0)$.

i.e. $$\underset{\delta x \to 0}{Lt} \frac{\delta y}{\delta x} = \underset{\delta x \to 0}{Lt} \frac{f(x+\delta x) - f(x)}{\delta x}$$

Thus, the instantaneous rate of change of '*y*' with respect to '*x*' is given by $\underset{\delta x \to 0}{Lt} \frac{\delta y}{\delta x}$. This is called the **differential coefficient** of '*y*' with respect to '*x*' and is denoted by $\frac{dy}{dx}$.

Integration

Integration is a mathematical process which is reverse of differentiation. In order to understand this concept, let a constant force **F** act on a body moving it through a distance S. Then, the work done by the force is calculated by the product W = **F**·**S**.

But, if the force is variable, ordinary algebra does not give any method to find the work done.

For example when a body is to be moved to a long distance up above the surface of the earth, the force of gravity on the body goes on changing as the body moves up. In such cases a method called integration is used to calculate the work done.

The work done by a variable force can be calculated as (see for details 6.2 work done by a variable force)

$$W = \Sigma F(x)\Delta x$$

For infinitesimally small values of Δx,

$$W = \sum_{\lim \Delta x \to 0} F(x)\,dx$$

This may be written as

$$W = \int F(x)\,dx$$

This expression is called integral of function $F(x)$ with respect to x, where the symbol '$\int$' denotes integration.

Some often used formulae of Integration and Differentiation

(i) $\int x^n dx = \frac{x^{n+1}}{n+1}$ (for n $\neq 1$)	(i) $\frac{d}{dx} x^n = nx^{n-1}$
(ii) $\int x^{-1} dx = \int \frac{1}{x} dx = \log x$	(ii) $\frac{d}{dx}(\log x) = \frac{1}{x}$
(iii) $\int dx = \int x^0 dx = \frac{x^1}{1} = x$	(iii) $\frac{d}{dx}(x) = 1$
(iv) $\int cx dx = c \int x dx$ (c is a constant)	(iv) $\frac{d}{dx}(cu) = c\frac{d}{dx}(u)$
(v) $\int (u+v+w) dx = \int u dx \pm \int v dx \pm \int w dx$	(v) $\frac{d}{dx}(u \pm v \pm w) = \frac{du}{dx} \pm \frac{dv}{dx} \pm \frac{dw}{dx}$
(vi) $\int e^x dx = e^x$	(vi) $\frac{d}{dx}(e^x) = e^x$
(vii) $\int \sin x dx = -\cos x$	(vii) $\frac{d}{dx}\{\sin(x)\} = +\cos x$
(viii) $\int \cos x dx = \sin x$	(viii) $\frac{d}{dx}(\cos x) = -\sin x$
(ix) $\int \sec^2 x dx = \tan x$	(ix) $\frac{d}{dx}(\tan x) = \sec x$
(x) $\int \text{cosec}^2 x dx = -\cot x$	(x) $\frac{d}{dx}(\cot x) = -\text{cosec}^2 x$

A close look at the table shows that Integration and differentiation are converse mathematical operations.

Take a pause and solve the following questions.

INTEXT QUESTIONS 2.4

1. A body starting from rest covers a distance of 40 m in 4s with constant acceleration along a straight line. Compute its final velocity and the time required to cover half of the total distance.

2. A car moves along a straight road with constant aceleration of 5 ms^{-2}. Initially at 5m, its velocity was 3 ms^{-1} Compute its position and velocity at $t = 2$ s.

3. With what velocity should a body be thrown vertically upward so that it reaches a height of 25 m? For how long will it be in the air?

4. A ball is thrown upward in the air. Is its acceleration greater while it is being thrown or after it is thrown?

WHAT YOU HAVE LEARNT

- The ratio of the displacement of an object to the time interval is known as average velocity.
- The total distance travelled divided by the time taken is average speed.
- The rate of change of the relative position of an object with respect to another object is known as the relative velocity of that object with respect to the other.
- The change in the velocity in unit time is called acceleration.
- The position-time graph for a body at rest is a straight line parallel to the time axis.
- The position-time graph for a uniform motion is a straight line inclined to the time axis.
- A body covering equal distance in equal intervals of time, however small, is said to be in uniform motion.
- The velocity of a particle at any one instant of time or at any one point of its path is called its instantaneous velocity.
- The slope of the position-time graph gives the average velocity.
- The velocity-time graph for a body moving with constant acceleration is a straight line inclined to the time axis.
- The area under the velocity-time graph gives the displacement of the body.
- The average acceleration of the body can be computed by the slope of velocity-time graph.
- The motion of a body can be described by following three equations :

  (i) $v = v_0 + at$

  (ii) $x = x_0 + v_0 t + \frac{1}{2} at^2$

  (iii) $v^2 = v_0^2 + 2a.(x - x_0)$

- Elementary ideas about concepts of differentiation and integration.

TERMINAL EXERCISE

1. Distinguish between average speed and average velocity.
2. A car C moving with a speed of 65 km h^{-1} on a straight road is ahead of motorcycle M moving with the speed of 80 km h^{-1} in the same direction. What is the velocity of M relative to A?
3. How long does a car take to travel 30m, if it accelerates from rest at a rate of 2.0 m s^2?
4. A motorcyclist covers half of the distance between two places at a speed of 30 km h^{-1} and the second half at the speed of 60 kmh^{-1}. Compute the average speed of the motorcycle.
5. A duck, flying directly south for the winter, flies with a constant velocity of 20 km h^{-1} to a distance of 25 km. How long does it take for the duck to fly this distance?
6. Bangalore is 1200km from New Delhi by air (straight line distance) and 1500 km by train. If it takes 2h by air and 20h by train, calculate the ratio of the average speeds.
7. A car accelerates along a straight road from rest to 50 kmh^{-1} in 5.0 s. What is the magnitude of its average acceleration?
8. A body with an initial velocity of 2.0 ms^{-1} is accelerated at 8.0 ms^{-2} for 3 seconds. (i) How far does the body travel during the period of acceleration? (ii) How far would the body travel if it were initially at rest?
9. A ball is released from rest from the top of a cliff. Taking the top of the cliff as the reference (zero) level and upwards as the positive direction, draw (i) the displacement-time graph, (ii) distance-time graph (iii) velocity-time graph, (iv) speed-time graph.
10. A ball thrown vertically upwards with a velocity v_0 from the top of the cliff of height h, falls to the beach below. Taking beach as the reference (zero) level, upward as the positive direction, draw the motion graphs. i.e., the graphs between (i) distance-time, (ii) velocity-time, (iii) displacement-time, (iv) speed - time graphs.
11. A body is thrown vertically upward, with a velocity of 10m/s. What will be the value of the velocity and acceleration of the body at the highest point?
12. Two objects of different masses, one of 10g and other of 100g are dropped from the same height. Will they reach the ground at the same time? Explain your answer.
13. What happens to the uniform motion of a body when it is given an acceleration at right angle to its motion?
14. What does the slope of velocity-time graph at any instant represent?

ANSWERS TO INTEXT QUESTIONS

2.1

1. Yes. When body returns to its initial postion its velocity is zero but speed is non-zero.
2. Average speed = $\dfrac{2+2}{\frac{2}{8}+\frac{2}{10}} = \dfrac{4}{9} \times 20 = 8.89$ km h^{-1}, average velocity = 0
3. Yes, two cars moving with same velocity in the same direction, will have zero relative velocity with respect to each other.
4. (a) 1 m s^{-1}
   (b) 2 m s^{-1}

2.2

1. See Fig. 2.2.
2. (i) A, (ii) B covers more distance, (iii) B, (iv) A, (v) When they are 3km from the starting point of B.
3. In the uniform motion.
4. (a) is wrong, because the distance covered cannot decrease with time or become zero.

2.3

1. (i) (a) The body starts with a zero velocity.
   (b) Motion of the body between start and 5th seconds is uniformly accelerated. It has been represented by the line OA.

$$a = \frac{15-0}{5-0} = 3 \text{ m s}^{-2}$$

   (c) Motion of the body between 5th and 10th second is a uniform motion (represented by AB). $a = \dfrac{15-15}{15-5} = \dfrac{0}{10} = 0$ m s^{-2}.

   (d) Motion between 15th and 25th second is uniformly retarded. (represented by the line BC). $a = \dfrac{0-15}{25-15} = -1.5$ m s^{-2}.

   (ii) (a) Average speed = $\dfrac{\text{Distance covered}}{\text{time taken}} = \dfrac{\text{Area of OA BC}}{(25-0)}$

$$= \frac{\left(\frac{1}{2}\times15\times5\right)+(15\times10)+\left(\frac{1}{2}\times15\times10\right)}{25} = \frac{525}{50} = 10.5 \text{ m s}^{-1}.$$

(b) Deccelerated Velocity decreases with time.

(c) Total distance covered $= \left(\frac{20\times15}{2}\right)\text{m} + \left(\frac{10\times7}{2}\right)\text{m} = 185$ m.

$\therefore$ average speed$= \left(\frac{185}{22}\right)\text{ms}^{-1} = 8.4$ ms^{-1}.

Total displacement $= \left(\frac{20\times15}{2}\right)\text{m} - \left(\frac{10\times7}{2}\right)\text{m} = 115$ m.

$\therefore$ average velocity $= \frac{115}{22}\text{ms}^{-1} = 5.22$ m s^{-1}.

2.4

1. Using $x = x_0 + \upsilon_0 t + ½\, at^2$

   $40 = \frac{1}{2} \times a \times 16$

   $\Rightarrow\ a = 5$ ms^{-2}

   Next using $\upsilon^2 = \upsilon_0^2 + 2a\,(x - x_0)$

   $\upsilon = 20$ m s^{-1},

   $20 = 0 + \frac{1}{2} \times 5 \times t^2 \ \Rightarrow\ t = 2\sqrt{2}$ s

2. Using Eqn.(2.9), $x = 21$m, and using Eqn.(2.6), $\upsilon = 13$ m s^{-1}.
3. At maximum height $\upsilon = 0$, using Eqn. (2.10), $\upsilon_0 = 7\sqrt{10}$ ms^{-1} = 22.6 m s^{-1}. The body will be in the air for the twice of the time it takes to reach the maximum height.
4. The acceleration of the ball is greater while it is thrown.

Answers to Terminal Exercises

2. 15 km h^{-1}
3. 5.47 s
4. 40 ms^{-1}
5. 1.25 h
6. 8 : 1
7. 2.8 m s^{-2} (or 3000 km h^{-2})
8. (i) 42 m (ii) 36 m
11. O and 9.8 m s^{-2}.

312en03

3

LAWS OF MOTION

In the previous lesson you learnt to describe the motion of an object in terms of its displacement, velocity and acceleration. But an important question is : what makes an object to move? Or what causes a ball rolling along the ground to come to a stop? From our everyday experience we know that we need to push or pull an object if we wish to change its position in a room. Similarly, a football has to be kicked in order to send it over a large distance. A cricket ball has to be hit hard by a batter to send it across the boundary for a six. You will agree that muscular activity is involved in all these actions and its effect is quite visible.

There are, however, many situations where the cause behind an action is not visible. For example, what makes rain drops to fall to the ground? What makes the earth to go around the sun? In this lesson you will learn the basic laws of motion and discover that force causes motion. The concept of force developed in this lesson will be useful in different branches of physics. Newton showed that force and motion are intimately connected. The laws of motion are fundamental and enable us to understand everyday phenomena.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain the significance of inertia;*
- *state Newton's laws of motion and illustrate them with examples;*
- *explain the law of conservation of momentum and illustrate it with examples;*
- *understand the concept of equilibrium of concurrent forces;*
- *define coefficient of friction and distinguish between static friction, kinetic friction and rolling friction;*

- *suggest different methods of reducing friction and highlight the role of friction in every-day life; and*
- *analyse a given situation and apply Newton's laws of motion using free body diagrams.*

3.1 CONCEPTS OF FORCE AND INERTIA

We all know that stationary objects remain wherever they are placed. These objects cannot move on their own from one place to another place unless forced to change their state of rest. Similarly, an object moving with constant velocity has to be forced to change its state of motion. **The property of an object by which it resists a change in its state of rest or of uniform motion in a straight line is called inertia. Mass of a body is a measure of its inertia.**

In a way, inertia is a fantastic property. If it were not present, your books or classnotes could mingle with those of your younger brother or sister. Your wardrobe could move to your friend's house creating chaos in life. You must however recall that the state of rest or of uniform motion of an object are not absolute. In the previous lesson you have learnt that an object at rest with respect to one observer may appear to be in motion with respect to some other observer. Observations show that **the change in velocity of an object can only be brought, if a net force acts on it.**

You are very familiar with the term force. We use it in so many situations in our everyday life. We are exerting force when we are pulling, pushing, kicking, hitting etc. Though a force is not visible, its effect can be seen or experienced. Forces are known to have different kinds of effects :

(a) **They may change the shape and the size of an object.** A balloon changes shape depending on the magnitude of force acting on it.

(b) **Forces also influence the motion of an object.** A force can set an object into motion or it can bring a moving object to rest. A force can also change the direction or speed of motion.

(c) **Forces can rotate a body about an axis.** You will learn about it in lesson seven.

3.1.1 Force and Motion

Force is a vector quantity. For this reason, when several forces act on a body simultaneously, a net equivalent force can be calculated by vector addition, as discussed in lesson 1.

Motion of a body is characterised by its displacement, velocity etc. We come across many situations where the velocity of an object is either continuously increasing or decreasing. For example, in the case of a body falling freely, the velocity of the body increases continuously, till it hits the ground. Similarly, in the case of a ball rolling on a horizontal surface, the velocity of the ball decreases continuously and ultimately becomes zero.

From experience we know that a net non-zero force is required to change the state of a body. ***For a body in motion, the velocity will change depending on the direction of the force acting on it.*** If a net force acts on a body in motion, its velocity will increase in magnitude, if the direction of the force and velocity are same. If the direction of net force acting on the body is opposite to the direction of motion, the magnitude of velocity will decrease. However, if a net force acts on a body in a direction perpendicular to its velocity, the magnitude of velocity of the body remains constant (see Sec 4.3). Such a force changes only the direction of velocity of the body. We may therfore conclude that ***velocity of a body changes as long as a net force is acting on it.***

3.1.2 First Law of Motion

When we roll a marble on a smooth floor, it stops after some time. It is obvious that its velocity decreases and ultimately it becomes zero. However, if we want it to move continuously with the same velocity, a force will have to be constantly applied on it.

We also see that in order to move a trolley at constant velocity, it has to be continuously pushed or pulled. Is there any net force acting on the marble or trolley in the situations mentioned here?

Motion and Inertia

Galileo carried out experiments to prove that in the absence of any external force, a body would continue to be in its state of rest or of uniform motion in a straight line. He observed that a body is accelerated while moving down an inclined plane (Fig. 3.1 a) and is retarded while moving up an inclined plane (Fig. 3.1 b). He argued that if the plane is neither inclined upwards nor downwards (i.e. if it is a horizontal plane surface), the motion of the body will neither be accelerated not retarded. That is, on a horizontal plane surface, a body will move with a uniform speed/velocity (if there is no external force).

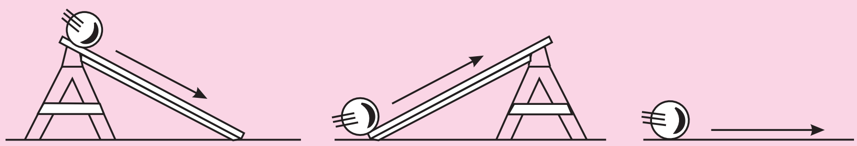


Fig. 3.1 : Motion of a body on inclined and horizontal planes

Notes

In another thought experiment, he considered two inclined planes facing each other, as shown in Fig. 3.2. The inclination of the plane PQ is same in all the three cases, whereas the inclination of the plane RS in Fig. 3.2 (a) is more than that in (b) and (c). The plane PQRS is very smooth and the ball is of marble. When the ball is allowed to roll down the plane PQ, it rises to nearly the same height on the face RS. As the inclination of the plane RS decreases, the balls moves a longer distance to rise to the same height on the inclined plane (Fig. 3.2b). When the plane RS becomes horizontal, the ball keeps moving to attain the same height as on the plane PQ, i.e. on a horizontal plane, the ball will keep moving if there is no friction between the plane and the ball.

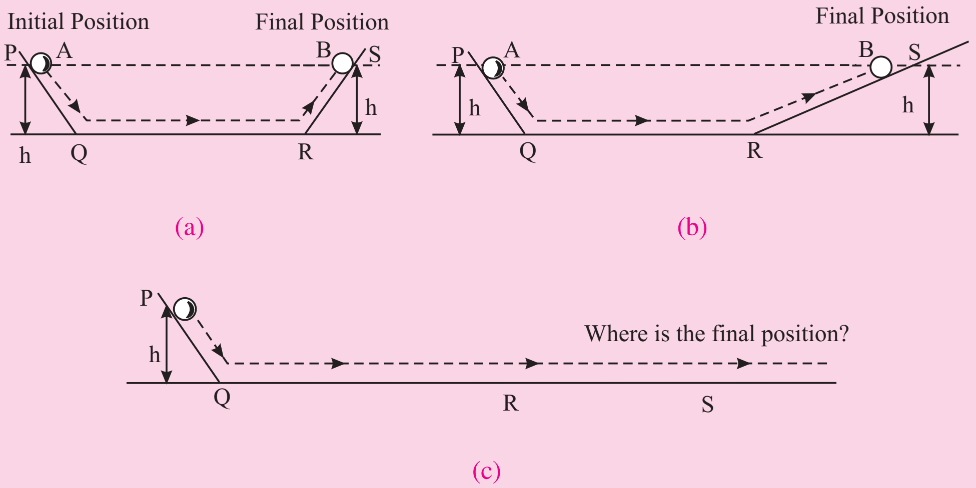


Fig. 3.2 : Motion of a ball along planes inclined to each other

Sir Issac Newton (1642–1727)

Newton was born at Wollsthorpe in England in 1642. He studied at Trinity College, Cambridge and became the most profound scientist. The observation of an apple falling towards the ground helped him to formulate the basic law of gravitation. He enunciated the laws of motion and the law of gravitation. Newton was a genius and contributed significantly in all fields of science, including mathematics. His contributions are of a classical nature and form the basis of the modern science. He wrote his book **"Principia"** in Latin and his book on optics was written in English.

You may logically ask : Why is it necessary to apply a force continuously to the trolley to keep it moving uniformly? We know that a forward force on the cart is needed for balancing out the force of friction on the cart. That is, the force of friction on the trolley can be overcome by continuously pushing or pulling it.

Isaac Newton generalised Galileo's conclusions in the form of a law known as ***Newton's first law of motion,*** which states that ***a body continues to be in a state of rest or of uniform motion in a straight line unless it is acted upon by a net external force.***

As you know, the state of rest or motion of a body depends on its relative position with respect to an observer. A person in a running car is at rest with respect to another person in the same car. But the same person is in motion with respect to a person standing on the road. For this reason, it is necessary to record measurements of changes in position, velocity, acceleration and force with respect to a chosen frame of reference.

A reference frame relative to which a body in translatory motion has constant velocity, if no net external force acts on it, is known as an ***inertial frame of reference.*** This nomenclature follows from the property of inertia of bodies due to which they tend to preserve their state (of rest or of uniform linear motion). A reference frame fixed to the earth (for all practical purposes) is considered an inertial frame of reference.

Now you may like to take a break and answer the following questions.

INTEXT QUESTIONS 3.1

1. Is it correct to state that a body always moves in the direction of the net external force acting on it?
2. What physical quantity is a measure of the inertia of a body?
3. Can a force change only the direction of velocity of an object keeping its magnitude constant?
4. State the different types of changes which a force can bring in a body when applied on it.

3.2 CONCEPT OF MOMENTUM

You must have seen that a fielder finds it difficult to stop a cricket ball moving with a large velocity although its mass is small. Similarly, it is difficult to stop a truck moving with a small velocity because its mass is large. These examples suggest that both, mass and velocity of a body, are important, when we study the effect of force on the motion of the body.

The product of mass m ***of a body and its velocity*** $\boldsymbol{v}$ ***is called its linear momentum*** **p. Mathmatically, we write**

$$\mathbf{p} = m\boldsymbol{v}$$

In SI units, momentum is measured in kg ms^{-1}. Momentum is a vector quantity. The direction of momentum vector is the same as the direction of velocity vector. Momentum of an object, therefore, can change on account of change in its magnitude or direction or both. The following examples illustrate this point.

Example 3.1: Aman weights 60 kg and travels with velocity 1.0 m s^{-1} towards Manoj who weights 40 kg, and is moving with 1.5 m s^{-1} towards Aman. Calculate their momenta.

Solution : For Aman

$$\begin{aligned} \text{momentum} &= \text{mass} \times \text{velocity} \\ &= (60 \text{ kg}) \times (1.0 \text{ m s}^{-1}) \\ &= 60 \text{ kg ms}^{-1} \end{aligned}$$

For Manoj

$$\begin{aligned} \text{momentum} &= 40 \text{ kg} \times (-1.5 \text{ ms}^{-1}) \\ &= -60 \text{ kg ms}^{-1} \end{aligned}$$

Note that the momenta of Aman and Manoj have the same magnitude but they are in opposite directions.

Example 3.2: A 2 kg object is allowed to fall freely at $t = 0$ s. Callculate its momentum at (a) $t = 0$, (b) $t = 1$ s and (c) $t = 2$ s during its free-fall.

Solution : (a) As velocity of the object at $t = 0$ s is zero, the initial momentum of the object will also be zero.

(b) At $t = 1$s, the velocity of the object will be 9.8 ms^{-1} [use $v = v_0 + at$] pointing downward. So the momentum of the object will be

$$p_1 = (2 \text{ kg}) \times (9.8 \text{ ms}^{-1}) = 19.6 \text{ kg ms}^{-1} \text{ pointing downward.}$$

(c) At $t = 2$ s, the velocity of the object will be 19.6 m s^{-1} pointing downward. So the momentum of the object will now be

$$p_2 = (2 \text{ kg}) \times (19.6 \text{ ms}^{-1}) = 39.2 \text{ kg ms}^{-1} \text{ pointing downward.}$$

Thus, we see that the momentum of a freely-falling body increases continuously in magnitude and points in the same direction. Now think what causes the momentum of a freely-falling body to change in magnitude?

Example 3.3: A rubber ball of mass 0.2 kg strikes a rigid wall with a speed of 10 ms^{-1} and rebounds along the original path with the same speed. Calculate the change in momentum of the ball.

Solution : Here the momentum of the ball has the same magnitude before and after the impact but there is a reversal in its direction. In each case the magnitude of momentum is (0.2 kg)×(10 ms^{-1}) i.e. 2 kg ms^{-1}.

If we choose initial momentum vector to be along + x axis, the final momentum vector will be along $-x$ axis. So p_i = 2 kg ms^{-1}, p_f = –2 kg ms^{-1}. Therefore, the change in momentum of the ball, $p_f - p_i$ = (–2 kgms^{-1}) – (2 kgms^{-1}) = – 4 kgms^{-1}.

Here negative sign shows that the momentum of the ball changes by 4 kg ms^{-1} in the direction of $-x$ axis. What causes this change in momentum of the ball?

In actual practice, a rubber ball rebounds from a rigid wall with a speed less than its speed before the impact. In such a case also, the magnitude of the momentum will change.

3.3 SECOND LAW OF MOTION

You now know that a body moving at constant velocity will have constant momentum. Newton's first law of motion suggests that no net external force acts on such a body.

In Example 3.2 we have seen that the momentum of a ball falling freely under gravity increases with time. Since such a body falls under the action of gravitational force acting on it, there appears to be a connection between change in momentum of an object, net force acting on it and the time for which it is acting. ***Newton's second law of motion*** gives a quantitative relation between these three physical quantities. It states that ***the rate of change of momentum of a body is directly proportional to the net force acting on the body. Change in momentum of the body takes place in the direction of net external force acting on the body.***

This means that if $\Delta\mathbf{p}$ is the change in momentum of a body in time Δt due to a net external force **F**, we can write

$$\mathbf{F} \propto \frac{\Delta\mathbf{p}}{\Delta t}$$

or

$$\mathbf{F} = k\,\frac{\Delta\mathbf{p}}{\Delta t}$$

where k is constant of proportionality.

By expressing momentum as a product of mass and velocity, we can rewrite this result as

$$\mathbf{F} = k\,m\left(\frac{\Delta\boldsymbol{v}}{\Delta t}\right)$$

$$\mathbf{F} = k\,m\,\mathbf{a} \qquad \left(\text{as } \frac{\Delta\boldsymbol{v}}{\Delta t} = \mathbf{a}\right) \tag{3.1}$$

The value of the constant k depends upon the units of m and **a**. If these units are chosen such that when the magnitude of m = 1 unit and a = 1 unit, the magnitude of **F** is also be 1 unit. Then, we can write

$$1 = k\,.\,1.\,1$$

i.e., $$k = 1$$

Using this result in Eqn. (3.1), we get

$$\mathbf{F} = m\,\mathbf{a} \tag{3.2}$$

In SI units, $m = 1$ kg, $a = 1$ m s^{-2}. Then magnitude of external force

$$\text{F} = 1\text{ kg} \times 1\text{ ms}^{-2} = 1\text{ kg ms}^{-2}$$

$$= 1 \text{ unit of force} \tag{3.3}$$

This unit of force (i.e., 1 kg m s^{-2}) is called one *newton.*

Note that the second law of motion gives us a unit for measuring force. The SI unit of force i.e., a newton may thus, be defined as the force which will produce an acceleration of 1 ms^{-2} in a mass of 1 kg.

Example 3.3: A ball of mass 0.4 kg starts rolling on the ground at 20 ms^{-1} and comes to a stop after 10s. Calculate the force which stops the ball, assuming it to be constant in magnitude throughout.

Solution : Given m= 0.4 kg, initial velocity u = 20 ms^{-1}, final velocity $\upsilon = 0$ m s^{-1} and t = 10s. So

$$|\mathbf{F}| = m|\mathbf{a}| = \frac{m(\upsilon - u)}{t} = \frac{0.4\text{ kg }(-20\text{ ms}^{-1})}{10\text{ s}}$$

$$= -0.8\text{ kg m s}^{-2} = -0.8\text{ N}$$

Here negative sign shows that force on the ball is in a direction opposite to that of its motion.

Example 3.4: A constant force of magnitude 50 N is applied to a body of 10 kg moving initially with a speed of 10 m s^{-1}. How long will it take the body to stop if the force acts in a direction opposite to its motion.

Solution : Given m = 10 kg, F = –50 N, υ_0 = 10 ms^{-1} and υ = 0. We have to calculate t. Since

$$\text{F} = ma$$

we can write

$$\text{F} = m\left(\frac{\upsilon - \upsilon_0}{t}\right)$$

$$\therefore \quad -50\text{ N} = 10\text{ kg}\left(\frac{0 - 10\text{ m s}^{-1}}{t}\right)$$

or $$t = \frac{-100 \text{ kgms}^{-1}}{-50 \text{ N}} = \frac{100 \text{ kg ms}^{-1}}{50 \text{ kg m s}^{-2}} = 2 \text{ s}.$$

It is important to note here that Newton's second law of motion, as stated here is applicable to bodies having constant mass. Will this law hold for bodies whose mass changes with time, as in a rocket?

INTEXT QUESTIONS 3.2

1. Two objects of different masses have the same momentum. Which of them is moving faster?
2. A boy throws up a ball with a velocity υ_0. If the ball returns to the thrower with the same velocity, will there be any change in
   (a) momentum of the ball?
   (b) magnitude of the momentum of the ball?
3. When a ball falls from a height, its momentum increases. What causes increase in its momentum?
4. In which case will there be larger change in momentum of the object?
   (a) A 150 N force acts for 0.1 s on a 2 kg object initially at rest.
   (b) A 150 N force acts for 0.2 s on a 2 kg. object initially at rest.
5. An object is moving at a constant speed in a circular path. Does the object have constant momentum? Give reason for your answer.

3.4 FORCES IN PAIRS

It is the gravitational pull of the earth, which allows an object to accelerate towards the earth. Does the object also pull the earth? Similarly when we push an almirah, does the almirah also push us? If so, why don't we move in the direction of that force? These situations compel us to ask whether a single force such as a push or a pull exists? It has been observed that actions of two bodies on each other are always mutual. Here, by action and reaction we mean 'forces of interaction'. So, whenever two bodies interact, they exert force on each other. One of them is called 'action' and the other is called 'reaction'. Thus, we can say that forces always exist in pairs.

3.4.1 Third Law of Motion

On the basis of his study of interactions between bodies, Newton formulated ***third law of motion: To every action, there is an equal and opposite reaction.***

Notes

Here by 'action' and 'reaction' we mean force. Thus, when a book placed on a table exerts some force on the table, the latter, also exerts a force of equal magnitude on the book in the upward direction, as shown in Fig. 3.3. Do the forces $\mathbf{F}_1$ and $\mathbf{F}_2$ shown here cancel out? It is important to note that $\mathbf{F}_1$ and $\mathbf{F}_2$ are acting on different bodies and therefore, they do not cancel out.

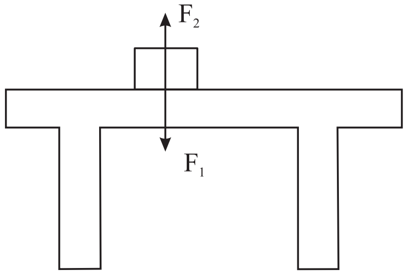


Fig 3.3 : A book placed on a table exerts a force F_1 (equal to its weight mg) on the table, while the table exerts a force F_2 on the book.

The action and reaction in a given situation appear as a pair of forces. Any one of them cannot exist without the other.

If one goes by the literal meaning of words, reaction always follows an action, whereas action and reaction introduced in Newton's third law exist simultaneously. For this reason, it is better to state Newton's third law as ***when two objects interact, the force exerted by one object on the other is equal in magnitude and opposite in direction to the force exerted by the latter object on the former.***

Vectorially, if $\mathbf{F}_{12}$ is the force which object 1 experiences due to object 2 and $\mathbf{F}_{21}$ is the force which object 2 experiences due to object 1, then according to Newton's third law of motion, we can write

$$\mathbf{F}_{12} = -\mathbf{F}_{21} \qquad (3.4)$$

3.4.2 Impulse

The effect of force applied for a short duration is called impulse. Impulse is defined as the product of force (**F**) and the time duration (Δt) for which the force is applied.

i.e., Impulse $= \mathbf{F}.\Delta t$

If the initial and final velocities of body acted upon by a force **F** are **u** and **v** respectively then we can write

$$\text{Impulse} = \frac{m\boldsymbol{v} - m\mathbf{u}}{\Delta t} \,.\, \Delta t$$

$$= m\boldsymbol{v} - m\,\boldsymbol{u}$$

$$= \mathbf{p}_f - \mathbf{p}_i$$

$$= \Delta\mathbf{p}$$

That is, impulse is equal to change in linear momentum.

Impulse in a vector quantity and its SI unit is kg ms^{-1} (or N s).

INTEXT QUESTIONS 3.3

1. When a high jumper leaves the ground, where does the force which throws the jumper upwards come from?
2. Identify the action - reaction forces in each of the following situations:
   (a) A man kicks a football
   (b) Earth pulls the moon
   (c) A ball hits a wall
3. "A person exerts a large force on an almirah to push it forward but he is not pushed backward because the almirah exerts a small force on him". Is the argument given here correct? Explain.

3.5 CONSERVATION OF MOMENTUM

It has been experimentally shown that if two bodies interact, the vector sum of their momenta remains unchanged, provided the force of mutual interaction is the only force acting on them. The same has been found to be true for more than two bodies interacting with each other. Generally, a number of bodies interacting with each other are said to be forming a system. If the bodies in a system do not interact with bodies outside the system, the system is said to be a closed system or an isolated system. ***In an isolated system***, ***the vector sum of the momenta of bodies remains constant.*** This is called the ***law of conservation of momentum.***

Here, it follows that it is the total momentum of the bodies in an isolated system remains unchanged but the momentum of individual bodies may change, in magnitude alone or direction alone or both. You may now logically ask : What causes the momentum of individual bodies in an isolated system to change? It is due to mutual interactions and their strengths.

Conservation of linear momentum is applicable in a wide range of phenomena such as collisions, explosions, nuclear reactions, radioactive decay etc.

3.5.1 Conservation of Momentum as a Consequence of Newton's Laws

According to Newton's second law of motion, Eqn. (3.1), the change in momentum $\Delta\mathbf{p}$ of a body, when a force $\mathbf{F}$ acts on it for time Δt, is

$$\Delta\mathbf{p} = \mathbf{F}\,\Delta t$$

This result implies that if no force acts on the body, the change in momentum of the body will be zero. That is, the momentum of the body will remain unchanged. This agrument can be extended to a system of bodies as well.

Notes

Newton's third law can also be used to arrive at the same result. Consider an isolated system of two bodies A and B which interact with each other for time Δt. If $\mathbf{F}_{AB}$ and $\mathbf{F}_{BA}$ are the forces which they exert on each other, then in accordance with Newton's third law

$$\mathbf{F}_{AB} = -\mathbf{F}_{BA}$$

or
$$\frac{\Delta \mathbf{p}_A}{\Delta t} = -\frac{\Delta \mathbf{p}_B}{\Delta t}$$

or
$$\Delta \mathbf{p}_A + \Delta \mathbf{p}_B = 0 \text{ or}$$

or
$$\Delta \mathbf{p}_{total} = 0$$

or
$$\mathbf{p}_{total} = \text{constant}$$

That is, there is no change in the momentum of the system. In other words, the momentum of the system is conserved.

3.5.2 A Few Illustrations of Conservation of Momentum

a) Recoil of a gun : When a bullet is fired from a gun, the gun recoils. The velocity v_2 of the recoil of the gun can be found by using the law of conservation of momentum. Let m be the mass of the bullet being fired from a gun of mass M. If v_1 is the velocity of the bullet, then momentum will be said to be conserved if the velocity v_2 of the gun is given by

$$m\boldsymbol{v}_1 + M\boldsymbol{v}_2 = 0$$

or
$$m\boldsymbol{v}_1 = -M\boldsymbol{v}_2$$

or
$$\boldsymbol{v}_2 = -\frac{m}{M}\boldsymbol{v}_1 \qquad (3.5)$$

Here, negative sign shows that $\boldsymbol{v}_2$ is in a direction opposite to $\boldsymbol{v}_1$. Since $m << M$, the recoil velocity of the gun will be considerably smaller than the velocity of the bullet.

b) Collision : In a collision, we may regard the colliding bodies as forming a system. In the absence of any external force on the colliding bodies, such as the force of friction, the system can be considered to be an isolated system. The forces of interaction between the colliding bodies will not change the total momentum of the colliding bodies.

Collision of the striker with a coin of carrom or collision between the billiared balls may be quite instructive for the study of collision between elastic bodies.

Example 3.5 : Two trolleys, each of mass m, coupled together are moving with initial velocity v. They collide with three identical stationary trolleys coupled together and continue moving in the same direction. What will be the velocity of the trolleys after the impact?

Solution : Let υ' be the velocity of the trolleys after the impact.

Momentum before collision = 2 $m\upsilon$

Momentum after collision = 5 $m\upsilon'$

In accordance with the law of conservation of momentum, we can write

$$2m\upsilon = 5\, m\upsilon'$$

or
$$\upsilon' = \frac{2}{5}\upsilon$$

c) Explosion of a bomb : A bomb explodes into fragments with the release of huge energy. Consider a bomb at rest initially which explodes into two fragments A and B. As the momentum of the bomb was zero before the explosion, the total momentum of the two fragments formed will also be zero after the explosion. For this reason, the two fragments will fly off in opposite directions. If the masses of the two fragments are equal, the velocities of the two fragments will also be equal in magnitude.

d) Rocket propulsion : Flight of a rocket is an important practical application of conservation of momentum. A rocket consists of a shell with a fuel tank, which can be considered as one body. The shell is provided with a nozzle through which high pressure gases are made to escape. On firing the rocket, the combustion of the fuel produces gases at very high pressure and temperature. Due to their high pressure, these gases escape from the nozzle at a high velocity and provide thrust to the rocket to go upward due to the conservation of momentum of the system. If M is the mass of the rocket and m is the mass of gas escaping per second with a velocity $\boldsymbol{\upsilon}$, the change in momentum of the gas in t second = $m\,\upsilon t$.

If the increase in velocity of the rocket in t second is **V**, the increase in its momentum = M**V**. According to the principle of conservation of momentum,

$$m\boldsymbol{\upsilon}t + M\mathbf{V} = 0$$

or
$$\frac{\mathbf{V}}{t} = \mathbf{a} = -\frac{m\boldsymbol{\upsilon}}{M}$$

i.e., the rocket moves with an acceleration

$$\mathbf{a} = -\frac{m\boldsymbol{\upsilon}}{M}$$

3.5.3 Equilibrium of Concurrent Forces

A number of forces acting simultaneously at a point are called **Concurrent Forces.** Such forces are said to be in equilibrium, if their resultant is zero.

Let $\mathbf{F}_1$, $\mathbf{F}_2$ and $\mathbf{F}_3$ be three concurrent forces acting at a point P, as shown in Fig. 3.4.

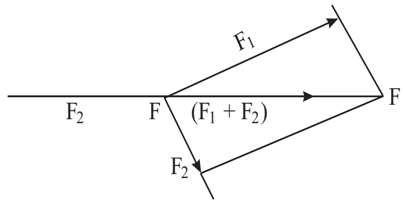


Fig. 3.4

The resultant of $\mathbf{F}_1$ and $\mathbf{F}_2$, obtained by the parallelogram law, is shown by **PA** (i.e. $\mathbf{PA} = \mathbf{F}_1 + \mathbf{F}_2$)

For equilibrium, the sum $(\mathbf{F}_1 + \mathbf{F}_2)$ must be equal and opposite to $\mathbf{F}_3$ i.e.

$$\mathbf{F}_3 = -(\mathbf{F}_1 + \mathbf{F}_2) \text{ or } \mathbf{F}_1 + \mathbf{F}_2 + \mathbf{F}_3 = 0$$

Or, the sum or resultant of two forces must be equal and opposite to the third force or for equilibrium, their vector sum must be zero.

3.6 FRICTION

You may have noticed that when a batsman hits a ball to make it roll along the ground, the ball does not continue to move forever. It comes to rest after travelling some distance. Thus, the momentum of the ball, which was imparted to it during initial push, tends to be zero. We know that some force acting on the ball is responsible for this change in its momentum. Such a force, called the ***frictional force***, exists whenever bodies in contact tend to move with respect to each other. It is the force of friction which has to be overcome when we push or pull a body horizontally along the floor to change its position.

Force of friction is a contact force and always acts along the surfaces in a direction opposite to that of the motion of the body. It is commonly known that friction is caused by roughness of the surfaces in contact. For this reason deliberate attempts are made to make the surfaces rough or smooth depending upon the requirement.

Friction opposes the motion of objects, causes wear and tear and is responsible for loss of mechanical energy. But then, it is only due to friction that we are able to walk, drive vehicles and stop moving vehicles. Friction thus plays a dual role in our lives. It is therefore said that friction is a necessary evil.

3.6.1 Static and Kinetic Friction

We all know that certain minimum force is required to move an object over a surface. To illustrate this point, let us consider a block resting on some horizontal surface, as shown in Fig.3.5. Let some external force $\mathbf{F}_{ext}$ be applied on the block. Initially the block does not move. This is possible only if some other force is acting on the block. The force is called the force of ***static friction*** and is represented by symbol f_s. As $\mathbf{F}_{ext}$ is increased, f_s also increases and remains equal to $\mathbf{F}_{ext}$ in magnitude until it reaches a critical value $f_s^{(max)}$. When $\mathbf{F}_{ext}$ is increased further, the block starts to slide and is then subject to ***kinetic friction.*** It is common experience that the force needed to set an object in motion is larger than the force needed to keep it moving at constant velocity. ***For this reason, the maximum value of static friction f_s between a pair of surfaces in contact will be larger than the force of kinetic friction f_k between them.*** Fig. 3.6 shows the variation of the force of friction with the external force.

For a given pair of surfaces in contact, you may like to know the ***factors*** on which $f_s^{(max)}$ and f_k depend? It is an experimental fact that $f_s^{(max)}$ is directly proportional to the normal force $\mathbf{F}_N$. i.e.

$$f_s^{(max)} \propto \mathbf{F}_N \quad \text{or} \quad f_s^{(max)} = \mu_s \mathbf{F}_N \qquad (3.6)$$

where μ_s is called the ***coefficient of static friction.*** The normal force $\mathbf{F}_N$ of the surface on the block can be found by knowing the force with which the block presses the surface. Refer to Fig. 3.5. The normal force $\mathbf{F}_N$ on the block will be *mg,* where *m* is mass of the block.

Since $\quad f_s = \mathbf{F}_{ext}$ for $f_s \le f_s^{max}$, we can write

$$f_s \le \mu_s \mathbf{F}_N.$$

It has also been experimentally found that ***maximum force of static friction between a pair of surfaces is independent of the area of contact.***

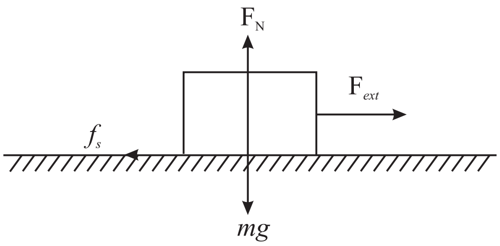


Fig. 3.5 : Forces acting on the block

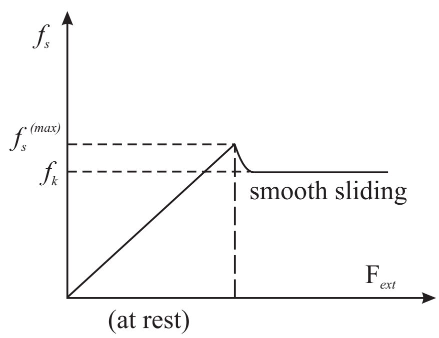


Fig. 3.6 : Variation of force of friction with external force

Similarly, we can write

$$f_k = \mu_k \mathbf{F}_N$$

where μ_k is the ***coefficient of kinetic friction***. In general, $\mu_s > \mu_k$. Moreover, coefficients μ_s and μ_k are not really constants for any pair of surfaces such as wood on wood or rubber on concrete, etc. ***Values of μ_s and μ_k for a given pair of materials depend on the roughness of surfaces, there cleanliness, temperature, humidity etc.***

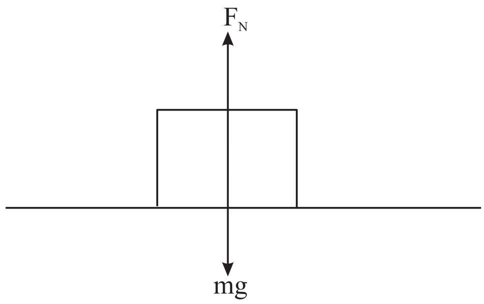


Fig. 3.7 : Normal force on the block

Example 3.6: A 2 kg block is resting on a horizontal surface. The coefficient of static friction between the surfaces in contact is 0.25. Calculate the maximum magnitude of force of static friction between the surfaces in contact.

Solution :

Here m = 2 kg and μ_s = 0.25. From Eqn. (3.6), we recall that

$$f_s^{(max)} = \mu_s \mathbf{F}_N$$
$$= \mu_s\, m\mathbf{g}$$
$$= (0.25)\ (2\text{ kg})\ (9.8\text{ ms}^{-2})$$
$$= 4.9\text{ N.}$$

Example 3.7: A 5 kg block is resting on a horizontal surface for which μ_k = 0.1. What will be the acceleration of the block if it is pulled by a 10 N force acting on it in the horizontal direction?

Solution :

As $f_k = \mu_k F_N$ and $F_N = mg$, we can write

$$f_k = \mu_k mg$$
$$= (0.1)\ (5\text{ kg})\ (9.8\text{ ms}^{-2})$$
$$= 4.9\text{ kg ms}^{-2} = 4.9\text{ N}$$

Net force on the block = $F_{ext} - f_k$

$$= 10\text{ N} - 4.9\text{ N}$$
$$= 5.1\text{ N}$$

Hence,

$$\text{acceleration} = a = \frac{F_{net}}{m} = \frac{5.1\text{ N}}{5\text{kg}} = 1.02\text{ ms}^{-2}$$

So the block will have an acceleration of 1.02 ms^{-2} in the direction of externally applied force.

3.6.2 Rolling Friction

It is a common experience that pushing or pulling objects such as carts on wheels is much easier. The motion of a wheel is different from sliding motion. It is a rolling motion. The friction in the case of rolling motion is known as ***rolling friction***. For the same normal force, rolling friction is much smaller than sliding friction. For example, when steel wheels roll over steel rails, rolling friction is about 1/100th of the sliding friction between steel and steel. Typical values for coefficient of rolling friction μ_r are 0.006 for steel on steel and 0.02 – 0.04 for rubber on concrete.

We would now like you to do a simple activity :

ACTIVITY 3.1

Place a heavy book or a pile of books on a table and try to push them with your fingers. Next put three or more pencils below the books and now push them again. In which case do you need less force? What do you conclude from your experience?

3.6.3 Methods of Reducing Friction

Wheel is considered to be greatest invention of mankind for the simple reason that rolling is much easier than sliding. Because of this, ball bearings are used in machines to reduce friction. In a ball-bearing, steel balls are placed between two co-axial cylinders, as shown in Fig.3.8. Generally one of the two cylinders is allowed to turn with respect to the other. Here the rotation of the balls is almost frictionless. Ball-bearings find application in almost all types of vehicles and in electric motors such as electric fans etc.

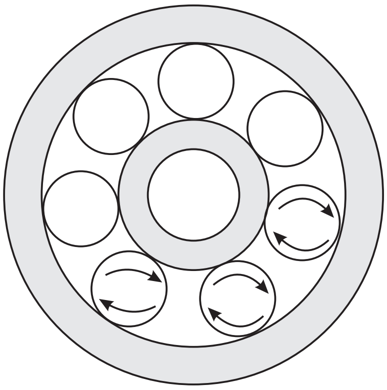


Fig. 3.8 : Balls in the ball-bearing

Use of ***lubricants*** such as grease or oil between the surfaces in contact reduces friction considerably. In heavy machines, oil is made to flow over moving parts. It reduces frictional force between moving parts and also prevents them from getting overheated. In fact, the presence of lubricants changes the nature of friction from dry friction to fluid friction, which is considerably smaller than the former.

Flow of compressed and purified air between the surfaces in contact also reduces friction. It also prevents dust and dirt from getting collected on the moving parts.

Fluid Friction

Bodies moving on or through a liquid or gas also face friction. Shooting stars (meteors) shine because of the heat generated by air-friction. Contrary to solid friction, fluid friction depends upon the shape of the bodies. This is why fishes have a special shape and fast moving aeroplanes and vehicles are also given a fish-like shape, called a stream-line shape. Fluid friction increases rapidly with increase in speed. If a car is run at a high speed, more fuel will have to be burnt to overcome the increased fluid (air) friction. Car manufactures advise us to drive at a speed of 40-45 km h^{-1} for maximum efficiency.

3.7 THE FREE BODY DIAGRAM TECHNIQUE

Application of Newton's laws to solve problems in mechanics becomes easier by use of the **free body diagram technique**. A diagram which shows all the forces acting on a body in a given situation is called a free body diagram (FBD). The procedure to draw a free body diagram, is described below :

1. Draw a simple, neat diagram of the system as per the given description.
2. Isolate the ***object of interest.*** This object will be called the ***Free Body*** now.
3. Consider all ***external forces*** acting on the free body and mark them by arrows touching the free body with their line of action clearly represented.
4. Now apply Newton's second law $\Sigma\mathbf{F} = m\,\mathbf{a}$

   (or $\Sigma F_x = m\,a_x$ and $\Sigma F_y = m\,a_y$)

Remember : (i) A net force must be acting on the object along the direction of motion. (ii) For obtaining a complete solution, you must have as many independent equations as the number of unknowns.

Example 3.8 : Two blocks of masses m_1 and m_2 are connected by a string and placed on a smooth horizontal surface. The block of mass m_2 is pulled by a force **F** acting parallel to the horizontal surface. What will be the acceleration of the blocks and the tension in the string connecting the two blocks (assuming it to be horizontal)?

Solution : Refer to Fig. 3.9. Let **a** be the acceleration of the blocks in the direction of **F** and let the tension in the string be **T**. On applying $\Sigma\mathbf{F} = m\mathbf{a}$ in the component form to the free body diagram of system of two bodies of masses m_1 and m_2, we get

$$N - (m_1 + m_2)\,g = 0$$

and

$$F = (m_1 + m_2)a$$

$$\Rightarrow \quad a = \frac{F}{m_1 + m_2}$$

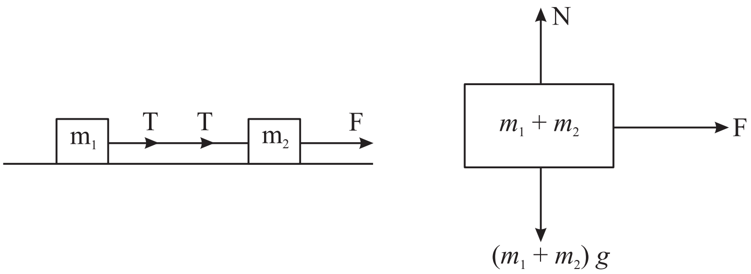


Fig 3.9: Free body diagram for two blocks connected by a string

On applying $\Sigma F = ma$ in the component form to the free body diagram of m_1 we get

$$N_1 - m_1 g = 0 \quad \text{and} \quad T = m_1 a$$

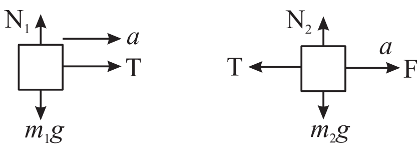


Fig 3.10

$$\Rightarrow \qquad T = m_1\left(\frac{F}{m_1 + m_2}\right)$$

or
$$T = \left(\frac{m_1}{m_1 + m_2}\right).F$$

Apply $\Sigma F = ma$ once again to the free body diagram of m_2 and see whether you get the same expressions for a and T.

Example 3.9 : Two masses m_1 and m_2 $(m_1 > m_2)$ are connected at the two ends of a light inextensible string that passes over a light frictionless fixed pulley. Find the acceleration of the masses and the tension in the string connecting them when the masses are released.

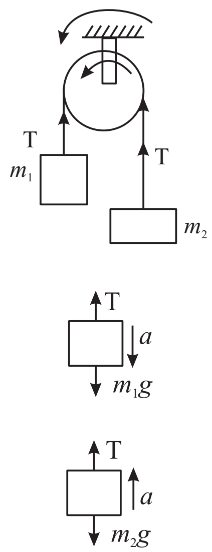


Fig 3.11

Solution : Let a be acceleration of mass m_1 downward. The acceleration of mass m_2 will also be a only but upward. (Why?). Let T be the tension in the string connecting the two masses.

On applying $\Sigma F = ma$ to m_1 and m_2 we get

$$m_1 g - T = m_1 a$$

$$T - m_2 g = m_2 a$$

On solving equations (1) and (2) for a and T we get

$$a = \left(\frac{m_1 - m_2}{m_1 + m_2}\right).g \qquad T = \left(\frac{2m_1 m_2}{m_1 + m_2}\right)a$$

At this stage you can check the prediction of the results thus obtained for the extreme values of the variables (i.e. m_1 and m_2). Either take $m_1 = m_2$ or $m_1 >> m_2$ and see whether a and T take values as expected.

Example 3.10 : A trolley of mass $M = 10$ kg is connected to a block of mass $m = 2$ kg with the help of massless inextensible string passing over a light frictionless pulley as shown in Fig. 3.12 (a). The coefficient of kinetic friction between the trolley and the surface $(\mu_k) = 0.02$. Find,

a) acceleration of the trolley, and

b) tension in the string.

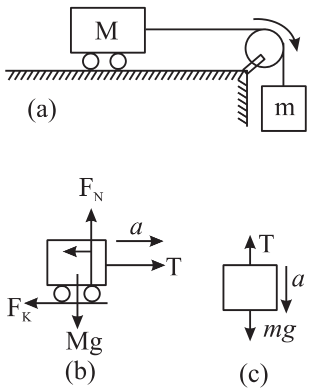


Fig. 3.12

Solution : Fig (b) and (c) shows the free body diagrams of the trolley and the block respectively. Let a be the acceleration of the block and the trolley.

For the trolley, $F_N = Mg$ and

$T - f_k = Ma$ where $f_k = \mu_k F_N$

$= \mu_k Mg$

So $\quad T - \mu_k Mg = Ma \quad$ ...(1)

For the block $\quad mg - T = ma \quad$ ...(2)

On adding equations (1) and (2) we get $mg - \mu_k Mg = (M + m)\,a$

or $$a = \frac{mg - \mu_k Mg}{M + m} = \frac{(2\text{ kg})(9.8\text{ ms}^{-2}) - (0.02)(10\text{ kg})(9.8\text{ ms}^{-2})}{(10\text{ kg} + 2\text{ kg})}$$

$$= \frac{19.6\text{ kg ms}^{-2} - 1.96\text{ kg ms}^{-2}}{12\text{ kg}} = 1.47\text{ ms}^{-2}$$

So $\quad a = 1.47\text{ ms}^{-2}$

From equation (2) $T = mg - ma = m(g - a)$

$= 2\text{ kg }(9.8\text{ ms}^{-2} - 1.47\text{ ms}^{-2})$

$= 2\text{ kg }(8.33\text{ ms}^{-2})$

So $\quad T = 16.66$ N

INTEXT QUESTIONS 3.4

1. A block of mass m is held on a rough inclined surface of inclination θ. Show in a diagram, various forces acting on the block.

2. A force of 100 N acts on two blocks A and B of masses 2 kg and 3 kg respectively, placed in contact on a smooth horizontal surface as shown. What is the magnitude of force which block A exerts on block B?

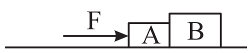


Fig. 3.13

3. What will be the tension in the string when a 5kg object suspended from it is pulled up with

   (a) a velocity of $2ms^{-1}$?

   (b) an acceleration of $2ms^{-2}$?

3.8 ELEMENTARY IDEAS OF INERTIAL AND NON INERTIAL FRAMES

To study motion in one dimension (i.e. in a straight line) a reference point (origin) is enough. But, when it comes to motions in two and three dimensions, we have to use a set of reference lines to specify the position of a point in space. This set of lines is called **frame of reference.**

Every motion is described by an observer. The description of motion will change with the change in the state of motion of the observer. For example, let us consider a box lying on a railway platform. A person standing on the platform will say that the box is at rest. A person in a train moving with a uniform velocity υ will say that the box is moving with velocity $-\upsilon$. But, what will be the description of the box by a person in a train having acceleration (a). He/she will find that the box is moving with an acceleration ($-a$). Obviously, the first law of motion is failing for this observer.

Thus a frame of reference is fixed with the observer to describe motion. If the frame is stationary or moving with a constant velocity with respect to the object under study (another frame of reference), then in this frame law of inertia holds good. Therefore, such frames are called inertial frames. On the other hand, if the observer's frame is accelerating, then we call it non-inertial frame.

For the motion of a body of mass m in a non-inertial frame, having acceleration (a), we may apply second law of motion by involving a psuedo force $m\,a$. In a rotating body, this force is called centrifugal force.

INTEXT QUESTIONS 3.5

1. A glass half filled with water is kept on a horizontal table in a train. Will the free surface of water remain horizontal as the train starts?

2. When a car is driven too fast around a curve it skids outwards. How would a passenger sitting inside explain the car's motion? How would an observer standing on a road explain the event?

3. A tiny particle of mass 6×10^{-10} kg is in a water suspension in a centrifuge which is being rotated at an angular speed of $2\pi \times 10^{3}$ rad s^{-1}. The particle is

at a distance of 4 cm from the axis of rotation. Calculate the net centrifugal force acting on the particle.

4. What must the angular speed of the rotation of earth so that the centrifugal force makes objects fly off its surface? Take $g = 10 \text{ m s}^{-2}$.

5. In the reference frame attached to a freely falling body of mass 2 kg, what is the magnitude and direction of inertial force on the body?

WHAT YOU HAVE LEARNT

- The ***inertia*** of a body is its tendency to resist any change in its state of rest or uniform motion.
- ***Newton's first law*** states that a body remains in a state of rest or of uniform motion in a straight line as long as net external force acting on it is zero.
- For a single particle of mass m moving with velocity $\boldsymbol{v}$ we define a vector quantity $\mathbf{p}$ called the linear momentum as $\mathbf{p} = m\,\boldsymbol{v}$.
- ***Newton's second law*** states that the time rate of change of momentum of a body is proportional to the resultant force acting on the body.
- According to Newton's second law, acceleration produced in a body of constant mass is directly proportional to net external force acting on the body : $\mathbf{F} = m\,\mathbf{a}$.
- ***Newton's third law*** states that if two bodies A and B interact with each other, then the force which body A exerts on body B will be equal and opposite to the force which body B exerts on body A.
- According to the law of conservation of momentum, if no net external force acts on a system of particles, the total momentum of the system will remain constant regardless of the nature of forces between them.
- A number of forces acting simultaneously at a point are called concurrent force. Such forces are said to be in equilibrium if their resultant is zero.
- Frictional force is the force which acts on a body when it attempts to slide, or roll along a surface. The force of friction is always parallel to the surfaces in contact and opposite to the direction of motion of the object.
- The maximum force of static friction $f_s^{(max)}$ between a body and a surface is proportional to the normal force $\mathbf{F}_N$ acting on the body. This maximum force occurs when the body is on the verge of sliding.
- For a body sliding on some surface, the magnitude of the force of kinetic friction f_k is given by $f_k = \mu_k \mathbf{F}_N$ where μ_k is the coefficient of kinetic friction for the surfaces in contact.

- Use of rollers and ball-bearings reduces friction and associated energy losses considerably as rolling friction is much smaller than kinetic friction.
- Newton's laws of motion are applicable only in an inertial frame of reference. An inertial frame is one in which an isolated object has zero acceleration.
- For an object to be in static equilibrium, the vector sum of all the forces acting on it must be zero. This is a necessary and sufficient conditions for point objects only.

TERMINAL EXERCISE

1. Which of the following will always be in the direction of net external force acting on the body?

   (a) displacement (b) velocity

   (c) acceleration (d) Change is momentum.

2. When a constant net external force acts on an object, which of the following may not change?

   (a) position (b) speed

   (c) velocity (d) acceleration

   Justify your answer with an example each.

3. A 0.5 kg ball is dropped from such a height that it takes 4s to reach the ground. Calculate the change in momentum of the ball.

4. In which case will there be larger change in momentum of a 2 kg object:

   (a) When 10 N force acts on it for 1s ?

   (b) When 10 N force acts on it for 1m ?

   Calculate change in momentum in each case.

5. A ball of mass 0.2 kg falls through air with an acceleration of 6 ms^{-2}. Calculate the air drag on the ball.

6. A load of mass 20 kg is lifted with the help of a rope at a constant acceleration. The load covers a height of 5 m in 2 seconds. Calculate the tension in the rope. In a rocket m changes with time. Write down the mathmatical form of Newton's law in this case and interpret it physically.

7. A ball of mass 0.1 kg moving at 10 m s^{-1} is deflected by a wall at the same speed in the direction shown. What is the magnitude of the change in momentum of the ball?

Notes

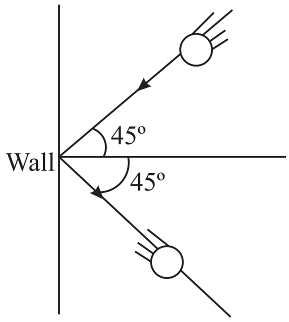


Fig. 3.14

8. Find the average recoil force on a machine gun that is firing 150 bullets per minute, each with a speed of 900 m s^{-1}. Mass of each bullet is 12 g.

9. Explain why, when catching a fast moving ball, the hands are drawn back while the ball is being brought to rest.

10. A constant force of magnitude 20 N acts on a body of mass 2 kg, initially at rest, for 2 seconds. What will be the velocity of the body after
(a) 1 second from start? (b) 3 seconds from start?

11. How does a force acting on a block in the direction shown here keep the block from sliding down the vertical wall?

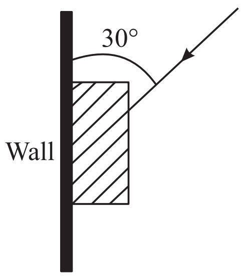


Fig 3.15

12. A 1.2 kg block is resting on a horizontal surface. The coefficient of static friction between the block and the surface is 0.5. What will be the magnitude and direction of the force of friction on the block when the magnitude of the external force acting on the block in the horizontal direction is

(a) 0 N ? (b) 4.9 N ? (c) 9.8 N ?

13. For a block on a surface the maximum force of static friction is 10N. What will be the force of friction on the block when a 5 N external force is applied to it parallel to the surface on which it is resting?

14. What minimum force F is required to keep a 5 kg block at rest on an inclined plane of inclination 30^0. The coefficient of static friction between the block and the inclined plane is 0.25.

Notes

15. Two blocks P and Q of masses m_1 = 2 kg and m_2 = 3 kg respectively are placed in contact with each other on horizontal frictionless surface. Some external force F = 10 N is applied to the block P in the direction parallel to the surface. Find the following
    (a) acceleration of the blocks
    (b) force which the block P exerts on block Q.
16. Two blocks P and Q of masses m_1 = 2 kg and m_2 = 4 kg are connected to a third block R of mass M as shown in Fig. 3.16 For what maximum value of M will the system be in equilibrium? The frictional force acting on each block is half the force of normal reaction on it.

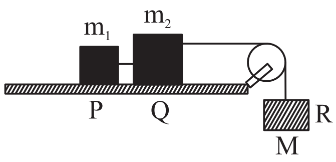


Fig. 3.16

17. Explain the role of friction in the case of bicycle brakes. What will happen if a few drops of oil are put on the rim?
18. A 2 kg block is pushed up an incline plane of inclination $\theta = 37^0$ imparting it a speed of 20 m s^{-1}. How much distance will the block travel before coming to rest? The coefficient of kinetic friction between the block and the incline plane is $\mu_k = 0.5$.

    Take g = 10 m s^{-2} and use sin 37^0 = 0.6, cos 37^0 = 0.8.

ANSWERS TO INTEXT QUESTIONS

3.1

1. No. The statement is true only for a body which was at rest before the application of force.
2. Inertial mass
3. Yes, as in uniform circular motion.
4. A force can change motion. It can also deform bodies.

3.2

1. Object of smaller mass
2. (a) Yes (b) No.
3. Momentum of the falling ball increases because gravitational force acts on it in the direction of its motion and hence velocity increases.

4. In case (b) the change in momentum will be larger. It is the F Δt product that gives the change in momentum. $\left(\text{as } F \propto \frac{\Delta p}{\Delta t}\right)$

5. No. Though the speed is constant, the velocity of the object changes due to change in direction. Hence its momentum will not be constant.

3.3

1. The jumper is thrown upwards by the force which the ground exerts on the jumper. This force is the reaction to the force which the jumper exerts on the ground.

2. (a) The force with which a man kicks a football is action and the force which the football exerts on the man will be its reaction.

   (b) The force with which earth pulls the moon is action and the force which the moon exerts on the earth will be its reaction.

   (c) If the force which the ball exerts on the wall is the action then the force which the wall exerts on the ball will be its reaction.

3. No. The arguement is not correct. The almirah moves when the push by the person exceeds the frictional force between the almirah and the floor. He does not get pushed backward due to a large force of friction that he experiences due to the floor. On a slippery surface, he will not be able to push the almirah foward.

3.4

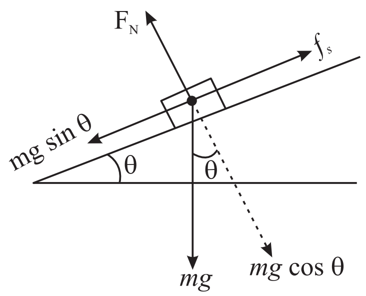


Fig. 3.17

2. 40 N

3. (a) (5×9.8) N

   (b) $F = (5 \times 2)\text{ N} + (5 \times 9.8)\text{ N} = 59\text{ N}$

3.5

(1) When the train starts it has an acceleration, say a. Thus the total force acting on water in the frame of reference attached to the train is

$$\mathbf{F} = m\,\mathbf{g} - m\,\mathbf{a}$$

where m is the mass of the water and the glass. (Fig. 3.16). The surface of the water takes up a position normal to F as shown.

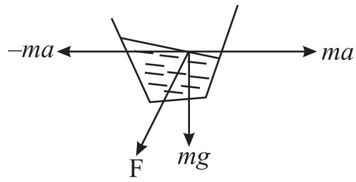


Fig. 3.18

(2) To the passenger sitting inside, a centrifugal force ($-mv^2/r$) acts on the car. The greater v is the larger r would be. To an observer standing on the road, the car moving in a curve has a centripetal acceleration given by v^2/r. Once again, the greater is v, the larger will be r.

(3) The net centrifugal force on the particle is $\mathbf{F} = m\omega^2 r = (6 \times 10^{-10}\,\text{kg}) \times (2\pi \times 10^3\,\text{rad s}^{-1})^2 \times (0.04\,\text{m}) = 9.6 \times 10^{-4}\,\text{N}$.

(4) For an object to fly off centrifugal force (= centripetal force) should be just more than the weight of a body. If r is the radius of the earth then $\frac{mv^2}{r} = mg$

as $v = r\omega$

$$\frac{r^2\omega^2}{r} = g$$

or, angular speed $\omega = \sqrt{g/r}$

$\therefore$ Any angular speed more than $\sqrt{g/r}$ will make objects fly off.

5. Zero (as it is a case of free fall of a body).

Answers to Terminal Problems

1. (d)

2. (a) if internal forces developed within the material counter bank the external force. A it happens in case of force applied on a wall.

   (b) It force is applied at right angles to the direction of motion of the body, the force changes the direction of motion of body and not to speed.

3. $v = 0 + (-g) \times 4$

$|v| = 40 \text{ m s}^{-1}$

$\therefore \Delta P = m\,(v - u) = (0.5 \times 40) = 20 \text{ kg m s}^{-2}$

4. When 10 N force acts for 1s.
5. 0.76 N
7. 250 N.
8. 27 N
10. (a) 10 m s^{-1} (b) 20 m s^{-1}
12. (a) 0 N (b) 4.9 N (c) ~7.5 N
13. 5 N
14. 14.2 N
15. (a) 2 m s^{-2} (b) 6 N
16. 3 kg
18. 20 m

312en04

4

MOTION IN A PLANE

In the preceding two lessons you have studied the concepts related to motion in a straight line. Can you describe the motion of objects moving in a plane, i.e, in two dimensions, using the concepts discussed so far. To do so, we have to introduce certain new concepts. An interesting example of motion in two dimensions is the motion of a ball thrown at an angle to the horizontal. This motions is called a *projectile motion.*

In this lesson you will learn to answer questions like : What should be the position and speed of an aircraft so that food or medicine packets dropped from it reach the people affected by floods or an earthquake? How should an athlete throw a discuss or a javelin so that it covers the maximum horizontal distance? How should roads be designed so that cars taking a turn around a curve do not go off the road? What should be the speed of a satellite so that it moves in a circular orbit around the earth? And so on.

Such situations arise in projectile motion and circular motion. Generally, circular motion refers to motion in a horizontal circle. However, besides moving in a horizontal circle, the body may also move in a vertical circle. We will introduce the concepts of angular speed, centripetal acceleration, and centripetal force to explain this kind of motion.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain projectile motion and circular motion and give their examples;*
- *explain the motion of a body in a vertical circle;*
- *derive expressions for the time of flight, range and maximum height of a projectile;*
- *derive the equation of the trajectory of a projectile;*
- *derive expressions for velocity and acceleration of a particle in circular motion; and*
- *define radial and tangential acceleration.*

4.1 PROJECTILE MOTION

The first breakthrough in the description of projectile motion was made by Galileo. He showed that the horizontal and vertical motions of a slow moving projectile are mutually independent. This can be understood by doing the following activity.

Take two cricket balls. Project one of them horizontally from the top of building. At the same time drop the other ball downward from the same height. What will you notice?

You will find that both the balls hit the ground at the same time. This shows that the downward acceleration of a projectile is the same as that of a freely falling body. Moreover, this takes place independent of its horizontal motion. Further, measurement of time and distance will show that the horizontal velocity continues unchanged and takes place independent of the vertical motion.

In other words, the two important properties of a projectile motion are :

(i) a constant horizontal velocity component

(ii) a constant vertically downward acceleration component.

The combination of these two motions results in the curved path of the projectile.

Refer to Fig. 4.1. Suppose a boy at A throws a ball with an initial horizontal speed. According to Newton's second law there will be no acceleration in the horizontal direction unless a horizontally directed force acts on the ball. Ignoring friction of air, the only force acting on the ball once it is free from the hand of the boy is the force of gravity.

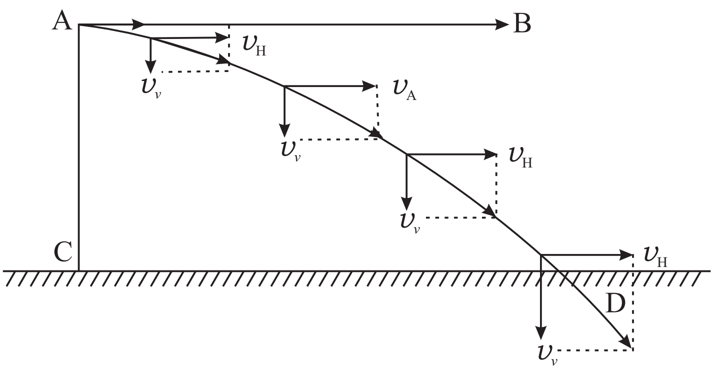


Fig. 4.1: Curved path of a projectile

Hence the horizontal speed υ_H of the ball does not change. But as the ball moves with this speed to the right, it also falls under the action of gravity as shown by the vector's υ_v representing the vertical component of the velocity. Note that $\upsilon = \sqrt{\upsilon_H^2 + \upsilon_v^2}$ and is tangential to the trajectory.

Having defined projectile motion, we would like to determine how high and how far does a projectile go and for how long does it remain in air. These factors are important if we want to launch a projectile to land at a certain target - for instance, a football in the goal, a cricket ball beyond the boundary and relief packets in the reach of people marooned by floods or other natural disasters.

4.1.1 Maximum Height, Time of Flight and Range of a Projectile

Let us analyse projectile motion to determine its maximum height, time of flight and range. In doing so, we will ignore effects such as wind or air resistance. We can characterise the initial velocity of an object in projectile motion by its vertical and horizontal components. Let us take the positive *x*-axis in the horizontal direction and the positive *y*-axis in the vertical direction (Fig. 4.2).

Let us assume that the initial position of the projectile is at the origin O at $t = 0$. As you know, the coordinates of the origin are $x = 0$, $y = 0$. Now suppose the projectile is launched with an initial velocity υ_0 at an angle θ_0, known as the ***angle of elevation,*** to the x-axis. Its components in the *x* and *y* directions are,

$$\upsilon_{ox} = \upsilon_o \cos \theta_0 \qquad (4.1\text{ a})$$

and

$$\upsilon_{oy} = \upsilon_o \sin \theta_0 \qquad (4.1\text{ b})$$

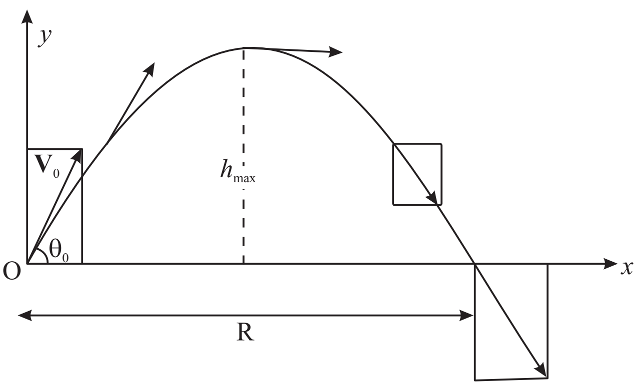


Fig 4.2 : Maximum height, time of flight and range of a projectile

Let a_x and a_y be the horizontal and vertical components, respectively, of the projectile's acceleration. Then

$$a_x = 0;\ a_y = -g = -9.8 \text{ m s}^{-2} \qquad (4.2)$$

The negative sign for a_y appears as the acceleration due to gravity is always in the negative *y* direction in the chosen coordinate system.

Notice that a_y is constant. Therefore, we can use Eqns. (2.6) and (2.9) to write expressions for the horizontal and vertical components of the projectile's velocity and position at time *t*. These are given by

Notes

Horizontal motion $\quad v_x = v_{ox}, \quad \text{since } a_x = 0$ (4.3a)

$x = v_{ox}t = v_0 \cos\theta_0 t$ (4.3b)

Vertical motion $\quad v_y = v_{oy} - g\,t = v_0 \sin\theta_0 - gt$ (4.3c)

$y = v_{oy}t - \tfrac{1}{2}g\,t^2 = v_0 \sin\theta_0 t - \tfrac{1}{2}g\,t^2$ (4.3d)

The vertical position and velocity components are also related through Eqn. (2.10) as

$$-g\,y = \tfrac{1}{2}\,(v_y^2 - v_{oy}^2) \qquad (4.3e)$$

You will note that the horizontal motion, given by Eqns. (4.3a and b), is motion with constant velocity. And the vertical motion, given by Eqns. (4.3c and d), is motion with constant (downward) acceleration. The vector sum of the two respective components would give us the velocity and position of the projectile at any instant of time.

Now, let us make use of these equations to know the maximum height, time of flight and range of a projectile.

(a) Maximum height : As the projectile travels through air, it climbs upto some maximum height (h) and then begins to come down. ***At the instant when the projectile is at the maximum height, the vertical component of its velocity is zero.*** This is the instant when the projectile stops to move upward and does not yet begin to move downward. Thus, putting $v_y = 0$ in Eqns. (4.3c and e), we get

$$0 = v_{oy} - g\,t,$$

Thus the time taken to rise taken to the maximum height is given by

$$t = \frac{v_{oy}}{g} = \frac{v_0 \sin\theta_0}{g} \qquad (4.4)$$

At the maximum height h attained by the projectile, the vertical velocity is zero. Therefore, applying $v^2 - u^2 = 2\,\text{a s} = 2\,g\,h,$ we get the expression for maximum height:

$$h = \frac{v_0^2 \sin^2\theta_0}{2g} \qquad (\text{as } v = 0 \text{ and } u = v_0 \sin\theta) \qquad (4.5)$$

Note that in our calculation we have ignored the effects of air resistance. This is a good approximation for a projectile with a fairly low velocity.

Using Eqn.(4.4) we can also determine the total time for which the projectile is in the air. This is termed as the ***time of flight.***

(b) Time of flight : ***The time of flight of a projectile is the time interval between the instant of its launch and the instant when it hits the ground.*** The time t given by Eq.(4.4) is the time for half the flight of the ball. Therefore, the total time of flight is given by

$$T = 2t = \frac{2\,\upsilon_0 \sin\theta_0}{g} \qquad (4.6)$$

Finally we calculate the distance travelled horizontally by the projectile. This is also called its ***range.***

(c) Range : The range R of a projectile is calculated simply by multiplying its time of flight and horizontal velocity. Thus using Eqns. (4.3b) and (4.4), we get

$$R = (\upsilon_{ox})\ (2\,t)$$

$$= (\upsilon_0 \cos\theta_0)\frac{(2\,\upsilon_0 \sin\theta_0)}{g}$$

$$= \upsilon_0^2\,\frac{(2\sin\theta_0 \cos\theta_0)}{g}$$

Since 2 sin θ cos θ = sin 2θ, the range R is given by

$$R = \frac{\upsilon_0^2 \sin 2\theta_0}{g} \qquad (4.7)$$

From Eqn. (4.7) you can see that the range of a projectile depends on

- its initial speed υ_0, and
- its direction given by θ_0.

Now can you determine the angle at which a disc, a hammer or a javelin should be thrown so that it covers maximum distance horizontally? In other words, let us find out the angle for which the range would be maximum?

Clearly, R will be maximum for any given speed when $\sin 2\theta_0 = 1$ or $2\theta_0 = 90^0$.

Thus, for R *to be maximum at a given speed* υ_0, θ_0 should be equal to 45^0.

Let us determine these quantities for a particular case.

Example 4.1 : In the centennial (on the occasion of its centenary) Olympics held at Atlanta in 1996, the gold medallist hammer thrower threw the hammer to a distance of 19.6m. Assuming this to be the maximum range, calculate the initial speed with which the hammer was thrown. What was the maximum height of the hammer? How long did it remain in the air? Ignore the height of the thrower's hand above the ground.

Solution : Since we can ignore the height of the thrower's hand above the ground, the launch point and the point of impact can be taken to be at the same height. We take the origin of the coordinate axes at the launch point. Since the distance covered by the hammer is the range, it is equal to the hammer's range for $\theta_0 = 45^0$. Thus we have from Eqn.(4.7):

$$R = \frac{v_0^2}{g}$$

or
$$v_0 = \sqrt{Rg}$$

It is given that $R = 19.6$ m. Putting $g = 9.8$ ms^{-2} we get

$$v_0 = \sqrt{(19.6\text{m}) \times (9.8\text{ ms}^{-2})} = 9.8\sqrt{2}\text{ ms}^{-1} = 14.01\text{ms}^{-1}$$

The maximum height and time of flight are given by Eqns. (4.5) and (4.6), respectively. Putting the value of v_0 and sin θ_0 in Eqns. (4.5) and (4.6), we get

$$\text{Maximum height, } h = \frac{\left(9.8\sqrt{2}\right)^2 \text{m}^2\text{s}^{-2} \times \left(\frac{1}{2}\right)^2}{2 \times 9.8\text{ms}^{-2}} = 4.9\text{ m}$$

$$\text{Time of flight, } T = \frac{2 \times \left(9.8\sqrt{2}\right)\text{m s}^{-1}}{9.8\text{ m s}^{-2}} \times \sqrt{\frac{1}{2}} = 2\text{ s}$$

Now that you have studied some concepts related to projectile motion and their applications, you may like to check your understanding. Solve the following problems.

INTEXT QUESTIONS 4.1

1. Identify examples of projectile motion from among the following situations:
   (a) An archer shoots an arrow at a target
   (b) Rocks are ejected from an exploding volcano
   (c) A truck moves on a mountainous road
   (d) A bomb is released from a bomber plane.
   [**Hint** : Remember that at the time of release the bomb shares the horizontal motion of the plane.]
   (e) A boat sails in a river.
2. Three balls thrown at different angles reach the same maximum height (Fig. 4.3):
   (a) Are the vertical components of the initial velocity the same for all the balls? If not, which one has the least vertical component?

(b) Will they all have the same time of flight?

(c) Which one has the greatest horizontal velocity component?

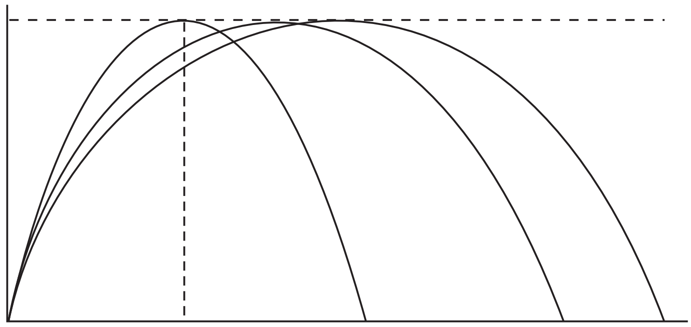


Fig. 4.3 : Trajectories of a projectile

3. An athelete set the record for the long jump with a jump of 8.90 m. Assume his initial speed on take off to be 9.5 ms^{-1}. How close did he come to the maximum possible range in the absence of air resistance?

   Take g = 9.78 ms^{-2}.

4.2 THE TRAJECTORY OF A PROJECTILE

The path followed by a projectile is called its trajectory. Can you recognise the shapes of the trajectories of projectiles shown in Fig. 4.1, 4.2 and 4.3.

Although we have discussed quite a few things about projectile motion, we have still not answered the question: What is the path or trajectory of a projectile? So let us determine the equation for the trajectory of a projectile.

It is easy to determine the equation for the path or trajectory of a projectile. You just have to eliminate t from Eqns. (4.3b) and (4.3d) for x and y. Substituting the value of t from Eqn. (4.3b) in Eqn.(4.3d) we get

$$y = \upsilon_{oy} \frac{x}{\upsilon_{ox}} - \frac{1}{2} \frac{g x^2}{\upsilon_{ox}^2} \left(\text{as } t = \frac{x}{\upsilon_{ox}} \right) \quad (4.8\text{ a})$$

Using Eqns. (4.1 a and b), Eqn (4.8a) becomes

$$y = (\tan \theta_0)\, x - \frac{g}{2(\upsilon_0 \cos \theta_0)^2} x^2 \quad (4.8\text{ b})$$

as $\upsilon_{oy} = \upsilon_0 \sin \theta$ and $\upsilon_{ox} = \upsilon_0 \cos \theta$.

Eqn. (4.8) is of the form $y = a\,x + b\,x^2$, which is the equation of a **parabola.** Thus, if air resistance is negligible, ***the path of any projectile launched at an angle to the horizontal is a parabola or a portion of a parabola.*** In Fig 4.3 you can see some trajectories of a projectile at different angles of elevation.

Eqns. (4.5) to (4.7) are often handy for solving problems of projectile motion. For example, these equations are used to calculate the launch speed and the angle of elevation required to hit a target at a known range. However, these equations do not give us complete description of projectile motion, if distance covered are very large. To get a complete description, we must include the rotation of the earth also. This is beyond the scope of this course.

Now, let us summarise the important equations describing projectile motion launched from a point (x_0, y_0) with a velocity υ_0 at an angle of elevation, θ_0.

Equations of Projectile Motion:

$$a_x = 0 \qquad a_y = -g \qquad (4.9\text{ a})$$

$$\upsilon_x = \upsilon_0 \cos\theta_0 \qquad \upsilon_y = \upsilon_0 \sin\theta - g\,t \qquad (4.9\text{ b})$$

$$x = x_0 + (\upsilon_0 \cos\theta_0)t \qquad y = y_0 + (\upsilon_0 \sin\theta)\,t - (½)\,g\,t^2 \qquad (4.9\text{ c})$$

Equation of trajectory:

$$y = y_0 + (\tan\theta)(x - x_0) - \frac{g}{2(\upsilon_0 \cos\theta_0)^2}(x - x_0)^2 \qquad (4.9\text{ d})$$

Note that these equations are more general than the ones discussed earlier. The initial coordinates are left unspecified as (x_0, y_0) rather than being placed at (0,0). Can you derive this general equation of the projectile trajectory? Do it before proceeding further?

Thus far you have studied motion of objects in a plane, which can be placed in the category of projectile motion. In projectile motion, the acceleration is constant both in magnitude and direction. There is another kind of two-dimensional motion in which acceleration is constant in magnitude but not in direction. This is uniform circular motion. Generally, circular motion refers to motion in a horizontal circle. However, motion in a vertical circle is also possible. You willl earn about them in the following section

Evangelista Torricelli
(1608 – 1647)

Italian mathematician and a student of Galelio Galili, he invented mercury barometer, investigated theory of projectiles, improved telescope and invented a primitive microscope. Disproved that nature abhors vacuum, presented torricellis theorem.

4.3 CIRCULAR MOTION

Look at Fig. 4.4a. It shows the position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ of a particle in uniform circular motion at two different instants of time t_1 and t_2, respectively. The word 'uniform' refers to constant speed. We have said that the speed of the particle is constant. What about its velocity? To find out velocity, recall the definition of average velocity and apply it to points P_1 and P_2:

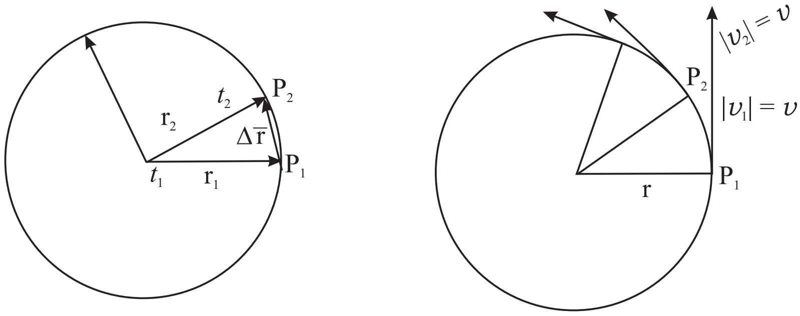
$$\boldsymbol{v}_{av} = \frac{\mathbf{r}_2 - \mathbf{r}_1}{t_2 - t_1} = \frac{\Delta \mathbf{r}}{\Delta t} \qquad (4.10\text{ a})$$

The motion of a gramophone record, a grinding wheel at constant speed, the moving hands of an ordinary clock, a vehicle turning around a corner are examples of circular motion. The movement of gears, pulleys and wheels also involve circular motion. The simplest kind of circular motion is uniform circular motion. The most familiar example of uniform circular motion are a point on a rotating fan blade or a grinding wheel moving at constant speed.

One of the example of uniform circular motion is an artificial satellite in circular orbit around the earth. We have been benefitted immensely by the INSAT series of satellites and other artificial satellites. So let us now learn about uniform circular motion.

4.3.1 Uniform Circular Motion

By definition, ***uniform circular motion is motion with constant speed in a circle.***



(a) (b)

Fig. 4.4: (a) Positions of a particle in uniform circular motion; (b) Uniform circular motion

The vector $\Delta\mathbf{r}$ is shown in Fig. 4.4a. Now suppose you make the time interval Δt smaller and smaller so that it approaches zero. What happens to $\Delta\mathbf{r}$? In particular, what is the direction of $\Delta\mathbf{r}$? It approaches the tangent to the circle at point P_1 as Δt tends to zero. Mathematically, we define the instantaneous velocity at point P_1 as

$$\boldsymbol{v} = \lim_{\Delta t \to 0} \frac{\Delta \mathbf{r}}{\Delta t} = \frac{d\mathbf{r}}{dt}$$

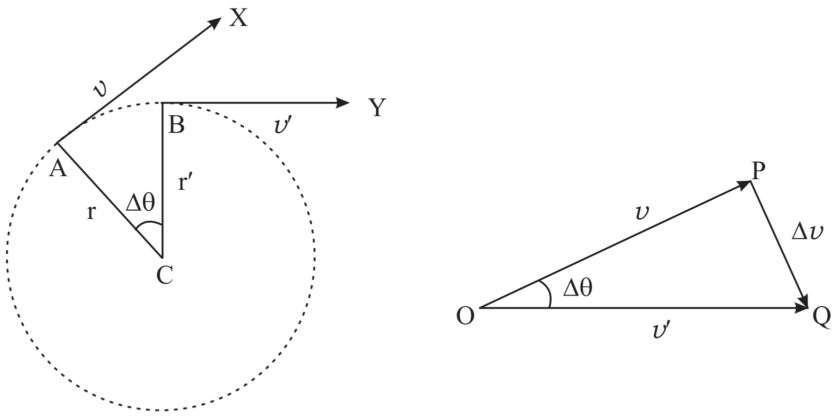
Thus, in uniform circular motion, the velocity vector changes continuously. Can you say why? This is because the direction of velocity is not constant. It goes on changing continuously as the particle travels around the circle (Fig. 4.4b). Because of ***this change in velocity, uniform circular motion is accelerated motion.*** The acceleration of a particle in uniform circular motion is termed as centripetal acceleration. Let us learn about it in some detail.

Centripetal acceleration : Consider a particle of mass m moving with a ***uniform speed*** υ in a circle. Suppose at any instant its position is at A and its motion is directed along AX. After a small time Δt, the particle reaches B and its velocity is represented by the tangent at B directed along BY.

Let **r** and **r′** be the position vectors and $\boldsymbol{\upsilon}$ and $\boldsymbol{\upsilon}'$; the velocities of the particle at A and B respectively as shown in Fig. 4.5 (a). The change in velocity $\Delta\boldsymbol{\upsilon}$ is obtained using the triangle law of vectors. As the path of the particle is circular and velocity is along its tangent, $\boldsymbol{\upsilon}$ is perpendicular to **r** and $\boldsymbol{\upsilon}'$ is perpendicular to Δ**r.** As the average acceleration $\left(\mathbf{a} = \frac{\Delta\boldsymbol{\upsilon}}{\Delta t}\right)$ is along $\Delta\boldsymbol{\upsilon}$, it (i.e., the average acceleration) is perpendicular to Δ**r**.

Let the angle between the position vectors **r** and **r′** be $\Delta\theta$. Then the angle between velocity vectors $\boldsymbol{\upsilon}$ and $\boldsymbol{\upsilon}'$ will also be $\Delta\theta$ as the velocity vectors are always perpendicular to the position vectors.

To determine the change in velocity $\Delta\boldsymbol{\upsilon}$ due to the change in direction, consider a point O outside the circle. Draw a line OP parallel to and equal to AX (or $\boldsymbol{\upsilon}$) and a line OQ parallel to and equal to BY (or $\boldsymbol{\upsilon}'$). As $|\boldsymbol{\upsilon}| = |\boldsymbol{\upsilon}'|$, OP = OQ. Join PQ. You get a triangle OPQ (Fig. 4.5b)



(a) (b)

Fig. 4.5

Now in triangle OPQ, sides OP and OQ represent velocity vectors $\boldsymbol{\upsilon}$ and $\boldsymbol{\upsilon}'$ at A and B respectively. Hence, their difference is represented by the side PQ in magnitude and direction. In other words the change in the velocity equal to PQ in magnitude and direction takes place as the particle moves from A to B in time Δt.

$\therefore$ Acceleration = Rate of change of velocity

$$\mathbf{a} = \frac{\mathbf{PQ}}{\Delta t} = \frac{\Delta \boldsymbol{v}}{\Delta t}$$

As Δt is very small AB is also very small and is nearly a straight line. Then Δ ACB and ΔPOQ are isosceles triangles having their included angles equal. The triangles are, therefore, similar and hence,

$$\frac{PQ}{AB} = \frac{OP}{CA}$$

or

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

[as magnitudes of velocity vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2 = v$ (say)]

or

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

But $\frac{\Delta v}{\Delta t}$ is the acceleration of the particle. Hence

Centripetal acceleration, $a = \frac{v^2}{r}$

Since $v = r\,\omega$, the magnitude of centripetal force in given by

$$F = m\,a = \frac{m\,v^2}{r} = m\,r\omega^2.$$

As Δt is very small, $\Delta\theta$ is also very small and $\angle$OPQ = $\angle$OQP = 1 right angle.

Thus PQ is perpendicular to OP, which is parallel to the tangent AX at A. Now AC is also perpendicular to AX. Therefore AC is parallel to PQ. It shows that the contripetal force at any point acts towards the centre along the radius.

It shows that some minimum centripetal force has to be applied on a body to make it move in a circular path. In the absence of such a force, the body will move in a straight line path.

To experience this, you can perform a simple activity.

ACTIVITY 4.1

Take a small piece of stone and tie it to one end of a string. Hold the other end with your fingers and then try to whirl the stone in a horizontal or vertical circle. Start with a small speed of rotation and increase it gradually. What happens when the speed of rotation is low? Do you feel any pull on your fingers when the stone is whirling. What happens to the stone when you leave the end of the string you were holding? How do you explain this?

ACTIVITY 4.2

Take an aluminium channel of length one metre and bend it in the form shown in the diagram with a circular loop in the middle. Take help of some technical preson if required.

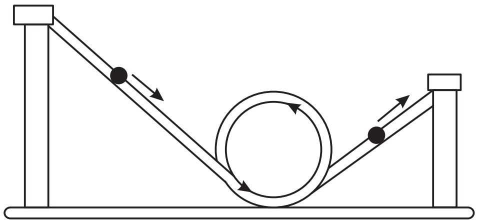


Fig. 4.6: The ball will loop if it starts rolling from a point high enough on the incline

Roll down a glass marble from different heights of the channel on the right hand side, and see whether the marble is able to loop the loop in each case or does it need some minimum height (hence velocity) below which the marble will not be able to complete the loop and fall down. How do you explain it?

Some Applications of Centripetal Force

(i) **Centrifuges :** These are spinning devices used for separating materials having different densities. When a mixture of two materials of different densities placed in a vessel is rotated at high speed, the centripetal force on the heavier material will be more. Therefore, it will move to outermost position in the vessel and hence can be separated. These devises are being used for uranium enrichment. In a chemistry laboratory these are used for chemical analysis.

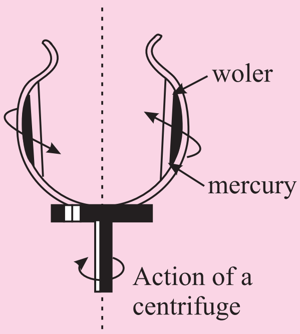


Fig. 4.7: When mercury and water are rotated in a dish, the water stays inside. Centripetal force, like gravitational force, is greater for the more dense substance.

(ii) Mud clings to an automobile tyre until the speed becomes too high and then it flies off tangentially (Fig. 4.8).

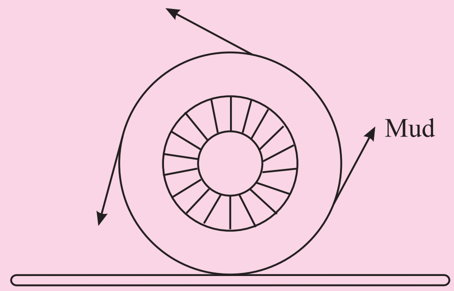


Fig. 4.8: Mud or water on a fast-turning wheel flies off tangentially

(iii) **Planetary motion :** The Earth and the other planets revolving round the sun get necessary centripetal force from the gravitational force between them and the sun.

Example 4.2 : Astronauts experience high acceleration in their flights in space. In the training centres for such situations, they are placed in a closed capsule, which is fixed at the end of a revolving arm of radius 15 m. The capsule is whirled around in a circular path, just like the way we whirl a stone tied to a string in a horizontal circle. If the arm revolves at a rate of 24 revolutions per minute, calculate the centripetal acceleration of the capsule.

Solution : The circumference of the circular path is $2\pi \times$ (radius) $= 2\pi \times 15$ m. Since the capsule makes 24 revolutions per minute or 60 s, the time it takes to go once around this circumference is $\frac{60}{24}$ s. Therefore,

speed of the capsule, $v = \frac{2\pi r}{T} = \frac{2\pi \times 15\ \text{m}}{(60/24)\ \text{s}} = 38\ \text{ms}^{-1}$

The magnitude of the centripetal acceleration

$$a = \frac{v^2}{r} = \frac{(38\ \text{ms}^{-1})^2}{15\ \text{m}} = 96\ \text{ms}^{-2}$$

Note that centripetal acceleration is about 10 times the acceleration due to gravity.

4.3.2 Motion In a Vertical Circle

When a body moves in a horizontal circle, the direction of its linear velocity goes on changing but the angular velocity remains constant. But, when a body moves in a vertical circle, the angular velocity, too, cannot remain constant on account of the acceleration due to gravity.

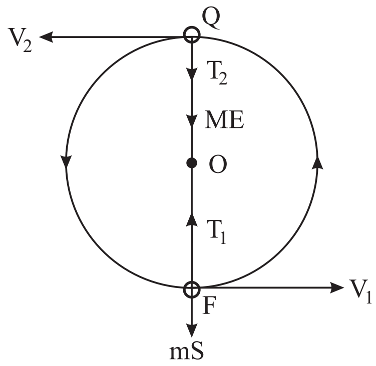


Fig. 4.9

Let a body of mass m tied to a string be rotated anticlockwise in a vertical circle of radius r about a point O. As the body rotates in the vertical circle, its speed is maximum at the lowest point P. It goes on decreasing as the body moves up to Q, and is minimum at the highest point Q. The speed goes on increasing as the body falls from Q to P along the circular path.

The forces acting on the body at P are weight of the body 'mg' and the tension T_1 of the string in the direction as shown in Fig. 4.9. Similarly, the forces acting on the body at Q are mg and the tension T_2 in the direction shown in Fig. 4.9. If v_1 and v_2 be the velocities of the body at P and Q, respectively, we have at P:

$$T_1 - mg = \frac{mv_1^2}{r}$$

or $$T_1 = \frac{mv_1^2}{r} + mg$$

Note that at P, the force $(T_1 - mg)$ acts along PO and provides the centripetal force.

Similarly at Q,

$$T_2 + mg = \frac{mv_2^2}{r}$$

or $$T_2 = \frac{mv_2^2}{r} - mg$$

For the body to move along the circle without any slaking of the string,

$$T_2 \geq 0$$

i.e. the minimum value of the tension should be zero at Q.

When, $T_2 = 0$, $$mg = \frac{mv_2^2}{r}$$

i.e. the minimum velocity at the highest point of the circle is, $\sqrt{gr}$

$$\because \qquad \omega_2 = \frac{v}{r} = \sqrt{g/r}$$

The minimum velocity (v_1) at the lowest point (P) of the circle should be such that the velocity (v_2) at the highest point (Q) becomes $\sqrt{gr}$

Using the relation, $v^2 - u^2 = 2as$, we have

$$v_2^2 - v_1^2 = -2g(2r) \qquad (s = 2r \text{ and g is negative})$$

or $$v_1^2 = v_2^2 + 4gr$$

$$v_1^2 = gr + 4gr = 5gr$$

or $$v_1 = \sqrt{5gr}$$

Hence, for a body to go around a vertical circle completely minimum velocity at the lowest point should be $\sqrt{5gr}$.

or $$w_1 = \sqrt{5g/r}$$

which shows that the angular velocity is also changing as the body moves in a vertical circle.

INTEXT QUESTIONS 4.2

1. In uniform circular motion, (a) Is the speed constant? (b) Is the velocity constant? (c) Is the magnitude of the acceleration constant? (d) Is acceleration constant? Explain.
2. In a vertical motion does the angular velocity of the body change? Explain.
3. An athlete runs around a circular track with a speed of 9.0 ms^{-1} and a centripetal acceleration of 3 ms^{-2}. What is the radius of the track?
4. The Fermi lab accelerator is one of the largest particle accelerators. In this accelerator, protons are forced to travel in an evacuated tube in a circular orbit of diameter 2.0 km at a speed which is nearly equal to 99.99995% of the speed of light. What is the centripetal acceleration of these protons?

   Take $c = 3 \times 10^8$ ms^{-1}.

4.4 APPLICATIONS OF UNIFORM CIRCULAR MOTION

So far you have studied that an object moving in a circle is accelerating. You have also studied Newton's laws in the previous lesson. From Newton's second law we can say that as the object in circular motion is accelerating, a net force must be acting on it.

What is the direction and magnitude of this force? This is what you will learn in this section. Then we will apply Newton's laws of motion to uniform circular motion. This helps us to explain why roads are banked, or why pilots feel pressed to their seats when they fly aircrafts in vertical loops.

Let us first determine the force acting on a particle that keeps it in uniform circular motion. Consider a particle moving with constant speed υ in a circle of radius r. From Newton's second law, the net external force acting on a particle is related to its acceleration by

$$\mathbf{F} = -\frac{m\upsilon^2}{r}\,\hat{\boldsymbol{r}}\,,\ |\mathbf{F}| = \frac{m\upsilon^2}{r} \qquad (4.19)$$

This net external force directed towards the centre of the circle with magnitude given by Eqn. (4.19) is called ***centripetal force***. ***An important thing to understand and remember is that the term 'centripetal force' does not refer to a type of force of interaction like the force of gravitation or electrical force.*** This term only tells us that the net force of a certain magnitude acting on a particle in uniform circular motion is directed towards the centre. It does not tell us how this force is provided.

Thus, the force may be provided by the gravitational attraction between two bodies. For example, in the motion of a planet around the sun, the centripetal force is provided by the gravitational force between the two. Similarly, the centripetal force for a car travelling around a bend is provided by the force of friction between the road and the car's tyres and/or by the horizontal component of normal reaction of banked road. You will understand these ideas better when we apply them in certain concrete situations.

4.4.1 Banking of Roads

While riding a bicycle and taking a sharp turn, you may have felt that something is trying to throw you away from your path. Have you ever thought as to why does it happen?

You tend to be thrown out because enough centripetal force has not been provided to keep you in the circular path. Some force is provided by the friction between the tyres and the road, but that may not be sufficient. When you slow down, the needed centripetal force decreases and you manage to complete this turn.

Consider now a car of mass m, travelling with speed υ on a curved section of a highway (Fig. 4.10). To keep the car moving uniformly on the circular path, a force must act on it directed towards the centre of the circle and its magnitude must be equal to $m\upsilon^2/r$. Here r is the radius of curvature of the curved section.

Now if the road is levelled, the force of friction between the road and the tyres provides the necessary centripetal force to keep the car in circular path. This causes a lot of wear and tear in the tyre and may not be enough to give it a safe turn. The roads at curves are, therefore, banked, where banking means the raising of the outer edge of the road above the level of the inner edge (Fig. 4.10). As a matter of fact, roads are designed to minimise reliance on friction. For example, when car tyres are smooth or there is water or snow on roads, the coefficient of friction becomes negligible. Roads are banked at curves so that cars can keep on track even when friction is negligible.

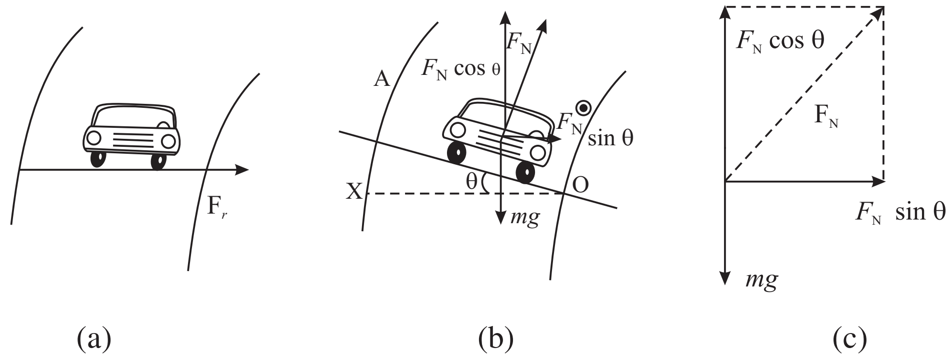


Fig. 4.10 : A car taking a turn (a) on a level road; (b) on a banked road; and (c) Forces on the car with F_N resolved into its rectangular components. Generally θ is not as large as shown here in the diagram.

Let us now analyse the free body diagram for the car to obtain an expression for the angle of banking, θ, which is adjusted for the sharpness of the curve and the maximum allowed speed.

Consider the case when there is no frictional force acting between the car tyres and the road. The forces acting on the car are the car's weight *mg* and $\mathbf{F}_N$, the force of normal reaction. The centripetal force is provided by the horizontal component of $\mathbf{F}_N$. Thus, resolving the force $\mathbf{F}_N$ into its horizontal and vertical components, we can write

$$F_N \sin\theta = \frac{m v^2}{r} \qquad (4.20a)$$

Since there is no vertical acceleration, the vertical component of $\mathbf{F}_N$ is equal to the car's weight:

$$F_N \cos\theta = m g \qquad (4.20b)$$

We have two equations with two unknowns, i.e., F_N and θ. To determine θ, we eliminate F_N. Dividing Eqn. (4.20 a) by Eqn. (4.20 b), we get

$$\tan\theta = \frac{m v^2 / r}{m g} = \frac{v^2}{r g}$$

or

$$\theta = \tan^{-1}\frac{v^2}{rg} \qquad (4.21)$$

How do we interpret Eqn. (4.21) for limits on υ and choice of θ? Firstly, Eqn.(4.21) tells us that the angle of banking is independent of the mass of the vehicle. So even large trucks and other heavy vehicles can ply on banked roads.

Secondly, θ should be greater for high speeds and for sharp curves (i.e., for lower values of r). For a given θ, if the speed is more than υ, it will tend to move towards the outer edge of the curved road. So a vehicle driver must drive within prescribed speed limits on curves. Otherwise, the will be pushed off the road. Hence, there may be accidents.

Usually, due to frictional forces, there is a range of speeds on either side of υ. Vehicles can maintain a stable circular path around curves, if their speed remains within this range. To get a feel of actual numbers, consider a curved path of radius 300 m. Let the typical speed of a vehicle be 50 ms^{-1}. What should the angle of banking be? You may like to quickly use Eqn.(4.21) and calculate θ.

$$\theta = \tan^{-1} \frac{(50 \text{ ms}^{-1})^2}{(300 \text{ m})(9.8 \text{ ms}^{-2})} = \tan^{-1}(0.017) = 1^0$$

You may like to consider another application.

4.4.2 Aircrafts in vertical loops

On Republic Day and other shows by the Indian Air Force, you might have seen pilots flying aircrafts in loops (Fig. 4.11a). In such situations, at the bottom of the loop, the pilots feel as if they are being pressed to their seats by a force several times the force of gravity. Let us understand as to why this happens. Fig. 4.11b shows the 'free body' diagram for the pilot of mass m at the bottom of the loop.

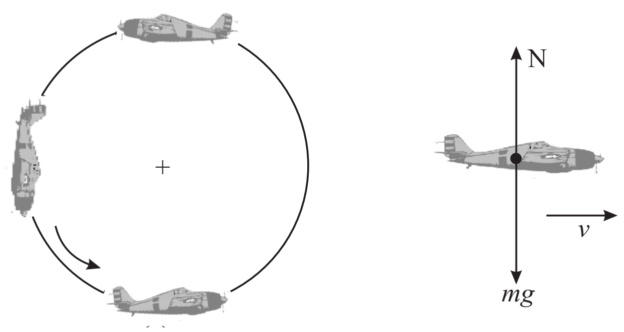


Fig. 4.11: (a) Aircrafts in vertical loops, (b) Free-body diagram for the pilot at the lowest point.

The forces acting on him are mg and the normal force N exerted by the seat. The net vertically upward force is N – mg and this provides the centripetal acceleration:

$$\mathrm{N} - mg = m\,a$$

or $$\mathrm{N} - mg = m\,\upsilon^2/r$$

or $$\mathrm{N} = m\,(g + \upsilon^2/r)$$

In actual situations, if $\upsilon = 200\ \mathrm{ms^{-1}}$ and $r = 1500$ m, we get

$$\mathrm{N} = m\,g\left[1 + \frac{(200\ \mathrm{m\,s^{-1}})^2}{(9.8\ \mathrm{m\,s^{-2}} \times 1500\ \mathrm{m})}\right] = m\,g \times 3.7$$

So the pilots feel as though force of gravity has been magnified by a factor of 3.7. If this force exceeds set limits, pilots may even black out for a while and it could be dangerous for them and for the aircraft.

INTEXT QUESTIONS 4.3

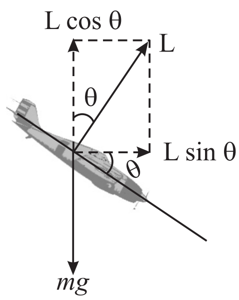


Fig. 4.12

1. Aircrafts usually bank while taking a turn when flying at a constant speed (Fig. 4.12). Explain why aircrafts do bank? Draw a free body diagram for this aircraft. ($\mathbf{F}_a$ is the force exerted by the air on the aircraft). Suppose an aircraft travelling at a speed $\upsilon = 100\ \mathrm{ms^{-1}}$ makes a turn at a banking angle of 30^0. What is the radius of curvature of the turn? Take $g = 10\ \mathrm{ms^{-2}}$.

2. Calculate the maximum speed of a car which makes a turn of radius 100 m on a horizontal road. The coefficient of friction between the tyres and the road is 0.90. Take $g = 10\ \mathrm{ms^{-2}}$.

3. An interesting act performed at variety shows is to swing a bucket of water in a vertical circle such that water does not spill out while the bucket is inverted at the top of the circle. For this trick to be performed sucessfully, the speed of the bucket must be larger than a certain minimum value. Derive an expression for the minimum speed of the bucket at the top of the circle in terms of its radius R. Calculate the speed for R = 1.0 m.

WHAT YOU HAVE LEARNT

- **Projectile motion** is defined as the motion which has constant velocity in a certain direction and constant acceleration in a direction perpendicular to that of velocity:

$$a_x = 0 \qquad\qquad a_y = -g$$

$$\upsilon_x = \upsilon_0 \cos\theta \qquad \upsilon_y = \upsilon_0 \sin\theta - g\,t$$

$$x = x_0 + (\upsilon_0 \cos\theta)\,t \qquad y = y_0 + (\upsilon_0 \sin\theta) -½ - g\,t^2$$

- Height $h = \dfrac{\upsilon_0^2 \sin 2\theta}{g}$
- Time of flight $T = \dfrac{2\upsilon_0 \sin\theta}{g}$
- Range of the projectile $R = \dfrac{\upsilon_0^2 \sin 2\theta}{g}$
- Equation of the Trajectory of a projectile $y = (\tan\theta_0)\,x - \dfrac{g}{2(\upsilon_0 \cos\theta_0)^2}\,x^2$
- **Circular motion** is uniform when the speed of the particle is constant. A particle undergoing ***uniform circular motion*** in a circle of radius r at constant speed υ has a ***centripetal acceleration*** given by

  $$\mathbf{a}_r = -\frac{\upsilon^2}{r}\,\hat{\mathbf{r}}$$

  where $\hat{\mathbf{r}}$ is the unit vector directed from the centre of the circle to the particle. The speed υ of the particle is related to its angular speed ω by $\upsilon = r\,\omega$.
- The ***centripetal force*** acting on the particle is given by

  $$\mathbf{F} = m\,\mathbf{a}_r = \frac{m\,\upsilon^2}{r}\,\hat{\mathbf{r}} = m\,r\,\omega^2$$

- When a body moves in a vertical circle, its angular velocity cannot remain constant.
- The minimum velocities at the highest and lowest points of a vetical circle are $\sqrt{gr}$ and $\sqrt{5gr}$ respectively

TERMINAL EXERCISE

1. Why does a cyclist bend inward while taking a turn on a circular path?
2. Explain why the outer rail is raised with respect to the inner rail on the curved portion of a railway track?
3. If a particle is having circular motion with constant speed, will its acceleration also be constant?
4. A stone is thrown from the window of a bus moving on horizontal road. What path will the stone follow while reaching the ground; as seen by a observer standing on the road?

5. A string can sustain a maximum force of 100 N without breaking. A mass of 1kg is tied to one end of the piece of string of 1m long and it is rotated in a horizontal plane. Compute the maximum speed with which the body can be rotated without breaking the string?
6. A motorcyclist passes a curve of radius 50 m with a speed of 10 m s^{-1}. What will be the centripetal acceleration when turning the curve?
7. A bullet is fired with an initial velocity 300 ms^{-1} at an angle of 30^0 with the horizontal. At what distance from the gun will the bullet strike the ground?
8. The length of the second's hand of a clock is 10 cm. What is the speed of the tip of this hand?
9. You must have seen actors in Hindi films jumping over huge gaps on horse backs and motor cycles. In this problem consider a daredevil motor cycle rider trying to cross a gap at a velocity of 100 km h^{-1}. (Fig. 4.13). Let the angle of incline on either side be 45^0. Calculate the widest gap he can cross.
10. A shell is fired at an angle of elevation of 30^0 with a velocity of 500 m s^{-1}. Calculate the vertical and horizontal components of the velocity, the maximum height that the shell reaches, and its range.
11. An aeroplane drops a food packet from a height of 2000 m above the ground while in horizontal flight at a constant speed of 200 kmh^{-1}. How long does the packet take to fall to the ground? How far ahead (horizontally) of the point of release does the packet land?

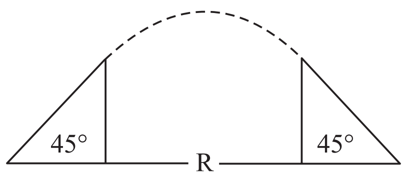


Fig. 4.13

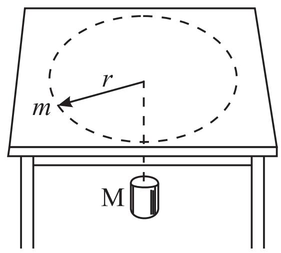


Fig. 4.14

12. A mass m moving in a circle at speed v on a frictionless table is attached to a hanging mass M by a string through a hole in the table (Fig. 4.14). Determine the speed of the mass m for which the mass M would remain at rest.
13. A car is rounding a curve of radius 200 m at a speed of 60 kmh^{-1}. What is the centripetal force on a passenger of mass $m = 90$ kg?

ANSWERS TO INTEXT QUESTIONS

4.1

(1) (a), (b), (d)

(2) (a) Yes (b) Yes (c) The ball with the maximum range.

(3) Maximum Range

$$\frac{v_0^2}{g} = \frac{(9.5\ \text{ms}^{-1})^2}{9.78\ \text{ms}^{-2}} = 9.23\ \text{m}$$

Thus, the difference is 9.23 m – 8.90 m = 0.33 m.

4.2

(1) (a) Yes (b) No (c) Yes (d) No

The velocity and acceleration are not constant because their directions are changing continuously.

(2) Yes. The angular velocity changes because of acceleration due to gravity

(3) Since

$$a = \frac{v^2}{r},\ r = \frac{v^2}{a} = \frac{(9.0\ \text{ms}^{-1})^2}{3\ \text{ms}^{-2}} = 27\ \text{m}$$

(4) $$a = \frac{c^2}{r} = \frac{(3\times10^8\ \text{ms}^{-1})^2}{10\times10^3\ \text{m}}$$

$$= 9 \times 10^{13}\ \text{ms}^{-2}$$

4.3

(1) This is similar to the case of banking of roads. If the aircraft banks, there is a component of the force L exerted by the air along the radius of the circle to provide the centripetal acceleration. Fig. 4.15 shows the free body diagram. The radius of curvature is

$$R = \frac{v^2}{g\tan\theta_0} = \left(\frac{100\ \text{ms}^{-1}}{10\ \text{ms}^{-2}\times\tan 30^\circ}\right)^2 = 10\sqrt{3}\ \text{m} = 17.3\ \text{m}$$

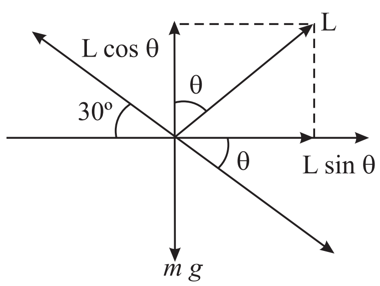


Fig. 4.15

(2) The force of friction provides the necessary centripetal acceleration :

$$F_s = \mu_s N = \frac{mv^2}{r}$$

Since the road is horizontal N – mg

Thus $$\mu_s mg = \frac{mv^2}{r}$$

or $$v^2 = \mu_s g r$$

or $$v = (0.9 \times 10 \text{ m s}^{-2} \times 100 \text{ m})^{1/2}$$

$$v = 30 \text{ ms}^{-1}.$$

(3) Refer to Fig. 4.16 showing the free body diagram for the bucket at the top of the circle. In order that water in the bucket does not spill but keeps moving in the circle, the force *mg* should provide the centripetal acceleration. At the top of the circle.

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

or $$v^2 = Rg$$

∴ $$v = \sqrt{Rg}$$

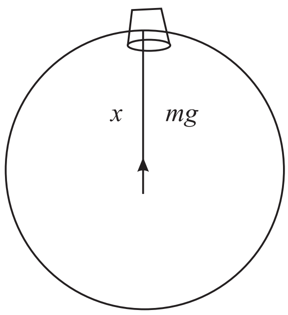


Fig. 4.16

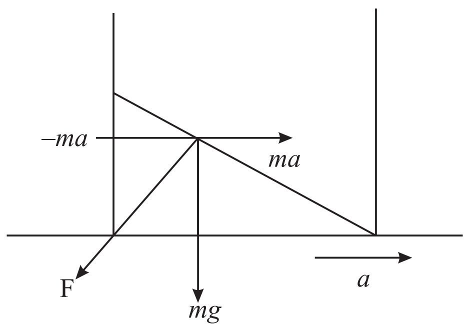


Fig. 4.17

This is the minimum value of the bucket's speed at the top of the vertical circle. For R = 1.0 m and taking g = 10 ms^{-2} we get

v = 10 m s^{-1} = 3.2 ms^{-1}

Answers to Terminal Problems

5. 10 ms^{-1}
6. 2 ms^{-2}
7. $900\sqrt{3}$ m

8. $1.05 \times 10^{-3}\ ms^{-1}$

9. 77.1 m

10. $\upsilon_x = 250\sqrt{3}\ ms^{-1}$

    $\upsilon y = 250\ ms^{-1}$

    Vertical height = 500 m

    Horizontal range = 3125 m

11. $t = 20$ s, 999.9 m

12. $\upsilon = \sqrt{\dfrac{m\,g\,r}{m}}$

13. 125 N

312en05

5

GRAVITATION

Have you ever thought why a ball thrown upward always comes back to the ground? Or a coin tossed in air falls back on the ground. Since times immemorial, human beings have wondered about this phenomenon. The answer was provided in the 17th century by Sir Isaac Newton. He proposed that the gravitational force is responsible for bodies being attracted to the earth. He also said that it is the same force which keeps the moon in its orbit around the earth and planets bound to the Sun. It is a universal force, that is, it is present everywhere in the universe. In fact, it is this force that keeps the whole universe together.

In this lesson you will learn Newton's law of gravitation. We shall also study the acceleration caused in objects due to the pull of the earth. This acceleration, called acceleration due to gravity, is not constant on the earth. You will learn the factors due to which it varies. You will also learn about gravitational potential and potential energy. You will also study Kepler's laws of planetary motion and orbits of artificial satellites of various kinds in this lesson. Finally, we shall recount some of the important programmes and achievements of India in the field of space research.

OBJECTIVES

After studying this lesson, you should be able to:

- *state the law of gravitation;*
- *calculate the value of acceleration due to gravity of a heavenly body;*
- *analyse the variation in the value of the acceleration due to gravity with height, depth and latitude;*
- *distinguish between gravitaitonal potential and gravitational potential energy;*
- *identify the force responsible for planetary motion and state Kepler's laws of planetary motion;*

- *calculate the orbital velocity and the escape velocity;*
- *explain how an artificial satellite is launched;*
- *distinguish between polar and equatorial satellites;*
- *state conditions for a satellite to be a geostationary satellite;*
- *calculate the height of a geostationary satellite and list their applications; and*
- *state the achievements of India in the field of satellite technology.*

5.1 LAW OF GRAVITATION

It is said that Newton was sitting under a tree when an apple fell on the ground. This set him thinking: since all apples and other objects fall to the ground, there must be some force from the earth acting on them. He asked himself: Could it be the same force which keeps the moon in orbit around the earth? Newton argued that at every point in its orbit, the moon would have flown along a tangent, but is held back to the orbit by some force (Fig. 5.1). Could this continuous 'fall' be due to the same force which forces apples to fall to the ground? He had deduced from Kepler's laws that the force between the Sun and planets varies as $1/r^2$. Using this result he was able to show that it is the same force that keeps the moon in its orbit around the earth. Then he generalised the idea to formulate the universal law of gravitation as.

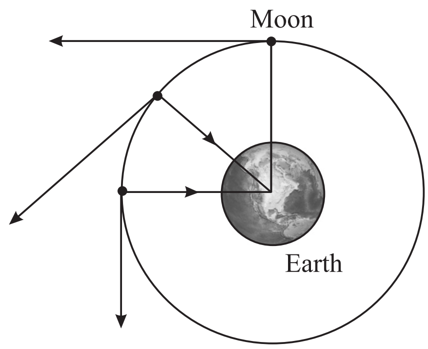


Fig. 5.1 : At each point on its orbit, the moon would have flown off along a tangent but the attraction of the earth keeps it in its orbit.

Every particle attracts every other particle in the universe with a force which varies as the product of their masses and inversely as the square of the distance between them. Thus, if m_1 and m_2 are the masses of the two particles, and r is the distance between them, the magnitude of the force F is given by.

$$F \propto \frac{m_1 m_2}{r^2}$$

or

$$F = G\frac{m_1 m_2}{r_2} \qquad (5.1)$$

The constant of proportionality, G, is called **the universal constant of gravitation. Its value remains the same between any two objects everywhere**

in the universe. This means that if the force between two particles is F on the earth, the force between these particles kept at the same distance anywhere in the universe would be the same.

One of the extremely important characteristics of the gravitational force is that it is always attractive. It is also one of the fundamental forces of nature.

Remember that the attraction is mutual, that is, particle of mass m_1 attracts the particle of mass m_2 and m_2 attracts m_1. Also, the force is along the line joining the two particles.

Knowing that the force is a vector quantity, does Eqn. (5.1) need modification? The answer to this question is that the equation should reflect both magnitude and the direction of the force. As stated, the gravitational force acts along the line joining the two particles. That is, m_2 attracts m_1 with a force which **is along the line joining the two particles** (Fig. 5.2). If the force of attraction exerted by m_1 on m_2 is denoted by $\mathbf{F}_{12}$ and the distance between them is denoted by $\mathbf{r}_{12}$, then the vector form of the law of gravitation is

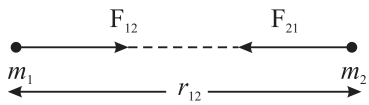


Fig. 5.2 : The masses m_1 and m_2 are placed at a distance r_{12} from eact other. The mass m_1 attracts m_2 with a force F_{12}.

$$\mathbf{F}_{12} = G\frac{m_1 m_2}{\mathbf{r}_{12}}\hat{\mathbf{r}}_{12} \qquad (5.2)$$

Here $\hat{r}_{12}$ is a unit vector from m_1 to m_2

In a similar way, we may write the force exerted by m_2 on m_1 as

$$\mathbf{F}_{21} = -G\frac{m_1 m_2}{\mathbf{r}_{21}^2}\hat{\mathbf{r}}_{21} \qquad (5.3)$$

As $\hat{\mathbf{r}}_{12} = -\hat{\mathbf{r}}_{21}$, from Eqns. (5.2) and (5.3) we find that

$$\mathbf{F}_{12} = -\mathbf{F}_{21} \qquad (5.4)$$

The forces $\mathbf{F}_{12}$ and $\mathbf{F}_{21}$ are equal and opposite and form a pair of forces of action and reaction in accordance with Newton's third law of motion. Remember that $\hat{\mathbf{r}}_{12}$ and $\hat{\mathbf{r}}_{21}$ have unit magnitude. However, the directions of these vectors are opposite to each other.

Unless specified, in this lesson we would use only the magnitude of the gravitational force.

The value of the constant G is so small that it could not be determined by Newton or his contemporary experimentalists. It was determined by Cavendish for the first time about 100 years later. Its accepted value today is 6.67×10^{-11} Nm^{2}kg^{-2}. It is because of the smallness of G that the gravitational force due to ordinary objects is not felt by us.

Example 5.1 : Kepler's third law states (we shall discuss this in greater details later) that if r is the mean distance of a planet from the Sun, and T is its orbital period, then r^3/T^2 = const. Show that the force acting on a planet is inversely proportional to the square of the distance.

Solution : Assume for simplicity that the orbit of a planet is circular. (In reality, the orbits are nearly circular.) Then the centripetal force acting on the planet is

$$F = \frac{m\upsilon^2}{r}$$

where υ is the orbital velocity. Since $\upsilon = r\omega = \frac{2\pi r}{\mathrm{T}}$, where T is the period, we can rewrite above expression as

$$F = m\left(\frac{2\pi r}{\mathrm{T}}\right)^2 \Big/ r$$

or $$F = \frac{4\pi^2 mr}{\mathrm{T}^2}$$

But $\mathrm{T}^2 \propto r^3$ or $\mathrm{T}^2 = Kr^3$ (Kelpler's 3rd law)

where K is a constant of proportionality. Hence

$$\therefore \quad F = \frac{4\pi^2 mr}{Kr^3} = \frac{4\pi^2}{K} \times \frac{m}{r^2} = \frac{4\pi^2 m}{K} \cdot \frac{1}{r^2}$$

or $$F \propto \frac{1}{r^2} \qquad \left(\because \frac{4\pi^2 m}{K} \text{ is constant for a planet}\right)$$

Before proceedins further, it is better that you check your progress.

INTEXT QUESTIONS 5.1

1. The period of revolution of the moon around the earth is 27.3 days. Remember that this is the period with respect to the fixed stars (the period of revolution with respect to the moving earth is about 29.5 days; it is this period that is used to fix the duration of a month in some calendars). The radius of moon's orbit is 3.84×10^8 m (60 times the earth's radius). Calculate the centripetal acceleration of the moon and show that it is very close to the value given by 9.8 ms^{-2} divided by 3600, to take account of the variation of the gravity as $1/r^2$.

2. From Eqn. (5.1), deduce dimensions of G.

3. Using Eqn. (5.1), show that G may be defined as the magnitude of force between two masses of 1 kg each separated by a distance of 1 m.

4. The magnitude of force between two masses placed at a certain distance is F. What happens to F if (i) the distance is doubled without any change in masses, (ii) the distance remains the same but each mass is doubled, (iii) the distance is doubled and each mass is also doubled?

5. Two bodies having masses 50 kg and 60 kg are seperated by a distance of 1m. Calculate the gravitational force between them.

5.2 ACCELERATION DUE TO GRAVITY

From Newton's second law of motion you know that a force **F** exerted on an object produces an acceleration **a** in the object according to the relation

$$\mathbf{F} = m\mathbf{a} \qquad (5.5)$$

The force of gravity, i.e., the force exerted by the earth on a body lying on or near its surface, also produces an acceleration in the body. The acceleration produced by the force of gravity is called the **acceleration due to gravity**. It is denoted by the symbol g. According to Eq. (5.1), the magnitude of the force of gravity on a particle of mass m on the earth's surface is given by

$$F = G\frac{mM}{R^2} \qquad (5.6)$$

where M is the mass of the earth and R is its radius. From Eqns. (5.5) and (5.6), we get

$$mg = G\frac{mM}{R^2}$$

or

$$g = G\frac{M}{R^2} \qquad (5.7)$$

Remember that the force due to gravity on an object is directed towards the center of the earth. It is this direction that we call vertical. Fig. 5.3 shows vertical directions at different places on the earth. The direction perpendicular to the vertical is called the **horizontal** direction.

Once we know the mass and the radius of the earth, or of any other celestial body such as a planet, the value of g at its surface can be calculated using Eqn. (5.7). On the surface of the earth, the value of g is taken as 9.8 ms^{-2}.

Given the mass and the radius of a satellite or a planet, we can use Eqn. (5.7) to find the acceleration due to the gravitational attraction of that satellite or planet.

Before proceeding further, let us look at Eqn. (5.7) again. The acceleration due to gravity produced in a body is independent of its mass. This means that a heavy ball and a light ball will fall with the same velocity. **If we drop these balls from a certain height at the same time, both would reach the ground simultaneously.**

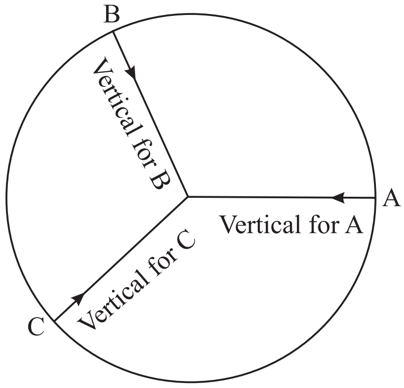


Fig. 5.3 : The vertical direction at any place is the direction towards the centre of earth at that point

ACTIVITY 5.1

Take a piece of paper and a small pebble. Drop them simultaneously from a certain height. Observe the path followed by the two bodies and note the times at which they touch the ground. Then take two pebbles, one heavier than the other. Release them simultaneously from a height and observe the time at which they touch the ground.

Fall Under Gravity

The fact that a heavy pebble falls at the same rate as a light pebble, might appear a bit strange. Till sixteenth century it was a common belief that a heavy body falls faster than a light body. However, the great scientist of the time, Galileo, showed that the two bodies do indeed fall at the same rate. It is said that he went up to the top of the Tower of Pisa and released simultaneously two iron balls of considerably different masses. The balls touched the ground at the same time. But when feather and a stone were made to fall simultaneously, they reached the ground at different times. Galileo argued that the feather fell slower because it experienced greater force of buoyancy due to air. He said that if there were no air, the two bodies would fall together. In recent times, astronauts have performed the feather and stone experiment on the moon and verified that the two fall together. Remember that the moon has no atmosphere and so no air.

Under the influence of gravity, a body falls vertically downwards towards the earth. For small heights above the surface of the earth, the acceleration due to gravity does not change much. Therefore, the equations of motion for initial and final velocities and the distance covered in time t are given by

$$v = u + gt$$

$$s = ut + (\frac{1}{2})gt^2$$

and $$v^2 = u^2 + 2gs. \qquad (5.8)$$

It is important to remember that g **is always directed vertically downwards, no matter what the direction of motion of the body is.** A body falling with an acceleration equal to g is said to be in **free-fall**.

From Eqn. (5.8) it is clear that if a body begins to fall from rest, it would fall a distance $h = (1/2)gt^2$ in time t. So, a simple experiment like dropping a heavy coin from a height and measuring its time of fall with the help of an accurate stop watch could give us the value of g. If you measure the time taken by a five-rupee coin to fall through a distance of 1 m, you will find that the average time of fall for several trials is 0.45 s. From this data, the value of g can be calculated. However, in the laboratory you would determine g by an indirect method, using a simple pendulum.

You must be wondering as to why we take radius of the earth as the distance between the earth and a particle on its surface while calculating the force of gravity on that particle. When we consider two discreet particles or mass points, the separation between them is just the distance between them. But when we calculate gravitational force between extended bodies, what distance do we take into account? To resolve this problem, the concept of **centre of gravity of a body is introduced**. This is a point such that, as far as the gravitational effect is concerned, we may replace the whole body by just this point and the effect would be the same. For geometrically regular bodies of uniform density, such as spheres, cylinders, rectangles, the geometrical center is also the centre of gravity. That is why we choose the center of the earth to measure distances to other bodies. For irregular bodies, there is no easy way to locate their centres of gravity.

Where is the center of gravity of metallic ring located? It should lie at the center the ring. But this point is outside the mass of the body. It means that the centre of gravity of a body may lie outside it. Where is your own centre of gravity located? Assuming that we have a regular shape, it would be at the centre of our body, somewhere beneath the navel.

Later on in this course, you would also learn about the **centre of mass** of a body. This is a point at which the whole mass of the body can be assumed to be concentrated. In a **uniform gravitational field**, the kind we have near the earth, the centre of gravity coincides with the centre of mass.

The use of centre of gravity, or the center of mass, makes our calculations extremely simple. Just imagine the amount of calculations we would have to do if we have to calculate the forces between individual particles a body is made of and then finding the resultant of all these forces.

You should remember that G and g represent different physical quantities. G **is the universal constant of gravitation** which remains the same everywhere, while g is **acceleration due to gravity,** which may change from place to place, as we shall see in the next section.

You may like to answer a few questions to check your progress.

INTEXT QUESTIONS 5.2

1. The mass of the earth is 5.97×10^{24} kg and its mean radius is 6.371×10^{6} m. Calculate the value of g at the surface of the earth.
2. Careful measurements show that the radius of the earth at the equator is 6378 km while at the poles it is 6357 km. Compare values of g at the poles and at the equator.
3. A particle is thrown up. What is the direction of g when (i) the particle is going up, (ii) when it is at the top of its journey, (iii) when it is coming down, and (iv) when it has come back to the ground?
4. The mass of the moon is 7.3×10^{22} kg and its radius is 1.74×10^{6} m. Calculate the gravitational acceleration at its surface.

5.3 VARIATION IN THE VALUE OF *G*

5.3.1 Variation with Height

The quantity R^2 in the denominator on the right hand side of Eqn. (5.7) suggests that **the magnitude of g decreases as square of the distance from the centre of the earth increases**. So, at a distance R from the surface, that is, at a distance $2R$ from the centre of the earth, the value of g becomes (1/4) th of the value of g at the surface. However, if the distance h above the surface of the earth, called **altitude**, is small compared with the radius of the earth, the value of g, denoted by g_h, is given by

$$g_h = \frac{\mathrm{G}M}{(R+h)^2}$$

$$= \frac{GM}{R^2\left(1+\dfrac{h}{R}\right)^2}$$

$$= \frac{g}{\left(1+\frac{h}{R}\right)^2} \qquad (5.9)$$

where $g = GM/R^2$ is the value of acceleration due to gravity at the surface of the earth. Therefore,

$$\frac{g}{g_h} = \left(1+\frac{h}{R}\right)^2 = 1 + \frac{2h}{R} + \left(\frac{h}{R}\right)^2$$

Since (h/R) is a small quantity, $(h/R)^2$ will be a still smaller quantity. So it can be neglected in comparison to (h/R). Thus

$$g_h = \frac{g}{\left(1+\frac{2h}{R}\right)} \qquad (5.10)$$

Let us take an example to understand how we apply this concept.

Example 5.2 : Modern aircrafts fly at heights upward of 10 km. Let us calculate the value of g at an altitude of 10 km. Take the radius of the earth as 6400 km and the value of g on the surface of the earth as 9.8 ms^{-2}.

Solution : From Eqn. (5.8), we have

$$g_h = \frac{g}{\left(1+\frac{2.(10)\ \text{km}}{6400\ \text{km}}\right)} = \frac{9.8\ \text{ms}^{-2}}{1.003} = 9.77\ \text{ms}^{-2}.$$

5.3.2 Variation of g with Depth

Consider a point P at a depth d inside the earth (Fig. 5.4). Let us assume that the earth is a sphere of uniform density ρ. The distance of the point P from the center of the earth is $r = (R - d)$. Draw a sphere of radius $(r - d)$. A mass placed at P will experience gravitational force from particles in (i) the shell of thickness d, and (ii) the sphere of radius r. It can be shown that the forces due to all the particles in the shell cancel each other. That is, the net force on the particle at P due to the matter in the shell is zero. Therefore, in calculating the acceleration due to gravity at P, we have to consider only the

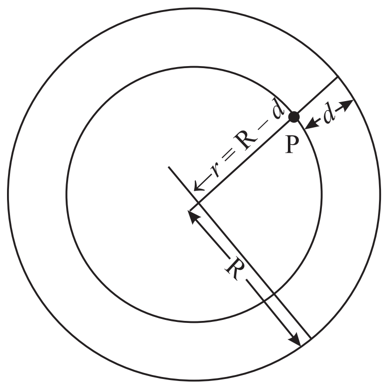


Fig. 5.4 : A point at depth d is at a distance $r = R - d$ from the centre of the earth

mass of the sphere of radius $(r - d)$. The mass M' of the sphere of radius $(r - d)$ is

$$M' = \frac{4\pi}{3} \rho (R - d)^3 \qquad (5.10)$$

The acceleration due to gravity experienced by a particle placed at P is, therefore,

$$g_d = G\frac{M'}{(R-d)^2} = \frac{4\pi G}{3} \rho (R - d) \qquad (5.11)$$

Note that as d increases, $(R - d)$ decreases. This means that **the value of g decreases as we go below the earth**. At d = R, that is, at the centre of the earth, the acceleration due to gravity will vanish. Also note that $(R - d) = r$ is the distance from the centre of the earth. Therefore, acceleration due to gravity is linearly proportional to r. The variation of g from the centre of the earth to distances far from the earth's surface is shown in Fig. 5.5.

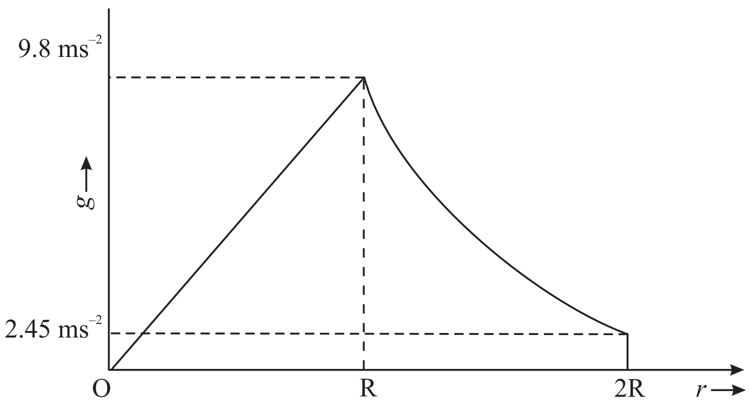


Fig. 5.5 : Variation of g with distance from the centre of the earth

We can express g_d in terms of the value at the surface by realizing that at $d = 0$, we get the surface value: $g = \frac{4\pi G}{3} \rho R$. It is now easy to see that

$$g_d = g\frac{(R-d)}{R} = g\left(1 - \frac{d}{R}\right),\ 0 \le d \le R \qquad (5.12)$$

On the basis of Eqns. (5.9) and (5.12), we can conclude that g **decreases with both height as well as depth.**

Internal Structure of the Earth

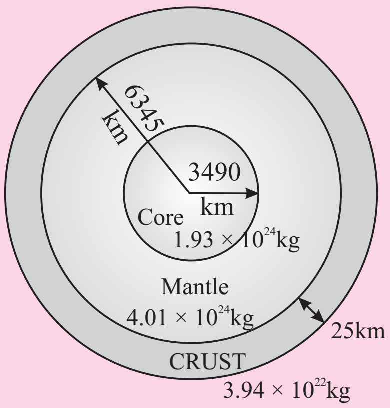


Fig. 5.6 : Structure of the earth (not to scale). Three prominent layers of the earth are shown along with their estimated masses.

Refer to Fig. 5.6 You will note that most of the mass of the earth is concentrated in its core. The top surface layer is very light. For very small depths, there is hardly any decrease in the mass to be taken into account for calculating *g*, while there is a decrease in the radius. So, the value of g increases up to a certain depth and then starts decreasing. It means that assumption about earth being a uniform sphere is not correct.

5.3.3 Variation of *g* with Latitude

You know that the earth rotates about its axis. Due to this, every particle on the earth's surface excecutes circular motion. In the absence of gravity, all these particles would be flying off the earth along the tangents to their circular orbits. Gravity plays an important role in keeping us tied to the earth's surface. You also know that to keep a particle in circular motion, it must be supplied centripetal force. A small part of the gravity force is used in supplying this centripetal force. As a result, the force of attraction of the earth on objects on its surface is slightly reduced. The maximum effect of the rotation of the earth is felt at the equator. At poles, the effect vanishes completely. We now quote the formula for variation in g with latitude without derivation. If g_λ denotes the value of g at latitude λ and g is the value at the poles, then

$$g_\lambda = g - R\omega^2 \cos\lambda, \qquad (5.13)$$

where ω is the angular velocity of the earth and R is its radius. You can easily see that at the poles, $\lambda = 90$ degrees, and hence $g_\lambda = g$.

Example 5.3 : Let us calculate the value of g at the poles.

Solution : The radius of the earth at the poles = 6357 km = 6.357×10^6 m

The mass of the earth = 5.97×10^{24} kg

Using Eqn. (5.7), we get

g at the poles = $[6.67 \times 10^{-11} \times 5.97 \times 10^{24} / (6.357 \times 10^{6})^{2}]$ ms^{-2}

= 9.853 ms^{-2}

Example 5.4 : Now let us calculate the value of g at $\lambda = 60^{\circ}$, where radius of earth is 6371 km.

Solution : The period of rotation of the earth, T = 24 hours = $(24 \times 60 \times 60)$ s

$\therefore$ frequency of the earth's rotation = $1/T$

angular frequency of the earth $\omega = 2\pi/T = 2\pi/(24 \times 60 \times 60)$

$= 7.27 \times 10^{-5}$

$\therefore R\omega^2 \cos\lambda = 6.371 \times 10^{6} \times (7.27 \times 10^{-5})^{2} \times 0.5 =$ 0.017 ms^{-2}

Since $g_0 = g - R\omega^2 \cos\lambda$, we can write

g_λ (at latitude 60 degrees) = 9.853 – 0.017 = 9.836 ms^{-2}

INTEXT QUESTIONS 5.3

1. At what height must we go so that the value of g becomes half of what it is at the surface of the earth?

2. At what depth would the value of g be 80% of what it is on the surface of the earth?

3. The latitude of Delhi is approximately 30 degrees north. Calculate the difference between the values of g at Delhi and at the poles.

4. A satellite orbits the earth at an altitude of 1000 km. Calcultate the acceleration due to gravity acting on the satellite (i) using Eqn. (5.9) and (ii) using the relation g is proportional to $1/r^2$, where r is the distance from the centre of the earth. Which method do you consider better for this case and why?

5.4 WEIGHT AND MASS

The force with which a body is pulled towards the earth is called its weight. If m is the mass of the body, then its weight W is given by

$$W = mg \tag{5.14}$$

Since weight is a force, its unit is newton. If your mass is 50 kg, your weight would be 50 kg $\times$ 9.8 ms^{-2} = 490 N.

Since g varies from place to place, weight of a body also changes from place to place.

The weight is maximum at the poles and minimum at the equator. This is because the radius of the earth is minimum at the poles and maximum at the equator. The weight decreases when we go to higher altitudes or inside the earth.

The mass of a body, however, does not change. Mass is an intrinsic property of a body. Therefore, it stays constant wherever the body may be situated.

Note: In everyday life we often use mass and weight interchangeably. Spring balances, though they measure weight, are marked in kg (and not in N).

5.4.1 Gravitational Potential and Potential energy

The Potential energy of an object under the influence of a conservative force may be defined as the energy stored in the body and is measured by the work done by an external agency in bringing the body from some standard position to the given position.

If a force F displaces a body by a small distance dr aganist the conservative force, without changing its speed, the small change in the potential energy dU is given by,

$$dU = -F.dr$$

In case of gravitational force between two masses M and m separated by a distance r,

$$F = \frac{GMm}{r^2}$$

∴ gravitational potential energy

$$dU = \frac{GMm}{r^2}dr$$

or

$$U = GMm\int_{\infty}^{r}\frac{1}{r^2}dr = -\frac{GMm}{r}$$

It shows that the gravitational potential energy between two particles of masses M and m separated by a distance r is given by

$$U = -\frac{GMm}{r} + \text{ a constant}$$

The gravitational potential energy is zero when r approaches infinity. So the constant is zero and $U = \frac{-GMm}{r}$

Gravitational Potential (V) of mass M is defined as the gravitational potential energy of unit mass. Hence,

Gravitational potential, $V = \frac{U}{m} = -\frac{GM}{r}$

It is a scalar quantity and its SI unit is J/kg.

ACTIVITY 5.2

Calculate the weight of an object of mass 50 kg at distances of 2R, 3R, 4R, 5R and 6R from the centre of the earth. Plot a graph showing the weight against distance. Show on the same graph how the mass of the object varies with distance.

Try the following questions to consolidate your ideas on mass and weight.

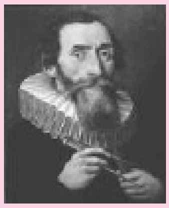
INTEXT QUESTIONS 5.4

1. Suppose you land on the moon. In what way would your weight and mass be affected?
2. Compare your weight at Mars with that on the earth? What happens to your mass? Take the mass of Mars = 6×10^{23} kg and its radius as 4.3×10^{6} m.
3. You must have seen two types of balances for weighing objects. In one case there are two pans. In one pan, we place the object to be weighed and in the other we place weights. The other type is a spring balance. Here the object to be weighed is suspended from the hook at the end of a spring and reading is taken on a scale. Suppose you weigh a bag of potatoes with both the balances and they give the same value. Now you take them to the moon. Would there be any change in the measurements made by the two balances?
4. State the SI unit of Gravitational potential.

5.5 KEPLER'S LAWS OF PLANETARY MOTION

In ancient times it was believed that all heavenly bodies move around the earth. Greek astronomers lent great support to this notion. So strong was the faith in the earth-centred universe that all evidences showing that planets revolved around the Sun were ignored. However, Polish Astronomer Copernicus in the 15th century proposed that all the planets revolved around the Sun. In the 16th century, Galileo, based on his astronomical observations, supported Copernicus. Another European astronomer, Tycho Brahe, collected a lot of observations on the motion of planets. Based on these observations, his assistant Kepler formulated **laws of planetary motion**.

Johannes Kepler



German by birth, Johannes Kepler, started his career in astronomy as an assistant to Tycho Brahe. Tycho religiously collected the data of the positions of various planets on the daily basis for more than 20 years. On his death, the data was passed on to Kepler who spent 16 years to analyse the data. On the basis of his analysis, Kepler arrived at the three laws of planetary motion.

He is considered as the founder of geometrical optics as he was the first person to describe the working of a telescope through its ray diagram.

For his assertion that the earth revolved around the Sun, Galileo came into conflict with the church because the Christian authorities believed that the earth was at the centre of the universe. Although he was silenced, Galileo kept recording his observations quietly, which were made public after his death. Interestingly, Galileo was freed from that blame recently by the present Pope.

Kepler formulated three laws which govern the motion of planets. These are:

1. The orbit of a planet is an ellipse with the Sun at one of the foci (Fig. 5.7). (An ellipse has two foci.)

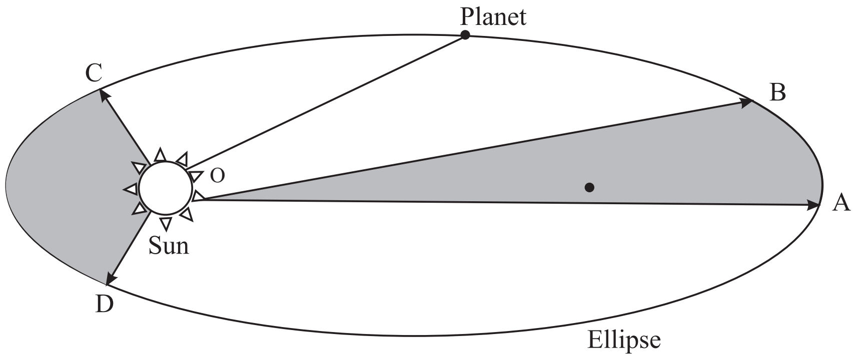


Fig. 5.7 : The path of a planet is an ellipse with the Sun at one of its foci. If the time taken by the planet to move from point A to B is the same as from point C to D, then according to the second law of Kepler, the areas AOB and COD are equal.

2. The area swept by the line joining the planet to the sun in unit time is constant through out the orbit (Fig 5.7)

3. The square of the period of revolution of a planet around the sun is proportional to the cube of its average distance from the Sun. If we denote the period by T and the average distance from the Sun as r, $T^2 \alpha r^3$.

Let us look at the third law a little more carefully. You may recall that Newton used this law to deduce that the force acting between the Sun and the planets varied as $1/r^2$ (Example 5.1). Moreover, if T_1 and T_2 are the orbital periods of two planets and r_1 and r_2 are their mean distances from the Sun, then the third law implies that

$$\frac{T_1^2}{T_2^2} = \frac{r_1^3}{r_2^3} \qquad (5.15)$$

The constant of proportionality cancels out when we divide the relation for one planet by the relation for the second planet. This is a very important relation. For example, it can be used to get T_2, if we know T_1, r_1 and r_2.

Example 5.5 : Calculate the orbital period of planet mercury, if its distance from the Sun is 57.9×10^9 m. You are given that the distance of the earth from the Sun is 1.5×10^{11} m.

Solution : We know that the orbital period of the earth is 365.25 days. So, T_1 = 365.25 days and $r_1 = 1.5 \times 10^{11}$ m. We are told that $r_2 = 57.9\times10^9$ m for mercury. Therefore, the orbital period of mercury is given by T_2

$$\frac{T_2^2}{T_1^2} = \frac{r_2^3}{r_1^3}$$

On substituting the values of various quantities, we get

$$T_2 = \sqrt{\frac{T_1^2 r_2^3}{r_1^3}} = \sqrt{\frac{(365.25)^2 \times (57.9\times10^9)^3\,\text{m}^3}{(1.5\times10^{11})^3\,\text{m}^3}} \text{ days}$$

$$= 87.6 \text{ days.}$$

In the same manner you can find the orbital periods of other planets. The data is given below. You can also check your results with numbers in Table 5.1.

Table 5.1: Some data about the planets of solar system

Name of the planet	Mean distance from the Sun (in terms of the distance of earth)	Radius ($\times10^3$ km)	Mass (Earth Masses)
Mercury	0.387	2.44	0.53
Venus	0.72	6.05	0.815
Earth	1.0	6.38	1.00
Mars	1.52	3.39	0.107
Jupiter	5.2	71.40	317.8
Saturn	9.54	60.00	95.16
Uranus	19.2	25.4	14.50
Neptune	30.1	24.3	17.20
Pluto	39.4	1.50	0.002

Kepler's laws apply to any system where the force binding the system is gravitational in nature. For example, they apply to Jupiter and its satellites. They also apply to the earth and its satellites like the moon and artificial satellites.

Example 5.6 : A satellite has an orbital period equal to one day. (Such satellites are called geosynchronous satellites.) Calculate its height from the earth's surface, given that the distance of the moon from the earth is 60 R_E (R_E is the radius of the earth), and its orbital period is 27.3 days. [This orbital period of the moon is with respect to the fixed stars. With respect to the earth, which itself is in orbit round the Sun, the orbital period of the moon is about 29.5 day.]

Solution : A geostationary satellite has a period T_2 equal to 1 day. For moon T_1 = 27.3 days and r_1 = 60 R_E, T_2 = 1 day. Using Eqn. (5.15), we have

$$r_2 = \left[\frac{r_1^3 T_2^2}{T_1^2}\right]^{1/3} = \left[\frac{(60^3 R_E^3)(1^2 \text{ day}^2)}{27.3^2 \text{ day}^2}\right]^{1/3} = 6.6\ R_E.$$

Remember that the distance of the satellite is taken from the centre of the earth. To find its height from the surface of the earth, we must subtract R_E from 6.6 R_E. The required distance from the earth's surface is 5.6 R_E. If you want to get this distance in km, multiply 5.6 by the radius of the earth in km.

5.5.1 Orbital Velocity of Planets

We have so for talked of orbital periods of planets. If the orbital period of a planet is T and its distance from the Sun is r, then it covers a distance $2\pi r$ in time T. Its orbital velocity is, therefore,

$$v_{orb} = \frac{2\pi r}{T} \qquad (5.16)$$

There is another way also to calculate the orbital velocity. The centripetal force experienced by the planet is mv_{orb}^2/r, where m is its mass. This force must be supplied by the force of gravitation between the Sun and the planet. If M is the mass of the Sun, then the gravitational force on the planet is $\frac{GmM_s}{r^2}$. Equating the two forces, we get

$$\frac{mv_{orb}^2}{r} = \frac{GM_s}{r^2},$$

so that,

$$v_{orb} = \sqrt{\frac{GM_s}{r^2}} \qquad (5.17)$$

Notes

Notice that the mass of the planet does not enter the above equation. The orbital velocity depends only on the distance from the Sun. Note also that if you substitute υ from Eqn. (5.16) in Eqn. (5.17), you get the third law of Kepler.

INTEXT QUESTIONS 5.5

1. Many planetary systems have been discovered in our Galaxy. Would Kepler's laws be applicable to them?
2. Two artificial satellites are orbiting the earth at distances of 1000 km and 2000 km from the surface of the earth. Which one of them has the longer period? If the time period of the former is 90 min, find the time period of the latter.
3. A new small planet, named Sedna, has been discovered recently in the solar system. It is orbiting the Sun at a distance of 86 AU. (An AU is the distance between the Sun and the earth. It is equal to 1.5×10^{11} m.) Calculate its orbital period in years.
4. Obtain an expression for the orbital velocity of a satellite orbiting the earth.
5. Using Eqns. (5.16) and (5.17), obtain Kepler's third law.

5.6 ESCAPE VELOCITY

You now know that a ball thrown upwards always comes back due to the force of gravity. If you throw it with greater force, it goes a little higher but again comes back. If you have a friend with great physical power, ask him to throw the ball upwards. The ball may go higher than what you had managed, but it still comes back. You may then ask: Is it possible for an object to escape the pull of the earth? The answer is 'yes'. The object must acquire what is called the **escape velocity. It is defined as the minimum velocity required by an object to escape the gravitational pull of the earth.**

It is obvious that the escape velocity will depend on the mass of the body it is trying to escape from, because the gravitational pull is proportional to mass. It will also depend on the radius of the body, because smaller the radius, stronger is the gravitational force.

The escape velocity from the earth is given by

$$\upsilon_{esc} = \sqrt{\frac{2GM}{R}} \qquad (5.18)$$

where M is the mass of the earth and R is its radius. For calculating escape velocity from any other planet or heavenly body, mass and radius of that heavenly body will have to be substituted in the above expression.

It is not that the force of gravity ceases to act when an object is launched with escape velocity. The force does act. Both the velocity of the object as well as the force of gravity acting on it decrease as the object goes up. It so happens that the force becomes zero before the velocity becomes zero. Hence the object escapes the pull of gravity.

Try the following questions to grasp the concept.

INTEXT QUESTIONS 5.6

1. The mass of the earth is 5.97×10^{24} kg and its radius is 6371 km. Calculate the escape velocity from the earth.
2. Suppose the earth shrunk suddenly to one-fourth its radius without any change in its mass. What would be the escape velocity then?
3. An imaginary planet X has mass eight times that of the earth and radius twice that of the earth. What would be the escape velocity from this planet in terms of the escape velocity from the earth?

5.7 ARTIFICIAL SATELLITES

A cricket match is played in Sydney in Australia but we can watch it live in India. A game of Tennis played in America is enjoyed in India. Have you ever wondered what makes it possible? All this is made possible by artificial satellites orbiting the earth. You may now ask : How is an artificial satellite put in an orbit?

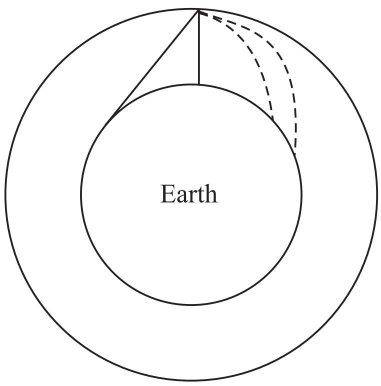


Fig. 5.8 : A projectile to orbit the earth

You have already studied the motion of a projectile. If you project a body at an angle to the horizontal, it follows a parabolic path. Now imagine launching bodies with increasing force. What happens is shown in Fig. 5.8. Projectiles travel larger and larger distances before falling back to the earth. Eventually, the projectile goes into an orbit around the earth. It becomes an **artificial satellite**. Remember that such satellites are man-made and launched with a particular purpose in mind. Satellites like the moon are natural satellites.

In order to put a satellite in orbit, it is first lifted to a height of about 200 km to minimize loss of energy due to friction in the atmosphere of the earth. Then it is given a horizontal push with a velocity of about 8 kms^{-1}.

The orbit of an artificial satellite also obeys Kepler's laws because the controlling force is gravitational force between the satellite and the earth. The orbit is elliptic in nature and its plane always passes through the center of the earth.

Remember that the orbital velocity of an artificial satellite has to be less than the escape velocity; otherwise it will break free of the gravitational field of the earth and will not orbit around the earth. From the expressions for the orbital velocity of a satellite close to the earth and the escape velocity from the earth, we can write

$$\upsilon_{orb} = \frac{\upsilon_{sec}}{\sqrt{2}} \qquad (5.19)$$

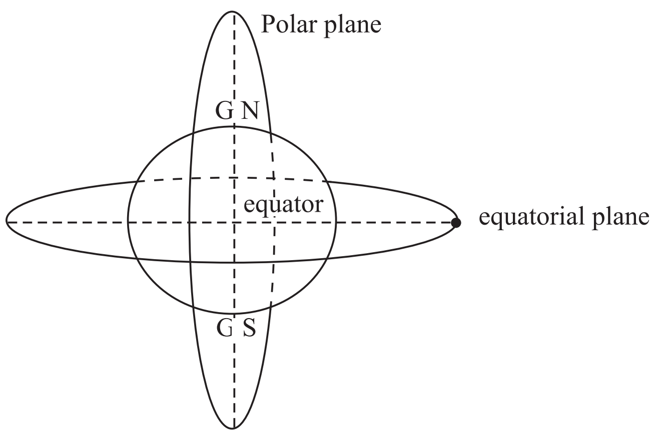


Fig. 5.9: Equitorial and polar orbits

Artificial satellites have generally two types of orbits (Fig. 5.9) depending on the purpose for which the satellite is launched. Satellites used for tasks such as remote sensing have **polar orbits**. The altitude of these orbits is about 800 km. If the orbit is at a height of less than about 300 km, the satellite loses energy because of friction caused by the particles of the atmosphere. As a result, it moves to a lower height where the density is high. There it gets burnt. The time period of polar satellites is around 100 minutes. It is possible to make a polar satellite **sun-synchronous**, so that it arrives at the same latitude at the same time every day. During repeated crossing, the satellite can scan the whole earth as it spins about its axis (Fig. 5.10). Such satellites are used for collecting data for weather prediction, monitoring floods, crops, bushfires, etc.

Satellites used for communications are put in equatorial orbits at high altitudes. Most of these satellites are **geo-synchronous, the ones which have the same orbital period as the period of rotation of the earth**, equal to 24 hours. Their height, as you saw in Example 5.6 is fixed at around 36000 km. Since their orbital period matches that of the earth, they appear to be hovering above the same spot on the earth. A combination of such satellites covers the entire globe, and signals can be sent from any place on the globe to any other place. Since a geo-synchronous satellite observes the same spot on the earth all the time, it can also be used for monitoring any peculiar happening that takes a long time to develop, such as severe storms and hurricanes.

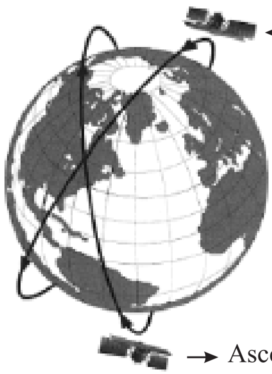


Fig. 5.10: A sun synchromous satellite scanning the earth

Applications of Satellites

Artificial satellites have been very useful to mankind. Following are some of their applications:

1. **Weather Forecasting :** The satellites collect all kinds of data which is useful in forecasting long term and short term weather. The weather chart that you see every day on the television or in newspapers is made from the data sent by these satellites. For a country like India, where so much depends on timely rains, the satellite data is used to watch the onset and progress of monsoon. Apart from weather, satellites can watch unhealthy trends in crops over large areas, can warn us of possible floods, onset and spread of forest fire, etc.

2. **Navigation :** A few satellites together can pinpoint the position of a place on the earth with great accuracy. This is of great help in locating our own position if we have forgotten our way and are lost. Satellites have been used to prepare detailed maps of large chunks of land, which would otherwise take a lot of time and energy.

3. **Telecommunication :** We have already mentioned about the transmission of television programmes from anywhere on the globe to everywhere became possible with satillites. Apart from television signals, telephone and radio signals are also transmitted. The communication revolution brought about by artificial satellites has made the world a small place, which is sometimes called a global village.

4. **Scientific Research :** Satellites can be used to send scientific instruments in space to observe the earth, the moon, comets, planets, the Sun, stars and galaxies. You must have heard of Hubble Space Telescope and Chandra X-Ray Telescope. The advantage of having a telescope in space is that light from distant objects does not have to go through the atmosphere. So there is hardly any reduction in its intensity. For this reason, the pictures taken by Hubble Space Telescope are of much superior quality than those taken by terrestrial telescopes.

   Recently, a group of Europeon scientists have observed an earth like planet out-side our solar system at a distance of 20 light years.

5. **Monitoring Military Activities :** Artificial satellites are used to keep an eye on the enemy troop movement. Almost all countries that can afford cost of these satellites have them.

Vikram Ambalal Sarabhai

Born in a family of industrialists at Ahmedabad, Gujarat, India. Vikram Sarabhai grew to inspire a whole generation of scientists in India. His initial work on time variation of cosmic rays brought him laurels in scientific fraternity. A founder of Physical Research Laboratory, Ahmedabad and a pioneer of space research in India, he was the first to realise the dividends that space research can bring in the fields of communication, education, metrology, remote sensing and geodesy, etc.

5.7.1 Indian Space Research Organization

India is a very large and populous country. Much of the population lives in rural areas and depends heavily on rains, particularly the monsoons. So, weather forecast is an important task that the government has to perform. It has also to meet the communication needs of a vast population. Then much of our area remains unexplored for minerals, oil and gas. Satellite technology offers a cost-effective solution for all these problems. With this in view, the Government of India set up in 1969 the Indian Space Research Organization (ISRO) under the dynamic leadership of Dr. Vikram Sarabhai. Dr. Sarabhai had a vision for using satellitis for educating the nation. ISRO has pursued a very vigorous programme to develop space systems for communication, television broadcasting, meteorological services, remote sensing and scientific research. It has also developed successfully launch vehicles for polar satellites (***PSLV***) (Fig. 5.11) and geo-synchronous satellites (***GSLV***) (Fig. 5.12). In fact, it has launched satellites for other countries like Germany, Belgium and Korea. and has joined the exclusive club of five countries. Its scientific programme includes studies of

(i) climate, environment and global change,

(ii) upper atmosphere,

(iii) astronomy and astrophysics, and

(iv) Indian Ocean.

Recently, ISRO launched an exclusive educational satillite EduSat, first of its kind in the world. It is being used to educate both young and adult students living in remote places.

It is now making preparation for a mission to the moon.

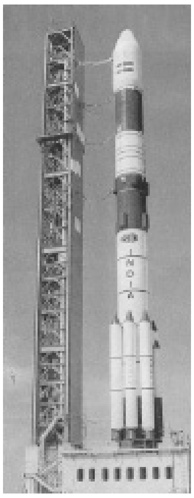


Fig. 5.11: PSLV

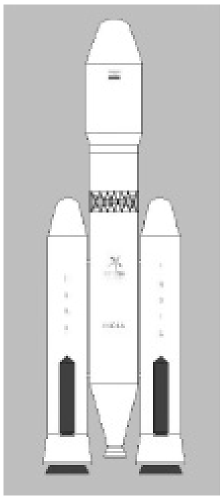


Fig. 5.12: GSLV

INTEXT QUESTIONS 5.7

1. Some science writers believe that some day human beings will establish colonies on the Mars. Suppose people living this desire to put in orbit a Mars synchronous satellite. The rotation period of Mars is 24.6 hours. The mass and radius of Mars are 6.4×10^{23} kg and 3400 km, respectively. What would be the height of the satellite from the surface of Mars?

2. List the advantages of having a telescope in space.

WHAT YOU HAVE LEARNT

- The force of gravitation exists between any two particles in the universe. It varies as the product of their masses and inversely as the square of distance between them.
- The constant of gravitation, G, is a universal constant.
- The force of gravitation of the earth attracts all bodies towards it.
- The acceleration due to gravity near the surface of the earth is 9.8 ms^{-2}. It varies on the surface of the earth because the shape of the earth is not perfectly spherical.
- The acceleration due to gravity varies with height, depth and latitude.
- The weight of a body is the force of gravity acting on it.
- Tthe gravitational potential energy between two particles of masses M and m separated by a distance r given by $U = -\dfrac{GMm}{r}$
- Kepler's first law states that the orbit of a planet is elliptic with sun at one of its foci.
- Kepler's second law states that the line joining the planet with the Sun sweeps equal areas in equal intervals of time.
- Kepler's third law states that the square of the orbital period of a planet is proportional to the cube of its mean distance from the Sun.
- A body can escape the gravitational field of the earth if it can acquire a velocity equal to or greater than the escape velocity.
- The orbital velocity of a satellite depends on its distance from the earth.

Notes

TERMINAL EXERCISE

1. You have learnt that the gravitational attraction is mutual. If that is so, does an apple also attract the earth? If yes, then why does the earth not move in response?
2. We set up an experiment on earth to measure the force of gravitation between two particles placed at a certain distance apart. Suppose the force is of magnitude F. We take the same set up to the moon and perform the experiment again. What would be the magnitude of the force between the two particles there?
3. Suppose the earth expands to twice its size without any change in its mass. What would be your weight if your present weight were 500 N?
4. Suppose the earth loses its gravity suddenly. What would happen to life on this plant?
5. Refer to Fig. 5.6 which shows the structure of the earth. Calculate the values of g at the bottom of the crust (depth 25 km) and at the bottom of the mantle (depth 2855 km).
6. Derive an expression for the mass of the earth, given the orbital period of the moon and the radius of its orbit.
7. Suppose your weight is 500 N on the earth. Calculate your weight on the moon. What would be your mass on the moon?
8. A polar satellite is placed at a height of 800 km from earth's surface. Calculate its orbital period and orbital velocity.

ANSWERS TO INTEXT QUESTIONS

5.1

1. Moon's time period $T = 27.3\text{d}$

$$= 27.3 \times 24 \times 3600 \text{ s}$$

Radius of moon's orbit $R = 3.84 \times 10^8$ m

$$\text{Moon's orbital speed } \upsilon = \frac{2\pi R}{T}$$

$$\text{Centripetal accleration} = \upsilon^2/R$$

$$= \frac{4\pi^2 R^2}{T^2} \cdot \frac{1}{R} = \frac{4\pi^2 R}{T^2}$$

$$= \frac{4\pi^2 \times 3.84 \times 10^8 \text{ m}}{(27.3 \times 24 \times 3600)^2 \text{ s}^2}$$

$$= \frac{4\pi^2 \times 3.84}{(27.3 \times 2.4 \times 3.6)^2} \times 10^{-2} \text{ ms}^{-2}$$

$$= .00272 \text{ ms}^{-2}$$

If we calculate centripetal acceleration on dividing g by 3600, we get the same value :

$$= \frac{9.8}{3600} \text{ms}^{-2}$$

$$= 0.00272 \text{ ms}^{-2}$$

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

   F is force $\therefore G = \frac{\text{Force} \times r^2}{(\text{mass})^2} = \frac{\text{Nm}^2}{\text{kg}^2}$

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

   If $m_1 = 1\text{kg}$, $m_2 = 1\text{kg}$, $r = 1\text{m}$, then F = G

   or G is equal to the force between two masses of 1kg each placed at a distance of 1m from each other

4. (i) $F \,\alpha\, 1/r^2$, if r is doubled, force becomes one-fourth.

   (ii) $F \,\alpha\, m_1 m_2$, if m_1 and m_2 are both doubled then F becomes 4 times.

   (iii) $F \,\alpha\, \frac{m_1\, m_2}{r^2}$,

   if each mass is doubled, and distance is also doubled, then

   F remains unchanged.

5. $F = G\frac{50 \text{ kg} \times 60 \text{ kg}}{1 \text{ m}^2}$; $\quad G = 6.68 \times 10^{-11} \frac{\text{Nm}^2}{\text{kg}^2}$

$$= 6.67 \times 10^{-11} \frac{\text{Nm}^2}{\text{kg}^2} \cdot \frac{3000 \text{ kg}^2}{1 \text{ m}^2}$$

$$= 6.67 \times 10^{-11} \times 3 \times 10^3 \text{ N}$$

$$= 2 \times 10^{-7} \text{ N}$$

5.2

1. $g = \dfrac{GM}{R^2}$

$$= 6.67 \times 10^{-11} \frac{\text{Nm}^2}{\text{kg}^2} \cdot \frac{5.97 \times 10^{24}\,\text{kg}}{(6.371 \times 10^6)^2\,\text{m}^2}$$

$$= \frac{6.97 \times 59.7}{6.371 \times 6.371} \frac{\text{N}}{\text{kg}} = 9.81 \text{ m s}^{-2}$$

2. g at poles

$$g_{\text{pole}} = \frac{GM}{R^2_{\text{pole}}}$$

$$= 6.67 \times 10^{-11} \frac{\text{Nm}^2}{\text{kg}^2} \cdot \frac{5.97 \times 10^{24}\,\text{kg}}{(6.371 \times 10^6)^2\,\text{m}^2}$$

$$= \frac{6.97 \times 59.7}{6.371 \times 6.371} \frac{\text{N}}{\text{kg}} = 9.81 \text{ ms}^{-2}$$

Similarly,

$$g_{\text{eguator}} = \frac{6.97 \times 59.7}{6.378 \times 6.378} \frac{\text{N}}{\text{kg}} = 9.79 \text{ ms}^{-2}$$

3. The value of g is always vertically downwards.

4. $g_{\text{moon}} = 6.67 \times 10^{-11} \dfrac{\text{Nm}^2}{\text{kg}^2} \times \dfrac{7.3 \times 10^{22}\,\text{kg}}{(1.74 \times 10^6)^2\,\text{m}^2}$

$$= \frac{6.67 \times 7.3}{1.74 \times 1.74} \times 10^{-1} \frac{\text{N}}{\text{kg}} = 1.61 \text{ m s}^{-2}$$

5.3

1. Let g at distance r from the centre of the earth be called g_1.

Outside the earth,

then $\dfrac{g}{g_1} = \dfrac{r^2}{\text{R}^2}$

If $g_1 = g/2 \Rightarrow r^2 = 2R^2 \Rightarrow r = \sqrt{2}R = 1.412\ R$

$\therefore$ Height from earth's surface $= 1.4142\ R - R$

$= 0.4142\ R$

2. Inside the earth g varies as distance from the centre of the earth. Suppose at depth d, g is called g_d.

Then $\qquad \dfrac{g_d}{g} = \dfrac{\mathrm{R} - d}{\mathrm{R}}$

If $\qquad g_d = 80\%$, then

$$\frac{0.8}{1} = \frac{\mathrm{R} - d}{\mathrm{R}}$$

$\therefore \qquad d = 0.2\ R$

3. In example 5.3, we calculated $\omega = 7.27 \times 10^{-5}$ rad s^{-1}

$\therefore$ $\mathrm{R}\omega^2 \cos 30° = 6.37 \times 10^6 \times (7.27 \times 10^{-5})^2\ s^{-2} \cdot \sqrt{3}/2 = 0.029\ ms^{-2}$

g at poles $= 9.853$ m s^2

(Calculated in example 5.2)

$\therefore$ g at Delhi $= 9.853\ ms^{-2} - 0.029\ ms^{-2}$

$= 9.824\ ms^{-2}$

4. Using formula (5.9),

$$g_h = \frac{g}{1 + \frac{2h}{R}} = \frac{9.81\ \mathrm{m\,s}^{-2}}{1 + \frac{2000\ \mathrm{km}}{6371\ \mathrm{km}}}$$

$$= \frac{9.81\ \mathrm{m\,s}^{-2}}{\frac{28371\ \mathrm{km}}{6371\ \mathrm{km}}} = 7.47\ \mathrm{m\ s}^{-2}$$

Using variation with r

$$g = \frac{GM}{(R + h)^2}$$

$$= 6.67 \times 10^{-11}\ \frac{\mathrm{Nm}^2}{\mathrm{kg}^2} \cdot \frac{5.97 \times 10^{24}\ \mathrm{kg}}{(7.371 \times 10^6)^2\ \mathrm{m}^2}$$

$= 7.33\ ms^{-2}$

This gives more accurate results because formula (5.9) is for the case $h << R$. In this case h is not $<< R$.

5.4

1. On the moon the value of *g* is only 1/6th that on the earth. So, your weight on moon will become 1/6th of your weight on the earth. The mass, however, remains constant.

2. Mass of Mars = 6×10^{23} kg

   Radius of Mars = 4.3×10^{6} m

$$\therefore \quad g_{\text{Mars}} = G\frac{M}{R^2} = 6.67 \times 10^{-11}\ \frac{\text{Nm}^2}{\text{kg}^2} \cdot \frac{6\times10^{23}\text{kg}}{(4.3\times10^{6})^2\text{m}^2} = 2.16$$

$$\frac{\text{Weight on Mars}}{\text{Weight on Earth}} = \frac{\text{m}\,.\,2.16}{\text{m}\,.\,9.81} = 0.22$$

   So, your weight will become roughly 1/4th that on the earth. Mass remains constant.

3. Balances with two pans actually compare masses because *g* acts on both the pans and gets cancelled. The other type of balance, spring balance, measures weight. The balance with two pans gives the same reading on the moon as on the earth. Spring balance with give weight as 1/6th that on the earth for a bag of potatoes.

4. SI unit of Gravational potential is J/kg.

5.5

1. Yes. Wherever the force between bodies is gravitational, Kepler's laws will hold.

2. According to Kepler's third law

$$\frac{T_1^2}{T_2^2} = \frac{r_1^3}{r_2^3} \quad \text{or} \quad T^2 \propto r^3 \Rightarrow T \propto r^{3/2}$$

   So, the satellite which is farther off has higher period.

   Let $T_1 = 90$ min, $\quad r_1 = 1000 \text{ km} + 6371 \text{ km}$

   $r_2 = 2000 \text{ km} + 6371 \text{ km}$

   [From the centre of the earth]

$$\therefore \quad T_2^2 = \frac{T_1^2 . r_2^3}{r_1^3} = (90 \text{ min})^2 \left(\frac{8371 \text{ km}}{7371 \text{ km}}\right)^3$$

$$T_2 = 108.9 \text{ min}$$

3. According to Kepler's third law

$$\frac{T_{earth}^2}{T_{sedna}^2} = \frac{r_{earth}^3}{r_{sedna}^3} \qquad \text{[Distance from the Sun]}$$

$T_{earth} = 1$ yr., $r_{earth} = 1$ AU

$$T_{sedna}^2 = \frac{(1\text{yr})^2 (86\text{ AU})^3}{(1\text{AU})^3} (86)^3 \text{ yr}^2$$

$\therefore \quad T_{sedna} = 797.5$ yr

4. If υ is the orbital velocity of the satellite of mass m at a distance r from the centre of the earth, then equating centripltal force with the gravitational force, we have

$$\frac{m\upsilon^2}{r} = \frac{GmM}{r^2} \Rightarrow \upsilon = \sqrt{\frac{GM}{r}}$$

where M is the mass of the earth.

5. From Eqs. (5.16) and (5.17),

$$\frac{4\pi^2 r^2}{T^2} = \frac{GM}{r} \Rightarrow T^2 = \frac{4\pi^2 . r^3}{GM.}$$

or $\quad T^2 \alpha r^3$.

5.6

1. $\upsilon_{esc} = \sqrt{\frac{2GM}{R}}$

$$= \sqrt{2 \times 6.67 \times 10^{-11} \frac{\text{Nm}^2}{\text{kg}^2} . \frac{5.97 \times 10^{24} \text{kg}}{6.371 \times 10^6 \text{ m}}}$$

$$= \sqrt{\frac{2 \times 6.67 \times 5.97 \times 10}{6.371}} \; 10^3 \text{ ms}^{-1}$$

$= 11.2 \times 10^3$ ms^{-1} = 11.3 kms^{-1}

2. $\upsilon_{esc} \; \alpha \; \sqrt{\frac{1}{R}}$

If R becomes 1/4th, υ_{esc} becomes double.

3. $\upsilon_{esc} \propto \sqrt{\frac{M}{R}}$

If M becomes eight times, and R twice,

then

$\upsilon_{esc} \propto \sqrt{4}$ or υ_{esc} becomes double.

5.7

1. $(R+h)\dfrac{4\pi^2}{T^2} = \dfrac{GM}{(R+h)^2}$

$$\Rightarrow (R+h)^3 = \frac{GM}{4\pi^2}T^2$$

$$= \frac{6.67\times10^{-11}\times6.4\times10^{23}\times(14.6\times3600)^2}{4\times(3.14)^2}$$

$$= 8370 \times 10^{18} \text{ m}$$

$$R + h = 20300 \text{ km}$$

$$h = 26900 \text{ km}$$

2. (a) Images are clearer

(b) x-ray telescopy etc. also work.

Answers to Terminal Problems

3. 125 N

5. $\square\, g$, 5.5 ms^{-2}

7. Weight $= \dfrac{500}{6}$ N, mass 50 kg on moon as well as on earth

8. $T \,\square\, 1\frac{1}{2}$ h, $\upsilon = 7.47$ km s^{-1}

312en06

6

WORK ENERGY AND POWER

You know that motion of objects arises due to application of force and is described by Newton's laws of motion. You also know how the velocity (speed and direction) of an object changes when a force acts on it. In this lesson, you will learn the concepts of work and energy. Modern society needs large amounts of energy to do many kinds of work. Primitive man used muscular energy to do work. Later, animal energy was harnessed to help people do various kinds of tasks. With the invention of various kinds of machines, the ability to do work increased greatly. Progress of our civilization now critrcally depends the on the availability of usable energy. Energy and work are, therefore, closely linked.

From the above discussion you will appreciate that the rate of doing work improved with newer modes, i.e. as we shifted from humans $\rightarrow$ animals $\rightarrow$ machines to provide necessary force. The rate of doing work is known as **power**.

OBJECTIVES

After studying this lesson, you should be able to:

- *define work done by a force and give unit of work;*
- *calculate the work done by an applied force;*
- *state work-energy theorem;*
- *define power of a system;*
- *calculate the work done by gravity when a mass moves from one point to another;*
- *explain the meaning of energy;*
- *obtain expressions for gravitational potential energy and elastic potential energy;*

- *apply the principle of conservation of energy for physical system; and*
- *apply the laws of conservation of momentum and energy in elastic collisions.*

6.1 WORK

The word 'work' has different meaning for different people. When you study, you do mental work. When a worker carries bricks and cement to higher floors of a building, he is doing physical work against the force of gravity. But in science, work has a definite meaning. The technical meaning of work is not always the same as the common meaning. The work is defined in the following way :

Let us suppose that a **constant force F** acting on an object results in displacement **d** i.e. moves it by a distance *d* along a straight line on a horizontal surface, as shown in Fig. 6.1. *The work done by a force is the product of the magnitude of force component in the direction of displacement and the displacement of this object.*

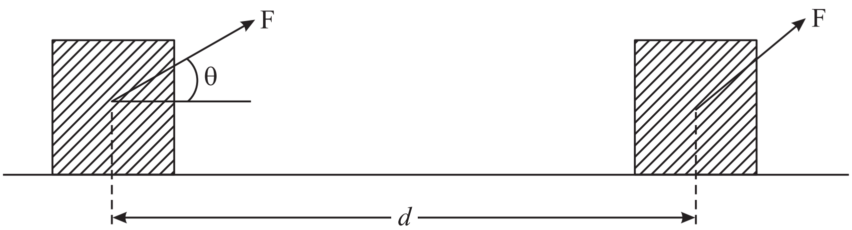


Fig 6.1 : A force F on a block moves it by a horizontal distance d. The direction of force makes an angle θ with the horizontal direction.

If force **F** is acting at angle θ with respect to the displacement **d** of the object, its component along **d** will be $F \cos\theta$. Then work done by force **F** is given by

$$W = F\cos\theta . d \qquad (6.1)$$

In vector form, the work done is given by:

$$W = F.\mathbf{d} \qquad (6.2)$$

Note that if $d = 0$, $W = 0$. That is, no work is done by a force, whatever its magnitude, if there is no displacement of the object. Also note that though both force and displacement are vectors, work is a scalar.

ACTIVITY 6.1

You and your friends may try to push the wall of a room. Irrespective of the applied force, the wall will not move. Thus we say that no work is done.

The unit of work is defined using Eqn.(6.2). If the applied force is in newton and displacement is in metre, then the unit of work is joule.

(Unit of Force)×(Unit of displacement) = newton . metre = Nm (6.3)

This unit is given a special name, **joule,** and is denoted by J.

One joule is defined, as the work done by a force of one newton when it produces a displacement of one metre. Joule is the SI unit of work.

Example 6.1 : Find the dimensional formula of work.

Solution :

$$W = \text{Force} \times \text{Distance}$$

$$= \text{Mass} \times \text{Acceleration} \times \text{distance}$$

$$\text{Dimension of work} = [M] \times [LT^{-2}] \times [L]$$

$$= [ML^2T^{-2}]$$

In electrical measurements, kilowatt-hour (kW h) is used as unit of work. It is related to joule as

$$1\text{kW h} = 3.6 \times 10^6 \text{ J}$$

You will study the details of this unit later in this lesson.

Example 6.2 : A force of 6 N is applied on an object at an angle of 60° with the horizontal. Calculate the work done in moving the object by 2m in the horizontal direction.

Solution : From Eqn. (6.2) we know that

$$W = Fd \cos\theta$$

$$= 6 \times 2 \times \cos 60^\circ$$

$$= 6 \times 2 \times (½)$$

$$= 6 \text{ J}$$

Example 6.3 : A person lifts 5 kg potatoes from the ground floor to a height of 4m to bring it to first floor. Calculate the work done.

Solution : Since the potatoes are lifted, work is being done against gravity. Therefore, we can write

$$\text{Force} = mg$$

$$= 5 \text{ kg} \times 9.8 \text{ m s}^{-2}$$

$$= 49 \text{ N}$$

$$\text{Work done} = 49 \times 4 \text{ (N m)}$$

$$= 196 \text{ J}$$

6.1.1 Positive and Negative Work

As you have seen, work done is defined by Eqn.(6.2), where the angle θ between the force and the displacement is also important. In fact, it leads us to the situation in which work becomes a positive or a negative quantity. Consider the examples given below:

Fig. 6.2 (a) shows a car moving in + x direction and a force F is applied in the same direction. The speed of the car keeps increasing. The force and the displacement both are in the same direction, i.e. $\theta = 0°$. Therefore, the work done is given by

$$W = Fd \cos 0°$$

$$= Fd \qquad (6.4)$$

The work is this case is positive.

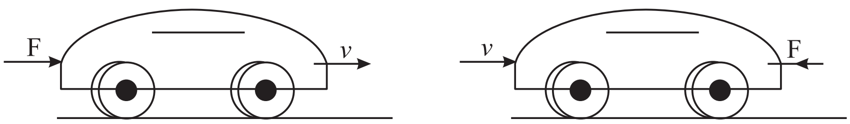


Fig. 6.2 : A car is moving on a horizontal road. a) A force F is applied in the direction of the moving car. It gets accelerated. b) A force F is applied in opposite direction so that the car comes to rest after some distance.

Figure 6.2 (b) shows the same car moving in the +x direction, but the force **F** is applied in the opposite direction to stop the car. Here, angle $\theta = 180°$. Therefore,

$$W = Fd \cos 180°.$$

$$= -Fd \qquad (6.5)$$

Hence, the work done by the force is negative. In fact, the work done by a force shall be negative for θ lying between 90° and 270°.

From the above examples, we can conclude that

a) When we press the accelerator of the car, the force is in the direction of motion of the car. As a result, we increase the speed of the car. The work done is positive.

b) When we apply brakes of a car, the force is applied in a direction opposite to its motion. The car loses speed and may finally come to rest. Negative work is said to have been done.

c) In case the applied force and displacement are as right angles, i.e. $\theta = 90°$, no work is said to be done.

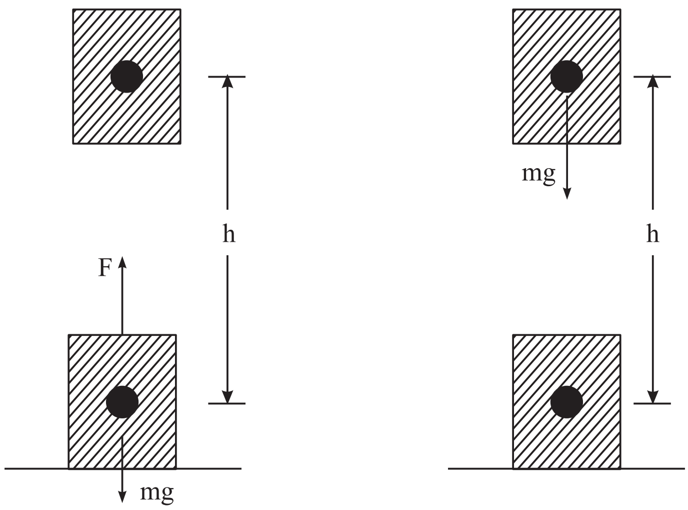
6.1.2 Work Done by the Force of Gravity

Fig.6.3(a) shows a mass *m* being lifted to a height *h* and Fig. 6.3(b) shows the same mass being lowered by a distance *h*. The weight of the object is *mg* in both cases. You may recall from the previous lesson that weight is a force.

In Fig. 6.3 (a), the work is done against the force *mg* (downwards) and the displacement is upward ($\theta = 180°$). Therefore,

$$W = Fd \cos 180°$$

$$= -\, mgh$$



(a) (b)

Fig 6.3 : (a) The object is lifted up against the force of gravity, (b) The object is lowered towards the earth.

In the Fig. 6.3(b), the mass is being lowered. The force *mg* and the displacement *d* are in the same direction ($\theta = 0°$). Therefore, the work done

$$W = Fd \cos 0°$$

$$= +\, mgh \qquad (6.6)$$

You must be very careful in interpreting the results obtained above. When the object is lifted up, the work done by the gravitational force is **negative** but the work done by the person lifting the object is **positive.** When the object is being lowered, the work done by the gravitational force is **positive** but the workdone by the person lowering the object is **negative**. In both of these cases, it is assumed that the object is being moved without acceleration.

INTEXT QUESTIONS 6.1

1. When a particle rotates in a circle, a force acts on the particle. Calculate the work done by this force on the particle.
2. Give one example of each of the following. Work done by a force is
   (a) zero
   (b) negative
   (c) positive
3. A bag of grains of mass 2 kg. is lifted through a height of 5m.
   (a) How much work is done by the lift force?
   (b) How much work is done by the force of gravity?
4. A force $\mathbf{F} = (2\hat{\mathbf{i}} + 3\hat{\mathbf{j}})$ N produces a displacements $d = (-\hat{\mathbf{i}} + 2\hat{\mathbf{j}})$ m. Calculate the work done.
5. A force $\mathbf{F} = (5\hat{\mathbf{i}} + 3\hat{\mathbf{j}})$ acts on a particle to give a displacement $\mathbf{d} = (3\hat{\mathbf{i}} + 4\hat{\mathbf{j}})$m
   (a) Calculate the magnitude of displacement
   (b) Calculate the magnitude of force.
   (c) How much work is done by the force?

6.2 WORK DONE BY A VARIABLE FORCE

You have so far studied the cases where the force acting on the object is constant. This may not always be true. In some cases, the force responsible for doing work may keep varying with time. Let us now consider a case in which the magnitude of force F(x) changes with the position x of the object. Let us now calculate the work done by a variable force. Let us assume that the displacement is from x_i to x_f, where x_i and x_f are the initial and final positions. In such a situation, work is calculated over a large number of small intervals of displacements Δx. In fact, Δx is taken so small that the force F(x) can be assumed to be constant over each such interval. The work done during a small displacements Δx is given by

$$\Delta W = F(x)\,\Delta x \qquad (6.7)$$

$F(x)\,\Delta x$ is numerically equal to the small area shown shaded in the Fig. 6.4(a). The total work done by the force between x_i and x_f is the sum of all such areas (area of all strips added together):

$$W = \Sigma \Delta W$$
$$= \Sigma F(x)\,\Delta x \qquad (6.8)$$

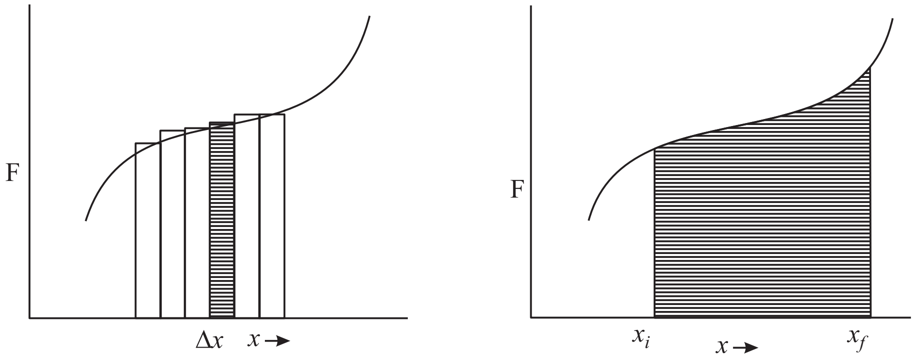


Fig 6.4 : A varying force F moves the object from the initial position x_i to final position x_f. The variation of force with distance is shown by the solid curve (arbitrary) and work done is numerically equal to the shaded area.

The width of the strips can be made as small as possible so that the areas of all strips added together are equal to the total area enclosed between x_i and x_f. It will give the total work done by the force between x_i and x_f :

$$W = \sum_{\lim \Delta x \to 0} F(x)\,\Delta x \qquad (6.9)$$

6.2.1 Work done by a Spring

A very simple example of a variable force is the force exerted by a spring. Let us derive the expression for work done in this case.

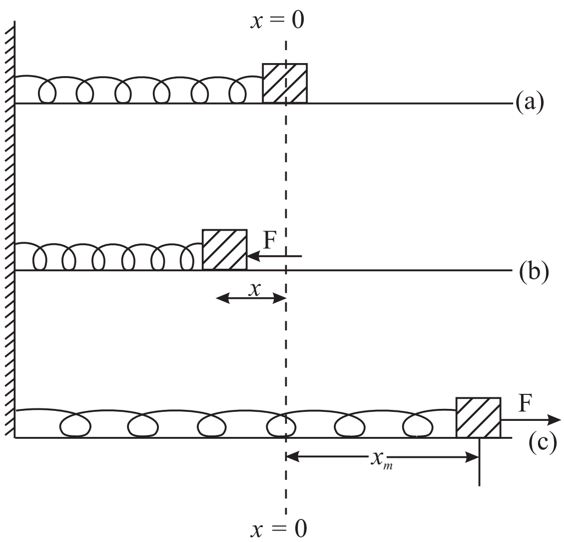


Fig. 6.5 : A spring-mass system whose one end is rigidly fixed and mass m, rests on a smooth horizontal surface. (a) The relaxed position of the spring's, free end at x = 0; (b) The spring is compressed by applying external force F and (c) Pulled or elongated by an external force F. The maximum compression/ elongation is x_m.

Fig. 6.5(a) shows the equilibrium position of a light spring whose one end is attached to a rigid wall and the other end is attached to a block of mass m. The system is placed on a smooth horizontal table. We take x-axis along the horizontal direction. Let mass m be at position $x = 0$. The spring is now compressed (or elongated) by an external force **F.** An internal force $\mathbf{F}_s$ is called into play in the spring due to its elastic property. This force $\mathbf{F}_s$ keeps increasing with increasing x and becomes equal to **F** when the compression (or elongation) is maximum at $x = x_m$.

According to Hooke's law (true for small x only), $|\mathbf{F}_s| = kx$, where k is known as spring constant. Since the direction of $\mathbf{F}_s$ is always opposite to compression (or extension), it is written as :

$$\mathbf{F} = \mathbf{F}_s = -k\mathbf{x} \qquad (6.10)$$

Let us now calculate the work done and also examine, if it is positive or negative. In the event of compression of the spring, the external force **F** is directed towards left and the displacement **x** is also towards left. Hence, the work done by the **external force** is positive. However, for the same direction of displacement, the restoring force generated in the spring is towards right, i.e. **F** and **x** are oppositely directed. The work done by the **spring force** is negative. You can yourself examine the case of extension of the spring and arrive at the same result: "***the work done by the external force is positive but the work done by the spring force is negative and its magnitude is*** (½) kx_m^2"

A simple calculation can be done to derive an expression for the work done. At $x = 0$, the force $\mathbf{F}_s = 0$. As x increases, the force $\mathbf{F}_s$ increases and becomes equal to **F** when $x = x_m$. Since the variation of the force is linear with displacement, the average force during compression (or extension) can be approximated to $\left(\frac{0 + \mathbf{F}_s}{2}\right)$ $= \frac{\mathbf{F}_s}{2}$. The work done by the force is given by

$$W = \text{force . displacement}$$

$$= \frac{\mathbf{F}_s}{2} \cdot \mathbf{x},$$

But $|\mathrm{F}_s| = k\,|\,x_m|$. Hence

$$W = \frac{1}{2} k\, x_m \times x_m$$

$$= \frac{1}{2} k\, x^2{}_m \qquad (6.11)$$

The work done can also be obtained graphically. It is shown in Fig. 6.6.

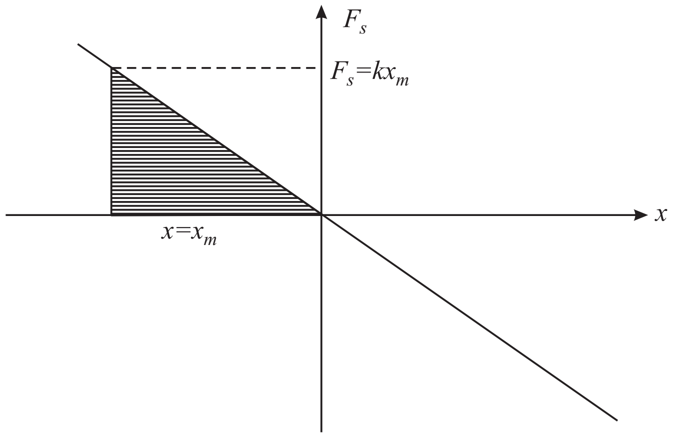


Fig. 6.6: The work done is numerically equal to the area of the shaded triangle.

The area of the shaded triangle is:

$$= \frac{1}{2}\ \text{base} \times \text{height}$$

$$W = \frac{1}{2} x_m \times kx_m$$

$$= \frac{1}{2} kx_m^{\ 2} \qquad (6.12)$$

This is the same as that obtained analytically in Eqn. (6.11)

ACTIVITY 6.2

Measuring spring constant

Suspend the spring vertically, as shown in Fig. 6.7 (a). Now attach a block of mass *m* to the lower end of the spring. On doing so, the spring extends by some distance. Measure the extension. Suppose it is *s*, as shown in Fig 6.7 (b). Now think why does the spring not extend further. This is because the spring force (restoring force) acting upwards balances the weight *mg* of the block in equilibrium state. You can calculate the spring constant by putting the values in

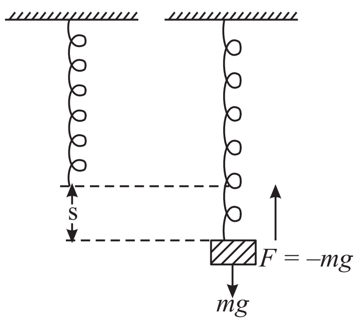


Fig. 6.7 : Extension in a spring under a load.

$$F_s = k.s$$

or $$mg = k.s$$

Thus, $$k = \frac{mg}{s} \qquad (6.13)$$

Example 6.4: A mass of 2 kg is attached to a light spring of force constant $k = 100\ \text{Nm}^{-1}$. Calculate the work done by an external force in stretching the spring by 10 cm.

Solution:

$$W = \frac{1}{2}kx^2$$

$$= \frac{1}{2} \times 100 \times (0.1)^2$$

$$= 50 \times 0.01 = 0.5\ \text{J}$$

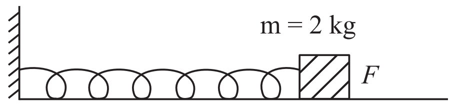


Fig. 6.8: A mass m = 2 kg is attached to a spring on a horizontal surface.

As explained earlier, the work doen by the restioning force in the spring = – 0.5 J.

INTEXT QUESTIONS 6.2

1. Define spring constant. Give its SI unit.
2. A force of 10 N extends a spring by 1cm.How much force is needed to extend this spring by 5 cm? How much work will be done by this force?

6.3 POWER

You have already learnt to calculate the work done by a force. In such calculations, we did not consider whether the work is done in one second or in one hour. In our daily life, however, the time taken to perform a particular work is important. For example, a man may take several hours to load a truck with cement bags, whereas a machine may do this work in much less time. Therefore, it is important to know the rate at which work is done. **The rate at which work is done is called power.**

If ΔW work is done in time Δt, the average power is defined as

$$\text{Average Power} = \frac{\text{Work done}}{\text{time taken}}$$

Mathematically, we can write

$$P = \frac{\Delta W}{\Delta t} \qquad (6.14)$$

If the rate of doing work is not constant, this rate may vary. In such cases, we may define instantaneous power P

$$P = \underset{\Delta t \to 0}{\text{limit}} \left(\frac{\Delta W}{\Delta t}\right) = \frac{dW}{dt} \qquad (6.15)$$

The definition of power helps us to determine the SI unit of power:

$$P = \frac{\Delta W}{\Delta t}$$

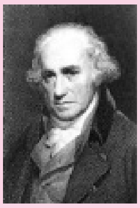
$$= \text{joule/ second} = \text{watt}$$

Thus, the SI unit of power is watt. It is abbreviated asW.

The power of an agent doing work is 1W, if one joule of work is done by it in one second. The more common units of power are kilowatt (kW) and megawatt (MW).

$$1 \text{ kW} = 10^3 \text{ W}, \quad \text{and} \quad 1 \text{ MW} = 10^6 \text{W}$$

James Watt
(1736–1819)



Scottish inventor and mechanical engineer, James Watt is renowned for improving the efficiency of a steam engine. This paved the way for industrial revolution.

He, introduced horse power as the unit of power. SI unit of power watt is named in his honour. Some of the important inventions by James Watt are : a steam locomotive and an attachment that adapted telescope to measure distances.

Example 6.5 : Determine the dimensions of power.

Solution : Since

$$P = \frac{\text{Work}}{\text{Time}}$$

$$= \text{Force} \times \frac{\text{Distance}}{\text{Time}}$$

$$\therefore \text{Dimension of } P = [\text{Mass}] \times [\text{Acceleration}] \times \frac{[\text{Distance}]}{[\text{Time}]}$$

$$= [\text{M}] \times \left[\frac{\text{L}}{\text{T}^2}\right] \times \left[\frac{\text{L}}{\text{T}}\right]$$

$$= [\text{ML}^2\text{T}^{-3}]$$

You may have heard electricians discussing the power of a machine in terms of the horse power, abbreviated as hp. This unit of power was under British system. It is a larger unit:

$$1\text{hp} = 746 \text{ W} \tag{6.16}$$

The unit of power is used to define a new unit of work (energy). One such unit of work is **kilowatt hour.** This unit is commonly used in electrical measurement.

$$\text{kilowatt. hour (kWh)} = 1 \text{ kW} \times 1 \text{ hour}$$

$$= 10^3 \text{ W} \times 3600 \text{ s}$$

$$= \frac{10^3 \text{ J}}{1\text{s}} \times 3600 \text{ s}$$

$$= 36,00,000 \text{ J} = 3.6 \times 10^6 \text{ J}$$

Or $$1 \text{ kW h} = 3.6 \text{ MJ (mega joules)} \tag{6.17}$$

The electrical energy that is consumed in homes is measured in kilowatt-hour. In common man's language : 1kW h = 1 Unit of electrical energy consumption.

INTEXT QUESTIONS 6.3

1. A body of mass 100 kg is lifted through a distance of 8 m in 10 s. Calculate the power of the lifter.
2. Convert 10 horse power into kilowatt.

6.4 WORK AND KINETIC ENERGY

As you know, the capacity to do work is called energy. If a system (object) has energy, it has ability to do work. An automobile moving on a road uses chemical energy of fuel (CNG, petrol, diesel). It can push an object which comes on its way to some distance. Thus it can do work. All moving objects possess energy because they can do work before they come to rest. We call this kind of energy as **kinetic energy**. Kinetic energy is the energy of an object because of its motion.

Let us consider an object of mass m moving along a straight line when a constant force of magnitude F acts on it along the direction of motion. This force produces a uniform acceleration a such that $F = ma$. Let υ_1 be the speed of the object at time t_1. This speed becomes υ_2 at another instant of time t_2. During this interval of time $t = (t_2 - t_1)$, the object covers a distance, s. Using Equations of Motion, we can write

$$\upsilon_2^2 = \upsilon_1^2 + 2as$$

or $$a = \frac{\upsilon_2^2 - \upsilon_1^2}{2s} \qquad (6.18)$$

Combining this result with Newton's second law of motion, we can write

$$F = m \times \frac{\upsilon_2^2 - \upsilon_1^2}{2s}$$

We know that work done by the force is given by

$$W = Fs$$

Hence, $$W = m \times \frac{\upsilon_2^2 - \upsilon_1^2}{2s} s$$

$$= \frac{1}{2} m\upsilon_2^2 - \frac{1}{2} m\upsilon_1^2$$

$$= K_2 - K_1 \qquad (6.19)$$

where $K_2 = \frac{1}{2} m\upsilon_2^2$ and $K_1 = \frac{1}{2} m\upsilon_1^2$ respectively denote the final and initial kinetic energies.

$(K_2 - K_1)$ denotes the change in kinetic energy, which is equal to the work done by the force.

Kinetic Energy is a scalar quantity. It depends on the product of mass and the square of the speed. It does not matter which one of the two (m and υ) is small and which one is large. It is the total value $\frac{1}{2} m\upsilon^2$ that determines the kinetic energy.

Work-Energy Theorem

The work-energy theorem states that the work done by the resultant of all forces acting on a body is equal to the change in kinetic energy of the body.

Example 6.6 : A body of mass 10 kg is initially moving with a speed of 4.0 m s^{-1}. A force of 30 N is now applied on the body for 2 seconds.

(i) What is the final speed of the body after 2 seconds?

(ii) How much work has been done during this period?

(iii) What is the initial kinetic energy?

(iv) What is the final kinetic energy?

(v) What is the distance covered during this period?

(vi) Show that the work done is equal to the change in kinetic energy?

Solution :

(i) Force $(F) = ma$

or $a = F/m$

$= 30/10$

$= 3 \text{ m s}^{-2}$

The final speed $v_2 = v_1 + at$

$= 4 + (3 \times 2) = 10 \text{ m s}^{-1}$

(ii) The distance covered in 2 seconds:

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

$$= (4\times2) + \frac{1}{2}(3\times4)$$

$= 8 + 6 = 14 \text{ m}$

Work done $W = F \times S$

$= 30 \times 14 = 420 \text{ J}$

(iii) The initial Kinetic Energy

$$K_1 = \frac{1}{2}mv_1^2$$

$$= \frac{1}{2}(10\times16) = 80 \text{ J}$$

(iv) The final kinetic energy

$$K_2 = \frac{1}{2}mv_2^2$$

$$= \frac{1}{2}(10\times100) = 500 \text{ J}$$

(v) The distance covered as calculated above = 14m

(vi) The change in kinetic energy is:

$$K_2 - K_1 = (500 - 80) = 420 \text{ J}$$

As may be seen, this is same as wok done.

INTEXT QUESTION 6.4

1. Is it possible for a particle to have a negative value of kinetic energy? Why?
2. What happens to the kinetic energy of a particle if
   (a) The speed υ of the particle is made 2υ.
   (b) The mass m of the particle is made $m/2$?
3. A particle moving with a kinetic energy 3.6 J collides with a spring of force constant 180 N m^{-1}. Calculate the maximum compression of the spring.
4. A car of mass 1000 kg is moving at a speed of 90 km h^{-1}. Brakes are applied and the car stops at a distance of 15 m from the braking point. What is the average force applied by brakes? If the car stops in 25 s after braking , calculate the average power of the brakes?
5. If an external force does 375 J of work in compressing a spring, how much work is done by the spring itself?

6.5 POTENTIAL ENERGY

In the previous section we have discussed that a moving object has kinetic energy associated with it. Objects possess another kind of energy due to their position in space. This energy is known as **Potential Energy.** Familiar example is the **Gravitational Potential Energy possessed by a body in Gravitational Field**. Let us understand it now.

6.5.1 Potential Energy in Gravitational Field

Suppose that a person lifts a mass m from a given height h_1 to a height h_2 above the earth's surface. Let us also assume that the value of acceleration due to gravity remains constant. The mass has been displaced by a distance $h = (h_2 - h_1)$ against the force of gravity. The magnifude of this force is mg and it acts downwards. Therefore, the work done by the person is

$$W = \text{force} \times \text{distance}$$

$$= mgh \qquad (6.20)$$

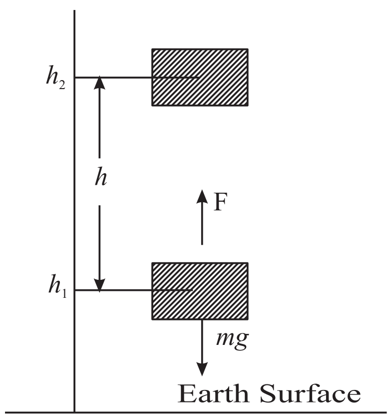


Fig. 6.9 : Object of mass m originally at height h_1 above the earth's surface is moved to a height h_2.

The work is positive and is stored in mass m as energy. This energy by virtue of the position in space is called **gravitational potential energy**. It has capacity to do work. If this mass is left free, it will fall down and during the fall it can be made to do work. For example, it can lift another mass if properly connected by a string, which is passing over a pulley.

The selection of the initial height h_1 is arbitrary. The important concept is the change in height, i.e. $(h_2 - h_1)$. We, therefore, say that the point of **zero potential energy** is arbitrary. Any point in space can be chosen as a point of zero potential energy. Normally, a point on the surface of the earth is assumed to be the reference point with zero potential energy.

Example 6.7 : A truck is loaded with sugar bags. The total mass of the load and the truck together is 100,000 kg. The truck moves on a winding path up a mountain to a height of 700 m in 1 hour. What average power must the engine produce to lift the material?

Solution :

$$W = mgh$$

$$= (100{,}000 \text{ kg}) \times (9.8 \text{ m s}^{-2} \times 700 \text{ m})$$

$$= 9.8 \times 7 \times 10^7 \text{ J}$$

$$= 68.6 \times 10^7 \text{J}$$

$$\text{Time taken} = 1 \text{ hour} = 60 \times 60 \text{ s}$$

$$= 3600 \text{ s}$$

$$\text{Average Power, } P = W/t$$

$$= \frac{68.6 \times 10^7 \text{ J}}{3600 \text{ s}}$$

$$= 1.91 \times 10^5 \text{ W}$$

We know that 746 W = 1 hp

$$\therefore \qquad P = \frac{1.91 \times 10^5}{746} = 2.56 \times 10^2 = 256 \text{ hp.}$$

Example 6.8 : Hydroelectric power generation uses falling water as a source of energy to turn turbine blades and generate electrical power. In a power station, 1000×10^3 kg water falls through a height of 51 m in one second.

(i) Calculate the work done by the falling water?

(ii) How much power can be generated under ideal conditions?

Solution :

(i) The potential energy of the water at the top = mgh

$$\text{P.E.} = (1000 \times 10^3 \text{ kg}) \times (9.8 \text{ m s}^{-2}) \times (51 \text{ m}).$$

$$= 9.8 \times 51 \times 10^6 \text{ J}$$
$$= 500 \times 10^6 \text{ J}$$

Water loses all its potential energy. The same is converted into work in moving the turbine blades. Therefore

$$W = \text{Force} \times \text{distance}$$
$$= mg \times h$$
$$= 1000 \times 10^3 \times 9.8 \times 51 \text{ J}$$
$$= 500 \times 10^6 \text{ J}$$
$$= 500 \text{ M J}$$

(ii) The work done per second is given by

$$P = W/t$$
$$= \frac{500 \text{ M J}}{1 \text{ s}}$$
$$= 500 \text{ MW}$$

Ideal conditions mean that there is no loss of energy due to frictional forces. In practice, there is the always some loss in machines. Such losses can be minimized but can never be eliminated.

6.5.2 Potential energy of springs

You now know that an external force is required to compress or stretch a given spring. These situations are shown in Fig. 6.5. Let there be a spring of force constant k. This spring is compressed by a distance x. From Eqn.(6.11) we recall that work done by the external force to compress the spring is given by

$$W = \frac{1}{2}kx^2$$

This work is stored in the spring as elastic potential energy. When the spring is left free, it bounces back and the elastic potential energy of the spring is converted into kinetic energy of the mass m.

6.5.3 Conservation of Energy

We see around us various forms of energy but we are familiar with some forms more than others. Examples are Electrical Energy, Thermal Energy, Gravitational Energy, Chemical Energy and Nuclear Energy etc. These forms of energy are very closely related in the sense that one can be changed to another. There is a very fundamental law about energy. It is known as **Law of Conservation of Energy**. It states, " **The total energy of an isolated system always remains constant.**" The energy may change its form. It can be converted from one form to other. But the total energy of the system remains unchanged. In an isolated system, if there is any loss of energy of one form, there is a gain of an equal amount of another form of energy. Thus energy is neither created nor destroyed. The universe is also an isolated system as there is nothing beyond this. It is therefore said that the total energy of the universe always remains constant in spite of the fact that variety of changes are taking place in the universe every moment. It is a law of great importance. It has led to many new discoveries in science and it has not been found to fail.

In a Thermal Power Station, the chemical energy of coal is changed into electrical energy. The electrical energy runs machines. In these machines, the electrical energy changes into mechanical energy, light energy or thermal energy.

The law of conservation of energy is more general than we can think of. It applies to systems ranging from big planets and stars to the smallest nuclear particles.

(a) Conservation of mechanical energy during the free fall of a body

We now wish to test the validity of the law of conservation of energy in case of mechanical energy, which is of immediate interest.

Let us suppose that an object of mass m lying on the ground is lifted to a height h. The work done is mgh, which is stored in the object as potential energy. This object is now allowed to fall freely. Let us calculate the energy of this object when it has fallen through a distance h_1. The height of the object now above the earth surface is $h_2 = h - h_1$ (Fig 6. 10). At this point P, the potential energy $= mgh_2$.

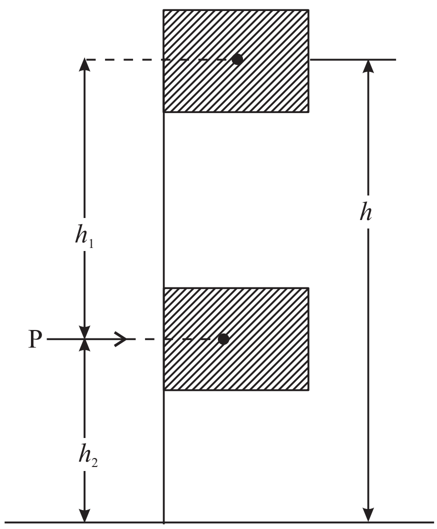


Fig. 6.10 : Mass m is lifted to a height h from earth's surface. It is then lowered to a height h_2 at point P. The total energy at P is same as that at the highest point.

When the object falls freely, it gets accelerated and gains in speed. We can calculate the speed of the object when it has fallen through a height h_1 from the top positions using the equation

$$v^2 = u^2 + 2gs \qquad (6.21)$$

where u is the initial speed at the height h_1, i.e. $u = 0$ and $s = h_1$. Then, we have

$$v^2 = 2gh_1$$

The kinetic energy at point P is given by

$$\text{K.E} = \frac{1}{2}mv^2$$

$$= \frac{m}{2} \times 2gh_1$$

$$= mgh_1 \qquad (6.22)$$

The total energy at the point P is

Kinetic Energy + Potential Energy = $mgh_1 + mgh_2$

$$= mgh \qquad (6.23)$$

This is same as the potential energy at the highest point. **Thus, the total Energy is conserved.**

(b) Conservation of Mechanical Energy for a Mass Oscillating on a Spring

Fig. 6.11 shows a spring whose one end is fixed to a rigid wall and the other end is connected to a wooden block lying on a smooth horizontal table. This free end is at x_0 in the relaxed position of the spring. A block of mass m moving with speed v along the line of the spring collides with the spring at the free end, and compresses it by x_m. This is the maximum compression. At x_0, the total energy of the spring-mass system is $\frac{1}{2}mv^2$. It is the kinetic energy of the mass. The potential energy of the spring is zero. At the point of extreme compression, the potential energy of the spring is $\frac{1}{2}k\,x_m^2$ and the kinetic energy of the mass is zero. The total energy now is $\frac{1}{2}k\,x_m^2$. Obviously, this means that

$$\frac{1}{2}k\,x_m^2 = \frac{1}{2}mv^2 \qquad (6.24)$$

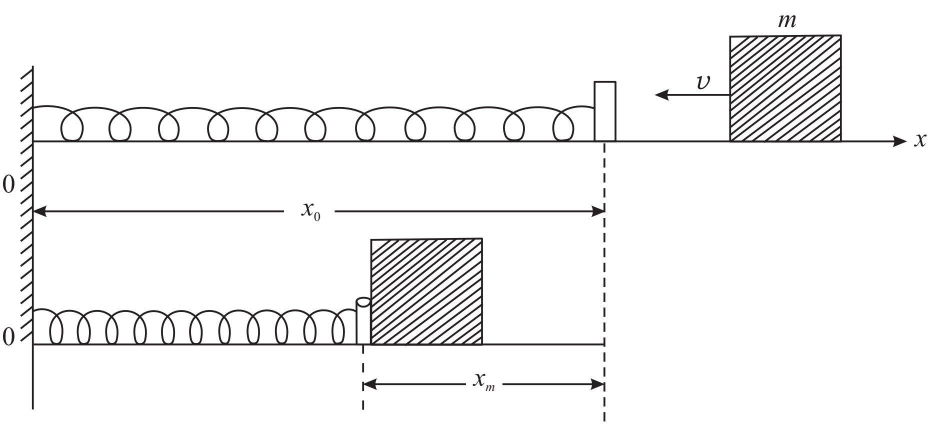


Fig. 6.11 : A block of mass m moving with velocity v on a horizontal surface collides with the spring. The maximum compression is x_m.

K.E + P.E (Before collision) = K.E. + P.E. (After collision)

$$\frac{1}{2}mv^2 + 0 = 0 + \frac{1}{2}kx_m^2 \quad (6.25)$$

i.e., *the total energy is conserved.*

Conservation of mass-energy in nuclear reactions

Nuclear energy is different from other forms of energy in the sense that it is not obtained by the transformation of some other form of energy. On the contrary, it is obtained by transformation of mass into energy.

Hence, in nuclear reacions, the law of conservation of mass and the law of conservation of energy merge into a single law of conservation of mass-energy.

Example 6.9 : A block of mass 0.5 kg slides down a smooth curved surface and falls through a vertical height of 2.5m to reach a horizontal surface at B (Fig 6.12). On the basis of energy conservtion, calculate, i) the energy of the block at point A, and ii) the speed of the block at point B.

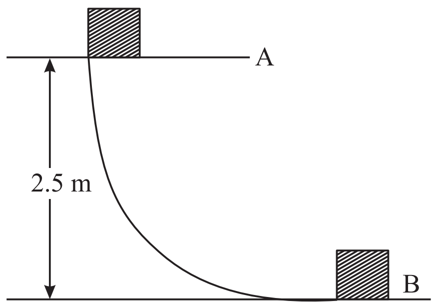


Fig. 6.12 : A block slides on a curved surface. The total energy at A (Potential only) gets converted into total energy at B (kinetic only).

Solution :

i) Potential energy at

A = mgh = (0.5) × (9.8) × 2.5 J

= 4.9 × 2.5 J

= 12.25 J

The kinetic energy at A = 0 and

Total Energy = 12.25 J

ii) The total energy of the block at A must be the same as the total energy at B.

The total energy (P.E. + K.E.) at A = 12.25 J

The total energy (P.E. + K.E.) at B = $\frac{1}{2}mv^2$

Since P.E. at B is zero, the total energy is only K.E.

$$\therefore \quad \frac{1}{2}mv^2 = 12.25$$

$$v^2 = \frac{12.25 \times 2}{0.5}$$

$$= 12.25 \times 4$$

$$v^2 = 49.00$$

Hence $\quad v = 7.0 \text{ ms}^{-1}$

Note: This can also be obtained from the equations of motion:

$$v^2 = v_0^2 + 2gx$$

$$= 0 + 2 \times 9.8 \times 2.5$$

$$v^2 = 49$$

$$v = 7 \text{ ms}^{-1}$$

6.5.4 Conservative and dissipative (Non conservative) Forces

(a) Conservative forces

We have seen that the work done by the gravitational force acting on an object depends on the product of the weight of the object and its vertical displacement. If an object is moved from a point A to a point B under gravity, (Fig 6.13), the work done by gravity depends on the vertical separation between the two points. It does not depend on the path followed to reach B starting from A. When a force obeys this rule, it is called **a conservative force**. Some of the examples of conservative forces are gravitational force, elastic force and electrostatic force.

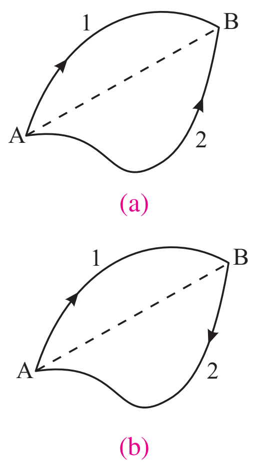


Fig. 6.13 : a) The object is moved from A to B along two different paths. b) It is taken from A to B along path 1 and brought back to A along path 2.

A conservative force has a property that *the work done by a conservative force is independent of path. In* Fig 6.13 (*a*)

$$W_{AB} \text{ (along 1)} = W_{AB} \text{ (along 2)}$$

Fig. 6.13 (b) shows the same two positions of the object. The object moves from A to B along the path 1 and returns back to A along the path 2. By definition, the work done by a conservative force along path 1 is equal and opposite to the work done along the path 2.

$$W_{AB} \text{ (along 1)} = -W_{BA} \text{ (along 2)}$$

or
$$W_{AB} + W_{BA} = 0 \qquad (6.27)$$

This result brings out an important property of the conservative force in ***that the work done by a conservative force on an object is zero when the object moves around a closed path and returns back to its starting point.***

(b) Non-conservative Forces

The force of friction is a good example of a non-conservative force. Fig. 6.14 shows a rough horizontal surface. A block of mass m is moving on this surface with a speed υ at the point A.

After moving a certain distance along a straight line, the block stops at the point B. The block had a kinetic energy $E = \frac{1}{2} m\upsilon^2$ at the point A. It has neither kinetic energy nor potential energy at the point B. It has lost all its energy. Do you know where did the energy go? It has changed its form. Work has been done against the frictional force or we can say that force of friction has done negative work on the block. The kinetic energy has changed to thermal energy of the system. The block with the same kinetic energy E is now taken from A to B through a longer path 2. It may not even reach the point B. It may stop much before reaching B. This obviously means that more work has to be done along this path. Thus, it canbe said that the work done depends on the path.

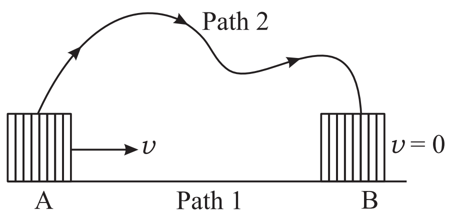


Fig. 6.14: A block which is given an initial speed v on a rough horizontal surface, moves along a straight line path 1 and comes to rest at B. It starts with the same speed υ at A but now moves along a different path 2.

INTEXT QUESTIONS 6.5

1. ABC is a triangle where AB is horizontal and BC is vertical. The length of the sides AB = 3m, BC = 4m and AC=5m. A block of mass 2 kg is at A. What is the change in potential energy of the block when

   a) it is taken from A to B

   b) from B to C

   c) from C to A

   d) How much work is done by gravitational force in moving the mass form B to C (positive or Negative work)?

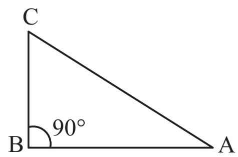


Fig. 6.15

Notes

2. A ball of mass 0.5 kg is at A at a height of 10m above the ground. Solve the following questions by applying work-energy principle. In free fall
   a) What is the speed of the ball at B?
   b) What is the speed of the ball at the point C?
   c) How much work is done by gravitational force in bringing the ball from A to C (give proper sign)?

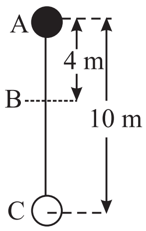


Fig. 6.16

3. A block at the top of an inclined plane slides down. The length of the plane BC = 2m and it makes an angle of 30° with horizontal. The mass of the block is 2 kg. The kinetic energy of the block at the point B is 15.6 J. How much of the potential energy is lost due to non-conservative forces (friction). How much is the magnitude of the frictional force?

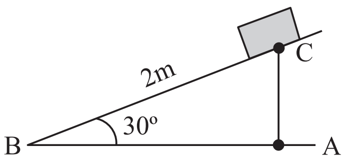


Fig. 6.17

4. The Figure shows two curves A and B between energy E and displacement x of the bob of a simple pendulum. Which one represents the P.E. of the bob and why?

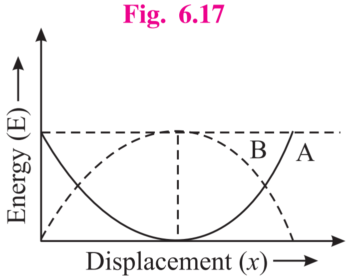


Fig. 6.18

5. When non- conservative forces work on a system, does the total mechanical energy remain constant?

6.6 ELASTIC AND INELASTIC COLLISIONS

Let us consider a system of two bodies. The system is a closed system which implies that no external force acts on it. The system may consist of two balls or two springs or one ball and one spring and so on. When two bodies interact, it is termed as ***collision***. There are no external forces acting on the system.

Let us start with a collision of two balls and to make the analysis simpler, let there be a "**head-on**" or "**central collision**". In such collisions, colliding bodies move along the line joining their centres. The collisions are of two kinds :

(i) **Perfectly Elastic Collision:** If the forces of interaction between the two bodies are **conservative,** the total kinetic energy is conserved i.e. the total kinetic energy before collision is same as that after the collision. Such collisions are termed as **completely elastic collisions**.

(ii) **Perfectly Inelastic collision:** When two colliding bodies stick together after the collision and move as one single unit, it is termed as **perfectly inelastic collision.** It is like motion of a bullet embedded in a target.

You should remember that the momentum is conserved in all types of collisions. Why? But kinetic energy is conserved in elastic collisions only.

6.6.1 Elastic Collision (Head-on)

Let two balls A and B having masses m_A and m_B respectively collide "head-on", as shown in Fig. (6.19). Let v_{Ai} and v_{Bi} be the velocities of the two balls before collision and v_{Af} and v_{Bf} be their velocities after the collision .

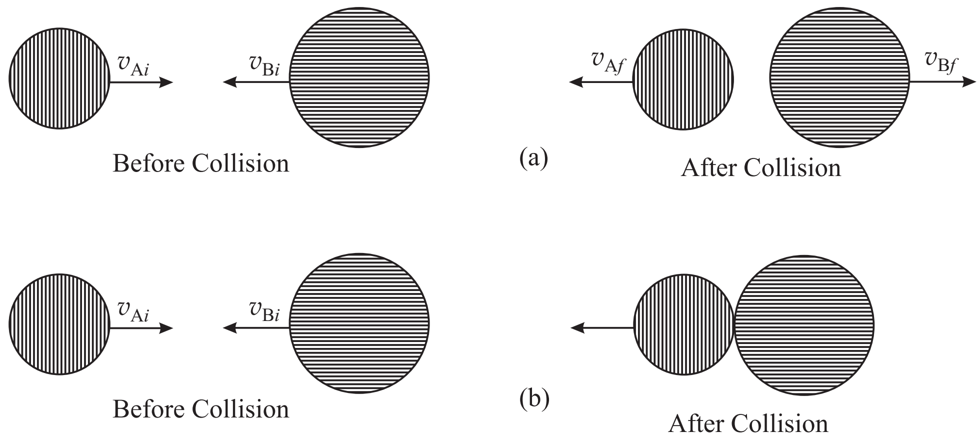


Fig. 6.19 : Schematic representation of Head-on collision (a) Elastic collision; (b) In elastic collision

Now applying the laws of conservation of momentum and kinetic energy, we get

For conservation of momentum

$$m_A v_{Ai} + m_A v_{Bi} = m_A v_{Af} + m_B v_{Bf} \qquad (6.28)$$

For conservation of kinetic energy

$$\frac{1}{2} m_A v_{Ai}^2 + \frac{1}{2} m_{Bi} = \frac{1}{2} m_A v_{Af}^2 + \frac{1}{2} m_B v_{Bf}^2 \qquad (6.29)$$

There are only two unknown quantities (velocities of the balls after collision) and there are two independent equations [Eqns. (6.28) and (6.29)]. The solution is not difficult, but a lengthy one. Therefore, we quote the results only

$$(v_{Bf} - v_{Af}) = -(v_{Bi} - v_{Ai}) \qquad (6.30)$$

$$v_{Af} = \frac{2m_B v_{Bi}}{m_A + m_B} + \frac{v_{Ai}(m_A - m_B)}{m_A + m_B} \qquad (6.31)$$

$$v_{Bf} = -\frac{2m_A v_{Ai}}{m_A + m_B} + \frac{(m_B - m_A)\, v_{Bi}}{(m_A + m_B)} \qquad (6.32)$$

Notes

We now discuss some special cases.

CASE I : Suppose that the two balls colliding with each other are identical i.e. $m_A = m_B = m$. Then the second term in Eqns. (6.31 and (6.32) will drop out resulting in

$$\upsilon_{Af} = \upsilon_{Bi} \qquad (6.33)$$

and

$$\upsilon_{Bf} = \upsilon_{Ai} \qquad (6.34)$$

That is, if two identical balls collide "head-on", their velocities after collision get interchanged.

After collision:

i) the velocity of A is same as that of B before collision.

ii) the velocity of B is same as that of A before collision.

Now, think what would happen if one of the balls is at rest before collision?

Let B be at rest so that $\upsilon_{Bi} = 0$. Then $\upsilon_{Af} = 0$ and $\upsilon_{Bf} = \upsilon_{Ai}$

After collision, A comes to rest and B moves with the velocity of A before collision.

Similar conclusion can be drawn about the kinetic energy of the balls after collision. Complete loss of kinetic energy or partial loss of kinetic energy $(m_A \# m_B)$ by A is same as the gain in the kinetic energy of B. These facts have very important applications in nuclear reactors in slowing down neutrons.

CASE II : The second interesting case is that of collision of two particles of unequal masses.

i) Let us assume that m_B is very large compared to m_A and particle B is initially at rest :

$$m_B >> m_A \text{ and } \upsilon_{Bi} = 0$$

Then, the mass m_A can be neglected in comparison to m_B. From Eqns. (6.31) and (6.32), we get

$$\upsilon_{Af} \approx -\upsilon_{Ai}$$

and

$$\upsilon_{Bf} \approx 0$$

After collision, the heavy particle continues to be at rest. The light particle returns back on its path with a velocity equal to its the initial velocity.

This is what happens when a child hits a wall with a ball.

These results find applications in Physics of atoms, as for example in the case where an α – particle hits a heavy nucleus such as uranium.

INTEXT QUESTIONS 6.6

1 Two hard balls collide when one of them is at rest.

a) Is it possible that both of them remain at rest after collision?

b) Is it possible that one of them remains at rest after collision?

2. There is a system of three identical balls A B C on a straight line as shown here. B and C are in contact and at rest. A moving with a velocity υ collides "head-on" with B. After collision, what will be the velocities of A, B and C separately? Explain.

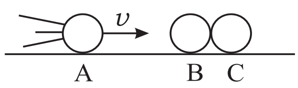


Fig. 6.20

3. Ball A of mass 2 kg collides head-on with ball B of mass 4 kg. A is moving in + x direction with speed 50 m s^{-1} and B is moving in $-x$ direction with speed 40 m s^{-1}. What are the velocities of A and B after collision? The collision is elastic.

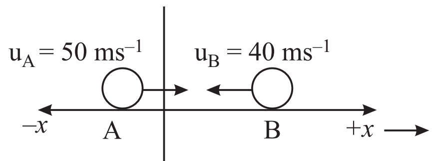


Fig. 6.21

4. A bullet of mass 1 kg is fired and gets embedded into a block of wood of mass 1 kg initially at rest. The velocity if the bullet before collision is 90m/s.

a) What is the velocity of the system after collision.

b) Calculate the kinetic energies before and after the collision?

c) Is it an elastic collision or inelastic collision?

d) How much energy is lost in collision?

5. In an elastic collision between two balls, does the kinetic energy of each ball change after collision?

WHAT YOU HAVE LEARNT

- Work done by a constant force F is

$$W = \mathbf{F.d} = Fd\cos\theta$$

Where θ is the angle between F and d. The unit of work is joule. Work is a scalar quantity.

- Work is numerically equal to the area under the F versus x graph.

- Work done by elastic force obeying Hooke's law is $W = \frac{1}{2}kx^2$ where k is force constant of the elastic material (spring or wire). The sign of W is **positive** for the external force acting on the spring and **negative** for the restoring force offered by spring. x is compression or elongation of the spring.

- The unit of k is newton per metre (N m^{-1}.)
- Power is the time rate of doing work. $P = W/t$ its unit is J/s i.e., watt (W)
- Mechanical energy of a system exists in two forms (i) kinetic energy and (ii) Potential energy.
- Kinetic energy of mass m moving with speed v is $E = \frac{1}{2}mv^2$. It is a scalar quantity.
- The **Work-Energy Theorem** states that the work done by all forces is equal to the change in the kinetic energy of the object.

$$W = K_f - K_i = \Delta K$$

- Work done by a conservative force on a particle is equal to the change in mechanical energy of the particle, that is change in the kinetic energy + the change in potential energy. In other words the mechanical energy is conserved under conservative forces.

$$\Delta E = (E_f - E_i) + (E_f - E_i)$$

$$= (\Delta E)_P + (\Delta E)_k$$

- Work done by a conservative force on an object is zero for a round trip of the object (object returning back to its starting point).
- Work done by a conservative force does not depend on the path of the moving object. It depends only on its initial and final positions.
- Work done is path dependent for a non-conservative force. The total mechanical energy is not conserved.
- The potential energy of a particle is the energy because of its position in space in a conservative field.
- Energy stored in a compressed or stretches spring is known as elastic potential energy. It has a value $\frac{1}{2}kx^2$, where k is spring constant and x is diplacement.
- The energy stored in a mass m near the earth's surface is mgh. It is called the gravitational potential energy. Here h is change in vertical co-ordinate of the mass. The reference level of zero potential energy is arbitrary.
- Energy may be transformed from one kind to another in an isolated system, but it can neither be created nor destroyed. The total energy always remains constant.

- Laws of conservation of momentum always hold good in any type of collision.
- The kinetic energy is also conserved in elastic collision while it is not conserved in inelastic collision.

TERMINAL EXERCISE

1. If two particles have the same kinetic energy, are their momenta also same? Explain.
2. A particle in motion collides with another one at rest. Is it possible that both of them are at rest after collision?
3. Does the total mechanical energy of a system remain constant when dissipative forces work on the system?
4. A child throws a ball vertically upwards with a velocity 20 m s^{-1}.

   (a) At what point is the kinetic energy maximum?

   (b) At what point is the potenital energy maximum?
5. A block of mass 3 kg moving with a velocity 20m s^{-1} collides with a spring of force constant 1200 N m^{-1}. Calculate the maximum compression of the spring.
6. What will be the compression of the spring in question 5 at the moment when kinetic energy of the block is equal to twice the elastic potential energy of the spring?
7. The power of an electric bulb is 60 W. Calculate the electrical energy consumed in 30 days if the bulb is lighted for 12 hours per day.
8. 1000 kg of water falls every second from a height of 120 m. The energy of this falling water is used to generate electricity. Calculate the power of the generator assuming no losses.
9. The speed of a 1200 kg car is 90 km h^{-1} on a highway. The driver applies brakes to stop the car. The car comes to rest in 3 seconds. Calculate the average power of the brakes.
10. A 400 g ball moving with speed 5 m s^{-1} has elastic head-on collision with another ball of mass 600 g initially at rest. Calculate the speed of the balls after collision.
11. A bullet of mass 10 g is fired with an initial velocity 500 m s^{-1}. It hits a 20 kg wooden block at rest and gets embedded into the block.

    (a) Calculate the velocity of the block after the impact

    (b) How much energy is lost in the collision?

12. An object of mass 6 kg. is resting on a horizontal surface. A horizontal force of 15 N is constantly applied on the object. The object moves a distance of 100m in 10 seconds.

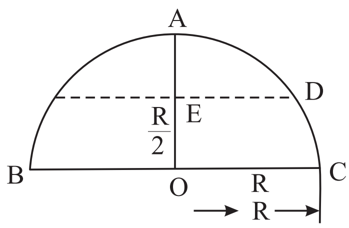
    (a) How much work does the applied force do?

    (b) What is the kinetic energy of the block after 10 seconds?

    (c) What is the magnitude and direction of the frictional force (if there is any)?

    (d) How much energy is lost during motion?

13. A, B, C and D are four point on a hemispherical cup placed inverted on the ground. Diameter BC = 50 cm. A 250 g particle at rest at A, slide down along the smooth surface of the cup. Calculate it's

    (a) Potential energy at A relative to B.

    (b) Speed at the point B (Lowest point).

    (c) Kinetic and potential energy at D.

    Do you find that the mechanical energy of the block is conserved? Why?

14. The force constant of a spring is 400 N/m. How much work must be done on the spring to stretch it (a) by 6.0 cm (b) from $x = 4.0$ cm to $x = 6.0$ cm, where $x = 0$ is the relaxed position of the spring.

15. The mass of a car is 1000 kg. It starts from rest and attains a speed of 15 m s^{-1} in 3.0 seconds. Calculate

    (a) The average power of the engine.

    (b) The work done on the car by the engine.

ANSWERS TO INTEXT QUESTIONS

6.1

1. The force always works at right angle to the motion of the particle. Hence no work is done by the force.

2. (a) Work done is zero (i) when there is no displacement of the object. (ii) When the angle between force and the displacement is 90º.

   When a mass moves on a horizontal plane the work done by gravitation force is zero.

(b) When a particle is thrown vericallly upwards, the work done by gravitational force is **negative**.

(c) When a particle moves in the direction of force, the work done by force is **positive.**

3. (a) $W = mgh = 2 \times 9.8 \times 5 = +98$ J

(b) The work done by gravity is –98 J

4. $\mathbf{F} = (2\hat{\mathbf{i}} + 3\hat{\mathbf{j}})$ $\quad\mathbf{d} = (-\hat{\mathbf{i}} + 2\hat{\mathbf{j}})$

$W = \mathbf{F.d}$ $\quad = (2\hat{\mathbf{i}} + 3\hat{\mathbf{j}}) . (-\hat{\mathbf{i}} + 2\hat{\mathbf{j}})$

$-2 + 6 = \mathbf{4}$

5. $\mathbf{F} = (5\hat{\mathbf{i}} + 3\hat{\mathbf{j}})$ $\quad\mathbf{d} = (3\hat{\mathbf{i}} + 4\hat{\mathbf{j}})$

(a) $|\mathbf{d}| = \sqrt{9+16} = \sqrt{25} = \mathbf{5}$ m

(b) $|\mathbf{F}| = \sqrt{25+9} = \sqrt{34} = 5.83$

(c) $W = \mathbf{F.d}$ $\quad = (5\hat{\mathbf{i}} + 3\hat{\mathbf{j}}) . (3\hat{\mathbf{i}} + 4\hat{\mathbf{j}})$

$= 15 + 12 = 27$ J

6.2

1. Spring constant is defined as the restoring force per unit displacement. Thus, it is measured in N m^{-1}.

2. $k = \dfrac{10\text{N}}{1\text{cm}} = \dfrac{10\text{N}}{1/100\text{m}} = 100$ N m

As $F = kx$ for $x = 50$ cm. $F = \left(100\dfrac{\text{N}}{\text{m}}\right)(0.5\text{ m})$

$= 50$ N.

$$W = \frac{1}{2}kx^2 = \frac{1}{2} \times \frac{100\text{N}}{\text{m}} \times \left(\frac{5}{100} \times \frac{5}{100}\right)\text{m}^2$$

$= 1.25$ N m $= 1.25$ J.

6.3

1. $P = \dfrac{\text{mgh}}{\text{t}} = \dfrac{(100 \times 9.8 \times 8)}{10\text{ s}}\text{J} = 784$ W.

2. 10 H.P $= (10 \times 746)$ W $= \dfrac{10 \times 746}{1000}$ W

$= 7.46$ kW

6.4

1. k.E. = $\frac{1}{2}mv^2$. It is never negative because

   (i) m is never negative

   (ii) v^2 is always positive.

2. (a) K.E = $\frac{1}{2}mv^2 = E$

   When v is made $2v$, K.E becomes 4 times and E becomes $4E$

   (b) When m becomes $\frac{m}{2}$, E becomes $\frac{E}{2}$

3. P.E. of spring = $\frac{1}{2}kx^2 = 3.6$ J

   $\therefore \quad x^2 = \frac{2 \times 3.6}{k} = \frac{2 \times 3.6}{180} = 0.04$ m

   and $\quad x = 0.2$ m = 20 cm.

4. $v^2 = u^2 - 2as$ Final velocity is zero and initial velocity is $\frac{90 \text{ km}}{\text{h}} = 25\text{m s}^{-1}$

   $\therefore \quad \frac{u^2}{2s} = a = \frac{25 \times 25}{2 \times 15} = 20.83 \text{ m s}^{-2}$

   $F = ma = 1000 \times 20.83 = 20830$ N.

   Power = $\frac{W}{t} = \frac{20830 \times 15}{25} = 12498$ W

5. Work done by external force = 375 J

   Work done by spring = – 375 J

6.5

1. (a) O, no change in P.E.

   (b) Change in P.E. = $mgh = 2 \times 9.8 \times 4 = 78.4$ J

   (c) Change in P.E. = 78.4 J.

   (d) – 78.4 J.

2. (a) Change in P.E. from = $mgh = 0.5 \times 9.8 \times 4 = 19.6$ J

$$\text{K.E. at B} = \frac{1}{2}m\upsilon^2 = 19.6 \text{ J}$$

$$\upsilon^2 = \frac{19.6 \times 2}{0.5}$$

$$\upsilon^2 = 78.4 \Rightarrow \upsilon = 8.85 \text{ m s}^{-1}$$

(b) $\upsilon = 14$ m s^{-1}

(c) $mgh = 0.5 \times 9.8 \times 10 = 49.0$ J (+ positive)

W = +49 J

3. BC = 2m

$$\frac{AC}{BC} = \sin 30^\circ$$

AC = BC sin30º

$$= 2 \times \frac{1}{2} = 1$$

Change in P.E. from C to B = mgh = $2 \times 9.8 \times 1 = 19.6$ J

But the K.E. at B is = 15.6 J

Energy lost = 19.6 – 15.6 = 4J

This loss is due to frictional force

4J = $F \times d = F \times 2$

F = 2N

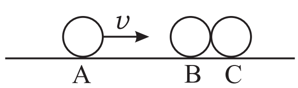
4. When the bob of a simple pendulum oscillates, its K.E. is max at $x = 0$ and min at $x = x_m$. The P.E. is min at $x = 0$ and max at $x = x_m$. Hence A represents the P.E. curve.

5. No.

6.6

1. (a) No, because, it will go against the low of conservation of linear momentum.

(b) yes.

2. 

$\upsilon_A = 0$, $\upsilon_B = 0$, $\upsilon_C = \upsilon$

∵ This condition only satisfies the laws of conservation of (i) linear momentum and (ii) total kinetic energy.

3. $$v_{Af} = \frac{2m_B v_{Bi}}{m_A + m_B} + \frac{v_{Ai}(m_A - m_B)}{m_A + m_B}$$

$$= \frac{2 \times 4 \times (-40)}{6} - \frac{50(-2)}{6}$$

$$= -\frac{320}{6} + \frac{100}{6}$$

$$= -\frac{220}{6}$$

$$= -35 \text{ms}^{-1}.$$

$$v_{Bf} = -\frac{2m_A v_{Ai}}{m_A + m_B} + \frac{(m_B - m_A)\, v_{Bi}}{(m_A + m_B)}$$

$$= \frac{2 \times 2 \times 50}{6} + \frac{(-40)(4-2)}{6}$$

$$= \frac{200}{6} - \frac{80}{6}$$

$$= \frac{120}{6} = 20 \text{ ms}^{-1}.$$

Thus ball A returns back with a velocity of 35 m s^{-1} and ball B moves on with a velocity of 20 m s^{-1}.

4. (a) 1.76 ms^{-1}.

   (b) 81 J and 1.58 J

   (c) Inelastic collision

   (d) 79.42 J

5. yes, but the total energy of both the balls together after collision is the same as it was before collision.

Answers to Terminal Problem

5. 1 m.

6. 0.707 m

7. 21.6 kW

8. 1.2 mega watt

9. 125 kW

10. $\frac{1}{4}$ m s^{-1}, $\frac{19}{6}$ m s^{-1}

11. (a) 0.25 m s^{-1} (b) 1249.4 J

12. (a) 1500 J (b) 1200 J

(c) 3 N opposite to the direction of motion

(d) 300 J

13. (a) 0.625 J (b) $\sqrt{5}$ m s^{-1} (c) 0.313 J

14. (a) 0.72 J (b) 0.4 J

15. (a) 37.5 kW (b) 1.125×10^5 J

312en07

7

MOTION OF RIGID BODY

So far you have learnt about the motion of a single object, usually taken as a point mass. This simplification is quite useful for learning the laws of mechanics. But in real life, objects consist of very large number of particles. A tiny pebble contains millions of particles. Do we then write millions of equations, one for each particle? Or is there a simpler way? While discovering answer to this question you will learn about centre of mass and moment of inertia, which plays the same role in rotational motion as does mass in translational motion.

You will also study an important concept of physics, the angular momentum. If no external force acts on a rotating system, its angular momentum in conserved. This has very important implications in physics. It enables us to understand how a swimmer is able to somersault while diving from a diving board into the water below.

OBJECTIVES

After studying this lesson, you should be able to :

- *define the centre of mass of a rigid body;*
- *explain why motion of a rigid body is a combination of translational and rotational motions;*
- *define moment of inertia and state theorems of parallel and perpendicular axes;*
- *define torque and find the direction of rotation produced by it;*
- *write the equation of motion of a rigid body;*
- *state the principle of conservation of angular momentum; and*
- *calculate the velocity acquired by a rigid body at the end of its motion on an inclined plane.*

7.1 RIGID BODY

As mentioned earlier, point masses are ideal constructs, brought in for simplicity in discussion. In practice, when extended bodies interact with each other and the distances between them are very large compared to their sizes, their sizes can be ignored and they may be treated as point masses. *Can you give two examples of such cases where the sizes of the bodies are not important?* Sizes of stars are small as compared to the size of the galaxy. So, stars can be considered as point masses. Similarly, in the earth-moon system, moon's size can be ignored. But when we have to consider the rotation of a body about an axis, the size of the body becomes important. When we consider the rotation of a system, we generally assume that during rotation, the distances between its constituent particles remain fixed. Such a system of particles is called a ***rigid body***.

A rigid body is one in which the separation between the constituent particles does not change with its motion.

This definition implies that the shape of a rigid body is preserved during its motion. However, like a point mass, a rigid body is also an idealisation because, if we apply large forces, the distances between particles do change, may be infinitesimally. Therefore, in nature there is nothing like a perfectly rigid body. For most purposes, a solid body is good enough approximation to a rigid body. A cricket ball, a wooden block, a steel disc, even the earth and the moon could be considered as rigid bodies in this lesson.

Can water in a bucket be considered a rigid body? Obviously, water in a bucket cannot be a rigid body because it changes shape as bucket is pushed around.

You may now like to check what you have understood about a rigid body.

INTEXT QUESTIONS 7.1

1. A frame is made of six wooden rods. The rods are firmly attached to each other. Can this system be considered a rigid body?
2. Can a heap of sand be considered a rigid body? Explain your answer.

7.2 CENTRE OF MASS (C.M.) OF A RIGID BODY

Before we deal with rigid bodies consisting of several particles, let us consider a simpler case. Suppose we have a system of two particles having same mass joined by weightless and inextensible rod. *Can we consider this system as a rigid body?*

In this system, the distance between the two particles is fixed. So it is a rigid body.

Suppose that the two particles are at heights z_1 and z_2 from a horizontal surface (Fig. 7.1). Suppose further that the gravitational force is uniform in the small region in which the two particles move about. The force on each particle will be mg. The total force acting on the system is therefore $2mg$. We have now to find a point C somewhere in the system so that if a force $2mg$ acts at that point located at a height z from the horizontal surface, the motion of the system would be the same as with two forces. The potiential energies of particles 1 and 2 are mgz_1 and mgz_2, respectively. The potential energy of the particle at C is $2mgz$. Since this must be equal to the combined potential energy of the two particles, can write

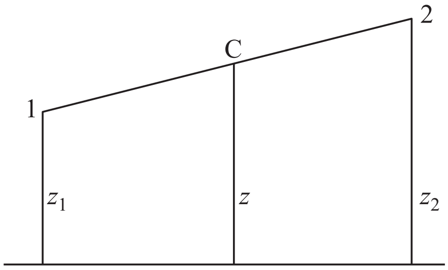


Fig. 7.1 : Two particle system

$$2\,mgz = mgz_1 + mgz_2 \tag{7.1}$$

or
$$z = \frac{z_1 + z_2}{2} \tag{7.2}$$

Note that the point C lies midway between the two particles. If the two masses were unequal, this point would not have been in the middle. If the mass of particle 1 is m_1 and that of particle 2 is m_2, Eqn. (7.1) modifies to

$$(m_1 + m_2)\,gz = m_1 g z_1 + m_2 g z_2 \tag{7.3}$$

so that
$$z = \frac{m_1 z_1 + m_2 z_2}{(m_1 + m_2)} \tag{7.4}$$

The point C is called the **centre of mass** (CM) of the system. As such, it is a mathematical tool and there is no physical point as CM.

To grasp this concept, study the following example carefully.

Example 7.1 : If in the above case, the mass of one particle is twice that of the other, let us locate the CM.

Solution : $m_1 = m$ and $m_2 = 2\,m$, Then Eqn. (7.4) gives

$$z = \frac{m\,z_1 + 2\,m\,z_2}{(m + 2\,m)} = \frac{z_1 + 2\,z_2}{3}$$

When a body consists of several particles, we generalise Eqn (7.4) to define its CM : *If the particle with mass m_1 has coordinates (x_1, y_1, z_1) with respect to some coordinate system, mass m_2 has coordinates (x_2, y_2, z_2) and so on* (*Fig*.7.2), *the coordinates of CM are given by*

$$x = \frac{m_1 x_1 + m_2 x_2 + ...}{m_1 + m_2 + ...}$$

$$= \frac{\sum_{i}^{N} m_i x_i}{\sum_{i=1}^{N} m_i} \qquad (7.5)$$

$$= \frac{\sum_{i=1}^{N} m_i x_i}{M}$$

Similarly $$y = \frac{\sum_{i=1}^{N} m_i y_i}{M} \qquad (7.6)$$

and $$z = \frac{\sum_{i=1}^{N} m_i z_i}{M} \qquad (7.7)$$

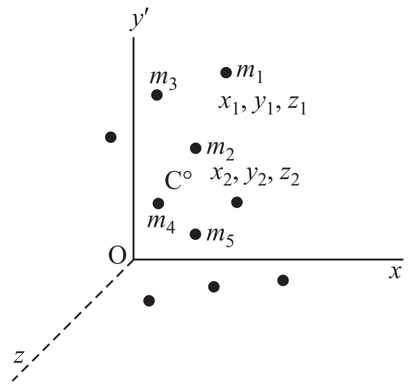


Fig. 7.2 : C.M. of a body consisting of several particles

where $\sum_{i=1}^{N} m_i$ denotes the sum over all the particles and, therefore, $\sum_{i=1}^{N} m_i$ is the total mass of the body, M.

Why should we define CM so precisely?

Recall that the rate of change of displacement is velocity, and the rate of change of velocity is acceleration. If a_{1x} denotes the component of acceleration of particle 1 along the x-axis and so on, from Eqn. (7.5), we can write

$$M\,a_x = m_1 a_{1x} + m_2 a_{2x} + ... \qquad (7.8)$$

where a_x is the acceleration of the centre of mass along x-axis. Similar equations can be written for accelerations along y- and z-axes. These equations can, however, be combined into a single equation using vector notation :

$$M\,\mathbf{a} = m_1 \mathbf{a_1} + m_2 \mathbf{a_2} + ... \qquad (7.9)$$

But the product of mass and acceleration is force. $m_1\, a_1$ is therefore the sum of all forces acting on particle 1. Similarly, $m_2\, \mathbf{a_2}$ gives the net force acting on particle 2. The right hand side is, thus, the total force acting on the body.

The forces acting on a body can be of two kinds. Some forces can be due to sources outside the body. These forces are called the ***external*** forces. A familiar example is the force of gravity. Some other forces arise due to the interaction among the particles of the body. These are called ***internal*** forces. A familiar example is cohesive force.

In the case of a rigid body, the sum of the internal forces is zero because they cancel each other in pairs. Therefore, the acceleration of individual particles of the body are due to the sum or resultant of the external forces. In the light of this, we may write Eqn. (7.9) as

$$M\,\mathbf{a} = \mathbf{F}_{ext} \qquad (7.10)$$

This shows that ***the CM of a body moves as though the entire mass of the body were located at that point and it was acted upon by the sum of all the external forces acting on the body.*** Note the simplification introduced in the derivation by defining the centre of mass. We donot have to deal with millions of individual particles now, only the centre of mass needs to be located to determine the motion of the given body. The fact that the motion of the CM is determined by the external forces and that the internal forces have no role in this at all leads to very interesting consequences.

You are familiar with the motion of a projectile. *Can you recall what path is traced by a projectile?*

The motion is along a parabolic path. Suppose the projectile is a bomb which explodes in mid air and breaks up into several fragments. The explosion is caused by the internal forces. There is no change in the external force, which is the force of gravity. The centre of mass of the projectile, therefore, continues to be the same parabola on which the bomb would have moved if it had not exploded (Fig. 7.3). The fragments may fly in all directions on different parabolic paths but the centre of mass of the various fragments will lie on the original parabola.

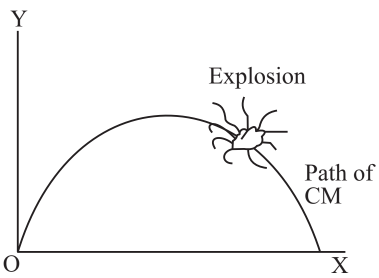


Fig. 7.3 : Centre of mass of a projectile

You might have now understood the importance of the concept of centre of mass of a rigid body. You will encounter more examples of importance in subsequent sections. Let us therefore see how the centre of mass of a system is obtained by taking a simple example.

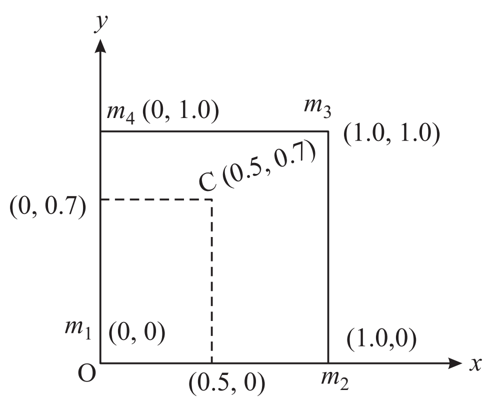


Fig. 7.4 : Locating CM of four masses placed at the corners of a square

Example 7.2 : Suppose four masses, 1.0 kg, 2.0 kg, 3.0 kg and 4.0 kg are located at the corners of a square of side 1.0 m. Locate its centre of mass?

Solution : We can always make the square lie in a plane. Let this plane be the (x,y) plane. Further, let us assume that one of the corners coincides with the origin of the coordinate system and the sides are along the x and y axes. The coordinates of the four masses are : m_1 (0,0), m_2 (1.0,0), m_3 (1.0,1.0) and m_4 (0, 1.0), where all distances are expressed in metres (Fig.7.4).

From Eqns. (7.5) and (7.6), we get

$$x = \frac{1.0 \times 0 + 2.0 \times 1.0 + 3.0 \times 1.0 + 4.0 \times 0}{1.0 + 2.0 + 3.0 + 4.0} \text{ m}$$

$$= 0.5 \text{ m}$$

and
$$y = \frac{1.0 \times 0 + 2.0 \times 0 + 3.0 \times 1.0 + 4.0 \times 1.0}{1.0 + 2.0 + 3.0 + 4.0} \text{ m}$$

$$= 0.7 \text{ m}$$

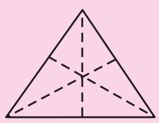
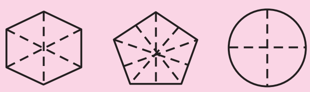
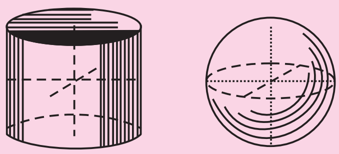
The CM has coordinates (0.5 m, 0.7 m) and is marked C in Fig.7.4. Note that the CM is not at the centre of the square although the square is a symmetrical figure.

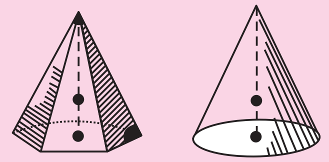
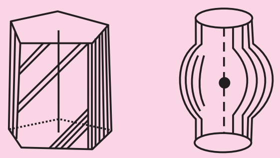
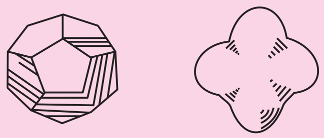
What could be the reason for the CM not being at the centre? To discover answer to this question, calculate the coordinates of CM if all masses are equal.

7.2.1 CM of Some Bodies

The position of centre of mass of extended bodies can not be easily calculated because a very large number of particles constituting the body have to be considered. The fact that all the particles of a rigid body have same mass and are uniformly distributed makes things somewhat simpler. If the body is regular in shape and possesses some symmetry, say it is cylindrical or spherical, the calculation is a little bit simplified. But even such calculations are beyond the scope of this course. However, keeping in mind the importance of CM, we give in Table 7.1 the position of CM of some regular, symmetrical bodies.

Table 7.1 Centres of Mass of some regular symmetrical bodies

Figure	Position of Centre of Mass
	*Triangular plate* Point of intersection of the three medians
	*Regular polygon and circular plate* At the geometrical centre of the figure
	*Cylinder and sphere* At the geometrical centre of the figure

	*Pyramid and cone* On line joining vertex with centre of base and at h/4 of the height measured from the base.
	*Figure with axial symmetry* Some point on the axis of symmetry
	*Figure with centre of symmetry* At the centre of symmetry

Two things must be remembered about the centre of mass : (i) It may be outside the body as in case of a ring. (ii) When two bodies revolve around each other, they actually revolve around their common centre of mass. For example, stars in a binary system revolve around their common centre of mass. The Earth-Sun system also revolves around its common centre of mass. But since mass of the Sun is very large as compared to the mass of earth, the centre of mass of the system is very close to the centre of the Sun.

Now it is time to check your progress.

INTEXT QUESTIONS 7.2

1. The grid shown here has particles A, B, C, D and E respectively have masses 1.0 kg, 2.0kg, 3.0 kg, 4.0 kg and 5.0 kg. Calculate the coordinates of the position of the centre of mass of the system (Fig. 7.5).

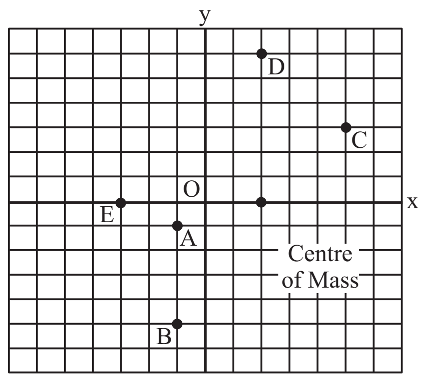


Fig. 7.5

2. If three particles of masses m_1 =1 kg, m_2 = 2 kg, and m_3 = 3 kg are situated at the corners of an equilateral triangle of side 1.0 m, obtain the position coordinates of the centre of mass of the system.

3. Show that the ratio of the distances of the two particles from their common centre of mass is inversely proportional to the ratio of their masses.

7.3 TRANSLATIONAL AND ROTATIONAL MOTION OF A RIGID BODY : A COMPARISON

When a rigid body moves in such a way that all its particles move along parallel paths (Fig.7.6), its motion is called **translational** motion. Since all the particles execute identical motion, its centre of mass must also be tracing out an identical path. Therefore, the translational motion of a body may be characterised by the motion of its centre of mass. We have seen that this motion is given by Eqn.(7.10):

$$M\,\mathbf{a} = \mathbf{F}_{\text{ext}}$$

Do you now see the advantage of defining the centre of mass of a body? With its help, the translational motion of body can be described by an equation for a single particle having mass equal to the mass of the whole body. It is located at the centre of mass and is acted upon by the sum of all the external forces which are acting on the rigid body. To understand the concept clearly, perform the following activities.

ACTIVITY 7.1

Take a wooden block. Make two or three marks on any of its surfaces. Now keep the marked surface in front of you and push the block along a horizontal floor. Note the paths traced by the marks. All these marks have paths parallel to the floor and, therefore, parallel to one another (Fig. 7.6). You can easily see that the lengths of the paths are also equal.

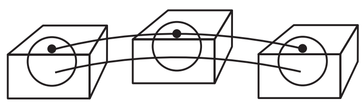


Fig. 7.6 : A wooden block moving along the floor performs translational motion.

ACTIVITY 7.2

Let us now perform another simple experiment. Take a cylindrical piece of wood. On its plane face make a mark or two. Now roll the cylinder slowly on the floor, keeping the plane face towards you. You would notice that the mark such as A in Fig. 7.7, has not only moved parallel to the floor, but has also performed circular motion. So, the body has performed both translational and rotational motion.

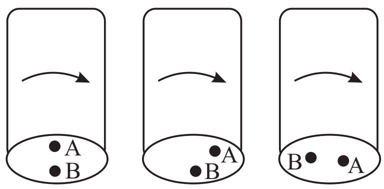


Fig. 7.7 : Rolling motion of a cylinder: The point A has not only moved parallel to the floor but also performed circular motion

While the general motion of a rigid body consists of both translation and rotation, it cannot have translational motion if one point in the body is fixed; it can then only rotate. The most convenient point to fix for this purpose is the CM of the body.

You might have seen a grinding stone (the chakki). The handle of the stone moves in a circular path. All the points on the stone also move in circular paths around an axis passing through the centre of the stone (Fig.7.8).

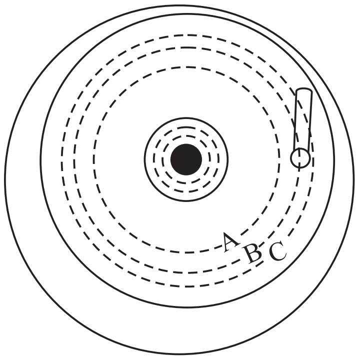


Fig. 7.8 : Pure rotation of a grinding stone

The motion of a rigid body in which all its constituent particles describe concentric circular paths is known as rotational motion.

We have noted above that translational motion of a rigid body can be described by an equation similar to that of a single particle. You are familiar with such equations. Therefore, in this lesson we concentrate only on the rotational motion of a rigid body. The rotational motion can be obtained by keeping a point of the body fixed so that it cannot have any translational motion. For the sake of mathematical convenience, this point is taken to be the CM. The rotation is then about an axis passing through the CM. A good example of rotational motion is the earth's rotation about its own axis (Fig. 7.9). You have studied in earlier lessons that the mass of the body plays a very important role. It determines the acceleration acquired by the body for a given force. Can we define a similar quantity for rotational motion also? Let us find out.

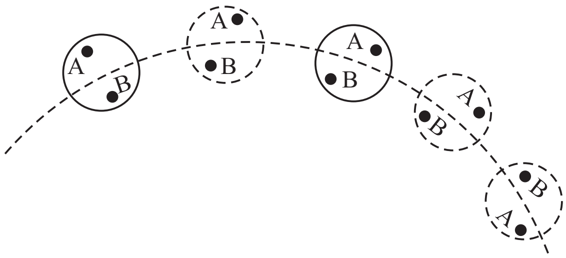


Fig. 7.9 : Rotation of the earth

7.3.1 Moment of Inertia

Let C be the centre of mass of a rigid body. Suppose it rotates about an axis through this point (Fig.7.10).

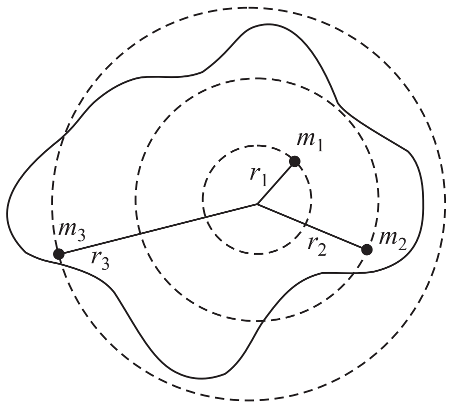


Fig. 7.10 : Rotation of a plane lamina about an axis passing through its centre of mass

Suppose particles of masses m_1, m_2, m_3...are located at distances r_1, r_2, r_3...from the axis of rotation and are moving with speeds υ_1,

υ_2, υ_3 respectively. Then particle 1 has kinetic energy (½) $m_1 \upsilon_1^2$. Similarly, the kinetic energy of particle of mass m_2 is (½) $m_2 \upsilon_2^2$. By adding the kinetic energies of all the particles, we get the total energy of the body. If T denotes the total kinetic energy of the body, we can write

$$T = (½)\, m_1 \upsilon_1^2 + (½)\, m_2 \upsilon_2^2 + ...$$

$$= \sum_{i=1}^{i=1} \left(\frac{1}{2}\right) m_i \upsilon_i^2 \qquad (7.11)$$

where $\sum_{i=1}^{i=n}$ indicates the sum over all the particles of the body.

You have studied in lesson 4 that angular speed (ω) is related to linear speed (υ) through the equation $\upsilon = r\,\omega$. Using this result in Eqn. (7.11), we get

$$T = \sum_{i=1}^{i=n} \left(\frac{1}{2}\right) m_i (r_i \omega)^2 \qquad (7.12)$$

Note that we have not put the subscript i with w because all the particles of a rigid body have the same angular speed. Eqn. (7.12) can now be rewritten as

$$T = \frac{1}{2}\left(\sum_{i=1}^{i=n} m_i\, r_i^2\right)\omega^2$$

$$= \frac{1}{2}\, I\, \omega^2 \qquad (7.13)$$

The quantity $$I = \sum_i m_i\, r_i^2 \qquad (7.14)$$

is called the ***moment of inertia*** of the body.

Example 7.3 : *F*our particles of mass m each are located at the corners of a square of side L. Calculate their moment of inertia about an-axis passing through the centre of the square and perpendicular to its plane.

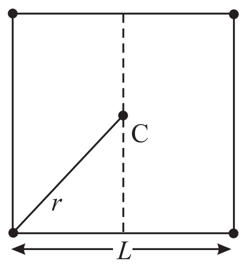


Fig. 7.11

Solution : Simple geometry tells us that the distance of each particle from the axis of rotation is $r = L\sqrt{2}$. Therefore, we can write

$$I = m\,r^2 + m\,r^2 + m\,r^2 + m\,r^2$$

$$= 4m\,r^2$$

Notes

$$= 4m\left(\frac{L}{\sqrt{2}}\right)^2 \qquad \left(\text{Since } r = \frac{L}{\sqrt{2}}\right).$$

$$= 2mL^2$$

It is important to remember that ***moment of inertia*** is defined with reference to an axis of rotation. Therefore, whenever you mention moment of inertia, the axis of rotation must also be specified. In the present case, I is the moment of inertia about an axis passing through the centre of the square and normal to the plane containing four perfect masses. (Fig. 7.10) The moment of inertia is expressed in kg m^2.

The moment of inertia of a rigid body is often written as

$$I = MK^2 \qquad (7.15)$$

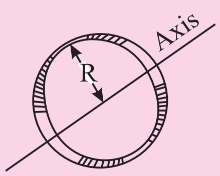
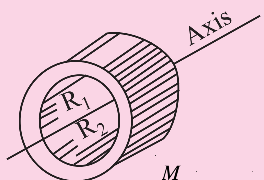
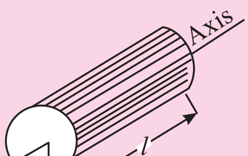
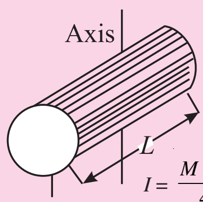
where M is the total mass of the body and K is called the ***radius of gyration*** of the body. ***The radius of gyration is that distance from the axis of rotation where the whole mass of the body can be assumed to be placed to get the same moment of inertia which the body actually has.*** It is important to remember that the moment of inertia of a body about an axis depends on the distribution of mass around that axis. If the distribution of mass changes, the moment of inertia will also change. This can be easily seen from Example 7.3. Suppose we place additional masses at one pair of opposite corners of amount m each. Then the moment of inertia of the system about the axis through C and perpendicular to the plane of square would be

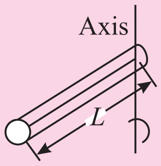
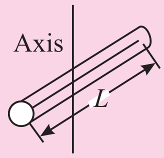
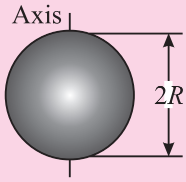
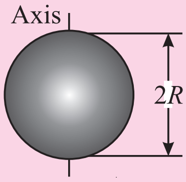
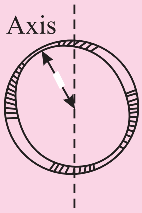
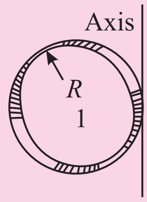
$$I = mr^2 + 2mr^2 + mr^2 + 2mr^2$$

$$= 6mr^2$$

Note that moment of inertia has changed from $2mL^2$ to $3\,mL^2$.

Table 7.2 Moments of inertia of a few regular and uniform bodies.

 Hoop about central axis $I = MR^2$	 Annular cylinder (or ring) about cylinder axis $I = \frac{M}{2}(R_1^2 + R_2^2)$
 Solid cylinder about cylindrical axis $I = \frac{MR^2}{2}$	 Solid cylinder (or disk) about a central diameter $I = \frac{MR^2}{4} + \frac{M^2\ell^2}{12}$

 $I = \frac{ML^2}{12}$ Thin rod about an axis passing through its centre and normal to its length	 $I = \frac{ML^2}{3}$ Thin rod about an axis passing through one end and perpendicular to length
 $I = \frac{2MR^2}{5}$ Solid sphere about any diameter	 $I = \frac{2MR^2}{3}$ Thin spherical shell about any diameter
 $I = \frac{MR^2}{2}$ Hoop about any diameter	 $I = \frac{3MR^2}{2}$ Hoop about any tangent line

Refer to Eqn.(7.13) again and compare it with the equation for kinetic energy of a body in linear motion. Can you draw any analogy? You will note that in rotational motion, the role of mass has been taken over by the moment of inertia and the angular speed has replaced the linear speed.

A. Physical significance of moment of inertia

The physical significance of moment of inertia is that it performs the same role in rotational motion that the mass does in linear motion.

Just as the mass of a body resists change in its state of linear motion, the moment of inertia resists a change in its rotational motion. This property of the moment of inertia has been put to a great practical use. Most machines, which produce rotational motion have as one of their components a disc which has a very large moment of inertia. Examples of such machines are the steam engine and the automobile engine. The disc with a large moment of inertia is called a ***flywheel***. To understand how a flywheel works, imagine that the driver of the engine wants to suddenly increase the speed. Because of its large moment of inertia, the flywheel resists this attempt. It allows only a gradual increase in speed. Similarly, it works against the attempts to suddenly reduce the speed, and allows only a gradual decrease in the speed. Thus , the flywheel, with its large moment of inertia, prevents jerky motion and ensures a smooth ride for the passengers.

We have seen that in rotational motion, angular velocity is analogous to linear velocity in linear motion. Since angular acceleration (denoted usually by α) is the rate of change of angular velocity, it must correspond to acceleration in linear motion.

B. Equations of motion for a uniformly rotating rigid body

Consider a lamina rotating about an axis passing through O and normal to its plane. If it is rotating with a constant angular velocity ω, as shown, then it will turn through an angle θ in t seconds such that

$$\theta = \omega t \qquad 7.16(a)$$

However, if the lamina is subjected to constant torque (which is the turning effect of force), it will undergo a constant angular acceleration. The following equations describe its rotational motion:

$$\omega_f = \omega_i + \alpha t \qquad 7.16(b)$$

where ω_i is initial angular velocity and ω_f is final angular velocity.

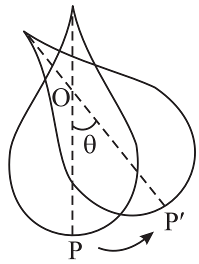


Fig. 7.12 : Rotation of a lamina about a fixed nail

Similarly, we can write

$$\theta = \omega_i t + \frac{1}{2} \alpha t^2 \qquad 7.16(c)$$

$$\omega_f^2 = \omega_i^2 + 2 \alpha \theta \qquad 7.16(d)$$

where θ is angular dispalcement in t seconds.

On a little careful scrutiny, you can recognise the similarity of these equations with the corresponding equations of kinematics for translatory motion.

Example 7.4 : A wheel of a bicycle is free to rotate about a horizontal axis (Fig. 7.11). It is initially at rest. Imagine a line OP drawn on it. By what angle would the line OP move in 2 s if it had a uniform angular acceleration of 2.5 rad s^{-2}.

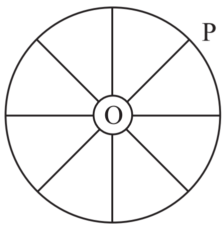


Fig. 7.13 : Rotation of bicycle wheel

Solution : Angular displacement of line OP is given by

$$\theta = \omega_0 t + (½)\,\alpha\, t^2$$
$$= 0 + (½) \times (2.5 \text{ rad s}^{-2}) . \times 4 \text{ s}^2$$
$$= 5 \text{ rad}$$

We have mentioned above that for rotational motion of a rigid body, its CM is kept fixed. However, it is just a matter of convenience that we keep CM fixed. But many a time, we consider points other than the CM. That is, a point in the body which can also be kept fixed and the body rotated about it. But then the axis of rotation will pass through this fixed point. The moment of inertia about this axis would be different from the moment of inertia about an axis passing through the CM. The relation between the two moments of inertia can be obtained using the theorems of moment of inertia.

7.3.2 Theorems of moment of inertia

There are two theorems which connect moments of inertia about two axes; one of which is passing through the CM of the body. These are :

(i) the theorem of parallel axes, and

(ii) the theorem of perpendicular axes.

Let us now learn about these theorems and their applications.

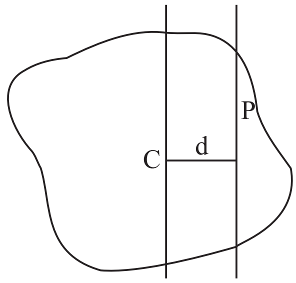


Fig. 7.14 : Parallel axes through CM and another point P

(*i*) Theorem of parallel axes

Suppose the given rigid body rotates about an axis passing through any point P other than the centre of mass. The moment of inertia about this axis can be found from a knowledge of the moment of inertia about a parallel axis through the centre of mass. Theorem of parallel axis states that *the moment of inertia about an axis parallel to the axis passing through its centre of mass is equal to the moment of inertia about its centre of mass plus the product of mass and square of the perpendicular distance between the parallel axes*. If I denotes the required moment of inertia and I_C denotes the moment of inertia about a parallel axis passing through the CM, then

$$I = I_C + M\,d^2 \qquad (7.17)$$

where M is the mass of the body and d is the distance between the two axes (Fig. 7.12). This is known as the **theorem of parallel axes.**

(*ii*) Theorem of perpendicular axes

Let us choose three mutually perpendicular axes, two of which, say x and y are in the plane of the body, and the third, the z–axis, is perpendicular to the plane (Fig.7.13). The perpendicular axes theorem states that ***the sum of the moments of inertia about x and y axes is equal to the moment of inertia about the z–axis.***

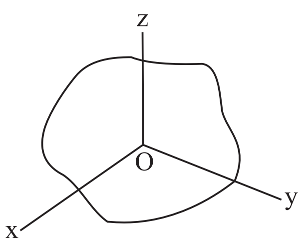


Fig. 7.15 : Theorem of perpendicular axes

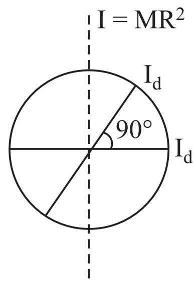


Fig. 7.16 : Moment of inertia of a hoop

That is,

$$I_z = I_x + I_y \tag{7.18}$$

We now illustrate the use of these theorems by the following example.

Let us take a hoop shown in Fig. 7.16. From Table 7.2 you would recall that moment of inertia of a hoop about an axis passing through its centre and perpendicular to the base is $M R^2$, where M is its mass and R is its radius. The theorem of perpendicular axes tells us that this must be equal to the sum of the moments of inertia about two diameters which are perpendicular to each other as well as to the central axis. The symmetry of the hoop tells us that the moment of inertia about any diameter is the same as about any other diameter. This means that all the diameters are **equivalent** and any two perpendicular diameters may be chosen. Since the moment of inertia about each is the same, say I_d, Eqn.(7.18) gives

$$M R^2 = 2 I_d$$

and therefore

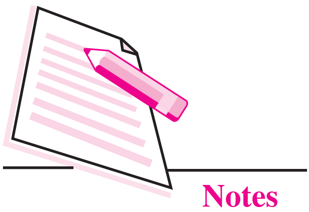
$$I_d = (½) MR^2$$

So, the moment of inertia of a hoop about any of its diameter is (½) $M R^2$.

Let us now take a point P on the rim. Consider a tangent to the hoop at this point which is parallel to the axis of the hoop. The distance between the two axes is obviously equal to R. The moment of inertia about the tangent can be calculated using the theorem of parallel axes. It is given by

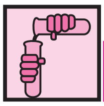
$$I_{\tan} = M R^2 + M R^2 = 2 M R^2.$$

It must be mentioned that many of the entries in Table 7.2 have been computed using the theorems of parallel and perpendicular axes.



Notes

7.3.3 Torque and Couple



ACTIVITY 7.3

Have you ever noticed that it is easy to open the door by applying force at a point far away from the hinges? What happens if you try to open a door by applying force near the hinges? Carry out this activity a few times. You would realise that much more effort is needed to open the door if you apply force near the hinges than at a point away from the hinges. Why is it so? Similarly, for turning a screw we use a spanner with a long handle. What is the purpose of keeping a long handle? Let us seek answers to these questions now.

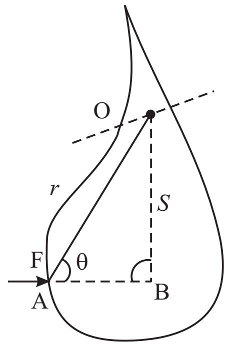


Fig. 7.17 : Rotation of a body

Suppose O is a fixed point in the body and it can rotate about an axis passing through this point (Fig.7.17). Let a force of magnitude F be applied at the point A along the line AB. If AB passes through the point O, the force F will not be able to rotate the body. The farther is the line AB from O, the greater is the ability of the force to turn the body about the axis through O. The ***turning effect of a force is called torque.*** Its magnitude is given by

$$\tau = F\,s = F\,r\sin\theta \qquad (7.19)$$

where s is the distance between the axis of rotation and the line along which the force is applied.

The units of torque are newton-metre, or Nm. The torque is actually a vector quantity. The vector from of Eqn.(7.19) is

$$\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F} \qquad (7.20)$$

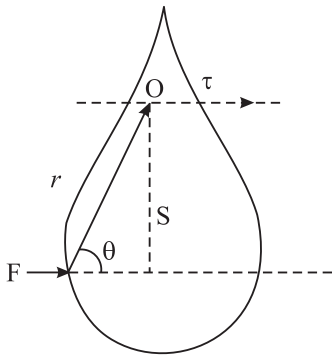


Fig. 7.18 : Right hand thumb rule

which gives both magnitude and direction of the torque. What is the direction in which the body would turn? To discover this, we recall the rules of vector product (refer to lesson 1) : τ is perpendicular to the plane containing vectors **r** and **F**, which is the plane of paper here (Fig.7.18). If we extend the thumb of the right hand at right angles to the fingers and curl the fingers so as to point from **r** to **F** through the smaller angle, the direction in which thumb points is the direction of $\boldsymbol{\tau}$.

Apply the above rule and show that the turning effect of the force in Fig. 7.18 is normal to the plane of paper downwards. This corresponds to clockwise rotation of the body.

Example 7.5 : Fig.7.19 shows a bicycle pedal. Suppose your foot is at the top and you are pressing the pedal downwards. (*i*) What torque do you produce? (ii) Where should your foot be for generating maximum torque?

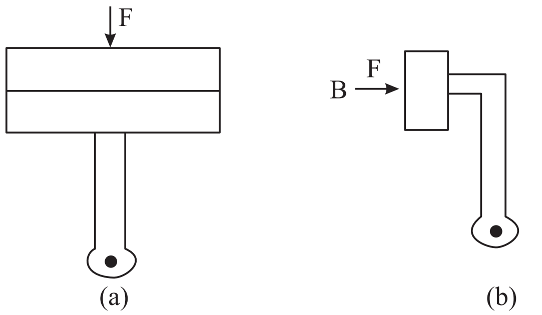


Fig. 7.19 : A bicycle pedal (a) at the top when $\tau = 0$; (b) when τ is maximum

Solution : (i) When your foot is at the top, the line of action of the force passes through the centre of the pedal. So, $\theta = 0$, and $\tau = Fr \sin\theta = 0$.

(ii) To get maximum torque, $\sin\theta$ must have its maximum value, that is θ must be 90º. This happens when your foot is at position B and you are pressing the pedal downwards.

If there are several torques acting on a body, the net torque is the vector sum of all the torques. Do you see any correspondence between the role of torque in the rotational motion and the role of force in the linear motion? Consider two forces of equal magnitude acting on a body in opposite directions (Fig.7.20). Assume that the body is free to rotate about an axis passing through O. The two torques on the body have magnitudes

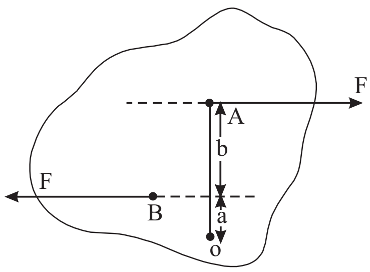


Fig. 7.20 : Two opposite forces acting on body

$$\boldsymbol{\tau}_1 = (a + b)\, F$$

and

$$\boldsymbol{\tau}_2 = a\, F.$$

You can verify that the turning effect of these torques are in the opposite directions. Therefore, the magnitude of the net turning effect on the body is in the direction of the larger torque, which in this case is $\boldsymbol{\tau}_1$:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_1 - \boldsymbol{\tau}_2 = \mathbf{bF} \qquad (7.21)$$

We may therefore conclude that ***two equal and opposite forces having different lines of action are said to form a couple whose torque is equal to the product of one of the forces and the perpendicular distance between them.***

There is another useful expression for torque which clarifies its correspondence with force in linear motion. Consider a rigid body rotating about an axis passing through a point O (Fig. 7.21). Obviously, a particle like P is rotating about the axis in a circle of radius r. If the circular motion is non–uniform, the particle experiences forces in the radial direction as well as in the tangential direction. The radial force is the centripetal force $m\,\omega^2 r$, which keeps the particle in the circular path. The tangential force is required to change the magnitude of υ, which at every instant is along the tangent to the circular path. Its magnitude is $m\,a$, where a is the tangential acceleration. ***The radial force does not produce any torque.*** *Do you know why? The tangential force produces a torque of magnitude m a r.* Since $a = r\,\alpha$, where α is the angular acceleration, the magnitude of the torque is $m\,r^2\alpha$. If we consider all the particles of the body, we can write

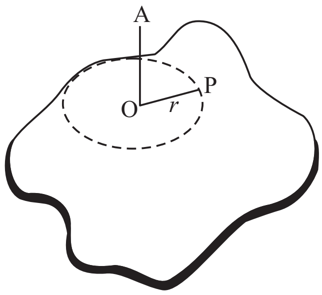


Fig. 7.21 : A rigid body rotating about on axis

$$\tau = \sum_{i=0}^{i=n} m_i r_i^2 \alpha = \left(\sum_i m_i r_1^2\right)\alpha$$

$$= I\,\alpha. \qquad (7.22)$$

because α is same for all the particles at a given instant.

The similarity between this equation and $\mathbf{F} = m\,a$ shows that $\boldsymbol{\tau}$ performs the same role in rotational motion as **F** does in linear motion. A list of corresponding quantities in rotational motion and linear motion is given in Table 7.3. With the help of this table, you can write any equation for rotational motion if you know its corresponding equation in linear motion.

Table 7.3 : Corresponding quantities in rotational and translational motions

Translational Motion		**Rotation about a Fixed Axis**	
Displacement	x	Angular displacement	θ
Velocity	$\upsilon = \frac{dx}{dt}$	Angular velocity	$\omega = \frac{d\theta}{dt}$
Acceleration	$a = \frac{dv}{dt}$	Angular acceleration	$\alpha = \frac{d\omega}{dt}$
Mass	M	Moment of inertia	I
Force	$F = m\,a$	Torque	$\tau = I\,\alpha$
Work	$W = \int F\,dx$	Work	$W = \int \tau\,d\theta$
Kinetic energy	$\frac{1}{2} M\,v^2$	Kinetic energy	$(\frac{1}{2})\,I\,\omega^2$
Power	$P = F\upsilon$	Power	$P = \tau\,\omega$
Linear momentum	$M\,\upsilon$	Angular momentum	$I\,\omega$

With the help of Eqn.(7.22) we can calculate the angular acceleration produced in a body by a given torque.

Example 7.6 : A uniform disc of mass 1.0 kg and radius 0.1m can rotate about an axis passing through its centre and normal to its plane without friction. A massless string goes round the rim of the disc and a mass of 0.1 kg hangs at its end (Fig.7.22). Calculate (i) the angular acceleration of the disc, (ii) the angle through which the disc rotates in one second, and (iii) the angular velocity of the disc after one second. Take $g = 10\ ms^{-2}$

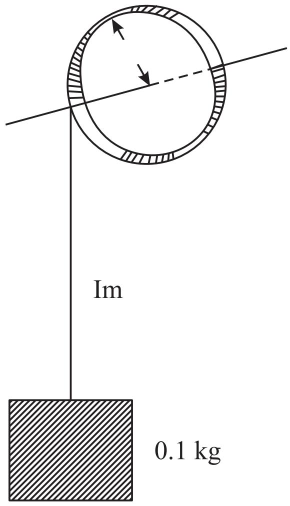


Fig. 7.22

Solution : (i) If R and M denote the radius and mass of the disc, from Table 7.2, we recall that its moment of inertia is given by $I = (½)\, M R^2$. If F denotes the magnitude of force (= $m\, g$) due to the mass at the end of the string then $\tau = F R$. Eqn. (7.22) now gives

$$\alpha = \tau / I = FR/I = 2F/MR$$

$$= \frac{2\times(0.1\text{kg})\times(10\text{ms}^{-2})}{(1.0\,\text{kg})\times(0.1\,\text{m})} = 20 \text{ rad s}^{-2}.$$

(ii) For angle θ through which the disc rotates, we use Eqn.(7.16). Since the initial angular velocity is zero, we have

$$\theta = (½) \times 20 \times 1.0 = 10 \text{ rad}$$

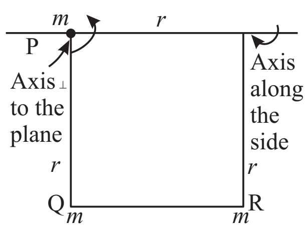
(iii) For the velocity after one second, we have

$$\omega = \alpha\, t = 20 \times 1.0 = 20 \text{ rad s}^{-1}$$

Now, you may like to check your progress.Try the following questions.

INTEXT QUESTIONS 7.3

1. Four particles, each of mass **m**, are fixed at the corners of a square whose each side is of length r. Calculate the moment of inertia about an axis passing through one of the corners and perpendicular to the plane of the square. Calculate also the moment of inertia about an axis which is along one of the sides. Verify your result by using the theorem of perpendicular axes.



2. Calculate the radius of gyration of a solid sphere if the axis is a tangent to the sphere. (You may use Table 7.2)

7.4 ANGULAR MOMENTUM

From Table 7.3 you may recall that rotational analogue of linear momentum is angular momentum. To understand its physical significance, we would like you to do an activity.

ACTIVITY 7.4

If you can get hold of a stool which can rotate without much friction, you can perform an interesting experiment. Ask a friend to sit on the stool with her arms folded. Make the stool rotate fast. Measure the speed of rotation. Ask your friend to stretch her arms and measure the speed again. Do you note any change in the speed of rotation of the stool? Ask her to fold her arms once again and observe the change in the speed of the stool.

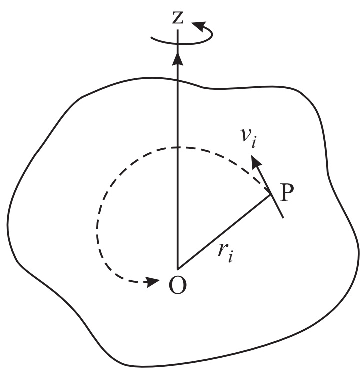


Fig. 7.23 : A rigid body rotating about an axis through 'O'

Let us try to understand why we expect a change in the speed of rotation of the stool in two cases : sitting with folded and stretched hands. For this, let us again consider a rigid body rotating about an axis, say z–axis through a fixed point O in the body. All the points of the body describe circular paths about the axis of rotation with the centres of the paths on the axis and have angular velocity ω. Consider a particle like P at distance r_i from the axis (Fig. 7.20). Its linear velocity is $v_i = r_i\omega$ and its momentum is therefore $m_i r_i \omega$. ***The product of linear momentum and the distance from the axis is called angular momentum, denoted by L.*** If we sum this product for all the particles of the body, we get

$$L = \sum_i m_i \omega r_i r_i = \left(\sum_i m_i r_i^2\right)\omega$$

$$= I\omega \qquad (7.23)$$

Remember that the angular velocity is the same for all the particles and the term within brackets is the moment of inertia. Like the linear momentum, the angular momentum is also a vector quantity. Eqn. (7.23) gives only the component of the vector **L** along the axis of rotation. It is important to remember that I must refer to the same axis. The unit of angular momentum is kg $m^2 s^{-1}$

Recall now that the rate of change of ω is α and I α = τ. Therefore, ***the rate of change of angular momentum is equal to torque.*** In vector notation, we write the equation of motion of a rotating body as

$$\frac{d\mathbf{L}}{dt} = \boldsymbol{\tau} = I\frac{d\omega}{dt} = I\,\alpha \qquad (7.24)$$

7.4.1 Conservation of angular momentum

Eqn. (7.24) shows that ***if there is no net torque acting on the body,*** $\frac{d\mathbf{L}}{dt} = 0$. ***This means that there is no change in angular momentum, i.e. the angular momentum is constant. This is the principle of conservation of angular momentum.*** Along with the conservation of energy and linear momentum, this is one of the most important principles of physics.

The principle of conservation of angular momentum allows us to answer questions such as : How the direction of toy umbrella floating in air remains fixed? The trick is to make it rotate and thereby impart it some angular momentum. Once it goes in air, there is no torque acting on it. Its angular momentum is then constant. Since angular momentum is a vector quantity, its constancy implies fixed direction and magnitude. Thus, the direction of the toy umbrella remains fixed while it is in air.

In the case of your friend on the rotating stool; when no net torque acts on the stool, the angular momentum of the stool and the person on it must be conserved. When the arms are stretched, she causes the moment of inertia of the system to increase. Eqn. (7.23) then implies that the angular velocity must decrease. Similarly, when she folds her arms, the moment of inertia of the system decreases. This causes the angular velocity to increase. Note that the change is basically caused by the change in the moment of inertia due to change in distance of particles from the axis of rotation.

Let us look at a few more examples of conservation of angular momentum. Suppose we have a spherical ball of mass M and radius R. The ball is set rotating by applying a torque on it. The torque is then removed. When there is no external torque, whatever angular momentum the ball has acquired must be conserved. Since moment of inertia of the ball is (2/5) MR^2 (Table 7.2), its angular momentum is given by

$$L = \frac{2}{5} M R^2 \omega \qquad (7.25)$$

where ω is its angular velocity. Imagine now that the radius of the ball somehow decreases. To conserve its angular momentum, ω must increase and the ball must rotate faster. This is what really happens to some stars, such as those which become pulsars (see Box on page 176).

What would happen if the radius of the ball were to increase suddenly?

You can again use Eqn.(7.25) to show that if R increases, ω must decrease to conserve angular momentum. If instead of radius, the moment of inertia of the system changes some how, ω will change again. For an interesting effect of this kind see Box below

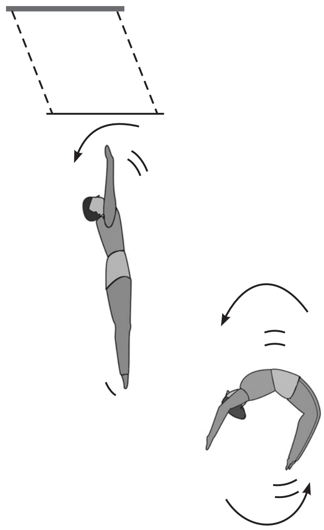


Fig. 7.24 : Diver, Sommer saulting after jumping off the diving boards.

The length of the day is not constant

Scientists have observed very small and irregular variations in the period of rotation of the earth about its axis, i.e. the length of the day. One of the causes that they have identified is weather. Due to changes in weather, large scale movement in the air of the earth's atmosphere takes place. This causes a change in the mass distribution around the axis of the earth, resulting in a change in the moment of inertia of the earth. Since the angular momentum of the earth $L = I\,\omega$ must be conserved, a change in I means a change in rotational speed of the earth, or in the length of the day.

Acrobats, skaters, divers and other sports persons make excellent use of the principle of conservation of angular momentum to show off their feats. You must have seen divers jumping off the diving boards during swimming events in national or international events such as Asian Games, Olympics or National meets. At the time of jumping, the diver gives herself a slight rotation, by which she acquires some angular momentum. When she is in air, there is no torque acting on her and therefore her angular momentum must be conserved. If she folds her body to decrease her moment of inertia (Fig. 7.24) her rotation must become faster. If she unfolds her body, her moment of inertia increases and she must rotate slowly. In this way, by controlling the shape of her body, the diver is able to demonstrate her feat before entering into pool of water.

Example 7.7 : Shiela stands at the centre of a rotating platform that has frictionless bearings. She holds a 2.0 kg object in each hand at 1.0 m from the axis of rotation of the system. The system is initially rotating at 10 rotations per minute. Calculate a) the initial angular velocity in rad s^{-1}, b) the angular velocity after the objects are brought to a distance of 0.2 m from the axis of rotation, and (c) change in the kinetic energy of the system. (d) If the kinetic energy has increased, what is the cause of this increase? (Assume that the moment of inertia of Shiela and platform I_{SP} stays constant at 1.0 kg m^2.)

Solution : (a) 1 rotation = 2π radian

$$\therefore \text{ initial angular velocity } \omega = \frac{10 \times 2\pi \text{ radian}}{60 \text{ s}} = 1.05 \text{ rad s}^{-1}$$

(b) The key idea here is to use the law of conservation of angular momentum. The initial moment of inertia $I_i = I_{SP} + m\,r_i^2 + m\,r_i^2$

$$= 1.0 \text{ kg m}^2 + (2.0 \text{ kg}) \times (1 \text{ m}^2) + (2.0 \text{ kg}) \times (1 \text{ m}^2)$$

$$= 5.0 \text{ kg m}^2.$$

After the objects are brought to a distance of 0.2 m, final moment of inertia.

$$I_t = I_{SP} + mr_f^2 + mr_f^2$$

$$= 1.0 \text{ kg m}^2 + 2.0 \text{ kg} \times (0.2)^2 \text{ m}^2 + 2.0 \text{ kg} \times (0.2)^2 \text{ m}^2$$

$$= 1.16 \text{ kg m}^2.$$

Conservation of angular momentum requires that

$$I_i\omega_i = I_f\,\omega_f$$

or

$$\omega_f = \frac{I_i\omega_i}{I_f}$$

$$= \frac{(5.0 \text{ kg m}^2) \times 1.05 \text{ rad s}^{-1}}{1.16 \text{ kg m}^2}$$

$$= 4.5 \text{ rad s}^{-1}$$

Suppose the change in kinetic energy of rotation is ΔE. Then

$$\Delta E = \frac{1}{2} I_f\,\omega_f^2 - \frac{1}{2} I_i\,\omega_i^2$$

$$= \frac{1}{2} \times 1.16 \text{ kg m}^2 \times (4.5)^2 (\text{rad s}^{-1})^2 - \frac{1}{2} \times 5.0 \text{ kg m}^2 \times (1.05)^2 (\text{rad s}^{-1})^2$$

$$= 9.05 \text{J}$$

Since final kinetic energy is higher than the initial kinetic energy, there is an increase in the kinetic energy of the system.

(d) When Shiela pulls the objects towards the axis, she does work on the system. This work goes into the system and increases its kinetic energy.

INTEXT QUESTIONS 7.4

1. A hydrogen molecule consists of two identical atoms, each of mass m and separated by a fixed distance d. The molecule rotates about an axis which is halfway between the two atoms, with angular speed ω. Calculate the angular momentum of the molecule.

2. A uniform circular disc of mass 2.0 kg and radius 20 cm is rotated about one of its diameters at an angular speed of 10 rad s^{-1}. Calculate its angular momentum about the axis of rotation.

3. A wheel is rotating at an angular speed ω about its axis which is kept vertical. Another wheel of the same radius but half the mass, initially at rest, is slipped on the same axle gently. These two wheels then rotate with a common speed. Calculate the common angular speed.

4. It is said that the earth was formed from a contracting gas cloud. Suppose some time in the past, the radius of the earth was 25 times its present radius. What was then its period of rotation on its own axis?

7.5 SIMULTANEOUS ROTATIONAL AND TRANSLATIONAL MOTIONS

We have already noted that if a point in a rigid body is not fixed, it can possess rotational motion as well as translational motion. The general motion of a rigid body consists of both these motions. Imagine the motion of an automobile wheel on a plane horizontal surface. Suppose you are observing the circular face (Fig.7.25). Fix your attention at a point P and at the centre C of the circular face. Remember that the centre of mass of the wheel lies at the centre of its axis and C is the end point of the axis. As it rolls, you would notice that point P rotates round the point C. The point C itself gets translated in the direction of motion. So the wheel has both the rotational and translational motions. If point C or the centre of mass gets translated with velocity υ_{cm}, the kinetic energy of translation is

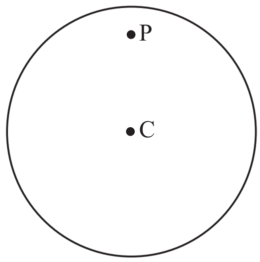


Fig. 7.25

$$(KE)_{tr} = \frac{1}{2} M \upsilon_{cm}^2 \quad (7.26)$$

where M is the mass. And if ω is the angular speed of rotation, the kinetic energy of rotation is

$$(KE)_{rot} = \frac{1}{2} I \omega^2 \quad (7.27)$$

where I is the moment of inertia. The total energy of the body due to translation and rotation is the sum of these two kinetic energies. An interesting case, where both translational and rotational motion are involved, is the motion of a body on an inclined plane.

Example 7.8 : Suppose a rigid body has mass M, radius R and moment of inertia I. It is rolling down an inclined plane of height h (Fig.7.26). At the end of its journey, it has acquired a linear speed υ and an angular speed ω. Assume that the loss of energy due to friction is small and can be neglected. Obtain the value of υ in terms of h.

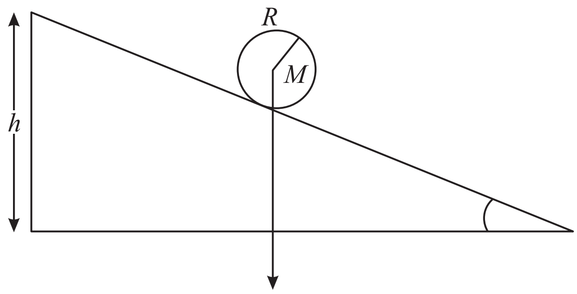


Fig. 7.26 : Motion of a rigid body on an inclined plane

Solution : The principle of conservation of energy implies that the sum of the kinetic energies due to translation and rotation at the foot of the inclined plane must be equal to the potential energy that the body had at the top of the inclined plane. Therefore,

$$\frac{1}{2} M\upsilon^2 + \tfrac{1}{2} I\,\omega^2 = M\,g\,h \qquad (7.28)$$

If the motion is pure rolling and there is no slipping, we can write $\upsilon = R\,\omega$. Inserting this expression is Eqn. (7.28), we get

$$\frac{1}{2} M\,\upsilon^2 + \frac{1}{2}\,I\,\frac{\upsilon^2}{R^2} = M\,g\,h \qquad (7.29)$$

To take a simple example, let the body be a hoop. Table 7.2 shows that its moment of inertia about its own axis is MR^2. Eqn.(7.29) then gives

$$\frac{1}{2} M\,\upsilon^2 + \frac{1}{2}\,\frac{M\,R^2\,\upsilon^2}{R^2} = M\,g\,h$$

or

$$\upsilon = \sqrt{g\,h} \qquad (7.30)$$

Do you notice any thing interesting in this equation? ***The linear velocity is independent of mass and radius of the hoop. Its means that a hoop of any material and any radius rolls down with the same speed on the inclined plane.***

INTEXT QUESTIONS 7.5

1. A solid sphere rolls down a slope without slipping. What will be its velocity in terms of the height of the slope?

2. A solid cylinder rolls down an inclined plane without slipping. What fraction of its kinetic energy is translational? What is the magnitude of its velocity after falling through a height *h*?

3. A uniform sphere of mass 2 kg and radius 10cm is released from rest on an inclined plane which makes an angle of 30^0 with the horizontal. Deduce its (a) angular acceleration, (b) linear acceleration along the plane, and (c) kinetic energy as it travels 2m along the plane.

Secret of Pulsars

An interesting example of the conservation of angular momentum is provided by pulsating stars. These are called pulsars. These stars send pulses of radiation of great intensity towards us. The pulses are periodic and the periodicity is extremely precise. The time periods range between a few milliseconds to a few seconds. Such short time periods show that the stars are rotating very fast. Most of the matter of these stars is in the form of neutrons. (The neutrons and protons are the building blocks of the atomic nuclei.) These stars are also called neutron stars. These stars represent the last stage in their life. The secret of their fast rotation is their tiny size. The radius of a typical neutron star is only 10 km. Compare this with the radius of the Sun, which is about 7×10^5 km. The Sun rotates on its axis with a period of about 25 days. Imagine that the Sun suddenly shrinks to the size of a neutron star without any change in its mass. In order to conserve its angular momentum, the Sun will have to rotate with a period as short as the fraction of a millisecond.

WHAT YOU HAVE LEARNT

- A rigid body can have rotational as well as translational motion.
- The equation of translational motion far a rigid body may be written in the same form as for a single particle in terms of the motion of its centre of mass.
- If a point in the rigid body is fixed, then it can possess only rotational motion.
- The moment of inertia about an axis of rotation is defined as $\sum_i m_i \, r_i^2$.
- The moment of inertia plays the same role in rotational motion as does the mass in linear motion.

- The turning effect of a force **F** on a rigid body is given by the torque $\boldsymbol{\tau} = \mathbf{r} \times \mathbf{F}$.
- Two equal and opposite forces constitute a couple. The magnitude of turning effect of torque is equal to the product of one of the forces and the perpendicular distance between the line of action of forces.
- The application of an external torque changes the angular momentum of the body.
- When no net torque acts on a body, the angular momentum of the body remains constant.
- When a cylindrical or a spherical body rolls down an inclined plane without slipping, its speed is independent of its mass and radius.

TERMINAL EXERCISE

1. The weight Mg of a body is shown generally as acting at the centre of mass of the body. Does this mean that the earth does not attract other particles?
2. Is it possible for the centre of mass of a body to lie outside the body? Give two examples to justify your answer?
3. In a molecule of carbon monoxide (CO), the nuclei of the two atoms are 1.13×10^{-10}m apart. Locate of the centre of mass of the molecule.
4. A grinding wheel of mass 5.0 kg and diameter 0.20 m is rotating with an angular speed of 100 rad s^{-1}. Calculate its kinetic energy. Through what distance would it have to be dropped in free fall to acquire this kinetic energy? (Take g = 10.0 m s^{-2}).
5. A wheel of diameter 1.0 m is rotating about a fixed axis with an initial angular speed of 2rev s^{-1}. The angular acceleration is 3 rev s^{-2}.

   (a) Compute the angular velocity after 2 seconds.

   (b) Through what angle would the wheel turned during this time?

   (c) What is the tangential velocity of a point on the rim of the wheel at t = 2 s?

   (d) What is the centripetal acceleration of a point on the rim of the wheel at t = 2 s?
6. A wheel rotating at an angular speed of 20 rads^{-1} is brought to rest by a constant torque in 4.0 seconds. If the moment of inertia of the wheel about

the axis of rotation is 0.20 kg m^2, calculate the work done by the torque in the first two seconds.

7. Two wheels are mounted on the same axle. The moment of inertia of wheel A is 5×10^{-2} kg m^2, and that of wheel B is 0.2 kg m^2. Wheel A is set spinning at 600 rev min^{-1}. while wheel B is stationary. A clutch now acts to join A and B so that they must spin together.
   (a) At what speed will they rotate?
   (b) How does the rotational kinetic energy before joining compare with the kinetic energy after joining?
   (c) What torque does the clutch deliver if A makes 10 revolutions during the operation of the clutch?

8. You are given two identically looking spheres and told that one of them is hollow. Suggest a method to detect the hollow one.

9. The moment of inertia of a wheel is 1000 kg m^2. Its rotation is uniformly accelerated. At some instant of time, its angular speed is 10 rad s^{-1}. After the wheel has rotated through an angle of 100 radians, the angular velocity of the wheel becomes 100 rad s^{-1}. Calculate the torque applied to the wheel and the change in its kinetic energy.

10. A disc of radius 10 cm and mass 1kg is rotating about its own axis. It is accelerated uniformly from rest. During the first second it rotates through 2.5 radians. Find the angle rotated during the next second. What is the magnitude of the torque acting on the disc?

ANSWERS TO INTEXT QUESTIONS

7.1

1. Yes, because the distances between points on the frame cannot change.

2. No. Any disturbance can change the distance between sand particles. So, a heap of sand cannot be considered a rigid body.

7.2

1. The coordinates of given five masses are A (–1, –1), B (–5, –1), C (6, 3), D (2, 6) and E (–3, 0) and their masses are 1 kg, 2kg, 3kg, 4kg and 5kg respectively.

   Hence, coordinates of centre of mass of the system are

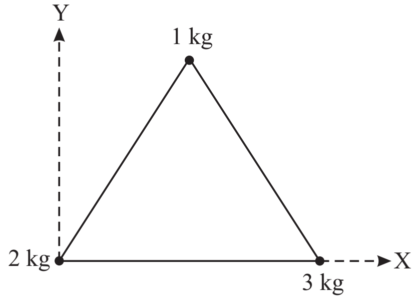
$$x = \frac{-1\times1-5\times2+6\times3+2\times4-3\times5}{1+2+3+4+5} = 0$$

$$y = \frac{-1\times1-1\times2+3\times3+4\times6+0\times5}{1+2+3+4+5} = \frac{30}{15} = 2.0$$

2. Let the three particle system be as shown in the figure here. Consider axes to be as shown with 2 kg mass at the origin.

$$x = \frac{2\times0+1\times0.5+3\times1}{1+2+3} = \frac{3.5}{6}\text{ m} = 0.5\text{ m}$$

$$y = \frac{2\times0+1\times\frac{\sqrt{3}}{2}+3\times0}{1+2+3} = \frac{\sqrt{3}}{12}\text{ m}$$



Hence, the co-ordinates of the centre of mass are $\left(\frac{3.5}{6}, \frac{\sqrt{3}}{12}\right)$

3. Let the two particles be along the x-axis and let their x-coordinates be o and x. The coordinate of CM is

$$X = \frac{m_1\times0+m_2\times x}{m_1+m_2} = \frac{m_2 x}{m_1+m_2}, \ Y = 0$$

X is also the distance of m_1 from the CM. The distance of m_2 from CM is

$$x - X = x - \frac{m_2 x}{m_1+m_2} = \frac{m_1 x}{m_1+m_2}$$

$$\therefore \frac{X}{x+X} = \frac{m_2}{m_1}$$

Thus, the distances from the CM are inversely proportional to their masses.

7.3

1. Moment of inertia of the system about an axis perpendicular to the plane passing through one of the corners and perpendicular to the plane of the square,

$= m\,r^2 + m\,(2\,r^2) + m\,r^2 = 4\,m\,r^2$

M.I. about the axis along the side $= m\,r^2 + m\,r^2 = 2\,m\,r^2$

Verification : Moment of inertia about the axis QP = $m\,r^2 + m\,r^2 + 2\,m\,r^2$. Now, according to the theorem of perpendicular axes, MI about SP ($2mr^2$) + MI about QP $2\,m\,r^2$ should be equal to MI about the axis through P and perpendicular to the plane of the square ($4\,m\,r^2$). Since it is true, the results are verified.

2. M.I. of solid sphere about an axis tangential to the sphere

$$= \frac{2}{5}\,M\,R^2 + M\,R^2 = \frac{7}{5}\,M\,R^2$$ according to the theorem of parallel axes.

If radius fo gyration is K, then $M\,K^2 = \frac{7}{5}\,M\,R^2$. So,

Radius of gyration $K = R\sqrt{\frac{7}{5}}$

7.4

1. Angular momentum

$$L = \left(m\frac{d^2}{4} + m\frac{d^2}{4}\right)\omega$$

$$L = \frac{m\,d^2\omega}{2}$$

2. Angular momentum about an axis of rotation (diameter).

$$L = I\,\omega = m\,\frac{r^2}{4} \times \omega$$

as M.I about a diameter $= \frac{m\,r^2}{4}$

$$\therefore\; L = 20\text{ kg} \times \frac{(0.2)^2 m^2}{4} \times 10\text{ rad s}^{-1} = 0.2\text{ kg m}^2\text{ s}^{-1}.$$

3. According to conservation of angular momentum

$$I_1\omega = (I_1 + I_2)\,\omega_1$$

where I_1 is M.I. of the original wheel and I_2 that of the other wheel, ω is the initial angular speed and ω_1 is the common final angular speed.

$$m\,r^2\,\omega = \left(m\,r^2 + \frac{m}{2}r^2\right)\omega_1$$

$$\omega = \frac{3}{2}\,\omega_1 \Rightarrow \omega_1 = \frac{2}{3}\,\omega$$

4. Let the present period of revolution of earth be T and earlier be T_0. According to the conservation of angular momentum.

$$\frac{2}{5}M\,(25\,R)^2 \times \left(\frac{2\pi}{T_0}\right) = \frac{2}{5}\,M\,R^2 \times \left(\frac{2\pi}{T}\right)$$

$$= \frac{2}{5} M R^2 \times \left(\frac{2\pi}{T}\right)$$

It gives, $T_0 = 6.25\ T$

Thus, period of revolution of earth in the past T_0 = 6.25 times the present time period.

7.5

1. Using ($I = \frac{2}{5} M R^2$), Eqn. (7.29) for a solid sphere

$$\frac{1}{2} m\upsilon^2 + \frac{1}{2} I \omega^2 = m g h$$

or, $$\frac{1}{2} m\upsilon^2 + \frac{1}{2} \times \frac{2}{5} m r^2 . \frac{\upsilon^2}{r^2} = m g h$$

$\because \omega = \upsilon/r.$

It gives $\upsilon = \sqrt{\frac{10}{7} . g h}$

2. For a solid cylinder, $I = \frac{m R^2}{2}$

$$\therefore \text{Total K.E } \frac{1}{2} m \upsilon^2 + \frac{1}{2} I \omega^2 = \frac{1}{2} m \upsilon^2 + \frac{1}{2} \frac{m R^2}{2} . \frac{\upsilon^2}{R^2} = \frac{3}{4} m \upsilon^2$$

$\because \omega = \upsilon/r$

Hence, fraction of translational K.E. $= \dfrac{\frac{1}{2} m \upsilon^2}{\frac{3}{4} m \upsilon^2} = \dfrac{2}{3}$

Proceeding as in Q.1 above : $\upsilon = \sqrt{\frac{4}{3} g h}$

Answers to Terminal Problems

3. At a distance 0.64 Å from carbon atom.
4. 125 J, 2.5 m
5. (a) 16 π rad s^{-1} (b) 20 π rad (c) 25 m s^{-1} (d) 1280 m s^{-2}
6. 30 J
7. (a) 4 π rad s^{-1} (b) $E_i = 5\ E_f$ (c) 49 π N m
9. T = 5 × 10^4 N m, KE = 5 × 10^6 J
10. 7.5 rad, τ = 5 × 10^{-2} J

8

312en08

ELASTIC PROPERTIES OF SOLIDS

In the previous lessons you have studied the effect of force on a body to produce displacement. The force applied on an object may also change its shape or size. For example, when a suitable force is applied on a spring, you will find that its shape as well as size changes. But when you remove the force, it will regain original position. Now apply a force on some objects like wet modelling clay or molten wax. Do they regain their original position after the force has been removed? They do not regain their original shape and size. Thus some objects regain their original shape and size whereas others do not. Such a behaviour of objects depends on a property of matter called **elasticity**.

The elastic property of materials is of vital importance in our daily life. It is used to help us determine the strength of cables to support the weight of bodies such as in cable cars, cranes, lifts etc. We use this property to find the strength of beams for construction of buildings and bridges. In this unit you will learn about nature of changes and the manner in which these can be described.

OBJECTIVES

After studying this lesson, you should be able to :

- *distinguish between three states of matter on the basis of molecular theory;*
- *distinguish between elastic and plastic bodies;*
- *distinguish between stress and pressure;*
- *study stress-strain curve for an elastic solid;*
- *define Young's modulus, bulk modulus, modulus of rigidity and Poisson's ratio; and*
- *derive an expression for the elastic potential energy of a spring.*

8.1 MOLECULAR THEORY OF MATTER : INTER-MOLECULAR FORCES

We know that matter is made up of atoms and molecules. The forces which act between them are responsible for the structure of matter. The interaction forces between molecules are known as **inter-molecular forces**.

The variation of inter molecular forces with inter molecular separation is shown in Fig. 8.1.

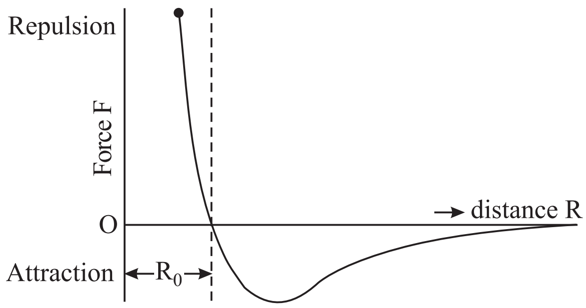


Fig. 8.1 : Graph between inter-molecular force and Inter molecular separation.

When the separation is large, the force between two molecules is attractive and weak. As the separation decreases, the net force of attraction increases up to a particular value and beyond this, the force becomes repulsive. At a distance $R = R_0$ the net force between the molecules is zero. This separation is called **equilibrium separation**. Thus, if inter-molecular separation $R > R_0$ there will be an attractive force between molecules. When $R < R_0$, a repulsive force will act between them.

In **solids**, molecules are very close to each other at their equilibrium separation ($\sim 10^{-10}$ m). Due to high intermolecular forces, they are almost fixed at their positions. You may now appreciate why a solid has a definite shape.

In **liquids**, the average separation between the molecules is somewhat larger ($\sim 10^{-8}$ m). The attractive force is weak and the molecules are comparatively free to move inside the whole mass of the liquid. You can understand now why a liquid does not have fixed shape. It takes the shape of the vessel in which it is filled.

In **gases**, the intermolecular separation is significantly larger and the molecular force is very weak (almost negligible). Molecules of a gas are almost free to move inside a container. That is why gases do not have fixed shape and size.

Ancient Indian view about Atom

Kanada was the first expounder of the atomic concept in the world. He lived around 6th century B.C. He resided at Prabhasa (near Allahabad).

According to him, everything in the universe is made up of **Parmanu** or Atom. They are eternal and indestructible. Atoms combine to form different molecules. If two atoms combine to form a molecule, it is called **duyanuka** and a triatomic molecule is called triyanuka. He was the author of "Vaisesika Sutra".

The size of atom was also estimated. In the biography of Buddha (Lalitavistara), the estimate of atomic size is recorded to be of the order 10^{-10} m, which is very close to the modern estimate of atomic size.

8.2 ELASTICITY

You would have noticed that when an external force is applied on an object, its shape or size (or both) change, i.e. deformation takes place. The extent of deformation depends on the material and shape of the body and the external force. When the deforming forces are withdrawn, the body tries to regain its original shape and size.

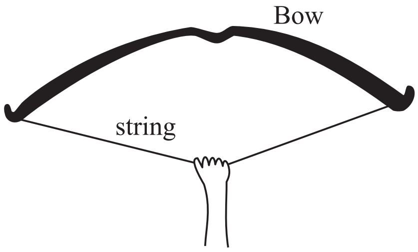


Fig 8.2 : Force applied on the string of a bow changes it shape

You may compare this with a spring loaded with a mass or a force applied on the string of a bow or pressing of a rubber ball. If you apply a force on the string of the bow to pull it (Fig 8.2), you will observe that its shape changes. But on releasing the string, the bow regains its original shape and size.

The property of matter to regain its original shape and size after removal of the deforming forces is called **elasticity**.

8.2.1 Elastic and Plastic Bodies

A body which regains its original state completely on removal of the deforming force is called **perfectly elastic**. On the other hand, if it completely retains its modified form even on removing the deforming force, i.e. shows no tendency to recover the deformation, it is said to be **perfectly plastic**. However, in practice the behaviour of all bodies is in between these two limits. There exists no perfectly elastic or perfectly plastic body in nature. The nearest approach to a perfectly elastic body is quartz fiber and to the perfectly plastic is ordinary putty. Here it can be added that the object which opposes the deformation more is more elastic. No doubt elastic deformations are very important in science and technology, but plastic deformations are also important in mechanical processes. You might have seen the processes such as stamping, bending and hammering of metal pieces. These are possible only due to plastic deformations.

The phenomenon of elasticity can be explained in terms of inter-molecular forces.

8.2.2 Molecular Theory of Elasticity

You are aware that a solid is composed of a large number of atoms arranged in a definite order. Each atom is acted upon by forces due to neighbouring atoms.

Due to inter-atomic forces, solid takes such a shape that each atom remains in a stable equilibrium. When the body is deformed, the atoms are displaced from their original positions and the inter-atomic distances change. If in deformation, the separation increases beyond their equilibrium separation (i.e., $R > R_0$), strong attractive forces are developed. However, if inter–atomic separation decreases (i.e. $R < R_0$), strong repulsive forces develop. These forces, called **restoring forces,** drive atoms to their original positions. The behaviour of atoms in a solid can be compared to a system in which balls are connected with springs.

Now, let us learn how forces are applied to deform a body.

8.2.3 Stress

When an external force or system of forces is applied on a body, it undergoes a change in the shape or size according to nature of the forces. We have explained that in the process of deformation, internal restoring force is developed due to molecular displacements from their positions of equilibrium. The internal restoring force opposes the deforming force. **The internal restoring force acting per unit area of cross-section of a deformed body is called *stress*.**

In equilibrium, the restoring force is equal in magnitude and opposite in direction to the external deforming force. Hence, stress is measured by the external force per unit area of cross-section when equilibrium is attained. If the magnitude of deforming force is F and it acts on area A, we can write

$$\text{Stress} = \frac{\text{restoring force}}{\text{area}} = \frac{\text{deforming force } (F)}{\text{area } (A)}$$

or

$$\text{Stress} = \frac{F}{A} \qquad (8.1)$$

The unit of stress is Nm^{-2}. The stress may be longitudinal, normal or shearing. Let us study them one by one.

(i) **Longitudinal Stress :** If the deforming forces are along the length of the body, we call the stress produced as **longitudinal stress,** as shown in its two forms in Fig 8.3 (a) and Fig 8.3 (b).

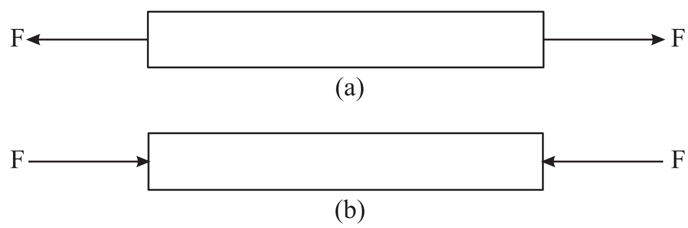
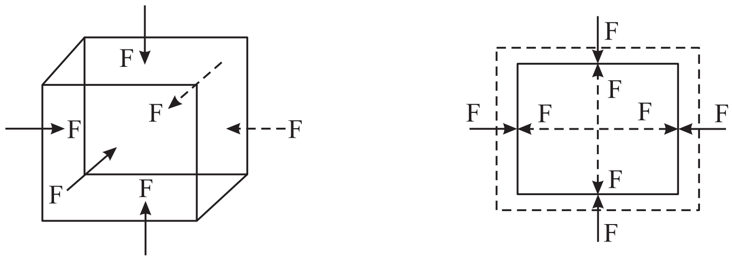


Fig. 8.3 (a) : Tensile stress; (b) Compressive stress

(ii) **Normal Stress :** If the deforming forces are applied uniformly and normally all over the surface of the body so that the change in its volume occurs without change in shape (Fig. 8.4), we call the stress produced as normal stress. You may produce normal stress by applying force uniformly over the entire surface of the body. Deforming force per unit area normal to the surface is called pressure while restoring force developed inside the body per unit area normal to the surface is known as **stress**.



(a) (b)

Fig. 8.4 : Normal stress

(iii) **Shearing Stress :** If the deforming forces act tangentially or parallel to the surface (Fig 8.5a) so that shape of the body changes without change in volume, the stress is called **shearing stress**. An example of shearing stress is shown in Fig 8.5 (b) in which a book is pushed side ways. Its opposite face is held fixed by the force of friction.

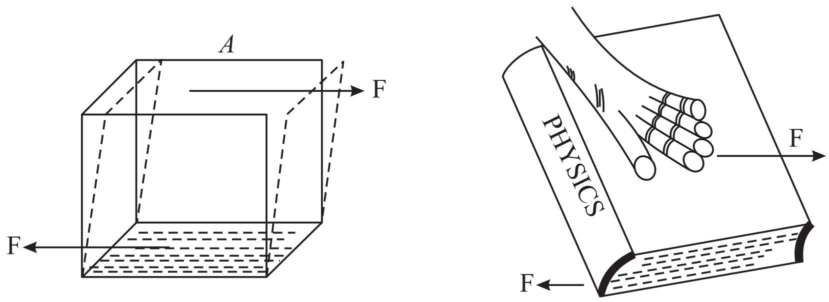


Fig. 8.5: (a) Shearing stress; (b) Pushing a book side ways

8.2.4 Strain

Deforming forces produce changes in the dimensions of the body. In general, the **strain is defined as the change in dimension (e.g. length, shape or volume) per unit dimension of the body.** As the strain is ratio of two similar quantities, it is a dimensionless quantity.

Depending on the kind of stress applied, strains are of three types : (i) linear strain,(ii) volume (bulk) strain, and (iii) shearing strain.

(i) **Linear Strain :** If on application of a longitudinal deforming force, the length ℓ of a body changes by $\Delta\ell$ (Fig. 8.6), then

Fig. 8.6: Linear strain

$$\text{linear strain} = \frac{\text{change in length}}{\text{original length}} = \frac{\Delta\ell}{\ell}$$

(ii) **Volume Strain :** If on application of a uniform pressure Δp, the volume V of the body changes by ΔV (Fig 8.7) without change of shape of the body, then

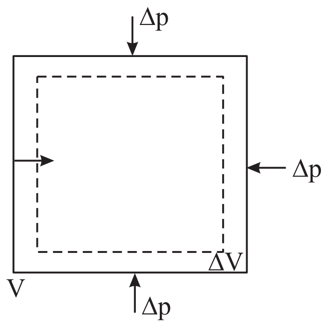
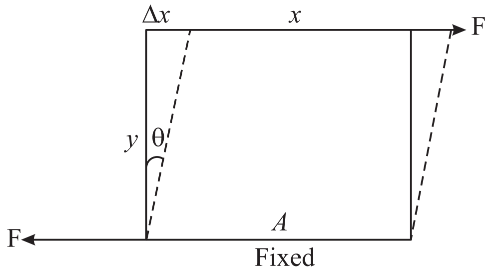


Fig. 8.7: Volume strain

$$\text{Volume strain} = \frac{\text{change in volume}}{\text{original volume}} = \frac{\Delta V}{V}$$

(i) **Shearing strain:** When the deforming forces are tangential (Fig 8.8), the shearing strain is given by the angle θ through which a line perpendicular to the fixed plane is turned due to deformation. (The angle θ is usually very small.) Then we can write



$$\theta = \frac{\Delta x}{y}$$

Fig. 8.8 : Shearing strain

8.2.5 Stress-strain Curve for a Metallic Wire

Refer to Fig. 8.9 which shows variation of stress with strain when a metallic wire of uniform cross-section is subjected to an increasing load. Let us study the regions and points

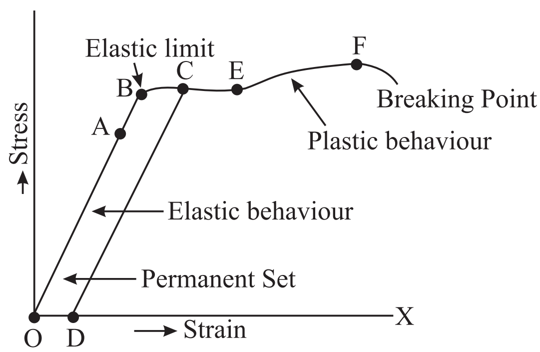


Fig. 8.9: Stress-strain curve for a steel wire

on this curve that are of particular importance.

(i) **Region of Proportionality:** OA is a straight line which indicates that in this region, stress is linearly proportional to strain and the body behaves like a perfectly elastic body.

(ii) **Elastic Limit** : If we increase the strain a little beyond A, the stress is not linearly proportional to strain. However, the wire still remains elastic, i.e. after removing the deforming force (load), it regains its original state. The maximum value of strain for which a body(wire) shows elastic property is called **elastic limit.** Beyond the elastic limit, a body behaves like a plastic body.

(iii) **Point** C : When the wire is stretched beyond the limit B, the strain increases more rapidly and the body becomes plastic. It means that even if the deforming load is removed, the wire will not recover its original length. The material follows dotted line CD on the graph on gradual reduction of load. The left over strain on zero load strain is known as a **permanent set**. After point E on the curve, no extension is recoverable.

(iv) **Breaking point** F : Beyond point E, strain increases very rapidly and near point F, the length of the wire increases continuously even without increasing of load. The wire breaks at point F. This is called the **breaking point** or **fracture point** and the corresponding stress is known as **breaking stress**.

The stress corresponding to breaking point F is called **breaking stress** or **tensile strength**. Within the elastic limit, the maximum stress which an object can be subjected to is called **working stress** and the ratio between working stress and breaking stress is called **factor of safety.** In U.K, it is taken 10, in USA it is 5. We have adopted UK norms. If large deformation takes place between the elastic limit and the breaking point, the material is called **ductile**. If it breaks soon after the elastic limit is crossed, it is called **brittle e.g.** glass.

8.2.6 Stress-Strain Curve for Rubber

When we stretch a rubber cord to a few times its natural length, it returns to its original length after removal of the forces. That is, the elastic region is large and there is no well defined plastic flow region. Substances having large strain are called **elastomers**. This property arises from their molecular arrangements. The stress-strain curve for rubber is distinctly different from that of a metallic wire. There are two important things to note from Fig. 8.10. Firstly, you can observe that there is no region of proportionality. Secondly, when the deforming force is gradually reduced, the original curve is not retraced, although the sample finally acquires its natural length. The work done by the material in returning to its original shape is less than the work done by the deforming force. This difference of energy is absorbed by the material and appears as heat. (You can feel it by touching the rubber band with your lips.) This phenomenon is called **elastic hysteresis**.

Elastic hysteresis has an important application in **shock absorbers**. A part of energy transferred by the deforming force is retained in a shock absorber and only a small part of it is transmitted to the body to which the shock absorber is attached.

8.2.7 Steel is more Elastic than Rubber

A body is said to be more elastic if on applying a large deforming force on it, the strain produced in the body is small. If you take two identical rubber and steel wires and apply equal deforming forces on both of them, you will see that the extension produced in the steel wire is smaller than the extension produced in the rubber wire. But to produce same strain in the two wires, significantly higher stress is required in the steel wire than in rubber wire. Large amount of stress needed for deformation of steel indicates that magnitude of internal restoring force produced in steel is higher than that in rubber. Thus, steel is more elastic than rubber.

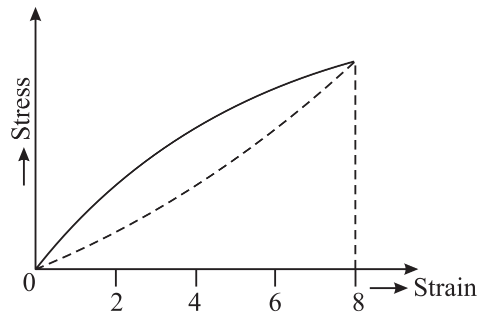


Fig. 8.10: Stress-strain curve for rubber

Example 8.1 : A load of 100 kg is suspended by a wire of length 1.0 m and cross sectional area 0.10 cm^2. The wire is stretched by 0.20 cm. Calculate the (i) tensile stress, and (ii) strain in the wire. Given, $g = 9.80$ ms^{-2}.

Solution :

$$\text{(i) Tensile stress} = \frac{F}{A} = \frac{Mg}{A}$$

$$= \frac{(100\,\text{kg})\,(9.80\,\text{ms}^{-2})}{0.10\times10^{-4}\,\text{m}^2}$$

$$= 9.8 \text{ x } 10^7\,\text{Nm}^{-2}$$

$$\text{(ii) Tensile strain} = \frac{\Delta\ell}{\ell} = \frac{0.20\times10^{-2}\,\text{m}}{1.0\,\text{m}}$$

$$= 0.20 \times 10^{-2}$$

Example 8.2 : Calculate the maximum length of a steel wire that can be suspended without breaking under its own weight, if its breaking stress = 4.0 x 10^8 Nm^{-2}, density = 7.9×10^3 kg m^{-3} and g = 9.80 ms^{-2}

Solution : The weight of the wire $W = A\ell\rho g$, where, A is area of cross section of the wire, ℓ is the maximum length and ρ is the density of the wire. Therefore, the breaking stress developed in the wire due to its own weight $\frac{W}{A} = \rho l g$. We are told that

breaking stress is 4.0 x 10^8 Nm^{-2}. Hence

$$\ell = \frac{4.0 \times 10^8 \, Nm^{-2}}{(7.9 \times 10^3 \, kg \, m^{-3})\,(9.8 \, ms^{-2})}$$

$$= 0.05 \times 10^5 \, m$$

$$= 5 \times 10^3 \, m = 5 \, km.$$

Now it is time to take a break and check your understanding

INTEXT QUESTIONS 8.1

1. What will be the nature of inter-atomic forces when deforming force applied on an object (i) increases, (ii) decreases the inter-atomic separation?
2. If we clamp a rod rigidly at one end and a force is applied normally to its cross section at the other end, name the type of stress and strain?
3. The ratio of stress to strain remains constant for small deformation of a metal wire. For large deformations what will be the changes in this ratio?
4. Under what conditions, a stress is known as breaking stress ?
5. If mass of 4 kg is attached to the end of a vertical wire of length 4 m with a diameter 0.64 mm, the extension is 0.60 mm. Calculate the tensile stress and strain?

8.3 HOOKE'S LAW

In 1678, Robert Hooke obtained the stress-strain curve experimentally for a number of solid substances and established a law of elasticity known as Hooke's law. According to this law: **Within elastic limit, stress is directly proportional to corresponding strain.**

i.e. stress α strain

or $$\frac{\text{stress}}{\text{strain}} = \text{constant } (E) \qquad (8.2)$$

This constant of proportionality E is a measure of elasticity of the substance and is called **modulus of elasticity**. As strain is a dimensionless quantity, the modulus of elasticity has the same dimensions (or units) as stress. Its value is independent of the stress and strain but depends on the nature of the material. To see this, you may like to do the following activity.

ACTIVITY 8.1

Arrange a steel spring with its top fixed with a rigid support on a wall and a metre scale along its side, as shown in the Fig. 8.11.

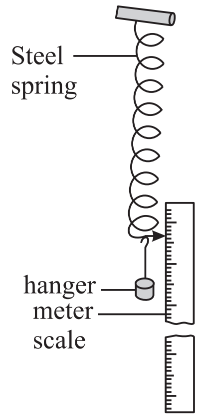


Fig. 8.11: Hooke's law apparatus

Add 100 g load at a time on the bottom of the hanger in steps. It means that while putting each 100 g load, you are increasing the stretching force by 1N. Measure the extension. Take the reading upto 500 g and note the extension each time.

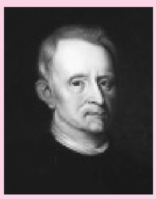
Plot a graph between load and extension. What is the shape of the graph? Does it obey Hooke's law?

The graph should be a straight line indicating that the ratio (load/ extension) is constant.

Repeat this activity with rubber and other materials.

You should know that the materials which obey Hooke's law are used in spring balances or as force measurer, as shown in the Fig. 8.11. You would have seen that when some object is placed on the pan, the length of the spring increases. This increase in length shown by the pointer on the scale can be treated as a measure of the increase in force (i.e., load applied).

Robert Hooke
(1635 – 1703)



Robert Hooke, experimental genius of seventeenth century, was a contemporary of Sir Isaac Newton. He had varied interests and contributed in the fields of physics, astronomy, chemistry, biology, geology, paleontology, architecture and naval technology. Among other accomplishments he has to his credit the invention of a universal joint, an early proto type of the respirator, the iris diaphragm, anchor escapement and balancing spring for clocks. As chief surveyor, he helped rebuild London after the great fire of 1666. He formulated Hooke's law of eleasticity and correct theory of combustion. He is also credited to invent or improve meteorological instruments such as barometer, anemometer and hygrometer.

8.3.1 Moduli of Elasticity

In previous sections, you have learnt that there are three kinds of strain. It is therefore clear that there should be three modulli of elasticity corresponding to these strains. These are **Young's modulus, Bulk Modulus and Modulus of rigidity** corresponding to linear strain, volume strain and shearing strain, respectively. We now study these one by one.

(i) **Young's Modulus:** The ratio of the longitudinal stress to the longitudinal strain is called Young's modulus for the material of the body.

Suppose that when a wire of length L and area of cross-section A is stretched by a force of magnitude F, the change in its length is equal to ΔL. Then

$$\text{Longitudinal stress} = \frac{F}{A}$$

and
$$\text{Longitudinal strain} = \frac{\Delta L}{L}$$

Hence,
$$\text{Young's modulus } Y = \frac{F/A}{\Delta L/L} = \frac{F \times L}{A \times \Delta L}$$

If the wire of radius r is suspended vertically with a rigid support and a mass M hangs at its lower end, then $A = \pi r^2$ and $F = M\,g$.

$$\therefore \qquad Y = \frac{M\,g\,L}{\pi\,r^2\,\Delta\,L} \qquad (8.3)$$

The SI unit of Y in is N m^{-2}. The values of Young's modulus for a few typical substances are given in Table. 8.1. Note that steel is most elastic.

Table 8.1. Young's modulus of some typical materials

Name of substance	Y (10^9Nm^{-2})
Aluminium	70
Copper	120
Iron	190
Steel	200
Glass	65
Bone	9
Polystyrene	3

(ii) **Bulk Modulus:** The ratio of normal stress to the volume strain is called bulk modulus of the material of the body.

If due to increase in pressure P, volume V of the body decreases by ΔV without change in shape, then

$$\text{Normal stress } = \Delta P$$

$$\text{Volume strain } = \Delta V/V$$

$$\text{Bulk modulus } B = \frac{\Delta P}{\Delta V/V} = V\frac{\Delta P}{\Delta V} \qquad (8.4)$$

The reciprocal of bulk modulus of a substance is called compressibility :

$$k = \frac{1}{B} = \frac{1}{V}\,\frac{\Delta V}{\Delta P} \qquad (8.5)$$

Gases being most compressible are least elastic while solids are most elastic or least compressible i.e. $B_{solid} > B_{liquid} > B_{gas}$

(iii) **Modulus of Rigidity or Shear Modulus:** The ratio of the shearing stress to shearing strain is called modulus of rigidity of the material of the body.

If a tangential force F acts on an area A and θ is the shearing strain, the modulus of rigidity

$$\eta = \frac{\text{Shearing stress}}{\text{Shearing strain}} = \frac{F/A}{\theta} = \frac{F}{A\theta} \qquad (8.6)$$

You should know that both solid and fluids have bulk modulus. However, fluids do not have Young's modulus and shear modulus because a liquid can not sustain a tensile or shearing stress.

Example 8.3 : Calculate the force required to increase the length of a wire of steel of cross sectional area 0.1 cm^2 by 50%. Given Y = 2 × 10^{11} N m^{-2}.

Solution : Increase in the length of wire = 50%. If ΔL is the increase and L is the normal length of wire then $\frac{\Delta L}{L} = \frac{1}{2}$

$$\therefore \qquad Y = \frac{F \times L}{A \times \Delta L}$$

or

$$F = \frac{Y \times A \times \Delta L}{L} = \frac{(2\times10^{11}\,\text{Nm}^{-2})\,(0.1\times10^{-4}\,\text{m}^2)\times1}{2}$$

$$= 0.1 \times 10^7 \text{ N} = 10^6 \text{ N}$$

Example 8.4 : When a solid rubber ball is taken from the surface to the bottom of a lake, the reduction in its volume is 0.0012 %. The depth of lake is 360 m, the density of lake water is 10^3 kgm^{-3} and acceleration due to gravity at the place is 10 m s^{-2}. Calculate the bulk modulus of rubber.

Solution :

Increase of pressure on the ball

$$P = h\rho g = 360\text{m} \times 10^3\,\text{kgm}^{-3} \times 10\text{ ms}^{-3}$$

$$= 3.6 \times 10^6\ \text{Nm}^{-2}$$

$$\text{Volume strain} = \frac{\Delta V}{V} = \frac{0.0012}{100} = 1.2 \times 10^{-5}$$

$$\text{Bulk Modulus } B = \frac{PV}{\Delta V} = \frac{3.6\times10^6}{1.2\times10^{-5}} = 3.0 \times 10^{11}\,\text{Nm}^{-2}$$

8.3.2 Poisson's Ratio

You may have noticed that when a rubber tube is stretched along its length, there is a contraction in its diameter (Fig.8.12). (This is also true for a wire but may not be easily visible.) While the length increases in the direction of forces, a contraction occurs in the perpendicular direction. The strain perpendicular to the applied force is called **lateral strain**. Poisson pointed out that within elastic limit, lateral strain is directly proportional to longitudinal strain i.e. the ratio of lateral strain to longitudinal strain is constant for a material body and is known as **Poisson's ratio**. It is denoted by a Greek letter σ (sigma). If α and β are the longitudinal strain and lateral strain respectively, then Poisson's ratio

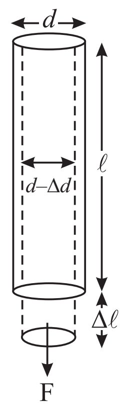


Fig. 8.12 : A stretched rubber tube.

$$\sigma = \beta/\alpha.$$

If a wire (rod or tube) of length ℓ and diameter d is elongated by applying a stretching force by an amount $\Delta\ell$ and its diameter decreases by Δd, then longitudinal strain

$$\alpha = \frac{\Delta\ell}{\ell}$$

lateral strain

$$\beta = \frac{\Delta d}{d}$$

and Possion's ratio

$$\sigma = \frac{\Delta d/d}{\Delta\ell/\ell} = \frac{\ell}{d}\,\frac{\Delta d}{\Delta\ell} \qquad (8.7)$$

Since Poisson's ratio is a ratio of two strains, it is a pure number.

The value of Poisson's ratio depends only on the nature of material and for most of the substances, it lies between 0.2 and 0.4. When a body under tension suffers no change in volume, i.e. the body is perfectly incompressible, the value of Poisson's ratio is maximum i.e. 0.5. Theoretically, the limiting values of Poisson's ratio are –1 and 0.5.

ACTIVITY 8.2

Take two identical wires. Make one wire to execute torsional vibrations for some time. After some time, set the other wire also in similar vibrations. Observe the rate of decay of vibrations of the two wires.

You will note that the vibrations decay much faster in the wire which was vibrating for longer time. The wire gets tired or fatigued and finds it difficult to continue vibrating. This phenomenon is known as **elastic fatigue**.

Some other facts about elasticity :

1. If we add some suitable impurity to a metal, its elastic properties are modified. For example, if carbon is added to iron or potassium is added to gold, their elasticity increases.
2. The increase in temperature decreases elasticity of materials. For example, carbon, which is highly elastic at ordinary temperature, becomes plastic when heated by a current through it. Similarly, plastic becomes highly elastic when cooled in liquid air.
3. The value of modulus of elasticity is independent of the magnitude of stress and strain. It depends only on the nature of the material of the body.

Example 8.5: A Metal cube of side 20 cm is subjected to a shearing stress of 10^4 Nm^{-2}. Calculate the modulus of rigidity, if top of the cube is displaced by 0.01 cm. with respect to bottom.

Solution : Shearing stress= 10^4 Nm^{-2}, Δx = 0.01 cm, and y = 20 cm.

$$\therefore \qquad \text{Shearing strain} = \frac{\Delta x}{y} = \frac{0.01\text{ cm}}{20\text{ cm}}$$

Hence, $= 0.005$

$$\text{Modulus of rigidity} \quad \eta = \frac{\text{Shearing stress}}{\text{Shearing strain}} = \frac{10^4\text{ Nm}^{-2}}{.0005}$$

$$= 2 \times 10^7 \text{ N m}^{-2}$$

Example 8.6 : A 10 kg mass is attached to one end of a copper wire of length 5 m long and 1 mm in diameter. Calculate the extension and lateral strain, if Poisson's ratio is 0.25. Given Young's modulus of the wire = 11×10^{10} N m^{-2}.

Solution : Here $L = 5$ m, $r = 0.05 \times 10^{-3}$ m, $y = 11 \times 10^{10}$ Nm^{-2} $F = 10 \times 9.8$ N, and $\sigma = 0.25$.

Extension produced in the wire

$$\Delta \ell = \frac{F.\ell}{\pi r^2 Y} = \frac{(10\,\text{kg}) \times (9.8\text{ms}^{-2}) \times (5\text{m})}{3.14\,(0.5 \times 10^{-3}\,\text{m})^2 \times (11 \times 10^{10}\,\text{Nm}^{-2})}$$

$$= \frac{490}{8.63 \times 10^4}\ \text{m}$$

$$= 5.6 \times 10^{-3}\ \text{m}$$

$$\text{longitudinal streain} = \alpha = \frac{\Delta \ell}{\ell}$$

$$= \frac{5.6 \times 10^{-3}\,\text{m}}{5\text{m}}$$

$$= 1.12 \times 10^{-2}$$

$$\text{Poission's ratio } (\sigma) = \frac{\text{lateral strain}(\beta)}{\text{longitudinal strain}(\alpha)}$$

$$\therefore \quad \text{lateral strain } \beta = \sigma \times \alpha$$

$$= 0.125 \times 1.12 \times 10^{-2}$$

$$= 0.14 \times 10^{-3}.$$

Now take a break to check your progress.

8.3.3 Elastic Energy

When a spring is either compressed or extended, it undergoes a change in its configuration and is capable of performing work.

Elastic energy is a kind of potential energy and it is the energy which is associated with the state of compression or extension of an elastic object like a spring. The force involved here is the spring force. If we compress or extend a spring, we change the relative locations of the coil of the spring. In case of a rubber like tube we change the relative locations of its different layers. A restoring force resists the change and result in work done by us due to which increases the elastic potential energy of the spring or such like objects increases.

Suppose the spring constant of a spring is k. If the spring is stretched through a distance x at any instant (Fig. 8.3.3), then the force applied is given by,

$$F = kx$$

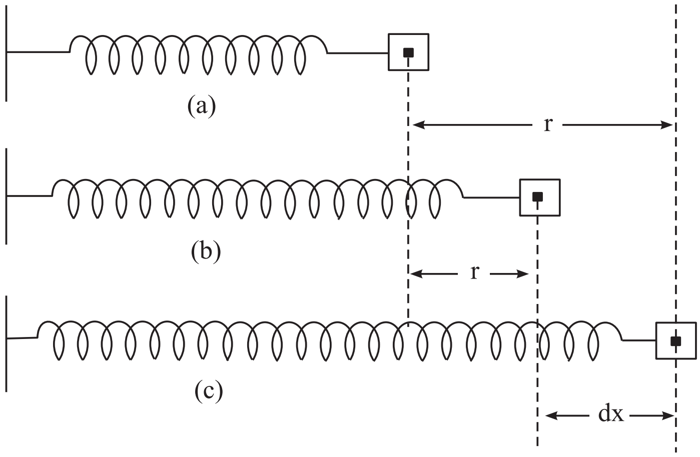


Fig. 8.13

If the spring is further stretched by a small distance dx as shown in the Fig. 8.13 then the small work done

$$dW = kx.dx$$

Therefore, the total work done in stretching the spring through a total distance r from its equilibrium position (Fig. 8.3.3) is given by

$$W = \int_0^r kxdx = k\left[\frac{x^2}{2}\right]_0^r = \frac{1}{2}kr^2$$

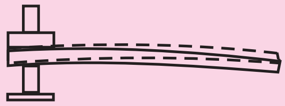
Hence the elastic potential energy $U = \frac{1}{2}kr^2$.

INTEXT QUESTIONS 8.2

1. Is the unit of longitudinal stress same as that of Young's modulus of elasticity? Give reason for your answer.
2. Solids are more elastic than liquids and gases. Justify
3. The length of a wire is cut to half. What will be the effect on the increase in its length under a given load?

4. Two wires are made of the same metal. The length of the first wire is half that of the second and its diameter is double that of the second wire. If equal loads are applied on both wires, find the ratio of increase in their lengths?
5. A wire increases by 10^{-3} of its length when a stress of 1×10^{8} Nm^{-2} is applied to it. Calculate Young's modulus of material of the wire.
6. Calculate the elastic potential energy stored in a spring of spring constant 200 Nm^{-1} when it is stretched through a distance of 10 cm.

Applications of Elastic Behaviour of Materials



(i)

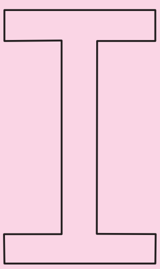
Elastic behaviour of materials plays an important role in our day to day life. Pillars and beams are important parts of our structures. A uniform beam clamped at one end and loaded at the other is called a Cantilever [Fig.(i)]. The hanging bridge of Laxman Jhula in Rishkesh and Vidyasagar Sethu in Kolkata are supported on cantilevers.

A cantilever of length l, breadth b and thickness d undergoes a depression δ at its free end when it is loaded by a weight of mass M :

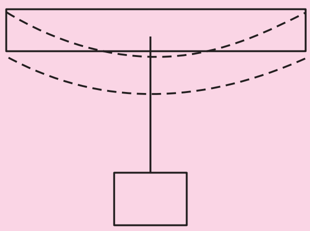
$$\delta = \frac{4M\ g\,\ell^3}{\gamma\, b\, d^3}$$

It is now easy to understand as to why the cross-section of girders and rails is kept I-shaped (Fig. ii). Other factors remaining same, $\delta \propto d^{-3}$. Therefore, by increasing thickness, we can decrease depression under the same load more effectively. This considerably saves the material without sacrificing strength for a beam clamped at both ends and loaded in the middle (Fig.iii), the sag in the middle is given by

$$\delta = \frac{M\ g\,\ell^3}{4\, b\, d^3 \gamma}$$

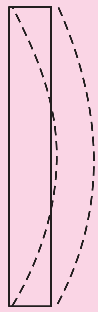


(ii)



(iii

Thus for a given load, we will select a material with a large Young's modulus Y and again a large thickness to keep δ small. However, a deep beam may have a tendency to buckle (Fig iv). To avoid this, a large load bearing surface is provided. In the form I-shaped cross-section, both these requirements are fulfilled.


(iv)

In cranes, we use a thick metal rope to lift and move heavy loads from one place to another. To lift a load of 10 metric tons with a steel rope of yield strength 300 mega pascal, it can be shown, that the minimum area of cross section required will be 10 cm or so. A single wire of this radius will practically be a rigid rod. That is why ropes are always made of a large number of turns of thin wires braided together. This provides ease in manufacturing, flexibility and strength.

Do you know that the maximum height of a mountain on earth can be ~ 10 km or else the rocks under it will shear under its load.

WHAT YOU HAVE LEARNT

- A force which causes deformation in a body is called deforming force.
- On deformation, internal restoring force is produced in a body and enables it to regain its original shape and size after removal of deforming force.
- The property of matter to restore its original shape and size after withdrawal of deforming force is called elasticity.
- The body which gains completely its original state on the removal of the deforming forces is called perfectly elastic.
- If a body completely retains its modified form after withdrawal of deforming force, it is said to be perfectly plastic.
- The stress equals the internal restoring force per unit area. Its units is Nm^{-2}
- The strain equals the change in dimension (e.g. length, volum or shape) per unit dimension. Strain has no unit.
- In normal state, the net inter-atomic force on an atom is zero. If the separation between the atoms becomes more than the separation in normal state, the interatomic forces become attractive. However, for smaller separation, these forces become repulsive.
- The maximum value of stress up to which a body shows elastic property is called its elastic limit. A body beyond the elastic limit behaves like a plastic body.
- Hooke's law states that within elastic limit, stress developed in a body is directly proportional to strain.

- Young's modulus is the ratio of longitudinal stress to longitudinal strain.
- Bulk modulus is the ratio of normal stress to volume strain.
- Modulus of rigidity is the ratio of the shearing stress to shearing strain.
- Poisson's ratio is the ratio of lateral strain to longitudinal strain.
- The work done in stretching a spring is stored as elastic potential energy of the spring.

TERMINAL EXERCISE

1. Define the term elasticity. Give examples of elastic and plastic objects.
2. Explain the terms stress, strain and Hooke's Law.
3. Explain elastic properties of matter on the basis of inter-molecular forces.
4. Define Young's modulus, Bulk modulus and modulus of rigidity.
5. Discuss the behaviour of a metallic wire under increasing load with the help of stress-strain graph.
6. Why steel is more elastic than rubber.
7. Why poission's ratio has no units.
8. In the three states of matter i.e., solid, liquid and gas, which is more elastic and why?
9. A metallic wire 4m in length and 1mm in diameter is stretched by putting a mass 4kg. Determine the alongation produced. Given that the Young's modulus of elasticity for the material of the wire is 13.78×10^{10} N m^{-2}.
10. A sphere contracts in volume by 0.02% when taken to the bottom of sea 1km deep. Calculate the bulk modulus of the material of the sphere. You make take density of sea water as 1000 kgm^{-3} and $g = 9.8$ms^{-2}.
11. How much force is required to have an increase of 0.2% in the length of a metallic wire of radius 0.2mm. Given $Y = 9 \times 10^{10}$ N m^{-2}.
12. What are shearing stress, shearing strain and modulus of rigidity?
13. The upper face of the cube of side 10cm is displaced 2mm parallel to itself when a tangential force of 5×10^5 N is applied on it, keeping lower face fixed. Find out the strain?
14. Property of elasticity is of vital importance in our lives. How does the plasticity helps us?
15. A wire of length L and radius r is clamped rigidly at one end. When the other end of wire is pulled by a force F, its length increases by x. Another wire of the same material of length $2L$ and radius $2r$, when pulled by a force $2F$, what will be the increase in its length.

ANSWERS TO INTEXT QUESTIONS

8.1

1. If $R > R_0$, the nature of force is attractive and if (ii) $R < R_0$ it is repulsive.
2. Longitudinal stress and linear strain.
3. The ratio will decrease.
4. The stress corresponding to breaking point is known as breaking stress.
5. 0.12×10^{10} N m^{-2}.

8.2

1. Both have same units since strain has no unit?
2. As compressibility of liquids and gases is more than solids, the bulk modulus is reciprocal of compressibility. Therefore solids are more elastic than liquid and gases.
3. Half.
4. 1 : 8
5. 1×10^{11} N m^{-2}.
6. 1 J

Answers To Terminal Problems

9. 0.15 m.

10. 4.9×10^{-10} N m^{-2}

11. 22.7 N

13. 2×10^{-2}

15. x.

312en09

9

PROPERTIES OF FLUIDS

In the previous lesson, you have learnt that interatomic forces in solids are responsible for determining the elastic properties of solids.Does the same hold for liquids and gases? (These are collectively called fluids because of their nature to flow in suitable conditions). Have you ever visited the site of a dam on a river in your area / state/ region? If so, you would have noticed that as we go deeper, the thickness of the walls increases. Did you think of the underlying physical principle? Similarly, can you believe that you can lift a car, truck or an elephant by your own body weight standing on one platform of a hydraulic lift? Have you seen a car on the platform of a hydraulic jack at a service centre? How easily is it lifted? You might have also seen that mosquitoes can sit or walk on still water, but we cannot do so. You can explain all these observations on the basis of properties of liquids like hydrostatic pressure, Pascal's law and surface tension. You will learn about these in this lesson.

Have you experienced that you can walk faster on land than under water? If you pour water and honey in separate funnels you will observe that water comes out more easily than honey. In this lesson we will learn the properties of liquids which cause this difference in their flow.

You may have experienced that when the opening of soft plastic or rubber water pipe is pressed, the stream of water falls at larger distance. Do you know how a cricketer swings the ball? How does an aeroplane take off? These interesting observations can be explained on the basis of Bernoulli's principle. You will learn about it in this lesson.

OBJECTIVES

After studying this lesson, you would be able to :

- *calculate the hydrostatic pressure at a certain depth inside a liquid;*

- *describe buoyancy and Archimedes Principle;*
- *state Pascal's law and explain the functioning of hydrostatic press , hydraulic lift and hydraulic brakes.;*
- *explain surface tension and surface energy ;*
- *derive an expression for the rise of water in a capillary tube;*
- *differentiate between streamline and turbulent motion of fluids;*
- *define critical velocity of flow of a liquid and calculate Reynold's number;*
- *define viscosity and explain some daily life phenomena based on viscosity of a liquid; and*
- *state Bernoulli's Principle and apply it to some daily life experiences.*

9.1 HYDROSTATIC PRESSURE

While pinning papers, you must have experienced that it is easier to work with a sharp tipped pin than a flatter one. If area is large, you will have to apply greater force. Thus we can say that for the same force, the effect is greater for smaller area. This effect of force on unit area is called *pressure.*

Refer to Fig. 9.1. It shows the shape of the side wall of a dam. Note that it is thicker at the base. Do we use similar shape for the walls of our house. No, the walls of rooms are of uniform thickness. Do you know the basic physical characteristic which makes us to introduce this change?

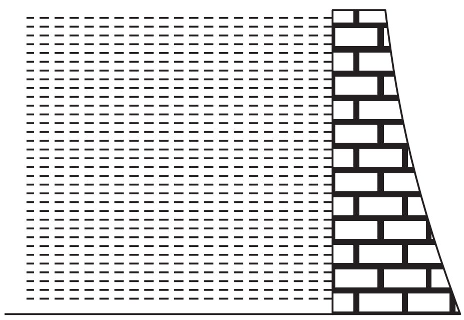


Fig. 9.1 : The structure of side wall of a dam

From the previous lesson you may recall that solids develop shearing stress when deformed by an external force, because the magnitude of inter-atomic forces is very large. But fluids do not have shearing stress and when an object is submerged in a fluid, the force due to the fluid acts normal to the surface of the object (Fig. 9.2). Also, the fluid exerts a force on the container normal to its walls at all points.

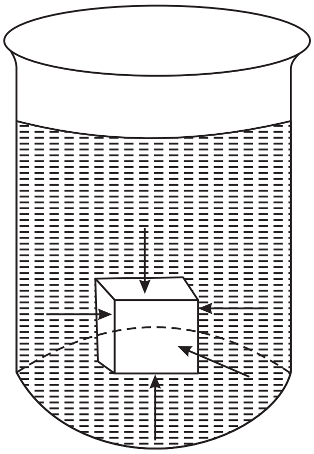


Fig. 9.2 : Force exerted by a fluid on a submerged object

The normal force or thrust per unit area exerted by a fluid is called pressure. We denote it by P :

$$P = \frac{\text{Thrust}}{\text{area}} \quad (9.1)$$

The pressure exerted by a fluid at rest is known as hydrostatic pressure

The SI Unit of pressure is Nm^{-2} and is also called pascal (Pa) in the honour of French scientist Blaise Pascal.

9.1.1 Hydrostatic Pressure at a point in side a liquid

Consider a liquid in a container and an imaginary right circular cylinder of cross sectional area *A* and height *h,* as shown in Fig. 9.3. Let the pressure exerted by the liquid on the bottom and top faces of the cylinder be P_1, and P_2, respectively. Therefore, the upward force exerted by the liquid on the bottom of the cylinder is P_1A and the downward force on the top of the cylinder is P_2A.

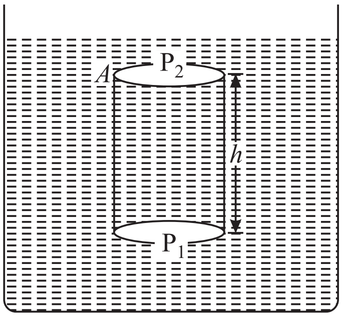


Fig. 9.3 : An imaginary cylinder of height *h* in a liquid.

∴ The net force in upward direction is $(P_1A - P_2A)$.

Now mass of the liquid in cylinder = density × volume of the cylinder

$= \rho . A . h$ where ρ is the density of the liquid.

∴ Weight of the liquid in the cylinder = $\rho . g . h . A$

Since the cylinder is in equilibrium, the resultant force acting or it must be equal to zero, i.e.

$$P_1 A - P_2 A - \rho g h A = 0$$

$$\Rightarrow \quad P_1 - P_2 = \rho g h \qquad (9.2)$$

So, the pressure P at the bottom of a column of liquid of height h is given by

$$P = \rho g h$$

That is, hydrostatic pressure due to a fluid increases linearly with depth. It is for this reason that the thickness of the wall of a dam has to be increased with increase in the depth of the dam.

If we consider the upper face of the cylinder to be at the open surface of the liquid, as shown in Fig.(9.4), then P_2 will have to be replaced by P_{atm} (Atmospheric pressure). If we denote P_1 by P, the absolute pressure at a depth below the surface will be

$$P - P_{atm} + \rho g h$$

or

$$P = P_{atm} + \rho g h \qquad (9.3)$$

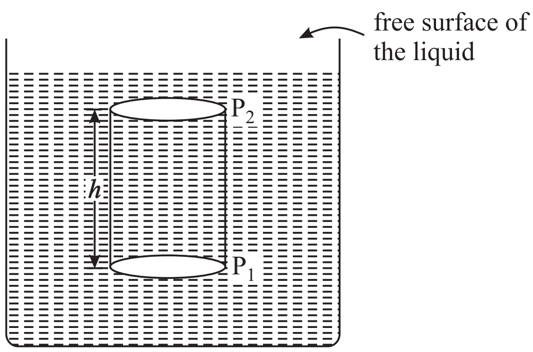


Fig. 9.4 : Cylinder in a liquid with one face at the surface of the liquid

Note that the expression given in Eqn. (9.3) does not show any term having area of the cylinder It means that pressure in a liquid at a given depth is equal, irrespective of the shape of the vessel (Fig 9.5).

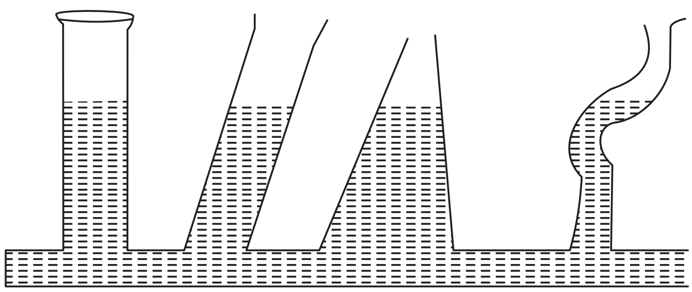


Fig. 9.5 : Pressure does not depend upon shape of the versel.

Example 9.1: A cemented wall of thickness one metre can withstand a side pressure of 10^5 Nm^{-2}. What should be the thickness of the side wall at the bottom of a water dam of depth 100 m. Take density of water = 10^3 kg m^{-3} and g = 9.8 ms^{-2}.

Solution: The pressure on the side wall of the dam at its bottom is given by

$$P = h\,d\,g$$

$$= 100 \times 10^3 \times 9.8$$

$$= 9.8 \times 10^5 \text{ Nm}^{-2}$$

Using unitary method, we can calculate the thickness of the wall, which will withstand pressure of 9.8×10^5 Nm^{-2}. Therefore thickness of the wall

$$t = \frac{9.8\times10^5}{10^5}\;\frac{\text{Nm}^{-2}}{\text{Nm}^{-2}}$$

$$= 9.8 \text{ m}$$

9.1.2 Atmospheric Pressure

We know that the earth is surrounded by an atmosphere upto a height of about 200 km. The pressure exerted by the atmosphere is known as the *atmospheric pressure*. A German Scientist O.V. Guericke performed an experiment to demonstrate the force exerted on bodies due to the atmospheric pressure. He took two hollow hemispheres made of copper, having diameter 20 inches and tightly joined them with each other. These could easily be separated when air was inside. When air between them was exhausted with an air pump, 8 horses were required to pull the hemispheres apart.

Toricelli used the formula for hydrostatic pressure to determine the magnitude of atmospheric pressure.

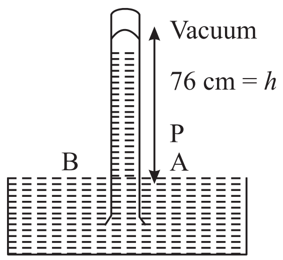


Fig: 9.6 : Toricelli's Barometer

He took a tube of about 1 m long filled with mercury of density 13,600 kg m^{-3} and placed it vertically inverted in a mercury tub as shown is Fig. 9.6. He observed that the column of 76 cm of mercury above the free surface remained filled in the tube.

In equilibrium, atmospheric pressure equals the pressure exerted by the mercury column. Therefore,

$$P_{atm} = h\,\rho\,g = 0.76 \times 13600 \times 9.8 \text{ Nm}^{-2}$$

$$= 1.01 \times 10^5 \text{ Nm}^{-2}$$

$$= 1.01 \times 10^5 \text{ Pa}$$

9.2 BUOYANCY

It is a common experience that lifting an object in water is easier than lifting it in air. It is because of the difference in the upward forces exerted by these fluids on these object. The upward force, which acts on an object when submerged in a fluid, is known as **buoyant force**. The nature of buoyant force that acts on objects placed inside a fluid was discovered by. Archimedes Based on his observations, he enunciated a law now known as Archimedes principle. It state that *when an object is submerged partially or fully in a fluid, the magnitude of the buoyant force on it is always equal to the weight of the fluid displaced by the object.*

The different conditions of an object under buoyant force is shown in Fig 9.7.

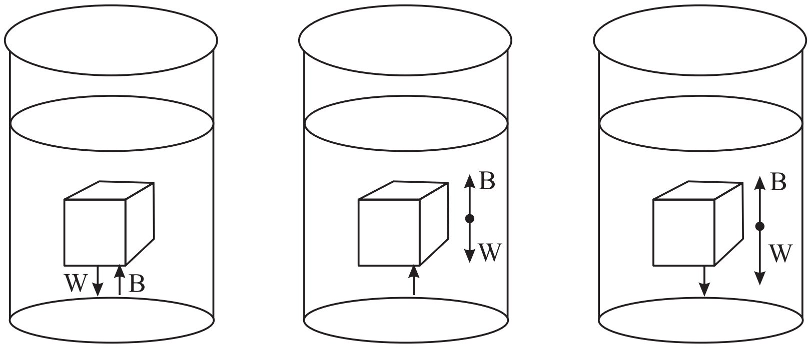


Fig. 9.7:

(a) : The magnitude of buoyant force B on the object is exactly equal to its weight in equilibrium.

(b) : A totally submerged object of density less than that of the fluid experiences a net upward force.

(c) : A totally submerged object denser than the fluid sinks.

Another example of buoyant force is provided by the motion of hot air balloon shown in Fig. 9.8. Since hot air has less density than cold air, a net upward buoyant force on the balloon makes it to float.

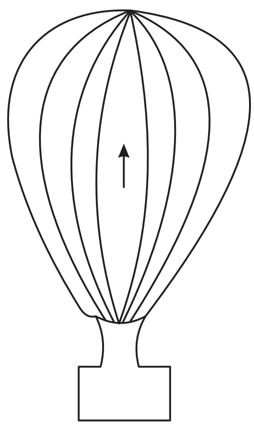
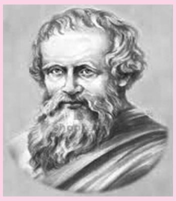


Fig. 9.8: Hot air balloon floating in air

Floating objects

You must have observed a piece of wood floating on the surface of water. Can you identify the forces acting on it when it is in equilibrium? Obviously, one of the forces is due to gravitational force, which pulls it downwards. However, the displaced water exerts buoyant force which acts upwards. These forces balance each other in equilibrium state and the object is then said to be floating on water. It means that a floating body displaces the fluid equal to its own weight.

Archimedes
(287- 212 B.C)



A Greek physicist, engineer and mathematician was perhaps the greatest scientist of his time. He is well known for discovering the nature of buoyant forces acting on objects. The Archimedes screw is used even today. It is an inclined rotating coiled tube used originally to lift water from the hold of ships. He also invented the catapult and devised the system of levers and pulleys.

Once Archimedes was asked by king Hieron of his native city Syracuse to determine whether his crown was made up of pure gold or alloyed with other metals without damaging the crown. While taking bath, he got a solution, noting a partial loss of weight when submerging his arm and legs in water. He was so excited about his discovery that he ran undressed through the streets of city shouting " Eureka, Eureka", meaning I have found it.

9.3 PASCAL'S LAW

While travelling by a bus, you must have observed that the driver stops the bus by applying a little force on the brakes by his foot. Have you seen the hydraulic jack or lift which can lift a car or truck up to a desired height? For this purpose you may visit a motor workshop. Packing of cotton bales is also done with the help of hydraulic press which works on the same principle.

These devices are based on Pascal's law, which states that *when pressure is applied at any part of an enclosed liquid, it is transmitted undiminished to every point of the liquid as well as to the walls of the container.*

This law is also known as the **law of transmission of liquid pressure.**

9.3.1 Applications of Pascal's Law

(A) Hydraulic Press/Balance/Jack/Lift

It is a simple device based on Pascal's law and is used to lift heavy loads by applying a small force. The basic arrangement is shown in Fig.9.9. Let a force F_1 be applied to the smaller piston of area A_1. On the other side, the piston of large area A_2 is attached to a platform where heavy load may be placed. The pressure on the smaller piston is transmitted to the larger piston through the liquid filled in-between the two pistons. Since the pressure is same on both the sides, we have

Pressure on the smaller piston, $P = \dfrac{\text{force}}{\text{area}} = \dfrac{F_1}{A_1}$

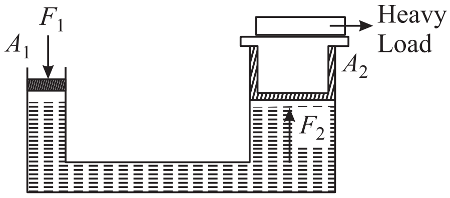


Fig. 9.9: Hydraulic lift

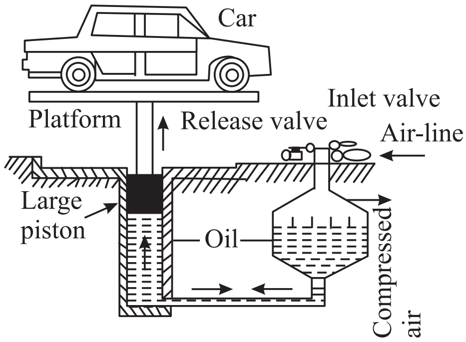


Fig. 9.10: Hydraulic jack

According to Pascal's law, the same pressure is transmitted to the larger cylinder of area A_2.

Hence the force acting on the larger piston

$$F_2 = \text{pressure} \times \text{area} = \frac{F_1}{A_1} \times A_2 \qquad (9.4)$$

It is clear from Eqn. (9.4) that force $F_2 > F_1$ by an amount equal to the ratio (A_2/A_1). With slight modifications, the same arrangement is used in hydraulic press, hydraulic balance, and hydraulic Jack, etc.

(B) Hydraulic Jack or Car Lifts

At automobile service stations, you would see that cars, buses and trucks are raised to the desired heights so that a mechanic can work under them (Fig 9.10). This is done by applying pressure, which is transmited through a liquid to a large surface to produce sufficient force needed to lift the car.

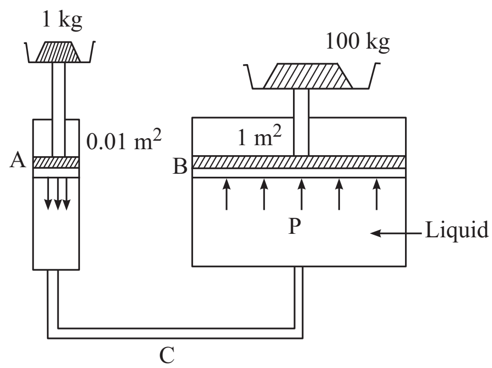


Fig. 9.11(a) : Hydraulic balance

(C) Hydraulic Brakes

While traveling in a bus or a car, we see how a driver applies a little force by his foot on the brake paddle to stop the vehicle. The pressure so applied gets transmitted through the brake oil to the piston of slave cylinders, which, in turn, pushes the break shoes against the break drum in all four wheels, simultaneously. The wheels stop rotating at the same time and the vehicle comes to stop instantaneously.

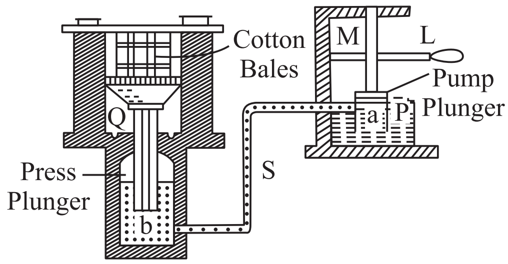


Fig. 9.11(b) : Hydraulic press

INTEXT QUESTIONS 9.1

1. Why are the shoes used for skiing on snow made big in size?
2. Calculate the pressure at the bottom of an ocean at a depth of 1500 m. Take the density of sea water 1.024×10^3 kg m^{-3}, atmospheric pressure=1.01×10^5 Pa and $g = 9.80$ ms^{-2}.
3. An elephant of weight 5000 kg f is standing on the bigger piston of area 10 m^2 of a hydraulic lift. Can a boy of 25 kg wt standing on the smaller piston of area 0.05m^2 balance or lift the elephant?
4. If a pointed needle is pressed against your skin, you are hurt but if the same force is applied by a rod on your skin nothing may happen. Why?
5. A body of 50 kg f is put on the smaller piston of area 0.1m^2 of a big hydraulic lift. Calculate the maximum weight that can be balanced on the bigger piston of area 10m^2 of this hydraulic lift.

9.4 SURFACE TENSION

It is common experience that in the absence of external forces, drops of liquid are always spherical in shape. If you drop small amount of mercury from a small height, it spreads in small spherical globules. The water drops falling from a tap or shower are also spherical. Do you know why it is so? You may have enjoyed the soap bubble game in your childhood. But you can not make pure water bubbles with same case? All the above experiences are due to a characteristic property of liquids, which we call **surface tension**. To appreciate this, we would like you to do the following activity.

ACTIVITY 9.1

1. Prepare a soap solution.
2. Add a small amount of glycerin to it.
3. Take a narrow hard plastic or glass tube. Dip its one end in the soap solution so that some solution enters into it.
4. Take it out and blow air at the other end with your mouth.
5. Large soap bubble will be formed.
6. Give a jerk to the tube to detach the bubble which then floats in the air.

To understand as to how surface tension arises, let us refresh our knowledge of intermolecular forces. In the previous lesson, you have studied the variation of intermolecular forces with distance between the centres of molecules/atoms.

The intermolecular forces are of two types: **cohesive** and **adhesive**. Cohesive forces characterise attraction between the molecules of the same substance, whereas force of adhesion is the attractive force between the molecules of two different substances. It is the force of adhesion which makes it possible for us to write on this paper. Gum, Fevicol etc. show strong adhesion.

We hope that now you can explain why water wets glass while mercury does not.

ACTIVITY 9.2

To show adhesive forces between glass and water molecule.

1. Take a clean sheet of glass
2. Put a few drops of water on it
3. Hold water containing side downward.
4. Observe the water drops.

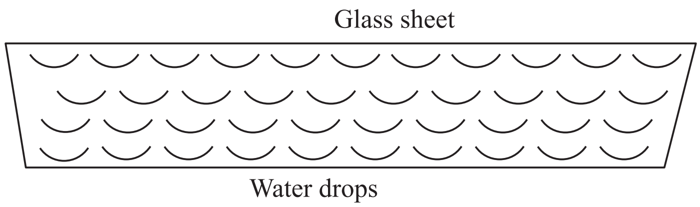


Fig. 9.12: Water drops remain stuck to the glass sheet

The Adhesive forces between glass and water molecules keep the water drops sticking on the glass sheet, as shown in Fig. 9.12.

9.4.1 Surface Energy

The surface layer of a liquid in a container exhibits a property different from the rest of the liquid. In Fig. 9.13, molecules are shown at different heights in a liquid. A molecule, say P, well inside the liquid is attracted by other molecules from all sides. However, it is not the case for the molecules at the surface.

Molecules S and R, which lie on the surface layer, experience a net resultant force downward because the number of molecules

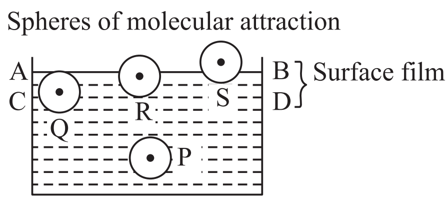


Fig 9.13 : Resultant force acting on P and Q is zero but molecules R and S experience a net vertically downward force.

in the upper half of sphere of influence attracting these molecules is less than those in the lower half. If we consider the molecules of liquid on the upper half of the surface of the liquid or liquid-air interface, even then the molecules will experience a net downward force because of less number of molecules of liquid. Therefore, if any liquid molecule is brought to the surface layer, work has to be done against the net inward force, which increases their potential energy. This means that surface layer possesses an additional energy, which is termed as *surface energy*.

For a system to be in equilibrium, its potential energy must be minimum. Therefore, the area of surface must be minimum. That is why free surface of a liquid at rest tends to attain minimum surface area. This produces a tension in the surface, called **surface tension.**

Surface tension is a property of the liquid surface due to which it has the tendency to decrease its surface area. As a result**,** the surface of a liquid acts like a stretched membrane You can visualise its existence easily by placing a needle gently on water surface and see it float.

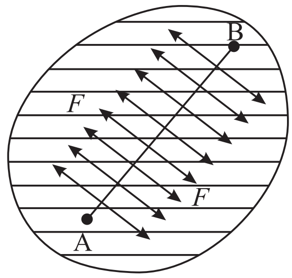


Fig. 9.14 : Direction of surface tension on a liquid surface

Let us now understand this physically. Consider an imaginary line AB drawn at the surface of a liquid at rest, as shown in Fig 9.14. The surface on either side of this line exerts a pulling force on the surface on the other side.

The **surface tension of a liquid can be defined as the force per unit length in the plane of liquid surface** :

$$T = F/L \qquad (9.5)$$

where surface tension is denoted by T and F is the magnitude of total force acting in a direction normal to the imaginary line of length L, (Fig 9.14) and tangential to the liquid surface. SI unit of surface tension is Nm^{-1} and its dimensions are $[MT^{-2}]$.

Let us take a rectangular frame, as shown in Fig. 9.15 having a sliding wire on one of its arms. Dip the frame in a soap solution and take out. A soap film will be formed on the frame and have two surfaces. Both the surfaces are in contact with the sliding wire, So we can say that surface tension acts on the wire due to both these surfaces.

Let T be the surface tension of the soap solution and L be the length of the wire.

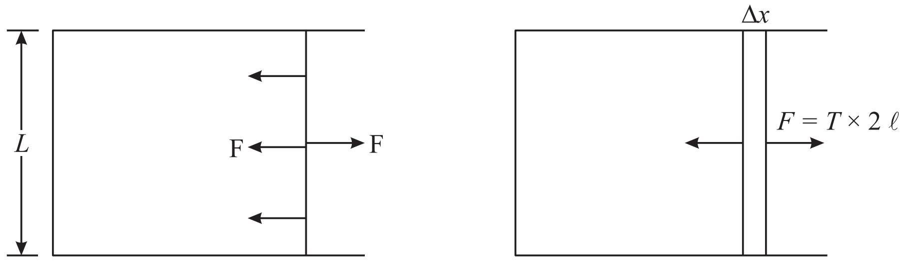


Fig. 9.15: A Film in equilibirum

The force exerted by each surface on the wire will be equal to $T \times L$. Therefore, the total force F on the wire = $2TL$.

Suppose that the surfaces tend to contract say, by Δx. To keep the wire in equilibrium we will have to apply an external uniform force equal to F. If we increase the surface area of the film by pulling the wire with a constant speed through a distance Δx, as shown in Fig. 9.15b, the work done on the film is given by

$$W = F \times \Delta x = T \times 2L \times \Delta x$$

where $2L \times \Delta x$ is the total increase in the area of both the surfaces of the film. Let us denote it by A. Then, the expressopm for work done on the film simplifies to

$$W = T \times A$$

This work done by the external force is stored as the potential energy of the new surface and is called as surface energy. By rearranging terms, we get the required expresion for surface tension :

$$T = W/A \tag{9.6}$$

Thus, we see that **surface tension of a liquid is equal to the work done in increasing the surface area of its free surface by one unit**. We can also say that **surface tension is equal to the surface energy per unit area**.

We may now conclude that surface tension

- is a property of the surface layer of the liquid or the interface between a liquid and any other substance like air;
- tends to reduce the surface area of the free surface of the liquid;
- acts perpendicular to any line at the free surface of the liquid and is tangential to its meniscus;
- has genesis in intermolecular forces, which depend on temperature; and
- decreases with temperature.

A simple experiment described below demonstrates the property of surface tension of liquid surfaces.

ACTIVITY 9.3

Take a thin circular frame of wire and dip it in a soap solution. You will find that a soap film is formed on it. Now take a small circular loop of cotton thread and put it gently on the soap film. The loop stays on the film in an irregular shape as shown in Fig. 9.16(a). Now take a needle and touch its tip to the soap film inside the loop. What do you observe?

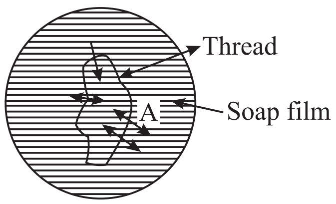


Fig 9.16 (a) : A soap film with closed loop of thread

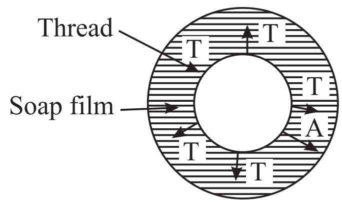


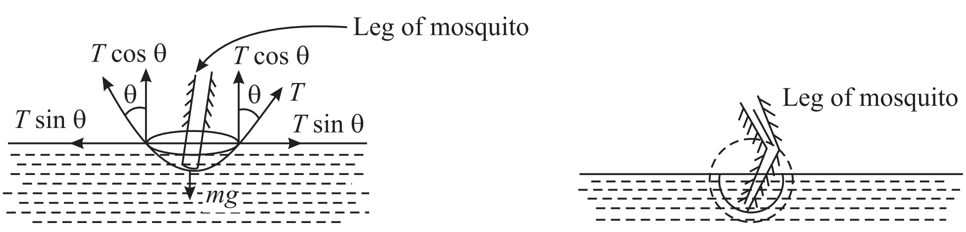
Fig. 9.16 (b) : The shape of the thread without inner soap film

You will find that the loop of cotton thread takes a circular shape as shown in Fig 9.16(b). Initially there was soap film on both sides of the thread. The surface on both sides pulled it and net forces of surface tension were zero. When inner side was punctured by the needle, the outside surface pulled the thread to bring it into the circular shape, so that it may acquire minimum area.

9.4.2 Applications of Surface Tension

(a) Mosquitoes sitting on water

In rainy reason, we witness spread of diseases like dengue, malaria and chickungunya by mosquito breeding on fresh stagnant water. Have you seen mosquitoes sitting on water surface? They do not sink in water due to surface tension. At the points where the legs of the mosquito touch the liquid surface, the surface becomes concave due to the weight of the mosquito. The surface tension



(a) (b)

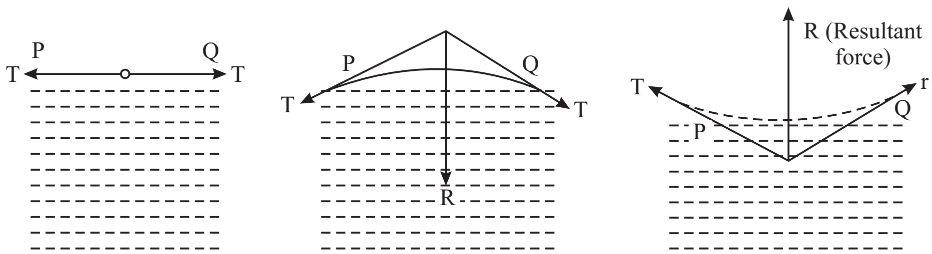
Fig. 9.17 : The weight of a mosquito is balanced by the force of surface tension $= 2\pi rT \cos\theta$ (a) Dip in the level to form concave surface, and (b) magnified image

acting tangentially on the free surface, therefore, acts at a certain angle to the horizontal. Its vertical component acts upwards. The total force acting vertically upwards all along the line of contact of certain length balances the weight of the mosquito acting vertically downward, as shown in Fig 9.17.

(b) Excess pressure on concave side of a spherical surface

Consider a small surface element with a line PQ of unit length on it, as shown in Fig. 9.18. If the surface is plane, i.e. $\theta = 90^0$, the surface tension on the two sides tangential to the surface balances and the resultant tangential force is zero [Fig. 9.18 (a)]. If, however, the surface is convex, [Fig. (9.18 (b)] or concave [Fig. 9.18 (c)], the forces due to surface tension acting across the sides of the line PQ will have resultant force **R** towards the center of curvature of the surface.

Thus, whenever the surface is curved, the surface tension gives rise to a pressure directed towards the center of curvature of the surface. This pressure is balanced by an equal and opposite pressure acting on the surface. Therefore, there is always an excess pressure on the concave side of the curved liquid surface [Fig. (9.18 b)].



(a) plane surface (b) convex surface (c) concave surface

Fig. 9.18

(i) Spherical drop

A liquid drop has only one surface i.e. the outer surface. (The liquid area in contact with air is called the surface of the liquid.) Let r be the radius of a small spherical liquid drop and P be excess pressure inside the drop (which is concave on the inner side, but convex on the outside). Then

$$P = (P_i - P_0)$$

where P_i and P_0 are the inside and outside pressures of the drop, respectively (Fig 9.19a)

If the radius of the drop increases by Δr due to this constant excess pressure P, then increase in surface area of the spherical drop is given by

$$\Delta A = 4\pi (r + \Delta r)^2 - 4\pi r^2$$

$$= 8\pi \, r \, \Delta r$$

where we have neglected the term containing second power of Δr.

The work done on the drop for this increase in area is given by

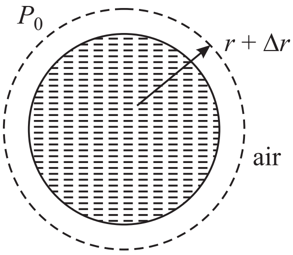


Fig. 9.19 (a) : A spherical drop

$$\mathbf{W} = \text{Extra surface energy}$$

$$= T\Delta A = T.\, 8\pi \, r \, \Delta r \qquad (9.7)$$

If the drop is in equilibrium, this extra surface energy is equal to the work done due to expansion under the pressure difference or excess pressure P:

$$\text{Work done} = P \, \Delta V = P.\, 4\pi \, r^2 \Delta r \qquad (9.8)$$

On combining Eqns. (9.7) and (9.8), we get

$$P.\, 4\pi \, r^2 \Delta r = T.8 \, \pi \, r \, \Delta r$$

Or $$P = 2\, T/\, r \qquad (9.9)$$

(ii) Air Bubble in water

An air bubble also has a single surface, which is the inner surface (Fig. 9.19b). Hence, the excess of pressure P inside an air bubble of radius r in a liquid of surface tension T is given by

$$P = 2T/\, r \qquad (9.10)$$

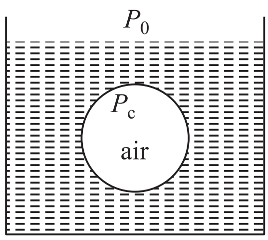


Fig. 9.19 b : Air Bubble

(iii) Soap bubble floating in air

The soap bubble has two surfaces of equal surface area (i.e. the outer and inner), as shown in Fig. 9.19(c). Hence, excess pressure inside a soap bubble floating in air is given by

$$P = 4T/\, r \qquad (9.11)$$

where T is suface tension of soap solution.

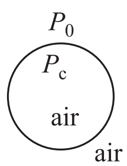


Fig. 9.19 (c)

This is twice that inside a spherical drop of same radius or an air bubble in water. Now you can understand why a little extra pressure is needed to form a soap bubble.

Example 9.3: Calculate the difference of pressure between inside and outside of a (i) spherical soap bubble in air, (ii) air bubble in water, and (iii) spherical drop of water, each of radius 1 mm. Given surface tension of water = 7.2×10^{-2} Nm^{-1} and surface tension of soap solution = 2.5×10^{-2} Nm^{-1}.

Solution:

(i) Excess pressure inside a soap bubble of radius r is

$$P = 4T/r$$

$$= \frac{4 \times 2.5 \times 10^{-2}}{1 \times 10^{-3}\,\text{m}}\ \text{Nm}^{-1}$$

$$= 100\ \text{Nm}^{-2}$$

(ii) Excess pressure inside an air bubble in water

$$= 2T'/r$$

$$= \frac{2 \times 7.2 \times 10^{-2}\,\text{Nm}^{-1}}{1 \times 10^{-3}\,\text{m}}$$

$$= 144\ \text{Nm}^{-2}$$

(iii) Excess pressure inside a spherical drop of water $= 2T'/r$

$$= 144\ \text{Nm}^{-2}$$

(c) Detergents and surface tension

You may have seen different advertisements highlighting that detergents can remove oil stains from clothes. Water is used as cleaning agent. Soap and detergents lower the surface tension of water. This is desirable for washing and cleaning since high surface tension of pure water does not allow it to penetrate easily between the fibers of materials, where dirt particles or oil molecules are held up.

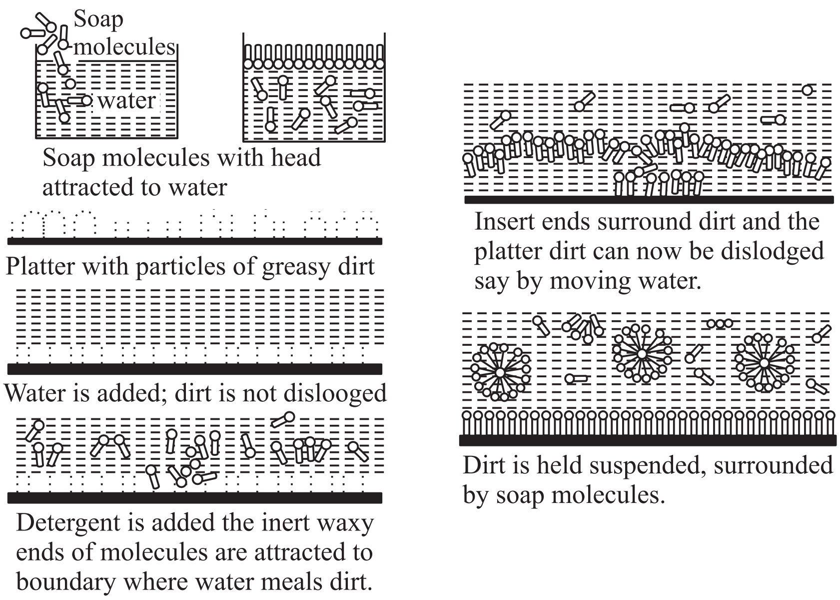


Fig: 9.20 : Detergent action

You now know that surface tension of soap solution is smaller than that of pure water but the surface tension of detergent solutions is smaller than that of soap solution. That is why detergents are more effective than soap. A detergent dissolved in water weakens the hold of dirt particles on the cloth fibers which therefore, get easily detached on squeezing the cloth.

The addition of detergent, whose molecules attract water as well as oil, drastically reduces the surface tension (T) of water-oil. It may even become favourable to form such interfaces, i.e. globes of dirt surrounded by detergent and then by water. This kind of process using surface active detergents is important for not only cleaning the clothes but also in recovering oil, mineral ores etc.

(d) Wax-Duck floating on water

You have learnt that the surface tension of liquids decreases due to dissolved impurities. If you stick a tablet of camphor to the bottom of a wax-duck and float it on still water surface, you will observe that it begins to move randomly after a minute or two. This is because camphor dissolves in water and the surface tension of water just below the duck becomes smaller than the surrounding liquid. This creates a net difference of force of surface tension which makes the duck to move.

Now, it is time for you to check how much you have learnt. Therefore, answer the following questions.

INTEXT QUESTIONS 9.2

1. What is the difference between force of cohesion and force of adhesion?
2. Why do small liquid drops assume a spherical shape.
3. Do solids also show the property of surface tension? Why?
4. Why does mercury collect into globules when poured on plane surface?
5. Which of the following has more excess pressure?
   (i) An air bubble in water of radius 2 cm. Surface tension of water is 727×10^{-3} Nm^{-1} or
   (ii) A soap bubble in air of radius 4 cm. Surface tension of soap solution is 25×10^{-3} Nm^{-1}.

9.5 ANGLE OF CONTACT

You can observe that the free surface of a liquid kept in a container is curved. For example, when water is filled in a glass jar, it becomes concave but if we fill water in a paraffin wax container, the surface of water becomes convex. Similarly, when mercury is filled in a glass jar, its surface become convex. Thus, we see that shape of the liquid surface in a container depends on the nature of the liquid, material of container and the medium above free surface of the liquid. To characterize it, we introduce the concept of angle of contact.

It is the angle **that the tangential plane to the liquid surface makes with the tangential plane to the wall of the container, to the point of contact, as measured from within the liquid, is known as angle of contact.**

Fig. 9.21 shows the angles of contact for water in a glass jar and paraffin jar. The angle of contact is acute for concave spherical meniscus, e.g. water with glass and obtuse (or greater than 90^0) for convex spherical meniscus e.g. water in paraffin or mercury in glass tube.

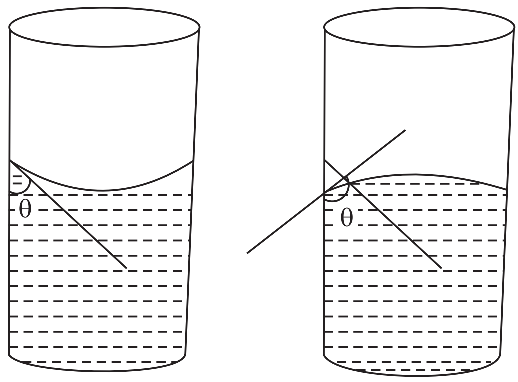


Fig 9.21 : Nature of free surface when water is filled in (a) glass jar, and (b) paraffin wax jar

Various forces act on a molecule in the surface of a liquid contained in a vessel near the boundary of the menisus. As the liquid is present only in the lower quadrant, the resultant cohesive force acts on the molecule at P symmetrically, as shown in the Fig.9.22(a). Similarly due to symmetry, the resultant adhesive force F_a acts outwards at right angles to the walls of the container vessel. The force F_c can be resolved into two mutually perpendicular components F_c cos

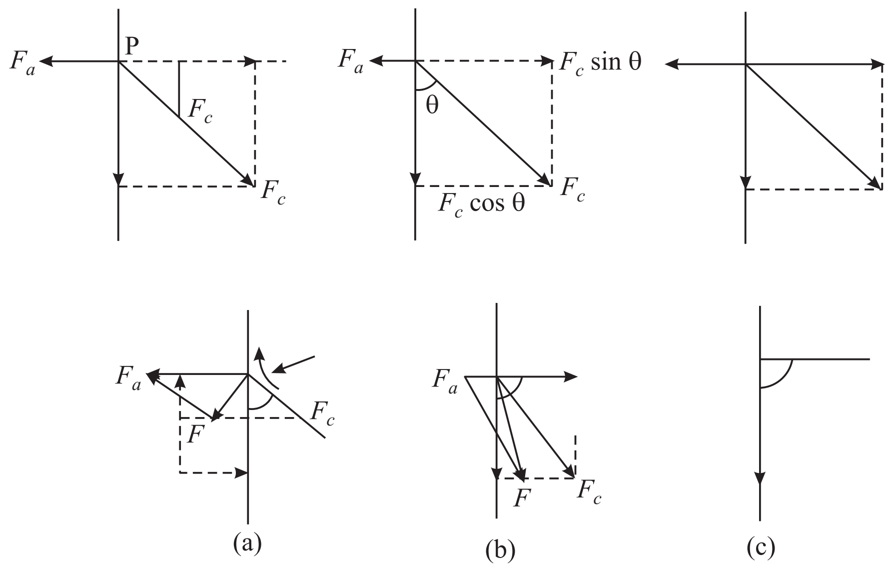


Fig. 9.22 : Different shapes of liquid meniscuses

θ acting vertically downwards and $F_c \sin\theta$ acting at right angled to the boundary, The value of the angle of contact depends upon the relative values of F_c and F_a.

CASE 1: If $F_a > F_c \sin\theta$, the net horizontal force is outward and the resultant of $(F_a - F_c \sin\theta)$ and $F_c \cos\theta$ lies outside the wall. Since liquids can not sustain constant shear, the liquid surface and hence all the molecules in it near the boundary adjust themselves at right angles to F_c so that no component of F acts tangential to the liquid surface. Obviously such a surface at the boundary is concave spherical (Since radius of a circle is perpendicular to the circumference at every point.) This is true in the case of water filled in a glass tube.

Case 2 : If $F_a < F_c \sin\theta$ the resultant F of $(F_c \sin\theta - F_a)$ acting horizontally and $F_c \cos\theta$ acting vertically down wards is in the lower quadrant acting into the liquid. The liquid surface at the boundary, therefore, adjusts itself at right angles to this and hence becomes convex spherical. This is true for the case of mercury filled in the glass tube.

Case 3 : When $F_a = F_c \sin\theta$, the resultant force acts vertically downwards and hence the liquid surface near the boundary becomes horizontal or plane.

9.6 CAPILLARY ACTION

You might have used blotting paper to absorb extra ink from your notebook. The ink rises in the narrow air gaps in the blotting paper. Similarly, if the lower end of a cloth gets wet, water slowly rises upward. Also water given to the fields rises in the innumerable capillaries in the stems of plants and trees and reaches the branches and leaves. Do you know that farmers plough their fields only after rains so that the capillaries formed in the upper layers of the soil are broken. Thus, water trapped in the soil is taken up by the plants. On the other hand, we find that when a capillary tube is dipped into mercury, the level of mercury inside it is below the outside level. Such an important phenomenon of the elevation or depression of a liquid in an open tube of small cross- section (i.e., capillary tube) is basically due to surface tension and is known as capillary action.

The phenomenon of rise or depression of liquids in capillary tubes is known as capillary action or capillarity.

9.6.1 Rise of a Liquid in a Capillary Tube

Let us take a capillary tube dipped in a liquid, say water. The meniscus inside the tube will be concave, as shown in Fig. 9.23 (a). This is essentially because the forces of adhesion between glass and water are greater than cohesive forces.

Let us consider four points A, B, C and D near the liquid-air interface Fig. 9.23(a). We know that pressure just below the meniscus is less than the pressure just above it by $2T/R$, i.e.

$$P_B = P_A - 2T/R \qquad (9.12)$$

where T is surface tension at liquid-air interface and R is the radius of concave surface.

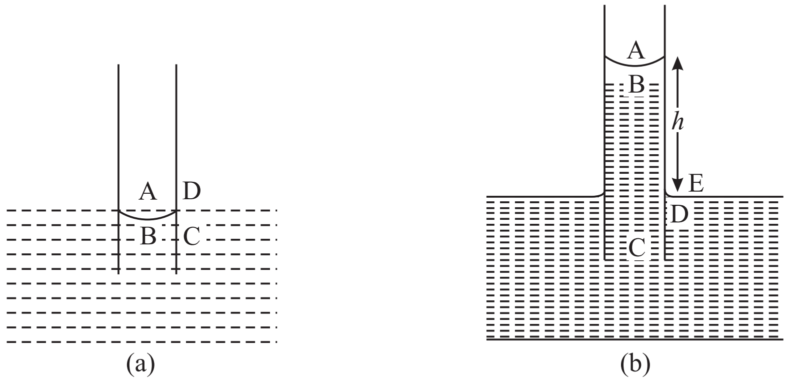


Fig. 9.23 : Capillary action

But pressure at A is equal to the pressure at D and is equal to the atmospheric pressure P (say). And pressure at D is equal to pressure at C . Therefore, pressure at B is less than pressure at D. But we know that the pressure at all points at the same level in a liquid must be same. That's why water begins to flow from the outside region into the tube to make up the deficiency of pressure at point B.

Thus liquid begins to rise in the capillary tube to a certain height h (Fig 9.23 b) till the pressure of liquid column of height h becomes equal to $2T/R$..Thereafter, water stops rising. In this condition

$$h \rho g = 2\, T/R \qquad (9.13)$$

where ρ is the density of the liquid and g is the acceleration due to gravity. If r be radius of capillary tube and θ be the angle of contact, then from Fig. 9.24, we can write

$$R = r / \cos\theta$$

Substituting this value of R in Equation (9.13)

$$h\, p\, g = 2T/\, r / \cos\theta$$

or $$h = 2T \cos\theta \;/\; r\, \rho\, g \qquad (9.14)$$

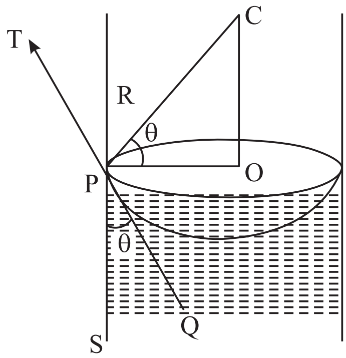


Fig. 9.24 : Angle of contact

It is clear from the above expression that if the radius of tube is less (i.e. in a very fine bore capillary), liquid rise will be high.

INTEXT QUESTIONS 9.3

1. Does the value of angle of contact depend on the surface tension of the liquid?
2. The angle of contact for a solid and liquid is less than the 90^0. Will the liquid wet the solid? If a capillary is made of that solid, will the liquid rise or fall in it?
3. Why it is difficult to enter mercury in a capillary tube, by simply dipping it into a vessel containing mercury while designing a thermometer.
4. Calculate the radius of a capillary to have a rise of 3 cm when dipped in a vessel containing water of surface tension 7.2×10^{-2} N m^{-1}. The density of water is 1000 kg m^{-3}, angle of contact is zero, and g = 10 m s^{-2}.
5. How does kerosene oil rise in the wick of a lantern?

9.7 VISCOSITY

If you stir a liquid taken in a beaker with a glass rod in the middle, you will note that the motion of the liquid near the walls and in the middle is not same (Fig.9.25). Next watch the flow of two liquids (e.g. glycerin and water) through identical pipes. You will find that water flows rapidly out of the vessel whereas glycerine flows slowly. Drop a steel ball through each liquid. The ball falls more slowly in glycerin than in water. These observations indicate a characteristic property of the liquid that determines their motion. This property is known as ***viscosity***. Let us now learn how it arises.

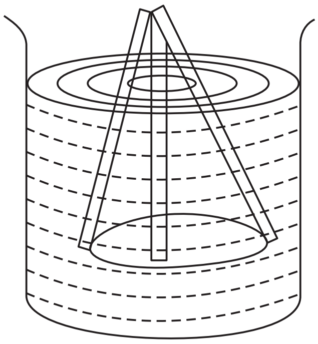


Fig. 9.25: Water being stirred with a glass rod

9.7.1 Viscosity

We know that when one body slides over the other, a frictional force acts between them. Similarly, whenever a fluid flows, two adjacent layers of the fluid exert a tangential force on each other; this force acts as a drag and opposes the relative motion between them. *The property of a fluid by virtue of which it opposes the relative motion in its adjacent layers is known as viscosity.*

Fig. 9.26 shows a liquid flowing through a tube. The layer of the liquid in touch with the wall of the tube can be assumed to be stationary due to friction between the solid wall and the liquid. Other layers are in motion and have different velocities Let υ be the velocity of the layer at a distance x from the surface and $\upsilon + d\upsilon$ be the velocity at a distance $x + dx$.

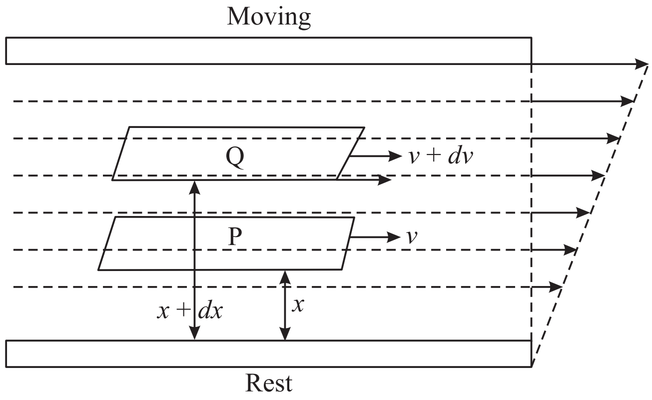


Fig. 9.26 : Flow of a liquid in a tube: Different layers move with different velocities

Thus, the velocity changes by dv in going through a distance dx perpendicular to it. The quantity dv/dx is called the **velocity gradient**.

The viscous force F between two layers of the fluid is proportional to

- area (A) of the layer in contact : $F \propto A$
- velocity gradient (dv/dx) in a direction perpendicular to the flow of liquid : $F \propto dv/dx$

Table 10.1 : Viscosity of a few typical fluids

Name of fluid	T [^{0}C]	Viscosity η (PR)
Water	20	1.0×10^{-3}
Water	100	0.3×10^{-3}
blood	37	2.7×10^{-3}
Air	40	1.9×10^{-5}

On combining these, we can write

$$F \propto A\, dv/dx$$

or

$$F = -\eta A\, (dv/dx) \quad (9.15)$$

where η is constant of proportionality and is called **coefficient of viscosity**. The negative sign indicates that force is frictional in nature and opposes motion.

The SI unit of coefficient of viscosity is Nsm^{-2}. In cgs system, the unit of viscosity is poise.

$$1 \text{ poise} = 0.1\ Nsm^{-2}$$

Dimensions of coefficient of viscosity are $[ML^{-1}T^{-1}]$

9.8 TYPES OF LIQUID FLOW

Have you ever seen a river in floods? Is it similar to the flow of water in a city water supply system? If not, how are the two different? To discover answer to such questions, let as study the flow of liquids.

9.8.1 Streamline Motion

The path followed by fluid particles is called line of flow. If every particle passing through a given point of the path follows the same line of flow as that of preceding particles, the flow is said to be *streamlined.* A streamline can be represented as the curve or path whose tangent at any point gives the direction of the liquid velocity at that point. In steady flow, the streamlines coincide with the line of flow (Fig. 9.27).

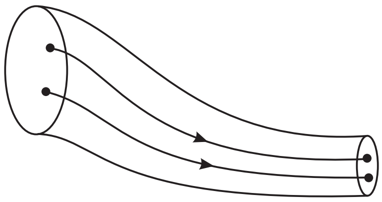


Fig. 9.27: Streamline flow

Note that streamlines do not intersect each other because two tangents can then be drawn at the point of intersection giving two directions of velocities, which is not possible.

When the velocity of flow is less than the critical velocity of a given liquid flowing through a tube, the motion is streamlined. In such a case, we can imagine the entire thickness of the stream of the liquid to be made up of a large number of plane layers (laminae) one sliding past the other, i.e. one flowing over the other. Such a flow is called *laminar flow*.

If the velocity of flow exceeds the critical velocity υ_c, the mixing of streamlines takes place and the flow path becomes zig-zag. Such a motion is said to be *turbulent*.

9.8.2 Equation of Continuity

If an incompressible, non-viscous fluid flows through a tube of non-uniform cross section, the product of the area of cross section and the fluid speed at any point in the tube is constant for a streamline flow. Let A_1 and A_2 denote the areas of cross section of the tube where the fluid is entering and leaving, as shown in Fig. 9.28. If υ_1 and υ_2 are the speeds of the fluid at the ends A and B respectively, and ρ is the density of the fluid, then the liquid entering the tube at A covers a distance υ_1 in one second. So volume of the liquid entering per second$= A_1 \times \upsilon_1$. Therefore

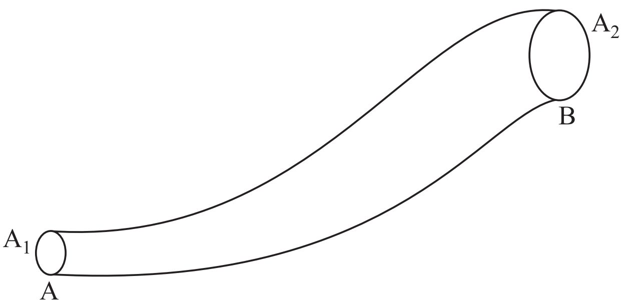


Fig. 9.28: Liquid flowing through a tube

Mass of the liquid entering per second at point A = $A_1 \upsilon_1 \rho$

Similarly, mass of the liquid leaving per second at point B = $A_2 \upsilon_2 \rho$

Since there is no accumulation of fluid inside the tube, the mass of the liquid crossing any section of the tube must be same. Therefore, we get

$$A_1 \upsilon_1 \rho = A_2 \upsilon_2 \rho$$

or
$$A_1 \upsilon_1 = A_2 \upsilon_2$$

This expression is called **equation of continuity**.

9.8.3 Critical Velocity and Reynolds's Number

We now know that when the velocity of flow is less than a certain value, valled *critical velocity*, the flow remains streamlined. But when the velocity of flow exceeds the critical velocity, the flow becomes turbulent.

The value of critical velocity of any liquid depends on the

- nature of the liquid, i.e. coefficient of viscosity (η) of the liquid;
- diameter of the tube (d) through which the liquid flows; and
- density of the liquid (ρ).

Experiments show that $\upsilon_c \,\alpha\, \eta$; $\upsilon_c \,\alpha\, \frac{1}{\rho}$ and $\upsilon_c \,\alpha\, \frac{1}{d}$.

Hence, we can write

$$\upsilon_c = R.\eta/\rho\, d \qquad (9.16)$$

where R is constant of proportionality and is called Reynolds's Number. It has no dimensions. Experiments show that if R is below 1000, the flow is laminar. The flow becomes unsteady when R is between 1000 and 2000 and the flow becomes turbulent for R greater than 2000.

Example 9.1: The average speed of blood in the artery (d =2.0 cm) during the resting part of heart's cycle is about 30 cm s^{-1}. Is the flow laminar or turbulent? Density of blood 1.05 g cm^{-3}; and $\eta = 4.0 \times 10^{-2}$ poise.

Solution: From Eqn. (9.16) we recall that Reynold's number $R = \upsilon_c\, \rho\, d/\eta$. On substituting the given values, we get

$$R = \frac{(30 \text{ cm s}^{-1}) \times 2\text{cm} \times (1.05 \text{g cm}^{-3})}{(4.0 \times 10^{-2} \text{g cm}^{-1} \text{s}^{-1})}$$

$$= 1575$$

Since 1575 < 2000, the flow is unsteady.

9.9 STOKES' LAW

George Stokes gave an empirical law for the magnitude of the tangential backward viscous force F acting on a freely falling smooth spherical body of radius r in a highly viscous liquid of coefficient of viscosity η moving with velocity υ. This is known as Stokes' law.

According to Stokes' law

$$F \propto \eta\ r\ \upsilon$$

or $$F = K\eta\ r\ \upsilon$$

where K is constant of proportionality. It has been found experimentally that $K = 6\pi$.

Hence Stokes' law can be written as

$$F = 6\pi\ \eta\ r\ \upsilon \qquad (9.17)$$

Stokes' Law can also be derived using the method of dimensions as follows:

According to Stokes, the viscous force depends on:

- coefficient of viscosity (η) of the medium
- radius of the spherical body (r)
- velocity of the body (υ)

Then $$F \propto \eta^a\ r^b\ \upsilon^c$$

or $$F = K\ \eta^a\ r^b\ \upsilon^c$$

where K is constant of proportionality

Taking dimensions on both the sides, we get

$$[MLT^{-2}] = [ML^{-1}T^{-1}]^a\ [L]^b\ [LT^{-1}]^c$$

or $$[MLT^{-2}] = [M^a\ L^{-a+b+c}\ T^{-a-c}]$$

Comparing the exponents on both the sides and solving the equations we get $a = b = c = 1$.

Hence $$F = K\ \eta\ r\ \upsilon$$

9.9.1 Terminal Velocity

Let us consider a spherical body of radius r and density ρ falling through a liquid of density σ.

The forces acting on the body will be

(i) Weight of the body **W** acting downward.

(ii) The viscous force **F** acting vertically upward.

(iii) The buoyant force **B** acting upward.

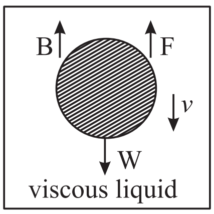


Fig. 9.29 : Force acting on a sphere falling in viscous fluid

Under the action of these forces, at some instant the net force on the body becomes zero, (since the viscous force increases with the increase of velocity). Then, the body falls with a constant velocity known as **terminal velocity**. We know that magnitude of these forces are

$$F = 6\pi\,\eta\, r\, \upsilon_0$$

where υ_0 is the terminal velocity.

$$W = (4/3)\,\pi\, r^3\ \rho g$$

and

$$B = (4/3)\,\pi\, r^3\ \sigma g$$

The net force is zero when object attains terminal velocity. Hence

$$6\pi\,\eta\, r\,\upsilon_0 = \frac{4}{3}\pi r^3\ \rho g - \frac{4}{3}\pi r^3\ \sigma g$$

Hence

$$\upsilon_0 = \frac{2r^2(\rho-\sigma)g}{9\eta} \qquad (9.18)$$

9.9.2 Applications of Stokes' Law

A. Parachute

When a soldier jumps from a flying aeroplane, he falls with acceleration due to gravity g but due to viscous drag in air, the acceleration goes on decreasing till he acquires terminal velocity. The soldier then descends with constant velocity and opens his parachute close to the ground at a pre-calculated moment, so that he may land safely near his destination.

B. Velocity of rain drops

When raindrops fall under gravity, their motion is opposed by the viscous drag in air. When viscous force becomes equal to the force of gravity, the drop attains a terminal velocity. That is why rain drops reaching the earth do not have very high kinetic energy.

Example 9.2: Determine the radius of a drop of rain falling through air with terminal velocity 0.12 ms^{-1}. Given $\eta = 1.8 \times 10^{-5}$ kg m^{-1} s^{-1}, $\rho = 1.21$ kg m^{-3}, $\sigma = 1.0 \times 10^3$ kg m^{-3} and $g = 9.8$ m s^{-2}.

Solution: We know that terminal velocity is given by

$$\upsilon_0 = \frac{2r^2(\rho-\sigma)g}{9\eta}$$

On rearranging terms, we can write

$$r = \sqrt{\frac{9\eta\,\upsilon_0}{2\,(\rho-\sigma)g}}$$

$$= \sqrt{\frac{9 \times 1.8 \times 10^{-5} \times 0.12}{2\,(1000 - 1.21)\,9.8}}\ \text{m}$$

$$= 10^{-5}\ \text{m}$$

INTEXT QUESTIONS 9.4

1. Differentiate between streamline flow and turbulent flow?
2. Can two streamlines cross each other in a flowing liquid?
3. Name the physical quantities on which critical velocity of a viscous liquid depends.
4. Calculate the terminal velocity of a rain drop of radius 0.01m if the coeflicient of viscosity of air is 1.8×10^{-5} Ns m^{-2} and its density is 1.2 kg m^{-3}. Density of water = 1000 kg m^{-3}. Take g = 10 m s^{-2}.
5. When a liquid contained in a tumbler is stirred and placed for some time, it comes to rest, Why?

Daniel Bernoulli (1700-1782)

Daniel Bernoulli, a Swiss Physicist and mathematician was born in a family of mathematicians on February 8, 1700. He made important contributions in hydrodynamics. His famous work, *Hydrodyanamica* was published in 1738. He also explained the behavior of gases with changing pressure and temperature, which led to the development of kinetic theory of gases.

He is known as the founder of mathematical physics. Bernoulli's principle is used to produce vacuum in chemical laboratories by connecting a vessel to a tube through which water is running rapidly.

9.10 BERNOULLI'S PRINCIPLE

Have you ever thought how air circulates in a dog's burrow, smoke comes quickly out of a chimney or why car's convertible top bulges upward at high speed? You must have definitely experienced the bulging upwards of your umbrella on a stormy- rainy day. All these can be understood on the basis of Bernoulli's principle.

Bernoulli's Principle states that **where the velocity of a fluid is high, the pressure is low and where the velocity of the fluid is low, pressure is high.**

9.10.1 Energy of a Flowing Fluid

Flowing fluids possess three types of energy. We are familiar with the kinetic and potential energies. The third type of energy possessed by the fluid is pressure energy. It is due to the pressure of the fluid. The pressure energy can be taken as the product of pressure difference and its volume. If an element of liquid of mass *m*, and density *d* is moving under a pressure difference *p*, then

$$\text{Pressure energy } = p \times (m/d) \text{ joule}$$

Pressure energy per unit mass = (p/d) J kg^{-1}

9.10.2 Bernoulli's Equation

Bernoulli developed an equation that expresses this principle quantitatively. Three important assumptions were made to develop this equation:

1. The fluid is incompressible, i.e. its density does not change when it passes from a wide bore tube to a narrow bore tube.
2. The fluid is non-viscous or the effect of viscosity is not to be taken into account.
3. The motion of the fluid is streamlined.

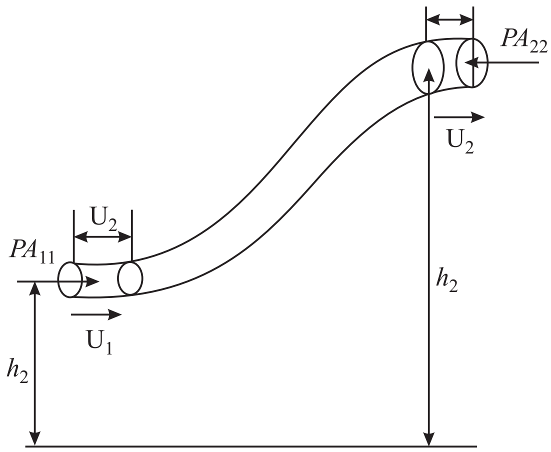


Fig. 9.30

We consider a tube of varying cross section shown in the Fig. 9.30. Suppose at point A the pressure is P_1, area of cross section A_1, velocity of flow υ_1, height above the ground h_1 and at B, the pressure is P_2 ,area of cross-section A_2 velocity of flow = υ_2, and height above the ground h_2.

Since points A and B can be any two points along a tube of flow, we write Bernoulli's equation

$$P + 1/2\, d\upsilon^2 + h\, dg = \text{Constant.}$$

That is, the sum of pressure energy, kinetic energy and potential energy of a fluid remains constant in streamline motion.

ACTIVITY 9.4

1. Take a sheet of paper in your hand.
2. Press down lightly on horizontal part of the paper as shown in Fig. 9.31 so that the paper curves down.
3. Blow on the paper along the horizontal line.

Watch the paper. It lifts up because speed increases and pressure on the upper side of the paper decreases.

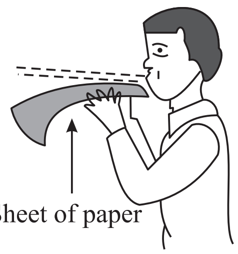


Fig. 9.31

9.10.3 Applications of Bernoulli's Theorem

Bernoulli's theorem finds many applications in our lives. Some commonly observed phenomena can also be explained on the basis of Bernoulli's theorem.

A. Flow meter or Venturimeter

It is a device used to measure the rate of flow of liquids through pipes. The device is inserted in the flow pipe, as shown in the Fig. 9.32

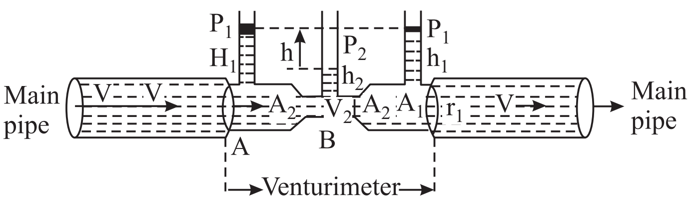


Fig. 9.32 : A Venturimeter

It consists of a manometer, whose two limbs are connected to a tube having two different cross-sectional areas say A_1 and A_2 at A and B, respectively. Suppose the main pipe is horizontal at a height *h* above the ground. Then applying Bernoulli's theorem for the steady flow of liquid through the venturimeter at A and B, we can write

Total Energy at A = Total Energy At B

$$\frac{1}{2} m\, \upsilon_1^2 + mgh + \frac{mp_1}{d} = \frac{1}{2} m\, \upsilon_2^2 + mgh + \frac{mp_2}{d}$$

On rearranging terms we can write,

$$(p_1 - p_2) = \frac{d}{2}\,(\upsilon_2^2 - \upsilon_1^2) = \frac{\upsilon_1^2 d}{2}\left[\left(\frac{\upsilon_2}{\upsilon_1}\right)^2 - 1\right] \qquad (9.19)$$

It shows that points of higher velocities are the points of lower pressure (because of the sum of pressure energy and K.E. remain constant). This is called *Venturi's Principle*.

For steady flow through the ventrurimeter, volume of liquid entering per second at A = liquid volume leaving per second at B. Therefore

$$A_1\upsilon_1 = A_2\upsilon_2 \qquad (9.20)$$

(The liquid is assumed incompressible i.e., velocity is more at narrow ends and vice versa.

Using this result in Eqn. (9.19), we conclude that pressure is lesser at the narrow ends;

$$p_1 - p_2 = \frac{\upsilon_1^2 d}{2}\left[\frac{A_1^2}{A_2^2} - 1\right]$$

$$= \frac{1}{2} d\upsilon_1^2 \left[\left(\frac{A_1}{A_2}\right)^2 - 1\right]$$

$$\upsilon_1 = \sqrt{\frac{2(p_1 - p_2)}{d\left(\dfrac{A_1^2}{A_2^2}\right) - 1}} \qquad (9.21)$$

If h denotes level difference between the two limbs of the venturimeter, then

$$p_1 - p_2 = h\,d\,g$$

and $$\upsilon_1 = \sqrt{2hg/[(A_1^2/A_2^2) - 1]}$$

From this we note that $\upsilon_1 \propto \sqrt{h}$ since all other parameters are constant for a given venturimeter. Thus

$$\upsilon_1 = K\sqrt{h};$$

where K is constant.

The volume of liquid flowing per second is given by

$$V = A_1\,\upsilon_1 = A_1 \times K\sqrt{h}$$

or $$V = K'h$$

where $K' = K A_1$ is another constant.

Bernaulli's principle has many applications in the design of many useful appliances like atomizer, spray gun, Bunsen burner, carburetor, Aerofoil, etc.

(i) Atomizer : An atomizer is shown in Fig. 9.33. When the rubber bulb A is squeezed, air blows through the tube B and comes out of the narrow orifice with larger velocity creating a region of low pressure in its neighborhood. The liquid (scent or paint) from the vessel is, therefore, sucked into the tube to come out to the nozzles N. As the liquid reaches the nozzle, the air stream from the tube B blows it into a fine spray.

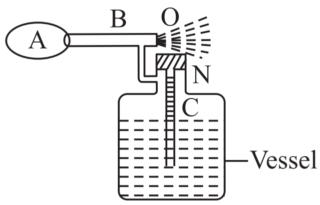


Fig. 9.33 : Atomizer

(ii) Spray gun : When the piston is moved in, it blows the air out of the narrow hole 'O' with large velocity creating a region of low pressure in its neighborhood. The liquid (e.g. insecticide) is sucked through the narrow tube attached to the vessel end having its opening just below 'O'. The liquid on reaching the end gets sprayed by out blown air from the piston (Fig. 9.34).

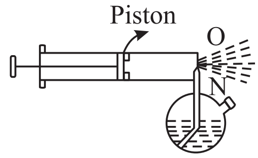


Fig. 9.34: Spray gun

(iii) Bunsen Burner : When the gas emerges out of the nozzle N, its velocity being high the pressure becomes low in its vicinity. The air, therefore, rushed in through the side hole A and gets mixed with the gas. The mixture then burns at the mouth when ignited, to give a hot blue flame (Fig.9.35).

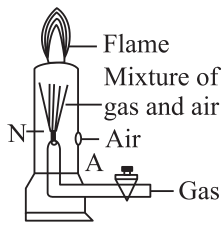


Fig. 9.35: Bunsen Burner

(iv) Carburetor : The carburetor shown in Fig. 9.36. is a device used in motor cars for supplying a proper mixture of air and petrol vapours to the cylinder of the engine. The energy is supplied by the explosion of this mixture inside the cylinders of the engine. Petrol is contained in the float chamber. There is a decrease in the pressure on the side A due to motion of the piston.This causes the air from outside to be sucked in with large velocity. This causes a low pressure near the nozzle B (due

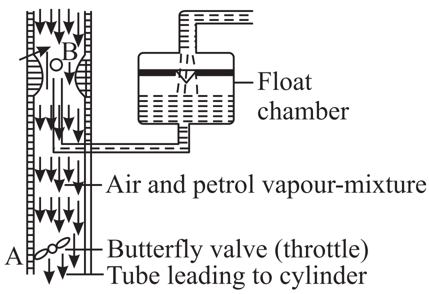


Fig. 9.36: Carburettor

to constriction, velocity of air sucked is more near B) and, therefore, petrol comes out of the nozzle B which gets mixed with the incoming. Air. The mixture of vaporized petrol and air forming the fuel then enters the cylinder through the tube A.

(Sometimes when the nozzle B gets choked due to deposition of carbon or some impurities, it checks the flow of petrol and the engine not getting fuel stops working. The nozzle has therefore, to be opened and cleaned.

(v) Aerofoil : When a solid moves in air , streamlines are formed . The shape of the body of the aeroplane is designed specially as shown in the Fig. 9.37. When the aeroplane runs on its runway, high velocity streamlines of air are formed. Due to crowding of more streamlines on the upper side, it becomes a region of more velocity and hence of comparatively low pressure region than below it. This pressure difference gives the lift to the aeroplane.

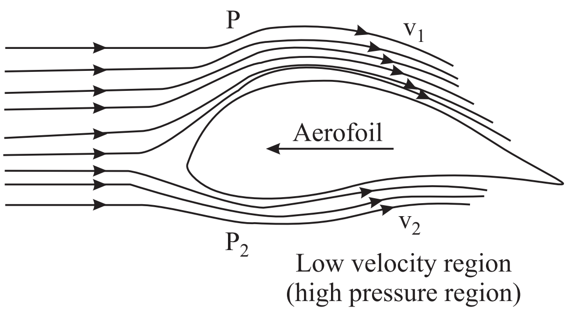


Fig. 9.37 : Crowding of streamlines on the upper side.

Based on this very principle i.e., the regions of high velocities due to crowding of steam lines are the regions of low pressure, following are interesting demonstrations.1

(a) Attracted disc paradox : When air is blown through a narrow tube handle into the space between two cardboard sheets [Fig. 9.38] placed one above the other and the upper disc is lifted with the handle, the lower disc is attracted to stick to the upper disc and is lifted with it. This is called attracted disc paradox,

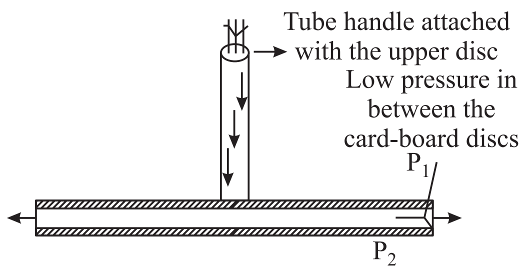


Fig. 9.38 : Attracted disc paradox

Notes

(b) Dancing of a ping pong ball on a jet of water:

If a light hollow spherical ball (ping-pong ball or table tennis ball) is gently put on a vertical stream of water coming out of a vertically upward directed jet end of a tube, it keeps on dancing this way and that way without falling to the ground (Fig.9.39). When the ball shifts to the lefts , then most of the jet streams pass by its right side thereby creating a region of high velocity and hence low pressure on its right side in comparison to that on the left side and the ball is again pushed back to the center of the jet stream .

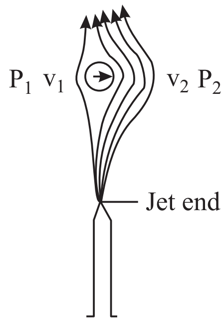


Fig. 9.39: Dancing Pring Pongball

(c) Water vacuum pump or aspirator or filter pump : Fig. 9.40 shows a filter pump used for producing moderately low pressures. Water from the tap is allowed to come out of the narrow jet end of the tube A . Due to small aperture of the nozzle, the velocity becomes high and hence a low-pressure region is created around the nozzle N. Air is, therefore, sucked from the vessel to be evacuated through the tube B; gets mixed with the steam of water and goes out through the outlet. After a few minutes., the pressure of air in the vessel is decreased to about 1 cm of mercury by such a pump

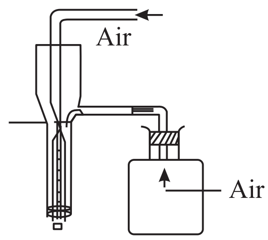


Fig. 9.40 : Filter Pump

(d) Swing of a cricket ball:

When a cricketer throws a spinning ball, it moves along a curved path in the air. This is called swing of the ball. It is clear from Fig. 9.41. That when a ball is moved forward, the air occupies the space left by the ball with a velocity υ (say). When the ball spins, the layer of air around it also moves with the ball, say with the velocity 'u'. So the resultant velocity of air above the ball becomes $(\upsilon - u)$ and below the ball becomes $(\upsilon + u)$. Hence, the pressure difference above and below the ball moves the ball in a curved path.

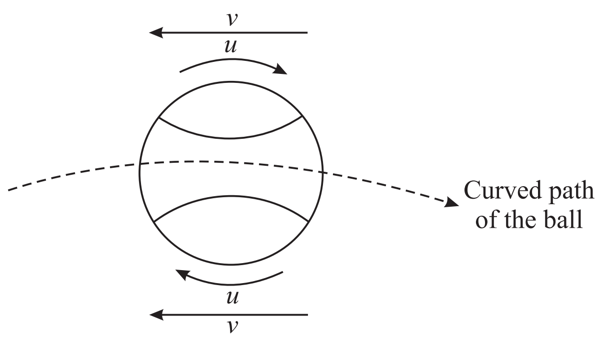


Fig. 9.41 : Swing of a cricket ball

Notes

Example 9.3: Water flows out of a small hole in the wall of a large tank near its bottom (Fig. 942). What is the speed of efflux of water when the height of water level in the tank is 2.5m?

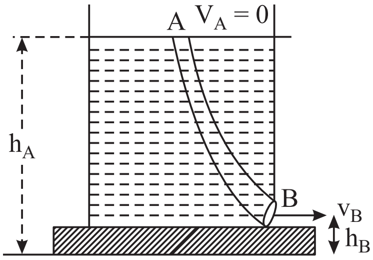


Fig. 9.42

Solution: Let B be the hole near the bottom. Imagine a tube of flow A to B for the water to flow from the surface point A to the hole B. We can apply the Bernoulli's theorem to the points A and B for the streamline flow of small mass m.

Total energy at B = Total energy at A

At A, $\upsilon_A = 0$, $p_A = p$ = atmospheric pressure, h = height above the ground.

At B, $\upsilon_B = \upsilon = ?$, $p_B = p$, h_B = height of the hole above the ground.

Let $h_A - h_B$ = H = height of the water level in the vessel = 2.5m

and d = density of the water.

Applying the Bernoulli's Principle and substituting the values we get,

$$\tfrac{1}{2} m\, \upsilon_B^2 = mg\,(h_A - h_B)$$

or

$$\upsilon_B = \sqrt{2g(h_A - h_B)}$$

$$= \sqrt{2 \times 9.8 \times 2.5}$$

$$= 7 \text{ m s}^{-1}$$

INTEXT QUESTIONS 9.5

1. The windstorm often blows off the tin roof of the houses, How does Bernoulli's equation explain the phenomenon?

2. When you press the mouth of a water pipe used for watering the plants, water goes to a longer distance, why?

3 What are the conditions necessary for the application of Bernoulli's theorem to solve the problems of flowing liquid?

4. Water flows along a horizontal pipe having non-uniform cross section. The pressure is 20 mm of mercury where the velocity is 0.20m/s. find the pressure at a point where the velocity is 1.50 m/s?

5. Why do bowlers in a cricket match shine only one side of the ball?

WHAT YOU HAVE LEARNT

- Hydrostatic pressure P at a depth *h* below the free surface of a liquid of density is given by

$$P = hdg$$

- The upward force acting on an object submerged in a fluid is known as buoyant force.
- According to Pascal's law, when pressure is applied to any part of an enclosed liquid, it is transmitted undiminished to every point of the liquid as well as to the walls of the container.
- The liquid molecules in the liquid surface have potential energy called surface energy.
- The surface tension of a liquid may be defined as force per unit length acting on a imaginary line drawn in the surface. It is measured in Nm^{-1}.
- Surface tension of any liquid is the property by virtue of which a liquid surface acts like a stretched membrane.
- Angle of contact is defined as the angle between the tangent to the liquid surface and the wall of the container at the point of contact as measured from within the liquid.
- The liquid surface in a capillary tube is either concave or convex. This curvature is due to surface tension. The rise in capillary is given by

$$h = \frac{2T\cos\theta}{r\,d\,g}$$

- The excess pressure P on the concave side of the liquid surface is given by

  $P = \frac{2T}{R}$, where *T* is surface tension of the liquid

  $P = \frac{2T}{R}$, for air bubble in the liquid and

  $P = \frac{4T'}{r}$, where T' is surface tension of soap solution, for soap bubble in air

- Detergents are considered better cleaner of clothes because they reduce the surface tension of water-oil.
- The property of a fluid by virtue of which it opposes the relative motion between its adjacent layers is known as viscosity.

- The flow of liquid becomes turbulent when the velocity is greater than a certain value called critical velocity (υ_c) which depends upon the nature of the liquid and the diameter of the tube i.e. (η.P and d).
- Coefficient of viscosity of any liquid may be defined as the magnitude of tangential backward viscus force acting between two successive layers of unit area in contact with each other moving in a region of unit velocity gradient.
- Stokes' law states that tangential backward viscous force acting on a spherical mass of radius r falling with velocity 'υ' in a liquid of coefficient of viscosity η is given by

$$F = 6\pi\,\eta\, r\, \upsilon.$$

- Bernoulli's theorem states that the total energy of an element of mass (m) of an incompressible liquid moving steadily remains constant throughout the motion. Mathematically, Bernoullis's equation as applied to any two points A and B of tube of flow

$$\frac{1}{2} m\, v_A^2 + m\, g\, h_A + \frac{m P_A}{d} = \frac{1}{2} m\, v_B^2 + m\, g\, h_B + \frac{m P_B}{d}$$

TERMINAL EXERCISES

1. Derive an expression for hydrostatic pressure due to a liquid column.
2. State pascal's law. Explain the working of hydraulic press.
3. Define surface tension. Find its dimensional formula.
4. Describe an experiment to show that liquid surfaces behave like a stretched membrane.
5. The hydrostatic pressure due to a liquid filled in a vessel at a depth 0.9 m is 3.0 N m^2. What will be the hydrostatic pressure at a hole in the side wall of the same vessel at a depth of 0.8 m.
6. In a hydraulic lift, how much weight is needed to lift a heavy stone of mass 1000 kg? Given the ratio of the areas of cross section of the two pistons is 5. Is the work output greater than the work input? Explain.
7. A liquid filled in a capillary tube has convex meniscus. If F_a is force of adhesion, F_c is force of cohesion and θ = angle of contact, which of the following relations should hold good?

   (a) $F_a > F_c \sin\theta$; (b) $F_a < F_c \sin\theta$; (c) $F_a \cos\theta = F_c$; (d) $F_a \sin\theta > F_c$
8. 1000 drops of water of same radius coalesce to form a larger drop. What happens to the temperature of the water drop? Why?

9. What is capillary action? What are the factors on which the rise or fall of a liquid in a capillary tube depends?

10. Calculate the approximate rise of a liquid of density 10^3 kg m^{-3} in a capillary tube of length 0.05 m and radius 0.2×10^{-3} m. Given surface tension of the liquid for the material of that capillary is 7.27×10^{-2} N m^{-1}.

11. Why is it difficult to blow water bubbles in air while it is easier to blow soap bubble in air?

12. Why the detergents have replaced soaps to clean oily clothes.

13. Two identical spherical balloons have been inflated with air to different sizes and connected with the help of a thin pipe. What do you expect out of the following observations?

    (i) The air from smaller balloon will rush into the bigger balloon till whole of its air flows into the later.

    (ii) The air from the bigger balloon will rush into the smaller balloon till the sizes of the two become equal.

    What will be your answer if the balloons are replaced by two soap bubbles of different sizes.

14. Which process involves more pressure to blow a air bubble of radius 3 cm inside a soap solution or a soap bubble in air? Why?

15. Differentiate between laminar flow and turbulent flow and hence define critical velocity.

16. Define viscosity and coefficient of viscosity. Derive the units and dimensional formula of coefficient of viscosity. Which is more viscous : water or glycerine? Why?

17. What is Reynold's number? What is its significance? Define critical velocity on the basis of Reynold's number.

18. State Bernoulli's principle. Explain its application in the design of the body of an aeroplane.

19. Explain Why :

    (i) A spinning tennis ball curves during the flight?

    (ii) A ping pong ball keeps on dancing on a jet of water without falling on to either side?

    (iii) The velocity of flow increases when the aperture of water pipe is decreased by squeezing its opening.

    (iv) A small spherical ball falling in a viscous fluid attains a constant velocity after some time.

(v) If mercury is poured on a flat glass plate; it breaks up into small spherical droplets.

20. Calculate the terminal velocity of an air bubble with 0.8 mm in diameter which rises in a liquid of viscosity of 0.15 kg m^{-1} s^{-1} and density 0.9 g m^{-3}. What will be the terminal velocity of the same bubble while rising in water? For water $\eta = 10^{-2}$ kg m^{-1} s^{-1}.

21. A pipe line 0.2 m in diameter, flowing full of water has a constriction of diameter 0.1 m. If the velocity in the 0.2 m pipe-line is 2 m s^{-1}. Calculate
    (i) the velocity in the constriction, and
    (ii) the discharge rate in cubic meters per second.

22. (i) With what velocity in a steel ball 1 mm is radius falling in a tank of glycerine at an instant when its acceleration is one-half that of a freely falling body?
    (ii) What is the terminal velocity of the ball? The density of steel and of glycerine are 8.5 gm cm^{-3} and 1.32 g cm^{-3} respectively; viscosity of glycerine is 8.3 Poise.

23. Water at 20ºC flows with a speed of 50 cm s^{-1} through a pipe of diameter of 3 mm.
    (i) What is Reynold's number?
    (ii) What is the nature of flow?

    Given, viscosity of water at 20ºC as = 1.005×10^{-2} Poise; and

    Density of water at 20ºC as = 1 g cm^{-3}.

24. Modern aeroplane design calls for a lift of about 1000 N m^{-2} of wing area. Assume that air flows past the wing of an aircraft with streamline flow. If the velocity of flow past the lower wing surface is 100 m s^{-1}, what is the required velocity over the upper surface to give a desired lift of 1000 N m^{-2}? The density of air is 1.3 kg m^{-3}.

25. Water flows horizontally through a pipe of varying cross-section. If the pressure of water equals 5 cm of mercury at a point where the velocity of flow is 28 cm s^{-1}, then what is the pressure at another point, where the velocity of flow is 70 cm s^{-1}? [Tube density of water 1 g cm^{-3}].

ANSWERS TO INTEXT QUESTIONS

9.1

1. Because then the weight of the person applies on a larger area hence pressure on snow decreases.

2. $P = P_a + \rho gh$

   $P = 1.5 \times 10^7$ Pa

3. Pressure applied by the weight of the boy $= \dfrac{2.5}{0.05} = 500$ N m^{-2}.

   Pressure due to the weight of the elephant $= \dfrac{5000}{10} = 500$ N m^{-2}.

   ∴ The boy can balance the elephant.
4. Because of the larger area of the rod, pressure on the skin is small.
5. $\dfrac{50}{0.1} = \dfrac{w}{10}$, $w = 5000$ kg wt.

9.2

1. Force of attraction between molecules of same substance is called force of cohesion and the force of attractive between molecules of different substance is called force of adhesion.
2. Surface tension leads to the minimum surface area and for a given volume, sphere has minimum surface area.
3. No, they have tightly bound molecules.
4. Due to surface tension forces.
5. For air bubble in water

   $$P = \frac{2T}{r} = \frac{2\times 727\times 10^{-3}}{2\times 10^{-2}} = 72.7 \text{ N m}^{-2}.$$

   For soap bubble in air

   $$P' = \frac{4T'}{r'} = \frac{4\times 25\times 10^{-3}}{4\times 10^{-2}} = 2.5 \text{ N m}^{-2}.$$

9.3

1. No.
2. Yes, the liquid will rise.
3. Mercury has a convex meniscus and the angle of contact is obtuse. The fall in the level of mercury in capillary makes it difficult to enter.

4. $r = \frac{2T}{h\rho g} = \frac{2 \times 7.2 \times 10^{-2}}{3 \times 1000 \times 10}$

$= 4.8 \times 10^{-6}$m.

5. Due to capillary action.

9.4

1. If every particle passing through a given point of path follows the same line of flow as that of preceding particle the flow is stream lined, if its zig-zag, the flow is turbulent.
2. No, otherwise the same flow will have two directions.
3. Critical velocity depends upon the viscous nature of the liquid, the diameter of the tube and density of the liquid.
4. .012 ms^{-1}
5. Due to viscous force.

9.5

1. High velocity of air creates low pressure on the upper part.
2. Decreasing in the area creates large pressure.
3. The fluid should be incompressible and non-viscous on (very less). The motion should be steamlined.
4. $(P_1 - P_2) = \frac{1}{2} d\,(\upsilon_2^2 - \upsilon_1^2)$
5. So that the stream lines with the two surfaces are different. More swing in the ball will be obtained.

Answers to the Terminal Exercises

5. 2.67 N m^{-2}.
6. 200 N, No.
20. 2.1 mm s^{-1}, 35 cm s^{-1}.
21. 8 m s^{-1}, 6.3×10^{-2} m^3 s^{-1}.
22. 7.8 mm s^{-1}, 0.19 m s^{-1}.
23. 1500, Unsteady.
24. 2 cm of mercury.

312en10

10

KINETIC THEORY OF GASES

As you have studied in the previous lessons, at standard temperature and pressure, matter exists in three states – solid, liquid and gas. These are composed of atoms/molecules which are held together by intermolecular forces. At room temperature, these atoms/molecules have finite thermal energy. If thermal energy increases, molecules begin to move more freely. This state of matter is said to be the gaseous state. In this state, intermolecular forces are very weak and very small compared to their kinetic energy.

Under different conditions of temperature, pressure and volume, gases exhibit different properties. For example, when the temperature of a gas is increased at constant volume, its pressure increases. In this lesson you will learn the kinetic theory of gases which is based on certain simplifying assumptions. You will also learn the kinetic interpretation of temperature and its relationship with the kinetic energy of the molecules. Why the gases have two types of heat capacities and concept of thermal expansion will also be explained in this lesson.

OBJECTIVES

After studying this lesson, you should be able to :

- *define heat capacity and specific heat;*
- *state principle of caorimetry;*
- *explain thermal expansion;*
- *derive relation between* α, β *and* γ;
- *state the assumptions of kinetic theory of gases;*
- *derive the expression for pressure* $P = \frac{1}{3}\rho\overline{c^2}$;
- *explain how rms velocity and average velocity are related to temperature;*
- *derive gas laws on the basis of kinetic theory of gases;*

- *give kinetic interpretation of temperature and compute the mean kinetic energy of a gas;*
- *explain degrees of freedom of a system of particles;*
- *explain the law of equipartition of energy;*
- *explain why a gas has two heat capacities; and*
- *derive the relation* $c_p - c_V = R/J$.

10.1 THERMAL ENERGY

During a year, the spring (Basant) season, when the temperature is not as high as in summer and not as low as in winter, is very pleasant. How does this change in temperature affect our day to day activities? How do things change their properties with change in temperature? Is there any difference between temperature and heat? All such questions will be discussed in the following sections.

The term temperature and heat are often used interchangeably in everyday language. In Physics, however, these two terms have very different meaning. Supply of heat does often increase the temperature but does it happen so when water boils or freezes? Why the wind in the coastal areas often reverses direction after the Sun goes down? Why does ice melt when kept on the palm of hand and why does the palm feel cool? All these facts will be explained in this chapter.

10.1.1 Heat Capacity and Specific Heat

When heat is supplied to a solid (or liquid), its temperature increases. The rise in temperature is found to be different in different solids in spite of having the same mass and being supplied the same quantity of heat. This simply implies that the rise in the temperature of a solid, when a certain amount of heat is supplied to it, depends upon the nature of the material of the solid. The nature of the solid is characterized by the term specific heat capacity or specific heat of the solid. The specific heat of the material of a solid (or a liquid) may be defined as the amount of heat required to raise the temperature of its unit mass through 1°C or 1K.

If an amount of heat ΔQ is required to raise the temperature of a mass m of the solid (or liquid) through $\Delta\theta$, then the specific heat may be expressed as

$$C = \frac{\Delta Q}{m\Delta\theta}$$

Thus, the amount of heat required to raise the temperature of a substance is given by:

$$\Delta Q = mC\Delta\theta$$

SI unit of specific heat is J kg^{-1} K^{-1}

10.1.2 Calorimetry

When two bodies at different temperatures are kept in contact, transfer of heat takes place from the body at higher temperature to the body at lower temperature till both the bodies acquire the same temperature. The specific heat of a material and other physical quanties related to this heat transfer are measured with the help of a device called calorimeter and the process of the measurement is called calcorimetry.

10.1.3 Principle of Calorimetry

Let two substances of mass m_1 and m_2, of specific heat capacities C_1 and C_2 and at temperatures θ_1 and $\theta_2(\theta_1 > \theta_2)$, respectively be kept in contact. Then, the heat will be transferred from the higher to the lower temperature and the substances will acquire the same temperature θ. (say) assuming that no energy loss takes place to the surroundings and applying the law of conservation of energy, we can say

$$\text{Heat lost} = \text{Heat gained.}$$

$$\Rightarrow \qquad m_1 C_1(\theta_1 - \theta) = m_2 C_2(\theta - \theta_2)$$

This is the principle of calorimetry. By using this relation the resultant temperature θ can be determined. Also, by knowing θ_1, θ_2 and θ the specific heat capacity of a substance can be determined if the specific heat capacity of the other substance is known.

10.1.4 Thermal Expansion

When heat is given to a substance it expands in length, area or volume. This is called thermal expansion. The expansion in length, area and volume are called linear, superficial and cubical expansion, respectively.

In linear expansion, the change in length is directly proportional to the original length and change in temperature.

$$\Delta l \propto l_0 \, \Delta\theta$$

or

$$\Delta l = \alpha \, l_0 \, \Delta\theta$$

where α is the coefficient of linear expansion or temperature coefficient of linear expansion. It is given by

$$\alpha = \frac{\Delta l}{l_0 \Delta\theta}$$

If, $\Delta\theta = 1°C$ and $l_0 = 1\text{m}$

Then $\alpha = \Delta l$

Thus, α is defined as the change in length of unit length of the substance whose temperature is increased by 1°C.

In superficial expansion, the change in area is directly proportional to the original area and change in temperature:

$$\Delta A \propto A_0 \, \Delta\theta$$

or $$\Delta A = \beta \, A_0 \, \Delta\theta$$

where β is the temperature coefficient of superficial expansion.

In cubical expansion, the change in volume is directly proportional to the change in temperature and original volume:

$$\Delta V \propto V_0 \, \Delta\theta$$

or $$\Delta V = \gamma V_0 \, \Delta\theta$$

where γ in the temperature coefficient of cubical expansion.

If $V_0 = 1 m^3$ and $\Delta\theta = 1°C$, then $\gamma = \Delta V$

Thus, coefficient of cubical expansion is defined as the change in volume of a unit volume of a substance whose temperature is increased by 1°C.

Relation between α, β and γ

Let there be a cube of side l whose temperature is increased by 1°C.

The change in length:

$$\Delta l = \alpha l \Delta\theta$$

$$= \alpha l \qquad (\because \Delta\theta = l°C)$$

or, new length $l' = l + \Delta l = l + \alpha l = l(1+\alpha)$

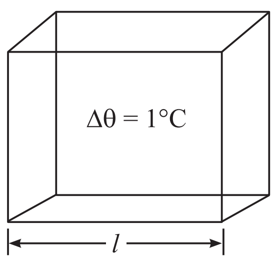


Fig. 10.1

Thus, $$\alpha = \frac{\Delta l}{l}$$

And $$\beta = \frac{\Delta A}{A} = \frac{l^2(1+\alpha)^2 - l^2}{l^2}$$

$$= 1 + \alpha^2 + 2\alpha - 1$$

Since α is very small therefore α^2 may be neglected. We therefore have

$$\beta = 2\alpha$$

similarly, $$\gamma = \frac{\Delta V}{V} = \frac{l^3(1+\alpha)^3 - l^3}{l^3}$$

or, $$\gamma = l^3 + \alpha^3 + 3\alpha^2 + 3\alpha - l^3$$

As α is very small, the term α^2 and α^3 may be neglected. We , therefore, have

$$\therefore \quad \gamma = 3\alpha.$$

10.1.5 Anomalous expansion in water and its effect

Generally, the volume of a liquid increases with increase in temperature. The coefficient of expansion of liquids is about 10 times that of solids. However the volume of water does not increase with temperature between 0 to 4°C.

As the temperature increases from 0°C to 4°C, the water contracts and hence the density of water reaches a maximum value of 1 g mL^{-1} or 1000 kg m^{-3} at 4°C. After that the volume starts increasing (while the density decreases) as shown in Fig. 10.2.

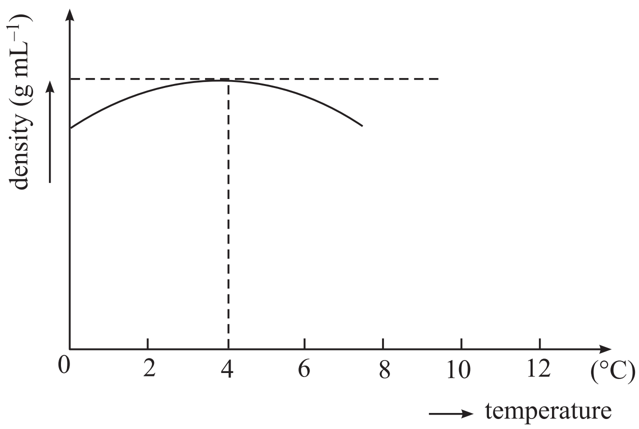


Fig. 10.2

Now, it can be understood why a pond or lake freezes at its surface whereas water may remain below it in liquid state. As the pond cools, the colder, denser water at the surface initially sinks to the bottom. When the temperature of the entire water body reaches 4°C, this flow stops. The temperating of surface water keeps on decreasing and freezes ultimately at 0°C. As water freezes at the surface, it remains there since ice is a bad conductor of heat; and since ice is less denser than water the ice continues to build up at the surface whereas water near the bottom remains at 4°C. If this had not happened fish and all the marine life would not have survived.

10.1.6 Thermal Expansion in Gases

When heat is supplied to the gases they also expand. This expansion is very large as compared to solids and liquids. But in case of a gas pressure and volume both may change simultaneously with rise in temperature. Hence we have to consider either expansion of the gas with temperature at constant pressure or the increase in its pressure at constant volume. Thus the coefficient of volume expansion of a gas at constant pressure is given by

$$\gamma_v = \left(\frac{V_2 - V_1}{V_1 \Delta\theta}\right)_{\Delta p=0}$$

and similarly

$$\gamma_p = \left(\frac{p_2 - p_1}{p_1 \Delta\theta}\right)_{\Delta v=0}$$

10.2 KINETIC THEORY OF GASES

You now know that matter is composed of very large number of atoms and molecules. Each of these molecules shows the characteristic properties of the substance of which it is a part. Kinetic theory of gases attempts to relate the macroscopic or bulk properties such as pressure, volume and temperature of an ideal gas with its microscopic properties such as speed and mass of its individual molecules. The kinetic theory is based on certain assumptions. (A gas whose molecules can be treated as point masses and there is no intermolecular force between them is said to be ideal.) A gas at room temperature and atmospheric pressure (low pressure) behaves like an ideal gas.

10.2.1 Assumptions of Kinetic Theory of Gases

Clark Maxwell in 1860 showed that the observed properties of a gas can be explained on the basis of certain assumptions about the nature of molecules, their motion and interaction between them. These resulted in considerable simplification. We now state these.

(i) A gas consists of a very large number of identical rigid molecules, which move with all possible velocities randomly. The intermolecular forces between them are negligible.

(ii) Gas molecules collide with each other and with the walls of the container. These collisions are perfectly elastic.

(iii) Size of the molecules is negligible compared to the separation between them.

(iv) Between collisions, molecules move in straight lines with uniform velocities.

(v) Time taken in a collision is negligible as compared to the time taken by a molecule between two successive collisions.

(vi) Distribution of molecules is uniform throughout the container.

To derive an expression for the pressure exerted by a gas on the walls of the container, we consider the motion of only one molecule because all molecules are identical (Assumption i). Moreover, since a molecule moving in space will have velocity components along x, y and z–directions, in view of assumption (vi)it is enough for us to consider the motion only along one dimension, say x-axis.(Fig. 10.1). Note that if there were N ($= 6 \times 10^{26}$ molecules m^{-3}), instead of considering $3N$ paths, the assumptions have reduced the roblem to only one molecule in one dimension. Let us consider a molecule having velocity C in the face LMNO. Its x, y and z components are u, v and w, respectively. If the mass of the molecule is m and it is moving with a speed u *along x–axis*, its momentum will be mu towards the wall and normal to it. On striking the wall, this molecule will rebound in the opposite direction with the same speed u, since the collision has been assumed to be perfectly elastic (Assumption ii). The momentum of the molecule after it rebounds is ($-mu$). Hence, the change in momentum of a molecule is

$$mu - (-mu) = 2mu$$

If the molecule travels from face LMNO to the face ABCD with speed u along *x–axis* and rebounds back without striking any other molecule on the way, it covers a distance $2l$ in time $2l/u$. That is, the time interval between two successive collisions of the molecules with the wall is $2l/u$.

According to Newton's second law of motion, the rate of change of momentum is equal to the impressed force. Therefore

$$\text{Rate of change of momentum at ABCD} = \frac{\text{Change in momentum}}{\text{Time}}$$

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

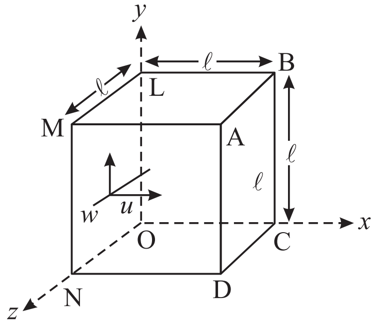


Fig. 10.3 : Motion of a molecule in a container

This is the rate of change of momentum of one molecule. Since there are N molecules of the gas, the total rate of change of momentum or the total force exerted on the wall ABCD due to the impact of all the N molecules moving along x-axis with speeds, $u_1, u_2, ..., u_N$ is given by

$$\text{Force on ABCD} = \frac{m}{l}\left(u_1^2 + u_2^2 + u_3^2 + \ldots + u_N^2\right)$$

We know that pressure is force per unit area. Therefore, the pressure P exerted on the wall ABCD of areas l^2 by the molecules moving along x-axis is given by

$$P = \frac{\frac{m}{l}\left(u_1^2 + u_2^2 + \ldots + u_N^2\right)}{l^2}$$

$$= \frac{m}{l^3}\left(u_1^2 + u_2^2 + \ldots + u_N^2\right) \qquad (10.1)$$

If $\overline{u^2}$ represents the mean value of the squares of all the speed components along x-axis, we can write

$$\overline{u^2} = \frac{u_1^2 + u_2^2 + u_3^2 + \ldots + u_N^2}{N}$$

or

$$N\overline{u^2} = u_1^2 + u_2^2 + u_3^2 + \ldots + u_N^2$$

Substituting this result in Eqn. (10.1), we get

$$P = \frac{Nm\overline{u^2}}{\ell^3} \qquad (10.2)$$

It can be shown by geometry that

$$c^2 = u^2 + \upsilon^2 + w^2$$

since u, v and w are components of c along the three orthogonal axes. This relation also holds for the mean square values, i.e.

$$\overline{c^2} = \overline{u^2} + \overline{\upsilon^2} + \overline{w^2}$$

Since the molecular distribution has been assumed to be isotropic, there is no preferential motion along any one edge of the cube. This means that the mean value of u^2, υ^2, w^2 are equal :

$$\overline{u^2} = \overline{\upsilon^2} = \overline{w^2}$$

so that

$$\overline{u^2} = \frac{\overline{c^2}}{3}$$

Substituting this result in Eqn. (10.2), we get

$$P = \frac{1}{3}\frac{Nm}{\ell^3}\overline{c^2}$$

But l^3 defines the volume V of the container or the volume of the gas. Hence, we get

$$PV = \frac{1}{3}Nm\overline{c^2} = \frac{1}{3}M\overline{c^2} \qquad (10.3)$$

Note that the left hand side has macroscopic properties i.e. pressure and volume and the right hand side has only microscopic properties i.e. mass and mean square speed of the molecules.

Eqn (10.3) can be re-written as

$$P = \frac{1}{3}\frac{Nm}{V}\overline{c^2}$$

If $\rho = \frac{mN}{V}$ is the density of the gas, we can write

$$P = \frac{1}{3}\rho\overline{c^2}$$

or

$$\overline{c^2} = \frac{3P}{\rho} \qquad (10.4)$$

If we denote the ratio *N*/V by number density *n*, Eqn. (10.3) can also be expressed as

$$P = \frac{1}{3}m\,n\,\overline{c^2} \qquad (10.5)$$

The following points about the above derivation should be noted:

(i) *From Eqn.* (10.4) *it is clear that the shape of the container does not play any role in kinetic theory; only volume is of significance. Instead of a cube we could have taken any other container. A cube only simplifies our calculations.*

(ii) We ignored the intermolecular collisions but these would not have affected the result, because, the average momentum of the molecules on striking the walls is unchanged by their collision; same is the cose when they collide with each other.

(iii) The mean square speed $\overline{c^2}$ *is not the same as the square of the mean speed.* This is illustrated by the following example.

Suppose we have five molecules and their speeds are 1, 2, 3, 4, 5 units, respectively. Then their mean speed is

$$\frac{1+2+3+4+5}{5} = 3 \text{ units}$$

Its square is 9 (nine).

On the other hand, the mean square speed is

$$\frac{1^2+2^2+3^2+4^2+5^2}{5} = \frac{55}{5} = 11$$

Thus we see that mean square speed is not the same as square of mean speed.

Example 10.1 : Calculate the pressure exerted by 10^{22} molecules of oxygen, each of mass 5×10^{-26} kg, in a hollow cube of side 10 cm where the average translational speed of molecule is 500 m s^{-1}.

Solution : Change in momentum $2m\,u = 2 \times (5 \times 10^{-26}\text{ kg}) \times (500\text{ m s}^{-1})$

$$= 5 \times 10^{-23}\text{ kg m s}^{-1}.$$

Time taken to make successive impacts on the same face is equal to the time spent in travelling a distance of 2×10 cm or 2×10^{-1} m. Hence

$$\text{Time} = \frac{2 \times 10^{-2}\text{ m}}{500\text{ ms}^{-1}} = 4 \times 10^{-4}\text{ s}$$

$$\therefore \text{ Rate of change of momentum} = \frac{5\times10^{-23}\text{kg ms}^{-1}}{4\times10^{-4}\text{s}} = 1.25 \times 10^{-19}\text{ N}$$

The force on the side due to one third molecules

and $$f = \frac{1}{3} \times 1.25 \times 10^{-19} \times 10^{22} = 416.7\text{ N}$$

$$\text{pressure} = \frac{\text{Force}}{\text{Area}} = \frac{417\text{ N}}{100\times10^{-4}\text{m}^2}$$

$$= 4.2 \times 10^{-4}\text{ N m}^{-2}$$

INTEXT QUESTIONS 10.1

1. (i) A gas fills a container of any size but a liquid does not. Why?

   (ii) Solids have more ordered structure than gases. Why?

2. What is an ideal gas?

3. How is pressure related to density of molecules?

4. What is meant by specific heat of a substance?
5. Define coefficient of cubical expansion.
6. A steel wire has a length of 2 m at 20°C. Its length becomes 2.01 m at 120°C. Calculate coefficient of linear expansion α of the material of wire.

10.3 KINETIC INTERPRETATION OF TEMPERATURE

From Eqn. (10.3) we recall that

$$PV = \frac{1}{3} m N \overline{c^2}$$

Also, for n moles of a gas, the equation of state is PV = n RT, where gas constant R is equal to 8.3 J mol^{-1} K^{-1}. On combining this result with the expression for pressure, we get

$$n R T = \frac{1}{3} m N \overline{c^2}$$

Multiplying both sides by $\frac{3}{2n}$ we have

$$\frac{3}{2} R T = \frac{1}{2} \frac{N m \overline{c^2}}{n} = \frac{1}{2} m N_A \overline{c^2}$$

where $\frac{N}{n} = N_A$ is Avogadro's number. It denotes the number of atoms or molecules in one mole of a substance. Its value is 6.023×10^{23} per gram mole. In terms of N_A, we can write

$$\frac{3}{2}\left(\frac{R}{N_A}\right) T = \frac{1}{2} m \overline{c^2}$$

But $\frac{1}{2} m \overline{c^2}$ is the mean kinetic energy of a molecule. Therefore, we can write

$$\frac{1}{2} m \overline{c^2} = \frac{3}{2}\left(\frac{R}{N_A}\right) T = \frac{3}{2} k T \qquad (10.6)$$

where

$$k = \frac{R}{N_A} \qquad (10.7)$$

is ***Boltzmamn constant.*** The value of k is 1.38×10^{-23} J K^{-1}.

In terms of k, the mean kinetic energy of a molecule of the gas is given as

$$\bar{\varepsilon} = \frac{1}{2} m \overline{c^2} = \frac{3}{2} k T \qquad (10.8)$$

Hence, kinetic energy of a gram mole of a gas is $\frac{3}{2}RT$

This relationship tells us that the kinetic energy of a molecule depends only on the absolute temperature T of the gas and it is quite independent of its mass. This fact is known as the ***kinetic interpretation of temperature.***

Clearly, at T = 0, the gas has no kinetic energy. In other words, all molecular motion ceases to exist at absolute zero and the molecules behave as if they are frozen in space. According to modern concepts, the energy of the system of electrons is not zero even at the absolute zero. The energy at absolute zero is known as *zero point energy*.

From Eqn.(10.5), we can write the expression for the square root of $\overline{c^2}$, called **root mean square speed :**

$$c_{rms} = \sqrt{\overline{c^2}} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3RT}{M}}$$

This expression shows that at any temperature T, the c_{rms} is inversely proportional to the square root of molar mass. It means that lighter molecule, on an average, move faster than heavier molecules. For example, the molar mass of oxygen is 16 times the molar mass of hydrogen. So according to kinetic theory, the hydrogen molecules should move 4 times faster then oxygen molecules. It is for this reason that lighter gases are in the above part of our atmosphere. This observed fact provided an early important evidence for the validity of kinetic theory.

10.4 DEDUCTION OF GAS LAWS FROM KINETIC THEORY

(i) Boyle's Law

We know that the pressure P exerted by a gas is given by Eqn. (11.3) :

$$PV = \frac{1}{3} M \overline{c^2}$$

When the temperature of a given mass of the gas is constant, the mean square speed is constant. Thus, both M and $\overline{c^2}$ on the right hand side of Eqn. (10.3) are constant. Thus, we can write

$$PV = \text{Constant} \qquad (10.9)$$

This is Boyle's law, which states that ***at constant temperature, the pressure of a given mass of a gas is inversely proportional to the volume of the gas.***

(ii) Charle's Law

From Eqn. (10.3) we know that

$$PV = \frac{1}{3} M \overline{c^2}$$

or

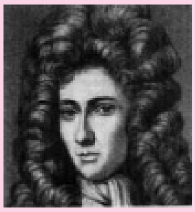
$$V = \frac{1}{3} \frac{M}{P} \overline{c^2}$$

i.e, $V \propto \overline{c^2}$, if M and P do not vary or M and P are constant. But $\overline{c^2} \propto T$

$$\therefore \quad V \propto T \qquad (10.10)$$

This is Charle's law : The volume of a given mass of a gas at constant pressure is directly proportional to temperature.

Robert Boyle
(1627 – 1691)



British experimentalist Robert Boyle is famous for his law relating the pressure and volume of a gas (PV = constant). Using a vacuum pump designed by Robert Hook, he demonstrated that sound does not travel in vacuum. He proved that air was required for burning and studied the elastic properties of air.

A founding fellow of Royal Society of London, Robert Boyle remained a bachalor throughout his life to pursue his scientific interests. Crater Boyle on the moon is named in his honour.

(iii) Gay Lussac's Law – According to kinetic theory of gases, for an ideal gas

$$P = \frac{1}{3} \frac{M}{V} \overline{c^2}$$

For a given mass (M constant) and at constant volume (V constant),

$$P \propto \overline{c^2}$$

But

$$\overline{c^2} \propto T$$

$$\therefore \quad P \propto T \qquad (10.11)$$

which is Gay Lussac's law. It states that **the pressure of a given mass of a gas is directly proportional to its absolute temperature T, if its volume remains constant.**

(iv) Avogadro's Law

Let us consider two different gases 1 and 2. Then from Eqn. (10.3), we recall that

$$P_1 V_1 = \frac{1}{3} m_1 N_1 \overline{c_1^2}$$

and $$P_2V_2 = \frac{1}{3} m_2 N_2 \overline{c_2^2}$$

If their pressure and volume are the same, we can write

$$P_1V_2 = P_2V_2$$

Hence $$\frac{1}{3} m_1 N_1 \overline{c_1^2} = \frac{1}{3} m_2 N_2 \overline{c_2^2}$$

Since the temperature is constant, their kinetic energies will be the same, i.e.

$$\frac{1}{2} m_1 \overline{c_1^2} = \frac{1}{2} m_2 \overline{c_2^2}$$

Using this result in the above expression, we get $N_1 = N_2$. (10.12)

That is*, equal volume of ideal gases under the same conditions of temperature and pressure contain equal number of molecules. This statement is Avogadro's Law.*

(v) Dalton's Law of Partial Pressure

Suppose we have a number of gases or vapours, which do not react chemically. Let their densities be ρ_1, ρ_2, ρ_3 ... and mean square speeds $\overline{c_1^2}, \overline{c_2^2}, \overline{c_3^2}$ ... respectively. We put these gases in the same enclosure. They all will have the same volume. Then the resultant pressure P will be given by

$$P = \frac{1}{3}\rho_1\overline{c_1^2} + \frac{1}{3}\rho_2\overline{c_2^2} + \frac{1}{3}\rho_3\overline{c_3^2} + ...$$

Here $\frac{1}{3}\rho_1\overline{c_1^2}, \frac{1}{3}\rho_2\overline{c_2^2}, \frac{1}{3}\rho_3\overline{c_3^2}$ ... signify individual (or partial) pressures of different gases or vapours. If we denote these by P_1, P_2, P_3, respectively we get

$$P = P_1 + P_2 + P_3 + ... \qquad (10.13)$$

In other words, ***the total pressure exerted by a gaseous mixture is the sum of the partial pressures that would be exerted, if individual gases occupied the space in turn. This is Dalton's law of partial pressures.***

(vi) Graham's law of diffusion of gases

Graham investigated the diffusion of gases through porous substances and found that ***the rate of diffusion of a gas through a porous partition is inversely proportional to the square root of its density. This is known as Graham's law of diffusion.***

On the basis of kinetic theory of gases, the rate of diffusion through a fine hole will be proportional to the average or root mean square velocity c_{rms}. From Eqn. (10.4) we recall that

$$\overline{c^2} = \frac{3P}{\rho}$$

or

$$\sqrt{\overline{c^2}} = c_{rms} = \sqrt{\frac{3P}{\rho}}$$

That is, the root mean square velocities of the molecules of two gases of densities ρ_1 and ρ_2 respectively at a pressure P are given by

$$(c_{rms})_1 = \sqrt{\frac{3P}{\rho_1}} \quad \text{and} \quad (c_{rms})_2 = \sqrt{\frac{3P}{\rho_2}}$$

Thus,

$$\frac{\text{Rate of diffusion of one gas}}{\text{Rate of diffusion of other gas}} = \frac{(c_{rms})_1}{(c_{rms})_2} = \sqrt{\frac{\rho_2}{\rho_1}} \qquad (10.14)$$

Thus, rate of diffusion of gases is inversely proportional to the square root of their densities at the same pressure, which is Graham's law of diffusion.

Example 10.2 : Calculate the root mean square speed of hydrogen molecules at 300 K. Take m(H_2) as 3.347×10^{-27} kg and k = 1.38×10^{-23} J mol^{-1} K^{-1}

Solution : We know that

$$c_{rms} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3 \times (1.38 \times 10^{-23}\,\text{J K}^{-1})\,(300\,\text{K})}{3.347 \times 10^{-27}\,\text{kg}}}$$

$$= 1927 \text{ m s}^{-1}$$

Example 10.3 : At what temperature will the root mean square velocity of hydrogen be double of its value at S.T.P., pressure being constant (STP = Standard temperature and pressure).

Solution : From Eqn. (10.8), we recall that

$$c_{rms} \propto \sqrt{T}$$

Let the rms velocity at S.T.P. be c_0.

If T K is the required temperature, the velocity $c = 2\,c_0$ as given in the problem

$$\therefore \quad \frac{c}{c_0} = \frac{2c_0}{c_0} = \sqrt{\frac{T}{T_0}}$$

Squaring both sides, we get

$$4 = \frac{T}{T_0}$$

or $$T = 4T_0$$

Since $T_0 = 273K$, we get

$$T = 4 \times 273K = 1092K = 819^0C$$

Example 10.4 : Calculate the average kinetic energy of a gas at 300 K. Given k = $1.38 \times 10^{-23} JK^{-1}$.

Solution : We know that

$$\frac{1}{2} M \overline{c^2} = \frac{3}{2} kT$$

Since $k = 1.38 \times 10^{-23}$ J K^{-1} and T = 300 K, we get

$$\therefore \quad \overline{E} = \frac{3}{2}(1.38 \times 10^{-23} J K^{-1})(300 K)$$

$$= 6.21 \times 10^{-21} J$$

INTEXT QUESTIONS 10.2

1. Five gas molecules chosen at random are found to have speeds 500 m s^{-1}, 600 m s^{-1}, 700 m s^{-1}, 800 m s^{-1}, and 900 m s^{-1}. Calculate their RMS speed.
2. If equal volumes of two non–reactive gases are mixed, what would be the resultant pressure of the mixture?
3. When we blow air in a balloon, its volume increases and the pressure inside is also more than when air was not blown in. Does this situation contradict Boyle's law?

10.4.1 Degrees of Freedom

Degrees of freedom of a system of particles are the number of independent ways in which the particles of the system can move.

Suppose you are driving along a road and several other roads are emanating from it towards left and right. You have the freedom to be on that road or to turn to the left or to the right you have two degrees of freedom. Now, say the road has a flyover at some point and you take the flyover route. Now, you do not have any freedom to turn left or right, which means that your freedom has get restricted. You can move only along the flyover and we say that your degree of freedom is '1'.

Refer to Fig. 10.4. A string is tied in a taut manner from one end A to other end B between two opposite walls of a room. An ant is moving on it. Then its degree of freedom is '1'.

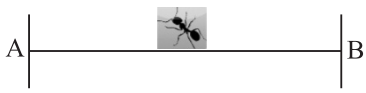


Fig. 10.4

Now suppose it falls on the floor of the room. Now, it can move along *x* or *y* direction independently. Hence its degrees of freedom is two. And if the ant has wings so that it can fly. Then it can move along *x*, *y* or *z* direction independently and its degree of freedom is '3'.

A monatomic molecule is a single point in space and like the winged ant in the above example has 3 degrees of freedom which are all translational. A diatomic molecule which is made up of two atoms, in addition to translator *y* motion can also rotate about two mutually perpendicular axes. Hence a diatomic molecule has (3 + 2 = 5) degrees of freedom: three translational and two rotational.

10.5 THE LAW OF EQUIPARTITION OF ENERGY

We now know that kinetic energy of a molecule of a gas is given by $\frac{1}{2}m\overline{c^2} = \frac{3}{2}\text{kT}$.

Since the motion of a molecule can be along *x*, *y*, and *z* directions equally probably, the average value of the components of velocity *c* (i.e., *u*, *v* and *w*) along the three directions should be equal. That is to say, for a molecule all the three directions are equivalent :

$$\overline{u} = \overline{v} = \overline{w}$$

and $$\overline{u^2} = \overline{v^2} = \overline{w^2} = \frac{1}{3}\overline{c^2}$$

Since $$c^2 = u^2 + v^2 + w^2$$

$$\overline{c^2} = \overline{u^2} + \overline{v^2} + \overline{w^2}$$

Multiplying throughout by $\frac{1}{2}m$, where m is the mass of a molecule, we have

$$\frac{1}{2}m\,\overline{u}^2 = \frac{1}{2}m\,\overline{v}^2 = \frac{1}{2}m\,\overline{w}^2$$

But $\frac{1}{2}m\,\overline{u}^2 = E =$ total mean kinetic energy of a molecule along x–axis. Therefore, $E_x = E_y = E_z$. But the total mean kinetic energy of a molecule is $\frac{3}{2}$ k T. Hence, we get an important result :

$$E_x = E_y = E_z = \frac{1}{2}\text{k T}$$

Since three velocity components u, v and w correspond to the three degree of freedom of the molecule, we can conclude that total kinetic energy of a dynamical system is equally divided among all its degrees of freedom and it is equal to $\frac{1}{2}$ k T for each degree of freedom. This is the law of equipartition of energy and was deduced by Ludwing Boltzmann. Let us apply this law for different types of gases.

So far we have been considering only translational motion. For a ***monoatomic molecule***, we have only translational motion because they are not capable of rotation (although they can spin about any one of the three mutually perpendicular axes if it is like a finite sphere). Hence, for one molecule of a ***monoatomic gas,*** total energy

$$E = \frac{3}{2}\text{k T} \qquad (10.15)$$

A ***diatomic molecule*** can be visualised as if two spheres are joined by a rigid rod. Such a molecule can rotate about any one of the three mutually perpendicular axes. However, the rotational inertia about an axis along the rigid rod is negligible compared to that about an axis perpendicular to the rod. It means that rotational energy consists of two terms such as $\frac{1}{2}\text{I}\,\omega_y^2$ and $\frac{1}{2}\text{I}\,\omega_z^2$.

Now the special description of the centre of mass of a diatomic gas molecules will require three coordinates. Thus, for a diatomic gas molecule, both rotational and translational motion are present but it has 5 degrees of freedom. Hence

$$E = 3\left(\frac{1}{2}\text{kT}\right) + 2\left(\frac{1}{2}\text{kT}\right)$$

$$= \frac{5}{2}\text{k T} \qquad (10.16)$$

Ludwing Boltzmann
(1844 – 1906)

Born and brought up in Vienna (Austria), Boltzmann completed his doctorate under the supervision of Josef Stefan in 1866. He also worked with Bunsen, Kirchhoff and Helmholtz. A very emotional person, he tried to commit suicide twice in his life and succeeded in his second attempt. The cause behind these attempts, people say, were his differences with Mach and Ostwald.

He is famous for his contributions to kinetic theory of gases, statistical mechanics and thermodynamics. Crater Bolzmann on moon is named in his memory and honour.

10.6 HEAT CAPACITIES OF GASES

We know that the temperature of a gas can be raised under different conditions of volume and pressure. For example, the volume or the pressure may be kept constant or both may be allowed to vary in some arbitrary manner. In each of these cases, the amount of thermal energy required to increase unit rise of temperature in unit mass is different. Hence, we say that a gas has two different heat capacities.

If we supply an amount of heat ΔQ to a gas to raise its temperature through ΔT, the heat capacity is defined as

$$\text{Heat capacity} = \frac{\Delta Q}{\Delta T}$$

The heat capacity of a body per unit mass of the body is termed as ***specific heat capacity*** of the substance and is usually denoted by c. Thus

$$\text{Specific heat capacity, c} = \frac{\text{heat capacity}}{m} \qquad (10.17)$$

Eqns. (10.16) and (10.17) may be combined to get

$$c = \frac{\Delta Q}{m\,\Delta T} \qquad (10.18)$$

Thus, specific heat capacity of a material is the heat required to raise the temperature of its unit mass by* 1 ºC *(or* 1 K*).

The SI unit of specific heat capacity is kilo calories per kilogram per kelvin (kcal $kg^{-1}K^{-1}$). It may also the expressed in joules per kg per K. For example the specific heat capacity of water is

$$1 \text{ kilo cal } kg^{-1}\,K^{-1} = 4.2 \times 10^3 \text{ J } kg^{-1}\,K^{-1}.$$

The above definition of specific heat capacity holds good for solids and liquids but not for gases, because it can vary with external conditions. In order to study the heat capacity of a gas, we keep the pressure or the volume of a gas constant. Consequently, we define two specific heat capacities :

(i) Specific heat at constant volume, denoted as c_V.

(ii) Specific heat at constant pressure, denoted as c_P.

(a) ***The specific heat capacity of a gas at constant volume*** (c_v) is defined as the amount of heat required to raise the temperature of unit mass of a gas through 1K, when its volume is kept constant :

$$c_v = \left(\frac{\Delta Q}{\Delta T}\right)_V \tag{10.19}$$

(b) ***The specific heat capacity of a gas at constant pressure*** (c_P) is defined as the amount of heat required to raise the temperature of unit mass of a gas through 1K when its pressure is kept constant.

$$c_p = \left(\frac{\Delta Q}{\Delta T}\right)_P \tag{10.20}$$

When 1 mole of a gas is considered, we define ***molar heat capacity.***

We know that when pressure is kept constant, the volume of the gas increases. Hence in the second case note that the heat required to raise the temperature of unit mass through 1 degree at constant pressure is used up in two parts :

(i) heat required to do external work to produce a change in volume of the gas, and

(ii) heat required to raise the temperature of the gas through one degree (c_v).

This means the specific heat capacity of a gas at constant pressure is greater than its specific heat capacity at constant volume by an amount which is thermal equivalent of the work done in expending the gas against external pressure. That is

$$c_p = W + c_v \tag{10.21}$$

10.7 RELATION BETWEEN C_P AND C_V

Let us consider one mole of an ideal gas enclosed in a cylinder fitted with a frictionless movable piston (Fig. 10.5). Since the gas has been assumed to be ideal (perfect), there is no intermolecular force between its molecules. When such a gas expands, some work is done in overcoming internal pressure.

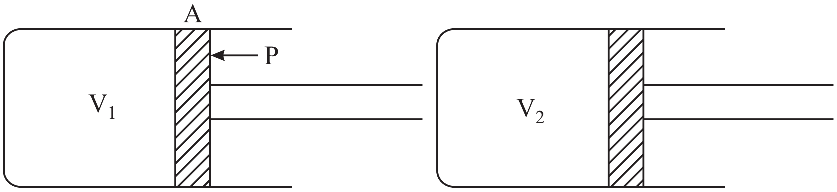


Fig. 10.5 : Gas heated at constant pressure

Let P be the external pressure and A be the cross sectional area of the piston. The force acting on the piston = P × A**.** Now suppose that the gas is heated at constant pressure by 1K and as a result, the piston moves outward through a distance x, as shown in Fig. 10.5. Let V_1 be the initial volume of the gas and V_2 be the volume after heating. Therefore, the work W done by the gas in pushing the piston through a distance x, against external pressure P is given by

$$W = P \times A \times x$$

$$= P \times (\text{Increase in volume})$$

$$= P\,(V_2 - V_1)$$

We know from Eqn. (10.22) that $c_p - c_v$ = Work done (W) against the external pressure in raising the temperature of 1 mol of a gas through 1 K, i.e.

$$c_p - c_v = P\,(V_2 - V_1) \qquad (10.22)$$

Now applying perfect gas equation to these two stages of the gas i.e. before and after heating, we have

$$PV_1 = RT \qquad (10.23)$$

$$PV_2 = R\,(T + 1) \qquad (10.24)$$

Substracting Eqn. (10.23) from Eqn.(10.24), we get

$$P\,(V_2 - V_1) = R \qquad (10.25)$$

Hence from Eqns. (10.19) and (10.22) we get

$$c_p - c_v = R \qquad (10.26)$$

where R is in J mol^{-1} K^{-1}

Converting joules into calories, we can write

$$c_p - c_v = \frac{R}{J} \qquad (10.27)$$

where J = 4.18 cal is the mechanical equivalent of heat.

Example 10.5 : Calculate the value of c_p and c_v for a monoatomic, diatomic and triatomic gas molecules.

Solution : We know that the average KE for 1 mol of a gas is given as

$$E = \frac{3}{2} R T$$

Now c_v is defined as the heat required to raise the temperature of 1 mole of a gas at constant volume by one degree i.e. if E_T denotes total energy of gas at T K and E_{T+1} signifies total energy of gas at (T + 1) K, then $c_v = E_{T+1} - E_T$.

(i) We know that for monoatomic gas, total energy = $\frac{3}{2}$ R T

$$\therefore \quad \text{monoatomic gas } c_V = \frac{3}{2} R (T+1) - \frac{3}{2} R T = \frac{3}{2} R.$$

Hence

$$c_p = c_V + R = \frac{3}{2} R + R = \frac{5}{2} R.$$

(ii) For diatomic gases, total energy = $\frac{5}{2}$ R T

$$\therefore \quad c_V = \frac{5}{2} R (T+1) - R \frac{5}{2} R T = \frac{5}{2} R$$

$$c_p = c_V + R = \frac{5}{2} R + R = \frac{7}{2} R.$$

(iii) You should now find out c_V and c_p for triatomic gas.

INTEXT QUESTIONS 10.3

1. What is the total energy of a nitrogen molecule?
2. Calculate the value of c_p and c_V for nitrogen (given, R = 8.3J mol^{-1} K^{-1}).
3. Why do gases have two types of specific heat capacities?

Brownian Motion and Mean Free Path

Scottish botanist Robert Brown, while observing the pollen grains of a flower suspended in water, under his microscope, found that the pollen grains were tumbling and tossing and moving about in a zigzag random fashion. The random motion of pollen grains, was initially attributed to live objects. But when motion of pollens of dead plants and particles of mica and stone were seen to exhibit the same behaviour, it became clear that the motion of the particles, now called **Brownian motion,** was caused by unbalanced forces due to impacts of water molecules. Brownian motion provided a direct evidence in favour of kinetic theory of matter. The Brownian displacement was found to depend on.

(i) Size of the particles of the suspension – smaller the particles, more the chances of inbalanced impacts and more pronounced the Brownian motion.

(ii) The Brownian motion also increases with the increase in the temperature and decreases with the viscosity of the medium.

Due to mutual collisions, the molecules of a fluid also move on zig-zag paths. The average distance between two successive collisions of the molecules is called mean free path. The mean free path of a molecule is given by

$$\sigma = \frac{1}{\sqrt{2}\, n \pi d^2}$$

where n is the number density and d the diameter of the molecules.

WHAT YOU HAVE LEARNT

- The specific heat of a substance is defined as the amount of heat required to raise the temperature of its unit mass through 1°C or 1 K.
- According to principle of calorimetry: Heat lost = Heat gained
- Kinetic theory assumes the existence of atoms and molecules of a gas and applies the law of mechanics to large number of them using averaging technique.
- Kinetic theory relates macroscopic properties to microscopic properties of individual molecules.
- The pressure of a gas is the average impact of its molecules on the unit area of the walls of the container.
- Kinetic energy of a molecule depends on the absolute temperature T and is independent of its mass.
- At absolute zero of temperature, the kinetic energy of a gas is zero and molecular motion ceases to exist.
- Gas law can be derived on the basis of kinetic theory. This provided an early evidence in favour of kinetic theory.
- Depending on whether the volume or the pressure is kept constant, the amount of heat required to raise the temperature of unit mass of a gas by 1ºC is different. Hence there are two specific heats of gas :

i) Specific heat capacity at constant volume (c_V)

ii) Specific heat capacity at constant pressure (c_p)

These are related as $c_p = W + c_V$

$$c_p - c_V = \frac{R}{J}$$

- The degrees of freedom of a system of particles are the number of independent ways in which the particles of the system can move.
- The law of equipartition of the energy states that the total kinetic energy of a dynamical system is distributed equally among all its degrees of freedom and it is equal to $\frac{1}{2}$ k T per degree of freedom.
- Total energy for a molecule of (i) a monatomic gas is $\frac{3}{2}$ k T, (ii) a diatomic gas is $\frac{5}{2}$, and (iii) a triatomic gas is 3 k T.

TERMINAL EXERCISE

1. Can we use Boyle's law to compare two different ideal gases?
2. What will be the velocity and kinetic energy of the molecules of a substance at absolute zero temperature?
3. If the absolute temperature of a gas is raised four times, what will happen to its kinetic energy, root-mean square velocity and pressure?
4. What should be the ratio of the average velocities of hydrogen molecules (molecular mass = 2) and that of oxygen molecules (molecular mass = 32) in a mixture of two gases to have the same kinetic energy per molecule?
5. If three molecules have velocities 0.5, 1 and 2 km s^{-1} respectively, calculate the ratio between their root mean square and average speeds.
6. Explain what is meant by the root-mean square velocity of the molecules of a gas. Use the concepts of kinetic theory of gases to derives an expression for the root-mean square velocity of the molecules in term of pressure and density of the gas.
7. i) Calculate the average translational kinetic energy of a neon atom at 25 ^{0}C.

   ii) At what temperature does the average energy have half this value?

8. A container of volume of 50 cm^3 contains hydrogen at a pressure of 1.0 Pa and at a temperature of 27 ^{0}C. Calculate (a) the number of molecules of the gas in the container, and (b)their root-mean square speed.

   ($R= 8.3\ J\ mol^{-1}\ K^{-1}$, $N = 6 \times 10^{23}\ mol^{-1}$. Mass of 1 mole of hydrogen molecule = 20×10^{-3} kg mol^{-1}).

9. A closed container contains hydrogen which exerts pressure of 20.0 mm Hg at a temperature of 50 K.

   (a) At what temperature will it exert pressure of 180 mm Hg?

   (b) If the root-mean square velocity of the hydrogen molecules at 10.0 K is 800 m s^{-1}, what will be their root-mean square velocity at this new temperature?

10. State the assumptions of kinetic theory of gases.

11. Find an expression for the pressure of a gas.

12. Deduce Boyle's law and Charle's law from kinetic the theory of gases.

13. What is the interpretation of temperature on the basis of kinetic theory of gases?.

14. What is Avagardo's law? How can it be deduced from kinetic theory of gases

15. Calculate the root-mean square of the molecules of hydrogen at 0°C and at 100 ^{0}C (Density of hydrogen at 0°C and 760 mm of mercury pressure = 0.09 kg m^{-3}).

16. Calculate the pressure in mm of mercury exerted by hydrogen gas if the number of molecules per m^3 is 6.8×10^{24} and the root-mean square speed of the molecules is 1.90×10 m s^{-1}. Avogadro's number 6.02×10^{23} and molecular weight of hydrogen = 2.02).

17. Define specific heat of a gas at constant pressure. Derive the relationship between c_p and c_V.

18. Define specific heat of gases at constant volume. Prove that for a triatomic gas $c_V = 3R$

19. Calculate c_p and c_V for argon. Given R = 8.3 J mol^{-1} K^{-1}.

ANSWERS TO INTEXT QUESTIONS

10.1

1. (i) Because in a gas the cohesive force between the molecules are extremely small as compared to the molecules in a liquid.

   (ii) Because the molecules in a solid are closely packed. The bonds between the molecules are stronger giving a ordered structure.

2. The gas which follows the kinetic theory of molecules is called as an ideal gas.

3. $P = \frac{1}{3}\rho\overline{c^2}$

4. The specific heat of a substance is the amount of heat required to raise the temperature of its unit mass through 1°C or 1K.

5. The coefficient of cubical expansion is defined as the change in volume per unit original volume per degree rise in temperature.

6. 0.00005 °C^{-1}

10.2

1. Average speed $\overline{c}$

$$= \frac{500+600+700+800+900}{5}$$

$= 700$ m s^{-1}

Average value of $\overline{c^2}$

$$= \frac{500^2+600^2+700^2+800^2+900^2}{5}$$

$= 510{,}000$ m^2 s^{-2}

$c_{rms} = \sqrt{\overline{c^2}} = \sqrt{510{,}000} = 714$ m s^{-1}

c_{rms} and $\overline{c}$ are not same

2. The resultant pressure of the mixture will be the sum of the pressure of gases 1 and 2 respectively i.e. $P = P_1 + P_2$.

3. Boyle's law is not applicable.

10.3

1. For each degree of freedom, energy = $\frac{1}{2}$k T

∴ for 5 degrees of freedom for a molecule of nitrogen, total energy = $\frac{5}{2}$k T.

2. c_V for a diatomic molecule = $\frac{5}{2}$ R

$c_V = \frac{5}{2} \times 8.3$ J mol^{-1} K^{-1} = 20.75 J mol^{-1} K^{-1}.

$c_p = c_V + R$ = 29.05 J mol^{-1} C^{-1}.

Answers to Terminal Problem

2. zero
3. becomes 4 times, doubles, becomes 4 time.
4. 4 : 1
5. 2
7. 6.18×10^{-21} m s^{-1}, – 124 ºC
8. 12×10^{20}, 7.9×10^{11} m s^{-1}
9. 2634ºC, 2560 m s^{-1}
15. 1800 m s^{-1}, 2088 m s^{-1}
16. 3.97×10^{3} N m^{-2}
17. 12.45 J mol^{-1} K^{-1}, 20.75 J mol^{-1} K^{-1}.

312en11

11

THERMODYNAMICS

You are familiar with the sensation of hotness and coldness. When you rub your hands together, you get the feeling of warmth. You will agree that the cause of heating in this case is mechanical work.This suggests that there is a relationship between mechanical work and thermal effect. A study of phenomena involving thermal energy transfer between bodies at different temperatures forms the subject matter of thermodynamics, which is a phenomenological science based on experience. A quantitative description of thermal phenomena requires a definition of temperature, thermal energy and internal energy. And the laws of thermodynamics provide relationship between the direction of flow of heat, work done on/by a system and the internal energy of a system.

In this lesson you will learn three laws of thermodynamics : the zeroth law, the first law and the second law of thermodynamics. These laws are based on experience and need no proof. As such, the zeroth, first and second law introduce the concept of temperature, internal energy and entropy, respectively. While the first law is essentially the law of conservation of energy for a thermodynamic system, the second law deals with conversion of heat into work and vice versa. You will also learn that Carnot's engine has maximum efficiency for conversion of heat into work.

OBJECTIVES

After studying this lesson, you should be able to :

- *draw indicator diagrams for different thermodynamic processes and show that the area under the indicator diagram represents the work done in the process;*
- *explain thermodynamic equilibrium and state the Zeroth law of thermodynamics;*

- *explain the concept of internal energy of a system and state first law of thermodynamics;*
- *apply first law of thermodynamics to simple systems and state its limitations;*
- *define triple point;*
- *state the second law of thermodynamics in different forms; and*
- *describe Carnot cycle and calculate its efficiency.*

11.1 CONCEPT OF HEAT AND TEMPERATURE

11.1.1 Heat

Energy has pervaded all facets of human activity ever since man lived in caves. In its manifestation as heat, energy is intimate to our existence. The energy that cooks our food, lights our houses, runs trains and aeroplanes originates in heat released in burning of wood, coal, gas or oil. You may like to ask : What is heat? To discover answer to this question, let us consider as to what happens when we inflate the tyre of a bicycle using a pump. If you touch the nozzle, you will observe that pump gets hot. Similarly, when you rub you hands together, you get the feeling of warmth. You will agree that in these processes heating is not caused by putting a flame or something hot underneath the pump or the hand. Instead, heat is arising as a result of mechanical work that is done in compressing the gas in the pump and forcing the hand to move against friction. These examples, in fact, indicate a relation between mechanical work and thermal effect.

We know from experience that a glass of ice cold water left to itself on a hot summer day eventually warms up. But a cup of hot coffee placed on the table cools down. It means that energy has been exchanged between the sysem – water or coffee – and its surrounding medium. This energy transfer continues till thermal equilibrium is reached. That is until both – the system and the suroundings – are as the same temperature. It also shows that the direction of energy transfer is always from the body at high temprature to a body at lower temperature. You may now ask : In what form is energy being transferred? In the above examples, energy is said to be transfered in the form of heat. So we can say that heat is the *form of energy transferred between two (or more) systems or a system and its surroundings because of temperature difference.*

You may now ask. What is the nature of this form of energy? The answer to this question was provided by Joule through his work on the equivalence of heat and mechanical work : Mechanical motion of molecules making up the system is associated with heat.

The unit of heat is calorie. One calorie is defined as the quantity of heat energy required to raise the temperature of 1 gram of water from 14.5°C to 15.5°C. It is denoted as cal.

Kilocalorie (k cal) is the larger unit of heat energy :

$$1 \text{ kcal} = 10^3 \text{ cal.}$$

Also $$1 \text{ cal} = 4.18 \text{ J}$$

11.1.2 Concept of Temperature

While studying the nature of heat, you learnt that energy exchange between a glass of cold water and its surroundings continues until thermal equilibrium was reached. All bodies in thermal equilibrium have a common property, called temperature, whose value is same for all of them. Thus, we can say that temperature of a body is the property which determines whether or not it is in thermal equilibrium with other bodies.

11.1.3 Thermodynamic Terms

(i) **Thermodynamic system :** A thermodynamic system refers to a definite quantity of matter which is considered unique and separated from everything else, which can influence it. Every system is enclosed by an arbitrary surface, which is called its boundary. The boundary may enclose a solid, a liquid or a gas. It may be real or imaginary, either at rest or in motion and may change its size and shape. The region of space outside the boundary of a system constitutes its surroundings.

   (a) **Open System :** It is a system which can exchange mass and energy with the surroundings. A water heater is an open system.

   (b) **Closed system :** It is a system which can exchange energy but not mass with the surroundings. A gas enclosed in a cylinder fitted with a piston is a closed system.

   (c) **Isolated system :** It is a system which can exchange neither mass nor energy with the surrounding. A filled thermos flank is an ideal example of an isolated system.

(ii) **Thermodynamic Variables or Coordinates :** In module–1, we have studied the motion of a body (or a system) in terms of its mass, position and velocity. To describe a thermodynamic system, we use its physical properties such on temperature (T), pressure (P), and volume (V). These are called thermodynamic variables.

(iii) **Indicator diagram :** You have learnt about displacement–time and velocity–time graphs in lesson 2. To study a thermodynamic system, we use a pressure-volume graph. This graph indicates how pressure (P) of a system varies with its volume (V) during a thermodynamic process and is known as an indicator diagram.

The indicator diagram can be used to obtain an expression for the work done. It is equal to the area under the P-V diagram (Fig. 11.1). Suppose that pressure is P at the start of a very small expansion ΔV. Then, work done by the system.

$$\Delta W = P\,\Delta V \qquad (11.1)$$

$$= \text{Area of a shaded strip ABCD}$$

Now total work done by the system when it expands from V_1 to V_2 = Area of $P_1P_2V_2V_1P_1$ Note that the area depends upon the shape of the indicator diagram.

The indicator diagram is widely used in calculating the work done in the process of expansion or compression. It is found more useful in processes where relationship between P and V is not known. The work done on the system increases its energy and work done by the system reduces it. For this reason, work done on the system is taken as negative. You must note that the area enclosed by an isotherm (plot of p versus V at constant temperature) depends on its shape. We may conclude that work done by or on a system depends on the path. That is, work does not depend on the initial and final states.

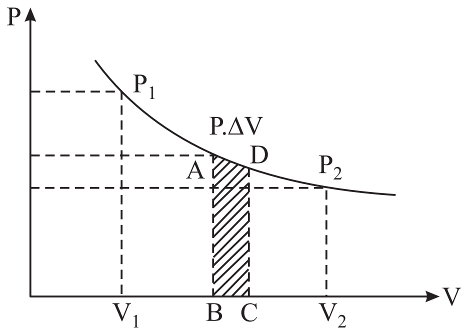


Fig. 11.1 : Indicater Diagram

11.2 THERMODYNAMIC EQUILIBRIUM

Imagine that a container is filled with a liquid (water, tea, milk, coffee) at 60° C. If it is left to itself, it is common experience that after some time, the liquid attains the room temperature. We then say that water in the container has attained thermal equilibrium with the surroundings.

If within the system, there are variations in pressure or elastic stress, then parts of the system may undergo some changes. However, these changes cease ultimately, and no unbalanced force will act on the system. Then we say that it is in mechanical equilibrium. Do you know that our earth bulged out at the equator in the process of attaining mechanical equilibrium in its formation from a molten state?

If a system has components which react chemically, after some time, all possible chemical reactions will cease to occur. Then the system is said to be in chemical equilibrium.

A system which exhibits thermal, mechanical and chemical equilibria is said to be in thermodynamic equilibrium. The macroscopic properties of a system in this state do not change with time.

11.2.1 Thermodynamic Process

If any of the thermodynamic variables of a system change while going from one equilibrium state to another, the system is said to execute a thermodynamic process. For example, the expansion of a gas in a cylinder at constant pressure due to heating is a thermodynamic process. A graphical representation of a thermodynamic process is called a path.

Now we will consider different types of thermodynamic processes.

(i) **Reversible process :** If a process is executed so that all intermediate stages between the inital and final states are equilibrium states and the process can be executed back along the same equilibrium states from its final state to its initial state, it is called reversible process. A reversible process is executed very slowly and in a controlled manner. Consider the following examples :

- Take a piece of ice in a beaker and heat it. You will see that it changes to water. If you remove the same quantity of heat of water by keeping it inside a refrigerator, it again changes to ice (initial state).
- Consider a spring supported at one end. Put some masses at its free end one by one. You will note that the spring elongates (increases in length). Now remove the masses one by one. You will see that spring retraces its initial positions. Hence it is a reversible process.

  As such, a reversible process can only be idealised and never achieved in practice.

(ii) **Irreversible process :** A process which cannot be retraced along the same equilibrium state from final to the initial state is called irreversible process.

All natural process are irrerersible. For example, heat poduced during friction, sugar dissolved in water, or rusting of iron in the air. It means that for irrerersible process, the intermediate states are not equilibrium states and hence such process can not be represented by a path. Does this mean that we can not analyse an irrerersible process? To do so, we use quasi-static process, which is infinitesimally close to the equilibrium state.

(iii) **Isothermal process :** A thermodynamic process that occurs at constant temperature is an isothermal process. The expansion and compression of a perfect gas in a cylinder made of perfectly conducting walls are isothermal processes. The change in pressure or volume is carried out very slowly so that any heat developed is transferred into the surroundings and the temperature of the system remains constant. The thermal equilibrium is always maintained. In such a process, ΔQ, ΔU and ΔW are finite.

(iv) **Adiabatic process :** A thermodynamic process in which no exchange of thermal energy occurs is an adiabatic process. For example, the expansion and compression of a perfect gas in a cylinder made of perfect insulating walls. The system is isolated from the surroundings. Neither any amount of heat leaves the system nor enters it from the surroundings. In this process, therefore $\Delta Q = 0$ and $\Delta U = -\Delta W$.

The change in the internal energy of the system is equal to the work done on the system. When the gas is compressed, work is done on the system. So, ΔU becomes positive and the internal energy of the system increases. When the gas expands, work is done by the system. It is taken as positive and ΔU becomes negative. The internal energy of the system decreases.

(v) **Isobaric process :** A thermodynamic process that occurs at constant pressure is an isobaric process. Heating of water under atmospheric pressure is an isobaric process.

(vi) **Isochoric process :** A thermodynamic process that occurs at constant volume is an isochoric process. For example, heating of a gas in a vessel of constant volume is an isochoric process. In this process, volume of the gas remains constant so that no work is done, i.e. $\Delta W = 0$. We therefore get $\Delta Q = \Delta U$.

In a **Cyclic Process** the system returns back to its initial state. It means that there is no change in the internal energy of the system. $\Delta U = 0$.

$$\therefore \Delta Q = \Delta W.$$

11.2.2 Zeroth Law of Thermodynamics

Let us consider three metal blocks A, B and C. Suppose block A is in thermal equilibrium with block B. Further suppose that block A is also in thermal equilibrium with block C. It means the temperature of the block A is equal to the temperature of block B as well as of block C. It follows that the temperatures of blocks B and C are equal. We summarize this result in the statement known as Zeroth Law of Thermodynamics :

If two bodies or systems A and B are separately in thermal equilibrium with a third body C, then A and B are in thermal equilibrium with each other.

Phase Change and Phase Diagram

You have learnt that at STP, matter exists in three states : solid, liquid and gas. ***The different states of matter are called its phases.*** For example, ice (solid), water (liquid) and steam (gas) are three phases of water. We can discuss these three phases using a three dimensional diagram drawn in pressure (P), temperature (T) and volume (V). It is difficult to draw three dimensional diagram. Thus, we discuss the three phases of matter by drawing a pressure-temperature diagram. This is called ***phase diagram.***

Notes

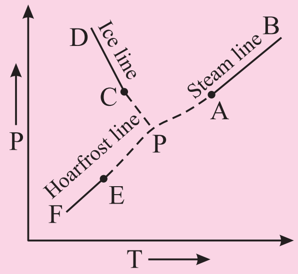


Fig. 11.2: Phase diagram of water

Refer to Fig. 11.2, which shows phase diagram of water. You can see three curves CD; AB and EF. Curve CD shows the variation of melting point of ice with pressure. It is known as a ***fusion curve***. Curve AB shows variation of boiling point of water with pressure. It is known as ***vaporization curve***. Curve EF shows change of ice directly to steam. It is known as a ***sublimation curve.*** This curve is also known as ***Hoarfrost Line.***

If you extend the curve AB, CD and EF (as shown in the figure with dotted lines), they meet at point P. This point is called ***triple point.*** At triple point, all three phases co-exist.

When we heat a solid, its temperature increases till it reaches a temperature at which it starts melting. This temperature is called **melting point** of the solid. During this change of state, we supply heat continuously but the temperature does not rise. The heat required to completely change unit mass of a solid into its corresponding liquid state at its melting point is called **latent heat of fusion of the solid.**

On heating a liquid, its temperature also rises till its **boiling point** is reached. At the boiling point, the heat we supply is used up in converting the liquid into its gaseous state. The amount of heat required to convert unit mass of liquid in its gaseous state at constant temperature is called **latent heat** of vaporization of the liquid.

11.2.3 Triple Point of Water

Triple point of a pure substance is a very stable state signified by precisely constant temperature and pressure values. For this reason, in kelvin's scale of thermometry, triple point of water is taken as the upper fixed point.

On increasing pressure, the melting point of a solid decreases and boiling point of the liquid increases. It is possible that by adjusting temperature and pressure, we can obtain all the three states of matter to co-exist simultaneously. These values of temperature and pressure signify the **triple point.**

Notes

INTEXT QUESTIONS 11.1

1. Fill in the blanks

   (i) Zeroth law of thermodynamics provides the basis for the concept of ....

   (ii) If a system A is in thermal equilibrium with a system B and B is in thermal equilibrium with another system C, then system A will also be in thermal equilibrium with system...............

   (iii) The unit of heat is

2. Fig. 11.3 is an indicator diagram of a thermodynamic process. Calculate the work done by the system in the process :

   (a) along the path ABC from A to C

   (b) If the system is returned from C to A along the same path, how much work is done by the system.

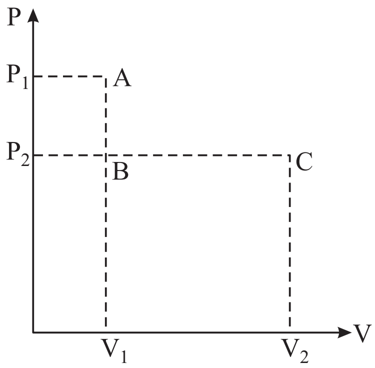


Fig. 11.3

3. Fill in the blanks.

   (i) A reversible process is that which can be .... in the opposite direction from its final state to its initial state.

   (ii) An .... process is that which cannot be retraced along the same equilibrium states from final state to the initial state.

4. State the basic difference between isothermal and adiabatic processes.

5. State one characteristic of the triple point.

11.3 INTERNAL ENERGY OF A SYSTEM

Have you ever thought about the energy which is released when water freezes into ice ? Don't you think that there is some kind of energy stored in water. This energy is released when water changes into ice. This stored energy is called the **internal energy**. On the basis of kinetic theory of matter, we can discuss the concept of internal energy as sum of the energies of individual components/ constituents. This includes kinetic energy due to their random motion and their potential energy due to interactions amongst them. Let us now discuss these.

(a) Internal kinetic energy : As you now know, according to kinetic theory, matter is made up of a large number of molecules. These molecules are in a state of constant rapid motion and hence possess kinetic energy. The **total**

kinetic energy of the molecules constitutes the internal kinetic energy of the body.

(b) Internal potential energy : The energy arising due to the inter-molecular forces is called the internal potential energy.

The internal energy of a metallic rod is made up of the kinetic energies of conduction electrons, potential energies of atoms of the metal and the vibrational energies about their equilibrium positions. The energy of the system may be increased by causing its molecules to move faster (gain in kinetic energy by adding thermal energy). It can also be increased by causing the molecules to move against inter-molecular forces, i.e., by doing work on it. **Internal energy** is denoted by the letter U.

Internal energy of a system = Kinetic energy of molecules + Potential energy of molecules

Let us consider an isolated thermodynamic system subjected to an external force. Suppose W amount of work is done on the system in going from initial state *i* to final state *f* adiabatically. Let U_i and U_f be internal energies of the system in its initial and final states respectively. Since work is done on the system, internal energy of final state will be higher than that of the initial state.

According to the law of conservation of energy, we can write

$$U_i - U_f = -W$$

Negative sign signifies that work is done on the system.

We may point out here that unlike work, internal energy depends on the initial and final states, irrespective of the path followed. We express this fact by saying that U is a function of state and depends only on state variables P, V, and T. Note that if some work is done by the system, its internal energy will decrease.

11.4 FIRST LAW OF THERMODYNAMICS

You now know that the zeroth law of thermodynamics tells us about thermal equilibrium among different systems characterised by same temperature. However, this law does not tell us anything about the non-equilibrium state. Let us consider two examples : (i) Two systems at different temperatures are put in thermal contact and (ii) Mechanical rubbing between two systems. In both cases, change in their temperatures occurs but it cannot be explained by the Zeroth law. To explain such processes, the first law of thermodynamics was postulated.

The first law of thermodynamics is, in fact, the law of conservation of energy for a thermodynamic system. It states that change in internal energy of a system during a thermodynamic process is equal to the sum of the heat given to it and the work done on it.

Suppose that ΔQ amount of heat is given to the system and $-\Delta W$ work is done on the system. Then increase in internal energy of the system, ΔU, according to the first law of thermodynamics is given by

$$\Delta U = \Delta Q - \Delta W \qquad (11.3\text{ a})$$

This is the mathematical form of the first law of thermodynamics. Here ΔQ, ΔU and ΔW all are in SI units.

The first law of thermodynamics can also be written as

$$\Delta Q = \Delta U + \Delta W \qquad (11.3\text{ b})$$

The signs of ΔQ, ΔU and ΔW are known from the following sign conventions :

1. Work done (ΔW) by a system is taken as positive whereas the work done on a system is taken as negative. The work is positive when a system expands. When a system is compressed, the volume decreases, the work done is negative. The work done does not depend on the initial and final thermodynamic states; it depends on the path followed to bring a change.
2. Heat gained by (added to) a system is taken as positive, whereas heat lost by a system is taken as negative.
3. The increase in internal energy is taken as positive and a decrease in internal energy is taken as negative.

If a system is taken from state 1 to state 2, it is found that both ΔQ and ΔW depend on the path of transformation. However, the difference ($\Delta Q - \Delta W$) which represents ΔU, remains the same for all paths of transformations.

We therefore say that the change in internal energy ΔU of a system does not depend on the path of the thermodynamic transformations.

11.4.1 Limitations of the First Law of Thermodynamics

The first law of thermodynamics asserts the equivalence of heat and other forms of energy. This equivalence makes the world around us work. The electrical energy that lights our houses, operates machines and runs trains originates in heat released in burning of fossil or nuclear fuel. In a sense, it is universal. It explains the fall in temperature with height; the adiabatic lapse rate in upper atmosphere. Its applications to flow process and chemical reations are also very interesting. However, consider the following processes :

- You know that heat always flows from a hot body to a cold body. But first law of thermodynamics does not prohibit flow of heat from a cold body to a hot body. It means that this law fails to indicate the direction of heat flow.
- You know that when a bullet strikes a target, the kinetic energy of the bullet is converted into heat. This law does not indicate as to why heat developed

in the target cannot be changed into the kinetic energy of bullet to make it fly. It means that this law fails to provide the conditions under which heat can be changed into work. Moreover, it has obvious limitations in indicating the extent to which heat can be converted into work.

Now take a pause and answer the following questions :

INTEXT QUESTIONS 11.2

1. Fill in the blanks
   (i) The total of kinetic energy and potential energy of molecules of a system is called its ....
   (ii) Work done = – W indicates that work is done .... the system.
2. The first law of thermodynamics states that ....

11.5 SECOND LAW OF THERMODYNAMICS

You now know that the first law of thermodynamics has inherent limitations in respect of the direction of flow of heat and the extent of convertibility of heat into work. So a question may arise in your mind : Can heat be wholly converted into work? Under what conditions this conversion occurs? The answers of such questions are contained in the postulate of *Second law of thermodynamics.* The second law of thermodynamics is stated in several ways. However, here you will study Kelvin-Planck and Clausius statements of second law of thermodynamics.

The Kelvin-Planck's statement is based on the experience about the performance of heat engines. (Heat engine is discussed in next section.) In a heat engine, the working substance extracts heat from the source (hot body), converts a part of it into work and rejects the rest of heat to the sink (environment). There is no engine which converts the whole heat into work, without rejecting some heat to the sink. These observations led Kelvin and Planck to state the second law of thermodynamics as

It is impossible for any system to absorb heat from a reservoir at a fixed temperature and convert whole of it into work.

Clausius statement of second law of thermodynamics is based on the performance of a refrigerator. A refrigerator is a heat engine working in the opposite direction. It transfers heat from a colder body to a hotter body when external work is done on it. Here concept of external work done on the system is important. To do this external work, supply of energy from some external source is a must. These observations led Clausius to state the second law of thermodynamics in the following form.

It is impossible for any process to have as its sole result to transfer heat from a colder body to a hotter body without any external work.

Thus, the second law of thermodynamics plays a unique role for practical devices like heat engine and refrigerator.

11.5.1 Carnot Cycle

You must have noticed that when water is boiled in a vessel having a lid, the steam generated inside throws off the lid. This shows that high pressure steam can be made to do useful work. ***A device which can convert heat into work is called a heat engine.*** Modern engines which we use in our daily life are based on the principle of heat engine. These may be categorised in three types : steam engine, internal combustion engine and gas turbine. However, their working can be understood in terms of Carnot's reversible engine. Let us learn about it now.

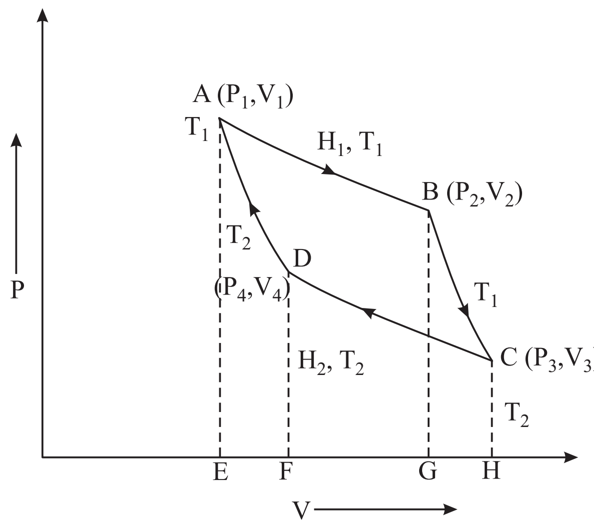


Fig. 11.4 : Indicator diagram of Carnot cycle

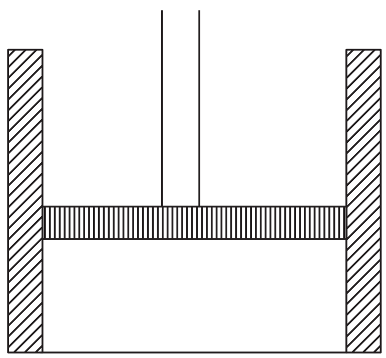


Fig. 11.5 : The cylinder with working substance

In Carnot cycle, the working substance is subjected to four operations : (a) isothermal expansion, (b) adiabatic expansion, (c) isothermal compression and (d) adiabatic compression. Such a cycle is represented on the P-V diagram in Fig. 11.4. To describe four operations of Carnot's cycle, let us fill one gram. mol. of the working substance in the cylinder (Fig. 11.5). Original condition of the substance is represented by point A on the indicator diagram. At this point, the substance is at temperature T_1, pressure P_1 and volume V_1.

(a) **Isothermal expansion :** The cylinder is put in thermal contact with the source and allowed to expand. The volume of the working substance increases to V_2. Thus working substance does work in raising the piston. In this way, the temperature of the working substance would tend to fall. But it is in thermal contact with the source. So it will absorb a quantity of heat H_1 from the source at temperature T_1. This is represented by the point B. At B, the values of pressure and volume are P_2 and V_2 respectively. On the indicator diagram (Fig. 11.4), you see that in going from A to B, temperature of the system remains constant and working substance expands. We call it ***isothermal expansion process***. H_1 is the amount of heat absorbed in the isothermal expansion process. Then, in accordance with the first law of thermodynamics, H_1 will be equal to the external work done by the gas during isothermal expansion from A to B at temperature T_1. Suppose W_1 is the external work done by the gas during isothermal expansion AB. Then it will be equal to the area ABGEA. Hence

$$W_1 = \text{Area ABGEA}$$

(b) **Adiabatic expansion :** Next the cylinder is removed from the source and placed on a perfectly non-conducting stand. It further decreases the load on the piston to P_3. The expansion is completely adiabatic because no heat can enter or leave the working substance. Therefore, the working substance performs external work in raising the piston at the expense of its internal energy. Hence its temperature falls. The gas is thus allowed to expand adiabatically until its temperature falls to T_2, the temperature of the sink. It has been represented by the adiabatic curve BC on the indicator diagram. We call it ***adiabatic expansion.*** If the pressure and volume of the substance are P_3 and V_3, respectively at C, and W_2 is the work done by the substance from B to C, then

$$W_2 = \text{Area BCHGB.}$$

(c) **Isothermal compression :** Remove the cylinder from the non-conducting stand and place it on the sink at temperature T_2. In order to compress the gas slowly, increase the load (pressure) on the piston until its pressure and volume become P_4 and V_4, respectively. It is represented by the point D on the indicator diagram (Fig. 11.4). The heat developed (H_2) due to compression will pass to the sink. Thus, there is no change in the temperature of the system. Therefore, it is called an isothermal compression process. It is shown by the curve CD (Fig. 11.4). The quantity of heat rejected (H_2) to the sink during this process is equal to the work done (say W_3) on the working substance. Hence

$$W_3 = \text{Area CHFDC}$$

(d) **Adiabatic compression :** Once again place the system on the non-conducting stand. Increase the load on the piston slowly. The substance will under go an adiabatic compression. This compression continues until the temperature rises to T_1 and the substance comes back to its original pressure P_1 and volume V_1. This is an adiabatic compression process and represented by the curve DA on the indicator diagram (Fig. 11.4). Suppose W_4 is the work done during this adiabatic compression from D to A. Then

$$W_4 = \text{Area DFEAD}$$

During the above cycle of operations, the working substance takes H_1 amount of heat from the source and rejects H_2 amount of heat to the sink. Hence the net amount of heat absorbed by the working substance is

$$\Delta H = H_1 - H_2$$

Also the net work done (say W) by the engine in one complete cycle

$$W = \text{Area ABCHEA} - \text{Area CHEADC}$$

$$= \text{Area ABCD}$$

Thus, the work done in one cycle is represented on a P-V diagram by the area of the cycle.

You have studied that the initial and final states of the substance are the same. It means that its internal energy remains unchanged. Hence according to the first law of thermodynamics

$$W = H_1 - H_2$$

Therefore, heat has been converted into work by the system, and any amount of work can be obtained by merely repeating the cycle.

11.5.2 Efficiency of Carnot Engine

Efficiency is defined as the ratio of heat converted into work in a cycle to heat taken from the source by the working substance. It is denoted as η:

$$\eta = \frac{\text{Heat converted into work}}{\text{Heat taken from source}}$$

or

$$\eta = \frac{H_1 - H_2}{H_1} = 1 - \frac{H_2}{H_1}$$

It can be shown that for Carnot's engine,

$$\frac{H_2}{H_1} = \frac{T_2}{T_1}$$

Hence, $$\eta = 1 - \frac{T_2}{T_1}$$

Note that efficiency of carnot engine does not depend on the nature of the working substance. Further, if no heat is rejected to the sink, η will be equal to one. But for H_2 to be zero, T_2 must be zero. It means that efficiency η can be100% only when $T_2 = 0$. The entire heat taken from the hot source is converted into work. This violates the second law of thermodynamics. Therefore, a steam engine can operate only between finite temperature limits and its efficiency will be less than one.

It can also be argued that the Carnot cycle, being a reversible cycle, is most efficient; no engine can be more efficient than a Carnot engine operating between the same two temperatures.

11.5.3 Limitation of Carnot's Engine

You have studied about Carnot's cycle in terms of isothermal and adiabatic processes. Here it is important to note that the isothermal process will take place only when piston moves very slowly. It means that there should be sufficient time for the heat to transfer from the working substance to the source. On the other hand, during the adiabatic process, the piston moves extremely fast to avoid heat transfer. In practice, it is not possible to fulfill these vital conditions. Due to these very reasons, all practical engines have an efficiency less than that of Carnot's engine.

INTEXT QUESTIONS 11.3

1. State whether the following statements are true or false.
   (i) In a Carnot engine, when heat is taken by a perfect gas from a hot source, the temperature of the source decreases.
   (ii) In Carnot engine, if temperature of the sink is decreased the efficiency of engine also decreases.
2. (i) A Carnot engine has the same efficiency between 1000K and 500K and between TK and 1000K. Calculate T.
   (ii) A Carnot engine working between an unknown temperature T and ice point gives an efficiency of 0.68. Deduce the value of T.

WHAT YOU HAVE LEARNT

- Heat is a form of energy which produces in us the sensation of warmth.
- The energy which flows from a body at higher temperature to a body at lower temperature because of temperature difference is called heat energy.
- The most commonly known unit of heat energy is calorie. 1 cal = 4.18 J and 1k cal = 10^3 cal.
- A graph which indicates how the pressure (P) of a system varies with its volume during a thermodynamic process is known as indicator diagram.
- Work done during expansion or compression of a gas is $P\Delta V = P(V_f - V_i)$.
- Zeroth law of thermodynamics states that if two systems are separately in thermal equilibrium with a third system, then they must also be in thermal equilibrium with each other.
- The sum of kinetic energy and potential energy of the molecules of a body gives the internal energy. The relation between internal energy and work is $U_i - U_f = -W$.
- The first law of thermodynamics states that the amount of heat given to a system is equal to the sum of change in internal energy of the system and the external work done.
- First law of thermodynamics tells nothing about the direction of the process.
- The process which can be retraced in the opposite direction from its final state to initial state is called a reversible process.
- The process which can not be retraced along the same equilibrium state from final to the initial state is called an irreversible process. A process that occurs at constant temperature is an isothermal process.
- Any thermodynamic process that occurs at constant heat is an adiabatic process.
- The different states of matter are called its phase and the pressure and temperature diagram showing three phases of matter is called a phase diagram.
- Triple point is a point (on the phase diagram) at which solid, liquid and vapour states of matter can co-exist. It is characterised by a particular temperature and pressure.
- According to Kelvin-Planck's statement of second law, it is not possible to obtain a continuous supply of work from a single source of heat.

- According to Clausius statement of second law, heat can not flow from a colder body to a hotter body without doing external work on the working substance.
- The three essential requirements of any heat engine are :
  - (i) source from which heat can be drawn
  - (ii) a sink into which heat can be rejected.
  - (iii) working substance which performs mechanical work after being supplied with heat.
- Carnot's engine is an ideal engine in which the working substance is subjected to four operations (i) Isothermal expansion (ii) adiabatic expansion (iii) isothermal compression and (iv) adiabatic compression. Such a cycle is called a Carnot cycle.
- Efficiency of a Carnot engine is given only

  $\eta = 1 - \frac{H_2}{H_1}$, H_1 = Amount of heat absorbed and H_2 = Amount of heat rejected.

  $= 1 - \frac{T_2}{T_1}$, T_1 = Temperature of the source, and T_2 = Temperature of the sink.
- Efficiency does not depend upon the nature of the working substance.

TERMINAL EXERCISE

1. Distinguish between the terms internal energy and heat energy.
2. What do you mean by an indicator diagram. Derive an expression for the work done during expansion of an ideal gas.
3. Define temperature using the Zeroth law of thermodynamics.
4. State the first law of thermodynamics and its limitations.
5. What is the difference between isothermal, adiabetic, isobaric and isochoric processes?
6. State the Second law of thermodynamics.
7. Discuss reversible and irreversible processes with examples.
8. Explain Carnot's cycle. Use the indicator diagram to calculate its efficiency.
9. Calculate the change in the internal energy of a system when (a) the system absorbs 2000J of heat and produces 500 J of work (b) the system absorbs 1100J of heat and 400J of work is done on it.

10. A Carnot's engine whose temperature of the source is 400K takes 200 calories of heat at this temperature and rejects 150 calories of heat to the sink. (i) What is the temperature of the sink. (ii) Calculate the efficiency of the engine.

ANSWERS TO INTEXT QUESTIONS

11.1

1. (i) Temperature (ii) C (iii) Joule or Calorie
2. (a) $P_2 (V_2 - V_1)$ (b) $-P_2 (V_2 - V_1)$
3. (i) retrace (ii) irreversible
4. An isothermal process occurs at a constant temperature whereas an adiabatic process occurs at constant heat.
5. At triple point all three states of matter i.e. solid, liquid and vapour can co-exist.

11.2

1. (i) Internal energy (ii) on
2. It states that the amount of heat given to a system is equal to the sum of the change in internal energy of the system and the external energy.

11.3

1. (i) False (ii) True
2. (i) 2000 K (ii) 8583.1K

Answers to Terminal Problems

9. (a) 1500 J (b) 1500 J.

10. 300K, 25%

312en12

12

HEAT TRANSFER AND SOLAR ENERGY

In the previous lesson you have studied the laws of thermodynamics, which govern the flow and direction of thermal energy in a thermodynamic system. In this lesson you will learn about the processes of heat transfer. The energy from the sun is responsible for life on our beautiful planet. Before reaching the earth, it passes through vacuum as well as material medium between the earth and the sun. Do you know that each one of us also radiates energy at the rate of nearly 70 watt? Here we will study the **radiation** in detail. This study enables us to determine the temperatures of stars even though they are very far away from us.

Another process of heat transfer is **conduction,** which requires the presence of a material medium. When one end of a metal rod is heated, its other end also becomes hot after some time. That is why we use handles of wood or similar other bad conductor of heat in various appliances. Heat energy falling on the walls of our homes also enters inside through conduction. But when you heat water in a pot, water molecules near the bottom get the heat first. They move from the bottom of the pot to the water surface and carry heat energy. This mode of heat transfer is called **convection**. These processes are responsible for various natural phenomena, like monsoon which are crucial for existence of life on the globe. You will learn more about these processes of heat transfer in this unit.

OBJECTIVES

After studying this lesson, you should be able to :

- *distinguish between conduction convection and radiation;*
- *define the coefficient of thermal conductivity;*
- *define the emissive power and the absorptive power of a body;*

Notes

- *describe green house effect and its consequencies for life on earth; and*
- *apply laws governing black body radiation.*

12.1 PROCESSES OF HEAT TRANSFER

You have learnt the laws of thermodynamics in the previous lesson. The second law postulates that the natural tendency of heat is to flow spontaneously from a body at higher temperature to a body at lower temperature. The transfer of heat continues until the temperatures of the two bodies become equal. From kinetic theory, you may recall that temperature of a gas is related to its average kinetic energy. It means that molecules of a gas at different temperatures have different average kinetic energies.

There are three processes by which transfer of heat takes place. These are : *conduction*, *convection* and *radiation.* In conduction and convection, heat transfer takes place through molecular motion. Let us understand how this happens.

Heat transfer through *conduction* is more common in solids. We know that atoms in solids are tightly bound. When heated, they can not leave their sites; they are constrained to vibrate about their respective equilibrium positions. Let us understand as to what happens to their motion when we heat a metal rod at one end (Fig.12.1). The atoms near the end A become hot and their kinetic energy increases. They vibrate about their mean positions with increased kinetic energy and being in contact with their nearest neighbouring atoms, pass on some of their kinetic energy (K.E.) to them. These atoms further transfer some K.E to their neighboures and so on. This process continues and kinetic energy is transferred to atoms at the other end B of the rod. As average kinetic energy is proportional to temperature, the end B gets hot. Thus, *heat* is *transferred from atom* to *atom by conduction. In this process, the atoms* do *not bodily move* but *simply vibrate about their mean equilbrium positions and pass energy from one to another.*

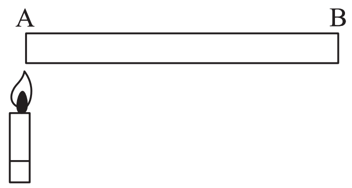


Fig. 12.1 : Heat conduction in a metal rod

In *convection,* molecules of fluids receive thermal energy and move up bodily. To see this, take some water in a flask and put some grains of potassium permanganate ($KMnO_4$) at its bottom. Put a bunsen flame under the flask. As the fluid near the bottom gets heated, it expands. The density of water decreases and the buoyant force causes it to move upward (Fig.12.2). The space occupied by hot water is taken

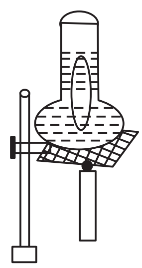


Fig. 12.2 : Convection currents are formed in water when heated

by the cooler and denser water, which moves downwards. Thus, a convection current of hotter water going up and cooler water coming down is set up. The water gradually heats up. These convection currents can be seen as $KMnO_4$ colours them red.

In *radiation,* heat energy moves in the form of waves. You will learn about the characteristics of these waves in a later section. These waves can pass through vacuum and do not require the presence of any material medium for their propagation. Heat from the sun comes to us mostly by radiation.

We now study these processes in detail.

12.1.1 Conduction

Consider a rectangular slab of area of cross-section **A** and thickness *d*. Its two faces are maintained at temperatures T_h and T_c ($< T_h$), as shown in Fig. 12.3. Let us consider all the factors on which the quantity of heat *Q* transferred from one face to another depends. We can intuitively feel that larger the area A, the greater will be the heat transferd ($Q \propto A$). Also, greater the thickness, lesser will be the heat transfer ($Q \propto 1/d$). Heat transfer will be more if the temperature difference between the faces, $(T_h - T_c)$, is large. Finally longer the time *t* allowed for heat transfer, greater will be the value of *Q*. Mathematically, we can write

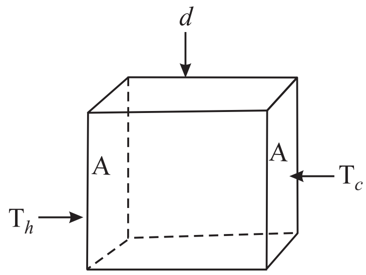


Fig. 12.3 : Heat conduction through a slab of thickness *d* and surface area *A*, when the faces are kept at temperatures T_h and T_c.

$$Q \propto \frac{A(T_h - T_c).t}{d}$$

$$Q = \frac{KA(T_h - T_c)\,t}{d} \qquad (12.1)$$

where *K* is a constant which depends on the nature of the material of the slab. It is called the coefficient of thermal conductivity,or simply, **thermal conductivity** of the material. **Thermal conductivity** of a material is defined as the amount of heat transferred in one second across a piece of the material having area of cross-section $1 m^2$ and edge 1m when its opposite faces are maintained at a temperature difference of 1 K. The SI unit of thermal conductivity is $W\,m^{-1}\,k^{-1}$. The value of K for some materials is given in Table 12.1

Table 12.1 : Thermal Conductivity of some materials

Material	Thermal conductivity ($Wm^{-1}\,K^{-1}$)
Copper	400
Aluminium	240
Concrete	1.2
Glass	0.8
Water	0.60
Body talc	0.20
Air	0.025
Thermocole	0.01

Example 12.1 : A cubical thermocol box, full of ice. has side 30 cm and thickness of 5.0 cm. If outside temperature is 45°C, estimate the amount of ice melted in 6 h. (K for thermocol is 0.01 J s^{-1} m^{-1} $°C^{-1}$ and latent heat of fusion of ice is 335 J g^{-1}.

Solution : The quantity of heat transferred into the box through its one face can be obtained using Eq. (12.1) :

$$Q = \frac{KA(T_h - T_c)t}{d}$$

$$= (0.01 \text{ J s}^{-1}\text{ m}^{-1}\text{°C}^{-1}) \times (900 \times 10^{-4}\text{ m}^2) \times (45\text{°C}) \times (6 \times 60 \times 60\text{ s}) / (5 \times 10^{-2}\text{ m})$$

$$= 10496 \text{ J}$$

Since the box has six faces, total heat passing into the box

$$Q = 10496 \times 6 \text{ J}$$

The mass of ice melted m, can be obtained by dividing Q by L :

$$m = Q/L$$

$$= \frac{10496\text{J}}{335\text{ Jg}^{-1}} \times 6$$

$$= 313 \times 6 \text{ g} = 1878 \text{ g}$$

We can see from Table 12.1 that metals such as copper and aluminium have high thermal conductivity. This implies that heat flows with more ease through copper.This is the reason why cooking vessels and heating pots are made of copper. On the other hand, air and thermocol have very low thermal conductivities. Substances having low value of K are sometimes called thermal insulators. We wear woollen clothes during winter because air trapped in wool fibres prevents heat loss from our body. Wool is a good thermal insulator because air is trapped between its fibres. The trapped heat gives us a feeling of warmth. Even if a few cotton clothes are put on one above another, the air trapped in-between layers stops cold. In the summer days, to protect a slab of ice from melting, we put it in a ice box made of thermocol. Sometimes we wrap the ice slab in jute bag, which also has low thermal conductivity.

12.1.2 Convection

It is common experience that while walking by the side of a lake or a sea shore on a hot day, we feel a cool breeze. Do you know the reason? Let us discover it.

Due to continuous evaporation of water from the surface of lake or sea, the temperature of water falls. Warm air from the shore rises and moves upwards (Fig.12.4). This creates low pressure area on the shore and causes cooler air from water surface to move to the shore. The net effect of these convection currents is the transfer of heat from the shore,which is hotter, to water, which is cooler. The rate of heat transfer depends on many factors.There is no simple equation for convection as for conduction. However, the ***rate of heat transfer by convection depends on the temperature difference between the surfaces and also on their areas.***

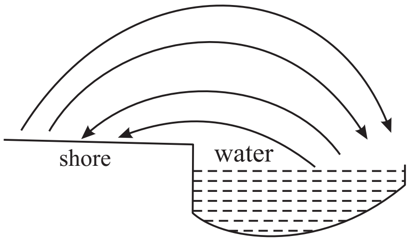


Fig. 12.4 : Convection currents. Hot air from the shore rises and moves towards cooler water. The convection current from water to the shores is experienced as cool breeze.

Now let us check how much you have learnt about the methods of heat transfer.

12.1.3 Radiation

Radiation refers to continuous emission of energy from the surface of a body. This energy is called radiant energy and is in the form of electromagnetic waves. These waves travel with the velocity of light (3×10^8 ms^{-1}) and can travel through vacuum as well as through air. They can easily be reflected from polished surfaces and focussed using a lens.

All bodies emit radiation with wavelengths that are chracteristic of their temperature. The sun, at 6000 K emits energy mainly in the visible spectrum. The earth at an ideal radiation temperature of 295 K radiates energy mainly in the far infra-red (thermal) region of electromagnetic spectrum. The human body also radiates energy in the infra-red region.

Let us now perform a simple experiment. Take a piece of blackened platinum wire in a dark room. Pass an electrical current through it. You will note that the wire has become hot. Gradually increase the magnitude of the current. After sometime, the wire will begin to radiate. When you pass a slightly stronger current, the wire will begin to glow with dull red light. This shows that the wire is just emitting red radiation of sufficient intensity to affect the human eye. This takes place at nearly 525°C. With further increase in temperature, the colour of the emitted rediation will change from dull red to cherry red (at nearly 900°C) to orange (at nearly 1100°C), to yellow (at nearly 1250°C) until at about 1600°C, it becomes white. What do you infer from this? It shows that the *temperature of a luminous body can be estimated from its colour.* Secondly, *with increase in temperature, waves of shorter wavelengths* (since red light is of longer wavelength than orange. yellow etc.) *are also emitted with sufficient intensity*. Considering in reverse order, you may argue that when the temperature of the wire is below 525°C, it emits waves longer than red but these waves can be detected only by their heating effect.

INTEXT QUESTIONS 12.1

1. Distinguish between conduction and convection.
2. Verify that the units of K are Js $^{-1}$ m^{-1} °C^{-1}.
3. Explain why do humans wrap themselves in woollens in winter season?
4. A cubical slab of surface area 1 m^2, thickness 1 m, and made of a material of thermal conductivity K. The opposite faces of the slab are maintained at 1°C temperature difference. Compute the energy transferred across the surface in one second. and hence give a numerical definition of K.
5. During the summer, the land mass gets very hot. But the air over the ocean does not get as hot. This results in the onset of sea breezes. Explain.

12.2 RADIATION LAWS

At any temperature, the radiant energy emitted by a body is a mixture of waves of different wavelengths.The most intense of these waves will have a particular wavelength (sayλ_m).At 400°C, the λ_m will be about 5×10^{-4} cm or 5 μm (1 micron (μ) =10^{-6}m) for a copper block.The intensity decreases for wavelengths either greater or less than this value (Fig. 12.5).

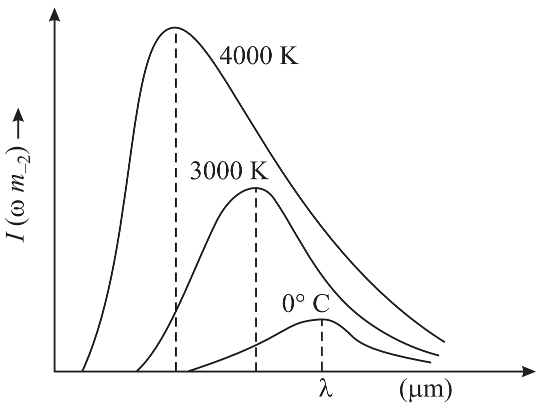


Fig. 12.5 : Variation in intensity with wavelength for a black body at different temperatures

Evidently area between each curve and the horizontial axis represents the total rate of radiation at that temperature. You may study the curves shown in Fig. 12.5 and verify the following two facts.

1) The rate of radiation at a particular temperature (represented by the area between each curve and the holizontal axis) increases rapidly with temperature.

2) Each curve has a definite energy maximum and a corresponding wavelength λ_m (i.e. wavelength of the most intense wave). The λ_m shifts towards shorter wavelengths with increasing temperature.

This second fact is expressed quantitatively by what is known as Wien's displacement law. It **states that λ_m *shifts towards shorter wavelengths as the temperature of a body is increased*.** This law is., strictly valid only for black bodies. Mathematically, ***we say that the product λ_m T is constant for a body emitting radiation at temperature* T*:***

$$\lambda_m T = \text{constant} \qquad (12.2)$$

The constant in Eqn. (12.2) has a value 2.884×10^{-3} mK. This law furnishes us with a simple method of determining the temperature of all radiating bodies including those in space. The radiation spectrum of the moon has a peak at λ_m = 14 micron. Using Eqn. (12.2), we get

$$T = \frac{2884 \text{ micron K}}{14 \text{ micron}} = 206\text{K}$$

That is, the temperature of the lunar surface is 206K

Wilhelm Wien
(1864 – 1928)

The 1911 Nobel Leureate in physics, Wilhelm Wien, was son of a land owner in East Prussia. After schooling at Prussia, he went to Germany for his college. At the University of Berlin, he studied under great physicist Helmholtz and got his doctorate on diffraction of light from metal surfaces in 1886.

He had a very brilliant professional carrer. In 1896, he succeeded Philip Lenard as Professor of Physics at Aix-la-chappelle. In 1899, he become Professor of Physics at University of Giessen and in 1900, he succeeded W.C. Roentgen at Wurzberg. In 1902, he was invited to succeed Ludwig Boltzmann at University of Leipzig and in 1906 to succeed Drude at University of Berlin. But he refused these invitations. In 1920, he was appointed Professor of Physics at munich and he remained there till his last.

12.2.1 Kirchhoff's Law

As pointed out earlier, when radiation falls on matter, it may be partly reflected, partly absorbed and partly transmitted. If for a particular wavelength λ and a given surface, r_λ, a_λ and t_λ, respectively denote the fraction of total incident energy reflected, absorbed and transmitted, we can write

$$1 = r_\lambda + a_\lambda + t_\lambda \qquad (12.3)$$

A body is said to be perfectly black, if $r_\lambda = t_\lambda = 0$ and $a_\lambda = 1$. It means that radiations incident on black bodies will be completely absorbed. As such, perfectly black body does not exist in nature. Lamp black is the nearest approximation to a black body. It absorbs about 96% of visible light and platinum black absorbs about 98%. It is found to transmit light of long wavelength.

A ***perfectly white body***, in contrast, defined as a body with $a_\lambda = 0$, $t_\lambda = 0$ and $r_\lambda = 1$. A piece of white chalk approximates to a perfectly white body.

This implies that good emitters are also good absorbers. But each body must either absorb or reflect the radiant energy reaching it. So we can say that a good absorber must be a poor reflector (or good emitter).

Designing a Black Body

Kirchoff's law also enables us to design a perfectly black body for experimental purposes. We go back to an enclosure at constant temperature containing radiations between wavelength range λ and $\lambda + d\lambda$. Now let us make a small hole in the enclosure and examine the radiation escaping out of it.This radiation undergoes multiple reflections from the walls. Thus, if the reflecting power of the surface of the wall is r, and emissive power is e_λ, the total radiation escaping out is given by

$$E_\lambda = e_\lambda + e_\lambda r_\lambda + e_\lambda r^2 + e_\lambda r_\lambda^3 + \ldots$$

$$= e_\lambda (1 + r_\lambda + r_\lambda^2 + r_\lambda^3 + \ldots)$$

$$= \frac{e_\lambda}{1 - r_\lambda} \qquad (12.4)$$

But from Kirchoffs Law $\dfrac{e_\lambda}{a_\lambda} = E_\lambda$

$$e_\lambda = E_\lambda a_\lambda \qquad (12.5)$$

where E_λ is the emission from a black body. If now walls are assumed to be opaque (i.e. $t = 0$), from Eqn. (12.3), we can write

$$a_\lambda = 1 - r_\lambda \qquad (12.6)$$

Substituting this result in Eqn. (12.5), we get

$$e_\lambda = E_\lambda (1 - r_\lambda)$$

or

$$E_\lambda = \frac{e_\lambda}{1 - r_\lambda} \qquad (12.7)$$

On comparing Eqns. (12.4) and (12.7), we note that the radiation emerging out of the hole will be identical to the radiation from a perfectly black emissive surface. Smaller the hole, the more completely black the emitted radiation is. So we see that the **uniformly heated enclosure with a small cavity behaves as a black body for emission.**

Such an enclosure behaves as a perfectly black body towards incident radiation also. Any radiation passing into the hole will undergo multiple reflections internally within the enclosure and will be unable to escape outside. This may be further improved by blackening the inside.Hence the enclosure is a perfect absorber and behaves as a perfectly black body.

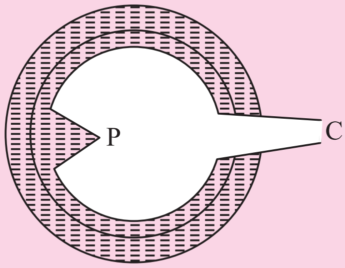


Fig. 12.6 : Fery's black body

Fig. 12.6 shows a black body due to Fery. There is a cavity in the form of a hollow sphere and its inside is coated with black material. It has a small conical opening O. Note the conical projection P opposite the hole O. This is to avoid direct radiation from the surface opposite the hole which would otherwise render the body not perfectly black.

ACTIVITY 12.1

You have studied that black surface absorbs heat radiations more quickly than a shiny white surface. You can perform the following simple experiment to observe this effect.

Take two metal plates A and B. Coat one surface of A as black and polish one surface of B. Take an electric heater. Support these on vertical stands such that the coated black surface and coated white surface face the heater. Ensure that coated plates are equidstant from the heater. Fix one cork each with wax on the uncoated sides of the plates.

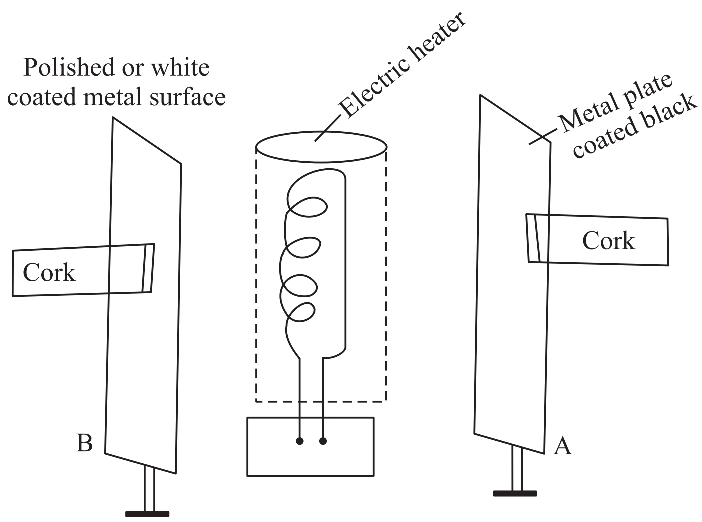


Fig. 12.7 : Showing the difference in heat absorption of a black and a shining surface

Switch on the electric heater. Since both metal plates are identical and placed at the same distance from the heater, they receive the same amount of radiation from it. You will observe that the cork on the blackened plate falls first. This is becasuse the black surface absorbs more heat than the white surface. This proves that black surfaces are good aborbers of heat radiations.

12.2.2 Emissive and Absorptive Power

Different bodies at the same temperature emit different amounts of thermal energy. The ability of a hot body to emit radiation is known as its **emissive power**. The total emissive power of a radiating body at a particular temperature is defined as the total amount of energy radiated per second per unit area of its surface. It also depends upon the temperature of the body above the surroundings. Its unit is $Jm^{-2}s^{-1}$. At the same temperature the total emissive power of a black body has the maximum value (E_b). The ratio of the total emissive power, E of a real body to the total emissive power E_b of a black-body at the same temperature is known as emissivity ε. Thus, emissivity,

$$\varepsilon = \frac{E}{E_b}$$

or

$$E = \varepsilon E_b$$

Note that both E and E_b are temperature dependent. Emissivity is also not a constant. It shows small variation with temperature.

When the radiant energy falls on a body, a part of the energy is absorbed. The ability of the body to absorb radiant energy falling on it is known as its **absorptive power**.

The total absorptive power of a body is defined as the ratio of the energy absorbed to the energy falling. The absorptive power (a) is the fraction of the incident energy which is absorbed. For a perfectly black body, $a = 1$.

Sometimes it is interesting to know the ability of a body to absorb radiation of a given wavelength. Under such situation, spectral absorptive power term, a_λ, is used. Thus, spectral absorptive power for perfectly black body $a_{b\lambda} = 1$.

It is experimentally found that the good emitters of thermal radiation are also good absorbers. This shows that the emissive power and absorptive power are closely related.

12.2.3 Stefan-Boltzmann Law

On the basis of experimental measurements, Stefan and Boltzmann concluded that the radiant energy emitted per second from a surface of area A is proportional to fourth power of temperature :

$$E = Ae\sigma T^4 \qquad (12.8)$$

where σ is ***Stefan-Boltzmann constant*** and has the value 5.672×10^{-8} $Jm^{-2}\, s^{-1}\, K^{-4}$. The temperature is expressed is kelvin, *e* is emissivity or relative emittance. It depends on the nature of the surface and temperature. The value of *e* lies between 0 and 1; being small for polished metals and 1 for perfectly black materials.

From Eqn. (12.8) you may think that if the surfaces of all bodies are continually radiating energy, why don't they eventually radiate away all their internal energy and cool down to absolute zero. They would have done so if energy were not supplied to them in some way. In fact, all objects radiate and absorb energy simultaneously. If a body is at the same temperaturture as its surroundings, the rate of emission is same as the rate of absorption; there is no net gain or loss of energy and no change in temperature. However, if a body is at a lower temperture than its surroundings, the rate of absorption will be greater than the rate of emission. Its temperature will rise till it is equal to the room temperature. Similarly, if a body is at higher temperature, the rate of emission will be greater than the rate of absorption. There will be a net energy loss. Hence, when a body at a temperature T_1 is placed in surroundings at temperature T_2, the amount of net energy loss per second is given by

$$E_{net} = Ae\sigma\ (T_1^4 - T_2^4) \text{ for } T_1 > T_2 \qquad (12.5)$$

Example 12.2 : Determine the surface area of the filament of a 100 W incandescent lamp at 3000 K. Given $\sigma = 5.7 \times 10^{-8}$ W m^{-2} K^{-4}, and emissivity *e* of the filament = 0.3.

Solution: According to Stefan-Boltzmann law

$$E = eA\sigma\ T^4$$

where *E* is rate at which energy is emitted, *A* is surface area, and T is tempeature of the surface. Hence we can rewrite it as

$$A = \frac{E}{e\sigma T^4}$$

On substituting the given data, we get

$$A = \frac{100\text{ W}}{0.3 \times (5.7 \times 10^{-8}\text{ Wm}^{-2}\text{ K}^{-4} \times (3000\text{K})^4}$$

$$= 7.25 \times 10^{-5}\text{ m}^2$$

Now it is time for you to check your understanding.

INTEXT QUESTIONS 12.2

1. At what wavelength does a cavity radiator at 300K emit most radiation?
2. Why do we wear light colour clothing during summer?
3. State the important fact which we can obtain from the experimental study of the spectrum of black body radiation.
4. A person with skin temperature 28°C is present in a room at temperature 22°C. Assuming the emissivity of skin to be unity and surface area of the person as 1.9 m^2, compute the radiant power of this person.
5. Define the emissive and absorptive power of a body. What is a perfectly black body?

12.3 SOLAR ENERGY

You have learnt in your previous classes that sun is the ultimate source of all energy available on the earth. The sun is radiating tremendous amount of energy in the form of light and heat and even the small fraction of that radiation received by earth is more than enough to meet the needs of living beings on its surface. The effective use of solar energy, therefore, may some day provide solution to our energy needs.

Some basic issues related with solar radiations are discussed below.

1. Solar Constant

To calculate the total solar energy reaching the earth, we first determine the amount of energy received per unit area in one second. The energy is called **solar constant**. Solar constant for earth is found to be 1.36×10^3 W m^{-2}. Solar constant multiplied by the surface area of earth gives us the total energy received by earth per second. Mathematically,

$$Q = 2\pi R_e^2 C$$

where R_e is radius of earth and C is solar constant

Note that Only half of the earth's surface has been taken into account as only this much of the surface is illuminated at one time. Therefore,

$$Q = 2 \times 3.14 \times (6.4 \times 10^6 \text{ m})^2 \times (1.36 \times 10^3 \text{ W m}^{-2})$$

$$\simeq 3.5 \times 10^{17} \text{ W}$$

$$\simeq 3.5 \times 10^{11} \text{ MW}$$

To determine solar constant for other planets of the solar system, we may make use of Stefan-Boltzman law, which gives the total energy emitted by the sun in one second :

$$\in = (4\pi r^2)\,\sigma T^4$$

where r is radius of sun and T is its temperature.

If R is radius of the orbit of the planet, then

$$E = \frac{\in}{4\pi R^2} = \left(\frac{r}{R}\right)^2 \sigma T^4 \qquad (12.6)$$

And the solar constant (E′) at any other planet orbiting at distance R′ from the sun would be

$$E' = \left(\frac{r}{R'}\right)^2 \sigma T^4 \qquad (12.7)$$

Hence

$$\frac{E'}{E} = \left(\frac{R}{R'}\right)^2 \qquad (12.8)$$

The distance of mars is 1.52 times the distance of earth from the sun. Therefore, the solar constant at mars

$$E' = E \times \left(\frac{1}{1.52}\right)^2$$

$$= 6 \times 10^2 \text{ W m}^{-2}$$

2. Greenhouse Effect

The solar radiations in appropriate amount are necessary for life to flourish on earth. The atmosphere of earth plays an important role to provide a comfortable temperature for the living organisms. One of the processes by which this is done is greenhouse effect.

In a greenhouse, plants, flowers, grass etc. are enclosed in a glass structure. The glass allows short wavelength radiation of light to enter. This radiation is absorbed by plants. It is subsequently re-radiated in the form of longer wavelength heat radiations – the infrared. The longer wavelength radiations are not allowed to escape from the greenhouse as glass is effectively opaque to heat. These heat radiations are thus trapped in the greenhouse keeping it warm.

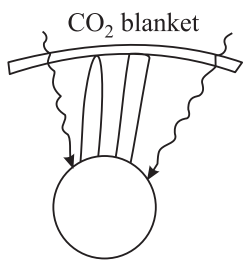


Fig. 12.8 : Green house effect

An analogous effect takes place in our atmosphere. The atmosphere, which contains a trace of carbon dioxide, is transparent to visible light. Thus, the sun's light passes through the atmosphere and reaches the earth's surface. The earth absorbs this light and subsequently emits it as infrared radiation. But carbon dioxide in air is opaque to infra-red radiations.CO_2 reflects these radiations back rather than allowing them to escape into the atmosphere. As a result, the temperature of earth increases. This effect is referred to as the ***greenhouse effect.***

Due to emission of huge quantities of CO_2 in our atmosphere by the developed as well as developing countries, the greenhouse effect is adding to global warming and likely to pose serious problems to the existence of life on the earth. A recent report by the UN has urged all countries to cut down on their emissions of CO_2, because glaciers have begun to shrink at a rapid rate. In the foreseable future, these can cause disasters beyond imagination beginning with flooding of major rivers and rise in the sea level. Once the glaciers melt, there will be scarcity of water and erosion in the quality of soil. There is a lurking fear that these together will create problems of food security. Moreover, changing weather patterns can cause droughts & famines in some regions and floods in others.

In Indian context, it has been estimated that lack of positive action can lead to serious problems in Gangetic plains by 2030. Also the sea will reclaim vast areas along our coast lie, inundating millions of people and bring unimaginable misery and devastation. How can you contribute in this historical event?

12.4 NEWTON'S LAW OF COOLING

Newton's law of cooling states that ***the rate of cooling of a hot body is directly proportional to the mean excess temperature of the hot body over that of its surroundings provided the difference of temperature is small. The law can be deduced from stefan-Boltzmann law.***

Let a body at temperature T be surrounded by another body at T_o. The rate at which heat is lost per unit area per second by the hot body is

$$E = e\sigma (T^4 - T_0^4)A \qquad (12.9)$$

As $T^4 - T_0^4 = (T^2 - T_0^2)(T^2 + T_0^2) = (T - T_0)(T - T_0)(T^2 + T_0^2)$. Hence (12.10)

$$E = e\sigma (T - T_0)(T^3 + T^2 T_0 + T T_0^2 + T_0^3)A$$

If $(T - T_0)$ is very small, each of the term T^3, T^2T_0, $T T_0^2$ and T_0^3 may be approximated to T_0^3. Hence

$$\therefore \quad E = e\sigma (T - T_0)\, 4T_0^3 A$$
$$= k (T - T_0)$$

where $k = 4e\sigma T_0^3 A$. Hence,

$$E \alpha (T - T_0) \qquad (12.11)$$

This is ***Newton's law of cooling.***

INTEXT QUESTIONS 12.3

1. Calculate the power received from sun by a region 40m wide and 50m long located on the surface of the earth?
2. What threats are being posed for life on the earth due to rapid consumption of fossil fuels by human beings?
3. What will be shape of cooling curve of a liquid?

WHAT YOU HAVE LEARNT

- Heat flows from a body at higher temperature to a body at lower temperature. There are three processes by which heat is transferred : conduction, convection and radiation.
- In conduction, heat is transferred from one atom/ molecule to another atom/ molecule which vibrate about their fixed positions.
- In convection, heat is transferred by bodily motion of molecules. In radiation, heat is transferred through electromagnetic waves.
- The quantity of heat transferred by conduction is given by

$$Q = \frac{K(T_h - T_c)\ At}{d}$$

- *Wien's Law.* The spectrum of energy radiated by a body at temperature T(K) has a maxima at wavelength λ_m' such that $\lambda_m T$ = constant (= 2880 μK)
- Stefan-Boltzmann Law. The rate of energy radiated by a source at T(K) is given by $E = e\sigma A T^4$

  The absorptive power a is defined as

$$\mathbf{a} = \frac{\text{Total amount of energy absorbed between } \lambda \text{ and } \lambda + d\lambda}{\text{Total amount of incident energy between } \lambda \text{ and } \lambda + d\lambda}$$

- The emissive power of a surface e_λ is the amount of radiant energy emitted per square metre area per second per unit wavelength range at a given temperature.
- The solar constant for the earth is $1.36 \times 10^3\ \text{Jm}^{-2}\ \text{s}^{-1}$
- Newton's Law of cooling states that the rate of cooling of a body is linearly proportional to the excess of temperature of the body above its surroundings.

TERMINAL EXERCISE

1. A thermosflask (Fig.12.9) is made of a double walled glass bottle enclosed in metal container. The bottle contains some liquid whose temperature we want to maintain, Look at the diagram carefully and explain how the construction of the flask helps in minimizing heat transfer due to conduction convection and radiation.

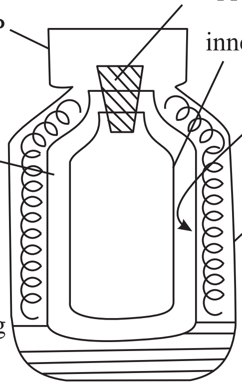


Fig. 12.9

2. The wavelength corresponding, to emission of energy maxima of a star is 4000 A°. Compute the temperature of the star.(1A° = 10^{-8} cm).
3. A blackened solid copper sphere of radius 2cm is placed in an evacuated enclosure whose walls are kept at 1000° C. At what rate must energy be supplied to the sphere to keep its temperature constant at 127° C.
4. Comment on the statement "A good absorber must be a good emitter"
5. A copper pot whose bottm surface is 0.5cm thick and 50 cm in diameter rests on a burner which maintains the bottom surface of the pot at 110°C. A steady heat flows through the bottom into the pot where water boils at atmospheric pressure. The actual temperature of the inside surface of the bottom of the pot is 105°C. How many kilograms of water boils off in one hour?
6. Define the coefficient of thermal conductivity. List the factors on which it depends.
7. Distinguish between conduction and convection methods of heat (transfer).
8. If two or more rods of equal area of cross-section are connected in series, show that their equivalent thermal resistance is equal to the sum of thermal resistance of each rod.

   [**Note :** Thermal resistance is reciprocal of thermal conductivity]
9. Ratio of coefficient of thermal conductivities of the different materials is 4:3. To have the same thermal resistance of the two rods of these materials of equal thickness. what should be the ratio of their lengths?

10. Why do we feel warmer on a winter night when clouds cover the sky than when the sky is clear?
11. Why does a peice of copper or iron appear hotter to touch than a smilar piece of wood even when both are at the same temperature?
12. Why is it more difficult to sip hot tea from a metal cup than from a china-clay cup?
13. Why are the woollen clothes warmer than cotton clothes?
14. Why do two layers of cloth of equal thickness provide warmer covering than a single layer of cloth of double the thickness?
15. Can the water be boiled by convection inside an earth satellite?
16. A. 500 W bulb is glowing. We keep our one hand 5 cm above it and other 5 cm below it. Why more heat is experienced at the upper hand?
17. Two vessels of different materials are identical in size and in dimensions. They are filled with equal quantity of ice at O°C. If ice in both vessles metls completely in 25 minutes and in 20 minutes respectively compare the (thermal conductivities) of metals of both vessels.
18. Calculate the thermal resistivity of a copper rod 20.0 cm. length and 4.0 cm. in diamter.

    Thermal conductivity of copper = 9.2×10^{-2} temperature different acrosss the ends of the rod be 50°C. Calculate the rate of heat flow.

ANSWERS TO INTEXT QUESTIONS

12.1

1. Conduction is the principal mode of transfer of heat in solids in which the particles transfer energy to the adjoining molecules.

   In convection the particles of the fluid bodily move from high temperature region to low temperature region and vice-versa.

2. $$K = \frac{Qd}{t\,A\,(Q_2 - Q_1)}$$

   $$= \frac{\text{J}}{\text{s}}\,\frac{\text{m}}{\text{m}^2\,°\text{C}}$$

   $$= \text{J s}^{-1}\,\text{m}^{-1}\,°\text{C}^{-1}$$

3. The trapped air in wool fibres prevents body heat from escaping out and thus keeps the wearer warm.
4. The coefficient of thermal conductivity is numerically equal to the amount of heat energy transferred in one second across the faces of a cubical slab of

surface area $1m^2$ and thickness 1m, when they are kept at a temperature difference of 1°C.

5. During the day, land becomes hotter than water and air over the ocean is cooler than the air near the land. The hot dry air over the land rises up and creates a low pressure region. This causes see breeze because the moist air from the ocean moves to the land. Since specific thermal capacity of water is higher than that of sand, the latter gets cooled faster and is responsible for the reverse process during the night causing land breezes.

12.2

1. $\lambda_m = \dfrac{\text{Wien's constant}}{\text{Temperature}}$

$= \dfrac{2880\mu K}{300K}$

$= 9.6\mu$

2. Hint: Because light colours absorb less heat.
3. Hint: (a) $\lambda_m T = S$ (b) $t = \sigma T^4$
4. 66.4 W.

12.3

1. Solar constant x .area

$= 2.7 \times 10^5$ W

2. Constant addition of CO_2 in air will increase greenhouse effect causing global warming due to which glaciers are likely to melt and flood the land mass of the earth.

3. Exponential decay

Answers to Terminal Problem

2. 7210 K

3. 71.6×50^{-11} W

5. 4.7×10^5 kg

9. 3 : 4

17. 4 : 5

18. 10.9 m °C^{-1} W^{-1}, 0.298 W

312en13

13

SIMPLE HARMONIC MOTION

You are now familiar with motion in a straight line, projectile motion and circular motion. These are defined by the path followed by the moving object. But some objects execute motion which are repeated after a certain interval of time. For example, beating of heart, the motion of the hands of a clock, to and fro motion of the swing and that of the pendulum of a bob are localised in space and repetitive in nature. Such a motion is called periodic motion. It is universal phenomenon.

In this lesson, you will study about the periodic motion, particularly the oscillatory motion which we come across in daily life. You will also learn about simple harmonic motion. Wave phenomena – types of waves and their characteristics– form the subject matter of the next lesson.

OBJECTIVES

After studying this lesson, you should be able to :

- *show that an oscillatory motion is periodic but a periodic motion may not be necessarily oscillatory;*
- *define simple harmonic motion and represent it as projection of uniform circular motion on the diameter of a circle;*
- *derive expressions of time period of a given harmonic oscillator;*
- *derive expressions for the potential and kinetic energies of a simple harmoic oscillator; and*
- *distinguish between free, damped and forced oscillations.*

13.1 PERIODIC MOTION

You may have observed a clock and noticed that the pointed end of its seconds hand and that of its minutes hand move around in a circle, each with a fixed period. The seconds hand completes its journey around the dial in one minute but the minutes hand takes one hour to complete one round trip. However, a pendulum bob moves to and fro about a mean position and completes its motion from one end to the other and back to its initial position in a fixed time. A motion which repeats itself after a fixed interval of time is called ***periodic motion***. There are two types of periodic motion : (i) ***non–oscillatory,*** and (ii) ***oscillatory***. The motion of the hands of the clock is non-oscillatory but the to and fro motion of the pendulum bob is oscillatory. However, both the motions are periodic. It is important to note that an oscillatory motion is normally periodic but a periodic motion is not necessarily oscillatory. Remember that a motion which repeats itself in equal intervals of time is periodic and if it is about a mean position, it is ***oscillatory***.

We know that earth completes its rotation about its own axis in 24 hours and days and nights are formed. It also revolves around the sun and completes its revolution in 365 days. This motion produces a sequence of seasons. Similarly all the planets move around the Sun in elliptical orbits and each completes its revolution in a fixed interval of time. These are examples of periodic non-oscillatory motion.

Jean Baptiste Joseph Fourier
(1768 – 1830)

French Mathematician, best known for his Fourier series to analyse a complex oscillation in the form of series of sine and consine functions.

Fourier studied the mathematical theory of heat conduction.
He established the partial differential equation governing heat diffusion and solved it by using infinite series of trigonometric functions.

Born as the ninth child from the second wife of a taylor, he was orphened at the age of 10. From the training as a priest, to a teacher, a revolutionary, a mathematician and an advisor to Nepolean Bonapart, his life had many shades.

He was a contemporary of Laplace, Lagrange, Biot, Poission, Malus, Delambre, Arago and Carnot. Lunar crator Fourier and his name on Eiffel tower are tributes to his contributions.

ACTIVITY 13.1

Suppose that the displacement y of a particle, executing simple harmonic motion, is represented by the equation :

$$y = a \sin \theta \qquad (13.1)$$

or

$$y = a \cos \theta \qquad (13.2)$$

From your book of mathematics, obtain the values of sin θ and cos θ for θ = 0, 30^0, 60^0, 90^0, 120^0, 150^0, 180^0, 240^0, 300^0, 330^0 and 360^0. Then assuming that a = 2.5cm, determine the values of y corresponding to each angle using the relation $y = a \sin\theta$. Choose a suitable scale and plot a graph between y and θ. Similarly, using the relation $y = a \cos\theta$, plot another graph between y and θ. You will note that both graphs represents an oscillation between $+a$ and $-a$. It shows that a certain type of oscillatory motion can be represented by an expression containing sine or cosine of an angle or by a combination of such expressions.

13.1.1 Displacement as a Function of Time

Periodic Motion

When an object repeats its motion after a definite interval of time, its motion is said to be periodic.

Let the position of an object change from O to B, from B to O; then from O to A and finally from A to O, after a fixed interval of time T.

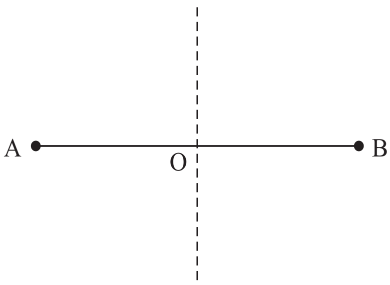


Fig. 13.1

Then, the changes in the position or displacement of the object can be expressed as a function of time:

$$x = af(t + T)$$

where a is a constant and T is the time after which the value of x is repeated

For each time interval T:

$$x = af(T) = 0 \text{ at } t = 0$$

$$x = af(T + T/4) = a \text{ at } t = T/4$$

$$x = af\left(T + \frac{T}{2}\right) = 0 \text{ at } t = \frac{T}{2}$$

$$x = af\left(T + \frac{3T}{4}\right) = -a \text{ at } t = \frac{3T}{4}$$

$$x = af(T + T) = 0 \text{ at } t = T$$

....

....

Thus, x is function of t and it repeats its motion after an interval T. Hence, the motion is periodic.

Now check your progress by answering the following questions.

INTEXT QUESTIONS 13.1

1. What is the difference between a periodic motion and an oscillatory motion?
2. Which of the following examples represent a periodic motion?
   (i) A bullet fired from a gun,
   (ii) An electron revolving round the nucleus in an atom
   (iii) A vehicle moving with a uniform speed on a road
   (iv) A comet moving around the Sun, and
   (v) Motion of an oscillating mercury column in a U-tube.
3. Give an example of (i) an oscillatory periodic motion and (ii)Non-oscillatory periodic motion.

13.2 SIMPLE HARMONIC MOTION : CIRCLE OF REFERENCE

The oscillations of a harmonic oscillator can be represented by terms containing sine and cosine of an angle. If the displacement of an oscillatory particle from its mean position can be represented by an equation $y = a \sin\theta$ or $y = a \cos\theta$ or $y = A \sin\theta + B \cos\theta$, where a, A and B are constants, the particle executes simple harmonic motion. We define simple harmonic motion as under :

A particle is said to execute simple harmonic motion if it moves to and fro about a fixed point periodically, under the action of a force* F *which is directly proportional to its displacement* x *from the fixed point and the direction of the force is opposite to that of the displacement. We shall restrict our discussion to linear oscillations. Mathematically, we express it as

$$F = -kx$$

where k is constant of proportionality.

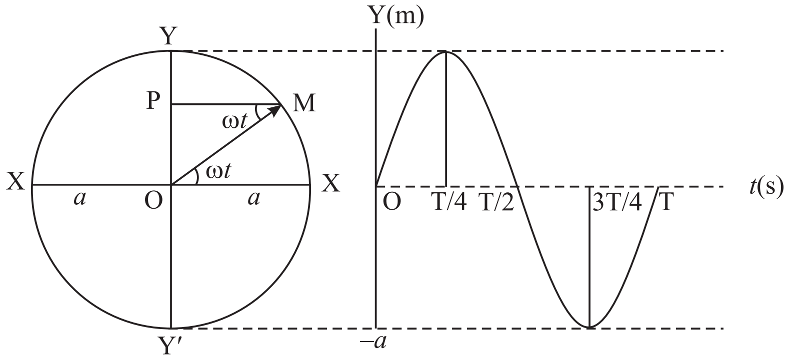


Fig. 13.2 : Simple harmonic motion of P is along YOY′

To derive the equation of simple harmonic motion, let us consider a point M moving with a constant speed υ in a circle of radius a (Fig. 13.2) with centre O. At $t = 0$, let the point be at X. The position vector OM specifies the position of the moving point at time t,. It is obvious that the position vector OM, also called the ***phaser,*** rotates with a constant angular velocity $\omega = \upsilon / a$. The acceleration of the point M is $\upsilon^2/a = a\,\omega^2$ towards the centre O. At time t, the component of this acceleration along OY = $a\omega^2 \sin \omega t$. Let us draw MP perpendicular to YOY′. Then P can be regarded as a particle of mass m moving with an acceleration $a\omega^2 \sin \omega t$. The force on the particle P towards O is therefore given by

$$F = ma\omega^2 \sin \omega t$$

But $\sin \omega t = y/a$. *Therefore*

$$F = m\omega^2 y \qquad (13.3)$$

The displacement is measured from O towards P and force is directed towards O. Therefore,

$$F = -\,m\omega^2 y$$

Since this force is directed towards O, and is proportional to displacement 'y' of P from O. we can say that the particle P is executing simple harmonic motion.

Let us put $m\omega^2 = k$, a constant. Then Eqn. (13.3) takes the form

$$F = -\,k\,y \qquad (13.4)$$

The constant k, which is force per unit displacement, is called ***force constant***. The angular frequency of oscillations is given by

$$\omega^2 = k / m \qquad (13.5)$$

In one complete rotation, OM describes an angle 2π and it takes time T to complete one rotation. Hence

$$\omega = 2\pi/T \qquad (13.6)$$

On combining Eqns. (13.5) and (13.6), we get an expression for time period :

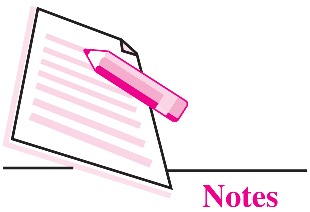
$$T = 2\pi\sqrt{k / m} \qquad (13.7)$$

This is the time taken by P to move from O to Y, then through O to Y′ and back to O. During this time, the particle moves once on the circle and the foot of perpendicular from its position is said to make an oscillation about O as shown in Fig.13.1.

Let us now define the basic terms used to describe simple harmonic motion.

13.2.1 Basic Terms Associated with SHM

Displacement is the distance of the harmonic oscillator from its mean (or equilibrium) position at a given instant.



Amplitude is the maximum displacement of the oscillator on either side of its mean position.

Time period is the time taken by the oscillator to complete one oscillation. In Fig. 13.1, OP, and OY respectively denote displacement and amplitude.

Frequency is the number of oscillations completed by an oscillator in one second. It is denoted by v. The SI unit of frequency is hertz (symbol Hz). Since v is the number of oscillations per second, the time taken to complete one oscillation is $1/v$. Hence $T = 1/v$ or $v = (1/T)\ s^{-1}$. As harmonic oscillations can be represented by expressions containing $\sin\theta$ and or $\cos\theta$, we introduce two more important terms.

Phase ϕ is the angle whose sine or cosine at a given instant indicates the position and direction of motion of the oscillator. It is expressed in radians.

Angular Frequency ω describes the rate of change of phase angle. It is expressed in radian per second. Since phase angle ϕ changes from 0 to 2π radians in one complete oscillation, the rate of change of phase angle is $\omega = 2\pi/T = 2\pi\, v$ or $\omega = 2\pi v$.

Example 13.1 : A tray of mass 9 kg is supported by a spring of force constant k as shown in Fig. 13.3. The tray is pressed slightly downward and then released. It begins to execute SHM of period 1.0 s. When a block of mass M is placed on the tray, the period increases to 2.0 s. Calculate the mass of the block.

Solution: The angular frequency of the system is given by $\omega = \sqrt{k/m}$, where m is the mass of the oscillatory system. Since $\omega = 2\pi/T$, from Eqn. (13.7) we get

$$4\pi^2/T^2 = \frac{k}{m}$$

or

$$m = \frac{kT^2}{4\pi^2}$$

When the tray is empty, $m = 9$ kg and $T = 1$ s. Therefore

$$9 = \frac{k(1)^2}{4\pi^2}$$

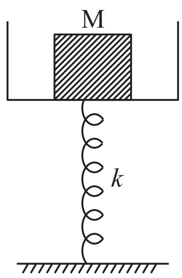


Fig. 13.3

On placing the block, $m = 9 + M$ and $T = 2$ s. Therefore, $9 + M = k \times (2)^2/4\pi^2$

From the above two equations we get

$$\frac{(9+M)}{9} = 4$$

Therefore, $M = 27$ kg.

Example 13.2 : A spring of force constant 1600 N m^{-1} is mounted on a horizontal table as shown in Fig. 13.4. A mass $m = 4.0$ kg attached to the free end of the spring is pulled horizontally towards the right through a distance of 4.0 cm and then set free. Calculate (i) the frequency (ii) maximum acceleration and (iii) maximum speed of the mass.

Solution : $\omega = \sqrt{k/m} = \sqrt{1600/4}$

$= 20 \text{ rad s}^{-1}$.

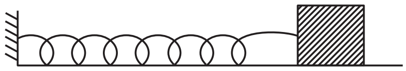


Fig. 13.4

Therefore $v = 20/2\pi = 3.18$ Hz. Maximum acceleration $= a\,\omega^2 = 0.04 \times 400 = 16 \text{ m s}^{-2}$, and $\upsilon_{max} = a\,\omega = 0.04 \times 20 = 0.8 \text{ m s}^{-1}$.

13.3 EXAMPLES OF SHM

In order to clarify the concept of SHM, some very common examples are given below.

13.3.1 Horizontal Oscillations of a Spring-Mass System

Consider a elastic spring of force constant k placed on a smooth horizontal surface and attached to a block P of mass m. The other end of the spring is attached to a rigid wall (Fig. 13.5)). Suppose that the mass of the spring is negligible in comparison to the mass of the block.

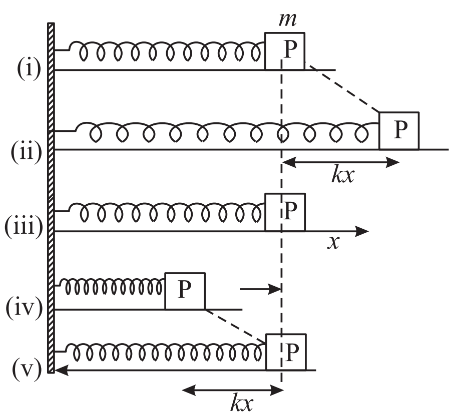


Fig.13.5 : Oscillations of a spring-mass system

Let us suppose that there is no loss of energy due to air resistance and friction. We choose *x–axis* along the horizontal direction. Initially, that is, at $t = 0$, the block is at rest and the spring is in relaxed condition [Fig.13.5(i)]. It is then pulled horizontally through a small distance [Fig. 13.5 (ii)]. As the spring undergoes an extension x, it exerts a force kx on the block. The force is directed against the extension and tends to restore the block to its equilibrium position. As the block returns to its initial position [Fig. 13.5 (iii)], it acquires a velocity υ and hence a kinetic energy $K = (1/2)\, m\, \upsilon^2$. Owing to inertia of motion, the block overshoots the mean position and continues moving towards the left till it arrives at the position shown in Fig. 13.5 (iv). In this position, the block again experiences a force kx which tries to bring it back to the initial position [Fig. 13.5 *v*]. In this way, the block continues oscillating about the mean position. The time period of oscillation is $2\pi\sqrt{m/k}$, where k is the force per unit extension of the spring.

13.3.2 Vertical Oscillations of a Spring–Mass System

Let us suspend a spring of force constant k from a rigid support [Fig.13.6(a)]. Then let us attach a block of mass m to the free end of the spring. As a result of this, the spring undergoes an extension, say l [Fig.13.6(b)]. Obviously, the force constant of the spring is $k = mg/l$. Let us now pull down the block through a small distance, y (Fig.13.6 (c)]. A force ky acts on the block vertically upwards. Therefore, on releasing the block, the force ky pulls it upwards. As the block returns to its initial position, it continues moving upwards on account of the velocity it has gained. It overshoots the equilibrium position by a distance y. The compressed spring now applies on it a restoring force downwards. The block moves downwards and again overshoots the equilibrium position by almost the same vertical distance y. Thus, the system continues to execute vertical oscillations. The angular frequency of vertical oscillations is

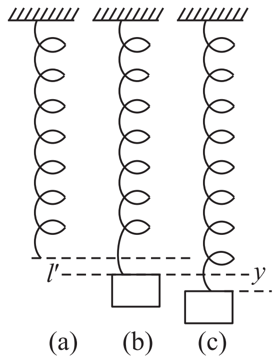


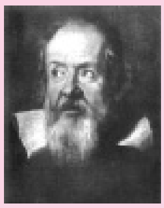
Fig. 13.6: Vertical oscillations of a a spring–mass system

$$\omega = \frac{2\pi}{T} = \sqrt{\frac{k}{m}}$$

Hence
$$T = 2\pi\sqrt{\frac{m}{k}} \qquad (13.8)$$

This result shows that acceleration due to gravity does not influence vertical oscillations of a spring–mass system.

Galileo Galilei
(1564-1642)



Son of Vincenzio Galilei, a wool merchant in Pisa, Italy, Galileo is credited for initiating the age of reason and experimentation in modern science. As a child, he was interested in music, art and toy making. As a young man, he wanted to become a doctor. To pursue the study of medicine, he entered the University of Pisa. It was here that he made his first discovery - the isochronosity of a pendulum, which led Christian Huygen to construct first pendulum clock.

For lack of money, Galileo could not complete his studies, but through his efforts, he learnt and developed the subject of mechanics to a level that the Grand Duke of Tuscany appointed him professor of mathematics at the University of Pisa.

Galileo constructed and used telescope to study celestial objects. Through his observations, he became convinced that Copernican theory of heliocentric universe was correct. He published his convincing arguments in the form of a book, "A Dialogue On The Two Principal Systems of The World", in the year 1632. The proposition being at variance with the Aristotelian theory of geocentric universe, supported by the Church, Galileo was prosecuted and had to apologize. But in 1636, he published another book "Dialogue On Two New Sciences" in which he again showed the fallacy in Aristotle's laws of motion.

Because sophisticated measuring devices were not available in Galileo's time, he had to apply his ingenuity to perform his experiments. He introduced the idea of thought-experiments, which is being used even by modern scientists, in spite of all their sophisticated devices.

13.3.3 Simple Pendulum

A simple pendulum is a small spherical bob suspended by a long cotton thread held between the two halves of a clamped split cork in a stand,as shown in Fig. 13.7. The bob is considered a point mass and the string is taken to be *inextensible*. The Pendulum can oscillate freely about the point of suspension.

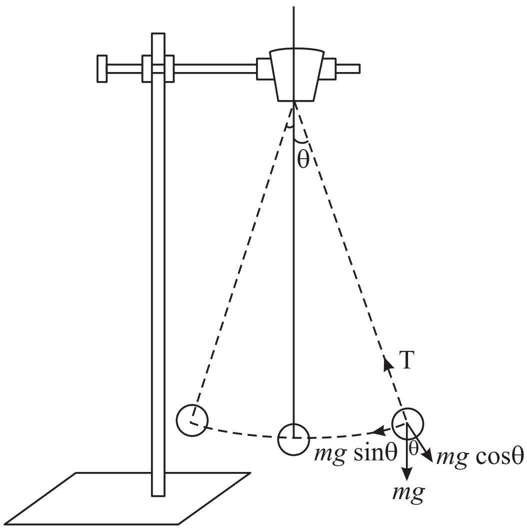


Fig.13.7 : Simple Pendulum

When the pendulum is displaced through a *small distance* from its equilibrium position and then let free, it executes angular oscillations in a vertical plane about its equilibrium position. The distance between the point of suspension and the centre of gravity of the bob defines the length of the pendulum. The forces acting on the bob of the pendulum in the displaced position shown in Fig. 13.7 are : (*i*) the weight of the bob $m\mathbf{g}$ vertically downwards, and (*ii*) tension in the string $\mathbf{T}$ acting upwards along the string.

The weight *mg* is resolved in two components : (a) $mg \cos\theta$ along the string but opposite to $\mathbf{T}$ and (b) $mg \sin\theta$ perpendicular to the string. The component *mg* cosθ balances the tension **T** and the component *mg* sinθ produces acceleration in the bob in the direction of the mean position. The restoring force, therefore, is *mg* sinθ. For small displacement x of the bob, the restoring force is $F = mg\theta = mg\, x/l$. The force per unit displacement $k = mg/l$ and hence

$$\omega = \sqrt{\frac{k}{m}} = \sqrt{\frac{mg/l}{m}} = \sqrt{\frac{g}{l}}$$

or

$$\frac{2\pi}{T} = \sqrt{\frac{g}{l}}$$

Hence,

$$T = 2\pi\sqrt{\frac{l}{g}} \qquad (13.9)$$

Measuring Weight using a Spring

We use a spring balance to measure weight of a body. It is based on the assumption that within a certain limit of load, there is equal extension for equal load, i.e., load/extension remains constant (force constant). Therefore, extension varies linearly with load. Thus you can attach a linear scale alongside the spring and calibrate it for known load values. The balance so prepared can be used to measure unknown weights.

Will such a balance work in a gravity free space, as in a space-rocket or in a satellite? Obviously not becuase in the absence of gravity, no extension occurs in the spring. Then how do they measure mass of astronauts during regular health check up? It is again a spring balance based on a different principle. The astronaut sits on a special chair with a spring attached to each side (Fig.13.8). The time period of oscillations of the chair with and without the astronaut is determined with the help of an electronic clock :

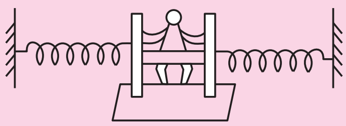


Fig. 13.8 : Spring balance for measuring the mass of an astronaut

$$T_1^2 = \frac{4\pi^2 m}{k}$$

where m is mass of the astronaut. If m_0 is mass of the chair, we can write

$$T_0^2 = \frac{4\pi^2 m_0}{k}$$

T_1 is time period of oscillation of the chair with the astronaut and T_0 without the astronaut.

On subtracting one from another, we get

Notes

$$T_1^2 - T_0^2 = \frac{4\pi^2}{k}(m - m_o)$$

$$\Rightarrow \qquad m = \frac{k}{4\pi^2}(T_1^2 - T_0^2) + m_o$$

Because the values of T_0 and k are fixed and known, a measure of T_1 itself shows the variation in mass.

Example 13.3 : Fig. 13.9 shows an oscillatory system comprising two blocks of masses m_1 and m_2 joined by a massless spring of spring constant k. The blocks are pulled apart, each with a force of magnitude F and then released. Calculate the angular frequency of each mass. Assume that the blocks move on a smooth horizontal plane.

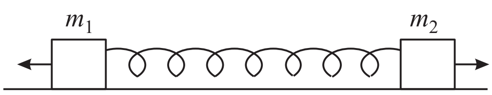


Fig. 13.9 : Oscillatory system of masses attached to a spring

Solution : Let x_1 and x_2 be the displacements of the blocks when pulled apart. The extension produced in the spring is $x_1 + x_2$. Thus the acceleration of m_1 is $k(x_1 + x_2)/m_1$ and acceleration of m_2 is $k(x_1 + x_2)/m_2$. Since the same spring provides the restoring force to each mass, hence the net acceleration of the system comprising of the two masses and the massless spring equals the sum of the acceleration produced in the two masses. Thus the acceleration of the system is

$$a = \frac{k(x_1 + x_2)}{\left(\frac{1}{m_1} + \frac{1}{m_2}\right)} = \frac{k\,x}{\mu}$$

where $x = x_1 + x_2$ is the extension of the spring and μ is the reduced mass of the system. The angular frequency of each mass of the system is therefore,

$$\omega = \sqrt{k/\mu} \qquad (13.10)$$

Such as analysis helps us to understand the vibrations of diatomic molecules like H_2, Cl_2, HCl, etc.

INTEXT QUESTIONS 13.2

1. A small spherical ball of mass m is placed in contact with the sunface on a smooth spherical bowl of radius r a little away from the bottom point. Calculate the time period of oscillations of the ball (Fig. 13.10).

2. A cylinder of mass m floats vertically in a liquid of density ρ. The length of the cylinder inside the liquid is l. Obtain an expression for the time period of its oscillations (Fig. 13.11).

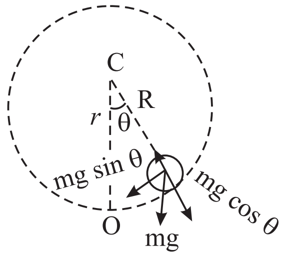


Fig. 13.10

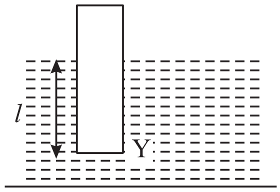


Fig.13.11

Fig. 13.12

3. Calculate the frequency of oscillation of the mass m connected to two rubber bands as shown in Fig. 13.12. The force constant of each band is k. (Fig. 13.12)

13.4 ENERGY OF SIMPLE HARMONIC OSCILLATOR

As you have seen, simple harmonic motion can be represented by the equation

$$y = a \sin \omega t \qquad (13.11)$$

When t changes to $\mathbf{t} + \Delta t$, y changes to $y + \Delta y$. Therefore, we can write

$$y + \Delta y = a \sin\omega (t + \Delta t) = a \sin (\omega t + \omega \Delta t)$$

$$= a [\sin\omega t \cos \omega\Delta t + \cos\omega t \sin \omega\Delta t]$$

As $\Delta t \rightarrow 0$, $\cos \omega\Delta t \rightarrow 1$ and $\sin \omega \Delta t \rightarrow \omega \Delta t$. Then

$$y + \Delta y = a \sin \omega t + a \omega\Delta t \cos \omega t. \qquad (13.12)$$

Subtracting Eqn. (13.11) from Eqn. (13.12), we get

$$\Delta y = \Delta t \, \omega a \cos \omega t$$

so that $$\Delta y/\Delta t = \omega a \cos \omega t$$

or $$\upsilon = \omega a \text{ cost } \omega t \qquad (13.13)$$

where $\upsilon = \Delta y/\Delta t$ is the velocity of the oscillator at time t. Hence, the ***kinetic energy*** of the oscillator at that instant of time is

$$K = (1/2) \, m\upsilon^2 = (1/2) \, \omega^2 a^2 \cos^2 \omega t \qquad (13.14)$$

Let us now calculate the potential energy of the oscillator at that time. When the displacement is y, the restoring force is ky, where k is the force constant. For this purpose we shall plot a graph of restoring force ky versus the displacement y. We get a straight line graph as shown in Fig. 13.13. Let us take two points P and Q and drop perpendiculars PM and QN on x–axis. As points P and Q are close to each other, trapezium PQNM can be regarded as a rectangle. The area of this rectangular strip is ($ky\,\Delta y$). This area equals the work done against the restoring force ky when the displacement changes by a small amount Δy. The area of the triangle OBC is, therefore, equal to the work done in the time displacement changes from O to OB (= y) = $\frac{1}{2}ky^2$. This work done against the conservative force is the ***potential energy*** U of the oscillator. Thus, the potential energy of the oscillator when the displacement is y is

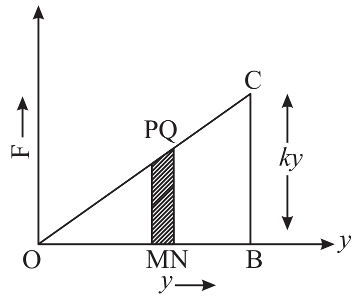


Fig.13.13 : Graph between the displacement y and the restoring force ky

$$U = \frac{1}{2}ky^2$$

But $\omega^2 = k/m$. Therefore, substituting $k = m\omega^2$ in above expression we get

$$U = \frac{1}{2}m\omega^2 y^2$$

Further as $y = a \sin \omega t$, we can write

$$U = \frac{1}{2}m\omega^2 a^2 \sin^2\omega t \qquad (13.15)$$

On combining this result with Eqn. (13.14), we find that total energy of the oscillator at any instant is given by

$$E = U + K$$

$$= \frac{1}{2}m\omega^2 a^2 (\sin^2\omega t + \cos^2\omega t)$$

$$= \frac{1}{2}ma^2\omega^2 \qquad (13.16)$$

The graph of kinetic energy K, potential energy U and the total energy E versus displacement y is shown in Fig.13.14. From the graph it is evident that for $y = 0$, $K = E$ and $U = 0$. As the displacement y from the mean position increases, the kinetic energy decreases but potential energy increases. At the mean position, the potential energy is zero but kinetic energy is maximum. At the extreme positions, the energy is wholly potential. However, the sum $K + U = E$ is constant.

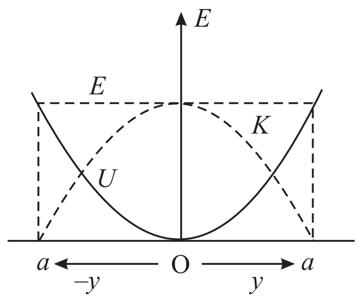


Fig.13.14 : Variation of potential energy U, kinetic energy K, and total energy E with displacement from equilibrium position

INTEXT QUESTIONS 13.3

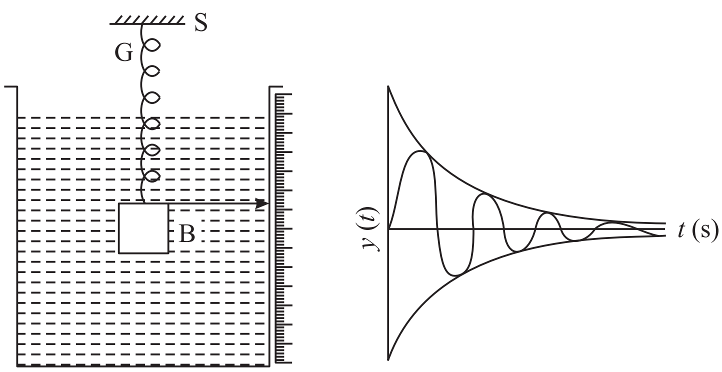
1. Is the kinetic energy of a harmonic oscillator maximum at its equilibrium position or at the maximum displacement position? Where is its acceleration maximum?
2. Why does the amplitude of a simple pendulum decrease with time? What happens to the energy of the pendulum when its amplitude decreases?

13.5 DAMPED HARMONIC OSCILLATIONS

Every oscillating system normally has a viscous medium surrounding it. As a result in each oscillation some of its energy is dissipated as heat. As the energy of oscillation decreases the amplitude of oscillation also decreases. The amplitude of oscillations of a pendulum in air decreases continuously. Such oscillations are called damped oscillations. To understand **damped oscillations** perform activity 13.2.

Activity 13.2

Take a simple harmonic oscillator comprising a metal block B suspended from a fixed support S by a spring G. (Fig. 13.15(a). Place a tall glass cylinder filled two thirds with water, so that the block is about 6 cm below the surface of water and about the same distance above the bottom of the beaker. Paste a millimetre scale (vertically) on the side of the cylinder just opposite the pointer attached to the block. Push the block a few centimetres downwards and then release it. After each oscillation, note down the uppermost position of the pointer on the millimetre scale and the time. Then plot a graph between time and the amplitude of oscillations. Does the graph [Fig. 13.15 (b)] show that the amplitude decreases with time. Such oscillations are said to be ***damped oscillations***.



(a) (b)

Fig. 13.15 : Damped vibrations : (a) experimental setup; (b) graphical representation

13.6 FREE AND FORCED VIBRATIONS : RESONANCE

To understand the difference between these phenomena, let us perform the following activity :

ACTIVITY 13.3

Take a rigid horizontal rod fixed at both ends. Tie a loose but strong thread and hang the four pendulums A,B,C,D, as shown in Fig. 13.16. The pendulums A and B are of equal lengths, whereas C has a shorter and D has a longer length than A and B. The pendulum B has a heavy bob. Set pendulum B into oscillations. You will observe that after a few minutes, the other three pendulums also begin to oscillate. (It means that if a no. of oscillators are coupled, they transfer their energy. This has an extremely important implication for wave propagation.) You will note that the amplitude of A is larger. Why? Each pendulum is an oscillatory system with natural frequency of its own. The pendulum B, which has a heavy bob, transmits its vibrations to each of the pendulums A, C and D. As a consequence, the pendulums C and D are forced to oscillate *not with their respective natural* frequency *but with the frequency of the pendulum B.* The phenomenon is called ***forced oscillation.*** By holding the bob of any one of these pendulums, you can force it to oscillate with the frequency of C or of D. Both C and D are forced to oscillate with the frequency of B. However, pendulum A on which too the oscillations of the pendulums B are impressed, oscillates with a relatively large amplitude with ***its natural frequency.*** This phenomenon is known as **resonance.**

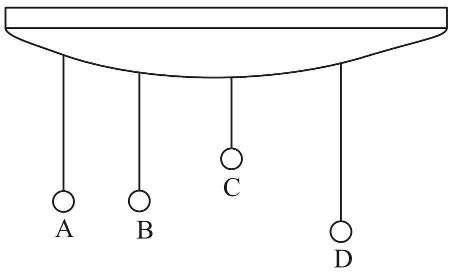


Fig. 13.16: Vibrations and resonance.

When the moving part of an oscillatory system is displaced from its equilibrium position and then set free, it oscillates to and fro about its equilibrium position with a frequency that depends on certain parameters of the system only. Such oscillations are known as ***free vibrations.*** The frequency with which the system oscillates is known as ***natural frequency.*** When a body oscillates under the influence of an external periodic force, the oscillations are called ***forced oscillations.*** In forced oscillations, the body ultimately oscillates with the frequency of the external force. The oscillatory system on which the oscillations are impressed is called driven and the system which applies the oscillating force is known as the driver. The particular case of forced oscillations in which natural frequencies of the driver and the driven are equal is known as ***resonance.*** In resonant oscillations, the driver and the driven reinforce each other's oscillations and hence their amplitudes are maximum.

INTEXT QUESTIONS 13.4

1. When the stem of a vibrating tuning fork is pressed against the top of a table, a loud sound is heard. Does this observation demonstrate the phenomenon of resonance or forced vibrations? Give reasons for your answer. What is the cause of the loud sound produced?

2. Why are certain musical instruments provided with hollow sound boards or sound boxes?

Mysterious happenings and resonance

1. Tacoma Narrows Suspension Bridge, Washington, USA collapsed during a storm within six months of its opening in 1940. The wind blowing in gusts had frequency equal to the natural frequency of the bridge. So it swayed the bridge with increasing amplitude. Ultimately a stage was reached where the structure was over stressed and it collapsed.

   The events of suspension bridge collapse also happened when groups of marching soldiers crossed them. That is why, now, the soldiers are ordered to break steps while crossing a bridge.

   The factory chimneys and cooling towers set into oscillations by the wind and sometimes get collapsed.

2. You might have heard about some singers with mysterious powers. Actually, no such power exists. When they sing, the glasses of the window panes in the auditorium are broken. They just sing the note which matches the natural frequency of the window panes.

3. You might have wondered how you catch a particular station you are interested in by operating the tuner of your radio or TV? The tuner in fact, is an electronic oscillator with a provision of changing its frequency. When the frequency of the tuner matches the frequency transmitted by the specific station, resonance occurs and the antenna catches the programme broadcasted by that station.

WHAT YOU HAVE LEARNT

- Periodic motion is a motion which repeats itself after equal intervals of time.
- Oscillatory motion is to and fro motion about a mean position. An oscillatory motion is normally periodic but a periodic motion may not *necessarily be oscillatory*.

- Simple harmonic motion is to and fro motion under the action of a restoring force, which is proportional to the displacement of the particle from its equilibrium position and is always directed towards the mean position.
- Time period is the time taken by a particle to complete one oscillation.
- Frequency is the number of vibrations completed by the oscillator in one second.
- Phase angle is the angle whose sine or cosine at the given instant indicates the position and direction of motion of the particle.
- Angular frequency is the rate of change of phase angle. Note that $\omega = 2\pi/T = 2\pi\nu$ where ω is the angular frequency in rads^{-1}, ν is the frequency in hertz (symbol : Hz) and T is the time period in seconds.
- Equation of simple harmonic motion is

$$y = a \sin(\omega t + \phi_0)$$

or

$$y = a \cos(\omega t + \phi_0)$$

where y is the displacement from the mean position at a time, ϕ_0 is the initial phase angle (at $t = 0$).

- When an oscillatroy system vibrates on its own, its vibrations are said to be free. If, however, an oscillatory system is driven by an external system, its vibrations are said to be forced vibrations. And if the frequency of the driver equals to the natural frequency of the driven, the phenomenon of resonance is said to occur.

TERMINAL EXERCISE

1. Distinguish between a periodic and an oscillatory motion.
2. What is simple harmonic motion?
3. Which of the following functions represent (i) a simple harmonic motion (ii) a periodic but not simple harmonic (iii) a non periodic motion? Give the period of each periodic motion.

   (1) $\sin \omega t + \cos \omega t$ (2) $1 + \omega^2 + \omega t$

   (3) $3 \cos (\omega t - \frac{\pi}{4})$
4. The time period of oscillations of mass 0.1 kg suspended from a Hooke's spring is 1s. Calculate the time period of oscillation of mass 0.9 kg when suspended from the same spring.
5. What is phase angle? How is it related to angular frequency?

Notes

6. Why is the time period of a simple pendulum independent of the mass of the bob, when the period of a simple harmonic oscillator is $T = 2\pi\sqrt{m/k}$?

7. When is the magnitude of acceleration of a particle executing simple harmonic motion maximum? When is the restoring force maximum?

8. Show that simple harmonic motion is the projection of a uniform circular motion on a diameter of the circle. Obtain an expression for the time period of a simple harmonic oscillator in terms of mass and force constant.

9. Obtain expressions for the instantaneous kinetic energy potential energy and the total energy of a simple harmonic oscillator.

10. Show graphically how the potential energy U, the kinetic energy K and the total energy E of a simple harmonic oscillator vary with the displacement from equilibrium position.

11. The displacement of a moving particle from a fixed point at any instant is given by $x = a \cos \omega t + b \sin \omega t$. Is the motion of the particle simple harmonic? If your answer is no, explain why? If your answer is yes, calculate the amplitude of vibration and the phase angle.

12. A simple pendulum oscillates with amplitude 0.04 m. If its time period is 10 s, calculate the maximum velocity.

13. Imagine a ball dropped in a frictionless tunnel cut across the earth through its centre. Obtain an expression for its time period in terms of radius of the earth and the acceleration due to gravity.

14. Fig. 13.17 shows a block of mass $m = 2$ kg connected to two springs, each of force constant $k = 400$ N m^{-1}. The block is displaced by 0.05 m from equilibrium position and then released. Calculate (a) The angular frequency ω of the block, (b) its maximum speed; (c) its maximum acceleration; and total energy dissipated against damping when it comes to rest.

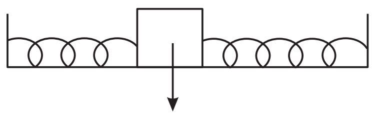


Fig.13.17

ANSWERS TO INTEXT QUESTIONS

13.1

1. A motion which repeats itself after some fixed interval of time is a periodic motion. A to and fro motion on the same path is an oscillatory motion. A periodic motion may or may not be oscillatory but oscillation motion is perodic.

2. (ii), (iv), (v);

3. (i) To and fro motion of a pendulum.

   (ii) Motion of a planet in its orbit.

13.2

1. Return force on the ball when displaced a distance x from the equilibrium position is $mg \sin \theta = mg\, \theta = mg\, x/r$. $\therefore\ \omega = \sqrt{g/r}$.

2. On being pushed down through a distance y, the cylinder experiences an upthrust $y\alpha\rho g$. Therefore $\omega^2 = \dfrac{\alpha\rho g}{m}$ and $m = \alpha p\rho$. From the law of flotation m = mass of black. Hence, $\omega^2 = g/l$ or $T = 2\pi\ \sqrt{l/g}$.

3. $\omega^2 = k/m$ and hence $v = 1/2\pi\sqrt{k/m}$. Note that when the mass is displaced, only one of the bands exerts the restoring force.

13.3

1. K.E is maximum at mean position or equilibrium position; acceleration is maximum when displacement is maximum.

2. As the pendulum oscillates it does work against the viscous resistance of air and friction at the support from which it is suspended. This work done is dissipated as heat. As a consequence the amplitude decreases.

13.4

1. When an oscillatory system called the driver applies is periodic of force on another oscillatory system called the driven and the second system is forced to oscillate with the frequency of the first, the phenomenon is known as forced vibrations. In the particular case of forced vibrations in which the frequency of the driver equals the frequency of the driven system, the phenomenon is known as resonance.

2. The table top is forced to vibrate not with its natural frequency but with the frequency of the tuning fork. Therefore, this observation demonstrates forced vibrations. Since a large area is set into vibrations, the intensity of the sound increases.

3. The sound board or box is forced to vibrate with the frequency of the note produced by the instrument. Since a large area is set into vibrations, the intensity of the note produced increases and its duration decreases.

Answers to Terminal Problems

4. 3s

11. $A = \sqrt{a^2 + b^2}, \theta = \tan^{-1}\left(\frac{a}{b}\right)$

12. $\frac{2}{\pi} \times 10^{-3}$ m s^{-1}

14. (*a*) 14.14 s^{-1}

    (*b*) 0.6 m s^{-1}

    (*c*) 0.3 m s^{-2}

    (*d*) 0.5 J

312en14

14

WAVE PHENOMENA

You would have noticed that when a stone is dropped into still water in a pond, concentric rings of alternate elevations and depressions emerge out from the point of impact and spread out on the surface of water. If you put a straw piece on the surface of water, you will observe that it moves up and down at its place. Here the particles of water are moving up and down at their places. But still there is something which moves outwards. We call it a *wave*. Waves are of different types : Progressive and stationary, mechanical and electro-magnetic. These can also be classified as longitudinal and transverse depending on the direction of motion of the material particles with respect to the direction of propagation of wave in case of mechnical waves and electric and magnetic vectors in case of e.m. waves. Waves are so intimate to our existence.

Sound waves travelling through air make it possible for us to listen. Light waves, which can travel even through vacuum make us see things and radio waves carrying different signals at the speed of light connect us to our dear ones through differents forms of communication. In fact, wave phenomena is universal.

The working of our musical instruments, radio, T.V require us to understand wave phenomena. Can you imagine the quality of life without waves? In this lesson you will study the basics of waves and wave phenomena.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain propagation of transverse and longitudinal waves and establish the relation* $\upsilon = \nu\lambda$;
- *write Newton's formula for velocity of longitudinal waves in a gas and explain Laplace's correction;*

- *discuss the factors on which velocity of longitudinal waves in a gas depends;*
- *explain formation of transverse waves on stretched strings;*
- *derive the equation of a simple harmonic wave;*
- *explain the phenomena of beats, interference and phase change of waves on the basis of principle of superposition*
- *explain formation of stationary waves and discuss harmonics of organ pipes and stretched strings;*
- *discuss Doppler effect and apply it to mechanical and optical systems;*
- *explain the properties of em waves, and*
- *state wavelength range of different parts of em spectrum and their applications.*

14.1 WAVE PROPAGATION

From the motion of a piece of straw, you may think that waves carry energy; these do not transport mass. A vivid demonstration of this aspect is seen in tidal waves. Do you recall the devastation caused by Tsunami waves which hit Indonesia, Thailand, Sri lanka and India caused by a deep sea quake waves of 20 m height were generated and were responsible for huge loss of life.

To understand how waves travel in a medium let us perform an activity.

ACTIVITY 14.1

Take a long coiled spring, called slinky, and stretch it along a smooth floor or bench, keeping one end fixed and the other end free to be given movements . Hold the free end in your hand and give it a jerk side–ways.[Fig 14.1(a)]. You will observe that a kink is produced which travels towards the fixed end with definite speed . This kink is a wave of short duration. Keep moving the free end continuously left and right. You will observe a train of pulses ravelling towards the fixed end. This is a *transverse wave* moving through the spring [Fig. 14.1 (b)].

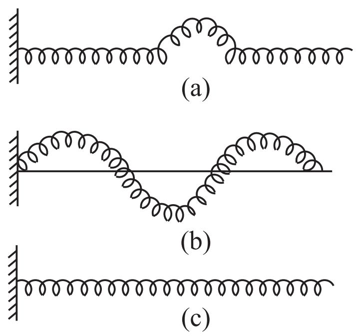


Fig. 14.1 : Wave motion on a slinky (a) pulse on a slinky, (b) transverse wave, and (c) longitudinal Wave

There is another type of wave that you can generate in the slinky . For this keep the slinky straight and give it a push along its length . A pulse of compression thus moves on the spring. By moving the hand

Notes

backwards and forwards at a constant rate you can see ulternate compressions and rarefactions travelling along its length . These are called *longitudinal waves* [Fig. 14.1(c)].

14.1.1 Propagation of Transverse Waves

Refer to Fig 14.2. It shows a mechanical model for wave propapation. It comprises a row of spherical balls of equal masses, evenly spaced and connected together by identical springs. Let us imagine that by means of suitable device, ball-1, from left, is made to execute S.H.M. in a direction perpendicular to the row of balls with a period T. All the balls, owing to inertia of rest will not begin to oscillate at the same time. The motion is passed on from one ball to the next one by one. Let

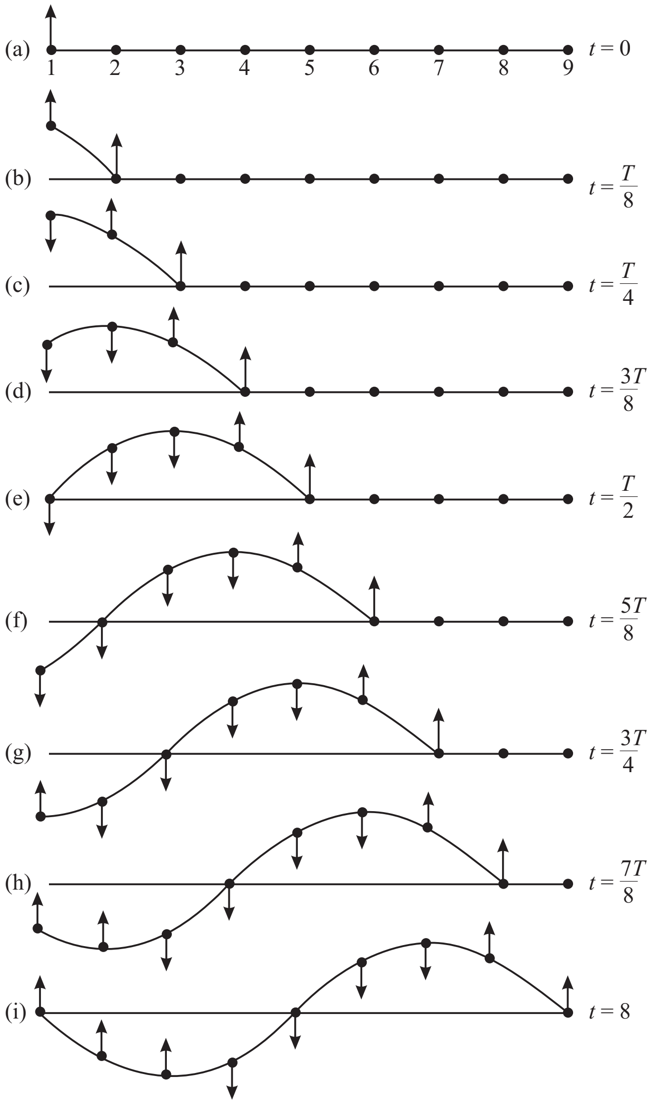


Fig. 14.2 : Instantaneous profiles at intervals of T/8 when a transverse wave is generated on a string.

us suppose that the time taken by the disturbance to travel from one ball to the next is $T/8$s. This means that in the interval $T/8$s, the disturbance propagates from the particle at mark 1 to the particale at mark 2. Similarly, in the next $T/8$ interval, the disturbance travels from the particle at mark 2 to the particle at mark 3 and so on. In parts (a)—(i) in Fig. 14.2 we have shown the instantaneous positions of particles at all nine marked positions at intervals of $T/8$. (The arrows indicate the directions of motion along which particles at various marks are about to move.) You will observe that

(i) At $t = 0$, all the particles are at their respective mean positions.

(ii) At $t = T$, the first, fifth and ninth particles are at their respective mean positions. The first and ninth particles are about to move upward whereas the fifth particle is about to move downward. The third and seventh particles are at position of maximum displacement but on opposite sides of the horizontal axis. The envelop joining the instantaneous positions of all the particles at marked positions in Fig. 14.2(a) are similar to those in Fig. 14.2(i) and represents a *transverse wave*. The positions of third and seventh particles denote a *trough and a crest*, respectively.

The important point to note here is that *while the wave moves along the string, all particles of the string are oscillating up and down about their respective equilibrium positions with the same period* (T) *and amplitude* (A). This wave remains *progressive* till it reaches the fixed end.

In a wave motion, the distance between the two nearest particles vibrating in the same phase is called a wavelength. *It is denoted by λ.*

It is evident that time taken by the wave to travel a distance λ is T. (See Fig. 14.2). Therefore, the velocity of the wave is

$$\upsilon = \frac{\text{Distance}}{\text{Time}} = \frac{\lambda}{T} \quad (14.1)$$

But, $1/T = \nu$, the frequency of the wave. Therefore,

$$\upsilon = \nu\lambda \quad (14.2)$$

Further, if two consecutive particles in same state of motion are separated by a distance λ, the phase difference between them is 2π. Therefore, the phase change per unit distance

$$k = \frac{2\pi}{\lambda} \quad (14.3)$$

We call k the propagation constant. You may recall that ω denotes phase change per unit time. But the phase change in time T is 2π hence

$$\omega = \frac{2\pi}{T} = 2\pi\nu \quad (14.4)$$

Dividing Eqn. (14.3) by Eqn. (14.4), we get an expression for the wave velocity:

$$\upsilon = \frac{\omega}{k} = \frac{2\pi\nu}{2\pi/\lambda}$$

or

$$\upsilon = \nu\lambda \qquad (14.5)$$

Let us now explain how the logitudinal waves propagate.

14.1.2 Propagation of a Longitudinal Wave

In longitudinal waves, the displacement of particles is along the direction of wave propagation In Fig. 14.3, the hollow circles represent the mean positions of equidistant particles in a medium. The arrows show their (rather magnified) longitudinal displacements at a given time. You will observe that the arrows are neither equal in length nor in the same direction. This is clear from the position of solid circles, which describe instantaneous positions of the particles corresponding to the heads of the arrows. The displacements to the right are shown in the graph towards + *y*-axis and displacements to the left towards the –*y*-axis.

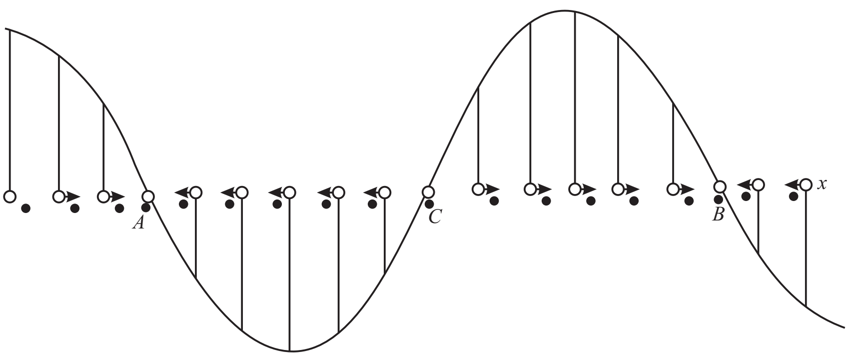


Fig. 14.3 : Graphical representation of a longtudinal wave.

For every arrow directed to the right, we draw a proportionate line upward. Similarly, for every arrow directed to the left, a proportionate line is drawn downward. On drawing a smooth curve through the heads of these lines, we find that the graph resembles the displacement-time curve for a transverse wave. If we look at the solid circles, we note that around the positions A and B, the particles have crowded together while around the position C, they have separated farther. These represent regions of *compression and rarefaction.* That is, there are alternate regions where density (pressure) are higher and lower than average.A sound wave propagating in air is very similar to the longitudinal waves that you can generate on your spring (Fig. 14.4).

Let us now derive equation of a simple harmonic wave.

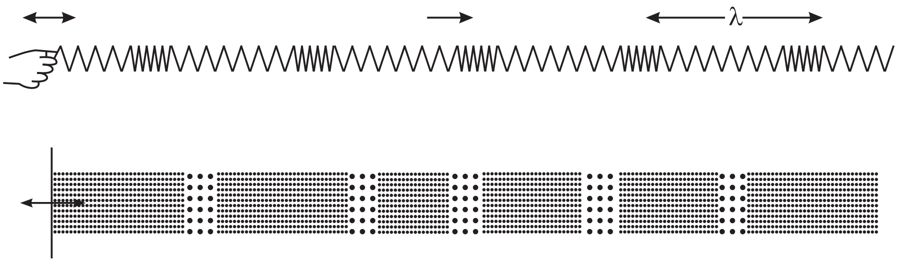


Fig. 14.4 : Longitudinal waves on a spring are analogous to sound waves.

14.1.3 Equation of a Simple Harmonic Wave in One Dimension

Let us consider a simple harmonic wave propagating along OX (Fig. 14.5). We assume that the wave is transverse and the vibrations of the particle are along YOY′. Let us represent the displacement at $t = 0$ by the equation

$$y = a \sin \omega t \tag{14.6}$$

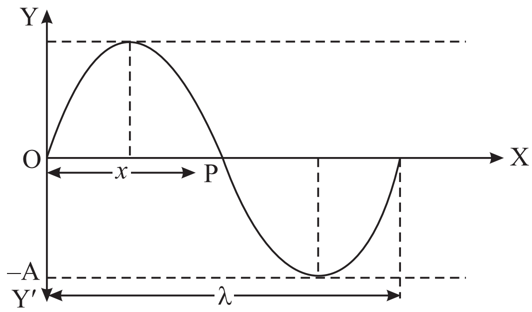


Fig. 14.5 : Simple harmonic wave travelling along x-direction

Then the phase of vibrations at that time at the point P lags behind by a phase, say ϕ. Then

$$y = a \sin (\omega t - \phi) \tag{14.7}$$

Let us put OP = x. Since phase change per unit distance is k, we can write $\phi = kx$. Hence,

Eqn. (14.7) take the form $\quad y(x, t) = a \sin (\omega t - kx) \qquad (14.8)$

Further as $\omega = 2\pi/t$ and $k = 2\pi/\lambda$, we can rewrite Eqn (14.8) as

$$y(x, t) = a \sin 2\pi \left(\frac{t}{T} - \frac{x}{\lambda}\right) \tag{14.9}$$

In terms of wave velocity ($\upsilon = \lambda/T$), this equation can be expressed as

$$y = a \sin \frac{2\pi}{\lambda} (\upsilon t - x) \tag{14.10}$$

In deriving Eqn. (14.8) we have taken initial phase of the wave at O as zero. However, if the initial phase angle at O is ϕ_0, the equation of the wave would be

$$y\,(x,t) = a \sin [(\omega t - kx) + \phi_0] \qquad (14.11)$$

Phase difference between two points on a wave

Let us consider two simple harmonic waves travelling along OX and represented by the equations

$$y = a \sin (\omega t - kx) \qquad (14.11a)$$

and

$$y = a \sin [\omega t - k\,(x + \Delta x)] \qquad (14.12)$$

The phase difference between them is

$$\Delta\phi = k\Delta x = \frac{2\pi}{\lambda}\Delta x = -\frac{2\pi}{\lambda}(x_2 - x_1) \qquad (14.13)$$

where Δx is called the path difference between these two points. Here the negative sign indicates that a point positioned later will acquire the same phase at a later time.

Phase difference at the same position over a time interval Δt :

We consider two waves at the same position at a time interval Δt. For the first wave, phase ϕ, is given by

$$\phi_1 = \frac{2\pi}{T}t_1 - \frac{2\pi}{\lambda}x$$

and for the another wave phase

$$\phi_2 = \frac{2\pi}{T}t_2 - \frac{2\pi}{\pi}x.$$

The phase difference between them

$$\Delta\phi = \phi_2 - \phi_1 = \frac{2\pi}{T}(t_2 - t_1)$$

$$= 2\pi v\,(t_2 - t_1) \qquad [14.13(a)]$$

$$= 2\pi v\,(\Delta t)$$

Example 14.1 : A progressive harmonic wave is given by $y = 10^{-4} \sin (100\pi t - 0.1\pi x)$. Calculate its (i) frequency, (ii) wavelength and (iii) velocity y and x are in metre.

Solution: comparing with the standard equation of progressive wave

$$y = A \sin \left(\frac{2\pi t}{T} - \frac{2\pi x}{\lambda}\right)$$

we get (i) $2\pi v = 100\,\pi \;\Rightarrow\; v = 50$ Hz

(ii) $\dfrac{2\pi}{\lambda} = 0.1\,\pi \;\Rightarrow\; \lambda = 20$ m

(iii) $\upsilon = v\lambda = 1000\ \text{ms}^{-1}$

14.1.4 Transverse and Longitudinal Waves

We now consider transverse and longitudinal waves and summerise the difference between them.

Transverse waves	Longitudinal waves
(i) Displacements of the particles are perpendicular to the direction of propagation of the wave.	(i) Displacements of the particles are along the direction of propagation of the wave.
(ii) Transverse waves look as crests and troughs propagating in the medium.	(ii) Longitudinal waves give the appearance of alternate compressions and rarefaction moving forward.
(iii) Transverse waves can only be transmitted in solids or on the surface of the liquids.	(iii) Longitudinal waves can travel in solids, liquids and gases.
(iv) In case of a transverse wave, the displacement - distance graph gives the actual picture of the wave itself.	(iv) In case of longitudinal waves, the graph only represents the displacement of the particles at different points at a given time.

Essential properties of the medium for propagation of longitudinal and transverse mechanical waves are: (i) the particles of the medium must possess mass, (ii) the medium must possess elasticity. Longitudinal waves for propagation in a medium require volume elasticity but transverse waves need modulus of rigidity. However, light waves and other electromagnetic waves, which are transverse, do not need any material medium for their propagation.

INTEXT QUESTIONS 14.1

1. State the differences between longitudinal and transverse waves?
2. Write the relation between phase difference and path difference.
3. Two simple harmonic waves are represented by equations $y_1 = a \sin(\omega t - kx)$ and $y_2 = a \sin[(\omega t - kx) + \phi]$. What is the phase difference between these two waves?

14.2 VELOCITY OF LONGITUDINAL AND TRANSVERSE WAVES IN AN ELASTIC MEDIUM

14.2.1 Newton's Formula for Velocity of Sound in a Gas

Newton to derive a relation for the velocity of sound in a gaseous medium, assumed that compression and rarefaction caused by the sound waves during their passage through the gas take place under isothermal condition. This means that the changes in volume and pressure take place at constant temperature. Under such conditions, Newton agreed that the velocity of sound wave in a gas is given by

$$\upsilon = \sqrt{\frac{P}{\rho}} \qquad (14.15)$$

For air, at standard temperature and pressure $P = 1.01 \times 10^5\ Nm^{-2}$ and $\rho = 1.29$ kg m^{-3}. On substituting these values in Eqn.(14.15) we get

$$\upsilon = \sqrt{1.01 \times 10^5 / 1.29} = 280\ ms^{-1}$$

Clouds collide producing thunder and lightening, we hear sound of thunder after the lightening. This is because the velocity of light is very much greater than the velocity of sound in air. By measuring the time interval between observing the lightening and hearing the sound, the velocity of sound in air can be determined. Using an improved technique, the velocity of sound in air has been determined as 333 ms^{-1} at 0^0C. The percent error in the value predicted by Newton's formula and that determined experimentally is $\frac{333-280}{333} \times 100\% = 16\%$. This error is too high to be regarded as an experimental error. Obviously there is something wrong with Newton's assumption that during the passage of sound, the compression and the rarefaction of air take place isothermally.

14.2.2 Laplace's Correction

Laplace pointed out that the changes in pressure of air layers caused by passage of sound take place under adiabatic condition owing to the following reasons.

(i) Air is bad conductor of heat and

(ii) Compression and rarefactions caused by the sound are too rapid to permit heat to flow out during compression and flow in during rarefaction.

Under adiabatic conditions

$$E = \gamma P,$$

Where $$\gamma = \frac{C_p}{C_v}$$

Hence, $$v = \sqrt{\frac{\gamma P}{\rho}} \qquad (14.16)$$

For air, $\gamma = 1.4$. Therefore, at STP, speed of sound is given by

$$v = \sqrt{1.4 \times 1.01 \times 10^5 / 1.29}$$

$$= 333 \text{ms}^{-1}$$

This value is very close to the experimentally observed value.

14.2.3 Factors affecting velocity of sound in a gas

(i) Effect of Temperature

From Laplace's formula

$$v = \sqrt{\frac{\gamma P}{\rho}}$$

Since density is ratio of mass perumit volume, this expression takes the form

$$= \sqrt{\frac{\gamma PV}{M}}$$

Using the equation of state PV = nRT, where n is number of moles in mass m of the gas

$$v = \sqrt{\frac{\gamma RT}{\frac{M}{n}}}$$

$$= \sqrt{\frac{\gamma RT}{m}} \qquad (14.17\text{ a})$$

Where m denotes the gram molecular mass. This result shows that

$$v \;\alpha\; \sqrt{T}$$

$$\Rightarrow \qquad v = v\text{o}\left(1 + \frac{t}{2 \times 273}\right) + \ldots\ldots\ldots\ldots\ldots$$

$$\simeq 333 + \frac{333}{546} t$$

$$\simeq 333 + 0.61t \qquad (14.17\text{b})$$

Note that for small temperature variations, velocity of sound in air increases by 0.61 ms^{-1} with every degree celsius rise in temperature.

(*ii*) Effect of pressure

When we increase pressure on a gas, it gets compressed but its density increases in the same proportion as the pressure i.e. P/ρ remains constant. It means that, pressure has no effect on the velocity of sound in a gas.

(*iii*) Effect of density

If we consider two gases under identical conditions of temperature and pressure, then

$$\upsilon \; \alpha \; \frac{1}{\sqrt{\rho}}$$

If we, compare the velocities of sound in oxygen and hydrogen, we get

$$\frac{\upsilon_{\text{oxygen}}}{\upsilon_{\text{hydrogen}}} = \sqrt{\frac{\rho_{\text{hydrogen}}}{\rho_{\text{oxygen}}}} = \sqrt{\frac{M_{\text{hydrogen}}}{M_{\text{oxygen}}}} = \sqrt{\frac{2}{32}} = \frac{1}{4}$$

This shows that velocity of sound in hydrogen is 4 times the velocity of sound in oxygen under identical conditions of temperature and pressure. Is this result valid for liquids and solids as well. You will discover answer to this question in the next sub–section.

(*iv*) Effect of humidity on velocity of sound in air

As humidity in air increases (keeping conditions of temperature and pressure constant), its density decreases and hence velocity of sound in air increases.

Example 14.2 : At what temperature is the speed of sound in air double of its value at S.T.P.

Solution : We know that $\frac{\upsilon}{\upsilon_0} = \sqrt{\frac{T}{m}} = 2 = \sqrt{\frac{T}{273}}$

On squaring both sides and rearranging terms, we get

∴ $$T = 273 \times 4 = 1092\text{k}$$

14.2.4 Velocity of Waves in Stretched Strings

The velocity of a transverse wave in a stretched string is given by

$$\upsilon = \sqrt{\frac{F}{m}} \qquad (14.18\text{ a})$$

Where F is tension in the string and *m* is mass per unit length of the wire. The velocity of longtudinal waves in an elastic medium is given by

$$\upsilon = \sqrt{E/\rho} \qquad (14.18\text{b})$$

where E is elasticity. It may be pointed out here that since the value of elasticity is more in solids, the velocity of longitudinal waves in solids is greater than that in gases and liquids. In fact, $\upsilon_g < \upsilon_\ell < \upsilon_s$.

INTEXT QUESTIONS 14.2

1. What was the assumption made by Newton in deriving his formula?
2. What was wrong with Newton's formula?
3. Show that for every 1^0C rise in temperature, the velocity of sound in air increases by 0.61 ms^{-1}.
4. Calculate the temperature at which the velocity in air is (3/2) times the velocity of sound at 7^0C?
5. Write the formula for the velocity of a wave on stretched string?
6. Let λ be the wavelength of a wave on a stretched string of mass per unit length m and n be its frequency. Write the relation between n, λ, F and m? Further if $\lambda = 2\ell$, what would be the relation between n, l, F and m?

14.3 SUPERPOSITION OF WAVES

Suppose two wave pulses travel in opposite directions on a slinky. What happens when they meet? Do they knock each other out? To answer these questions, let us perform an activity.

ACTIVITY 14.2

Produce two wave crests of different amptitudes on a stretched slinky, as shown in Fig. 14.6 and watch carefully. The crests are moving in the opposite directions. They meet and overlap at the point midway between them [Fig. 14.6(b)] and then separate out. Thereafter, they continue to move in the same direction in which they were moving before crossing each other. Moreover, their shape also does not change [Fig. 14.6(c)].

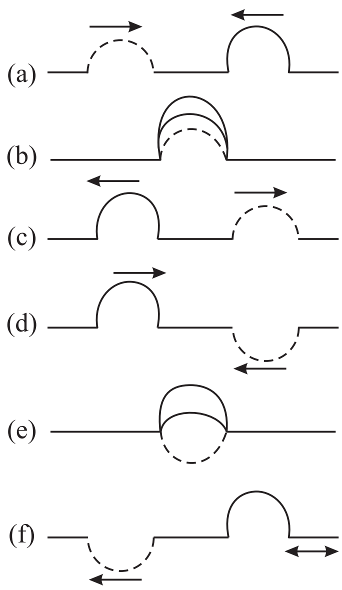


Fig. 14.6 : Illustrating principle of superpositionof waves

Now produce one crest and one trough on the slinky as shown in Fig. 14.6(d). The two are moving in the opposite direction. They meet [Fig. 14.6(e)], overlap and then separate out. Each one moves in the same direction in which it was moving before crossing and each one has the same shape as it was having before crossing. Repeat the experiment again and observe carefully what happens at the spot of overlapping of the two pulses [(Fig. 14.6(b) and (e)].

You will note that when crests overlap, the resultant is more and when crest overlaps the through, the resultant is on the side of crest but smaller size.We may summarize this result as : *At the points where the two pulses overlap, the resultant displacement is the vector sum of the displacements due to each of the two wave pulses.* ***This is called the principle of superposition.***

This activity demonstrates not only the principle of superposition but also shows that two or more waves can traverse the same space independent of each other. Each one travels as if the other were not present. This important property of the waves enable us to tune to a particular radio station even though the waves broadcast by a number of radio stations exist in space at the same time. We make use of this principle to explain the phenomena of *interference of waves, formation of beats and stationary or standing waves.*

14.3.1 Reflection and Transmission of Waves

We shall confine our discussion in respect of mechanical waves produced on strings and springs. What happens and why does it happen when a transverse wave crest propagates towards the fixed end of a string? Let us perform the following activity to understand it.

ACTIVITY 14.3

Fasten one end of a slinky to a fixed support as shown in (Fig. 14.7 (a). Keeping the slinky horizontal, give a jerk to its free end so as to produce a transverse wave pulse which travells towards the fixed end of the slinky (Fig. 14.7(a)). You will observe that the pulse bounces back from the fixed end. As it bounces back, the crest becomes a trough travels back in the opposite direction. Do you know the reason? As the pulse meets the fixed end, it exerts a force on the support. The equal and opposite reaction not only reverses the direction of propagation of the wave pulse but also reverses the direction of the displacement of the wave pulse (Fig. 14.7(b)). The support being much heavier than the slinky, it can be regarded as a denser medium. The wave pulse moving in the opposite direction is called the reflected wave pulse. So, we can say that ***when reflection takes place from a denser medium, the wave undergoes a phase change of π, that is, it suffers a phase reversal.***

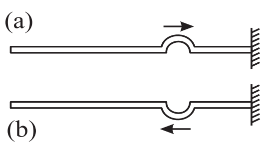


Fig. 14.7 : Reflection from a denser medium : a phase reversal.

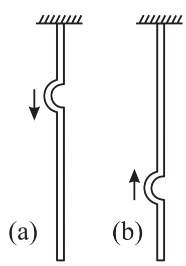


Fig.14.8(a) : A pulse travelling down towards the free end, (b) on reflection from the free end direction of its displacement remains unchanged

Let us now see what happens on reflection from a rarer medium. For this we perform another activity.

ACTIVITY 14.4

Suspend a fixed rubber tube from a rigid support (Fig. 14.8 a). Then generate a wave pulse travelling down the tube. On reflection from the free end, the wave pulse travels upward but without any change in the direction of its displacements i.e. crest returns as crest. Why? As the wave pulse reaches the free end of the tube, it gets reflected from a rarer boundary. (Note that air is rarer than the rubber tube.) Hence there is no change in the direction of displacement of the wave pulse. ***Thus on reflection from a rarer medium, no phase change takes place.***

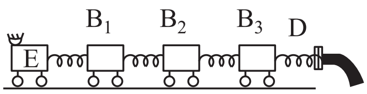


Fig. 14.9 : Longitudinal waves are reflected from a denser medium without change of type but with change of sign

You may now raise the question : Do longitudinal waves also behave similarly? Refer to Fig. 14.9, which shows a row of bogies. Now suppose that the engine E moves a bit towards the right. The buffer spring between the engine E and the first bogie gets compressed and pushes bogie B_1 towards the right. It then tries to go back to its original shape. As this compressed spring expands, the spring between the 1st and the 2nd bogie gets compressed. As the second compressed spring expands, it moves a bit towards the 3rd bogie. In this manner the compression arrives at the last buffer spring in contact with the fixed stand D. As the spring between the fixed stand and the last bogie expands, only the last bogie moves towards the left. As a result of this, the buffer spring between the next two bogies on left is compressed. This process continues, till the compression reaches between the engine and the first bogie on its right. Thus, a compression returns as a compression. But the bogies then move towards the left. In this mechanical model, the buffer spring and the bogies form a medium. The bogies are the particles of the medium and the spring between them shows the forces of elasticity.

Thus, when reflection takes place from a denser medium, the longitudinal waves are reflected without change of type but with change in sign. And on reflection from a rare medium, a longitudinal wave is reflected back without change of sign but with change of type. By 'change of type' we mean that rarefaction is reflected back as compression and a compression is reflected back as rarefaction.

INTEXT QUESTIONS 14.3

1. What happens when two waves travelling in the opposite directions meet?
2. What happens when two marbles each of the same mass travelling with the same velocity along the same line meet?

3. Two similar wave pulses travelling in the opposite directions on a string meet. What happens (i) when the waves are in the same phase? (ii) the waves are in the opposite phases?
4. What happens when a transverse wave pulse travelling along a string meets the fixed end of the string?
5. What happens when a wave pulse travelling along a string meets the free end of the string?
6. What happens when a wave of compression is reflected from (i) a rarer medium (ii) a denser medium?

14.4 SUPERPOSITION OF WAVES TRAVELLING IN THE SAME DIRECTION

Superposition of waves travelling in the same direction gives rise to two different phenomena (i) interference and (ii) beats depending on their phases and frequencies. Let us discuss these phenomena now.

14.4.1 Interference of waves

Let us compute the ratio of maximum and minimum intensities in an interference pattern obtained due to superposition of waves. Consider two simple harmonic waves of amplitudes a_1 and a_2 each of angular frequency ω, both propagating along $x - axis$, with the same velocity $v = \omega/k$ but differing in phase by a constant phase angle ϕ. These waves are represented by the equations

$$y_1 = a_1 \sin(\omega t - kx)$$

and

$$y_2 = a_2 \sin[(\omega t - kx) + \phi]$$

where $\omega = 2\pi/T$ is angular frequency and $k = \dfrac{2\pi}{\lambda}$ is wave number.

Since, the two waves are travelling in the same direction with the same velocity along the same line, they overlap. According to the principle of superposition, the resultant displacement at the given location at the given time is

$$y = y_1 + y_2 = a_1 \sin(\omega t - kx) + a_2 \sin[(\omega t - kx) + \phi]$$

If we put $(\omega t - kx) = \theta$, then

$$y = a_1 \sin\theta + a_2 \sin(\theta + \phi)$$

$$= a_1 \sin\theta + a_2 \sin\theta \cos\phi + a_2 \sin\phi \cos\theta$$

Let us put $a_2 \sin\phi = A \sin\alpha$

and $a_1 + a_2 \cos\phi = A \cos\alpha$

Then

$$y = A\cos\alpha \sin\theta + A\sin\alpha\cos\theta$$

$$= A\sin(\theta + \alpha)$$

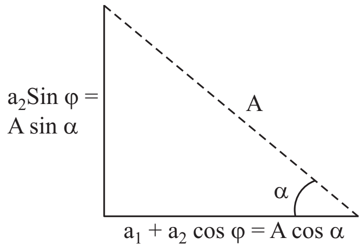


Fig. 14.10 : Calculating resultant amplitude A

Substituting for θ we get

$$y = A\sin[(\omega t - kx) + \alpha]$$

Thus, the resultant wave is of angular frequency ω and has an amplitude A given by

$$A^2 = (a_1 + a_2\cos\phi)^2 + (a_2\sin\phi)^2$$

$$= a_1^2 + a_2^2\cos^2\phi + 2a_1a_2\cos\phi + a_2^2\sin^2\phi$$

$$A^2 = a_1^2 + a_2^2 + 2a_1a_2\cos\phi \qquad (14.18)$$

In Eqn. (14.18), ϕ is the phase difference between the two superposed waves. If path difference, between the two waves corresponds to phase difference ϕ, then

$\phi = \dfrac{2\pi p}{\lambda}$, where $\dfrac{2\pi}{\lambda}$ is the phase change per unit distance.

When the path difference is an even multiple of $\dfrac{\lambda}{2}$, *i.e.*, $p = 2m\dfrac{\lambda}{2}$, then phase difference is given by $\phi = (2\pi/\lambda) \times (2m\,\lambda/2) = 2m\pi$. Since $\cos 2\pi = +1$, from Eqn. (14.18) we get

$$A^2 = a_1^2 + a_2^2 + 2a_1a_2 = (a_1 + a_2)^2$$

That is, when the collinear waves travelling in the same directions are in phase, the amplitude of the resultant wave on superposition is equal to sum of individual amplitudes.

As intensity of wave at a given position is directly proportional to the square of its amplitude, we have

$$I_{max} \propto (a_1 + a_2)^2$$

When $p = (2m + 1)\,\lambda/2$, then $\phi = (2m + 1)\,\pi$ and $\cos\phi = -1$. Then from Eqn. (14.18),

we get $$A^2 = a_1^2 + a_2^2 - 2a_1a_2 = (a_1 - a_2)^2$$

This shows that when phases of two collinear waves travelling in the same direction differ by an odd integral multiple of π, the amplitude of resultant wave generated by their superposition is equal to the difference of their individual amplituds.

Then $I_{min} \propto (a_1 - a_2)^2$

Thus $$\frac{I_{max}}{I_{min}} = \frac{(a_1 + a_2)^2}{(a_1 - a_2)^2} \qquad (14.19)$$

If $a_1 = a_2$, the intensity of resultant wave is zero. These results show that interference is essentially redistribution of energy in space due to superposition of waves.

14.4.2 Beats

We have seen that superposition of waves of same frequency propagating in the same direction produces interference. Let us now investigate what would be the outcome of superposition of waves of nearly the same frequency. First let us perform an activity.

ACTIVITY 14.5

Take two tuning forks of same frequency 512 Hz. Let us name them as A and B. Load the prong of the tuning fork B with a little wax. Now sound them together by a rubber hammer. Press their stems against a table top and note what you observe. You will observe that the intensity of sound alternately becomes maximum and minimum. These alternations of maxima and minima of intensity are called beats. One alternation of a maximum and a minimum is one beat. On loading the prong of B with a little more wax, you will find that no. of beats increase. On further loading the prongs of B, no beats may be heard. The reason is that our ear is unable to hear two sounds as separate produced in an interval less than one tenths of a second. Let us now explain how beats are produced.

(a) Production of beats : Suppose we have two tuning forks A and B of frequencies N and N + n respectively; n is smaller than 10. In one second, A completes N vibrations but B completes N + n vibrations. That is, B completes n more vibrations in one second than the tuning fork A. In other words, B gains n vibrations over A in 1s and hence it gains 1 vib. in (1/n) s. and half vibration over A in (1/2n) s. Suppose at $t = 0$, i.e. initially, both the tuning forks were vibrating in the same phase. Then after (1/2n)s, B will gain half vibration over A. Thus after $\frac{1}{2n}s$ it will vibrate in oposite phase. If A sends a wave of compression then B sends a wave of rarefaction to the observer. And, the resultant intensity received by the ear would be zero. After (1/n)s, B would gain one complete vibration. If now A sends a wave of compression, B too would send a wave of compression to the observer. The intensity observed would become maximum. After (3/2n)s, the two forks again vibrate in the opposite phase and hence the intensity would again become minimum. This process would continue. The observer would hear 1 beat in (1/n)s, and hence n beats in one second. Thus, the *number of beats heard in* *one second equals the difference in the frequencies of the two tuning forks.* If more than 10 beats are produced in one second, the beats are not heard as separate. The beat frequency is n and beat period is $1/n$.

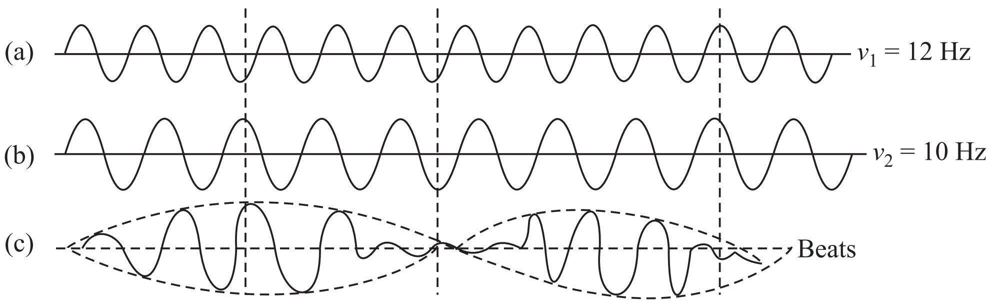


Fig.14.11 : (a) Displacement time graph of frequency 12 Hz. **(b)** displacement time graph of frequency 10 Hz. Superposition of the two waves produces 2 beats per second.

(b) Graphic method : Draw a 12 cm long line. Divide it into 12 equal parts of 1 cm. On this line draw 12 wavelengths each 1 cm long and height 0.5 cm. This represents a wave of frequency 12 Hz. On the line (b) draw 10 wavelengths each of length 1.2 cm and height 0.5 cm. This represents a wave of frequency 10Hz. (c) represents the resultant wave. Fig, 14.11 is not actual waves but the displacement time graphs. Thus, the resultant intensity alternately becomes maximum and minimum. The number of beats produced in one second is Δv. Hence, the beat frequency equals the difference between the frequencies of the waves superposed.

Example 14.3 : A tuning fork of unknown frequency gives 5 beats per second with another tuning of 500 Hz. Determine frequency of the unknown fork.

Solution : $v' = v \pm n = 500 \pm 5$

$\Rightarrow$ The frequency of unknown tuning fork is either 495 Hz or 505 Hz.

Example 14.4 : In an interference pattern, the ratio of maximum and minimum intensities is 9. What is the amplitude ratio of the superposing waves?

Solution : $\dfrac{I_{max}}{I_{min}} = \left(\dfrac{a_1 + a_2}{a_1 - a_2}\right)^2 \quad \Rightarrow \quad 9 = \left(\dfrac{1+r}{1-r}\right)^2$, where $r = \dfrac{a_2}{a_1}$.

Hence, are can write

$$\frac{1+r}{1-r} = 3$$

You can easily solve it to get $r = \dfrac{1}{2}$, i.e., amplitude of one wave is twice that of the other.

Notes

INTEXT QUESTIONS 14.4

1. If the intensity ratio of two waves is 1:16, and they produce interference, calculate the ratio I_{max}/I_{min}?
2. Waves of frequencies v and $v + 4$ emanating from two sources of sound are superposed. What will you observe?
3. Two waves of frequencies v and $v + \Delta v$ are supperposed, what would be the frequency of beats?
4. Two tuning forks A and B produce 5 beats per second. On loading one prong of A with a small ring, again 5 beats per second are produced. What was the frequency of A before loading if that of B is 512 Hz. Give reason for your answer.

14.5 SUPERPOSITION OF WAVES OF SAME FREQUENCY TRAVELLING IN THE OPPOSITE DIRECTIONS

So far we have discussed superposition of collinear waves travelling in the same direction. In such waves, crests, and troughs or rarefactions and compressions in a medium travel forward with a velocity depending upon the properties of the medium. Superposition of progressive waves of same wavelength and same amplitude travelling with the same speed along the same line in a medium in opposite direction gives rise to stationary or standing waves. In these waves crests and troughs or compressions and rarefactions remain stationary relative to the observer.

14.5.1 Formation of Stationary (Standing) Waves

To understand the formation of stationary waves, refer to Fig. 14.12 where we have shown the positions of the incident, reflected and resultant waves, each after T/4s, that is, after quarter of a period of vibration.

(i) Initially, at $t = 0$, [Fig. 14.12(i)], the incident wave, shown by dotted curve, and the reflected wave, shown by dashed curve, are in the opposite phases. Hence the resultant displacement at each point is zero. All the particles are in their respective mean positions.

(ii) At $t = T/4$s [Fig. 14.12(ii)], the incident wave has advanced to the right by $\lambda/4$, as shown by the shift of the point P and the reflected wave has advanced to the left by $\lambda/4$ as shown by the shift of the point P′. The resultant wave form has been shown by the thick continuous curve. It can be seen that the resultant displacement at each point is maximum. Hence the particle velocity at each point is zero and the strain is maximum

(iii) At t = T/2s [Fig. 14.12(iii)], the incident wave advances a distance λ/2 to the right as shown by the shift of the point P and the reflected wave advances a distance λ/2 to the left as shown by the shift of the point P′. At each point, the displacements being in the opposite directions, have a zero resultant shown by a thick line.

(iv) At t = 3T/4s [Fig. 14.12(iv)], the two waves are again in the same phase. The resultant displacement at each point is maximum. The particle velocity is zero but the strain is maximum possible.

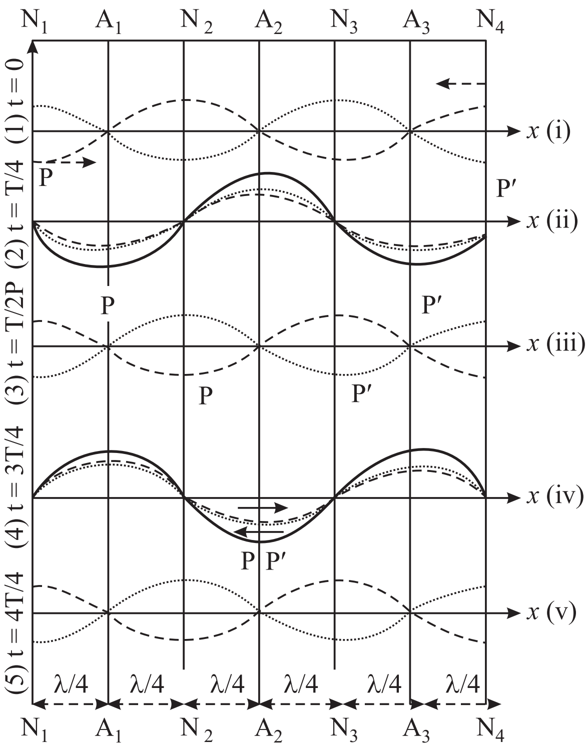


Fig. 14.12 : Showing formation of stationary waves due to superposition of two waves of same wave length, same amplitude travelling in opposite direction along the same line.

(v) At t = 4T/4s [Fig. 14.12(v)], the incident and reflected waves at each point are in the opposite phases. The resultant is a straight line (shown by an unbroken thick line). The strain $\Delta y/\Delta x$ at each point is zero.

Note that

- at points N_1, N_2, N_3 and N_4, the amplitude is zero but the strain is maximum. Such points are called **nodes;**
- at points A_1, A_2 and A_3, the amplitude is maximum but the strain is minimum. These points are called **antinodes;**
- the distance between two successive nodes or between two, successive antinode is λ/2;
- the distance between a node and next antinode is λ/4;
- the time period of oscillation of a stationary wave equals the time period of the two travelling waves whose superposition has resulted in the formation of the stationary wave; and
- the energy alternately surges back and forth about a point but on an average, the energy flow past a point is zero.

Superposition of two identical collinear waves travelling with the same speed in opposite directions leads to formation of stationary waves. They are called stationary waves, because the wave form does not move forward, but alternately shrinks and dilates. The energy merely surges back and forth and on an average, there is no net flow of energy past a point.

14.5.2 Equation of Stationary Wave

The equation of a simple harmonic wave travelling with velocity $\upsilon = \omega/k$ in a medium is

$$y_1 = -a \sin(\omega t - kx)$$

On reflection from a denser medium, suppose the wave travels along the same line along X–axis in the opposite direction with phase change of π. The equation of the reflected wave is therefore,

$$y_2 = a \sin(\omega t - kx)$$

Thus, owing to the superposition of the two waves, the resultant displacement at a given point and time is

$$y = y_1 + y_2$$

$$= a \sin(\omega t - kx) - a \sin(\omega t - kx)$$

Using the trigonometric identity. sin A – sin B = 2 Sin (A – B)/2, cos (A + B)/2, above expression simplifies to

$$y = -2a \sin kx \cos \omega t \qquad (14.20)$$

Let us put $-2a \sin kx = A$. Then we can write

$$y = A \cos \omega t$$

Eqn. (14.20) represents a resultant wave of angular frequency ω and amplitude $2a \sin kx$ This is the equation of stationary wave. The amplitude of the resultant wave, oscillates in space with an angular frequency ω, which is the phase change per metre. At such points where $kx = m\,\pi = m\lambda/2$, $\sin kx = \sin m\pi = 0$. Hence $A = 0$,

The points where the amplitude is zero are referred to as **nodes.** At these points $\Delta y/\Delta x$ = maximum, that is strain is maximum. Obviously, the spacing between two nearest points is $\lambda/2$.

At those points where

$$kx = (2m + 1)\,\pi/2 \text{ or } x = (2m + 1)\,\lambda/2 \times \lambda/2\pi = (2m + 1)\,\lambda/4$$

$$\text{Sin } kx = \sin(2m + 1)\,\pi/2 = \pm 1.$$

Hence, A is maximum. At these points the strain $\Delta y/\Delta x$ is zero. Obviously the spacing between two such neighbouring points is $\lambda/2$. These points where the amplitude is maximum but strain is zero are referred to as **antinodes.**

It may be pointed out here that at nodes, the particle velocity is zero and at antinodes, particle velocity $\Delta y/\Delta t$ is maximum. Therefore, it follows that the average flow of energy across any point is zero. The energy merely surges back and forth. That is why, these waves have been named stationary or standing waves.

14.5.3 Distinction between Travelling and Standing Waves

Let us summarise the main differences between travelling and standing waves.

Travelling Waves	Standing Waves
1. Particular conditions of the medium namely crests and troughs or compressions and rarefactions appear to travel with a definite spped depending on density and elasticity (or tension) of the medium.	Segments of the medium between two points called nodes appear to contract and dialate. Each particule or element of the medium vibrates to and fro like a pendulum.
2. The amplitude of vibration of all the particles is the same.	At nodes the amplitude is zero but at antinodes the amplitude is maximum.
3. All the particles pass through their mean positions with maximum velocity one after the other.	At nodes the particle velocity is zero and at antinodes it is maximum.
4. Energy is transferred from particle to particle with a definite speed.	The energy surges back and forth in a segment but does not move past a point.
5. In a travelling wave the particles attain their maximum velocity one after the other.	In a stationary wave the maximum velocity is different at different points. It is zero at nodes but maximum at antinodes. But all the particles attain their respective maximum velocity simultaneously.
6. In a travelling wave each region is subjected to equal strains one after the other.	In case of standing waves strain is maximum at nodes and zero at antinodes.
7. There is no point where there is no change of density.	Antinodes are points of no change of density but at nodes there is maximum change of density.

INTEXT QUESTIONS 14.5

1. Does energy flow across a point in case of stationary waves? Justify your answer.

2. What is the distance between two successive nodes, and between a node and next antinode?
3. Pressure nodes are displacement antinodes and pressure antinodes are displacement nodes. Explain.
4. Stationary waves of frequency 170Hz are formed in air. If the velocity of the waves is 340 ms^{-1}, what is the shortest distance between (i) two nearest nodes (ii) two nearest antinode (iii) nearest node and antinode?

14.6 CHARACTERISTICS OF MUSICAL SOUND

The characteristics of musical sounds help us to distinguish one musical sound from another.

These are pitch, londness and quality. We will now discuss these briefly.

14.6.1 Pitch

The term pitch is the characteristic of musical notes that enables us to classify a note as 'high' or 'low'. It is a subjective quantity which cannot be measured by an instrument. It depends on frequency. However, there does not exist any one-to-one correspondence between the two. A shrill, sharp or acute sound is said to be of high pitch. But a dull, grave and flat note is said to be of low pitch. Roaring of lion, though of high intensity, is of low pitch. On the other hand, the buzzing of mosquito, though of low intensity, is of high pitch.

14.6.2 Loudness

The loudness of sound is a subjective effect of intensity of sound received by listeners ear. The *intensity of waves is the average amount of energy transported by the wave per unit area per second normally across a surface at a given point.* There is a large range of intensities over which the ear is sensitive. As such, logarithmic scale rather than arithmetic intensity scale is more convenient.

Threshold of hearing and Intensity of Sound

The intensity level β of a sound wave is defined by the equation.

$$\beta = 10 \log I/I_0 \qquad (14.21)$$

where I_0 is arbitrarily chosen reference intensity, taken as 10^{-12} Wm^{-2}. This value corresponds to the faintest sound that can be heard. Intensity level is expressed in decibels, abbreviated *db*. If the intensity of a sound wave equals I_0 or 10^{-12} Wm^{-2}, its intensity level is then $I_0 = 0$ *db*. Within the range of audibility, sensitivity of human ear varies with frequency. ***The threshold audibility at any frequency is the minimum intensity of sound at that frequency, which can be detected.***

The standard of perceived loudness is the ***sone.*** A sone is the loudness experienced by a listener with normal hearing when 1kilo hertz tone of intensity 40db is presented to both ears.

The range of frequencies and intensities to which ear is sensitive have been represented in a diagram in Fig. 14.13, which is in fact a graph between frequency in hertz versus intensity level 1 in decibels. This is a graph of auditory area of good hearing. The following points may be noted.

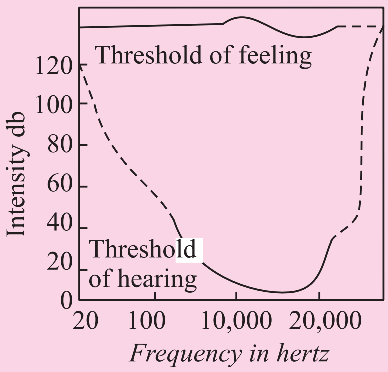


Fig. 14.13 : Auditory area between threshold of hearing and threshold of feeling

- The lower part of the curve shows that the ear is most sensitive for frequencies between 2000 Hz to 3000 Hz, where the threshold of hearing is about 5db. Threshold of hearing in general, is zero decibel.
- At intensities above those corresponding to the upper part of the curve, the sensation changes from one of hearing to discomfort and even pain. This curve represents the threshold of feeling.
- Loudness increases with intensity, but there is no definite relation between the two.
- Pure tones of same intensity but different frequencies do not necessarily produce equal loudness.
- The height of the upper curve is constant at a level of 120 db for all frequencies.

The intensity of sound waves depends on the following factors :

- ***Amplitude of vibration :*** *$I \propto a^2$ where a* is amplitude of the wave.
- ***Distance between the observer and the Source :*** $I \propto 1/r^2$ where *r* is the distance of the observer from the source (provided it is a point source).
- ***Intensity is directly proportional to the square of frequency of the wave*** ($I \propto v^2$).
- ***Intensity is directly proportional to the density of the medium*** ($I \propto \rho$).

14.6.3 Quality

It is the *characteristic of sound waves which enables us to distinguish between two notes of the same pitch and intensity but sounded by two different instruments.* No instrument, except a tuning fork, can emit a pure note; a note of one particular frequency. In general, when a note of frequency *n* is sounded, in addition to it, notes of higher frequencies $2n$, $3n$, $4n$ ... may also be produced. These notes, have different amplitudes and phase relations. The resultant wave form of the emitted waves determines the quality of the note emitted. Quality, like loudness and pitch is a subjective quantity. It depend on the resultant wave form.

14.6.4 Organ Pipes

It is the simplest form of a wind instrument. A wooden or metal pipe producing musical sound is known as organ pipe. Flute is an example of organ pipe. If both the ends of the pipe are open, we call it an ***open pipe.*** However, if one end is closed, we call it a ***closed pipe.*** When we blow in gently, almost a pure tone is heard. This pure tone is called a ***fundamental note***. But, when we blow hard, we also hear notes of frequencies which are integral multiple of the frequency of the fundamental note. You can differentiate between the sounds produced by water from a tap into a bucket. These frequencies are called **overtones.**

Note that:

- At the closed end of a pipe, there can be no motion of the air particles and the closed end must be node.
- At the open end of the pipe, the change in density must be zero since this end is in communication with atmosphere. Further, since the strain is zero, hence this end must be an antinode.

(a) Open pipe : The simplest mode of vibrations of the air column called fundamental mode is shown in Fig.14.14 (a). At each end, there is an antinode and between two antinodes, there is a node. Since the distance between a node and next antinode is $\lambda/4$, the length l of the pipe is

$$l = (\lambda/4) + (\lambda/4) = \lambda/2 \text{ or } \lambda = 2l.$$

The frequency of the note produced is

$$n_1 = \upsilon/\lambda = \upsilon/2l$$

The next mode of vibration of the air column is shown in Fig.14.14 (b). One more node and one more antinode has been produced. In this case

$$\lambda = (\lambda/4) + (\lambda/4) + (\lambda/4) + (\lambda/4) = l$$

The frequency of the note is

$$n_2 = \upsilon/\lambda = \upsilon/l = 2\upsilon/2l$$

$$n_2 = 2\upsilon/2l$$

That is $n_2 = 2n_1$

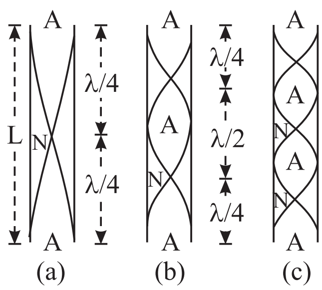


Fig. 14.14 : Harmonics of an open Organ pipe. The curves represent the wave of the longitudinal standing waves

The note produced is called second harmonic or 1st ***overtone***. To get the second harmonic you have to blow harder. But if you blow still harder one more node and one more antinode is produced [Fig.14.14{c)]. Thus, in this case

$$l = \frac{\lambda}{2} + \frac{\lambda}{4} + \frac{\lambda}{2} + \frac{\lambda}{4}$$

$$\lambda = \frac{2l}{3}$$

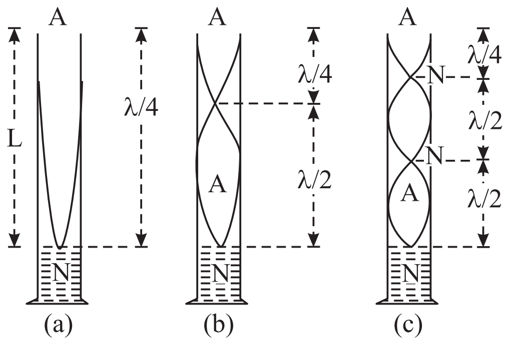


Fig. 14.15 : Harmonics of a closed organ pipe. The curves represented wave form of the longitudinal standing waves.

Therefore, the frequency of the note emitted is

$$n_3 = \frac{\upsilon}{\lambda} = \frac{3\upsilon}{2l} = 3n_1$$

The note produced is called the 3rd harmonic or 2nd overtone.

(b) Closed pipe : The simplest manner in which the air column can vibrate in a closed pipe is shown in Fig. 14.15(a). There is an antinode at the open end and a node at the closed end. The wave length of the wave produced is given by

$$l = \lambda/4 \text{ or } \lambda = 4l$$

Therefore, the frequency of the note emitted is

$$n_1 = \upsilon/\lambda = \upsilon/41$$

The note produced is called *fundamental* note. On blowing harder one more node and antinode will be produced (Fig. 14.15(b)). The wavelength of the note produced is given by

$$l = \frac{\lambda}{2} + \frac{\lambda}{4} = \frac{3\lambda}{4} \quad \text{or } \lambda = \frac{4l}{3}$$

The frequency of the note emitted will be

$$n_3 = \frac{\upsilon}{\lambda} = \frac{3\upsilon}{4l} = 3n_1$$

The note produced is called the first overtone or the 3rd harmonic of the fundamental, blowing still harder one more node and one more antinode will be produced Fig. 14.15(C). The wavelength of the note produced is then given by

$$l = \frac{\lambda}{2} + \frac{\lambda}{2} + \frac{\lambda}{4} + \frac{5\lambda}{4} \quad \text{or } \lambda = \frac{4l}{5}$$

The frequency of the note emitted then will be

$$n_3 = \frac{\upsilon}{\lambda} = \frac{5\upsilon}{4l} = 5n_1$$

The note produced is called the second overtone or the 5th harmonic of the fundamental. On comparison with the notes emitted by the open and closed pipe, you will find that the open pipe is richer in overtones. In closed pipe, the even order harmonics are missing.

Example 14.5 : Two organ pipes – one open and the other closed – are of the same length. Calculate the ratio of their fundamental frequencies.

Solution : $\dfrac{\text{Frequency of open pipe}}{\text{Frequency of closed pipe}} = \dfrac{\upsilon/2\ell}{\upsilon/4\ell} = 2$

∴ Frequency of note produced by open pipe = 2 × frequency of fundamental note produced by closed pipe.

INTEXT QUESTIONS 14.6

1. How pitch is related to frequency?
2. What is that characteristic of musical sounds which enables you distinguis between two notes of the same frequency, and same intensity but sounded by two different instruments?
3. Name the characteristic of sound which helps you identify the voice of your friend.
4. Out of open and closed organ pipes, which one is richer in overtones?
5. What is the ratio of the frequencies of the notes emitted (1) by an open pipe and (ii) by a closed pipe of the same length.
6. What will be the effect of temperature, if any, on the frequency of the fundamental note of an open pipe?

Noise Pollution

When the sensation of sound changes from one of hearing to discomfort, it causes noise pollution and if it persits for a long time, it has harmful effects on certain organ of human beings. Noise is also one of the by-products of industrialisation and misuse of modern amentities provided by science to human beings. We summarise here under the sources or description of noises and their effects as perceived by the human beings.

Table 14.1 : Sources of Noise and their Effects

Source	Intensity Level in decibels	Perceived Effect by human being
Threshold of hearing	0 (=10^{-12} Wm^{-2})	Just audible
Rustle of leaves	10	Quiet
Average whisper	20	Quiet
Radio at low volume	40	Quiet
Quiet automobile	50	moderately loud
Ordinary conversation	65	do
Busy street traffic	70 to 80	loud
Motor bike and heavy vehicles	90	very loud
Jet engine about 35m away	105	Uncomfortable
Lightening	120 (=1 Wm^{-2})	do
Jet plane at take off	150	Painful sound

(a) Effect of Noise Pollution

1. It causes impairment of hearing. Prolonged exposure of noise at 85 or more than 85db causes severe damage to the inner part of the ear.
2. It increases the rate of heart beat and causes dilation of the pupil of eye.
3. It causes emotional disturbance, anxiety and nervousness.
4. It causes severe headache leading to vomiting.

(b) Methods of Reducing Noise Pollution

1. Shifting of old industries and setting new ones away from the dwellings.
2. Better maintenance of machinery, regular oiling and lubrication of moving parts.
3. Better design of engines and machines.

4. Restriction on use of loudspeakers and amplifiers.
5. Restricting the use of fire crackers, bands and loud speakers during religious, political and marriage processions.
6. Planting trees on roads for intercepting the path of sound.
7. Intercepting the path of sound by sound absorbing materials.
8. Using muffs and cotton plugs.

Shock Waves

When a source of waves is travelling faster than the sound waves, shock waves are produced. The familiar example is the explosive sound heared by an observer when a supersonic plane flies past over the head of the observer. It may be pointed out that the object which moves with a speed greater than the speed of sound is itself a source of sound.

14.7 ELECTROMAGNETIC WAVES

You know that light is an e.m. wave. It has wavelength in the range 4000°A to 7500°A. A brief description of em waves is given below.

14.7.1 Properties of e.m. waves

The following properties of e.m. waves may be carefully noted.

(*i*) e.m. waves are transverse in nature

(*ii*) They consist of electric (**E**) and magnetic fields (**B**) oscillating at right angles to each other and perpendicular to the direction of propagation (k). Also $\mathbf{E} = c\mathbf{B}$. [see figures 14.16]

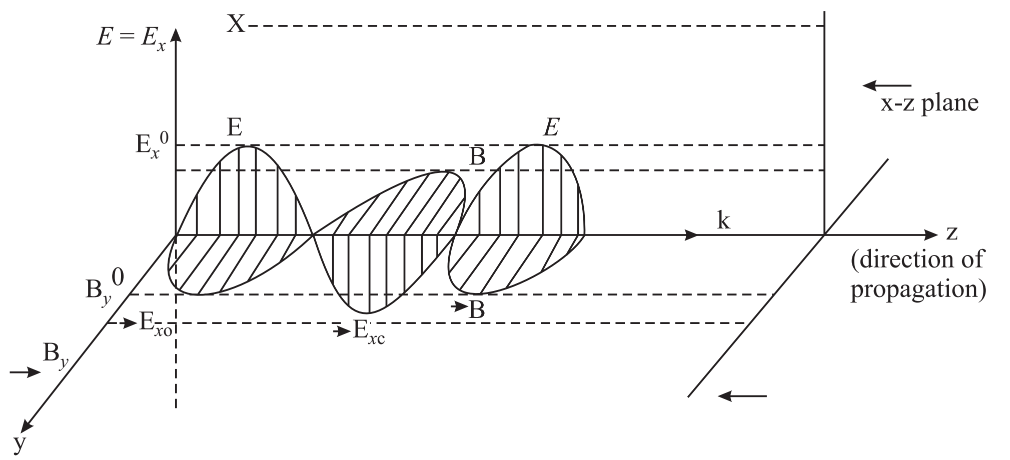


Fig. 14.16 : Electrical and Magnetic fields in em waves

(iii) They propagate through free space (in vacuum) with a uniform velocity = $\frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3 \times 10^8$ ms^{-1} = c (velocity of light). For a medium of permeability μ (= $\mu_0 . \mu_r$) and permittivity ε (= ε_0 . ε_r) the velocity becomes

$$\upsilon = \frac{1}{\sqrt{\mu\varepsilon}} = \frac{1}{\sqrt{\mu_0\varepsilon_0}\ \sqrt{\mu_r\varepsilon_r}} = \frac{c}{\sqrt{\mu_r\varepsilon_r}} < c$$

(iv) The nature and action of these waves depends on their frequency (or wavelength). Maxwell's theory placed no restriction on possible wavelengths for e.m. waves and hence e.m. waves of wavelengths ranging from 6×10^{-13} m have been successfully produced. There is no limit to very long wavelengths which correspond to radio broadcast waves. The whole range of e.m. waves from very long to very short wavelengths constitutes the *electromagnetic spectrum.*

James Clark Maxwell
(1831 – 1879)

Scottish Mathematician and physicist Maxwell is famous for his theories of electromagnetic fields. Through his equations of electromagnetic principles he showed that they implicitly indicated the existence of em waves which travelled with the speed of light, thus relating light and electromagnetism.

With clausius he developed the kinetic theory of gases. He developed a statistical theory of heat. A man of varied interests, he derived the theorem of equipartition of energy, showed that viscosity varies directly with temperature and tried to explain the rings of saturn.

14.7.2 Electromagnetic Spectrum

Maxwell gave the idea of e.m. waves while Hertz, J.C. Bose, Marconi and others successfully produced such waves of different wavelengths experimentally. However, in all the methods, *the source of e.m. waves is the accelerated charge.*

Electromagnetic waves are classified according to the method of their generation and are named accordingly. Overlapping in certain parts of the spectrum by different classes of e.m waves is also observed. This tells that the e.m. waves of wavelenths in the overlapping region can be produced by two different methods. It is ***important to remember that the physical properties of e.m. waves are determined by the frequencies or wavelengths and not by the method of their generation.*** A suitable classification of e.m. waves is called the electromagnetic spectrum.

There is no sharp dividing point between one class of e.m. waves and the next. The different parts are as follows :

(i) **The low frequency radiations** $\begin{Bmatrix} v = 60\text{Hz to }50\text{Hz} \\ \lambda = 5\times10^6\, m \textit{ to } 6\times10^6\, m \end{Bmatrix}$: generated from a.c. circuits are classified as power frequencies or power waves or electric power utility e.m. waves. These *weaves have the lowest* frequency.

(ii) **Radio Waves** $\begin{Bmatrix} \lambda = 0.3\text{m to }10^6\text{ m} \\ v = 10^9\text{ Hz } \textit{to } 300\text{Hz} \end{Bmatrix}$: Radio waves are generated when charges are accelerated through conducting wires. They are generated in such electronic devices as LC oscillators and are *used extensively in radio and television communications.*

(iii) **Microwaves** $\begin{Bmatrix} \lambda = 10^{-3}\text{ m to }0.3\text{m} \\ v = 10^{11}\text{ Hz } \textit{to } 10^9\text{ Hz} \end{Bmatrix}$: These are generated by oscillating currents in special vacuum tubes. Because of their short wavelengths, they are well suited for the radar system used in aircraft navigation, T.V. communication and for studying the atomic and molecular properties of matter. Microwave ovens use these radiations as heat waves. It is suggested that solar energy could be harnessed by beaming microwaves to Earth from a solar collector in space.

(iv) **Infra-red waves** $\begin{Bmatrix} \lambda = 7\times10^{-7}\text{ m to }10^{-3}\text{ m} \\ v = 4.3\times10^{14}\text{ Hz to }3\times10^{11}\text{ Hz} \end{Bmatrix}$: Infra-red waves, also called heat waves, are produced by hot bodies and molecules. These are readily absorbed by most materials. The temperature of the body, which absorbs these radiations, rises. Infrared radiations have many practical and scientific applications including physical therapy infrared photography etc. These are detected by a thermopile.

(v) **Visible light** $\begin{Bmatrix} \lambda = 4\times10^{-7}\text{ m to }7\times10^{-7}\text{ m} \\ v = 7.5\times10^{14}\text{ Hz to }4.3\times10^{14}\text{ Hz} \end{Bmatrix}$: These are the e.m. waves that human eye can detect or to which the human retina is sensitive. It forms a very small portion of the whole electromagnetic spectrum. These waves are produced by the rearrangement of electrons in atoms and molecules. When an electron-jumps from outer orbit to inner orbit of lower energy, the balance of energy is radiated in the form of visible radiation. The various wavelengths of visible lights are classified with colours, ranging from violet ($\lambda = 4 \times 10^{-7}$m) to red ($\lambda = 7 \times 10^{-7}$). Human eye is most sensitive to yellow-green light ($\lambda = 5 \times 10^{-7}$m). Light is the basis of our communitation with the world around us.

(vi) **Ultraviolet** $\left\{\begin{array}{l}\lambda = 3\times10^{-9}\,\text{m to } 4\times10^{-7}\,\text{m}\\ \nu = 10^{-17}\,\text{Hz to } 7.5\times10^{14}\,\text{Hz}\end{array}\right\}$: Sun is the important source of ultraviolet radiations, which is the main cause of suntans. Most of the ultraviolet light from Sun is absorbed by atoms in the upper atmosphere i.e. stratosphere, which contains ozone gas. This ozone layer then radiates out the absorbed energy as heat radiations. Thus, the lethal (harmful to living beings) radiations get converted into useful heat radiations by the ozone gas, which warms the stratosphere. These ultraviolet rays are used in killing the bacteria in drinking water, in sterilisation of operation theatres and also in checking the forgery of documents.

(vii) **X-rays** $\left\{\begin{array}{l}\lambda = 4\times10^{-13}\,\text{m to } 4\times10^{-8}\,\text{m}\\ \nu = 7.5\times10^{20}\,\text{Hz to } 7.5\times10^{15}\,\text{Hz}\end{array}\right\}$: These are produced when high energy electrons bombard a metal target (with high melting point) such as tungsten. X-rays find their important applications in medical diagnostics and as a treatment for certain forms of cancer. Because, they destroy living tissues, care must be taken to avoid over-exposure of body parts. X-rays are also used in study of crystal-structure. They are detected by photographic plates.

(viii) **Gamma rays** $\left\{\begin{array}{l}\lambda = 6\times10^{-17}\,\text{m to } 10^{-10}\,\text{m}\\ \nu = 5\times10^{24}\,\text{Hz to } 3\times10^{18}\,\text{Hz}\end{array}\right\}$: These are emitted by radioactive nuclei such as cobalt (60) and ceasium (137) and also during certain nuclear reactions in nuclear reactors. These are highly penetrating and cause serious damage when absorbed by living tissues. Thick sheets of lead are used to shield the objects from the lethal effects of gamma rays.

The energy (E) of e.m. waves is directly proportional to their frequency ν $\left(E = h\nu = \dfrac{hc}{\lambda}\right)$ and inversely proportional to their wave-length (λ). Thus gamma rays are the most energetic and penetrating e.m. waves, while the power frequencies, and the A.M. radio waves are the weakest radiations. Gamma rays are used to detect metal flaws in metal castings. They are detected by Geiger tube or scintillation counter.

Depending on the medium, various types of radiations in the spectrum will show different characteristic behaviours. For example, while whole of the human body is opaque to visible light, human tissues are transparent to X-rays but the bones are relatively opaque. Similarly Earth's atmosphere behaves differently for different types of radiations.

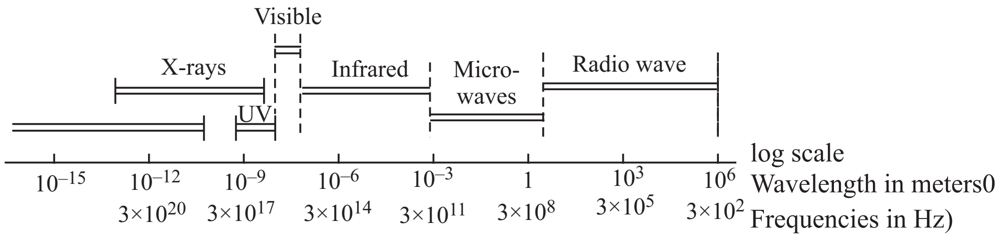


Fig. 14.17: Electromagnetic spectrum

INTEXT QUESTIONS 14.7

1. Fill in the blanks:

   (i) ....are generated by oscillating currents in special vacuum tubes.

   (ii) Human eye is most sensitive to....................................color light.

   (iii) ....is the important source of ultraviolet radiation.

   (iv) ....are used as the diagnostic tool in medical,

   (v) Infrared radiations can be detected by a...................................

2. Which of the e.m. waves are more energetic?

   (i) Ultraviolet or infrared.

   (ii) x-rays or γ-rays

3. Which of the e.m. waves are used in aircraft navigation by radar?

4. Which gas in the atmosphere absorbs ultraviolet radiations from the Sun before reaching the earth's surface?

5. How are the electric field and magnetic field oriented with respect to each other in an e.m. wave?

14.8 DOPPLER EFFECT

While waiting on a railway platform for the arrival of a train, you might have observed that the pitch of the whistle when the engine approaches you and when the engine moves away from you are different. You will note that the pitch is higher when the engine approaches but is lower when the engine moves away from you. Similarly, the pitch of the horn of a bus going up a hill changes constantly.

Apparent change of frequency observed due to the relative motion of the observer and the source is known as Doppler effect.

Let υ be velocity of the sound waves relative to the medium, (air), υ_s velocity of the source; and υ_o velocity of the observer.

Christian Doppler

(1803 – 1853)

C.J. Doppler, an Austrian physicist and mathematician, was born on Nov., 29, 1803 in a family of stone mesons. A pale and frail person, he was not considered good enough for his family business. So on recommendation of the professor of mathematics at Salzburg Lycousin, he was sent to Vienna Polytechnic from where he graduated in 1825.

A struggler through out his life, Doppler had to work for 18 months as a book-keeper at a cotton spinning factory. He could think of marrying in 1836 only when he got a permanent post at the technical secondary school at prague. He was once reprimanded for setting too harsh papers in maths for polytechnique students. But he pushed his way through all odds and finally got succes in getting the position of the first director of the new Institute of Physics at Vienna University.

The Doppler effect discovered by him made him famous overnight, because the effect had far reaching impact on acoustics and optics. The RADAR, the SONAR, the idea of expanding universe there are so many developments in science and technology which owe a lot to Doppler effect. He died on March 17, 1853 in Venice, Italy.

It is important to note that the wave originated at a moving source does not affect the speed of the sound. The speed υ is the property of the medium. The wave forgets the source as it leaves the source. Let us suppose that the source, the observer and the sound waves travel from left to right. Let us first consider the ***effect of motion of the source.*** A particular note which leaves the sources at a given time after one second arrives at the point A such that SA = υ. In this time, the source moves a distance υ_s. Hence all the n waves that the source had emitted in one second are contained in the space $x = \upsilon - \upsilon_s$. Thus length of each wave decreased to

$$\lambda' = \frac{\upsilon - \upsilon_s}{n} \qquad ...(14.22)$$

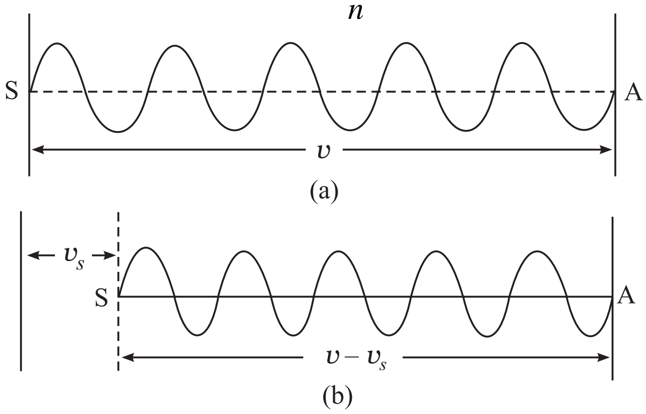


Fig. 14.18 : Crowding of waves when source is moving

Now let us consider the ***effect of motion of the observer.*** A particular wave which arrives at O at a particular time after one second will be at B such that OB = υ. But in the mean time, the observer moves from O to O′. Hence only the waves contained in the space O′B have passed across the observer in one second. The number of the waves passing across the observer in one second is therefore,

$$n' = (\upsilon - \upsilon_0)/\lambda' \qquad (14.23)$$

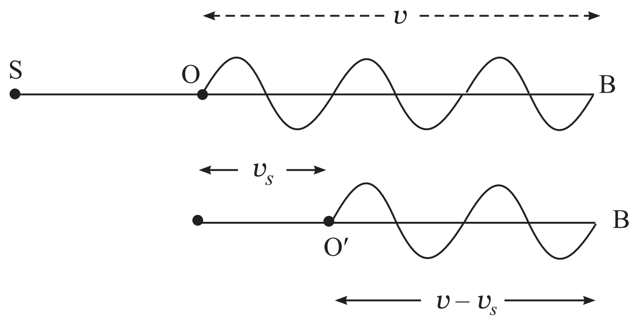


Fig. 14.19 : Waves received by a moving listner

Substituting for λ' from Eqn. (14.22) we get

$$n' = \frac{\upsilon - \upsilon_0}{\upsilon - \upsilon_s} n \qquad (14.24)$$

where n' is the observed frequency when both observer and source are moving in the direction from the source to the observer.

In using Eqn.(14.24) ***the velocity of sound is taken positive in the direction from the source to the observer. Similarly ,*** υ_0 and υ_s are taken positive if these are in the direction of υ and vice versa.

The utility of Doppler's effect arises from the fact that it is applicable to light waves as much as to sound waves. In particular, it led us to the concept of expansion of the universe.

The following examples will help you to understand this application of Doppler's effect.

Example 14.6 : The light from a star, on spectroscopic analysis, shows a shift towards the red end of the spectrum of a spectral line. If this shift, called the red shift, is 0.032%, calculate the velocity of recession of the star.

Solution : In this case, the source of waves is the star. The observer is at rest on the Earth. We have shown that in such a case

$$\lambda' = \frac{\upsilon - \upsilon_s}{n}$$

But $n = \upsilon/\lambda$ Therefore, $\lambda' = \dfrac{\upsilon - \upsilon_s}{\upsilon/\lambda}$

$$= \lambda \frac{(\upsilon - \upsilon_s)}{\upsilon}$$

$$= \lambda \left(1 - \frac{\upsilon_s}{\upsilon}\right)$$

On rearranging terms, we can write

$$\frac{\lambda' - \lambda}{\lambda} = -\frac{\upsilon_s}{\upsilon}$$

or
$$\frac{\Delta\lambda}{\lambda} = \frac{\upsilon_s}{\upsilon}$$

we are told that $\frac{\Delta\lambda}{\lambda}$ = 0.032/100. And since $\upsilon = c = 3 \times 10^8$ ms^{-1}, we get

$$\upsilon_s = \upsilon\frac{\Delta\lambda}{\lambda} = -(3 \times 10^8 \text{ ms}^{-1} \times 0.032/100) = -9.6 \times 10^4 \text{ ms}^{-1}.$$

The negative sign shows that the star is receding away. This made the astrophysists to conclude that the world is in a state of expansion

INTEXT QUESTIONS 14.8

1. A SONAR system fixed in a submarine operates at frequency 40.0kHz. An enemy submarine moves towards it with a speed of 100ms^{-1}. Calculate the frequency of the sound reflected by the sonar. Take the speed of sound in water to be 1450 ms^{-1}.
2. An engine, blowing a whistle of frequency 200Hz moves with a velocity 16ms^{-1} towards a hill from which a well defined echo is heard. Calculate the frequency of the echo as heard by the driver. Velocity of sound in air is 340ms^{-1}.

Constancy of Speed of Light

Aristotle, believed that light travels with infinite velocity. It was for the first time in September, 1876 that the Danish astronomer, Roemer, indicated in a meeting of Paris Academy of Sciences that the anomalous behaviour of the eclipse, times of Jupiter's inner satellite, Io, may be due to the finite speed of light. Feazeu, Focult, Michelson and many other scientists carried out experiments to determine the speed of light in air with more and more precision.

Albert Einstein, in his 1905 paper, on special theory of relativity, based his arguments on two postulates. One of the postulates was the constancy of speed of light in vacuum, irrespective of the wavelength of light, the velocity of the source or the observer. In 1983, the velocity of light in vacuum, was declared a universal constant with a value 299792458 ms^{-1}.

However, the Autralian researcher Barry Setterfield and Trevn Norwah have studied, the data of 16 different experiments on the speed of light in vacuum, carried out over the last 300 years, by different scientists at different places. According to them, the speed of light in vacuum is decreasing with time. If this hypothesis is sustained and coroborated by experiements, it will bring in thorough change in our world view. Major areas in which this change will be enormous are : Maxwell's laws, atomic structure, radioactive decay, gravitation, concepts of space, time and mass etc.

WHAT YOU HAVE LEARNT

- The distance between two nearest points in a wave motion which are in the same phase is called wavelength.
- The equation of a simple harmonic wave propagating along x–axis is $y = a \sin(\upsilon t - kx)$.
- The energy transmitted per second across a unit area normal to it is called intensity..
- If the vibrations of medium particle are perpendicular to the direction of propagation, the wave is said to be transverse but when the vibrations are along the direction of propagation the wave is said to be longitudinal. Velocities of transverse wave and longitudinal waves is given by $\upsilon = \sqrt{T/m}$ and $\upsilon = \sqrt{E/\rho}$ respectively.
- On reflection from a denser medium, phase is reversed by π. But there is no phase reversal on reflection from a rarer medium.
- When two waves are superposed, the resultant displacement at any point is vector sum of individual displacements at that point. Superposition of two colliner waves of same frequency but differing phases, when moving in the same direction results in redistribution of energy giving rise to interference pattern.
- Superposition of two collinear waves of the same frequency and same amplitude travelling in the opposite directions with the same speed results in the formation of stationary waves. In such waves, waveform does not move.
- In a stationary wave, the distance between two successive nodes or successive antinodes is $\lambda/2$. It is, therefore, obvious that between two nodes, there is an antinode and between two antinodes there is a node.

- The displacement is maximum at antinodes and minimum at nodes.
- Intensity level is defined by the equation $\beta = 10\log (I/I_0)$, where I_0 is an arbitrarily chosen reference intensity of 10^{-12} W m^{-2}. Intensity level is expressed in decibels (Symbol. db)
- Quality of a note is the characteristic of musical sounds which enable us to distinguish two notes of the same pitch and same loudness but sounded by two different instruments.
- Electromagnetic waves are transverse in nature, and do not require any medium for their propagation.
- Light is an e.m. wave with wavelength in the range 4000 Å – 7500 Å.
- The frequency of e.m. waves does not change with the change in the medium.
- e.m. waves are used for wireless radio communication, TV transmission, satellite communication etc.

TERMINAL EXERCISES

1. How will you define a wave in the most general form?
2. Explain using a suitable mechanical model, the propagation of (i) transverse waves (ii) longitudinal wave. Define the term wavelength and frequency.
3. Define angular frequency ω and propagation constant k and hence show that the velocity of the wave propagation is $v = \omega/k = n\lambda$.
4. Derive the equation of a simple harmonic wave of angular frequency of (i) transverse (ii) longitudinal waves.
5. What are the essential properties of the medium for propagation of (i) transverse waves (ii) longitudinal waves.
6. Derive an expression for the intensity of the wave in terms of density of the medium, velocity of the wave, the amplitude of the wave and the frequency of the wave.
7. Write Newton's formula for the velocity of sound in a gas and explain Laplace's correction.
8. When do two waves interfere (i) constructively (ii) destructively?
9. Show using trigonometry that when two simple harmonic waves of the same angular frequency ω and same wavelenght λ but of amplitudes a_1 and a_2 are superposed, the resultant amplitude is $A = \sqrt{a_2^2 + a_2^2 + 2a_1a_2 \cos\theta}$, where θ is the phase difference between them. What would be the value of A, for θ = 0, (ii) for θ = 2π, and (iii) for θ = (2m + 1)π?

10. What are beats? How are they formed? Explain graphically.
11. Discuss graphically the formation of stationary waves. Why are these wave called stationary waves? Define nodes and antinodes.
12. State three differences between stationary and travelling waves.
13. Derive the equation of a stationary wave and show that displacement nodes are pressure antinodes and displacement antinodes are pressure nodes?
14. What are the characteristics of musical sounds. Explain.
15. What is a decibel (symbol) db)? What is meant by 'threshold of hearing' and 'threshold of feeling'?
16. What is meant by quality of sound? Explain with examples?
17. Discuss the harmonics of organ pipes. Show that an open pipe is richer in harmonics.
18. Show that (i) the frequency of open organ pipes. is two times the frequency of the fundamental note of a closed pipe of same length (ii) to produce a fundamental note of same frequency, the length of the open pipe must be two times the length of the closed pipe.
19. Describe an experiment to demonstrate existence of nodes and antinodes in an organ pipes?
20. State the causes of noise pollution, its harmful effects and methods of minimising it.
21. Explain Doppler's effect and derive an expression for apparent frequency. How does this equation get modified if the medium in which the sound travels is also moving.
22. Discuss the applications of Doppler's effect in (i) measuring the velocity of recession of stars, (ii) velocity of enemy plane by RADAR and (iii) velocity of enemy boat by SONAR?
23. Calculate the velocity of sound in a gas in which two waves of wavelengths 1.00m and 1.01m produce 10 beats in 3 seconds.
24. What will be the length of a closed pipe if the lowest note has a frequency 256Hz at 20C. Velocity of sound at 0C = 332 ms^{-1}.
25. The frequency of the sound waves emitted by a source is 1 kHz. Calculate the frequency of the waves as perceived by the observer when (a) the source and the observer are stationary, (b) the source is moving with a velocity of 50ms^{-1} towards the observer, and (c) the source is moving with a velocity of 50ms^{-1} away from the observer. Velocity of sound in air is 350ms^{-1}.
26. Write the characteristic properties of e.m. waves which make them different from sound waves.

27. How does the velocity of e.m. waves depend upon the permeability μ and permittivity ε of the medium through which they pass?

28. Give the range of wavelengths of the following e.m. waves:

    (i) Radio Waves (ii) Microwaves : (iii) Ultraviolet; (iv) x-rays.

29. How are x-rays produced?

30. Can e.m. waves of all frequencies propagate through vacuum?

31. Fill in the blanks.

    (i) A changing electric field produces a______________ in the adjacent region.

    (ii) ___________ are more harmful to our eyes than x-rays.

    (iii) __________ are emitted from radio active nuclei of cobalt.

    (iv) Infra red rays are less energies than_______________

    (v) In an e.m. wave propagating along z-direction, if the E field oscillates in the X,Z plane then the B field will oscillate in the _______________ plane.

    (vi) The ratio $\frac{E}{H}$ in free space of e.m. wave is called ______________.

    (vii) The frequency range of F.M. band is ________________.

    (viii)___________signal is frequency modulated in T.V. broadcasting.

ANSWERS TO INTEXT QUESTIONS

14.1

1. See section 14.1.4.

2. If p be the path difference, then the phase difference is $\theta = \frac{2\pi}{\lambda} p$.

3. ϕ

14.2

1. Newton assumed that compression and rarefaction caused by sound waves takes place under isothermal condition.

3. Newton assumed that isothermal conditions instead of adiabatic conditions for sound propagation.

4. 357^{0}C.

Notes

5. $v = \sqrt{\frac{T}{m}}$

6. Therefore, $n = \frac{1}{\lambda}\sqrt{\frac{T}{m}}$

Further, for the simplest mode of vibration, at the two ends of the string, there are nodes and in between the two nodes is an antinode. Therefore, $l = l/2$ or $\lambda = 2l$, hence $n = \lambda/2l \sqrt{\frac{T}{m}}$. If the string vibrates in p segments, the $\lambda = p\ l/2$ or $\lambda = 2l/p$. Then $n = (p/2l)\sqrt{\frac{T}{m}}$.

14.3

For answers to all questions see text.

14.4

1. 25/9.
2. Beats with frequency 4Hz are produced.
3. Frequency of beat is Δv.
4. 517, on loading the frequency of A decreases from 517 to 507.

14.5

1. No energy swings back and forth in a segment.
2. Distance between two successive nodes is $\lambda/2$, and between a node and antinode is $\lambda/4$.
4. (i) 1m, (ii) 1m, (iii) 1/4m.

14.6

1. Pitch increases with increase in frequency.
2. Timbre
3. Timbre
4. Open pipe
5. For a closed pipe in case of fundamental note $l = \lambda/4$ or $\lambda = 4l$, therefore $n = \upsilon/\lambda = \upsilon/4l$.

For an open pipe $\ell = \lambda/2$. Therefore $n' = \upsilon/2l$.

Comparing (i) and (ii) we find that $n' = 2n$

6. $n = \frac{\upsilon}{2\ell}$. As υ increases with increase in temperature n also increases.

14.7

(*i*) microwaves.

(*ii*) yellow–green ($\lambda = 5 \times 10^{-7}$ m)

(*iii*) Sun.

(*iv*) X – rays.

(*v*) thermopile.

2. (*i*) ultra voilet

   (ii) r – rays.

3. Microwaves

4. Ozone.

5. Perpendicular to each other.

14.8

1. $n' = n \frac{c - \upsilon_0}{c}$

$$= 40 \times 10^3 \times \frac{1450 - 100}{1450}$$

$$= 40 \times \frac{135}{145} \times 10 = 37.2 \text{ KHz.}$$

2. $n' = 200 \times \frac{340 + 16}{340 - 16}$

$$= 200 \times \frac{356}{224} = 220 \text{ Hz.}$$

Answer to Terminal Problems

23. 337 ms^{-1}

24. ~30 cm.

25. (*a*) 1 kHz

    (*b*) 857 Hz

    (*c*) 1143 Hz.

312en15

15

ELECTRIC CHARGE AND ELECTRIC FIELD

So far you have learnt about mechanical, thermal and optical systems and various phenomena exhibited by them. The importance of electricity in our daily life is too evident. The physical comforts we enjoy and the various devices used in daily life depend on the availability of electrical energy. An electrical power failure demonstrates directly our dependence on electric and magnetic phenomena; the lights go off, the fans, coolers and air-conditioners in summer and heaters and gysers in winter stop working. Similarly, radio, TV, computers, microwaves can not be operated. Water pumps stop running and fields cannot be irrigated. Even train services are affected by power failure. Machines in industrial units can not be operated. In short, life almost comes to a stand still, sometimes even evoking public anger. It is, therefore, extremely important to study electric and magnetic phenomena.

In this lesson, you will learn about two kinds of electric charges, their behaviour in different circumstances, the forces that act between them, the behaviour of the surrounding space etc. Broadly speaking, we wish to study that branch of physics which deals with electrical charges at rest. This branch is called **electrostatics**.

OBJECTIVES

After studying this lesson, you should be able to :

- *state the basic properties of electric charges;*
- *explain the concepts of quantisation and conservation of charge;*
- *explain Coulomb's law of force between electric charges;*
- *define electric field due to a charge at rest and draw electric lines of force;*
- *define electric dipole, dipole moment and the electric field due to a dipole;*

- *state Gauss' theorem and derive expressions for the electric field due to a point charge, a long charged wire, a uniformly charged spherical shell and a plane sheet of charge; and*
- *describe how a van de Graaff generator functions.*

15.1 FRICTIONAL ELECTRICITY

The ancient Greeks observed electric and magnetic phenomena as early as 600 B.C. They found that a piece of amber, when rubbed, becomes electrified and attracts small pieces of feathers. The word **electric** comes from Greek word for amber meaning **electron**.

You can perform simple activities to demonstrate the existence of charges and forces between them. If you run a comb through your dry hair, you will note that the comb begins to attract small pieces of paper. Do you know how does it happen? Let us perform two simple experiments to understand the reason.

ACTIVITY 15.1

Take a hard rubber rod and rub it with fur or wool. Next you take a glass rod and rub it with silk. Suspend them (rubber rod and a glass rod) separately with the help of non-metallic threads, as shown in Fig. 15.1.

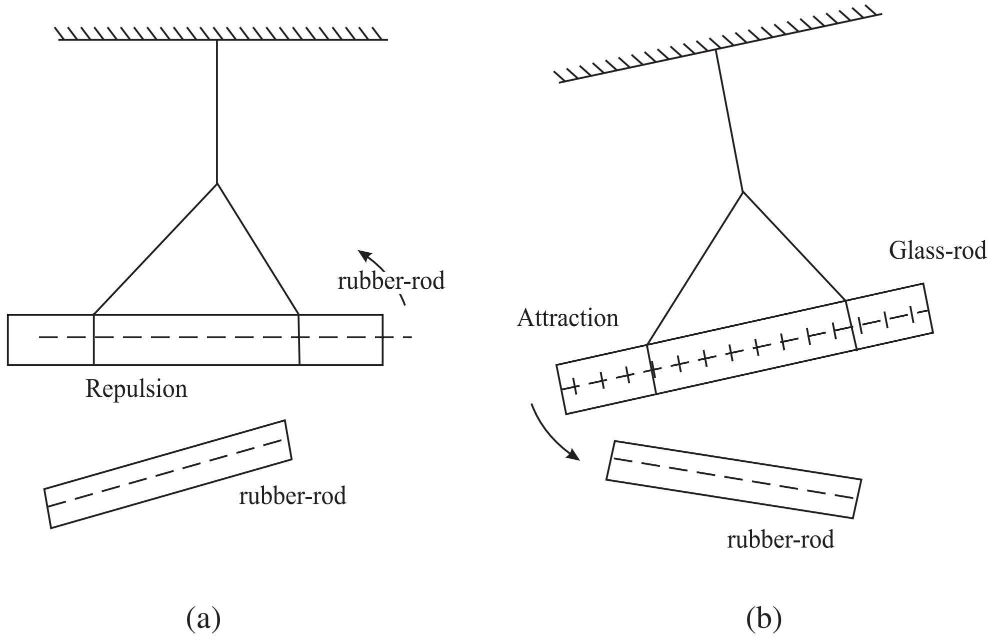


Fig. 15.1: Force of attraction/repulsion between charges: a) a charged rubber rod repels another charged rubber rod : like charges repel each other; and b) a charged glass rod attracts a charged rubber rod : unlike charges attract each other.

Now bring rubber rod rubbed with wool near these rods one by one. What do you observe? You will observe that

- when a charged rubber rod is brought near the charged (suspended) rubber rod, they show repulsion [Fig. 15.1(a)]; and
- when the charged rubber rod is brought near the (suspended) charged glass rod, they show attraction [Fig 15.1(b)].

Similar results will be obtained by bringing a charged glass rod.

On the basis of these observations, we can say that

- A charged rubber rod attracts a charged glass rod but repels a charged rubber rod.
- A charged glass rod repels a charged glass rod but attracts a charged rubber rod.

From these activities we can infer that the rubber rod has acquired one kind of electricity and the glass rod has acquired another kind of electricity. Moreover, **like charges repel and unlike charges attract each other.**

Franklin (Benjamin Franklin, 1706 -1790) suggested that the charge on glass rod is to be called **positive** and that on the rubber rod is to be called **negative.** We follow this convention since then.

Once a body is charged by friction, it can be used to charge other conducting bodies by

conduction, i.e., by touching the charged body with an uncharged body; and

induction, i.e., by bringing the charged body close to an uncharged conductor and earthing it. Subsequently, the charged body and the earthing are removed simultaneously.

15.1.1 Conservation of Charge

In Activity 15.1, you have seen that when a glass rod is rubbed with silk, the rod acquires positive charge and silk acquires negative charge. Since both materials in the normal state are neutral (no charge), the positive charge on the glass rod should be equal in magnitude to the negative charge on silk. This means that the total charge of the system (glass + silk) is conserved. It is neither created nor destroyed. It is only transferred from one body of the system to the other. The transfer of charges takes place due to increase in the thermal energy of the system when the glass rod is rubbed; the less tightly bound electrons from the glass rod are transferred to silk. The glass rod (deficient in electrons) becomes positively charged and silk, which now has excess electrons, becomes negatively charged. When rubber is rubbed with fur, electrons from the fur are transferred to rubber.

That is, rubber gains negative charge and fur gains an equal amount of positive charge. Any other kind of charge (other than positive and negative) has not been found till today.

15.1.2 Quantisation of Charge

In 1909, Millikan (Robert Millikan, 1886-1953) experimentally proved that charge always occurs as some integral multiple of a fundamental unit of charge, which is taken as the charge on an electron. This means that if Q is the charge on an object, it can be written as $Q = Ne,$ where N is an integer and e is charge on an electron. Then we say that charge is quantised. It means that a charged body cannot have 2.5e or 6.4e amount of charge. In units 24-26, you will learn that an electron has charge $-e$ and a proton has charge $+e$. Neutron has no charge. Every atom has equal number of electrons and protons and that is why it is neutral. From this discussion, we can draw the following conclusions :

- There are only two kinds of charges in nature; positive and negative.
- Charge is conserved.
- Charge is quantised.

INTEXT QUESTIONS 15.1

1. A glass rod when rubbed with silk cloth acquires a charge $q = +3.2\times10^{-17}$ C.

   i) Is silk cloth also charged?

   ii) What is the nature and magnitude of the charge on silk cloth?

2. There are two identical metallic spheres A and B. A is given a charge $+Q$. Both spheres are then brought in contact and then separated.

   (i) Will there be any charge on B ?

   (ii) What will the magnitude of charge on B, if it gets charged when in contact with A.

3. A charged object has $q = 4.8\times10^{-16}$ C. How many units of fundamental charge are there on the object? (Take $e = 1.6\times10^{-19}$ C).

15.2 COULOMB'S LAW

You have learnt that two stationary charges either attract or repel each other. The force of attraction or repulsion between them depends on their nature. Coulomb studied the nature of this force and in 1785 established a fundamental law governing it. From experimental observations, he showed that the electrical force between two static point charges q_1 and q_2 placed some distance apart is

- directly proportional to their product ;
- inversely proportional to the square of the distance r between them;
- directed along the line joining the two charged particles ; and
- repulsive for same kind of charges and attractive for opposite charges.

The magnitude of force F can then be expressed as

$$F = k\frac{q_1 \times q_2}{r^2} \qquad (15.1)$$

For free space, we write

$$F = \frac{1}{4\pi\varepsilon_0} \times \frac{q_1 \times q_2}{r^2} \qquad (15.2)$$

where constant of proportionality $k = \frac{1}{4\pi\varepsilon_0}$ for free space (vacuum) and $k = \frac{1}{4\pi\varepsilon}$ for a material medium. ε_0 is called **permittivity** of free space and ε is the permittivity of the medium. It means that if the same system of charges is kept in a material medium, the magnitude of Coulomb force will be different from that in free space.

The constant k has a value which depends on the units of the quantities involved. The unit of charge in SI system is coulomb (C). The coulomb is defined in terms of the unit of current, called **ampere.** (You will learn about it later.) In SI system of units, the value of k is

$$k = \frac{1}{4\pi\varepsilon_0} = 9 \times 10^9 \ \frac{\text{N m}^2}{\text{C}^2} \qquad (15.3)$$

since $\varepsilon_0 = 8.85 \times 10^{-12}$ $C^2 N^{-1} m^{-2}$.

Thus in terms of force, one coulomb charge can be defined as ***: If two equal charges separated by one metre experience a force of* 9 × 10⁹ N, *each charge has a magnitude of one coulomb.*** The value of electronic charge *e is* 1.60×10^{-19} C.

Note that

- Coulomb's law is also an inverse square law just like Newton's law of Gravitation, which you studied in lesson 6.
- Coulomb's law holds good for point charges only.
- Coulomb's force acts at a distance, unlike mechanical force.

How Big is One Coulomb?

The unit of electrical charge is coulomb. Have you ever thought : How big a coulomb is? To know this, let us calculate the magnitude of force between two charges, each of one coulomb, placed at a distance of one metre from one another:

$$|\mathbf{F}| = k \times \frac{q_1 \times q_2}{r^2}$$

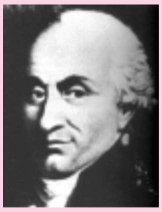
$$= 9.0 \times 10^9 \times \frac{1 \times 1}{1^2}$$

$$= 9.0 \times 10^9 \approx 10^{10}\,\text{N}$$

If the mass of a loaded passenger bus is 5000 kg, its weight mg = (5000 × 10) N (assume $g \approx 10$ m s^{-2}) = 5 × 10^4 N.

Let us assume that there are 10,000 such loaded buses in Delhi. The total weight of all these buses will be 5 × 10^4 × 10,000 = 5 × 10^8 N. If there are 10 cities having same number of buses as those in Delhi, the total weight of all these loaded buses will be 5 × 10^9 N. It means that the force between two charges, each of 1C and separated by on metre is equivalent to the weight of about two hundred thousand buses, each of mass 5000 kg.

Charles Augustin de Coulomb (1736–1806)



A French physicist, Coulomb started his career as military engineer in West Indies. He invented a torsional balance and used it to perform experiments to determine the nature of interaction forces between charges and magnets. He presented the results of these experiments in the form of Coulomb's law of electrostatics and Coulomb's law of magnetostatics. The SI unit of charge has been named in his honour.

You now know that the ratio of forces between two point charges q_1 and q_2 separated by a distance r, when kept in free space (vacuum) and material medium, is equal to $\varepsilon/\varepsilon_0$:

$$\frac{F_0 \text{ (in vaccum)}}{F \text{ (in medium)}} = \frac{\varepsilon}{\varepsilon_0} = \varepsilon_r$$

where ε_r is known as relative permittivity or **dielectric constant.** Its value is always greater than one. We will define dielectric constant in another form later.

15.2.1 Vector Form of Coulomb's Law

You know that force is a vector quantity. It means that force between two charges should also be represented as a vector. That is, Eqn. (15.1) should be expressed in vector form. Let us learn to do so now.

Let there be two point charges q_1 and q_2 separated by a distance r (Fig. 15.3). Suppose that $\mathbf{F}_{12}$ denotes the force experienced by q_1 due to the charge q_2 and $\mathbf{F}_{21}$ denotes the force on q_2 due to charge q_1. We denote the unit vector pointing from q_1 to q_2 by $\hat{\mathbf{r}}_{12}$. Then from Fig. 15.3 (a), it follows that

$$\mathbf{F}_{12} = k\,\frac{q_1\, q_2}{|r_{12}^2|}\,\hat{\mathbf{r}}_{12} \tag{15.4}$$

Similarly, for charges shown in Fig. 15.3 (b), we can write

$$\mathbf{F}_{21} = -\,k\,\frac{q_1\, q_2}{|r_{12}^2|}\,\hat{\mathbf{r}}_{12} \tag{15.5}$$

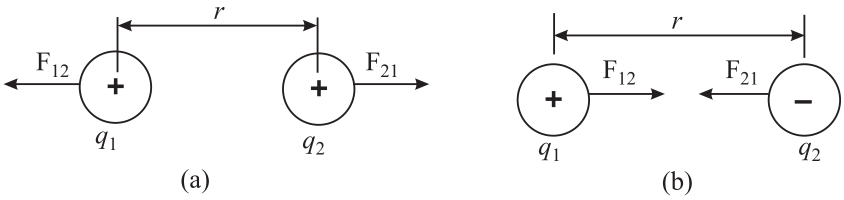


Fig. 15.3 : Two point charges q_1 and q_2 separated by a distance r : a) the direction of forces of repulsion between two positive charges, and b) the direction of forces of attraction between a positive and a negative charge.

The positive sign in Eqn. (15.4) indicates that the force is repulsive and the negative sign in Eqn. (15.5) indicates that the force is attractive.

The Coulomb's law obeys the principle of action and reaction between two charges q_1 and q_2. Therefore,

$$\mathbf{F}_{12} = -\,\mathbf{F}_{21} \tag{15.6}$$

In general, we can write the expression for force between two charges as

$$\mathbf{F}_{12} = k \times \frac{q_1\, q_2}{r^2}\,\hat{\mathbf{r}}_{12} \tag{15.7}$$

15.2.2 Principle of Superposition

If there are more than two charges, we can calculate the force between any two charges using Eqn. (15.7). Suppose now that there are several charges q_1, q_2, q_3, q_4, etc. The force exerted on q_1 due to all other charges is given by Eqn. (15.7):

$$\mathbf{F}_{12} = k \frac{q_1 q_2}{|r_{12}^2|} \hat{\mathbf{r}}_{12}$$

$$\mathbf{F}_{13} = k \frac{q_1 q_3}{|r_{13}^2|} \hat{\mathbf{r}}_{13}$$

and $$\mathbf{F}_{14} = k \frac{q_1 q_4}{|r_{14}^2|} \hat{\mathbf{r}}_{14} \qquad (15.8)$$

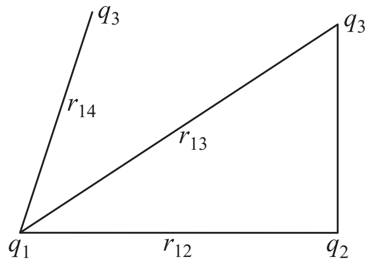


Fig. 15.4: Principle of superposition

The resultant of all these forces, i.e., the total force **F** experienced by q_1 is their vector sum:

$$\mathbf{F} = \mathbf{F}_{12} + \mathbf{F}_{13} + \mathbf{F}_{14} + \ldots \qquad (15.9)$$

This is known as principle of superposition.

Example 15.1 : A charge $+q_1 = 12$C is placed at a distance of 4.0 m from another charge $+q_2 = 6$C, as shown in the Fig. 15.5. Where should a negative charge q_3 be placed on the line joining q_1 and q_2 so that the charge q_3 does not experience any force?

Solution : Let q_3 be placed between q_1 and q_2 at a distance of x metre from q_1. (It can be easily seen that on placing q_3 on the left of q_1 or on the right of q_2 or at any position other than the one between the line joining q_1 and q_2, the resultant force can not be zero.) The force exerted on q_3 by q_1 will be

$$\mathbf{F}_{31} = k \frac{q_1 q_3}{r_{31}^2} \hat{\mathbf{r}}_{31} \text{ towards } q_1$$

$$\therefore \qquad |\mathbf{F}_{31}| = k \frac{q_3 q_1}{x^2}$$

The magnitude of force on q_3 due to q_2 is given by

$$|\mathbf{F}_{32}| = k \frac{q_3 q_2}{(4-x)^2} \text{ towards } q_2$$

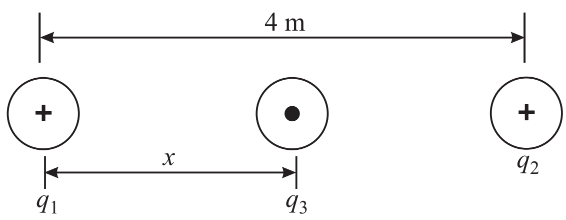


Fig. 15.5 : Three point charges q_1, q_2 and q_3 placed in a straight line

The resultant force on q_3 will be zero when $\mathbf{F}_{31} = \mathbf{F}_{32}$. Therefore, on substituting the numerical values, we get

$$k \times \frac{12q_3}{x^2} = k \times \frac{6q_3}{(4-x)^2}$$

Note that $6q_3k$ is common on both sides and cancels out. Therefore, on simplification, we get

$$\frac{2}{x^2} = \frac{1}{(4-x)^2}$$

or $\qquad 2(4-x)^2 = x^2$

$\Rightarrow \qquad x^2 - 16x + 32 = 0$

On solving this, we get two values of x : 2.35 m and 13.65 m. The latter value is inadmissible because it goes beyond q_2. Therefore, the charge q_3 should be placed at a distance of 2.35 m from q_1.

It is a reasonable solution qualitatively also. The charge q_1 is stronger than q_2. Hence the distance between q_1 and q_3 should be greater than that between q_2 and q_3.

> The roots of a quadratic equation of the form
>
> $$ax^2 + bx + c = 0$$
>
> are given by
>
> $$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$
>
> In this case, $a = 1$, $b = -16$ and $c = 32$.
>
> $$\therefore \quad x = \frac{16 \pm \sqrt{256 - 4 \times 32}}{2}$$
>
> $$= 2.35,\ 13.65$$

Example 15.2 : Two charges, each of 6.0×10^{-10} C, are separated by a distance of 2.0 m. Calculate the magnitude of Coulomb force between them.

Solution : We know that the magnitude of Coulomb force between two charges is given by Eqn. (15.2) :

$$F = k\frac{q_1 \cdot q_2}{r^2}$$

Given, $q_1 = q_2 = 6.0 \times 10^{-10}$C and $r = 2.0$ m, Therefore on putting these values, we get

$$F = \frac{(9\times10^{9}\ \text{N m}^2\text{C}^{-2})\times(6.0\times10^{-10}\text{C})^2}{2^2\ \text{m}^2}$$

$$= \frac{9 \times 10^9 \times 36.0 \times 10^{-20}}{4}\ \text{N}$$

$$= 81 \times 10^{-11}\ \text{N}$$

INTEXT QUESTIONS 15.2

1. Two charges q_1 = 16μC and q_2 = 9 μC are separated by a distance 12m. Determine the magnitude of the force experienced by q_1 due to q_2 and also the direction of this force. What is the direction of the force experienced by q_2 due to q_1?

2. There are three point charges of equal magnitude q placed at the three corners of a right angle triangle, as shown in Fig. 15.2. $AB = AC$. What is the magnitude and direction of the force exerted on $-q$?

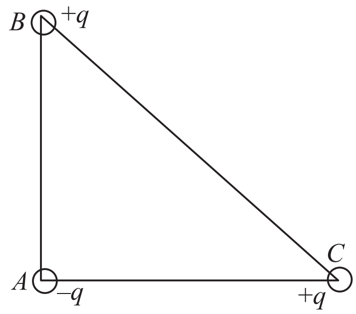


Fig. 15.2 : Three charges placed at the three corners of a right angle triangle.

15.3 ELECTRIC FIELD

To explain the interaction between two charges placed at a distance, Faraday introduced the concept of electric field. The electric field **E** at a point is defined as the electric force **F** experienced by a positive test charge q_0 placed at that point divided by the magnitude of the test charge. Mathematically, we write

$$\mathbf{E} = \frac{\mathbf{F}}{q_0} \qquad (15.10)$$

This is analogous to the definition of acceleration due to gravity, $\mathbf{g} = \mathbf{F}/m_0$, experienced by mass m_0 in the gravitational field **F**.

The electric field **E** is a vector quantity and has the same direction as the electric force **F**. Note that the electric field is due to an external charge and not due to the test charge. The test charge q_0 should, therefore, be so small in magnitude that it does not disturb the field due to external charge. (In practice, however, even the smallest test charge will disturb the external field.) Strictly speaking, mathematical definition given below is more accurate :

$$\mathbf{E} = \lim_{q_0 \to 0} \frac{\mathbf{F}}{q_0} \qquad (15.11)$$

In SI system, the force is in newton and the charge is in coulomb. Therefore, according to Eqn.(15.10), the electric field has the unit newton per coulomb. The direction of **E** is same as that of **F**. Note that *the action of electric force is mediated through electric field.*

Let us now examine why the test charge q_0 should be infinitesimally small.

Refer to Fig. 15.6. It shows a uniformly charged metallic sphere with charge q and a test charge $q_0 (<< q)$. It means that charge density per unit area is same around points A, B, C and D. The test charge q_0 must measure the force **F** without disturbing the charge distribution on the sphere. Fig. 15.6 (b) shows the situation when $q \simeq q_0$. In this case, the presence of the test charge modifies the surface

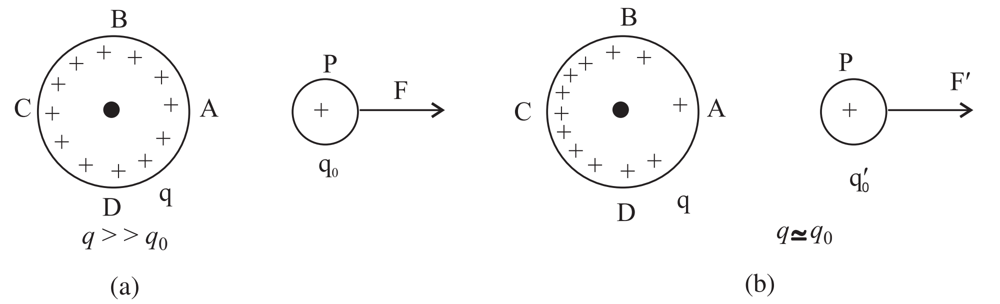


Fig. 15.6 : a) uniformly charged metallic sphere and a test charge, and **b)** redistribution of charge on the sphere when another charge is brought near it.

charge density. As a result, the electrical force experienced by the test charge q_0 will also change, say from **F** to **F′**. That is, the force in the presence of test charge is different from that in its absence. But without q_0, the force cannot be measured. If q_0 is infinitesimally small in comparison to q, the charge distribution on the sphere will be minimally affected and the results of measurement will have a value very close to the true value. That is, **F′** will be very nearly equal to **F**. We hope you now appreciate the point as to why the test charge should be infinitesimally small.

Let there be a point charge q. A test charge q_0 is placed at a distance r from q. The force experienced by the test charge is given by

$$\mathbf{F} = k\frac{qq_0}{r^2}\hat{\mathbf{r}} \tag{15.12}$$

The electric field is defined as the force per unit charge. Hence

$$\mathbf{E} = k \times \frac{q}{r^2}\hat{\mathbf{r}} \tag{15.13}$$

If q is positive, the field **E** will be directed away from it. If q is negative, the field **E** will be directed towards it. This is shown in Fig. 15.7.

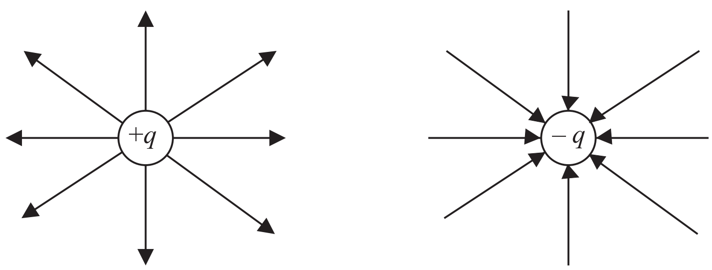


Fig. 15.7 : Direction of electric field due to positive and negative charges

The principle of superposition applies to electric field also. If there are a number of charges q_1, q_2, q_3, ..., the corresponding fields at a point P according to Eqn. (15.13) are

$$\mathbf{E}_1 = k \times \frac{q_1}{r_1^2}\,\hat{\mathbf{r}}_1, \quad \mathbf{E}_2 = k \times \frac{q_2}{r_2^2}\,\hat{\mathbf{r}}_2 \text{ and } \mathbf{E}_3 = k \times \frac{q_3}{r_3^2}\,\hat{\mathbf{r}}_3$$

The total field at point P due to all charges is the vector sum of all fields. Thus,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 + \ldots$$

or

$$\mathbf{E} = k \sum_{i=1}^{N} \frac{q_i\,\hat{\mathbf{r}}_i}{r_i^2} \qquad (15.15)$$

where r_i is the distance between P and charge q_i and $\hat{\mathbf{r}}_i$ is the unit vector directed from q_i to P. The force on a charge q in an electric field **E** is

$$\mathbf{F} = q\,\mathbf{E} \qquad (15.16)$$

Example 15.3 : The electric force at some point due to a point charge $q = 3.5\mu C$ is 8.5×10^{-4} N. Calculate the strength of electric field at that point.

Solution : From Eq. (15.16) we can write

$$E = \frac{F}{q} = \frac{8.5\times10^{-4}\,\text{N}}{3.5\times10^{-6}\,\text{C}}$$

$$= 2.43 \times 10^2\ \text{NC}^{-1}$$

Example 15.4 : Three equal positive point charges are placed at the three corners of an equilateral triangle, as shown in Fig. 15.8. Calculate the electric field at the centroid P of the triangle.

Solution : Suppose that a test charge q_0 has been placed at the centroid P of the triangle. The test charge will experienced force in three directions making same angle between any two of them. The resultant of these forces at P will be zero. Hence the field at P is zero.

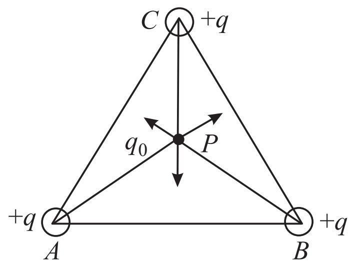


Fig. 15.8 : Electric field at the centroid of an equilateral triangle due to equal charges at its three corners is zero.

INTEXT QUESTIONS 15.3

1. A charge + Q is placed at the origin of co-ordinate system. Determine the direction of the field at a point P located on

   a) + x-axis   b) + y-axis   c) x = 4 units and y = 4 units

2. The ΔABC is defined by $AB = AC = 40$ cm. And angle at A is 30°. Two charges, each of magnitude 2×10^{-6} C but opposite in sign, are placed at B and C, as shown in Fig. 15.9. Calculate the magnitude and direction of the field at A.

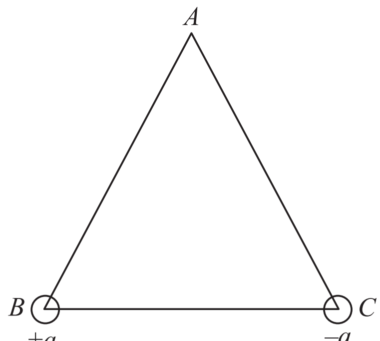


Fig. 15.9

3. A negative charge is located in space and the electric field is directed towards the earth. What is the direction of the force on this charge?

4. Two identical charges are placed on a plane surface separated by a distance d between them. Where will the resultant field be zero?

15.3.1 Electric Field due to a Dipole

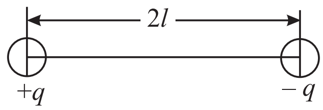


Fig. 15.10: Two unlike charges of equal magnitude separated by a small distance form a dipole.

If two equal and opposite charges are separated by a small distance, the system is said to form a dipole**.** The most familiar example is H_2O. Fig 15.10 shows charges $+q$ and $-q$ separated by a small distance $2l$. The product of the magnitude of charge and separation between the charges is called **dipole moment,** p **:**

$$p = q \times 2l \qquad (15.17)$$

Its SI unit is coulomb-metre.

The dipole moment is a vector quantity. Eqn. (15.17) gives its magnitude and its direction is from negative charge to positive charge along the line joining the two charges (axis of the dipole). Having defined a dipole and dipole moment, we are now in a position to calculate the **electric field due to a dipole.** The calculations are particularly simple in the following cases.

CASE I : *Electric field due to a dipole at an axial point : End–on position*

To derive an expression for the electric field of a dipole at a point P which lies on the axis of the dipole, refer to Fig. 15.11.This is known as **end-on position.** The point charges $-q$ and $+q$ at points A and B are separated by a distance $2l$. The point O is at the middle of AB. Suppose that point P is at a distance r from the mid point O. Then electric field at P due to $+q$ at B is given by

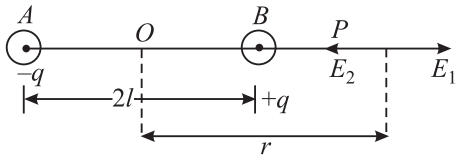


Fig. 15.11 : Field at point P on the dipole axis

$$\mathbf{E}_1 = k \times \frac{q}{(r-l)^2} \text{ in the direction } AP$$

Similarly, the electric field $\mathbf{E}_2$ at P due to $-q$ is given by

$$\mathbf{E_2} = k \times \frac{q}{(r+l)^2} \text{ in the direction } PA$$

The resultant field $\mathbf{E}$ at P will be in the direction of $\mathbf{E}_1$, since $\mathbf{E}_1$ is greater than $\mathbf{E}_2$ [as $(r - l)$ is less than $(r + l)$]. Hence

$$\mathbf{E} = \frac{kq}{(r-l)^2} - \frac{kq}{(r+l)^2}$$

$$= kq\left[\frac{1}{(r-l)^2} - \frac{1}{(r+l)^2}\right]$$

$$= kq\left[\frac{(r+l)^2 - (r-l)^2}{(r^2-l^2)^2}\right]$$

$$= kq \times \frac{4lr}{(r^2-l^2)^2}$$

$(a + b)^2 - (a - b)^2 = 4ab$
$(a + b)(a - b) = a^2 - b^2$

$$= k\frac{(2lq)\,2r}{(r^2-l^2)^2}$$

$$= k\frac{2\mathbf{p}r}{(r^2-l^2)^2}$$

where dipole moment $\mathbf{p} = 2lq$. Since $k = 1/4\pi\varepsilon_0$, we can rewrite it as

$$\mathbf{E} = \frac{2\mathbf{p}}{4\pi\varepsilon_0} \times \frac{r}{r^4\,(1 - l^2/r^2)^2}$$

If $r >> l$, l^2/r^2 will be very small compared to 1. It can even be neglected and the expression for electric field then simplifies to

$$\mathbf{E} = \frac{2\mathbf{p}}{4\pi\varepsilon_0 r^3} \quad (15.18)$$

It shows that electric field is in the direction of **p** and its magnitude is inversely proportional to the third power of distance of the observation point from the centre of the dipole.

CASE II : *Electric field due to a dipole at a point on the perpendicular bisector : Broad-on position*

Suppose that point P lies on the perpendicular bisector of the line joining the charges shown in Fig. 15.12. Note that $AB = 2l$, $OP = r$, and $AO = OB = l$.

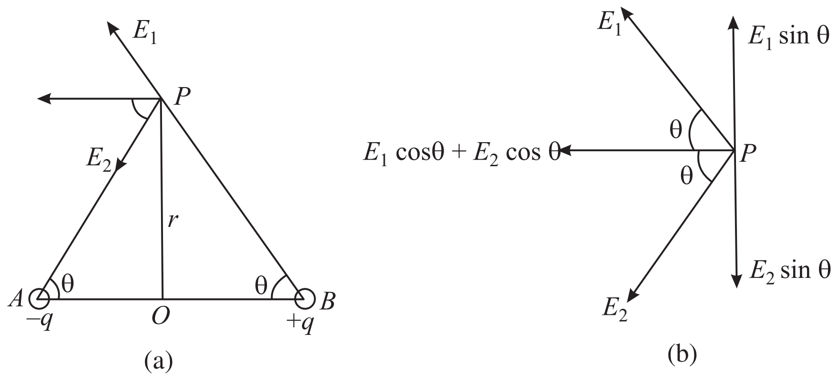


Fig. 15.12 : a) Field at point P on the perpendicular bisector of the line joining the charges, and b) resolution of field in rectangular components.

The angle θ is shown in Fig. 15.12(a). From right angled Δs *PAO* and *PBO*, we can write

$$AP = BP = \sqrt{l^2 + r^2}$$

The field at P due to charge $+q$ at B in the direction of BP can be written as

$$\mathbf{E}_1 = k\frac{q}{l^2 + r^2}$$

Similarly, the field at P due to charge at A in the direction of PA is given as

$$\mathbf{E}_2 = k\frac{q}{l^2 + r^2}$$

Note that the magnitudes of $\mathbf{E}_1$ and $\mathbf{E}_2$ are equal.

Let us resolve the fields $\mathbf{E}_1$ and $\mathbf{E}_2$ parallel and perpendicular to AB. The components parallel to AB are $\mathbf{E}_1\cos\theta$ and $\mathbf{E}_2\cos\theta$, and both point in the same direction.

The components normal to AB are $\mathbf{E}_1\sin\theta$ and $\mathbf{E}_2\sin\theta$ and point in opposite directions. (Fig. 15.12b) Since these component are equal in magnitude but opposite in direction, they cancel each other. Hence, the magnitude of resultant electric field at P is given by

$$E = \mathrm{E}_1\cos\theta + \mathrm{E}_2\cos\theta$$

$$= k\frac{q}{l^2 + r^2}\cos\theta + k\frac{q}{l^2 + r^2}\cos\theta$$

But $\cos\theta = \dfrac{l}{\sqrt{(l^2 + r^2)}}$. Using this expression in the above result, the electric field at P is given by

$$E = \frac{kq}{(l^2 + r^2)} \times \frac{2l}{\sqrt{(l^2 + r^2)}}$$

$$= k\frac{2lq}{(l^2 + r^2)^{3/2}}$$

$$= k\frac{2lq}{r^3(1 + l^2/r^2)^{3/2}}$$

But $p = 2lq$. If $r^2 >> l^2$, the factor l^2/r^2 can be neglected in comparison to unity. Hence

$$E = \frac{p}{4\pi\varepsilon_0 r^3} \tag{15.19}$$

Note that electric field due to a dipole at a point in broad-on position is inversely proportional to the third power of the perpendicular distance between P and the line joining the charges.

If we compare Eqns. (15.18) and (15.19), we note that the electric field in both cases is proportional to $1/r^3$. But there are differences in details:

- The magnitude of electric field in end-on-position is twice the field in the broad-on position.
- The direction of the field in the end-on position is along the direction of dipole moment, whereas in the broad-on position, they are oppositely directed.

15.3.2 Electric Dipole in a Uniform Field

A uniform electric field has constant magnitude and fixed direction. Such a field is produced between the plates of a charged parallel plate capacitor. Pictorially, it is represented by equidistant parallel lines. Let us now examine the behaviour of an electric dipole when it is placed in a uniform electric field (Fig 15.13).

Let us choose x-axis such that the electric field points along it. Suppose that the dipole axis makes an angle θ with the field direction. A force $q\mathbf{E}$ acts on charge $+q$ along the $+x$ direction and an equal force acts on charge $-q$ in the $-x$ direction. Two equal, unlike and parallel forces form a couple and tend to rotate the dipole in clockwise direction. This couple tends to align the dipole in the direction of the external electric field **E**. The magnitude of torque τ is given by

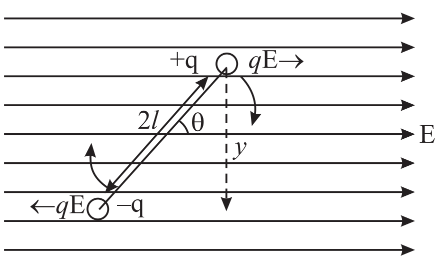


Fig. 15.13 : A dipole in a uniform electric field. The forces on the dipole form a couple and tend to rotate it.

$$\tau = \text{Force} \times \text{arm of the couple}$$

$$= qE \times y$$

$$= qE \times 2l \sin\theta$$

$$= pE \sin\theta$$

In vector form, we can express this result to

$$\boldsymbol{\tau} = \mathbf{p} \times \mathbf{E} \qquad (15.20)$$

We note that

- when $\theta = 0$, the torque is zero, and
- for $\theta = 90°$, the torque on the dipole is maximum, equal to pE. So we may conclude that the electric field tends to rotate the dipole and align it along its own direction.

Example 15.5 : Two charges $+q$ and $-q$, each of magnitude 6.0×10^{-6} C, form a dipole. The separation between the charges is 4×10^{-10} m. Calculate the dipole moment. If this dipole is placed in a uniform electric field $E = 3.0 \times 10^{2}$ NC^{-1} at an angle 30° with the field, calculate the value of torque on the dipole.

Solution : The dipole moment $p = qd$

$$= (6.0 \times 10^{-6} \text{C}) \times (4.0 \times 10^{-10} \text{m})$$

$$= 24 \times 10^{-16} \text{Cm.}$$

Since torque $\tau = p\text{E} \sin \theta$, we can write

$$\tau = (24 \times 10^{-16}\text{cm}) \times 3.0 \times 10^{2} \text{ NC}^{-1}) \sin 30°$$

$$= \frac{72}{2} \times 10^{-14} \text{ Nm}$$

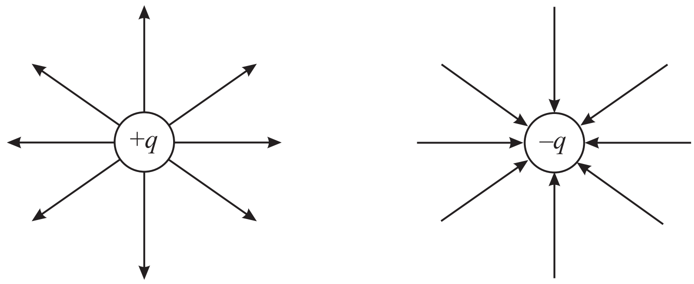
$$= 36 \times 10^{-14} \text{Nm}$$

If a dipole is placed in a non-uniform electric field, the forces on the charges $-q$ and $+q$ will be unequal. Such as electric field will not only tend to rotate but also displace the dipole in the direction of the field.

15.3.3 Electric Lines of Force (Field Lines)

A very convenient method for depicting the electric field (or force) is to draw lines of force pointing in the direction of the field. The sketch of the electric field lines gives us an idea of the magnitude and direction of the electric field. **The number of field lines passing through a unit area of a plane placed perpendicular the direction of the field is proportional to the strength of the field.** A tangent at any point on the field lines gives the direction of the field at that point.

Note that the *electric field lines are only fictitious construction to depict the field. No such lines really exist*. But the behaviour of charges in the field and the interaction between charges can be effectively explained in terms of field lines. Some illustrative examples of electric field lines due to point charges are shown in Fig 15.14. The field lines of a stationary positive charge point radially in outward direction. But for stationary negative charge, the lines start from infinity and terminate at the point charge in radially inward direction (towards the point charge). You must understand that the electric field lines in both cases are in all directions in the space. Only those which are in the plane containing the charge are shown here.

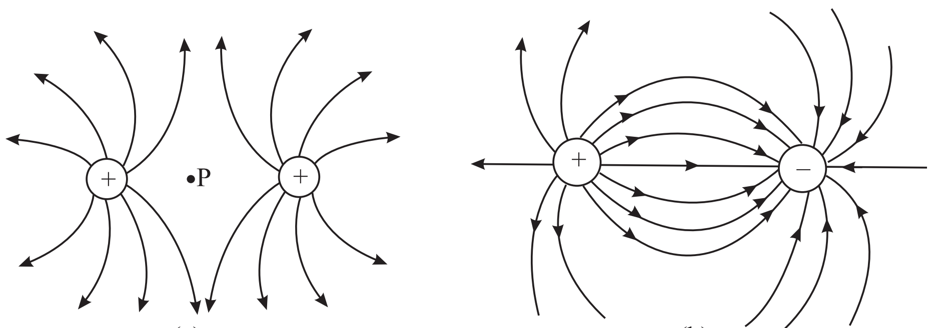


(a) (b)

Fig. 15.14 : Electrical field lines of single point charges : a) The field lines of positive charge, and b) the field lines of negative charge.

Fig 15.15(a) shows a sketch of electric field lines of two equal and similar positive charges placed close to each other. The lines are almost radial at points very close to the positive charges and repel each other, bending outwards. There is a point *P* midway between the charges where no lines are present. The fields of the two charges at this point cancel each other and the resultant field at this point is zero.

Fig. 15.15(b) depicts the field lines due to a dipole. The number of lines leaving the positive charge is equal to the number of lines terminating on the negative charge.



(a) (b)

Fig. 15.15 : Electric field lines due to a system of two point charges : **a)** Two positive charges at rest, and **b)** The field lines due to a dipole start from the positive charge and terminate on the negative charge.

You must remember the following properties of the electric field lines :

- The field lines start from a positive charge radially outward in all directions and terminate at infinity.
- The field lines start from infinity and terminate radially on a negative charge.

- For a dipole, field lines start from the positive charge and terminate on the negative charge.
- A tangent at any point on field line gives the direction of electric field at that point.
- The number of field lines passing through unit area of a surface drawn perpendicular to the field lines is proportional to the field strength on this surface.
- Two field lines never cross each other.

15.4 ELECTRIC FLUX AND GAUSS' LAW

Let us consider a sphere of radius r having charge $+q$ located at its center. The magnitude of electric field at every point on the surface of this sphere is given by

$$E = k \times \frac{q}{r^2}$$

The direction of the electric field is normal to the surface and points outward. Let us consider a small element of area Δs on the spherical surface. **Δs** is a vector whose magnitude is equal to the element of area Δs and its direction is perpendicular to this element (Fig.15.16). The electric flux Δϕ is defined as the scalar product of **Δs** and **E :**

$$\Delta\phi = \mathbf{E} \cdot \mathbf{\Delta s}$$

The total flux over the entire spherical surface is obtained by summing all such contributions:

$$\phi_E = \sum_{\Delta s_i \to 0} \mathbf{E_i} \cdot \mathbf{\Delta s_i} \qquad (15.21)$$

Since the angle between **E** and Δ**s** is zero, the total flux through the spherical surface is given by

$$\phi_E = k \times \frac{q}{r^2} \Sigma \Delta s$$

The sum of all elements of area over the spherical surface is $4\pi r^2$. Hence the net flux through the spherical surface is

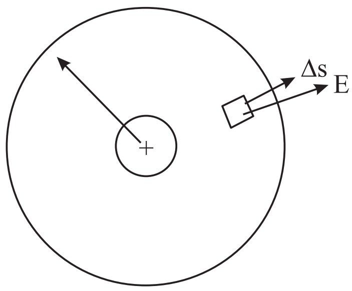


Fig. 15.16

$$\phi_E = k \times \frac{q}{r^2} \times 4\pi r^2$$

$$= 4\pi k \times q$$

On substituting for $k = 1/4\pi\varepsilon_0$, we get

$$\phi_E = \frac{1}{4\pi\varepsilon_0} \times 4\pi q$$

$$= q/\varepsilon_0 \qquad (15.22)$$

The spherical surface of the sphere is referred to as *Gaussian surface*. Eqn. (15.22) is known as Gauss' law. It states that *the net electric flux through a closed gaussian surface is equal to the total charge q inside the surface divided by* ε_0.

Gauss' law is a useful tool for determining the electric field. You must also note that gaussian surface is an imaginary mathematical surface. It may not necessarily coincide with any real surface.

Carl Friedrich Gauss
(1777 – 1855)

German genius in the field of physics and mathematics, Gauss has been one of the most influential mathematicians. He contributed in such diverse fields as optics, electricity and magnetism, astronomy, number theory, differential geometry, and mathematical analysis.

As child prodigy, Gauss corrected an error in his father's accounts when he was only three year old. In primary school, he stunned his teacher by adding the integers 1 to 100 within a second.

Though he shun interactions with scientific community and disliked teaching, many of his students rose to become top class mathematicians – Richard Dedekind, Berhard Riemann, Friedrich Bessel and Sophie Germain are a few among them. Germany issued three postal stamps and a 10 mark bank note in his honour. A crater on moon called Gauss crater, and asteroid 100 called Gaussia have been named after him.

15.4.1 Electric Field due to a Point Charge

Let us apply Gauss' law to calculate electric field due to a point charge. Draw a spherical surface of radius r with a point charge at the centre of the sphere, as shown in Fig. 15.17.

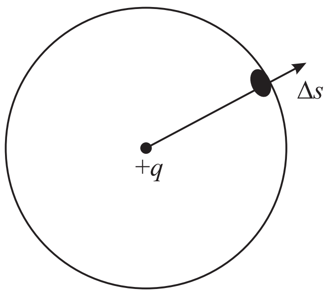


Fig. 15.17 : Electric field on a spherical surface due to a charge $+q$ at its centre

The electric field **E** is along the radial direction pointing away from the centre and normal to the surface of the sphere at every point. The normal to the element of area **Δs** is parallel to **E.** According to Gauss' law, we can write

$$\phi_E = \sum_i \mathbf{E}_i . \Delta \mathbf{s}_i = q/\varepsilon_0$$

Since cos θ = 1 and **E** is same on all points on the surface, we can write

$$\phi_E = E \times 4\pi r^2$$

or

$$q/\varepsilon_0 = E \times 4\pi r^2$$

$$\Rightarrow \qquad E = \frac{q}{4\pi\varepsilon_0 r^2} \qquad (15.23)$$

If there is a second charge q_0 placed at a point on the surface of the sphere, the magnitude of force on this charge would be

$$F = q_0 \times E$$

so that

$$F = \frac{q q_0}{4\pi\varepsilon_0 r^2} \qquad (15.24)$$

Do you recogmise this result? It is expression for Coulomb's force between two static point charges.

15.4.2 Electric Field due to a Long Line Charge

A line charge is in the form of a **thin charged wire of infinite length** with a uniform linear charge density σ_l (charge per unit length). Let there be a charge $+q$ on the wire. We have to calculate the electric field at a point P at a distance r. Draw a right circular cylinder of radius r with the long wire as the axis of the cylinder. The cylinder is closed at both ends. The surface of this cylinder is the gaussian surface and shown in Fig. 15.18. The magnitude of the electric field E is same at every point on the curved surface of the cylinder because all points are at the same distance from the charged wire. The electric field direction and the normal to area element Δs are parallel.

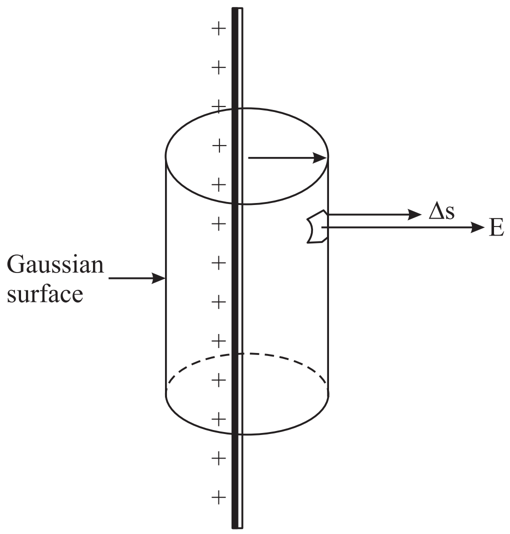


Fig. 15.18 : Electric field due to an infinite line of charges having uniform linear charge density. The gaussian surface is a right circular cylinder.

Let the length of the gaussian cylinder be l. The total charge enclosed in the cylinder is $q = \sigma_l\, l$. The area of the curved surface of the cylinder is $2\,\pi rl$.

For the flat surfaces at the top and bottom of the cylinder, the normals to these areas are perpendicular to the electric field (cos 90° = 0). These surfaces, therefore, do not contribute to the total flux. Hence

$$\phi_E = \Sigma\, \mathbf{E}.\, \Delta\mathbf{s}$$

$$= E \times 2\,\pi rl$$

According to Gauss' law, $\phi_E = q/\varepsilon_0$. Hence

$$E \times 2\,\pi rl = q/\varepsilon_0 = \sigma_l\, l/\varepsilon_0$$

or

$$E = \frac{\sigma_l}{2\pi\varepsilon_0 r} \qquad (15.25)$$

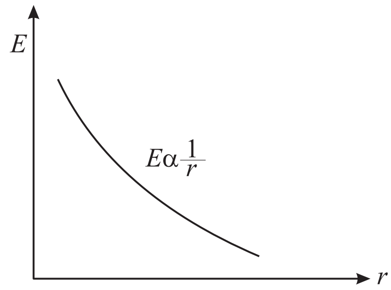
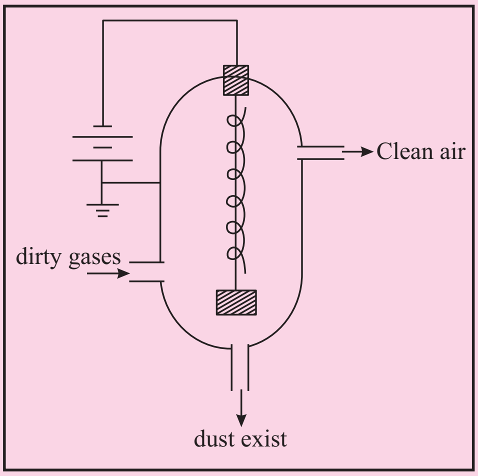


Fig. 15.19 : Variation of E with r for a line charge

This shows that electric field varies inversely with distance. This is illustrated in Fig. 15.19.

Electrostatic Filter

You must have seen black smoke and dirt particles coming out of a chimney of a thermal power station or brick klin. The smoke consists of not only gases but large quantities of small dust (coal) particles. The smoke along with the dirt is discharged into the atmosphere. The dust particles settle down on earth and pollute the soil. The gases contribute to global warming. These are extremely injurious to living systems (health). It is therefore essential that the dirt is removed from smoke before it is discharged into the atmosphere.



A very important application of electrical discharge in gases by application of high electric field is the construction of a device called *Electrostatic Filter or Precipitator*.

The basic diagram of the device is shown here. The central wire inside a metallic container is maintained at a very high negative potential (about 100 kV). The wall of the container is connected to the positive terminal of a high volt battery and is **earthed.** A weight W keeps the wire straight in the central part. The electric field thus created is from the wall towards the wire. The dirt and gases are passed through the container. An electrical discharge takes place because of the high field near the wire. Positive and negative ions and electrons are generated. These negatively charged particles are accelerated towards the wall. They collide with dust particles and charge them. Most of the dust particles become negatively charged because they capture electrons or negative ions. They are attracted towards the wall of the container. The container is periodically shaken so that the particles leave the surface and fall down at the bottom of the container. These are taken out through the exit pipe.

The undesirable dust particles are thus removed from the gases and the clean air goes out in the atmosphere. Most efficient systems of this kind are able to remove about 98% of the ash and dust from the smoke.

15.4.3 Electric Field due to a Uniformly Charged Spherical Shell

A spherical shell, by definition, is a hollow sphere having infinitesimal small thickness. Consider a spherical shell of radius R carrying a total charge Q which is uniformly distributed on its surface. We shall calculate the electric field due to the spherical charge distribution at points external as well as internal to the shell.

(a) Field at an external point

Let P be an external point distant r from the center O of the shell. Draw a spherical surface (called Gaussian surface) passing though P and concentric with the charge distribution. By symmetry, the electric field is radial, being directed outward as shown in Fig. 15.20.

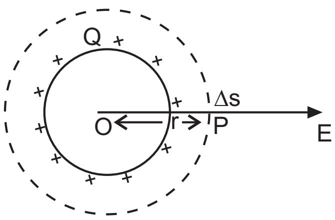


Fig. 15.20

The electric field $\vec{E}$ is normal to the surface element everywhere. Its magnitude at all points on the Gaussian surface has the same value E.

According to Gauss' law,

$$\Sigma E \Delta s \cos 0° = \frac{Q}{\varepsilon_0}$$

or
$$\Delta E \cdot 4\pi r^2 = \frac{Q}{\varepsilon_0}$$

or
$$E = \frac{Q}{4\pi\varepsilon_0 r^2}$$

From the result we can conclude that for a point external to the spherical shell, the entire charge on the shell can be treated as though located at its centre. The electric field decreases with distance.

Instead of a spherical shell if we had taken a charged solid conducting sphere, we would have obtained the same result. This is because the charge of a conductor always resides on its outer surface.

(b) Field at an Internal Point

Let P' be an internal point distant r from the centre of the shell. Draw a concentric sphere passing through the point P'.

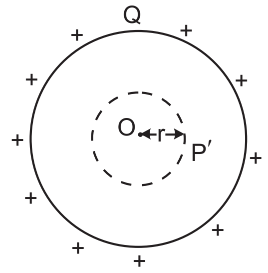


Fig. 15.21

Applying Gauss' Law,

$$\Sigma E \Delta s \cos 0° = \frac{Q}{\varepsilon_0}$$

or
$$E \cdot 4\pi r^2 = \frac{Q}{\varepsilon_0}$$

$\Rightarrow$ $E = 0$ as $Q = 0$

the electric field at an internal point of the shell is zero. The same result is applicable to a charged solid conducting sphere.

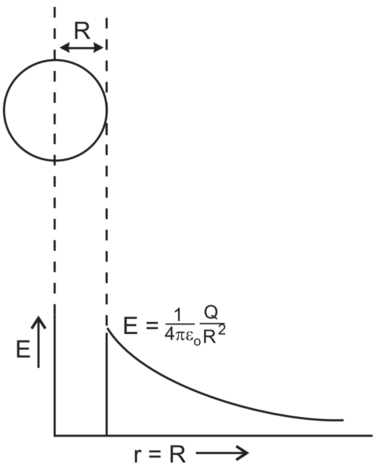


Fig. 15.22

The variation of the electric field with the radial distance r has been shown in Fig 15.22.

15.4.4 Electric Field due to a Plane Sheet of Charge

Consider an infinite plane sheet of charge $ABCD$, charged uniformly with surface charge density σ.

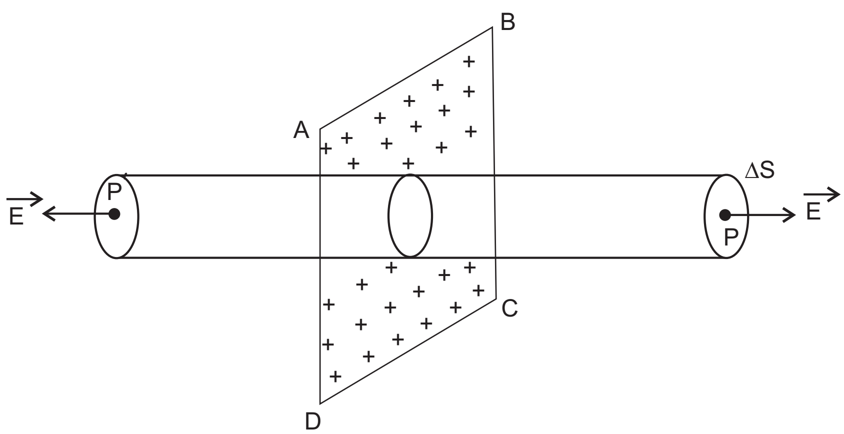


Fig. 15.23

For symmetry reasons, the electric field will be perpendicular to the sheet, directed away from it, if $\sigma > 0$. Let P be the point in front of the sheet where we want to find the electric field. Draw a Gaussian surface in the form of a cylinder with its axis parallel to the field and one of its circular caps passing through P. The other circular cap of the cylinder lies symmetrically opposite at P', on the other side of the sheet, being situated at the same distance as P.

The electric flux through both the circular caps is

$$\vec{E} \cdot \Delta \vec{s} + \vec{E} \cdot \Delta \vec{s} = E\Delta s + E\Delta s$$
$$= 2E\Delta s$$

The electric flux through the curved surface of the Gaussian surface is $\vec{E} \cdot \Delta \vec{s} = E\Delta s \cos 90° = 0.$ Hence, the total electric flux through the Gaussian cylinder is

$$\phi_E = \sum \vec{E} \cdot \Delta \vec{s}$$
$$= 2E\Delta s$$

As the charge enclosed by the Gaussian cylinder is $\sigma\Delta s$, using Gauss' Law we have

$$2E\Delta s = \frac{1}{\varepsilon_0} \sigma \Delta s$$

or $$E = \frac{\sigma}{2\varepsilon_0}$$

Please note that the electric field is independent of the distance from the sheet.

15.5 VAN DE GRAAFF GENERATOR

Van de Graaff Generator is an electrostatic device that can produce potential differences of the order of a few million volts. It was named after its designer Robert J. van de Graaff.

It consists of a large hollow metallic sphere S mounted on an insulating stand. A long narrow belt, made of an insulating material, like rubber or silk, is wound around two pulleys P_1 and P_2 as shown in Fig. 15.5. The pulley P_2 is mounted at the centre of the sphere S while the pulley P_1 is mounted near the bottom. The belt is made to rotate continuously by driving the pulley P_1 by an electric motor M. Two comb-shaped conductors C_1 and C_2, having a number of sharp points in the shape of metallic needles, are mounted near the pulleys.

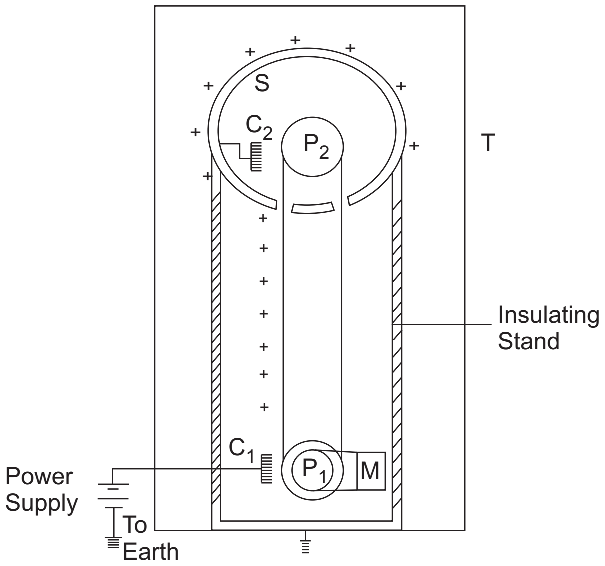


Fig. 15.24

The needles point towards the belt. The comb-shaped conductor C_1 is maintained at a high positive potential (~10^4 V) relative to the ground with the help of a power supply. The upper comb C_2 is connected to the inner surface of the metallic sphere *S*.

Near the sharp points of the comb-shaped conductor C_1, the charge density and electrostatic field are very high. Large electrostatic field near their pointed ends causes dielectric breakdown of the air, producing ions (both positive and negative) in the process. This phenomenon is known as corona discharge. The negative charges from the air move torwards the needles and the positive charge towards the belt. The negative charges neutralize some of the positive charges on the comb C_1. However, by supplying more positive charges to C_1, the power supply maintains its positive potential. As the belt carrying the positive charges moves towards C_2, the air near it becomes conducting due to corona discharge. The negative charges of the air move towards the belt neutralizing its positive charges while the positive charges of the air move towards the needles of the comb C_2. These positive charges are then transferred to the conducting sphere *S* which quickly moves them to its outer surface.

The process continues and positive charges keep on accumulating on the sphere *S* and it acquires a high potential.

As the surrounding air is at ordinary pressure, the leakage of charge from the sphere takes place. In order to prevent the leakage, the machine is surrounded by an earthed metallic chamber *T* whose inner space is filled with air at high pressure.

Notes

By using Van de Graaff generator, voltage upto 5 million volts (MV) have been achieved. Some generators have even gone up to creation of such high voltages as 20 MV.

Van de Graaff generator is used to accelerate the ion beams to very high energies which are used to study nuclear reactions.

INTEXT QUESTIONS 15.4

1. If the electric flux through a gaussian surface is zero, does it necessarily mean that
   (a) the total charge inside the surface is zero?
   (b) the electric field is zero at every point on the surface?
   (c) the electric field lines entering into the surface is equal to the number going out of the surface?
2. If the electric field exceeds the value 3.0×10^6 NC^{-1}, there will be sparking in air. What is the highest value of charge that a metallic sphere can hold without sparking in the surrounding air, if the radius of the sphere is 5.0 cm?
3. What is the magnitude and direction of the net force and net torque on a dipole placed along a a) uniform electric field, and b) non-uniform field.

WHAT YOU HAVE LEARNT

- Electric charge is produced when glass rod is rubbed with silk or rubber is rubbed with fur.
- By convention, the charge on glass rod is taken **positive** and that on rubber is taken **negative.**
- Like charges repel and unlike charges attract each other.
- Coulomb's law gives the magnitude and direction of force between two point charges :

$$\mathbf{F} = k \frac{q_1 \times q_2}{r^2} \hat{\mathbf{r}}$$

where $k = \dfrac{1}{4\pi\varepsilon_0} = 9.0 \times 10^9$ Nm^{2}C^{-2}.

- The smallest unit of charge in nature is the charge on an electron :

$$e = 1.60 \times 10^{-19}\,\text{C}$$

- Charge is conserved and quantised in terms of electronic charge.
- The electric field **E** due to a charge q at a point in space is defined as the force experienced by a unit test charge q_0 :

$$\mathbf{E} = \mathbf{F}/q_0 = k \times \frac{q}{r^2}\,\hat{\mathbf{r}}$$

- Superposition principle can be used to obtain the force experienced by a charge due to a group of charges. It is also applicable to electric field at a point due to a group of charges.
- Electric dipole is a system of two equal and unlike charges separated by a small distance. It has a dipole moment $|\mathbf{p}| = qr$; the direction of **p** is from negative charge to positive charge along the line joining the two charges.
- The electric field due to a dipole in end-on position and broad-on position is respectively given by

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\frac{2\mathbf{p}}{r^3}$$

and

$$\mathbf{E} = -\frac{1}{4\pi\varepsilon_0}\frac{\mathbf{p}}{r^3}.$$

- Electric field lines (line of force) are only a pictorial way of depicting field.
- Electric flux is the total number of electric lines of force passing through an area and is defined as $\phi_E = \mathbf{E} \cdot \mathbf{A}$.
- Gauss's law states that the total flux passing through a closed area is $\frac{1}{\varepsilon_0}$ times the total charge enclosed by it.
- The electric field due to a line charge is given by $E = \frac{\sigma_l}{2\pi\varepsilon_0 r}$.

TERMINAL EXERCISE

1. A + 12μ C charge is at x =20 cm and a –18 μ C (– q) charge is at x = 29 cm on the x-axis. Calculate the magnitude and direction of the force on a charge of 18μC. What is the direction of force on 12μC charge?
2. Two point Charges q_1 and q_2 separated by a distance of 3.0 m experience a mutual force of 16×10^{-15} N. Calculate the magnitude of force when $q_1 = q_2 = q$. What will be the magnitude of force if separation distance is changed to 6.0 m?

3. There are two points A and B separated by a distance x. If two point charges $+q$ each are on the points A and B, the force between them is F. The point charges are now replaced by two identical metallic spheres having the same charge $+q$ on each. The distance between their centers is again x only. Will the force between them change? Give reasons to support your answer.

4. The force of repulsion between two point chargers placed 16 cm apart in vaccum is 7.5×10^{-10} N. What will be force between them, if they are placed in a medium of dielectric constant $k = 2.5$?

5. Compare the electrical force with the gravitational force between two protons separated by a distance x. Take charge on proton as 1.60×10^{-19} C, mass of proton as 1.67×10^{-27} kg and Gravitational constant $G = 6.67 \times 10^{-11}$ $Nm^2 kg^{-2}$.

6. Four identical point charges $+q$ each are placed at the four corners (one q at one corner) of a square of side 1. Find the force experienced by a test charge q_0 placed at the center of the square.

7. When are the electric field lines parallel to each other?

8. How many electrons should be removed from a metallic sphere to give it a positive charge $= 6.4 \times 10^{-7}$C.

9. Consider an electric dipole of $q = 3.0 \times 10^{-6}$ C and $2l = 4 \times 10^{-10}$ m. Calculate the magnitude of dipole moment. Calculate electric field at a point $r = 6 \times 10^{-6}$ m on the equatorial plane.

10. A Charge $-q = 15 \times 10^{-6}$ C is placed on a metallic sphere of radius R=3.0 mm. Calculate the magnitude and direction of the electric field at a point r =15 cm from the center of the sphere. What will be the magnitude and direction of the field at the same point if 3.0 mm sphere is replaced by 9.0 mm sphere having the same Charge.

11. A charge of +15μC is located at the center of a sphere of radius 20 cm. Calculate the electric flux through the surface of the sphere.

12. A proton is placed in a uniform electric field $E = 8.0 \times 10^4$ NC^{-1}. Calculate the acceleration of the proton.

13. Two point charges q_1 and q_2 are 3.0 cm apart and $(q_1 + q_2) = 20$μC. If the force of repulsion between them is 750N, calculate q_1 and q_2.

ANSWERS TO INTEXT QUESTIONS

15.1

1. (i) Yes (ii) Charge $= 3.2 \times 10^{-17}$ C.

2. A has charge $+Q$. When A and B are brought in contact, charge will get distributed equally.
   (i) Yes., (ii) $+Q/2$
3. $q = 4.8 \times 10^{-16}$

   Since $Ne = q$, we get

   $$N = \frac{4.8 \times 10^{-16}}{1.6 \times 10^{-19}} = 3.0 \times 10^{3} \text{charges}$$

15.2

1. $Q_1 = 16\mu C$, $Q_2 = q\mu C$ and $r = 12m$

Since

$$F = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r^2}$$

$$= \frac{(9\times10^9 \text{ Nm}^2\text{C}^{-2})(16\times10^{-6}\text{ C})\times(12\times^{-6}\text{ C})}{144\text{m}^2}$$

$= 9\times 10^{-3}$ N

(i) direction from q_2 to q_1

(ii) direction from q_1 to q_2

2. The force at A due to charge at B, $F_1 = k\frac{q^2}{a^2}$ where $AB = a$

   Since $AB = AC$, the force at A due to charge at B is

   $$F_2 = k\frac{q^2}{a^2}$$

   $R^2 = F_1^2 + F_2^2 = 2F^2$

   $R = F\sqrt{2}$ at 45°

15.3

1. (a) E along the $+x$ axis.
   (b) along the $+y$ axis.
   (c) at 45° with the x axis
2. $AB = AC = 40$ cm

   $$|\mathbf{E}_1| = \frac{kq}{r^2} = |\mathbf{E}_2| = \frac{9\times10^9 \text{ Nm}^2\text{C}^{-2}\times(2\times10^{-6}\text{ C})}{(0.40\text{m})^2} = 1.125\times10^5 \text{ NC}^{-1}$$

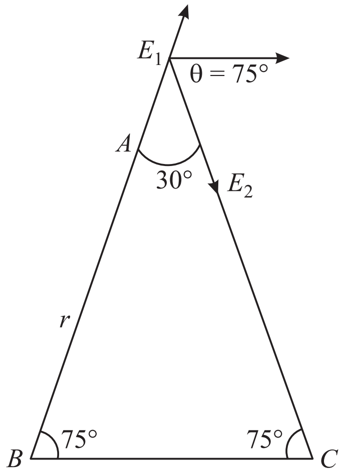
The resultant of $\mathbf{E}_1$ and $\mathbf{E}_2$ will be parallel to BC. Hence

$$R^2 = E_1^2 + E_2^2 + 2E_1 E_2 \cos 150$$

$$= 2E^2 + 2E^2 \cos(180\text{-}30)$$

$$= 2E^2 - 2E^2 \times \cos 30 = 2E^2\left(1+\frac{\sqrt{3}}{2}\right) = 4.723 \times 10^{10}\ \text{N}^2\text{C}^{-2}.$$

Direction will be parallel to BC in the direction $B \rightarrow C$.



3. $\mathbf{E}$ is directed towards the earth. The force on –ve charge will be vertically upwards.
4. The field will be zero at the mid point between the charge.

15.4

1. (i) Yes (ii) not necessarily (iii) Yes.

2. $E = \dfrac{Q}{4\pi\varepsilon_0 r^2}$

$\therefore\ Q = 4\pi\varepsilon_0 r^2 E$

$$= (3 \times 10^6\ \text{NC}^{-1}) \times \frac{1}{(9\times10^9\ \text{Nm}^2\text{C}^{-2})} \times (25 \times 10^{-4}\text{m}^2)$$

$$= 8.3 \times 10^{-7}\ \text{C}$$

3. (a) $\mathbf{F} = 0, \boldsymbol{\tau} = 0$

   (b) $\mathbf{F} \neq 0\ \boldsymbol{\tau} = 0$

Answers to Problems in Terminal Exercise

1. 240 N towards negative x–direction force on +12 μC charge is towards + x direction.
2. $q = 4 \times 10^{-3}$C   4. 3×10^{-10}N
5. Electric force is 10^{36} times the gravitational force.
6. zero.   8. 4×10^{12} electrons
9. 12×10^{-16} Cm. 0.5×10^{15} or Nc^{-1}
10. 6×10^{6} NC^{-1} towards the centre, same field.
11. 1.7×10^{6} μm   12. 7.6×10^{12} ms^{-2}
13. 15 μC and 5 μC.

312en16

16

ELECTRIC POTENTIAL AND CAPACITORS

In modules 2 and 3, you learnt about the direction of flow of fluids and thermal energy. You may recall that the level of water in a container determines the direction in which it flows. If the level of water in one container is higher than that in the other, water will flow from higher level to lower level, irrespective of the quantity of water in the containers. Temperature plays a similar role in case of flow of thermal energy from one object to another. Thermal energy always flows from a body at higher temperature to the one at lower temperature. Here also, the direction of flow does not depend on the quantity of thermal energy possessed by an object.

Electric potential plays a similar role in the flow of charges from one point to another. The positive charge always moves from a point at higher potential to a point at lower potential. A positive test charge, when left free in an electric field, moves in the direction of the electric field. From this behaviour of a positive test charge, you may be tempted to say that the **electric field (E)** and **electric potential** (V) are closely related. In this lesson, you will learn to establish a relation between these physical quantities. You will also learn about a device called capacitor, which is used to store charge, filter alternating current and finds wide applications in electronic circuitory as well as power transmission.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain the meaning of electric potential at a point and potential difference;*
- *derive expressions for electric potential due to a point charge and a dipole;*
- *explain the principle of capacitors and state their applications;*
- *derive an expression for the capacitance of a parallel plate capacitor;*
- *obtain equivalent capacitance in grouping of capacitors;*

Notes

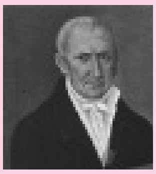
- *calculate the energy stored in a capacitor; and*
- *explain polarization of dielectric materials in an electric field.*

16.1 LECTRIC POTENTIAL AND POTENTIAL DIFFERENCE

When a charged particle is made to move in an electrostatic field in a direction opposite to the direction of the field, work is done by an external agency. This work is stored as potential energy of charge in accordance with the law of conservation of energy. So, we can say that an electric charge placed at a point in an electric field has potential energy, which is a function of its position. We can visualize the potential energy of charge in the field as a scalar function of position and for a unit charge call it potential. It means that different points in an electric field would be at different potentials. And if a positively charged particle is placed in an electric field, it will tend to move from higher to lower potential to minimize its potential energy. In the next lesson, you will learn how the concept of potential difference leads to flow of current in electric circuits.

*The **electric potential** at any point in an electric field is equal to the work done against the electric force in moving a unit positive charge from outside the electric field to that point.* Electric potential is a scalar quantity, as it is related to work done.

Alessandro, Conte Volta (1745-1827)



Born at Como, Italy, Volta was a professor at Pavia for more than 20 years. A well travelled man, he was known to many famous men of his times. He decisively proved that animal electricity observed by Luigi Galvani in frog muscles was a general phenomenon taking place between two dissimilar metals seperated by acidic or salt solutions. On the basis of this observation, he invented first electro-chemical cell, called voltaic cell. The unit of potential difference is named volt in his honour.

The potential at a point is taken positive when work is done against the field by a positive charge but negative when work is done by the electric field in moving the unit positive charge from infinity to the point in the field.

Consider two points *A* and *B* in an electric field (Fig. 16.1). If a test charge q_0 is moved from point *A* to point *B* along any path by an external force, the amount of work done by the external force is given by

$$W_{AB} = q_0 (V_B - V_A) \qquad (16.1)$$

Thus, potential difference between points *A* and *B* will be

$$V_{AB} = V_B - V_A = \frac{W_{AB}}{q_0} \qquad (16.2)$$

Here V_A and V_B are potentials at points *A* and *B*, respectively.

A potential difference is said to exist between two points in an electric field, if work is done against the electric force in moving a positive test charge from one point to the other. Note that this work is independent of the path. (*For this reason, the electric field is said to be a conservative field*). The SI unit of potential and potential difference is **volt** :

1 **volt** = 1 **joule**/1 **coulomb**

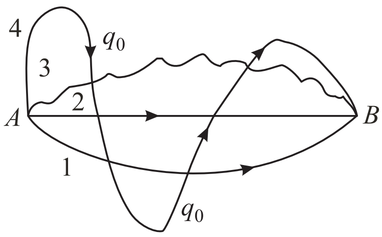


Fig. 16.1 : The work done in moving a test charge from one point to another in an electric field is independent of the path followed.

If one joule of work is done in taking a test charge of one coulomb from one point to the other in an electric field, the potential difference between these points is said to be one volt. If one joule of work is done in bringing a test charge of one coulomb from infinity to a point in the field, the potential at that point is one volt.

Note that potential at a point is not a unique quantity as its value depends on our choice of zero potential energy (infinity). However, the potential difference between two points in a stationary field will have a unique value. Let us now learn to calculate potential at a point due to a single charge.

16.1.1 Potential at a point due to a Point Charge

Suppose we have to calculate electric potential at point P due to a single point charge $+q$ situated at O (Fig. 16.2), where $OP = r$. The magnitude of electric field at P due to the point charge is given by

$$E_p = \frac{1}{4\pi\varepsilon_0} \times \frac{q}{r^2} \qquad (16.3)$$

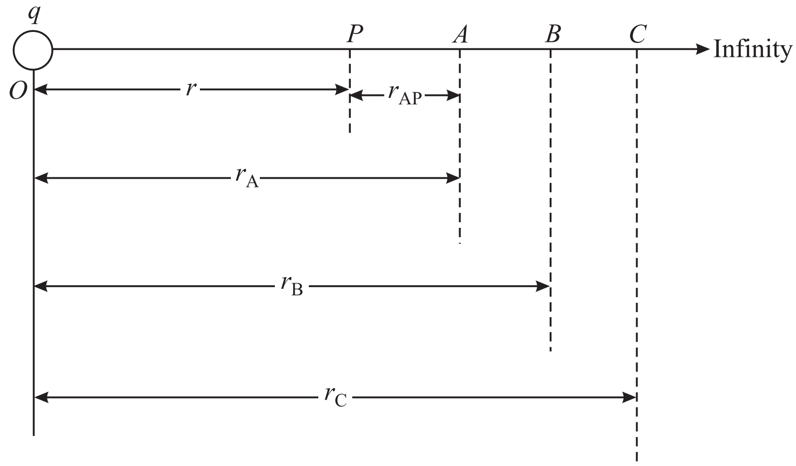


Fig. 16.2 : Work done per unit charge in moving a charge q_0 from infinity to a point P in an electric field E is the potential at that point.

Similarly, the electric field at point *A* will be

$$E_A = \frac{1}{4\pi\varepsilon_0}\frac{q}{r_A^2} \tag{16.4}$$

If points *P* and *A* are very close, the average field E_{AP} between these points can be taken as the geometric mean of E_p and E_A :

$$E_{AP} = \sqrt{E_A \times E_p}$$

$$= \sqrt{\frac{1}{4\pi\varepsilon_o}\frac{q}{r_A^2} \times \frac{1}{4\pi\varepsilon_o}\frac{q}{r^2}}$$

$$= \frac{1}{4\pi\varepsilon_0}\frac{q}{r_A r} \tag{16.5}$$

Therefore, the magnitude of force experienced by a test charge q_0 over this region will be

$$F_{AP} = q_o E_{AP} = \frac{1}{4\pi\varepsilon_0}\frac{q\,q_0}{r_A r} \tag{16.6}$$

and the work done in moving charge q_0 from *A* to *P* is given by

$$W_{AP} = F_{AP} \times r_{AP}$$

$$= \frac{1}{4\pi\varepsilon_0}\frac{q\,q_o}{r_A r} \times (r_A - r)$$

$$= \frac{q\,q_0}{4\pi\varepsilon_0} \times \left(\frac{1}{r} - \frac{1}{r_A}\right) \tag{16.7}$$

where r_{AP} is the distance between points *A* and *P*.

Similarly, work done in moving this charge from *B* to *A* will be given by

$$W_{BA} = \frac{q\,q_0}{4\pi\varepsilon_0} \times \left(\frac{1}{r_A} - \frac{1}{r_B}\right) \tag{16.8a}$$

And work done in moving the test charge from *C* to *B* will be

$$W_{CB} = \frac{q\,q_0}{4\pi\varepsilon_0} \times \left(\frac{1}{r_B} - \frac{1}{r_C}\right) \tag{16.8b}$$

and so on. The total work done in moving the charge from infinity to the point *P* will be

$$W = \frac{q\,q_0}{4\pi\varepsilon_0} \times \left(\frac{1}{r} - \frac{1}{r_A} + \frac{1}{r_A} - \frac{1}{r_B} + \frac{1}{r_B} - \frac{1}{r_C} ... + ... - \frac{1}{\infty}\right)$$

$$= \frac{q\,q_0}{4\pi\varepsilon_0} \times \left(\frac{1}{r} - \frac{1}{\infty}\right)$$

$$= \frac{q\,q_0}{4\pi\varepsilon_0 r} \qquad (16.9)$$

By definition, potential at a point is given by

$$V_P = \frac{W}{q_0}$$

$$= \frac{1}{4\pi\varepsilon_0} \frac{q}{r} \qquad (16.10)$$

Note that potential is inversely proportional to distance. It is positive or negative depending on whether q is positive or negative.

If there are several charges of magnitudes q_1, q_2, q_3,…, the electric potential at a point is the scalar sum of the potentials due to individual charges (Fig.16.3) :

$$V = V_1 + V_2 + V_3 + ...$$

$$= \sum_{i=1}^{\infty} \frac{q_i}{4\pi\varepsilon_0 r_i} \qquad (16.11)$$

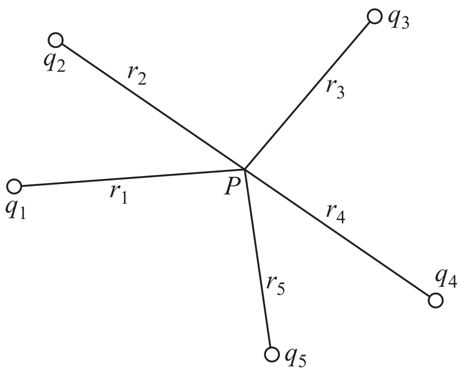


Fig. 16.3 : Potential at a point P due to a system of charges

16.1.2 Potential at a Point due to an Electric Dipole

Let us consider an electric dipole consisting of two equal and opposite point charges $-q$ at A and $+q$ at B, separated by a distance $2l$ with centre at O. We wish to calculate potential at a point P, whose polar co-ordinates are (r,θ); i.e. $OP = r$ and $\angle BOP = \theta$, as shown in Fig. 16.4. Here $AP = r_1$ and $BP = r_2$. We can easily calculate potential as P due to point charges at A and B using Eqn.(16.10) :

$$V_1 = \frac{1}{4\pi\varepsilon_0} \times \frac{(-q)}{r_1}$$

and

$$V_2 = \frac{1}{4\pi\varepsilon_0} \times \frac{q}{r_2}$$

Total potential at *P* due to both the charges of the dipole is given by

$$V = V_1 + V_2$$

That is,

$$V = \frac{q}{4\pi\varepsilon_0}\left[\frac{1}{r_2} - \frac{1}{r_1}\right] \quad (16.12)$$

To put this result in a more convenient form, we draw normals from *A* and *B* on the line joining *O* and *P*. From Δ *BOD*, we note that *OD* = *l* cos θ and from Δ *OAC* we can write *OC* = *l* cos θ. For a small dipole (*AB*<<*OP*), from Fig. 16.4, we can take *PB* = *PD* and *PA* = *PC* . Hence

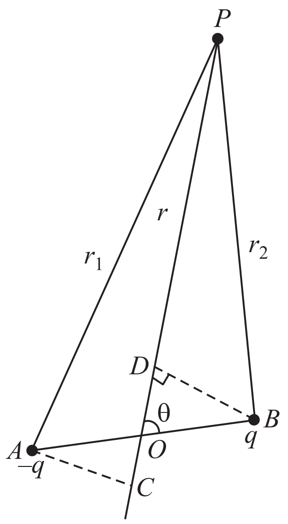


Fig. 16.4 : Electric potential at a point *P* due to an electric dipole.

$$r_1 = r + l\cos\theta$$

$$r_2 = r - l\cos\theta$$

Using these results in Eqn (16.12), we get

$$V = \frac{q}{4\pi\varepsilon_0}\left[\frac{1}{(r - \ell\cos\theta)} - \frac{1}{(r + \ell\cos\theta)}\right]$$

$$= \frac{q}{4\pi\varepsilon_0}\left[\frac{(2\ell\cos\theta)}{(r^2 - \ell^2\cos^2\theta)}\right]$$

$$= \frac{q \times 2\ell\cos\theta}{4\pi\varepsilon_0 r^2}$$

where we have neglected the term containing second power of *l* since *l*<<r.

In terms of dipole moment ($p = q \times 2\ell$), we can express this result as

$$V = \frac{p\cos\theta}{4\pi\varepsilon_0 r^2} \quad (16.13)$$

This result shows that unlike the potential due to a point charge, the potential due to a dipole is inversely proportional to the square of the distance.

Let us now consider its special cases.

Special Cases

Case I : When point *P* lies on the axial line of the dipole on the side of positive charge, θ = 0 and cos θ =1. Then Eqn. (16.13) reduces to

$$V_{\text{AXIS}} = \frac{p}{4\pi\varepsilon_0 r^2} \quad (16.14)$$

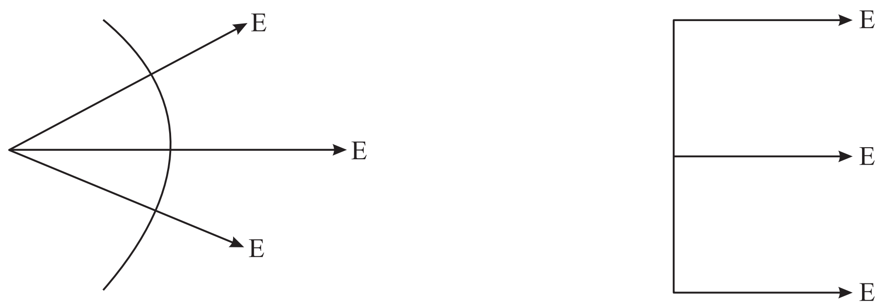
Case II : When point *P* lies on the axial line of the dipole but on the side of negative charge, θ = 180º and cos θ = –1. Hence

$$V_{\text{AXIS}} = -\frac{p}{4\pi\varepsilon_0 r^2} \qquad (16.15)$$

Case III : When point *P* lies on the equatorial line of the dipole (perpendicular bisector of *AB*), θ = 90º and cos θ = 0. Then

$$V_{\text{equatorial}} = 0 \qquad (16.16)$$

That is, electric potential due to a dipole is zero at every point on the equatorial line of the dipole. When a dipole is kept in 3D space, the equatorial line will lie in the plane of the paper. The potential at all points in this plane will be same, i.e. zero. Such a surface is referred to as *equipotential surface*. The electric field is always perpendicular to an equipotential surface. No work is done in moving a charge from one point to another on the equipotential surface.



(a) Spherical equipotential surface

(b) Plane equipotential surface

Fig. 16.5 : Equipotential surfaces and electric field directions

16.1.3 Potential Energy of a System of Point Charges

The electric potential energy is the energy possessed by a system of point charges by virtue of their being in an electric field. When charges are infinite distance apart, they do not interact and their potential energy is zero. If we want to assemble a charge system, i.e. bring charges near each other, work will have to be done. This work is stored in the form of potential energy in the system of these charges. This is called the electric potential energy of the charge system. Hence, we can define *potential energy of a system of point charges as the total amount of work done in bringing various point charges of the system to their respective positions from infinitely large mutual separations.*

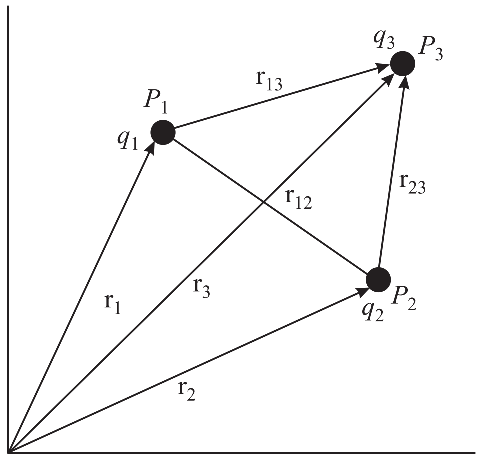


Fig. 16.6 : Potential energy of a system of point charges separated by a distance

Suppose that a point charge q_1 is located at a point P_1 with position vector $\mathbf{r}_1$ in space. Assume that point charge q_2 is at infinity. This is to be brought to the point P_2 having position vector $\mathbf{r}_2$ where $P_1P_2 = \mathbf{r}_{12}$, as shown in Fig. 16.6. We know that electric potential at P_2 due to charge q_1 at P_1 is

$$V_2 = \frac{1}{4\pi\varepsilon_0} \frac{q_1}{|\mathbf{r}_{12}|} \qquad (16.17)$$

From the definition of potential, work done in bringing charge q_2 from infinity to point P_2 is

$$W = (\text{Potential at } P_2) \times \text{value of charge}$$

This work is stored in the system of charges q_1 and q_2 in the form of electric potential energy U. Thus,

$$U = \frac{q_1 \times q_2}{4\pi\varepsilon_0 |r_{12}|} \qquad (16.18)$$

In case the two charges have same sign, work is done against the repulsive force to bring them closer and hence, electric potential energy of the system increases. On the other hand, in separating them from one another, work is done by the field. As a result, potential energy of the system decreases. If charges are of opposite sign, i.e. one is positive and the other is negative, the potential energy of the charge system decreases in bringing the charges closer and increases in separating them from one another.

For a three point charge system (Fig. 16.6), Eqn. (16.18) can be written as

$$U = \frac{1}{4\pi\varepsilon_0}\left(\frac{q_1q_2}{r_{12}} + \frac{q_1q_3}{r_{13}} + \frac{q_2q_3}{r_{23}}\right) \qquad (16.19)$$

Proceeding in the same way, we can calculate the potential energy of a system of any number of charges.

By combining Eqns. (16.3) and (16.13), the potential energy of a dipole in a uniform electric field can be written as

$$U_\theta = -pE\cos\theta = -\mathbf{p.E} \qquad (16.20)$$

where **p** is the dipole moment in electric field **E** and θ is the angle between **p** and **E**.

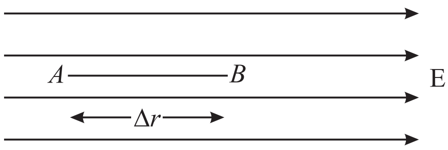
16.2 RELATION BETWEEN ELECTRIC FIELD AND POTENTIAL

Consider two points A and B in a uniform electric field **E**, separated by a small distance Δr. By definition, potential difference ΔV between A and B is equal to the work done in moving a unit positive test charge from A to B:

ΔV = (Force on unit positive charge) × (AB)

$= \mathbf{E}.\ \Delta\mathbf{r} = E(\Delta r) \cos 180^{O}$

$= -\ \mathbf{E}\ \Delta\mathbf{r}$



or
$$\mathbf{E} = -\frac{\Delta V}{\Delta r} \tag{16.21}$$

The negative sign indicates that work is done against the electric field.

Hence, at any point, the electric field is equal to negative rate of change of potential with distance (called ***potential gradient***) at that point in the direction of field. Remember that electric potential is a scalar quantity but electric potential gradient is a vector as it is numerically equal to electric field.

From the above relation, for a uniform electric field, we can write

$$E = \frac{V_A - V_B}{d} \tag{16.22}$$

Here V_A and V_B are potentials at points A and B, respectively separated by a distance d.

Example 16.1 : In a 10 volt battery, how much work is done when a positively charged particle having charge 1.6×10^{-19} C is moved from its negative terminal to the positive terminal?

Solution : According to Eqn. (16.2)

$$V_{AB} = W_{AB} / q_O$$

Since V_{AB}= 10 V and $q_0 = 1.6 \times 10^{-19}$ C, we get

$$W_{AB} = (10\text{V}) \times (1.6 \times 10^{-19}\text{C})$$

$$= 1.6 \times 10^{-18}\ \text{J}$$

Example 16.2 : A point charge q is at the origin of Cartesian co-ordinate system. The electric potential is 400 V and the magnitude of electric field is 150 N C^{-1} at a point x. Calculate x and q.

Solution : The electric field

$$E = \frac{V}{x}$$

On inserting the numerical values, we get

$$150 = \frac{400}{x}$$

or
$$x = 2.67 \text{ m}$$

Recall that electric field is given by the expression

$$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{x^2}$$

We subsitute $\frac{1}{4\pi\varepsilon_0} = 9 \times 10^9 \text{ N C}^{-2} \text{ m}^2$, $E = 150 \text{ N C}^{-1}$ and $x = 2.67$ m and obtain

$$q = \frac{(150 \text{ N C}^{-1}) \times (2.67\text{m})^2}{9 \times 10^9 \text{ N C}^{-2}}$$

$$= 11.9 \times 10^{-8} \text{ C}$$

INTEXT QUESTIONS 16.1

1. A metallic sphere of radius R has a charge $+q$ uniformly distributed on its surface. What is the potential at a point r ($> R$) from the centre of the sphere?
2. Calculate the work done when a point charge is moved in a circle of radius r around a point charge q.
3. The electric potential V is constant in a region. What can you say about the electric field **E** in this region ?
4. If electric field is zero at a point, will the electric potential be necessarily zero at that point.
5. Can two equipotential surfaces intersect?

On the basis of charge conduction, substances are broadly classified as **conductors** and **insulators**. In solids, conduction of electricity usually takes place due to free electrons, whereas in fluids, it is due to ions. Conductors have free charge carriers through which electric currents can be established on applying an electric field. Metals are good conductors. Substances having no free charge carriers are called ***insulators***. The common insulators are wood, ebonite, glass, quartz, mica etc. Substances which have electrical conductivity in between those of conductors and insulators are called ***semiconductors***. The ratio of electrical conductivities of good conductors and good insulators is of the order of 10^{20}. Let us now learn how conductors behave in an electric field.

16.2.1 Behaviour of Conductors in an Electric Field

Conductors have electrons which are not bound tightly in their atoms. These are free to move within the conductor. However, there is no net transfer of electrons (charges) from one part of the conductor to the other in the absence of any applied electric field. The conductor is said to be in electrostatic equilibrium.

Refer to Fig. 16.7(a) which shows a conductor placed in an external electric field **E**. The free electrons are accelerated in a direction opposite to that of the electric field. This results in build up of electrons on the surface *ABCD* of the conductor. The surface *FGHK* becomes positively charged because of removal of electrons. These charges (-ve on surface *ABCD* and +ve on surface *FGHK*) create their own fields, which are in a direction opposite to **E**. The transfer of electrons from *FGHK* to *ABCD* continues till **E** becomes equal to $\mathbf{E_1}$. Such a state of electrostatic equilibrium is reached usually in 10^{-16} s. We then say that equilibrium is reached almost instantaneously. If there is a cavity inside a conductor, the electric field inside the cavity is zero (Fig. 16.7(b)).

These results are true for a charged conductor or when charges are induced on a neutral conductor by an external electric field.

This property of a conductor is used in **Electrostatic Shielding** — a phenomenon of protecting a certain region of space from external electric fields. To protect delicate instruments from external electric fields, they are enclosed in hollow conductors. That is why in a thunder storm accompanied by lightning, it is safer to be inside a car or a bus than outside. The metallic body of the car or bus provides electrostatic shielding from lightning.

Conductors in electrostatic equilibrium exhibit the following properties :

- There is no electric field inside a conductor.
- The electric field outside a charged conductor is perpendicular to the surface of the conductor, irrespective of the shape of the conductor.
- Any charge on the conductor resides on the surface of the conductor.

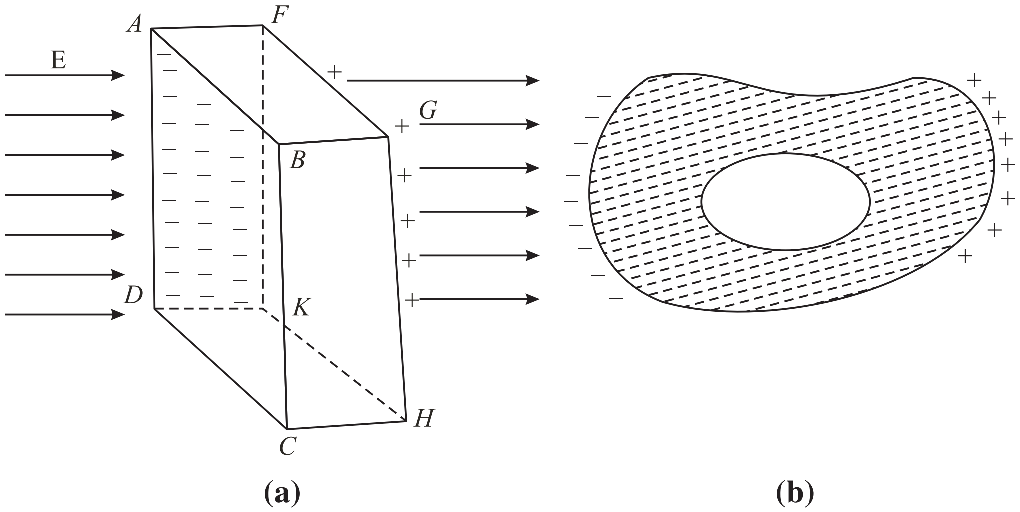


Fig.16.7 : Electrostatic shielding: **(a)** External electric field E pulls free electrons on the surface *ABCD*. The surface *FGHK,* which is deficient in electrons, becomes positively charged;the net field inside the conductor is zero. **(b)** If there is a cavity inside a conductor, the field inside the cavity is zero.

16.3 CAPACITANCE

Let us consider two conductors having equal but opposite charges $+Q$ and $-Q$ on them. There is a potential difference V between them. Such a system of conductors is called a ***capacitor***. Experimentally it is found that the potential difference is directly proportional to charge on a conductor. As charge increases, the potential difference between them also increases but their ratio remains constant. This ratio is termed as capacitance of the capacitor:

$$C = Q / V \qquad (16.23)$$

The capacitance is defined as the ratio between the charge on either of the conductors and the potential difference between them. It is a measure of the capability of a capacitor to store charge.

In SI system of units, capacitance is measured in farad (F). The capacitance is one farad, if a charge of one coulomb creates a potential difference of one volt :

$$1 \text{ farad} = \frac{1 \text{ coulomb}}{1 \text{ volt}} \qquad (16.24)$$

You may recall from the previous unit that coulomb is a very large unit of charge. It means that farad is also a very large unit of capacitance. Usually we use capacitors of values in microfarad or picofarad:

$$1 \text{ microfarad} = 10^{-6} \text{ farad, written as } \mu\text{F}$$

$$1 \text{ picofarad} = 10^{-12} \text{ farad, written as pF}$$

In an electrical circuit, a capacitor is represented by two parallel lines.

16.3.1 Capacitance of a Spherical Conductor

Suppose that a sphere of radius r is given charge q. Let the potential of the sphere be V. Then

$$V = \frac{1}{4\pi\varepsilon_0} \frac{q}{r}$$

Since $C = q/V$, we find that

$$C = \frac{q}{q / 4\pi\varepsilon_0 r} = 4\pi\varepsilon_0 r = \frac{r}{9\times10^9} \qquad (16.25)$$

This shows that capacitance of a spherical conductor is directly proportional to its radius. In fact, it is numerically equal to its radius divided by 9×10^9, where radius is taken in metre. For example, the capacitance of a sphere of radius 0.18 m is

$$C = \frac{0.18}{9} \times 10^{-9} \text{ F} = 20 \text{ pF}$$

16.3.2 Types of Capacitors

You will come across many types of capacitors in your physics laboratory. The power supply system of your city also uses capacitors. These also form important components of devices such as radio, T.V., amplifires and oscillators. A capacitor essentially consists of two conductors, one charged and the other usually earthed. To understand the principle of a capacitor, let us consider an insulated metal plate A and give it positive charge (q) till its potential (V) becomes maximum. (Any

Notes

further charge given to it would leak out.) The capacitance of this plate is equal to q/V.

Now bring another insulated metal plate B near plate A. By induction, negative charge is produced on the nearer face of B and equal positive charge develops on its farther face (Fig. 16.8a). The induced negative charge tends to decrease whereas induced positive charge tends to increase the potential of A. If plate B is earthed (Fig.16.8b), the induced positive charge on it, being free, flows to earth. (In reality, it is the negative charge that flows from the earth to the plate. Positive charges in the plate are immobile.) But negative charge will stay as it is bound to positive charge on A. Due to this induced negative charge on B, the potential of A decreases and its capacitance increases.

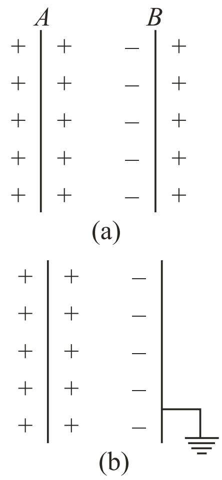


Fig.16.8 : Working principle of a capacitor

Hence, we can say that *capacitance of an insulated conductor can be increased by bringing near it an uncharged earthed conductor.* This is the basic principle of a capacitor. Capacitors are used for storing large amounts of electric charge and hence electrical energy in a small space for a small interval of time.

A Parallel Plate Capacitor

A parallel plate capacitor is one of the simplest capacitors in which two parallel metallic plates, each of area A, are separated from one another by a small distance d. An insulating medium like air, paper, mica, glass etc separates the plates. The plates are connected to the terminals of a battery, as shown in Fig. 16.9. Suppose that these plates acquire $+q$ and $-q$ charge when the capacitor is fully charged. These charges set up a uniform electric field **E** between the plates. When the separation d is small compared to the size of the plates, distortion of electric field at the boundaries of the plates can be neglected.

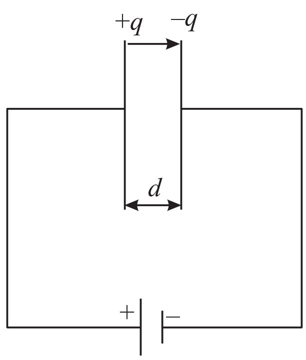


Fig. 16.9 : Working principle of a capacitor

If σ is surface charge density on either plate, the magnitude of electric field between the plates is given by

$$E = \frac{\sigma}{\varepsilon_0} = \frac{q}{\varepsilon_0 A}$$

and the potential difference between the plates is given by

$$V = Ed$$

Hence, capacitance of a parallel plate capacitor, whose plates are separated by *d* and have air in-between them is given by

$$C_0 = \frac{q}{V} = \frac{q}{qd/\varepsilon_0 A}$$

$$= \frac{\varepsilon_0 A}{d} \qquad (16.26)$$

It shows that capacitance of a parallel plate capacitor is directly proportional to the area of the plates and inversely proportional to their separation. It means that to obtain high capacitance, area of the plates should be large and separation between them should be small.

If the plates of a capacitor are separated by a dielectric material other than air or vacuum, the capacitance of a parallel plate capacitor is given by

$$C = \frac{\varepsilon A}{d} = \frac{k\,\varepsilon_0 A}{d}$$

where ε is called permittivity of the medium. Therefore, we find that capacitance of a dielectric filled parallel plate capacitor becomes *K* times the capacitance with air or vacuum as dielectric :

$$C = KC_0 \qquad (16.27)$$

16.3.3 Relative Permittivity or Dielectric Constant

We can also define dielectric constant by calculating the force between the charges. According to Coulomb's law, the magnitude of force of interaction between two charges q_1 and q_2 separated by a distance *r* in vacuum is :

$$F_V = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2} \qquad (16.28)$$

where ε_0 is the permittivity of free space.

If these charges are held at the same distance in a material medium, the force of interaction between them will be given by

$$F_m = \frac{1}{4\pi\varepsilon}\frac{q_1 q_2}{r^2} \qquad (16.29)$$

On combining Eqns. (16.28) and (16.29), we get

$$\frac{F_v}{F_m} = \frac{\varepsilon}{\varepsilon_0} = \varepsilon_r \qquad (16.30)$$

where ε_r (or K) is relative permittivity. It is also termed as dielectric constant of the medium. Note that it is the ratio of permittivity of the material medium to the permittivity of free space. We can also define the dielectric constant of a medium as the ratio of the electrostatic force of interaction between two point charges held at certain distance apart in air or vacuum to the force of interaction between them held at the same distance apart in the material medium.

The dielectric constant can also be expressed as

$$K = \frac{\text{Capacitance with dielectric between the plates}}{\text{Capacitance with vacuum between the plates}}$$

$$= \frac{C_m}{C_0}$$

Thus

$$C_m = KC_0 \qquad (16.31)$$

For metals, $K = \infty$, for mica $K \approx 6$, and for paper $K = 3.6$.

16.4. GROUPING OF CAPACITORS

Capacitors are very important elements of electrical and electronic circuits. We need capacitors of a variety of capacitances for different purposes. Sometimes a capacitance of a proper value may not be available. In such situations, grouping of capacitors helps us to obtain desired (smaller or larger) value of capacitance with available capacitors. Two most common capacitor groupings are :

- Series grouping, and
- Parallel grouping

Let us learn about these now.

16.4.1 Parallel Grouping of Capacitors

In parallel grouping, one plate of each capacitor is connected to one terminal and the other plate is connected to another terminal of a battery, as shown in Fig. 16.10. Let V be the potential difference applied to the combination between points A and B. Note that *in parallel combination, potential difference across each capacitor is the same*. Therefore, charge on these will be different, say q_1, q_2 and q_3 such that

$$q_1 = C_1V$$

$$q_2 = C_2V$$

$$q_3 = C_3V \qquad (16.32)$$

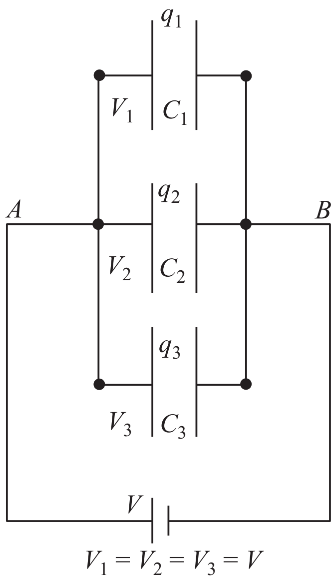


Fig. 16.10 : Capacitors joined in parallel

Total charge on all the capacitors of the combination is :

$$q = q_1 + q_2 + q_3$$

$$q = (C_1 + C_2 + C_3 + \ldots)V \qquad (16.33)$$

Let C_p be the equivalent capacitnace in parallel combination. Then

$$q = C_p V$$

From these relations, we get

$$q = C_p V = (C_1 + C_2 + C_3)V$$

In general, we can write

$$C_p = C_1 + C_2 + C_3 = \sum_{i=1}^{n} C_i \qquad (16.34)$$

Thus, we see that *equivalent capacitance of a number of capacitors joined in parallel is equal to the sum of the individual capacitances.*

Remember that in parallel combination, all the capacitors have the same potential difference between their plates but charge is distributed in proportion to their capacitances. Such a combination is used for charge accumulation.

16.4.2 Series Grouping of Capacitors

In the series combination of capacitors, the first plate of the first capacitor is connected to

the electrical source. The second plate of the first capacitor is connected to the first plate of the second capacitor. The second plate of second capacitor is connected to first plate of the next capacitor of the combination and so on. The second plate of last capacitor of the combination is connected to the electrical source, as shown in Fig.16.11. Let $+q$ unit of charge be given to the first plate of capacitor C_1 from the source. Due to electrical induction, as expalined in the principle of capacitor, $-q$ charge appears on the inner side of right plate of C_1 and $+q$ charge develops on the outer side of the second plate of C_1. The $+q$ unit of charge flows to the first plate of C_2 and so on. Thus, each capacitor receieves the same charge of magnitude q. As their capacitances are different, potential difference across these capacitors will be

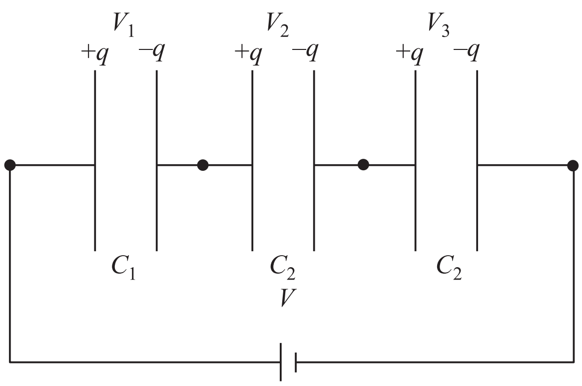


Fig.16.11 : Capacitors in series grouping. The amount of charge on each capacitor plate is same.

$$V_1 = \frac{q}{C_1},\ V_2 = \frac{q}{C_2},\ V_3 = \frac{q}{C_3} \qquad (16.35)$$

If C_s is the total capacitance of the series grouping, then

$$V = \frac{q}{C_s}$$

and $$V = V_1 + V_2 + V_3 \qquad (16.36)$$

Hence $$\frac{q}{C_s} = \frac{q}{C_1} + \frac{q}{C_2} + \frac{q}{C_3}$$

or $$\frac{1}{C_s} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \qquad (16.37)$$

For n capacitors joined in series, we can write

$$\frac{1}{C_s} = \sum_{i=1}^{n} \frac{1}{C_i}$$

Types of Capacitors

There are three common varieties of capacitors in commercial use. Their schematic diagrams are shown in Fig.16.12.

1. **Paper capacitor:** Several large thin sheets of paraffin impregnated paper or mylar are cut in proper size (rectangular). Several sheets of metallic foils are also cut to the same size. These are spread one over the other alternately. The outer sheet is mylar, then over it a sheet of metal foil, again over it a sheet of mylar and then a sheet of metal foil and so on. The entire system is then rolled in the form of a cylinder to form a small device.

2. **Metal plate capacitors:** A large number of metals are alternately joined to two metal rods as shown in Fig.16.12 (b). The entire plate system is immersed in silicon oil which works as dielectric material between the plates. High voltage capacitors are usually of this type. Variable capacitors of micro farad capacitance are usually of this type and use air as dielectric. One set of plates is fixed and the other set is movable. The movable plates, when rotated, change their effective area, thereby changing the capacitance of the system. You might see such capacitors in a radio receiver. Variable capacitance helps in tuning to different radio stations.

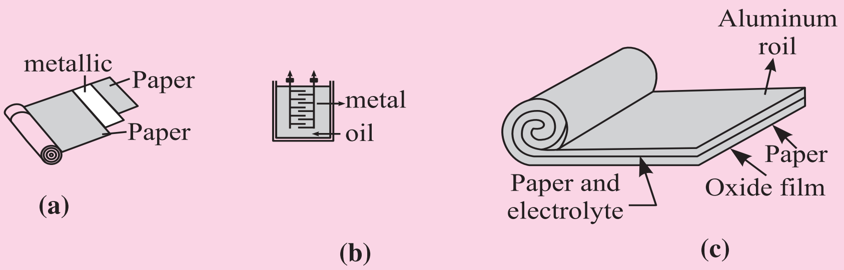


Fig.16.12 : Different types of capacitors : **a)** paper capacitor, **b)** variable capacitor, and **c)** electrolytic capacitor

3. **Electrolytic capacitor:** An electrolytic capacitor is shown Fig. 16.12(c). A metal foil is rolled in the shape of a cylinder with increasing diameter so that there is always a space between one surface and the other. The system is immersed in an electrolyte in the form of a solution. This solution is conducting because of ions in the solution. A voltage is applied between the electrolyte and the metallic foil. Because of the conducting nature of the electrolyte, a thin layer of metal oxide, which is an insulator, is formed on the foil. The oxide layer works as dielectric material. Since the dielectric layer is extremely thin, the system provides a very high value of capacitance. It is important in this type of capacitor to mark the **positive** and **negative** terminals. A wrong connection of positive and negative terminals removes the oxide layer. (The capacitor then starts conducting.) This type of capacitor is used in storing large amount of charge at low voltage.

Thus, *the reciprocal of equivalent capacitance of any number of capacitors connected in series is equal to the sum of the reciprocals of individual capacitances*. From the above relation, you will agree that C_s is less than the least of C_1, C_2, and C_3.

Note that *all the capacitors in series grouping have the same amount of charge but the potential difference between their plates are inversely proportional to their capacitances.* It means that the capacitor with minimum capacitance of the combination will have maximum potential difference between its plates.

Example 16.3 : The capacitance of a parallel plate air capacitor is 22.0 μF. The separation between the plates is d. A dielectric slab of thickness $d/2$ is put in-between the plates. Calculate the effective capacitance, if the dielectric constant $K = 5$.

Solution: The Capacitance of the air capacitor is given by

$$C_0 = \frac{\varepsilon_0 A}{d} = 22.0\ \mu\text{F}$$

The new system can be considered as a series combination of two capacitors:

$$C_1 = \frac{K\varepsilon_0 A}{d/2} = \frac{2K\varepsilon_0 A}{d} = 2\,KC_0$$

and

$$C_2 = \frac{\varepsilon_0 A}{d/2} = \frac{2\varepsilon_0 A}{d} = 2C_0$$

The effective capacitance C is given by

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}$$

or

$$C = \frac{C_1 C_2}{C_1 + C_2}$$

$$= \frac{2KC_0 \times 2C_0}{2KC_0 + 2C_0}$$

$$= \frac{2KC_0}{K+1}$$

$$= \frac{10 \times 22 \times 10^{-6}\text{F}}{6}$$

$$= 36.7\ \mu\text{F}$$

INTEXT QUESTIONS 16.2

1. Write the dimensions of capacitance.
2. What is the potential difference between two points separated by a distance d in a uniform electric field **E** ?
3. The usual quantities related with an air capacitor are C_0, E_0 and V_0. How are these related with C, **E** and V of the same capacitor filled with dielectric constant K ?
4. Calculate the area of air filled capacitor plate when the separation between the plates is 50 cm and capacitance is 1.0 μF .

16.4.3 Energy Stored in a Capacitor

The charging of a capacitor can be visualized as if some external agent, say a battery, pulls electrons from the positive plate of a capacitor and transfers them to the negative plate. Some work is done in transferring this charge, which is stored in the capacitor in the form of electrostatic potential energy. This energy is obtained from the battery (stored as chemical energy). When this capacitor is discharged through a resistor, this energy is released in the form of heat.

Let us assume that an uncharged capacitor, when connected to a battery, develops a maximum charge q. The charging takes place slowly. The initial potential difference between the capacitor plates is zero and the final potential difference is V. The average potential difference during the entire process of charging is

$$\frac{0+V}{2} = \frac{V}{2} = \frac{q}{2C}$$

The work done during charging is given by

$$W = \text{Charge} \times \text{potential difference}$$

$$= q \, \frac{q}{2C} = \frac{1}{2}\frac{q^2}{C}$$

Hence potential energy

$$U = \frac{1}{2}qV = \frac{1}{2}\frac{q^2}{C} = \frac{1}{2}CV^2 \qquad (16.38)$$

This energy is stored in the electric field between the plates. The stored energy is directly proportional to the capacitance. It also increases as potential difference increases. However, every capacitor can store only a limited amount of energy. An automatic discharge will take place when the potential difference becomes more than its threshold value.

It is dangerous to touch the plates of a charged capacitor. The capacitor may get discharged through your body resulting in an electric shock. Such a shock could be fatal for high value capacitors when fully charged.

16.5 DIELECTRICS AND DIELECTRIC POLARIZATION

We know that dielectrics are insulating materials, which transmit electric effects without conducting. Dielectrics are of two types : **non-polar** and **polar**. We now learn about these.

(a) Non-polar dielectrics

In the molecules of non–polar dielectrics, the centre of positive charge coincides with the centre of negative charge. Each molecule has zero dipole moment in its normal state. These molecules are mostly symmetrical such as nitrogen, oxygen, benzene, methane, CO_2, etc.

(b) Polar dielectrics

Polar dielectrics have asymmetric shape of the molecules such as water, NH_3, HCl etc. In such molecules, the centres of positive and negative charges are separated through a definite distance and have finite permanent dipole moment.

When a non-polar dielectric is held in an external electric field, the centre of positive charge in each molecule is pushed in the direction of **E** and the centre of negative charge is displaced in the direction opposite to **E**. Because of external electric field, centres of positive and negative charges in the non-polar dielectric molecules are separated. Dielectric is then said to be polarized and a tiny dipole moment develops in each molecule. In fact, the force due to external electric field pulling the charge centres apart balances the force of mutual attraction between the centres (i.e. equilibrium is set) and the molecule is said to be polarised. Induced dipole moment **p** acquired by the molecule may be written as

$$\mathbf{p} = \alpha \, \varepsilon_0 \mathbf{E}$$

where α is constant of proportionality and is called atomic/molecular polarizability.

Let us now consider a non-polar slab *ABCD* placed in an electric field **E** maintained between the plates of a capacitor. As shown in Fig.16.13, the dielectric slab gets polarised. The nuclei of dielectric molecules are displaced towards the negative plate and electrons towards the positive plate. Because of polarisation, an electric field $\mathbf{E}_p$ is produced within the dielectric, which is opposite to **E**. Hence, due to the presence of a non-polar dielectric, the field between the plates is reduced, i.e. effective electric field in a polarised dielectric is given by

$$\mathbf{E}(\text{effective}) = \mathbf{E} - \mathbf{E}_p \qquad (16.39)$$

Thus, the potential difference between the capacitor plates is correspondingly reduced (as $V = Ed$), increasing the value of capacitance of the capacitor (as $C = q/V$).

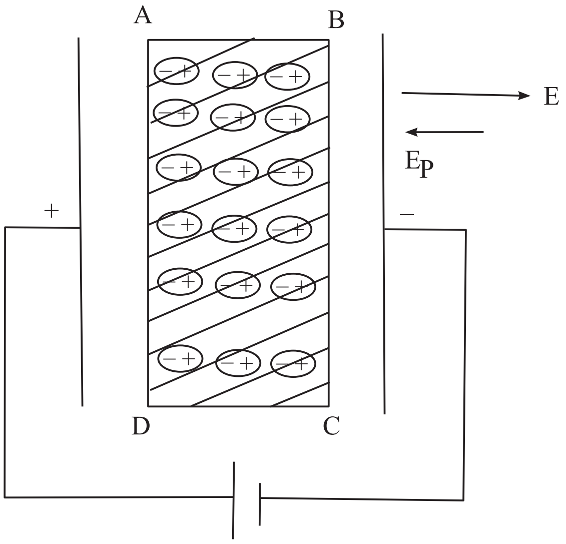


Fig.16.13 : A dielectric slab between the charged capacitor plates.

Applications of Electrostatics

Electrostatics provides basis for the theory of electromagnetics, apart from useful assistance in many fields of science and technology.

- Capacitors are essential parts of most electronic and electrical circuitry. These play a very crucial role in power transmission.
- Gold leaf electroscope – the simple device used for detecting charge, paved the way for cosmic ray research.
- Lightning conductor devised by Benjamin Franklin is still used to protect sky-scrappers from the strokes of lightning and thunder.
- The working of photocopiers, so common these days, is based on the principle of electrostatics.

INTEXT QUESTIONS 16.3

1. Two capacitors $C_1 = 12$ mF and $C_2 = 4$ mF are in group connections. Calculate the effective capacitance of the system when they are connected (a) in series (b) in parallel.
2. Four capacitors are connected together as shown in Fig.16.14. Calculate the equivalent capacitance of the system.
3. An air capacitor $C = 8$ mF is connected to a 12V battery. Calculate
   (a) the value of Q when it is fully charged?
   (b) the charge on the plates, when slab of dielectric constant $K = 5$ fills the gap between the plates completely.
   (c) potential difference between the plates; and
   (d) capacitance of the new capacitor

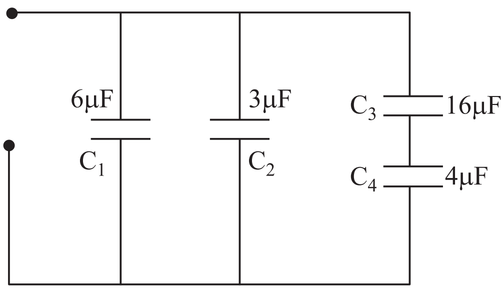


Fig.16.14 : Grouping of capacitors

4. A parallel plate capacitor of capacitance C_0 is connected to a battery and charged to a potential difference V_0. After disconnecting the battery, the gap between the plates is completely filled with a slab of dielectric constant K. How much energy is stored in the capacitor (a) in the first state? (b) in the second state? and (c) which one is larger and why?

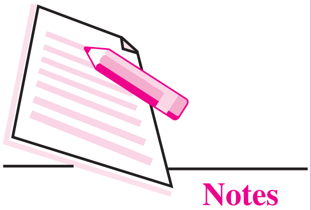
WHAT YOU HAVE LEARNT

- The potential at any point in an electric field is equal to the work done against the electric field in moving a unit charge from infinity to that point.
- Work done in transferring a charge from one point to another in an electrostatic field is path independent.
- If one joule of work is done in bringing a test charge of one coulomb from infinity to a point in the field, we say that potential at that point is one volt.
- Electric potential due to a dipole is zero at every point on the equatorial line of the dipole.
- In an equipotential surface, every point has same electric potential.
- At any point in an electric field, the negative rate of change of potential with distance (called potential gradient) gives the field.
- Electrostatic shielding is the phenomenon of protecting a region of space from electric field.

- Capacitance of a conductor depends on its shape, size and nature of medium, rather than its material.
- The capacitance of a dielectric filled parallel plate capacitor becomes K times the capacitance with air or vacuum as dielectric.
- Relative permittivity is the ratio of capacitance with dielectric between the plates to the capacitance with air or vacuum between the plates.
- In series combination of capacitors, the equivalent capacitance is less than the least of any of the individual capacitances.
- In parallel combination of capacitors, the equivalent capacitance is equal to the sum of individual capacitances.
- Due to the presence of a non-polar dielectric, the field between the plates of a capacitor is reduced.

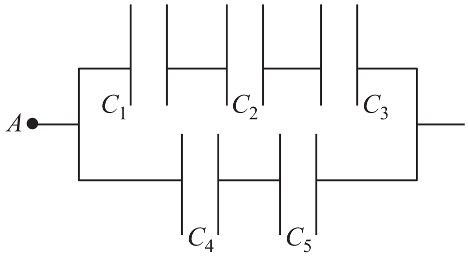
TERMINAL EXERCISES

1. Calculate the potential at a point P at a distance of 30 cm from a point charge $q = 20\ \mu C$
2. Three point charges q_1, q_2 and q_3, each of magnitude 200 μC, are placed at the corners A, B and C respectively of an equilateral triangle. The length of the side is 10cm. Calculate the potential energy of the system.
3. The potential difference between the plates of a capacitor separated by 3mm is 12.0 V. Calculate the magnitude of E between the plates?
4. Two ions having charges $+e$ and $-e$ are 4.0×10^{-10} m apart. Calculate the potential energy of the system.
5. The plates A and B of a parallel plate capacitor have a potential difference of 15 V. A proton ($m = 1.67 \times 10^{-27}$ kg) is moved from the positive plate A to B. Calculate the speed of the proton near plate B.
6. Show that dimensionally the quantities Vq and $(½)mv^2$ are equivalent. The symbols carry the usual meaning.
7. Under what condition, the electric field between the plates of a parallel plate capacitor is uniform?
8. A metallic sphere of radius r has a charge $+q$. Calculate the work done in moving a test charge q_0 from one end of a diameter to its other end.
9. A parallel plate air capacitor of value C_0 is charged to a potential V_0 between the plates and $+q_0$ is charge on one plate. Separation between plates is d. A dielectric of dielectric constant $K = 3$ fills the space between the plates. Which of these quantities will change and why. (i) capacitance (ii) charge (iii) potential difference and (iv) field density?



Notes

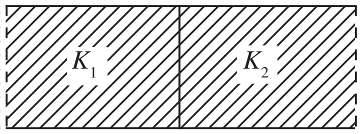
10. Examine the following network of capacitors. The potential difference between A and B is 16V :



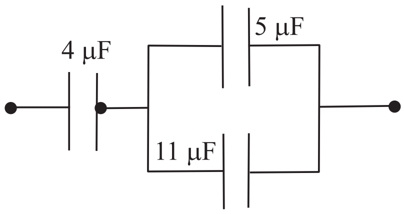
$$C_1 = 2\ \mu F,\ C_2 = 4\ \mu F,\ C_3 = 8\ \mu F,\ C_4 = 3\ \mu F,\ C_5 = 3\mu F$$

Calculate (a) the effective capacitance between A and B, (b) the charge on each capacitor, and (c) the potential difference across each capacitor.

11. The value of capacitance of an air capacitor is 8μF. Two dielectrics of identical size fill the space between the plates as shown. Dielectric constants are K_1 = 3.0 and K_2 = 6.0. Calculate the value of the new capacitance.



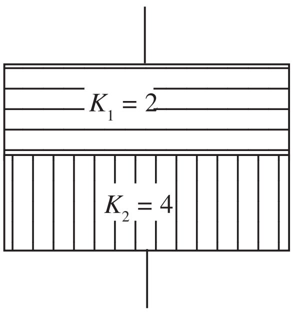
12. Calculate the equivalent capacitance of the following system.



13. A 3.0 μF air capacitor is charged to a potential 12.0 V. A slab of dielectric constant K = 7 is made to fill the space. Calculate the ratio of the energies stored in the two systems.

14. A dipole of dipole moment $P = 3.5 \times 10^{-15}$ Cm is placed in a uniform electric field $E = 2.0 \times 10^4$ NC^{-1}. The dipole makes an angle of 60° with the field. Calculate the (a) Potential energy of the dipole and (b) the torque on the dipole.

15. The capacitance of a parallel plate air capacitor is 12μF. The separation between the plates is 8mm. Two dielectric slabs of the same size fill the air space. Calculate the new value of capacitance.



ANSWERS TO INTEXT QUESTIONS

16.1

1. The potential at r $(r > R)$

$$V = \frac{1}{4\pi\varepsilon_0}\frac{q}{r}$$

2. The field around a point charge possesses spherical symmetry. Thus every point on the surface of the sphere is equipotential. And no work is done when a charge moves on an equipotantial surface

3. $E = -\frac{dV}{dr}$ Since V is constant, E is zero.

   We can obtain the same result using Eqn. (16.22) :

   $E = \frac{V_A - V_B}{d}$. Since $V_A = V_B, E$ is zero

4. No. Not necessarily. When $E = 0$, the potential is either constant or zero.

5. Two equipotential surfaces never intersect. If they do so, at the point of intersection we can draw two normals giving directions of electric field.

16.2

1. $$C = \frac{Q}{V} = \frac{Q}{\frac{\text{Work done}}{\text{Charge}}} = \frac{Q \times Q}{\text{Work done}}$$

   $$= \frac{\text{Q}^2}{\text{N.m.}}$$

   The basic unit is

   $$\text{A} = \frac{\text{C}}{\text{s}}$$

$\therefore$ $\text{C}^2 = \text{A}^2\,\text{s}^2$ and newton = mass × acc = $\text{kg}\,\frac{\text{m}}{\text{s}^2}$

$$\text{Capacitance} = \frac{\text{C}^2}{\text{Nm}} = \frac{\text{A}^2\text{s}^2}{\text{kg}\frac{\text{m}^2}{\text{s}^2}} = \frac{\text{A}^2\text{s}^4}{\text{kg m}^2}$$

$$= \text{A}^2\,\text{s}^4\,(\text{kg m}^2)^{-1}$$

2. In a capacitor, E is uniform between this plates. Potential difference between the plates

   $V_A - V_B = E \times d.$

3. C_0, E_0, V_0 for air capacitor and
   C, E, V for dielectric capacitor. Then

$$k = \frac{C}{C_0},\ k = \frac{V_0}{V}\ \ k = \frac{E_0}{E}.$$

4. $C = 1.0\ \mu\ \text{F} = 1.0 \times 10^{-6}$ F.
   $d = 50$ cm $= 0.5$m.

$$C = \frac{\varepsilon_0 A}{d}$$

$\therefore\ A = \frac{Cd}{\varepsilon_0}$. Since $\varepsilon_0 = 8.85 \times 10^{-12}$,

$$A = \frac{1.0 \times 10^{-6} \times 0.5}{8.85 \times 10^{-12}}$$

$$= \frac{5 \times 10^{-7}}{8.85 \times 10^{-12}}$$

$$= 0.56 \times 10^{5}\ \text{m}^2$$

16.3

1 (a) 3 mF (b) 16 mF 2. 12.2μF

3. (a) 96mC (b) 0.480 C (c) 12 v (d) 40 mF

4. (a) $\frac{1}{2} C_0 V_0^2$ (b) $\frac{1}{2}\frac{(C_0 V_0)^2}{C_0 R} = \frac{1}{2k} C_0 V_0^2$

   (c) The energy in the first case is more, because same energy is used up for sucking in the dielectric slab.

Answers to Problems in Terminal Exercises

1. 6×10^5 V.
2. 1.08×10^4 J.
3. $4 \times 103\ \text{Vm}^{-1}$
4. -5.76×10^{-19} J
5. $1.4 \times 109\ \text{ms}^{-2}$

10. (a) $\frac{37}{14}$ μF, (b) $\frac{128}{7}$ μC, $\frac{128}{7}$ μC, $\frac{128}{7}$ μC, 24μC,

    (c) $\frac{64}{7}$ V, $\frac{32}{7}$ V, $\frac{16}{7}$ V, 8V, 8V

11. 36 μF.
12. $\frac{16}{5}$ μF.
13. 1 : 7
14. (a) 3.5×10^{-11} J (b) 6×10^{-11}Nm.
15. 32 μF

312en17

17

ELECTRIC CURRENT

In our daily life we use electricity for various activities. The electric lamps and tubes light our houses, we listen music on a tape recorder or radio, see different programmes on television, enjoy cool breeze from electric fan or cooler, and use electric pump to irrigate fields. In fact, electricity is a unique gift of science to mankind. We can not imagine life without electricity in the modern world. At home you might have observed that as soon as you switch on an electrtc lamp, it starts glowing. Why does it happen? What is the function of a switch?

In the preceding lessons of this module, you have studied about static electric charges and forces between them. In this lesson, you will learn about electric charges in motion. You will also learn that the rate of flow of charge through a conductor depends on the potential difference across it. You will also study the distribution of current in circuits and Kirchhoff's laws which govern it. Elementary idea of primary and secondary cells will also be discussed in this lesson.

Physics is an experimental science and the progress it has made to unfold laws of nature became possible due to our ability to verify theoretical predictions or reproduce experimental results. This has led to continuous improvement in equipment and techniques. In this lesson you will learn about potentiometer, which is a very versatile instrument. It can be used to measure resistance as well as electro-motive force using null method.

OBJECTIVES

After studying this lesson, you should be able to :

- *state Ohm's law and distinguish between ohmic and non-ohmic resistances;*
- *obtain equivalent resistance for a series and parallel combination of resistors;*
- *distinguish between primary and secondary cells;*

- *apply Kirchhoff's rules to closed electrical circuits;*
- *apply Wheatstone bridge equation to determine an unknown resistance; and*
- *explain the principle of potentiometer and apply it to measure the e.m.f and internal resistance of a cell.*

Free and Bound Electrons

An atom is electrically neutral, i.e. as many negatively charged electrons revolve around the nucleus in closed orbits as there are positively charged protons inside it. The electrons are bound with the nucleus through Coulomb (attractive) forces.

Farther the electrons from the nucleus, weaker is the Coulomb force. The electrons in the outermost orbit are, therefore, most loosely bound with the nucleus. These are called *valence electrons*. In metallic solids, the valence electrons become free to move when a small potential difference is applied.

17.1 ELECTRIC CURRENT

You have studied in the previous lesson that when a potential difference is applied across a conductor, an electric field is set up within it. The free electrons move in a direction opposite to the field through the conductor. This constitutes an electric current. Conventionally, *the direction of current is taken as the direction in which a positive charge moves*. The electrons move in the opposite direction. To define current precisely, let us assume that the charges are moving perpendicular to a surface of area *A*, as shown in Fig. 17.1. The current is the rate of flow of charge through a surface area placed perpendicular to the direction of flow. If charge Δq flows in time Δt, the average current is defined as :

$$I_{av} = \frac{\Delta q}{\Delta t} \qquad (17.1)$$

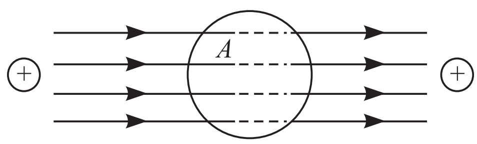


Fig. 17.1 : Motion of charges inside a conductor of surface area *A*

If the rate of flow of charge varies with time, the current also varies with time. The instantaneous current is expressed as :

$$I = \frac{dq}{dt} \qquad (17.2)$$

The electric current through a conductor is the rate of transfer of charge across a surface placed normal to the direction of flow.

The SI unit of current is ampere. Its symbol is A :

$$1 \text{ ampere} = \frac{1 \text{ coulomb}}{1 \text{ second}} \qquad (17.3)$$

The smaller units of current are milliampere, 1 mA = 10^{-3}A, and microampere, 1μA = 10^{-6}A. The current can arise due to flow of negative charges (electrons), as in metals. In a semiconductor, flow of electrons (negative charge) and holes constitutes current. Holes are vacancies in a crystal. These are taken as positively charged particles having the same amount of charge as that on an electron. You will study about these particles in more detail in lesson 28.

Let us consider a conductor of cross sectional area A shown in Fig. 17.2. The volume element for a length Δx is $A\ \Delta x$. If n is the number of electrons per unit volume, the number of electrons in this volume element will be $nA\Delta x$. The total charge in this volume

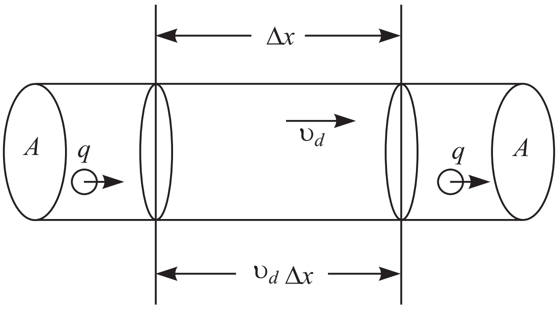


Fig. 17.2 : The charges move with a speed υ_d through a surface of area A. The number of charges in a length Δx is $nA\ \upsilon_d\ \Delta t$.

element is $\Delta q = nA\Delta xe$, where e is charge on the electron. If electrons drift with a speed υ_d due to thermal energy, the distance travelled in time Δt is $\Delta x = \upsilon_d \Delta t$. On substituting this value of Δx in the expression for Δq, we find that total charge in the volume element under consideration is given by

$$\Delta q = nAe\ \upsilon_d\ \Delta t$$

so that

$$\frac{\Delta q}{\Delta t} = I = nAe\upsilon_d \qquad (17.4)$$

You will learn more about the drift velocity in sec.17.9.

17.2 OHM'S LAW

In 1828, Ohm studied the relation between current in a conductor and potential difference applied across it. He expressed this relation in the form of a law, known as **Ohm's law**.

George Simon Ohm
(1787-1854)

German physicist, George Simon Ohm is famous for the law named after him. He arrived at the law by considering an analogy between thermal and electrical conduction. He also contributed to theory of sirens, interference of polarised light in crystals etc. Ohm, the practical unit of resistance, is named in his honour.

According to Ohm's law, **the electric current through a conductor is directly proportional to the potential difference across it, provided the physical conditions such as temperature and pressure remain unchanged.**

Let V be the potential difference applied across a conductor and I be the current flowing through it. According to Ohm's law,

$$V \propto I$$

or $$V = RI$$

$$\Rightarrow \qquad \frac{V}{I} = R \qquad (17.5)$$

where constant of proportionality R signifies the electrical resistance offered by a conductor to the flow of electric current. *Resistance is the property of a conductor by virtue of which it opposes the flow of current through it.* The I–V graph for a metallic conductor is a straight line (Fig. 17.3(a)).

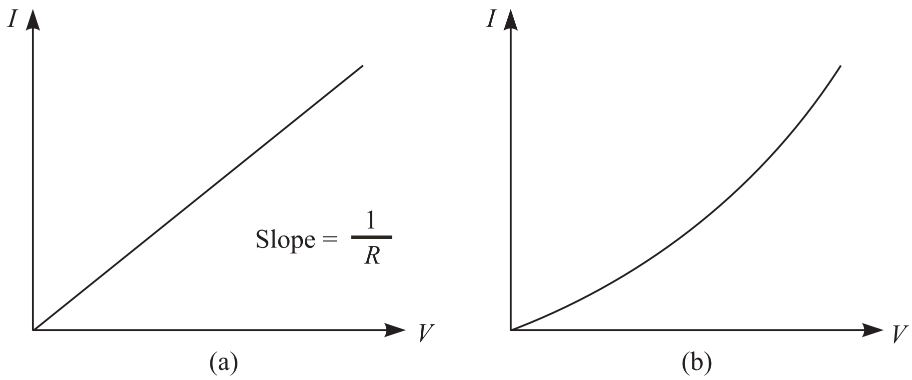


Fig. 17.3 : Current-voltage graph for **a)** an ohmic device, and **b)** a semiconductor diode

The SI unit of resistance is ohm. It is expressed by symbol Ω (read as omega)

1 ohm = 1 volt/1 ampere

Most of the metals obey Ohm's law and the relation between voltage and current is linear. Such resistors are called *ohmic*. Resistors which do not obey Ohm's law are called *non-ohmic*. Devices such as vacuum diode, semiconductor diode, transistors show non ohmic character. For semiconductor diode, Ohm's law does not hold good even for low values of voltage. Fig. 17.3(b) shows a non-linear *I–V* graph for a semiconductor diode.

ACTIVITY 17.1

Aim : To study conduction of electricity through an electrolyte.

Material Required Ammeter, Voltmeter, a jar containing copper sulphate solution, two copper plates, a battery, plug key, connecting wires and a rheostat.

How to Proceed :

1. Set up the apparatus as shown in Fig. 17.4.
2. Plug in the key and note ammeter and voltmeter readings.
3. Change the value of ammeter reading by moving the sliding contact of rheostat and note voltmeter reading again.
4. Repeat step –3 at least five times and record ammeter and voltmeter readings each time.

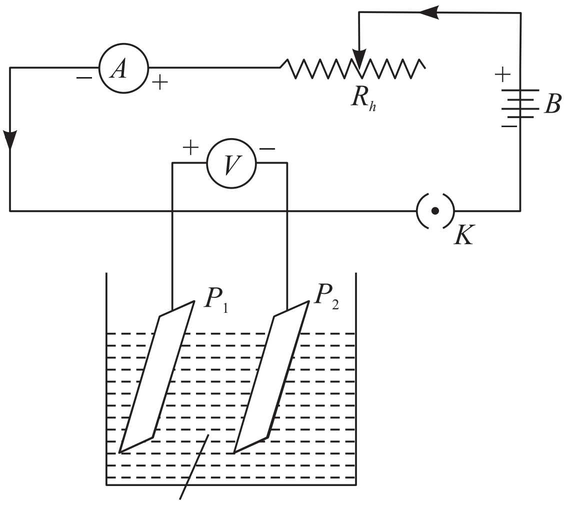


Fig. 17.4 : Electrical conduction through an electrolyte

5. Repeat the experiment by changing (a) seperation between P_1 and P_2, (b) plate area immersed in electrolyte, and (c) concentration of electrolyte.
6. Plot *I–V* graph in each case.

What do you conclude?

- If *I–V* graph is a straight line passing through the orgin, as shown in Fig. 17.5, we say that ionic solution behaves as an ohmic resistor.
- The slope of the graph changes steeply with change in volume of electrolyte between the plates. It means that resistivity of an electrolyte depends not only on its nature but also on the area of the electrodes and the seperation between them.

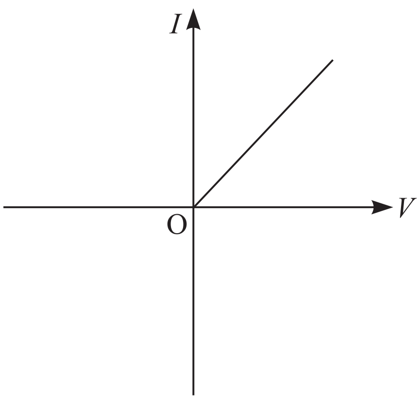


Fig. 17.5 : *I–V* graph for an ionic solution

17.2.1 Resistance and Resistivity

Let us now study the factors which affect the resistance of a conductor. You can perform two simple experiments. To do so, set up a circuit as shown in Fig. 17.6.

ACTIVITY 17.2

Take a long conducting wire of uniform cross section. Cut out pieces of different lengths, say l_1, l_2, l_3, etc from it. This makes sure that wires have same area of cross-section. Connect l_1 between *A* and *B* and note down the current through this wire. Let this current be *I*. Perform the same experiment with wires of lengths l_2 and l_3, one by one. Let the currents in the wires be I_2 and I_3 respectively.

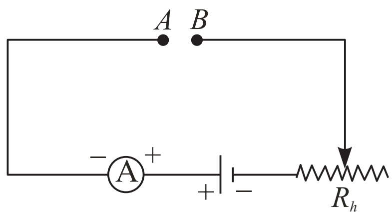


Fig. 17.6 : Electrical circuit to study factors affecting resistance of conductors

Plot a graph between l^{-1} and *I*. You will find that the graph is a straight line and longer wires allow smaller currents to flow. That is, longer wires offer greater resistance [Fig.17.7(a)]. Mathematically, we express this fact as

$$R \propto l \qquad (17.6)$$

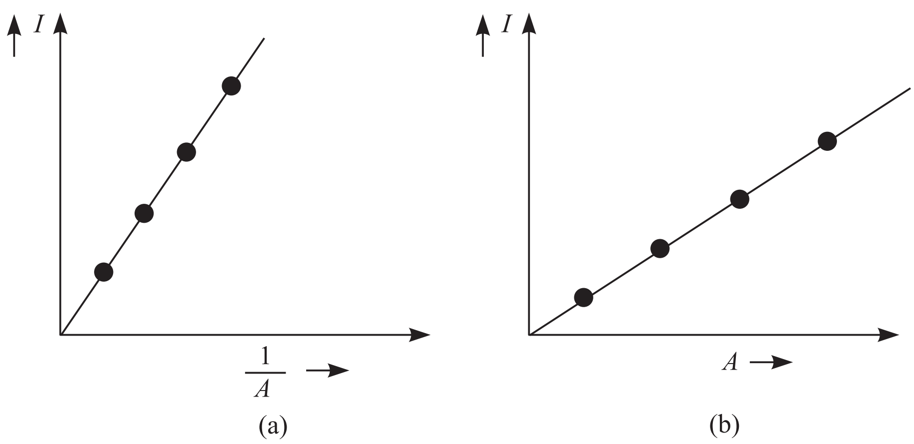


Fig. 17.7 : **a)** The graph between *I* and 1/*l* for wires of uniform cross-section and **b)** the graph between current and area of cross section for wires of same length

ACTIVITY 17.3

Take wires of the same length of a given material but having different areas of cross section, say A_1, A_2, A_3 etc. Connect the wires between *A* and *B* one by one and note down the currents I_1, I_2, I_3 etc. in each case. A plot of *I* and *A* will give

a straight line. Wires of greater cross sectional area allow greater currents to flow. You may say that wires of larger area of cross-section offer smaller resistance [Fig. 17.7 (b)]. Mathematically, we can write

$$R \propto \frac{1}{A} \qquad (17.7)$$

On combining Eqns.(17.6) and (17.7), we can write

$$R \propto \frac{\ell}{A}$$

or

$$R = \rho \frac{\ell}{A} \qquad (17.8)$$

where ρ is a constant for the material at constant temperature. It is called the **specific resistance** or **resistivity** of the material. By rearranging terms, we can write

$$\rho = \frac{RA}{\ell} \qquad (17.9)$$

If ℓ = 1m and A = 1m^2, then ρ = R ohm-metre. Thus *resistivity of a material is the resistance offered by a wire of length one metre and area of cross section one m^2*. The unit of resistivity is ohm metre (Ωm)

Reciprocal of resistivity is called conductivity (specific conductance) and is denoted by σ :

$$\sigma = \frac{1}{\rho} \qquad (17.10)$$

Unit of **conductivity** is Ohm^{-1} metre^{-1} or mho-metre^{-1} or Sm^{-1}.

Resistivity depends on the nature of the material rather than its dimensions, whereas the resistance of a conductor depends on its dimensions as well as on the nature of its material.

You should now study the following examples carefully.

Example 17.1 : In our homes, the electricity is supplied at 220V. Calculate the resistance of the bulb if the current drawn by it is 0.2A.

Solution :

$$R = \frac{V}{I} = \frac{220 \text{ volt}}{0.2 \text{ amp.}} = 1100 \ \Omega$$

Example 17.2 : A total of 6.0 × 10^{16} electrons pass through any cross section of a conducting wire per second. Determine the value of current in the wire.

Solution : Total charge passing through the cross-section in one second is

$\Delta Q = ne = 6.0 \times 10^{16} \times 1.6 \times 10^{-19}$ C $= 9.6 \times 10^{-3}$C

$$\therefore \qquad I = \frac{\Delta Q}{\Delta t} = \frac{9.6\times10^{-3}\,\mathrm{C}}{1\,\mathrm{s}}$$

$$= 9.6 \times 10^{-3}\ \mathrm{A}$$

$$= 9.6\ \mathrm{mA}$$

Example 17.3 : Two copper wires *A* and *B* have the same length. The diameter of *A* is twice that of *B*. Compare their resistances.

Solution : From Eqn. (17.8) we know that

$$R_A = \rho\frac{\ell}{\pi r_A^2} \quad \text{and} \quad R_B = \rho\frac{\ell}{\pi r_B^2}$$

$$\therefore \qquad \frac{R_A}{R_B} = \frac{r_B^2}{r_A^2}$$

Since diameter of $A = 2 \times$ diameter of B, we have $r_A = 2r_B$. Hence

Resistance of *B* will be four times the resistance of *A*.

Example 17.4 : The length of a conducting wire is 60.0 m and its radius is 0.5cm. A potential difference of 5.0 V produces a current of 2.5 A in the wire. Calculate the resistivity of the material of the wire.

Solution : $$R = \frac{V}{I} = \frac{5.0\,V}{2.5\,A} = 2.0\ \Omega$$

Radius of the wire = 0.5 cm = 5.0×10^{-3} m

Area of cross section $A = \pi R^2 = 3.14 \times (5.0 \times 10^{-3})^2\ \mathrm{m}^2 = 78.5 \times 10^{-6}\ \mathrm{m}^2$

$$\therefore \qquad \rho = \frac{2.0\times78.5\times10^{-6}\ \Omega\mathrm{m}^2}{60.0\ \mathrm{m}} = 2.6 \times 10^{-6}\ \Omega\mathrm{m}$$

INTEXT QUESTIONS 17.1

1. (a) A current *I* is established in a copper wire of length ℓ. If the length of the wire is doubled, calculate the current due to the same cell.

   (b) What happens to current in an identical copper wire if the area of cross section is decreased to half of the original value?

2. The resistivity of a wire of length *l* and area of cross section *A* is $2 \times 10^{-8}\Omega$m. What will be the resistivity of the same metallic wire of length 2*l* and area of cross section 2*A*?

3. A potential difference of 8 V is applied across the ends of a conducting wire of length 3m and area of cross section 2cm^2. The resulting current in the wire is 0.15A. Calculate the resistance and the resistivity of the wire.
4. Do all conductors obey Ohm's law? Give examples to support your answer.
5. 5×10^{17} electrons pass through a cross-section of a conducting wire per second from left to right. Determine the value and direction of current.

17.3 GROUPING OF RESISTORS

An electrical circuit consists of several components and devices connected together. Some of these are batteries, resistors, capacitors, inductors, diodes, transistors etc. (They are known as circuit elements.) These are classified as resistive and reactive. The most common resistive components are resistors, keys, rheostats, resistance coils, resistance boxes and connecting wires. The reactive components include capacitors, inductors and transformers. In addition to many other functions performed by these elements individually or collectively, they control the current in the circuit. In the preceding lesson you learnt how grouping of capacitors can be used for controlling charge and voltage. Let us now discuss the role of combination of resistors in controlling current and voltage.

Two types of groupings of resistors are in common use. These are : **series grouping and parallel grouping. We define equivalent resistance of the combination as a single resistance which allows the same current to flow as the given combination when the same potential difference is applied across it.**

17.3.1 Series Combination

You may connect many resistors in series by joining them end-to-end such that the same current passes through all the resistors. In Fig. 17.8, two resistors of resistances R_1 and R_2 are connected in series. The combination is connected to a battery at the ends *A* and *D*. Suppose that current *I* flows through the series combination when it is connected to a battery of voltage *V*. Potential differences V_1 and V_2 develop across R_1 and R_2, respectively. Then $V_1 = IR_1$ and $V_2 = IR_2$. But sum of V_1 and V_2 is equal to *V*, i.e.

$$\Rightarrow \qquad V = V_1 + V_2 = I R_1 + I R_2$$

If equivalent resistance of this series combination is *R*, then

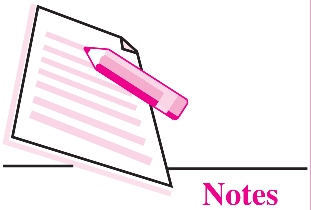
$$V = IR = I (R_1 + R_2)$$

so that

$$R = R_1 + R_2$$

This arrangement may be extended for any number of resistors to obtain

$$R = R_1 + R_2 + R_3 + R_4 + \cdots \qquad (17.11)$$



Notes

That is, the equivalent resistance of a series combination of resistors is equal to the sum of individual resistances. If we wish to apply a voltage across a resistor (say electric lamp) less than that provided by the constant voltage supply source, we should connect another resistor (lamp) in series with it.

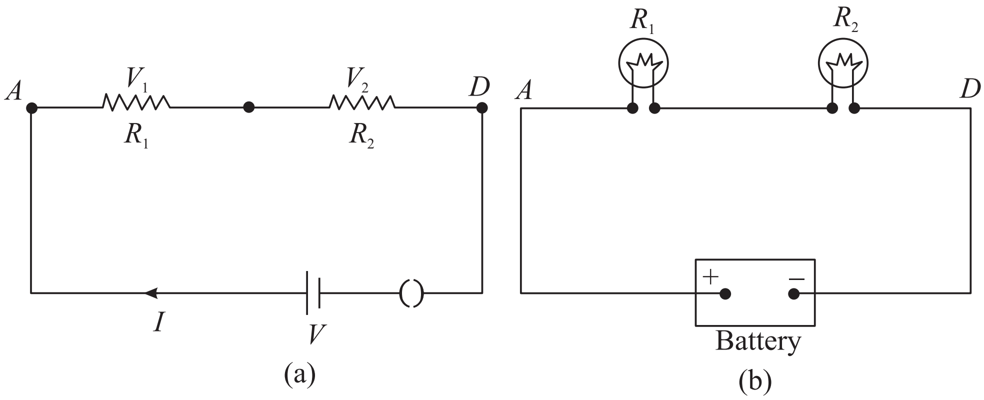


Fig. 17.8 : **a)** Two resistors connected in series to a battery, and **b)** two lamps joined in series connected to *a dc* source.

17.3.2 Parallel Combination

You may connect the resistors in parallel by joining their one end at one point and the other ends at another point. In parallel combination, **same potential difference exists across all resistors.** Fig. 17.9 shows a parallel combination of two resistors R_1 and R_2. Let the combination be connected to a battery of voltage V and draw a current I from the source.

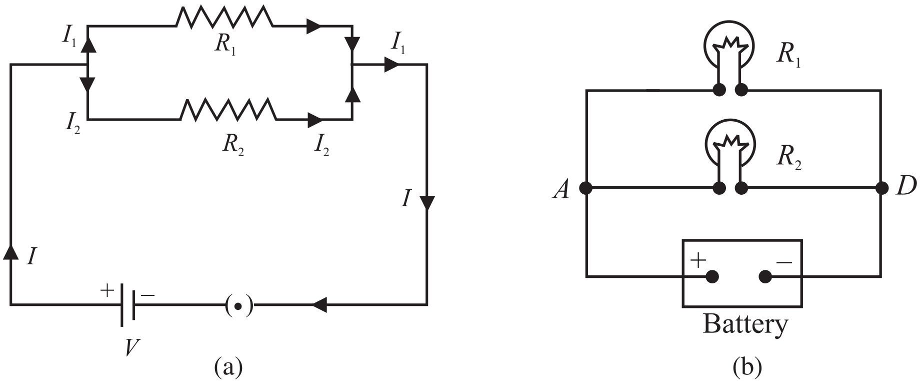


Fig. 17.9 : **a)** Two resistors connected in parallel. The battery supplies the same voltage to both resistors, and **b)** lamps connected in parallel to a battery.

The main current divides into two parts. Let I_1 and I_2 be the currents flowing through resistors R_1 and R_2, respectively. Then $I_1 = V/R_1$ and $I_2 = V/R_2$.

The main current is the sum of I_1 and I_2. Therefore, we can write

$$\Rightarrow \qquad I = I_1 + I_2 = \frac{V}{R_1} + \frac{V}{R_2}$$

If the equivalent resistance of combination is R, we write $V = IR$ or $I = V/R$:

$$I = \frac{V}{R} = \frac{V}{R_1} + \frac{V}{R_2}$$

$$\therefore \quad \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} \quad (17.12a)$$

or

$$R = \frac{R_1 R_2}{R_1 + R_2} \quad (17.12b)$$

From Eqn. (17.12a) we note that **reciprocal of equivalent resistance of parallel combination is equal to the sum of the reciprocals of individual resistances.** The process may be extended for any number of resistors, so that

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4} + \ldots \quad (17.13)$$

Note that the equivalent resistance of parallel combination is smaller than the smallest individual resistance. You may easily see this fact by a simple electrical circuit having a resistor of 2 Ω connected across a 2V battery. It will draw a current of one ampere. When another resistor of 2 Ω is connected in parallel, it will also draw the same current. That is, total current drawn from the battery is 2A. Hence, resistance of the circuit is halved. As we increase the number of resistors in parallel, the resistance of the circuit decreases and the current drawn from the battery goes on increasing.

In our homes, electrical appliances such as lamps, fans, heaters etc. are connected in parallel and each has a separate switch. Potential difference across each remains the same and their working is not influenced by others. As we switch on bulbs and fans, the resistance of the electrical circuit of the house decreases and the current drawn from the mains goes on increasing (Fig.17.10).

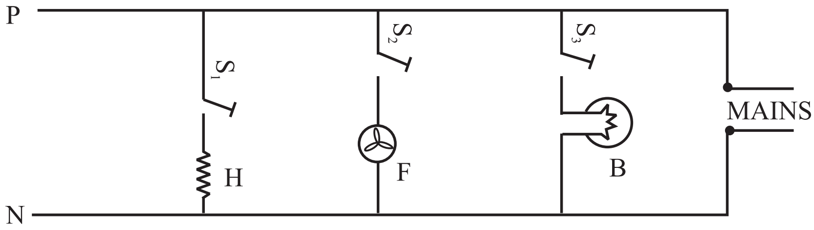


Fig. 17.10 : Arrangement of appliances in our homes. These are connected in parallel so that every appliance is connected to 220 V main supply. The total current drawn from the mains is the sum of the currents drawn by each appliance.

Example 17.5 : For the circuit shown in Fig. 17.11, calculate the value of resistance R_2, and current I_2 flowing through it.

Solution: If the equivalent resistance of parallel combination of R_1 and R_2 is R, then

$$R = \frac{R_1 R_2}{R_1 + R_2} = \frac{10 R_2}{10 + R_2}$$

According to Ohm's law,

$$R = \frac{50}{10} = 5\Omega$$

$$\therefore \quad \frac{10\,R_2}{10 + R_2} = 5$$

$$\Rightarrow \; 10\,R_2 = 50 + 5\,R_2 \quad \text{or} \quad R_2 = 10\,\Omega$$

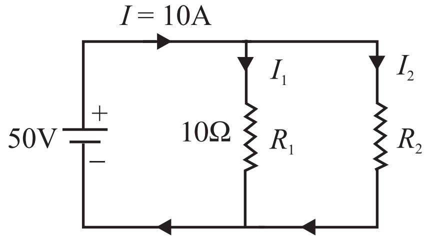


Fig. 17.11 : Two resistors in parallel

Since R_1 and R_2 are equal, current will be equally divided between them. Hence, $I_2 = 5$A

Example 17.6 : For the circuit shown in Fig. 17.12, calculate the equivalent resistance between points *a* and *d*.

Solution : 15Ω and 3Ω resistors are connected in parallel. The equivalent resistance of this combination is

$$R_1 = \frac{15 \times 3}{15 + 3}\Omega = \frac{45}{18} = \frac{5}{2} = 2.5\Omega$$

Now we can regard the resistances 5Ω, R_1 = 2.5Ω and 7Ω as connected in series. Hence, equivalent resistance between points *a* and *d* is

$$R = (5 + 2.5 + 7) = 14.5\ \Omega$$

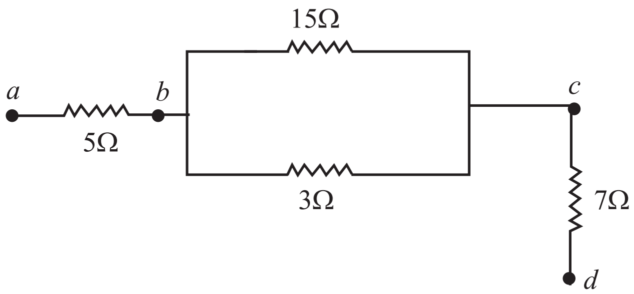


Fig. 17.12: A combination of series and parallel groupings

Example 17.7 : Refer to the network shown in Fig. 17.13. Calculate the equivalent resistance between the points (i) *b* and *c* (ii) *c* and *d*, and (iii) *a* and *e*.

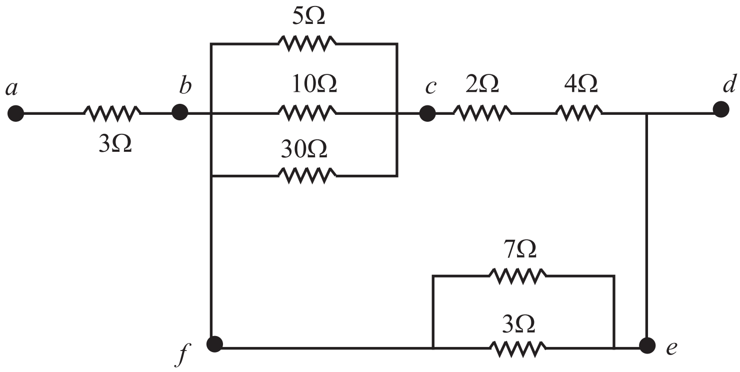


Fig. 17.13 : A combination of series and parallel groupings of resistors

Solution :

(i) Three resistors (5Ω, 10Ω and 30Ω) are connected in parallel. Therefore, equivalent resistance is given by

$$\frac{1}{R_1} = \frac{1}{5}+\frac{1}{10}+\frac{1}{30}=\frac{6+3+1}{30}=\frac{10}{30}\Omega$$

or $$R_1 = 3\ \Omega$$

(ii) The resistors with resistances 2Ω and 4Ω are in series. The equivalent resistance

$$R_2 = (2 + 4) = 6\Omega$$

(iii) The resistances 7Ω and 3Ω are in parallel. So equivalent resistance

$$\frac{1}{R_3} = \left(\frac{1}{7}+\frac{1}{3}\right)=\frac{3+7}{21}=\frac{10}{21}$$

or, $$R_3 = \frac{21}{10}\Omega = 2.1\ \Omega$$

Now we can treat equivalent resistance R_1 and R_2 to be in series. Therefore

$$R_4 = R_1 + R_2 = (3 + 6) = 9\Omega$$

Now R_4 and R_3 are in parallel. Therefore equivalent resistance

$$\frac{1}{R_5} = \frac{1}{R_4}+\frac{1}{R_3}$$

$$=\frac{1}{9}+\frac{1}{2.1}$$

$$=\frac{1}{9}+\frac{10}{21} = \frac{37}{63}$$

$$R_5 = \frac{63}{57}\Omega = 1.70\ \Omega$$

(iv) Finally R_5 and 3Ω (between *a* and *b*) are in series. Hence

$$R = (1.70 + 3) = 4.79\ \Omega$$

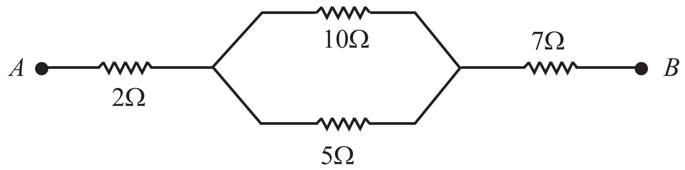
Note : For ease and convenience, you should draw a new equivalent circuit after every calculation.

INTEXT QUESTIONS 17.2

1. There are two bulbs and a fan in your bed room. Are these connected in series or in parallel? Why?

2. The electric supply in a town is usually at 220 V. Sometimes the voltage shoots upto 300 V and may harm your T V set and other gadgets. What simple precaution can be taken to save your appliances?

3. Calculate the equivalent resistance between points *A* and *B* for the following circuit :



17.4 TYPES OF RESISTORS

We use resistors in all electrical and electronic circuits to control the magnitude of current. Resistors usually are of two types :

- carbon resistors
- wire wound resistors

In a wire wound resistor, a resistance wire (of manganin, constantan or nichrome) of definite length, which depends on the required value of resistance, is wound two-fold over an insulating cylinder to make it non-inductive. In carbon resistors, carbon with a suitable binding agent is molded into a cylinder. Wire leads are attached to the cylinder for making connections to electrical circuits. Resistors are colour coded to give their values :

$$R = AB \times 10^{C}\ \Omega,\ D$$

where *A*, *B* and *C* are coloured stripes. The values of different colours are given in Table 17.1. As may be noted,

– first two colours indicate the first two digits of the resistance value;

– third colour gives the power of ten for the multiplier of the value of the resistance; and

– fourth colour (the last one) gives the tolerance of the resistance, which is 5% for golden colour, 10% for silver colour and 20% for body colour.

Table 17.1 : Colour codes of resistors

Colour	Number	Multiplier
Black	0	1
Brown	1	10^1
Red	2	10^2
Orange	3	10^3

Yellow	4	10^4
Green	5	10^5
Blue	6	10^6
Violet	7	10^7
Grey	8	10^8
White	9	10^9

Suppose that four colours on a resistor are Blue, Grey, Green and Silver. Then

The first digit will be 6 (blue)

The second digit will be 8 (Grey)

The third colour signifies multiplier 10^5 (Green)

The fourth colour defines tolerance = 10% (Silver)

Hence value of the resistance is

$$68 \times 10^5 \pm 10\%$$
$$= 68 \times 10^5 \pm (68 \times 10^5 \times 10/100)$$
$$= 68 \times 10^5 \pm 68 \times 10^4$$
$$= (6.8 \pm 0.68)\ \text{M}\Omega$$

17.5 TEMPERATURE DEPENDENCE OF RESISTANCE

The resistivity of a conductor depends on temperature. For most metals, the resistivity increases with temperature and the change is linear over a limited range of temperature :

$$\rho = \rho_0 [1 + \alpha (T - T_0)] \qquad (17.14)$$

where ρ and ρ_0 are the resistivities at temperatures T and T_0, respectively. The temperatures are taken in °C and T_0 is the reference temperature. α is called the temperature co-efficient of resistivity. Its unit is per degree celcius.

Superconductors

Temperature dependence of resistivity led scientists to study the behaviour of materials at very low temperatures. They observed that certain metals and their alloys lost their resistivity completely below a certain temperature, called **transition temperature,** which is specific to the material. In such materials, current, once set up, remained, unchanged for ever without the use of an external source to maintain it. Such materials were termed as **superconductors.**

It was soon realised that superconductors, if they may exist near room temperature, will bring in revolutionary changes in technology. (These have been termed as *high temperature superconductors*.) For example, energy efficient powerful electromagnets made of superconducting coils may levitate vehicles above a magnetic track and make a high speed transportation system possible.

Efforts are being made to develop high temperature superconductors. The work done so far suggests that oxides of copper, barium and ytterium are showing good possibilities. A superconductor ($T_2\ Ba_2\ Ca_2\ Cu_3O_{10}$) which can exist at –153°C has been developed. India is a front runner in this area of research.

Eqn. (17.14) can be rearranged to obtain an expression for temperature coefficient of resistivity :

$$\rho = \rho_0 + \rho_0\, \alpha\, (T - T_0)$$

or

$$\alpha = \frac{(\rho - \rho_0)}{\rho_0\, (T - T_0)} = \frac{1}{\rho_0}\frac{\Delta\rho}{\Delta T}$$

where $\Delta\rho = (\rho - \rho_0)$ and $\Delta T = T - T_0$.

The resistivity versus temperature graph for a metal like copper is shown in Fig. 17.14(a). The curve is linear over a wide range of temperatures.

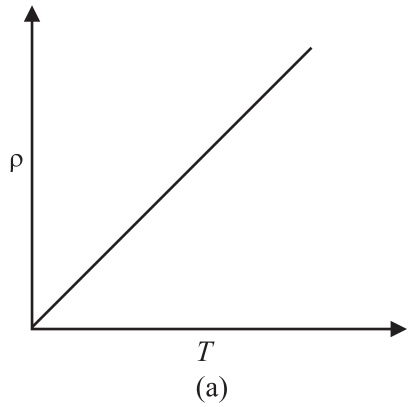


Fig. 17.14 : Typical resistivity–temperature graph for a metal

You may recall that resistance of a conductor is proportional to its resistivity. Therefore, temperature variation of resistance can written as :

$$R = R_0\, [1 + \alpha\, (T - T_0)] \qquad (17.15)$$

The resistances corresponding to two different temperatures T_1 and T_2 are given by

$$R_1 = R_0\, [1 + \alpha\, (T_1 - T_0)] \qquad (17.16)$$

and

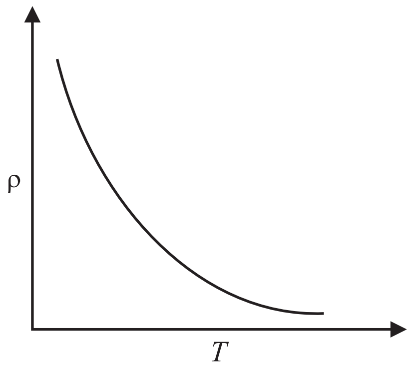
$$R_2 = R_0\, [1 + \alpha\, (T_2 - T_0)] \qquad (17.17)$$

On combining these equations, we can write an expression for temperature coefficient ofresistivity :

$$\alpha = \frac{(R_2 - R_1)}{R_0 (T_2 - T_1)} = \frac{1}{R_0}\frac{\Delta R}{\Delta T} \qquad (17.18)$$

If $R_0 = 1\Omega$ and $(T_2 - T_1) = 1°C$, then $\alpha = (R_2 - R_1)$. Thus **temperature coefficient of resistance is numerically equal to the change in resistance of a wire of resistance 1Ω at 0° C when the temperature changes by 1° C. This property of metals is used in making resistance thermometers.**

The resistivity of alloys also increases with increase in temperature. But the increase is very small compared to that for metals. For alloys such as **manganin, constantan** and **nichrome,** the temperature coefficient of resistivity is vanishingly small ($\sim 10^{-6}\,°C^{-1}$)and resistivity is high. That is why these materials are used for making resistance wires or standard resistances.



(b)

Fig. 17.14(b) : Resistivity of semiconductors decreases with temperature

Semiconductors such as germanium and silicon have resistivities which lie between those of metals and insulators. The resistivity of semiconductors usually decreases with increase in temperature [Fig.17.14(b)]. This gives a negative temperature coefficient of resistance. This will be discussed in detail in the lesson on semiconductors.

17.6 ELECTROMOTIVE FORCE (EMF) AND POTENTIAL DIFFERENCE

EMF is the short form of electromotive force. EMF of a cell or battery equals the potential difference between its terminals when these are not connected (open circuit) externally. You may easily understand the difference between e.m.f. and potential difference of a cell by performing the following activity.

ACTIVITY 17.4

Connect a cell in a circuit having a resistor *R* and key *K*. A voltmeter of very high resistance is connected in parallel to the cell, as shown in Fig.17.15. When key *K* is closed, voltmeter reading will decrease. Can you give reasons for this decrease in the voltmeter reading? Actually when key *K* is open, no current flows through the loop having cell and voltmeter: (The resistance in the circuit is infinite.) Hence the voltmeter reading gives e.m.f. *E* of the cell, which is the potential difference between the terminals of the cell when no current is drawn from it. When key *K* is closed, current flows outside and inside the cell. The cell

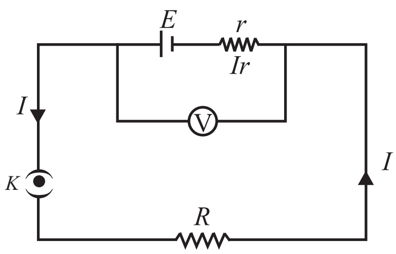


Fig. 17. 15

introduces a resistance r, called **internal resistance** of the cell. Let current I be flowing in the circuit. Potential drop Ir across internal resistance r due to current flow acts opposite to the e.m.f. of the cell. Hence, the voltmeter reading will be

$$E - Ir = V$$

or

$$E = V + Ir \qquad \{17.19)$$

Thus while drawing current from a cell, e.m.f. of the cell is always greater than the potential difference across external resistance, unless internal resistance is zero.

E.M.F. of a cell depends on :

- the electrolyte used in the cell;
- the material of the electrodes; and
- the temperature of the cell.

Note that the e.m.f. of a cell does not depend on the size of the cell, i.e. on the area of plates and distance between them. This means that if you have two cells of different sizes, one big and one small, the e.m.f.s can be the same if the material of electrods and electrolyte are the same. However, cells of larger size will offer higher resistance to the passage of current through it but can be used for a longer time.

Example 17.8 : When the current drawn from a battery is 0.5A, potential difference at the terminals is 20V. And when current drawn from it is 2.0A, its voltage reduces to 16V. Calculate the e.m.f. and internal resistance of the battery.

Solution : Let E and r be the e.m.f. and internal resistance of battery. When current I is drawn from it, the potential drop across internal resistance of the cell is Ir. Then we can write

$$V = E - Ir$$

For $I = 0.5$A and $V = 20$ volt, we have

$$20 = E - 0.5\, r \qquad \text{(i)}$$

For $I = 2.0$A and $V = 16$ volt, we can write

$$16 = E - 2r \qquad \text{(ii)}$$

We can rewrite Eqns. (i) and (ii) as

$$2E - r = 40$$

and

$$E - 2r = 16$$

Solving these,we get

$$E = 21.3 \text{ V and } r = 2.67\Omega$$

17.6.1 Elementary Idea of Primary and Secondary cells

We have seen that to pass electric current through a conductor continuously we have to maintain a potential difference between its ends. For the purpose, generally, we use a device called chemical cell.

Chemical calls are of two types :

(i) **Primary Cells :** In these cells, the chemical energy is directly converted into electrical energy. The material of a primary cell is consumed as we use the cell and, therefore, it cannot be recharged and reused. Dry cell, Daniel Cell, Voltaic Cell etc are examples of primary cells.

(ii) **Secondary Cells :** These are chemical cells in which electrical energy is stored as a reversible chemical reaction. When current is drawn from the cells the chemical reaction runs in the reverse direction and the original substances are obtained. These cells, therefore, can be charged again and again. Acid-accumulator, the type of battery we use in our inverter or car, is a set of secondary cells.

17.7 KIRCHHOFF'S RULES

You now know that Ohm's law gives current–voltage relation for resistive circuits. But when the circuit is complicated, it is difficult to know current distribution by Ohm's law. In 1842, Kirchhoff formulated two rules which enable us to know the distribution of current in complicated electrical circuits or electrical networks.

Gustav Robert Kirchhoff
(1824-1887)

The fundamental contributions of German physicist Kirchhoff were in the fields of black body radiation and spectroscopy. But he also contributed in many other fields. His rules that you will study in this lesson enable us to analyse complex electric networks. With the help of Bunsen spectrum analysis, he discovered elements Rubidium and Cesium.

(i) **Kirchhoff's First Rule (Junction Rule) :** It states that *the sum of all currents directed towards a junction (point) in an electrical network is equal to the sum of all the currents directed away from the junction.*

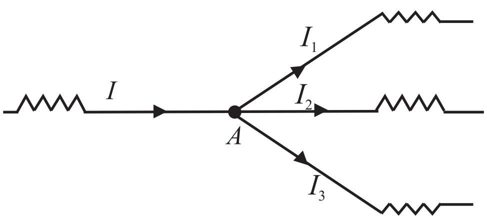


Fig. 17.16 : Kirchhoff's first rule : Sum of currents coming to a junction is equal to the sum of currents going away from it.

Refer to Fig. 17.16.If we take currents approaching point A as positive and those leaving it as negative, then we can write

$$I = I_1 + I_2 + I_3$$

or $$I - (I_1 + I_2 + I_3) = 0 \qquad (17.20)$$

In other words, the algebraic sum of all currents at a junction is zero.

Kirchhoff's first rule tells us that there is no accumulation of charge at any point if steady current flows in it. The net charge coming towards a point should be equal to that going away from it in the same time. In a way, it is an extension of continuity theorem in electrical circuits.

(ii) **Kirchhoff's Second Rule (Loop Rule)** : This rule is an application of law of conservation of energy for electrical circuits. It tells us that *the algebraic sum of the products of the currents and resistances in any closed loop of an electrical network is equal to the algebraic sum of electromotive forces acting in the loop.*

While using this rule, we start from a point on the loop and go along the loop either clockwise or anticlockwise to reach the same point again. The product of current and resistance is taken as positive when we traverse in the direction of current. The e.m.f is taken positive when we traverse from negative to positive electrode through the cell. Mathematically, we can write

$$\sum IR = \sum E \qquad (17.21)$$

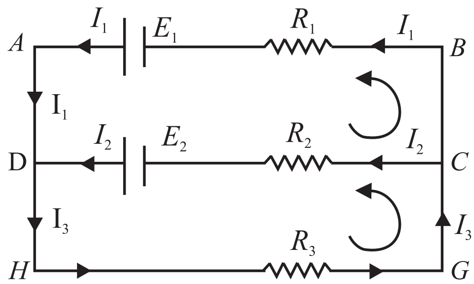


Fig. 17.17 : A network to illustrate Kirchhoff's second rule

Let us consider the electrical network shown in Fig. 17.17. For closed mesh *ADCBA*, we can write

$$I_1R_1 - I_2R_2 = E_1 - E_2$$

Similarly, for the mesh *DHGCD*

$$I_2R_2 + (I_1 + I_2)\, R_3 = E_2$$

And for mesh *AHGBA*

$$I_1R_1 + I_3\, R_3 = E_1$$

At point D $\qquad I_1 + I_2 = I_3$

In more general form, Kirchhoff's second rule is stated as : *The algebraic sum of all the potential differences along a closed loop in a circuit is zero.*

Example 17.9 : Consider the network shown in Fig. 17.18. Current is supplied to the network by two batteries. Calculate the values of currents I_1, I_2 and I_3. The directions of the currents are as indicated by the arrows.

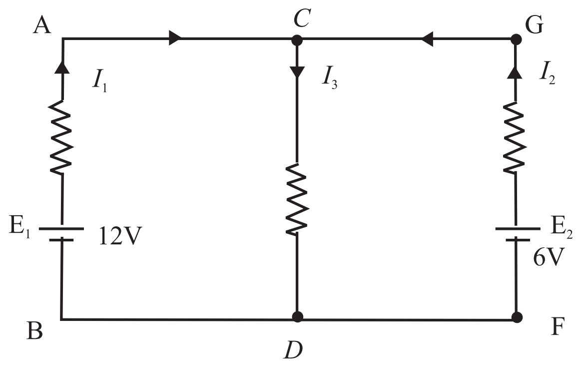


Fig. 17.18 : Calculation of currents in a network of resistors and batteries.

Solution : Applying Kirchhoff's first rule to junction *C*, we get

$$I_1 + I_2 - I_3 = 0 \quad \text{(i)}$$

Applying Kirchhoff's second rule to the closed loops *ACDBA* and *GCDFG*, we get

$$5I_1 + 2I_3 = 12 \quad \text{(ii)}$$

and $$3I_2 + 2I_3 = 6 \quad \text{(iii)}$$

On combining these equations, we get

$$5I_1 - 3I_2 = 6 \quad \text{(iv)}$$

Multiply (i) by 2 and add to (ii) to obtain

$$7I_1 + 2I_2 = 12 \quad \text{(v)}$$

On multiplying Eqn. (iv) by 2 and Eqn. (v) by 3 and adding them, we get

$$31I_1 = 48$$

or $$I_1 = 1.548\text{A}$$

Putting this value of I_1 in eqn. (v), we get

$$I_2 = 0.582\text{A}$$

And from (i), we get

$$I_3 = I_1 + I_2 = 2.13\text{A}$$

17.7.1 Wheatstone Bridge

You have learnt that a resistance can be measured by Ohm's law using a voltmeter and an ammeter in an electrical circuit. But this measurement may not be accurate for low resistances. To overcome this difficulty, we use a wheatstone bridge. It is an arrangement of four resistances which can be used to measure one of them in terms of the other three.

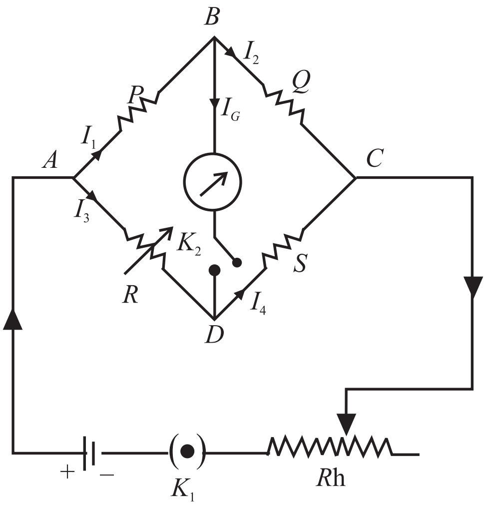


Fig. 17.19 : A wheatstone bridge.

Consider the circuit shown in Fig. 17.19 where

(i) *P* and *Q* are two adjustable resistances connected in arms *AB* and *BC*.

(ii) *R* is an adjustable known resistance.

(iii) *S* is an unknown resistance to be measured.

(iv) A sensitive galvanometer *G* along with a key K_2 is connected in the arm *BD*.

(v) A battery *E* along with a key K_1 is connected in the arm *AC*.

On closing the keys, in general, some current will flow through the galvanometer and you will see a deflection in the galvanometer. It indicates that there is some potential difference between points *B* and *D*. We now consider the following three possibilities:

(i) **Point *B* is at a higher potential than point *D* :** Current will flow from *B* towards *D* and the galvanometer will show a deflection in one direction, say right

(ii) **Point *B* is at a lower potential than point *D* :** Current will flow from point *D* towards *B* and the galvanometer will show a deflection in the opposite direction.

(iii) **Both points *B* and *D* are at the same potential:** In this case, no current will flow through the galvanometer and it will show no deflection, i.e. the galvanometer is in null condition. In this condition, the Wheatstone bridge is said to be in the *state* of *balance.*

The points B and D will be at the same potential only when the potential drop across P is equal to that across R. Thus

$$I_1 P = I_3 R \qquad (1\ 7.22)$$

But

$$I_1 = I_2 + I_G$$

and

$$I_4 = I_3 + I_G \qquad (17.23)$$

Applying Kirchhoff's first rule at junctions B and D in the null condition ($I_G = 0$), we get

$$I_1 = I_2$$

and

$$I_3 = I_4 \qquad (1\ 7.24)$$

Also potential drop across Q will be equal to that across S. Hence

$$I_2 Q = I_4 S \qquad (17.25)$$

Dividing Eqn. (17.22) by Eqn. (17.25), we obtain

$$\frac{I_1 P}{I_2 Q} = \frac{I_3 R}{I_4 S} \qquad (17.26)$$

Using Eqn. (17.24), we get

$$\frac{P}{Q} = \frac{R}{S} \qquad (17.27)$$

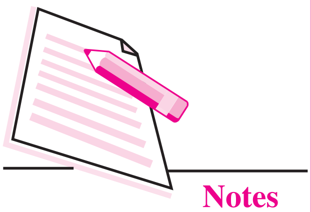
This is the condition for which a Wheatstone bridge will be balanced. From Eqn. (17.27), we find that the unknown resistance S is given by

$$S = \frac{QR}{P}$$

You can easily see that measurement of resistance by Wheatstone bridge method has the following merits.

(i) The balance condition given by Eqn. (17.27) *at null position is independent* of *the applied voltage V. In other words, even if you change the e.m.f* of *the cell, the balance condition will not change.*

(ii) The measurement of *resistance does not depend on the accuracy of calibration* of *the galvanometer. Galvanometer is used only* as a *null indicator (current detector).*



Notes

The main factor affecting the accuracy of measurement by Wheatstone bridge is its sensitivity with which the changes in the null condition can be detected. It has been found that the bridge has the greatest sensitivity when the resistances in all the arms are nearly equal.

Example 17.9: Calculate the value of R shown in Fig.17.20. when there is no current in 50Ω resistor.

Solution: This is Wheatstone bridge where galvanometer has been replaced by 50Ω resistor. The bridge is balanced because there is no current in 50Ω resistor. Hence,

$$\frac{20}{10} = \frac{40}{R}$$

or $$R = \frac{40 \times 10}{20} = 20\Omega$$

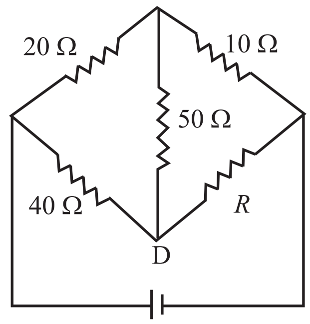
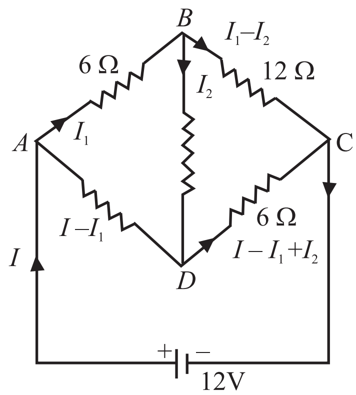


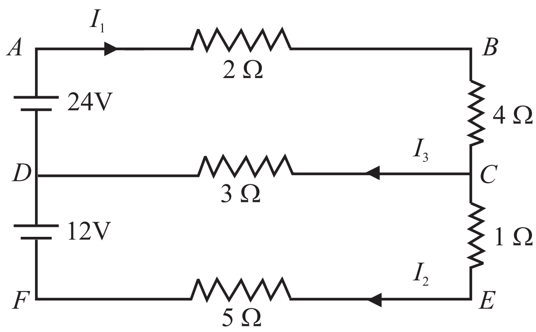
Fig. 17.20 : When there is no current through 50Ω resistor, the bridge is balanced.

INTEXT QUESTIONS 17.3

1. Refere to figure below. Calculate the value of currents in the arms AB, AD and BD.



2. Examine the following circuit containing resistors and batteries. Calculate the current I_1, I_2 and I_3.



17.8 POTENTIOMETER

You now know how to measure e.m.f. of a source or potential difference across a circuit element using a voltmeter. (An ideal voltmeter should have infinite resistance so that it does not draw any current when connected across a source of e.m.f.) Practically it is not possible to manufacture a voltmeter which will not draw any current. To overcome this difficulty, we use a potentiometer, which draws no current from it. It employs a null method. The potentiometer can also be used for measurement of internal resistance of a cell, the current flowing in a circuit and comparison of resistances.

17.8.1 Description of a Potentiometer

A potentiometer consists of a wooden board on which a number of resistance wires (usually ten) of uniform cross-sectional area are stretched parallel to each other. The wire is of maganin or nichrome. These wires are joined in series by thick copper strips. In this way, these wires together act as a single wire of length equal to the sum of the lengths of all the wires. The end terminals of the wires are provided with connecting screws.

A metre scale is fixed on the wooden board parallel to wires. A jockey (a sliding contact maker) is provided with the arrangement. It makes a knife edge contact at any desired point on a wire. Jockey has a pointer which moves over the scale. It determines the position of the knife edge contact. In Fig. 17.21 a ten wire potentiometer is shown. *A* and *B* are ends of the wire. *K* is a jockey and *S* is a scale. Jockey slides over a rod *CD*.

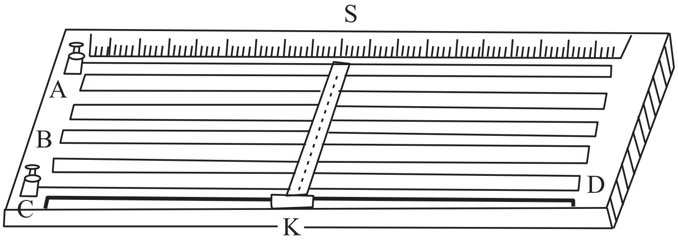


Fig. 17.21 : An illustrative diagram of a potentiometer

17.8.2 Measurements with a Potentiometer

Let us suppose that a steady source of e.m.f. *E* (say an accumulator) is connected across a uniform wire *AB* of length *l*. Positive terminal of the accumulator is connected at end *A* (Fig.17.22). A steady current *I* flows through the wire. The potential difference across *AB* is given by

$$V_{AB} = RI$$

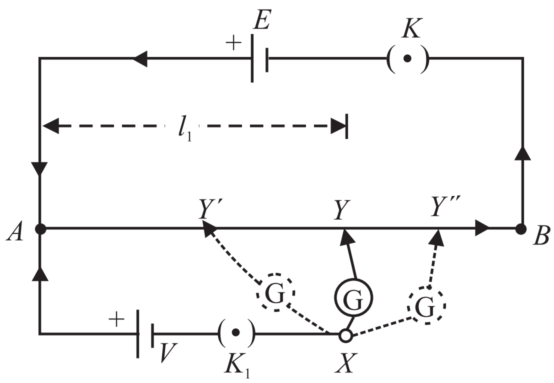


Fig. 17.22 : Potentiometer circuit to measure potential difference between the terminals of a cell.

If r is the resistance per unit length of the wire, and k is the potential drop across unit length of the wire, then

$$R = r\ell$$

and $$E = k\,\ell$$

or $$k = \frac{E}{\ell}$$

For length ℓ_1 of wire, potential drop is given by

$$V_1 = k\ell_1 = \frac{E}{\ell}\ell_1 \qquad (17.28)$$

Thus potential falls linearly with distance along the wire from the positive to the negative end.

We wish to measure an unknown voltage V. The positive terminal of the cell is connected to end A of the wire and negative terminal through a galvanometer to the jockey having variable contact Y. Note that for $V > E$, it will not be possible to obtain a null point. So we use a standard cell of emf E ($> V$), as shown in Fig.17.22. To check this, insert keys K and K_1 and tap at ends A and B. The galvanometer should show deflection in opposite directions. If so, all is well with the circuit.

Insert key K_1 and start moving jockey from A towards B. Suppose that at position Y' potential drop across the length AY of the wire is less than voltage V. The current in the loop $AY'XA$ due to voltage V exceeds the current due to potential difference across AY'. Hence galvanometer shows some deflection in one direction. Then jockey is moved away, say to Y'' such that potential drop across AY'' is greater than the voltage V. If galvanometer shows deflection in the other direction, the voltage drop across AY'' is greater than that across AY'. Therefore, the jockey is moved slowly between Y' and Y''. A stage is reached, say at point Y, where potential drop across AY is equal to voltage V. Then points X and Y will be at the same potential and hence the galvanometer will not show any deflection, i.e. null point is achieved. If l_1 is the length between A and Y, then

$$V = k\ell_1 = \frac{E\ell_1}{\ell} \qquad (17.29)$$

Thus, the unknown voltage V is measured when no current is drawn

The measurements with potentiometer have following advantages :

- When the potentiometer is balanced, no current is drawn from the circuit on which the measurement is being made.
- It produces no change in conditions in a circuit to which it is connected.
- It makes use of null method for the measurement and the galvanometer used need not be calibrated.

17.8.3 Comparison of E.M.Fs of two Cells

You have learnt to measure the e.m.f. of a cell using a potentiometer. We shall now extend the same technique for comparison of e.m.fs of two cells. Let us take, for example, a Daniel cell and a Leclanche cell and let E_1 and E_2 be their respective e.m.fs.

Refer to circuit diagram shown in Fig.17.23. The cell of e.m.f. E_1 is connected in the circuit through terminals 1 and 3 of key K_1. The balance point is obtained by moving the jockey on the potentiometer wire as explained earlier. Note that e.m.f of cell E should be greater than the emfs of E_1 and E_2 seperately. (Otherwise, balance point will not be obtained.) Let the balance point on potentiometer be at point Y_1 and length $AY_1 = l_1$. The cell of e.m.f. E_2 is connected in the circuit through terminals 2 and 3 of the key K_2. Suppose balance is obtained at point Y_2 and length $AY_2 = l_2$.

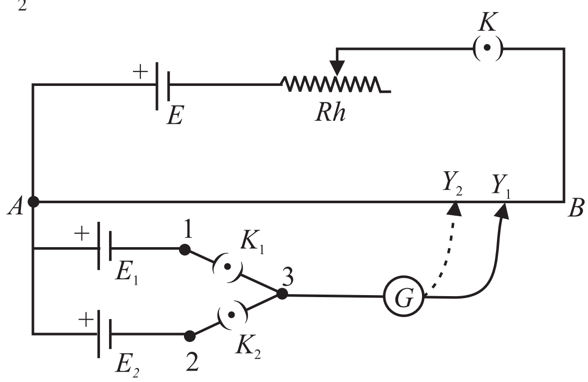


Fig. 17.23 : Circuit diagram for comparison of e.m.fs of two cells E_1 and E_2.

Applying potentiometer principle, we can write

$$E_1 = kl_1 \quad \text{and} \quad E_2 = kl_2$$

where k is the potential gradient along the wire AB. Hence

$$\frac{E_1}{E_2} = \frac{\ell_1}{\ell_2} \tag{17.30}$$

17.8.4 Determination of Internal Resistance of a Cell

You have learnt that cells always offer resistance to the flow of current through them, which is often very small. This resistance is called the internal resistance of the cell and depends on the size of the cell, i.e. the area of the plates immersed in the liquid, the distance between the plates and strength of electrolyte used in the cell.

Let us now learn how to measure internal resistance of a cell using a potentiometer. Refer to Fig. 17.24, which shows the circuit diagram for measuring internal resistance 'r' of a cell of emf E_1. A resistance box R with a key K_1 is connected in parallel with the cell. The primary circuit has a standard cell, a rheostat and a one way key K. As soon as key K is closed, a current I begins to flow through the wire

AB. The key K_1 is kept open and on moving the jockey, a balance is obtained with the cell E_1 at point, say Y_1. Let $AY_1 = l_1$. Then we can write

$$E_1 = kl_1 \qquad (17.31a)$$

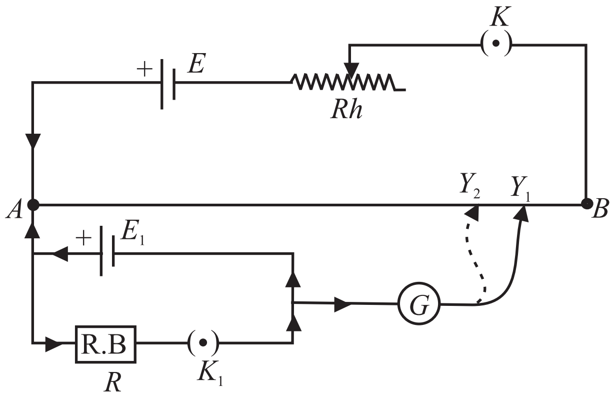


Fig. 17.24 : Measurement of the internal resistance r of a cell

Now key K_1 is closed. This introduces a resistance across the cell. A current, say I_1, flows in the loop $E_1RK_1E_1$ due to cell E. Using Ohm's law, we can write

$$I_1 = \frac{E_1}{R+r}$$

where r is internal resistance of the cell. It means that terminal potential difference V_1 of the cell will be less than E_1 by an amount I_1r. The value of V_1 is

$$V_1 = I_1R = \frac{E_1}{R+r}\ R$$

Then, potential difference V_1 is balanced on the potentiometer wire without change in current I. Let the balance point be at point Y_2 such that $AY_2 = l_2$. Then

$$V_1 = kl_2$$

(17.31b)

Using Eqns. (17.31a,b) we get

$$\frac{E_1}{V_1} = \frac{\ell_1}{\ell_2} = \frac{R+r}{R}$$

or

$$r = R\left(\frac{l_1}{l_2} - 1\right) \qquad (17.32)$$

Thus by knowing R_1, l_1 and l_2, the value of r can be easily calculated.

Example 17.10 : Length of a potentiometer wire is 5 m. It is connected with a battery of fixed e.m.f. Null point is obtained for the Daniel cell at 100 cm. If the length of the wire is kept 7 m, what will be the position of null point?

Solution: Let e.m.f. of battery be E volt. The potential gradient for 5 m length is

$$k_1 = \frac{E}{5} \text{ Vm}^{-1}$$

When the length of potentiometer wire is 7 m, potential gradient is

$$k_2 = \frac{E}{7} \text{ Vm}^{-1}$$

Now, if null point is obtained at length l_2, then

$$E_1 = k_2\ell_2 = \frac{E}{7}\ell_2$$

Here same cell is used in two arrangements. Hence

$$\frac{E}{5} = \frac{E}{7}\ell_2$$

$$\Rightarrow \qquad \ell_2 = 7/5 = 1.4\text{m}$$

17.9 DRIFT VELOCITY OF ELECTRONS

Let us now understand the microscopic picture of electrical conduction in a metal. The model presented here is simple but its strength lies in the fact that it conforms to Ohm's law.

We assume that a metallic solid consists of atoms arranged in a regular fashion. Each atom usually contributes free electrons, also called conduction electrons. These electrons are free to move in the metal in a random manner, almost the same way as atoms or molecules of a gas move about freely in the a container. It is for this reason that sometimes conduction electrons are referred to as **electron gas**. The average speed of conduction electrons is about 10^6ms^{-1}.

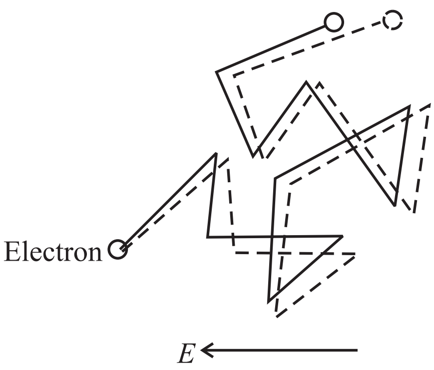


Fig. 17.25 : Motion of electrons in a conductor placed in an electric field.

We know that no current flows through a conductor in the absence of an electric field, because the **average velocity** of free electrons is zero. On an average, the number of electrons moving in $+x$ direction is same as number of electrons moving in $-x$ direction. There is no net flow of charge in any direction.

The conduction electrons frequently collide with the atoms in the solid. The free electrons drift slowly in a direction opposite to the direction of the applied electric field. The average drift velocity is of the order of 10^{-4}ms^{-1}. This is very small compared to the average speed of free electrons between two successive collisions (10^6ms^{-1}). On applying an electric field, the conduction electrons get accelerated.

The excess energy gained by the electrons is lost during collisions with the atoms. The atoms gain energy and vibrate more vigorously. The conductor gets heated up. Fig. 17.25 shows how the motion of electrons is modified when an electric field is applied is applied.

Let us now obtain an expression for the drift velocity of conduction electrons. Let e and m be the charge and mass respectively of an electron. If E is the electric field, the force on the electron is eE. Hence acceleration experienced by the electron is given by

$$\mathbf{a} = \frac{e\mathbf{E}}{m}$$

If τ is the average time between collisions, we can write the expression for velocity of drifting electrons in terms of electric field as

$$\boldsymbol{v}_{\mathrm{d}} = \frac{e\mathbf{E}}{m}\tau$$

On combining this result with Eqn. (17.4), we obtain the expression for current :

$$I = -\,neAv_d$$

$$= -neA\,\frac{eE}{m}\tau$$

$$= -\frac{Ane^2E}{m}\tau$$

Since electric field is negative spatial gradient of potential $\left(E = -\frac{\partial V}{\partial r}\right)$ we can rewrite the expression for current as

$$I = +\frac{ne^2A}{m}\frac{V}{\ell}\tau \qquad (17.33)$$

$$\Rightarrow \qquad \frac{V}{I} = \frac{m}{ne^2\tau}\frac{\ell}{A} = R \qquad (17.34)$$

Eqn. (17.34) implies that conduction current obeys Ohm's law.

On combining this result with Eqn. (17.9), we get

$$\rho = \frac{1}{\sigma} = \frac{m}{ne^2\tau} \qquad (17.35)$$

17.10 POWER CONSUMED IN AN ELECTRICAL CIRCUIT

Let us examine the circuit in Fig. 17.26 where a battery is connected to an external resistor R . The positive charges (so to say) flow in the direction of the current in the resistor and from negative to positive terminal inside the battery. The potential difference between two points gives kinetic energy to the charges. These moving charges collide with the atoms (ions) in the resistor and thus lose a part of their kinetic energy. This energy increases with the temperature of the resistor. The loss of energy by moving charges is made up at the expense of chemical energy of the battery.

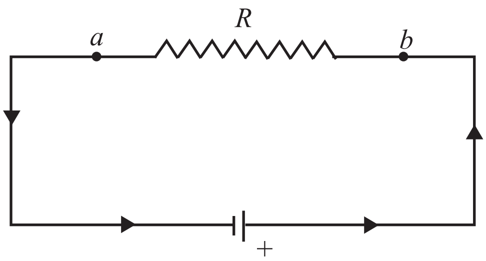


Fig. 17.26 : A circuit containing a battery and a resistor. The power consumed depends on the potential difference between the points *a* and *b*, the current through the resistor.

The rate of loss of potential energy by moving charge ΔQ in going through the resistor is

$$\frac{\Delta U}{\Delta t} = V\frac{\Delta Q}{\Delta t} = VI \qquad (17.\ 36)$$

where I is the current in the circuit and V is potential difference between the ends of the resistor.

It is assumed that the resistance of the connecting wires is negligible. The total loss is in the resistor R only. Rate of loss of energy is defined as power :

$$P = VI$$

Since $V = IR$, we can write

$$P = I^2R = V^2/R \qquad (17.37)$$

The SI unit of power is watt (W).

The electrical power lost in a conductor as heat is called *joule heat*. The heat produced is proportional to : (i) square of current (I), (ii) resistance of conductor (R), and (iii) time for which current is passed (t).

The statement $Q = I^2 Rt$, is called Joule's law for heating effect of current.

Example: 17.11 : A 60W lamp is connected to 220V electricity supply in your home. Calculate the power consumed by it, the resistance of its filament and the current through it.

Solution : We know that $I = P/V$

$$\therefore \qquad = \frac{60\text{W}}{220\text{V}} = \frac{3}{11}\text{A} = 0.27\text{A}$$

Resistance of the lamp

$$R = \frac{V}{I}$$

$$= \frac{220V}{3/11A}$$

$$= \frac{220 \times 11}{3}\Omega = 807\Omega$$

The lamp consumes 60J of energy per second. It will consume 60 Wh energy in one hour and $60 \times 24 = 1440$ Wh energy in one day.

Energy consumed per day = 1.440 kWh

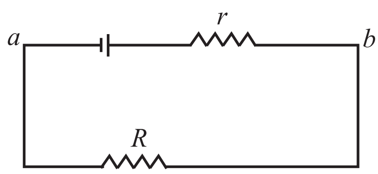
In common man's language, it is known as 1.4 unit of energy.

INTEXT QUESTIONS 17.4

1. When current drawn from a cell increases, the potential difference between the cell electrodes decreases. Why?

2. A metallic wire has a resistance of 30Ω at 20° C and 30.16Ω at 40° C. Calculate the temperature coefficient of resistance.

3. The e.m.f of a cell is 5.0 V and R in the circuit is 4.5Ω. If the potential difference between the points a and b is 3.0 V, calculate the internal resistance r of the cell.



4. In a potentiometer circuit, balance point is obtained at 45 cm from end A when an unknown e.m.f is measured. The balance point shifts to 30 cm from this end when a cell of 1.02 V is put in the circuit. Standard cell E always supplies a constant current. Calculate the value of unknown e.m.f.

5. A potentiometer circuit is used to compare the e m.f. of two cells E_1 and E_2. The balance point is obtained at lengths 30 cm and 45 cm, respectively for E_1 and E_2. What is the e.m.f of E_1, if E_2 is 3.0 V?

6. A current of 0.30 A flows through a resistance of 500Ω. How much power is lost in the resistor?

7. You have two electric lamps. The printed specifications on them are 40W, 220V and 100W, 220 V. Calculate the current and resistance of each lamp when put in a circuit of 220 V supply line.

WHAT YOU HAVE LEARNT

- Drift velocity is the average velocity with which electrons move opposite to the field when an electric field exists in a conductor.
- Electric current through any cross-sectional area is the rate of transfer of charge from one side to other side of the area. Unit of current is ampere and is denoted by A.
- Ohm's law states that the current flowing through a conductor is proportional to the potential difference when physical conditions like pressure and temperature remain unchanged.
- Ratio V/I is called resistance and is denoted by R. Unit of resistance is ohm (denoted by Ω)
- Resistivity (or specific resistance) of a material equals the resistance of a wire of the material of one metre length and one m^2 area of cross section. Unit of resistivity is ohm metre.
- For a series combination of resistors, the equivalent resistance is sum of resistances of all resistors.
- For a parallel combination of resistors, inverse of equivalent resistance is equal to the sum of inverses of all the resistances.
- Primary cells cannot be recharged and reused, whereas, secondary cells can be charged again and again.
- Kirchhoff's rules help us to study systematically the complicated electrical circuits. The first rule states that the sum of all the currents directed towards a point in an electrical network is equal to the sum of all currents directed away from the point. Rule II : The algebraic sum of all potential differences along a closed loop in an electrical network is zero.
- The Wheatstone bridge circuit is used to measure accurately an unknown resistance (S) by comparing it with known resistances (P, Q and R). In the balance condition,

  $P/Q = R/S$.
- The e.m.f. of a cell is equal to the potential difference between its terminals when a circuit is not connected to it.
- A potentiometer measures voltages without drawing current. Therefore, it can be used to measure e.m.f. of a source that has appreciable internal resistance.
- Drift velocity of electrons in a conductor is given by $v_d = -\frac{eE}{m}\tau$.
- Power consumed in an electrical circuit through Joule heating is given by

$$p = VI = I^2R = \frac{V^2}{R}.$$

TERMINAL EXERCISES

1. Explain the drift of free electrons in a metallic conductor under external electric field. Derive an expression for drift velocity.

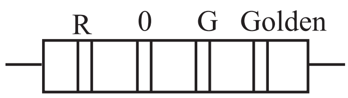
2. Define electric current and discuss Ohm's law.

3. Define resistivity of a conductor. How does the resistance of a wire depend on the resistivity of its material, its length and area of cross-section?

4. Define electrical conductivity. Write its unit. How does electrical conductivity depend on free electron concentration of the conductor?

5. Explain the difference between ohmic and non-ohmic resistances. Give some examples of non-ohmic resistances.

6. What is the effect of temperature on the resistivity of a material? Why does electrical conductivity of a metal decrease with increase in temperature?



7. The colours on the resistor shown here are red, orange, green and gold as read from left to right. How much is the resistance according to colour code?

8. Three resistors of resistances R_1, R_2 and R_3 are connected (i) in series, and (ii) in parallel. Calculate the equivalent resistance of combination in each case.

9. What is the difference between emf and potential difference between the electrodes of a cell. Derive relation between the two.

10. Expalin the difference between primary cells and secondary cells.

11. State Kirchhoff's rules governing the currents and electromotive forces in an electrical network?

12. Give theory of Wheatsone's bridge method for measuring resistances.

13. Discuss the theory of potentiometer.

14. How will you measure unknown potential difference with the help of a potentiometer?

15. Describe potentiometer method of comparing e.m.f. of two cells.

16. How will you determine internal resistance of a cell with the help of a potentiometer? What factors are responsible for internal resistance of a cell ?

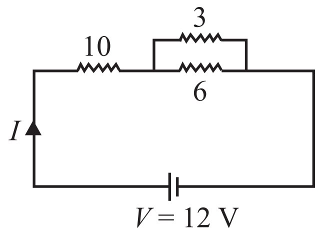
17. A wire of length 1 m and radius 0.1 mm has a resistance of 100Ω. Calculate the resistivity of the material.

18. Consider a wire of length 4m and cross-sectional area 1mm^2 carrying a current of 2A. If each cubic meter of the material contains 10^{29} free electrons, calculate the average time taken by an electron to cross the length of the wire.

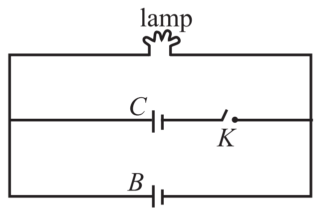
19. Suppose you have three resistors, each of value 30Ω. List all the different resistances that you can obtain by combining them.

20. The potential difference between the terminals of a battery of e.m.f. 6.0V and internal resistance 1Ω drops to 5.8V when connected across an external resistor. Find the resistance of the external resistor.

21. For the circuit shown here, calculate the value of current *I* and equivalent resistance *R*.



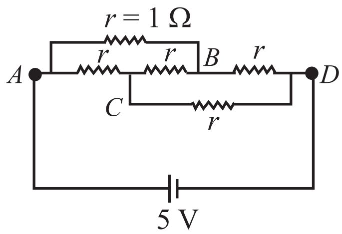
22. Examine the following network containing a lamp, a capacitor and a battery. The lamp is lighted when connected directly to the battery. What happens to it in this circuit when the switch is closed.



23. The following Wheatstone's bridge is balanced. Calculate

    (a) the value of equivalent resistance *R* in the circuit, and

    (b) the current in the arms *AB* and *DC*.



ANSWERS TO INTEXT QUESTIONS

17.1

1. (a) The current reduces to half as resistance of the wire is doubled.

   (b) The current is doubled as resistance is halved.

2. Resistivity is a property of the material of wire. It will not change with change in length and area of cross-section.

   $\rho = 2 \times 10^{-8}\ \Omega\text{m}$

3. $R = \dfrac{V}{I} = \dfrac{8}{0.15} = \dfrac{800}{15} = 53.3\ \Omega$

   $R = \dfrac{P\ell}{A} \Rightarrow \dfrac{800}{15} = \rho \dfrac{3}{2\times10^{-4}} \Rightarrow \rho = \dfrac{800\times2\times10^{-4}}{15\times3} = 35.5 \times 10^{-4}\ \Omega\text{m}.$

4. No. Only metallic conductor obey Ohm's law upto a certain limit. Semiconductors and electrolytes do not obey Ohm's law.

5. $I = \frac{q}{\tau} = \frac{n|e|}{t} = \frac{5\times10^{17}\times1.6\times10^{-19}}{1}\text{A} = 0.8\times10^{-3}\text{A} = 0.8 \text{ mA}$

The direction of current is opposite to the direction of flow of electrons, i.e., from right to left.

17.2

1. In parallel. They may draw different currents needed for their operation and are operated separately using different switches.
2. We use a voltage stablizer
3. $R = R_1 + \frac{R_2R_3}{R_2 + R_3} + R_4$

   $= 2 + \frac{10}{3} + 7$

   $= 12.3\ \Omega$

17.3

1. Applying Kirchhoff's second rule on loop *ABCDA,* we get

   $2I_1 + 4I_1 + 3I_3 = 24$

   $6I_1 + 3I_3 = 24 \quad ...(1) \quad \Rightarrow \quad 2I_1 + I_3 = 8 \quad ....(1)$

   Similarly, for loop *DCBFD,* we can write

   $-3I_3 + 6I_2 = 12 \quad \Rightarrow \quad 2I_2 - I_3 = 4 \quad ...(2)$

   Also applying Kirchoff's first rule at junction *D* we get

   $I_2 + I_3 = I_1$

   Substituting in (1) we get

   $2I_2 + 3I_3 = 8$

   $2I_2 - I_3 = 4$

   $4I_3 = 4$

   $I_3 = 1A$

   Substituting in (2)

   $2I_2 = 5 \Rightarrow I_2 = 2.5 \text{ A}$

2. $\frac{P}{Q} = \frac{6}{12} = \frac{1}{2}$ and $\frac{R}{S} = \frac{3}{6} = \frac{1}{2}$

$$\frac{P}{Q} = \frac{R}{S} \qquad \therefore \text{ bridge is balanced}$$

Hence $V_B = V_D$ and $I_2 = 0$

$$I_1 = \frac{V}{I} = \frac{12}{18} = \frac{2}{3}\text{A}$$

and

$$I - I_1 = \frac{12}{9} = \frac{4}{3}\text{A}$$

17.4

1. $V = E - Ir$ as I increases V decreases.

2. $R_{20} = R_0\,(1 + 20\,\alpha)$

$$R_{40} = R_0\,(1 + 40\alpha)$$

$$\frac{R_{40}}{R_{20}} = \frac{1 + 40\,\alpha}{1 + 20\,\alpha}$$

$$\frac{1 + 40\,\alpha}{1 + 20\,\alpha} = \frac{30.16}{30} = 1 + \frac{0.16}{30}$$

$$1 + \frac{20\,\alpha}{1 + 20\,\alpha} = 1 + \frac{0.16}{30}$$

$$\frac{20\,\alpha}{1 + 20\,\alpha} = \frac{0.16}{30}$$

On cross-multiplication, we get $600\alpha = 0.16 + 3.2\alpha$

$$\Rightarrow \quad \alpha \simeq \frac{0.16}{600} = 2.67 \times 10^{-4}\,\text{K}^{-1}$$

3. $I = \dfrac{V}{R} = \dfrac{3}{4.5} = \dfrac{30}{45} = \dfrac{2}{3}\text{A}$

$$V = \Sigma - Ir \;\Rightarrow\; 3 = 5 - \frac{2}{3}r$$

$$\therefore\; r = \frac{2 \times 3}{2} = 3\Omega$$

4. $\frac{E_2}{E_1} = \frac{\ell_2}{\ell_1} \Rightarrow \frac{1.02}{E_1} = \frac{30}{45} \Rightarrow E_1 = 0.51 \times \frac{3}{2} = 1.53\text{V}$

5. $\frac{E_2}{E_1} = \frac{\ell_2}{\ell_1}$

$\frac{E_1}{3} = \frac{2}{3}$

$E_1 = 2\ V$

6. $P = IV$

$= 3 \times 0.3 \times 500$

$= 45$ WaH.

7. $I = \frac{P}{V} \Rightarrow I_1 = \frac{40}{220} = \frac{2}{11}\text{A}$ and $I_2 = \frac{100}{220} = \frac{5}{11}\text{A}$

$R = \frac{V^2}{P} \Rightarrow I_1 = \frac{40}{220} = \frac{2}{11}\text{A}\ \frac{V^2}{P} \Rightarrow R_1 = \frac{220 \times 220}{40} = 1210\Omega$

and $R_2 = \frac{220 \times 220}{100} = 484\Omega$

Answers to Problems in Terminal Excercises

17. $3.14 \times 10^{-6}\ \Omega$m. 18. 32 ms.

19. (i) All resistance in series; equivalent resistance 90Ω

(ii) All resistances in parallel; equivalent resistance 10Ω

(iii) One resistance in series with two others which are connected in parallel; equivalent resistance 45Ω

(iv) Two resistances in series and one resistance in parallel to them; equivalent resistance 20Ω.

20. 29Ω 21. $I = 1\text{A}$, $R = 12\Omega$

23. (a) $R = \text{r} = 1\Omega$ (b) $I = 2.5\text{A}$

312en18

18

MAGNETISM AND MAGNETIC EFFECT OF ELECTRIC CURRENT

In lesson 15, you learnt how charged rods attract each other or small bits of paper. You might have also played with magnets – the substances having the property of attracting small bits of iron. But did you ever think of some relation between electricity and magnetism? Such a relationship was discovered by Oersted in 1820. Now we know, for sure, how intimately magnetism and electricity are related.

In this lesson, you will learn the behaviour of magnets and their uses as also the magnetic effects of electric current. The behaviour of current carrying conductors and moving charges in a magnetic field are also discussed. On the basis of these principles, we will discuss the working of electric devices like motors and measuring devices like an ammeter, a voltmeter and a galvanometer.

OBJECTIVES

After studying this lesson, you should be able to :

- *define magnetic field and state its SI unit;*
- *list the elements of earth's magnetic field and write the relation between them;*
- *describe the magnetic effect of electric current : Oersted's experiment;*
- *state Biot-Savart's law and explain its applications;*
- *explain Ampere's circuital law and its application;*
- *describe the motion of a charged particle in uniform electric field and magnetic field;*
- *explain the construction and working of a cyclotron;*

- *derive an expression for the force experienced by a current carrying conductor placed in a uniform magnetic field;*
- *derive an expression for the force between two infinitely long current carrying conductors placed parallel to each other; and*
- *explain the working principle of a galvanometer, an ammeter and a voltmeter.*

18.1 MAGNETS AND THEIR PROPERTIES

The phenomenon of magnetism was known to Greeks as early as 600 B.C. They observed that some stones called magnetite (Fe_3O_4) attracted iron pieces. The pieces of naturally occurring magnetite are called **natural magnets**. Natural magnets are weak, but materials like iron, nickel, cobalt may be converted into strong permanent magents. All magnets–natural or artificial – have same properties. You must be familiar with basic properties of magnets. However, for completeness, we recapitulate these.

(i) **Directive Property :** A small bar magnet, when suspended freely on its center of mass so as to rotate about a vertical axis, always stays in approximately geographical north-south direction.

(ii) **Attractive Property :** A magnet attracts small pieces of magnetic materials like iron, nickel and cobalt. The force of attraction is maximum at points near the ends of the magnet. These points are called **poles** of the magnet. In a freely suspended magnet, the pole which points towards the geographical north is called is **north pole** and the one which points towards the geographical south is called **south pole**. Do directive and attractive properties suggest that our earth also acts like a magnet? Yes, it does.

(iii) Unlike poles of two magnets attract each other and like poles repel (Fig.18.1).

(iv) The poles of a magnet are inseparable, i.e. the simplest specimen providing magnetic field is a magnetic dipole.

(v) When a magnet is brought close to a piece of iron, the nearer end of the piece of iron acquires opposite polarity and the farther end acquires same polarity. This phenomenon is called **magnetic induction**.

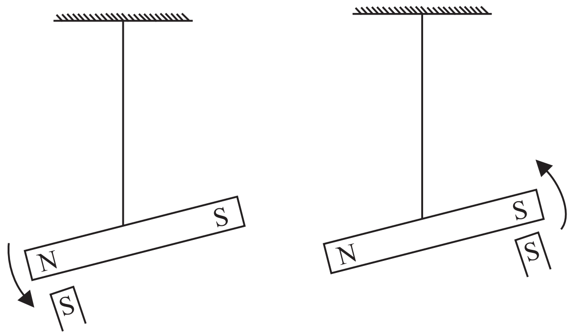


Fig. 18.1 : Unlike poles of two magnets attract each other and like poles repel.

18.1.1 Magnetic Field Lines

Interactions between magnets or a magnet and a piece of iron essentially represent action at a distance. This can be understood in terms of magnetic field. A very convenient method to visualize the direction and magnitude of a field is to draw the field lines :

- The direction of magnetic field vector **B** at any point is given by the tangent to the field line at that point.
- The number of field lines that pass through unit area of a surface held perpendicular to the lines is proportional to the strength of magnetic field in that region. Thus, the magnetic field **B** is large where the field lines are closer together and smaller where they are far apart.

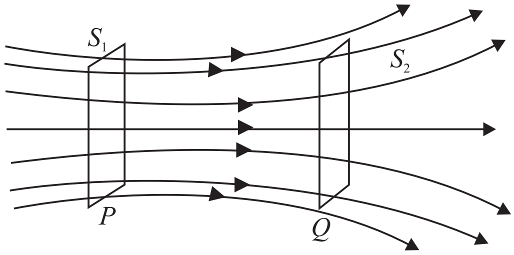


Fig 18.2: Magnetic field lines passing through two parallel surfaces

Fig 18.2 shows a certain number of field lines passing through parallel surfaces S_1 and S_2. The surface area of S_1 is same as that of S_2 but the number of field lines passing through S_1 is greater than those passing through S_2. Hence, the number of lines per unit area passing through S_1 is greater than that through S_2. We can, therefore, say that the magnetic field in the region around P is stronger than that around Q.

- Outside the magnet, the field lines run from north pole to south pole and inside it, these run from south pole to north pole forming closed curves (Fig. 18.3).
- Two magnetic field lines can never cross each other.

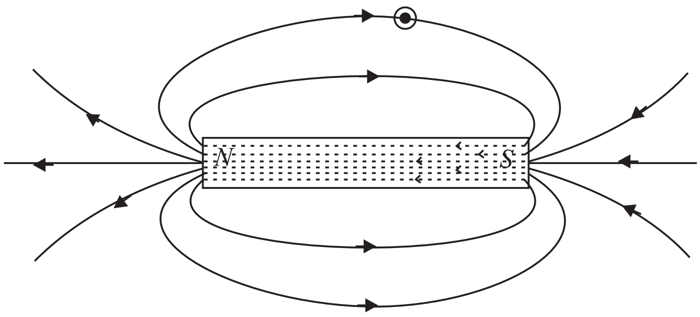


Fig. 18.3 : Magnetic field lines of a bar magnet

INTEXT QUESTIONS 18.1

1. You are given a magnet. How will you locate its north pole?

2. You are provided two identical looking iron bars. One of these is a magnet. Using just these two, how will you identify which of the two is a magnet.
3. You are given a thread and two bar magnets. Describe a method by which you can identify the polarities of the two magnets.

Magnetic field of the Earth

The directive property of magnets could be explained by considering that the earth acts as a magnet, i.e., as if a large bar magnet is placed inside the earth. The south pole of this magnet is considered near the geographical north pole and the magnetic north pole near the geographical south pole. RR_1 is the rotation axis of earth and MM_1 is the magnetic axis of the earth.

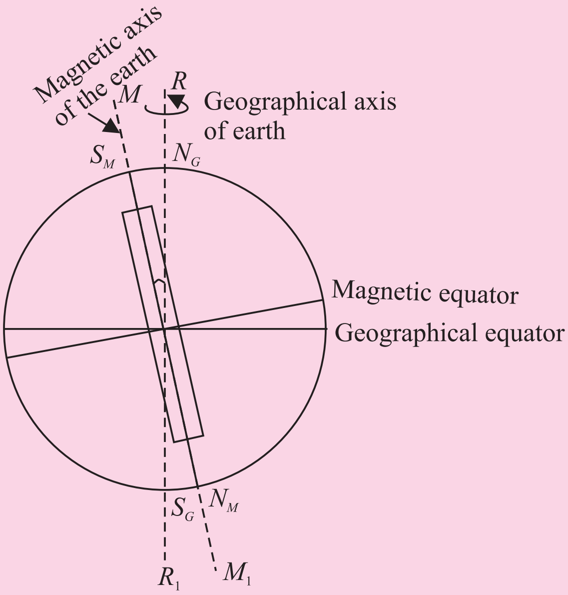


Fig. 18.4 : Magnetic field of the earth

ACTIVITY 18.1

Let us perform an experiment with a magnetic needle. (You can actually perform the experiment with a globe containing a bar magnet along its axis of rotation with north pole of the magnet pointing south.) Suspend the needle freely in such a manner that it can rotate in horizontal as well as vertical planes. If the needle is near the equator on earth's surface, it rests in horizontal plane. Suppose this needle is taken to places in the northern hemisphere. The needle rotates in the vertical plane and the north pole dips towards the earth, as we move towards geographical north pole. Finally at a point very near to Hudson bay in Canada, the north pole of the needle will point vertically downward. This place, located at 6° east of north, is considered to be the south pole of the earth's magnet.

This place is about 650 km away from the earth's geographical north pole. If we take the same magnetic needle to places in the southern hemisphere, the south pole of the needle will dip downward and point vertically downward at a point 650 km west of the geographical south. This point could be considered as the *N* ***pole of the earth's magnet. From this we conclude that the magnetic axis of the earth does not coincide with the geographical axis.***

An important aspect of earth's magnetic field is that it does not remain constant; its magnitude and direction change with time.

Elements of the Earth's Magnetic Field

Three measurable quantities are used to describe the magnetic field of earth. These are called elements of earth's magnetic field :

(a) Inclination or dip (δ);

(b) Declination (θ); and

(c) Horizontal component of the earth's field ($\mathbf{B}_M$).

(a) Inclination or Dip

If you suspend a magnetic needle freely at a place, you will observe that the needle does not rest in the horizontal plane. It will point in the direction of the resultant intensity of earth's field.

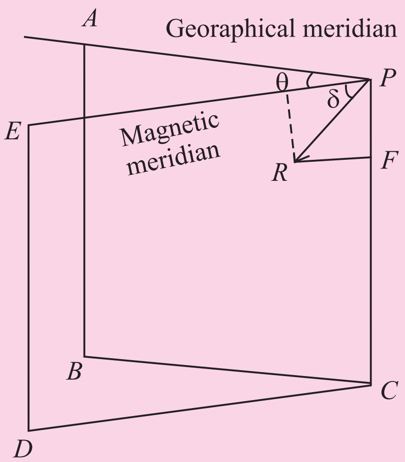


Fig. 18.5: Elements of earth's magnetic field

Fig. 18.5 shows the plane *PCDE*, which is the magnetic meridian at the point *P* (i.e. the vertical plane passing through the north and south poles of the earth's magnet) on the surface of the earth and *PABC* is the geographic meridian (i.e. the vertical plane passing through the geographical north and south poles of the earth). Suppose that *PR* represents the magnitude and direction of the earth's magnetic field at the point *P*. Note that *PR* makes an angle δ with the horizontal direction. This angle is known as inclination or dip at *P* on the surface of the earth.

The angle which the earth's magnetic field makes with the horizontal direction in the magnetic meridian is called the ***dip or inclination.***

(b) Declination

Refer to Fig 18.5 again. The plane *PCDE* contains the magnetic field vector (**PR**) of the earth. The angle between the planes *PCDE* and *PABC* is called the declination at the point *P*. It is shown as angle θ.

The angle which the megnetic meridian at a place makes with the geographical meridian is called the declination at that place.

(c) Horizontal component

Fig. 18.5 shows that **PR** is the resultant magnetic field at the point *P*. *PH* represents the horizontal component and *PF* the vertical component of the earth's magnetic field in magnitude and direction. Let the magnetic field at the point *P* be **B**. The horizontal component

$$B_H = B \cos \delta \qquad (18.1)$$

and the vertical component

$$B_V = B \sin \delta \qquad (18.2)$$

By squarring and adding Eqns. (18.1) and (18.2), we get

$$B_H^2 + B_V^2 = B^2 \cos^2\delta + B^2 \sin^2\delta = B^2 \qquad (18.3)$$

On dividing Eqn. (18.2) by Eqn. (18.1), we have

$$\frac{B_V}{B_H} = \tan \delta \qquad (18.4)$$

18.2 ELECTRICITY AND MAGNETISM : BASIC CONCEPTS

You now know that flow of electrons in a conductor due to a potential difference across it constitutes electric current. The current flowing in a conductor is seen to exert a force on a free magnetic needle placed in a region around it. A magnetic needle is also affected by a magnet and hence we say that a current carrying conductor has a magnetic field around it. The magnetic field *B* is visualized by magnetic field lines. You will learn about these and some more terms such as magnetic permeability later in this lesson.

18.2.1 Magnetic Field around an Electric Current

Let us do a simple experiment.

ACTIVITY 18.2

Take a 1.5 volt battery, a wire about 1 m in length, a campass needle and a match box. Wind 10-15 turns of the electric wire on its base. Under the windings, place a campass needle, as shown in Fig. 18.6. Place the match box on the table so as to have the wires running along the north – south direction. Connect the free ends of the wire to the battery. What happens to the needle? You will observe that needle shows deflection. This means that there is a magnetic field in and around the coil. The deflection will reverse if you reverse the direction of current by changing the terminals of the battery. When there is no

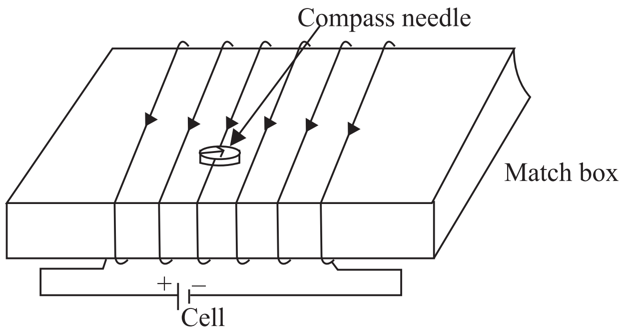
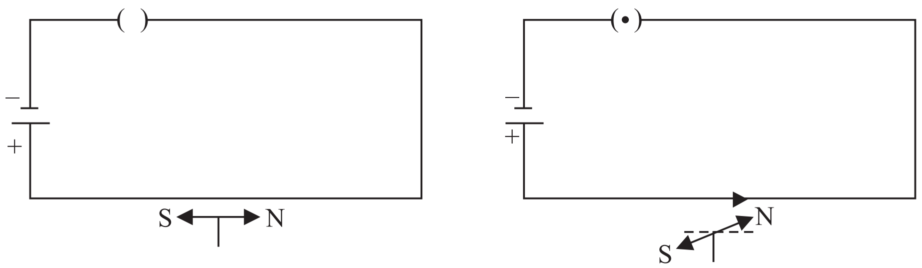


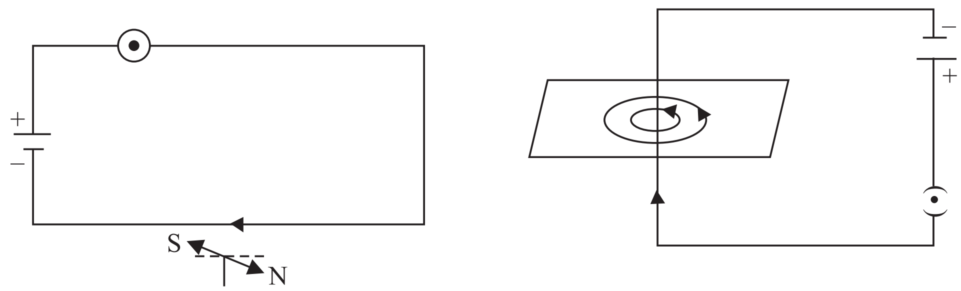
Fig. 18.6 : Demonstration of magnetic field due to electric current

current in the wire, the compass needle points in the north – south direction (Fig. 18.7 a, b & c). When a magnetic needle is brought close to a vertical current carrying wire, the magnetic field lines are concentric circles around the wire, as shown in Fig 18.7 (d).



(a) No current, No deflection north

(b) Current towards north deflection of pole towards west



(c) When direction of current is reversed, direction of deflection is reversed

(d) Circular field lines around a straight current carrying conductor

Fig. 18.7 : Magnetic field around a current carrying conductor

In 1820 Hans Christian Oersted, Professor of Physics at Copenhaegen in Denmark performed similar experiments and established that there is a magnetic field around a current carrying conductor.

18.3 BIOT-SAVART'S LAW

Biot-Savart's law gives a quantitative relationship between current in conductor and the resulting magnetic field at a point in the space around it. Each part of a current carrying conductor contributes to magnetic field around it. The net value of **B** at a point is thus the combined effect of all the individual parts of the conductor. As shown in Fig. 18.8, the net magnetic field due to any current carrying conductor is the vector sum of the contributions due to the current in each infinitesimal element of length $\Delta \ell$.

Experiments show that the field B due to an element $\Delta \ell$ depends on

- current flowing through the conductor, I;
- length of the element $\Delta \ell$;
- inversely proportional to the square of the distance of observation point P from the element $\Delta \ell$; and
- the angle between the element and the line joing the element to the observation point.

Thus, we can write

$$|\Delta B_0| \propto \frac{I \Delta \ell \sin\theta}{r^2}$$

$$= \frac{\mu_0}{4\pi} \frac{I\, d\ell \sin\theta}{r^2} \qquad (18.5)$$

where μ_0 is permeability of vacuum. Its value is $4\pi \times 10^{-7}$ WA^{-1}m^{-1}. The value of permeability of air is also nearly equal to μ_0

If the conductor is placed in a medium other than air, the value of the field is altered and is given by $|\mathbf{B}| = \mu |\mathbf{B_0}|$. Here μ represents the permeabilty of the medium.

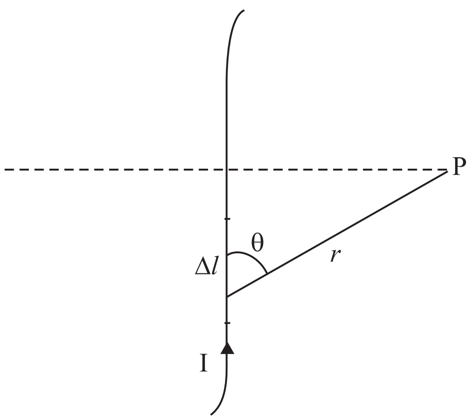
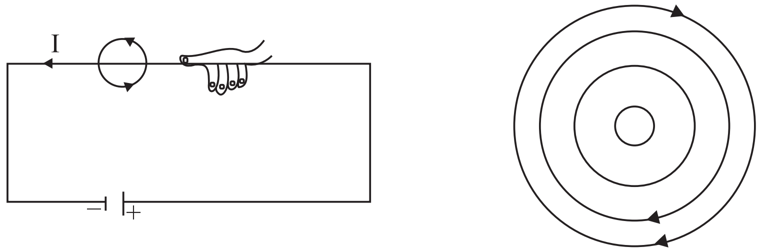


Fig. 18.8 : Magnetic field at P due to a current element Δl

Direction of B : Magnetic field at a point is a vector quantity. The direction of **B** may be determined by applying the right hand grip rule. To apply this rule, let us consider the direction of the field produced in some simple cases. As shown in the Fig. 18.9 (a), grasp the wire in your right hand so that the thumb points in the direction of the current. Then the curled fingers of the hand will point in the direction of the magnetic field. To represent the magnetic field on paper, let us consider that current is flowing into the plane of the paper. Then according to the right hand rule, the field lines shall be in the plane of the paper (Fig.18.9 b).



(a) (b)

Fig. 18.9 : Direction of magnetic field : **a)** Right hand rule : thumb in the direction of current, field lines in the direction of curling fingers, and **b)** when current is in the plane of paper, the field lines shall be in the plane of paper, according to the right hand rule.

18.3.1 Applications of Biot-Savart's Law

You now know that Biot-Savart's law gives the magnitude of the magnetic field. Let us now apply it to find the field around conductors of different shapes. Note that to calculate the net field due to different segments of the conductor, we have to add up the field contributions due to each one of them. We first consider a circular coil carrying current and calculate magnetic field at its centre.

(a) Magnetic field at the centre of a circular coil carrying current : Refer to Fig.18.10. It shows a circular coil of radius r carrying current I. To calculate magnetic field at its centre O, we first consider a small current element $\Delta \ell$ of the circular coil. Note that the angle between current element $\Delta\ell$ and r is 90º. From Eqn. (18.5) we know that the field at the centre O due to $\Delta\ell$ is

$$|\Delta \mathbf{B}| = \frac{\mu_0}{4\pi} I \frac{\Delta\ell}{r^2} \sin 90^\circ$$

$$= \frac{\mu_0}{4\pi} I \frac{\Delta\ell}{r^2} \qquad \text{(as } \sin 90^\circ = 1\text{)}$$

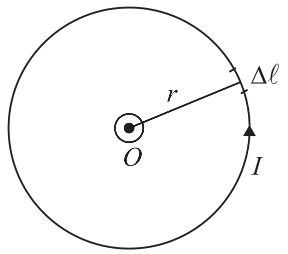


Fig. 18.10: Circular coil carrying current

The direction of $\Delta\mathbf{B}$ is normal to the plane of the coil. Since the field due to every element of the circular coil will be in the same direction, the resultant is obtained by adding all the contributions at the centre of the loop. Therefore

$$|\mathbf{B}| = \Sigma|\Delta\mathbf{B}| = \frac{\mu_0 I}{4\pi r^2}\ \Sigma\ \Delta\ell = \frac{\mu_0 I}{4\pi r^2}\ .\ 2\pi r$$

Hence, magnetic field at the centre of a coil of radius r carrying current I is given by

$$|\mathbf{B}| = \frac{\mu_0 I}{2r} \qquad (18.6)$$

In case there is more than one loop of wire (say there are n turns), the field is given by

$$|\mathbf{B}| = \frac{\mu_0 \, nI}{2r}$$

You can check the direction of the net field using the rule given in Fig. 18.7. You can use right hand rule in any segment of the coil and will obtain the same result. (Another simple quick rule to identify the direction of magnetic field due to a current carrying coil is the so called End-rule, illustrated in Fig. 18.11 (a, b).

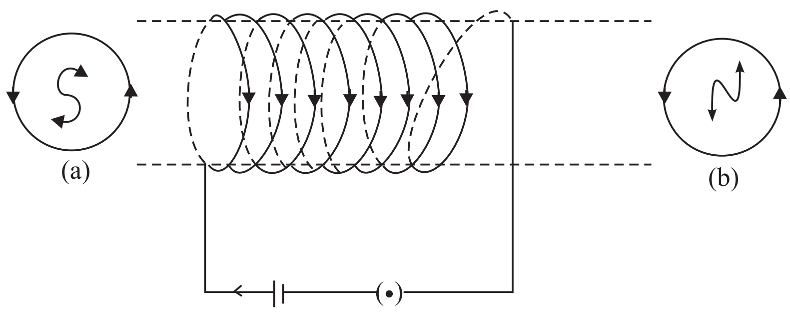


Fig 18.11: Direction of magnetic field : End-rule

When an observer looking at the circular coil at its either end finds the current to be flowing in the clockwise sense, the face of the coil behaves like the south pole of the equivalent magnet, i.e., B is directed inwards. On the other hand, if the current is seen to flow in the anticlockwise sense, the face of the coil behaves like the north pole of the equivalent magnet or the field is directed out of that end.

INTEXT QUESTIONS 18.2

1. What can you say about the field developed by

   (i) a stationary electron ?

   (ii) a moving electron ?

2. Electrons in a conductor are in constant motion due to thermal energy. Why do they not show magnetism till such time that a potential difference is applied across it ?

3. A current is flowing in a long wire. It is first shaped as a circular coil of one turn, and then into a coil of two turns of smaller radius. Will the magnetic field at the centre coil change? If so, how much ?

18.4 AMPERE'S CIRCUITAL LAW

Ampere's circuital law provides another way of calculating magnetic field around a current carrying conductor in some simple situations.

Ampere's circuital law states that the line integral of the magnetic field **B** around a closed loop is μ_0 times the total current, *I*. Mathematically, we write

$$\oint \mathbf{B} \cdot \mathbf{dl} = \mu_0 I \qquad (18.7)$$

Note that this is independent of the size or shape of the closed loop.

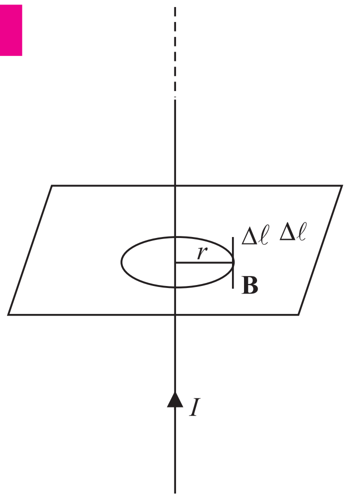


Fig. 18.12 : Ampere's circuital law

Andre Marie Ampere (1775 – 1836)

French Physicist, mathematician and chemist, Ampere was a child prodigy. He mastered advanced mathematics at the age of 12. A mix of experimental skills and theoretical acumen, Ampere performed rigorous experiments and presented his results in the form of a theory of electrodynamics, which provides mathematical formulation of electricity and its magnetic effects. Unit of current is named in his honour. Lost in his work and ideas, he seldom cared for honours and awards. Once he forgot an invitation by emperor Nepoleon to dine with him. His gravestone bears the epitaph : Tendun felix (Happy at last), which suggests that he had to face a very hard and unhappy life. But it never lowered his spirit of creativity.

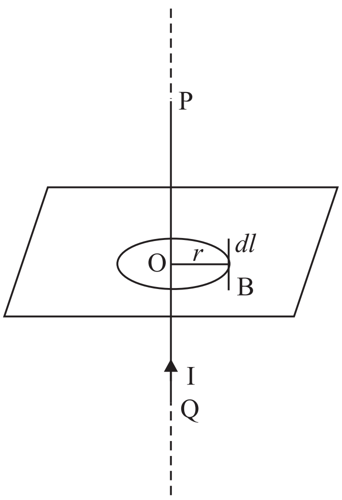


Fig. 18.13: Infinitely long current carrying conductor

18.4.1 Applications of Ampere's Circuital Law

We now apply Ampere's circuital law to obtain magnetic field in two simple situations.

(a) Magnetic field due to an infinitely long current carrying conductor

Refer to Fig. 18.13. It shows an infinitely long current carrying conductor *POQ* carrying current *I*. Consider a circular loop of radius *r* around it in the plane as shown. Then

$$\Sigma \mathbf{B}.d\ell = B\, 2\pi r$$

By applying Ampere's circuital law, we can write

$$|\mathbf{B}|\ 2\pi r = \mu_o I$$

or $$|\mathbf{B}| = \frac{\mu_0 I}{2\pi r} \qquad (18.8)$$

This gives the magnetic field around an infinitely long straight current carrying conductor.

Solenoids and toroids are widely used in motors, generators, toys, fan-windings, transformers, electromagnets etc. They are used to provide uniform magnetic field. When we need large fields, soft iron is placed inside the coil.

(b) Magnetic field due to a solenoid

A solenoid is a straight coil having a large number of loops set in a straight line with a common axis, as shown in Fig. 18.14. We know that a current I flowing through a wire, sets up a magnetic field around it. Suppose that the length of the solenoid is ℓ and it has N number of turns. To calculate the magnetic field inside the solenoid along its axis (Fig 18.14), we can treat it to be a section of a toroidal solenoid of a very large radius. Thus :

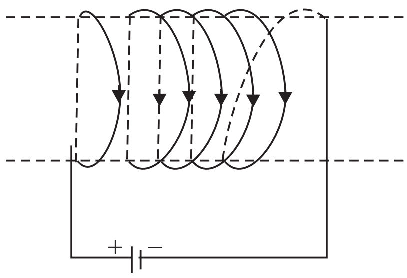


Fig. 18.14 : A solenoid

$$|\mathbf{B}| = \mu_0 \, nI$$

The direction of the field is along the axis of the solenoid. A straight solenoid is finite. Therefore, $|\mathbf{B}| = \mu_0 \, nI$ should be correct well inside the solenoid, near its centre.

For solenoids of small radius, the magnitude of B at the ends is given by

$$|\mathbf{B}| = \frac{\mu_0 nI}{2} \qquad (18.9)$$

The solenoid behaves like a bar magnet and the magnetic field is as shown in Fig. 18.15.

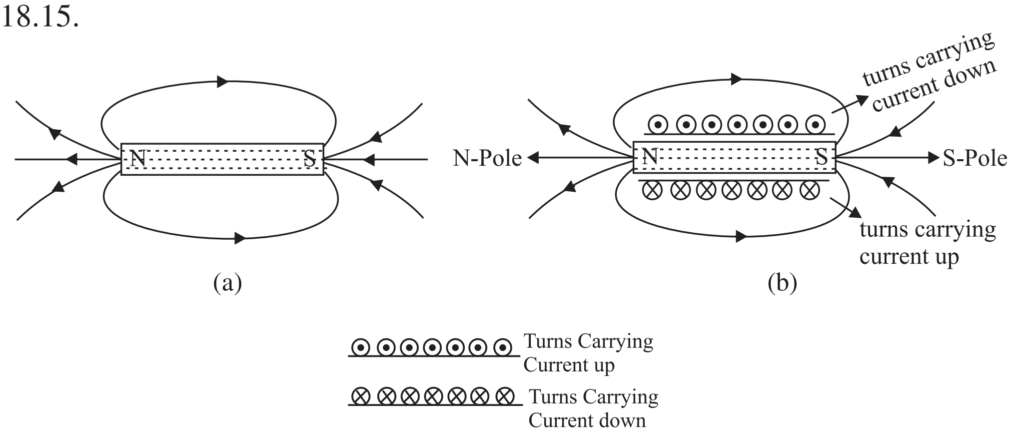


Fig. 18.15: Solenoid behaves like a bar magnet : **a)** Magnetic field due to a bar magnet, and **b)** magnetic field due to a current carrying solenoid

18.4.2 Application of Ampere's Circuital Law

(b) Magnetic Field due to a Straight Solenoid

A solenoid is a straight coil having a large number of loops set in a straight line with a common axis, as shown in Fig. 18.4.2. We know that a current I flowing through a wire, sets up a magnetic field around it. Suppose that the length of the solenoid is ℓ and it has N number of turns.

The magnetic field inside the solenoid, in its middle, is uniform and parallel to its axis. Outside the solenoid, however, the field is negligibly weak. These statements hold true, strictly speaking, if the length of the solenoid is very large as compared to its diameter. For a long solenoid, whose windings are very tightly and uniformly wound, the magnetic field inside it is fairly uniform everywhere and is zero outside it.

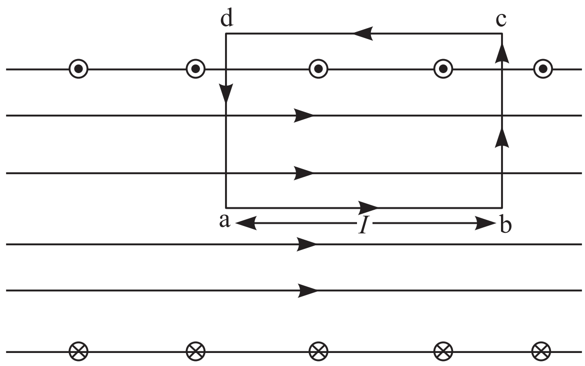


Fig. 18.4.2

Let us take a rectangular loop abcd as shown in Fig 18.4.2. Along the path ab, the magnetic field is uniform. Hence, for this path $\mathbf{B}\cdot d\ell = B\ell$. Along the paths *cd*, as the magnetic field is weak it may be taken as zero. Hence, for this path $\mathbf{B}\cdot d\ell = 0$. The two short sides *bc* and *da* also do not contribute anything to $\mathbf{B}\cdot d\ell$ as $\mathbf{B}$ is either zero (outside the solenoid), or perpendicular to $d\ell$ (inside the solenoid).

If n be the number of turns per unit length along the length of the solenoid, then the number of turns enclosed by the rectangular loop of length ℓ is $n\ell$. If each turn of the solenoid carries a current i, then the total current threading the loop is $n\ell i$. Hence, from Ampere's circuital law,

$$\sum \mathbf{B}\cdot d\ell = \mu_0 (nli)$$

or $$B\ell = \mu_0 nli$$

or $$B = \mu_0 ni$$

(c) Magnetic field due to a toroid

A toroid is basically an endless solenoid which may be formed by bending a straight solenoid so as to give it a circular shape.

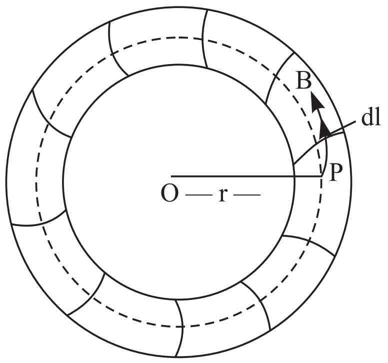


Fig. 18.4.3

Suppose, we want to find the magnetic field at a point *P*, inside the toroid, whose distance from the centre *O* is *r*. Draw a circle passing through the point *P* and concentric with the toroid. The magnetic field will everywhere be tangential to the circle, its magnitude being the same at all points of it. So, we can write:

$$\sum \mathbf{B} \cdot d\ell = \sum B d\ell = B \sum d\ell$$

But $\sum d\ell = 2\pi r$, the circumference of the circular path.

Therefore,

$$\sum \mathbf{B} \cdot d\ell = 2\pi r B$$

If *N* be the total number of turns and *i* the current flowing through the windings of the toroid, then the total current threaded by the circular path of radius *r* is *Ni*. Hence, from Ampere's circuital law,

$$\sum \mathbf{B} \cdot d\ell = \mu_0 N i$$

or $$2\pi r B = \mu_0 N i$$

or $$B = \frac{\mu_0 N i}{2\pi r}$$

18.4.3 Electromagnets and Factors Affecting their Strength

We have seen that a current-carrying solenoid behaves as a bar magnet, with one end behaving as north pole and the other as south pole depending on the direction of flow of current. The polarity of such magnets is determined by the end rule and the strength of the magnetic field is given by

$$|\mathbf{B}| = \mu_0 nI$$

where μ_0 is the permeability of free space, n is the number of turns per unit length and I is the current flowing through the solenoid.

It is clear that the solenoid remains a magnet as long as the a current is flowing through it. Thus, a current-carrying solenoid is called an electromagnet.

Its strength depends on :

(i) Number of turns per unit length of the solenoid, and

(ii) The current flowing through it.

It may also be noted that the strength of the magnetic field of an electromagnet increases when a soft iron core is introduced inside it.

18.4.4 Concept of Displacement Current

The concept of displacement current was introduced by Maxwell. As we know, magnetic field is produced due to the conduction current. However, according to Maxwell in empty space (where no conduction current exists), the magnetic field is produced due to the displacement current which, unlike conduction current, is not associated with the motion of charges.

Consider a simple circuit consisting of a small parallel-plate capacitor being charged by a current I.

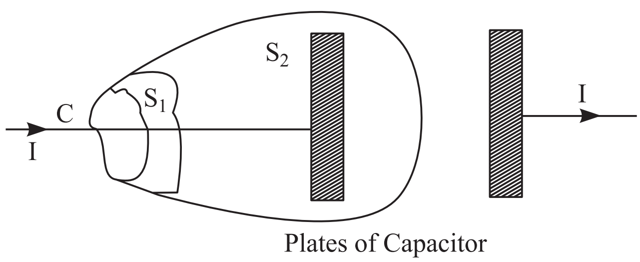


Fig. 18.4.4

Applying Ampere's circuital law to the contour C and the surface S_1, we find

$$\oint \mathbf{B} \cdot d\ell = \mu_0 I$$

However, applying Ampere's circuital law to the contour C and the surface S_2, as there is no current through this surface, we get

$$\oint \mathbf{B} \cdot d\ell = 0$$

The above two equations are mutually contradictory. To avoid this contradiction, Maxwell assumed that a current exists between the capacitor plates. He called this current displacement current and showed that this current arises due to the variation of electric field with time.

A simple expression for the displacement current can be derived as follows. Consider a parallel plate capacitor. Let q be the charge on the capacitor plates at any instant t.

The electric field inside the capacitor is given by

$$E = \frac{q}{A\varepsilon_0}$$

When A is the surface area of the plates. Therefore, the electric flux through the capacitor is

$$\phi_E = EA = \frac{q}{\varepsilon_0}$$

The rate of change of the instantaneous flux can be written as

$$\frac{\Delta\phi_E}{\Delta t} = \frac{1}{\varepsilon_0}\frac{\Delta q}{\Delta t} = \frac{I}{\varepsilon_0}$$

So, we can write

$$\varepsilon_0 \frac{\Delta\phi_E}{\Delta t} = I$$

The expression on the left hand side is equivalent to a current, which though equal to the conduction current I is actually different from it as it is not associated with the motion of free charges. It is called displacement current. Unlike the conduction current I, the displacement current arises whenever the electric field and hence the electric flux changes with time.

Adding displacement current to the conduction currents I, Maxwell modified the Ampere's circuital law in the form,

$$\sum \mathbf{B} \cdot d\ell = \mu_0 \left(I + \varepsilon_0 \frac{\Delta\phi_E}{\Delta t} \right)$$

Maxwell's modification of Ampere's law tells us that, in addition to conduction current, a time–varying electric field can also produce magnetic field.

Example 18.1 : A 50 cm long solenoid has 3 layers of windings of 250 turns each. The radius of the lowest layer is 2cm. If the current through it is 4.0 A, calculate the magnitude of **B** (a) near the centre of the solenoid on and about the axis; (b) near the ends on its axis; and (c) outside the solenoid near the middle.

Solution :

a) At the centre or near it

$$B = \mu_0 nI$$
$$= 4\pi \times 10^{-7} \times \frac{3\times 250}{0.5} \times 4$$
$$= 16\pi \times 1500 \times 10^{-7}\ \text{T}$$
$$= 24\pi \times 10^{-4}\ \text{T}$$

b) At the ends

$$B_{ends} = \frac{1}{2}\ B_{centre} = 12\pi \times 10^{-4}\ \text{T}$$

c) Outside the solenoid the field is zero.

Example 18.2: Calculate the distance from a long straight wire carrying a current of 12A at which the magnetic field will be equal to 3×10^{-5} T.

Solution : $$B = \frac{\mu_0 I}{2\pi r} \quad \Rightarrow r = \frac{\mu_0 I}{2\pi B}$$

∴ $$r = \frac{2\times 10^{-7}\times 12}{3\times 10^{-5}} = 0.25\ \text{m}$$

INTEXT QUESTIONS 18.3

1. A drawing of the lines of force of a magnetic field provides information on
   a) direction of field only
   b) magnitude of field only
   c) both the direction and magnitude of the field
   d) the force of the field
2. What is common between Biot-Savart's law and Ampere's circuital law ?
3. In the following drawing of lines of force of a non-unifrom magnetic field, at which piont is the field (i) uniform, (ii) weakest, (iii) strongest?

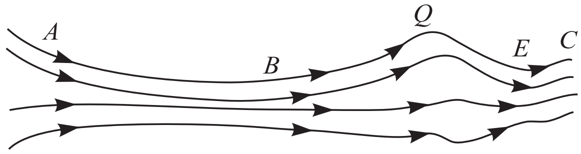


Fig. 18.16 : A typical magnetic field

4. A 10 cm long solenoid is meant to have a magnetic field 0.002T inside it, when a current of 3A flows through it. Calculate the required no. of turns.
5. Derive an expression for the field due to a toroid using Ampere's circuital law.

18.5 FORCE ON A MOVING CHARGE IN A MAGNETIC FIELD

When a charged body moves in a magnetic field, it experiences a force. Such a force experienced by a moving charge is called the *Lorentz force*. The Lorentz force on a particle with a charge $+q$ moving with a velocity $\boldsymbol{v}$ in a magnetic field **B** is given by

$$\mathbf{F} = q(\boldsymbol{v} \times \mathbf{B})$$

or

$$|\mathbf{F}| = q\, v B \sin\theta \qquad (18.10)$$

where θ is the angle between the directions of v and **B**. The direction of **F** is given by Fleming's left hand rule.

Fleming's left hand rule states that if we stretch the fore finger, the central finger and the thumb of our left hand at right angles to each other and hold them in such a way that the fore finger points in the direction of magnetic field and the central finger points in the direction of motion of positively charged particle, then the thumb will point in the direction of the Lorentz force (Fig. 18.17).

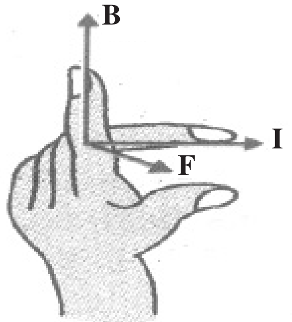


Fig. 18.17 : Fleming's left hand rule

Some important points to note

- **F** is a mechanical force resulting in a pull or a push.
- The direction of force is given by Fleming's left hand rule.
- In case of negative charges, the central finger should point opposite to the direction of its motion.
- If the charge stops, the force becomes zero instantly.
- Force is zero when charges move along the field **B**.
- Force is maximum when charges move perpendicular to the field : $F = qvB$

18.5.1 Force on a Current Carrying Conductor in a Uniform Magnetic Field

The concept of Lorentz force can be easily extended to current carrying conductors placed in uniform magnetic field **B**. Suppose that the magnetic field is parallel to the plane of paper and a conductor of length $\Delta\ell$ carrying current I is placed normal to the field. Suppose further that the current is flowing downward with a drift velocity $\boldsymbol{v}_d$ and hence each free electron constituting the current experiences a Lorentz force $\mathbf{F} = e\, \boldsymbol{v}_d . \mathbf{B}$

If there are N free electrons in the conductor, the net force on it is given by

$$F = N e v_d B = nA \Delta\ell\, ev_d B \qquad (18.11)$$

where n denotes the number of free electrons per unit volume. But $neAv_d = I$. Hence

$$\therefore \qquad F = I \Delta\ell\, B \qquad (18.12)$$

If conductor makes an angle θ with **B**, then $|\mathbf{F}| = I\, \Delta\ell\, B \sin\theta$.

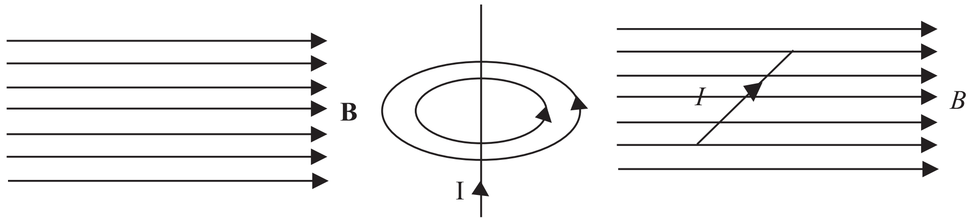


Fig. 18.18: **a)** Uniform magnetic field, **b)** field due to current carrying inductor, and **c)** force on a current carrying conductor

The direction of the force is again given by Fleming's left hand rule.

Eqn. (18.12) can be used to define the unit of magnetic field in terms of the force experiencd by a current carrying conductor. By rearranging terms, we can write

$$B = \frac{F}{I\Delta\ell}$$

Since F is taken in newton, I in ampere and $\Delta\ell$ in metre, the unit of B will be $NA^{-1}\,m^{-1}$. It is called tesla (T).

18.5.2 Force Between two Parallel Wires Carrying Current

You now know that every current carrying conductor is surrounded by a magnetic field. It means that it will exert force on a nearby current carrying conductor. The force between two current carrying conductors placed parallel to each other is mutual and magnetic in origin. A current carrying wire has no net electric charge, and hence cannot interact electrically with another such wire.

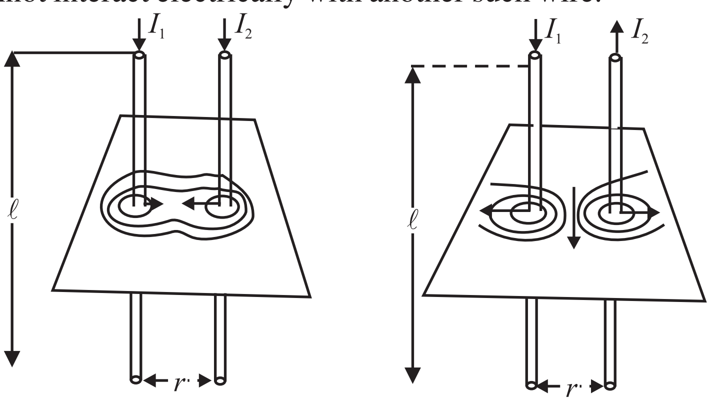


Fig. 18.19: Experimental demonstration of force between two parallel wires carrying current

Fig 18.19 shows two parallel wires separated by distance r and carrying currents I_1 and I_2, respectively. The magnetic field due to one wire at a distance r from it is $B_1 = \frac{\mu_0 I_1}{2\pi r}$.

Similarly, the field due to second wire at a distance r from it will be $B_2 = \frac{\mu_0 I_2}{2\pi r}$.

These fields are perpendicular to the length of the wires and therefore the force on a length l, of the other current carrying conductor is given by

$$F = B\,I\,\ell = \frac{\mu_0 I_1}{2\pi r}\, I_2\, \ell$$

or force per unit length

$$\frac{F}{\ell} = \frac{\mu_0\, I_1\, I_2}{2\pi r} \qquad (18.13)$$

The forces are attractive when the currents are in the same direction and repulsive when they are in opposite directions.

Eqn (18.13) can be used to define the unit of current. If $I_1 = I_2 = 1$A, $l = 1$m and $r = 1$m, then

$$F = \frac{\mu_0}{2\pi} = 2 \times 10^{-7}\,\text{N}$$

Thus, ***if two parallel wires carrying equal currents and placed 1 m apart in vacuum or air experience a mutual force of 2×10^{-7} N m^{-1}, the current in each wire is said to be one ampere***.

18.5.3 Motion of a Charged Particle in a Uniform Field

We can now think of various situations in which a moving charged particle or a current carrying conductor in a magnetic field experiences Lorentz force. The work done by a force on a body depends on its component in the direction of motion of the body. When the force on a charged particle in a magnetic field is perpendicular to its direction of motion, no work is said to be done. Hence the particle keeps the same speed and kinetic energy which it had while moving in the field, even though it is deflected. On the other hand, the speed and energy of a charged particle in an electrical field is always affected due to the force by the field on the particle. A charged particle moving perpendicular to a magnetic field follows a circular path (Fig. 18.20) because it experience a force at right angles to the direction of motion at every position.

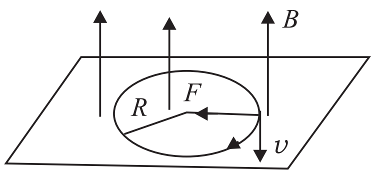


Fig. 18.20: Path of a charged particle in a uniform magnetic field

To know the radius R of the circular path of the charged particle, we note that the magnetic force $q\,\upsilon B$ provides the particle with the centripetal force ($m\,\upsilon^2/R$) that keeps it moving in a circle. So we can write

$$q\ \upsilon\ B = \frac{m\upsilon^2}{R}$$

On rearrangement, we get

$$R = \frac{m\upsilon}{q\,B} \qquad (18.14)$$

The radius of the path traced by a charged particle in a uniform magnetic field is directly proportional to its momentum ($m\upsilon$) and inversely proportional to its charge and the magnetic field. It means that greater the momentum, larger the circle, and stronger the field, the smaller the circle. The time period of rotation of the particle in a circular path is given by

$$T = \frac{2\pi R}{\upsilon} = \frac{2\pi\, m}{Bq} \qquad (18.14\ a)$$

Note that the time period is independent of velocity of the particle and radius of the orbit. It which means that once the particle is in the magnetic field, it would go round and round in a circle of the same radius. If m, B, q, remain constant, the time period does not chnage even if υ and R are changed.

Now think, what happens to R and T if a) field B is made stronger; b) field B is made weaker; c) field B ceases to exist; d) direction of B is changed; d) the particle is made to enter the magnetic field at a higher speed; f) the particle enters at an angle to B; and g) the charged particle loses its charge.

18.5.4 Motion of a Charged Particle in uniform Electric Field and Magnetic Field

(a) Motion in Electric Field

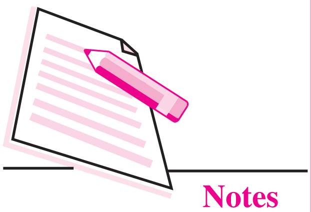
When a charged particle q is placed in a uniform electric field E, it experiences a force,

$$\mathbf{F} = q\mathbf{E}$$

Thus, the charged particle will be accelerated under the influence of this force. The acceleration is given by

$$\mathbf{a} = \frac{\mathbf{F}}{m} = \frac{q\mathbf{E}}{m}$$

The acceleration will be in the direction of the force. If it is a positive charge, it will accelerate in the direction of the field and if it is a negative charge it will



Notes

accelerate in a direction opposite that of the field. The velocity and displacement of charged particle can also be calculated by using the equations of motion:

$$v = u + \left(\frac{qE}{m}\right)t$$

$$s = ut + \frac{1}{2}\left(\frac{qE}{m}\right)t^2$$

where t denotes time.

(b) Motion in magnetic field

In article 18.5 (Page 114, Book 2), it has been discussed that the force experienced by a charged particle in a magnetic field is given by

$$F = qBv\sin\theta$$

Where θ is angle between the velocity and magnetic field.

If $\theta = 0$, $F = 0$ and charged particle will move along a straight line with constant speed.

If $\theta = 90°$, F will be maximum and its direction, according to Fleming's left hand rule, will be perpendicular to the plane of **v** and **B** and the charged particle will move along a circular path with a constant speed and frequency.

If $\theta \neq 0° \neq 90°$, then the velocity of the charged particle will be $v\sin\theta$ perpendicular to the field and $v\cos\theta$ parallel to the field. The particle, therefore, moves along a helical path.

What we note from the above discussion is that a magnetic field does not change the speed of a moving charge, it only changes its direction of motion.

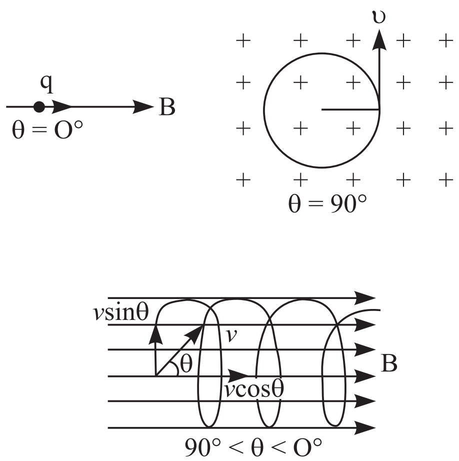
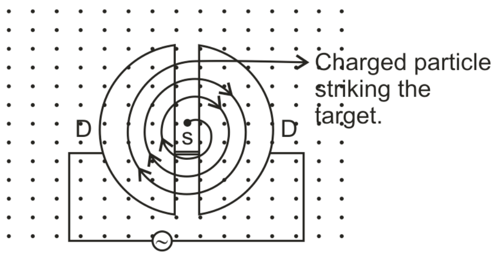


Fig. 18.5.4

18.5.5 Cyclotron

The cyclotron is a device invented by E.O. Lawrence in 1929, that is used for accelerating charged particles (such as protons, deutron or α–particles) to high velocities. It consists of two semi-circular hollow metallic disks *DD*, called dees, on account of their shape resembling the letter *D* of English alphabet. They are insulated from one another with a small gap between them. The dees are placed in an evacuated chamber.



18.5.5

A magnetic field perpendicular to the plane of the dees (out of the paper in Fig) is maintained with the help of an electromagnet having flat pole–pieces. A rapidly oscillating potential difference is applied between the dees with the help of an oscillator. This produces an oscillating electric field in the gap between the dees.

Consider a charged particle of mass m and charge q in the gap between the dees. The particle is accelerated by the electric field towards one of them. Inside the dees, it moves with constant speed in a semicircle in a clockwise direction. If the frequency of the oscillator is equal to the frequency of revolution of the charged particle, then it reaches the gap at the instant when the opposite dee becomes negative because of the reversal of the direction of electric field.

The frequency of revolution of the charged particle is given by (see Eq. 18. 14a):

$$\nu = \frac{1}{T} = \frac{v}{2\pi R} = \frac{Bq}{2\pi m}$$

where B is the magnetic field.

It is also called cyclotron frequency and denoted by ν_c. When $\nu_c = \nu_o$, the frequency of the oscillator, the particle reaches the gap when the electric potential at the opposite 'D' has just reversed its sign. This condition is also known as cyclotron resonance condition. On account of this, the particle gains energy and, therefore, it moves in a circle of larger radius. This energy gain can be repeated many times.

Thus, the energy and the radius of the path of the particle keep on increasing progressively. However, the maximum radius which the path can have is limited by the radius R of the dees. The high energy charged particle finally comes out through an opening in the dee.

Example 18.3 : Refer to Fig. 18.21 and calculate the force between wires carrying current 10A and 15A, if their length is 5m. What is the nature of this force ?

Solution : When currents flow in two long parallel wires in the same direction, the wires attract each other and the force of attraction is given by

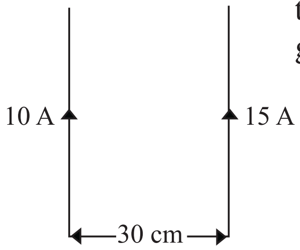


Fig. 18.21

$$\frac{F}{\ell} = \frac{\mu_0 I_1 I_2}{2\pi r} = \frac{2\times10^{-7}\times10\times15}{3} = 10^{-4}\ \text{N m}^{-1}$$

$$\therefore F = 5\times10^{-4}\ \text{N}$$

The force is attractive in nature.

Example 18.4 : An electron with velocity $3\times10^7\ \text{ms}^{-1}$ describes a circular path in a uniform magnetic field of 0.2T, perpendicular to it. Calculate the radius of the path.

Solution :

We know that $$R = \frac{mv}{Bq}$$

Here, $m_e = 9\times10^{-31}$ kg, $e = 1.6\times10^{-19}$ C, $v = 3\times10^7$ m s^{-1} and $B = 0.2$ T. Hence

$$R = \frac{9\times10^{-31}\times3\times10^7}{0.2\times1.6\times10^{-19}}$$

$$= 0.85\times10^{-3}\ \text{m}$$

$$= 8.5\times10^{-4}\ \text{m}$$

INTEXT QUESTION 18.4

1. A stream of protons is moving parallel to a stream of electrons but in the opposite direction. What is the nature of force between them ?
2. Both electrical and magnetic fields can deflect an electron. What is the difference between them?
3. A body is suspended from a vertical spring. What shall be the effect on the position of the body when a current is made to pass through the spring.
4. How does a cyclotron accelerate charged particles?

18.6 CURRENT LOOP AS A DIPOLE

From Eqn. (18,6) you will recall that the field at the centre of a coil is given by

$$B = \frac{\mu_0 I}{2r}$$

On multiplying the numerator and denominator by $2\pi r^2$, we can rewrite it as

$$B = \frac{\mu_0 \, 2I \,.\, \pi r^2}{4\pi r^3} = \frac{\mu_0 \, 2\,I\,A}{4\pi r^3} = \frac{\mu_0}{4\pi}\frac{2M}{r^3}$$

where A is area of coil and M is magnetic moment. This shows that a current carrying coil behaves like a magnetic dipole having north and south poles. One face of the loop behaves as north pole while the other behaves as south pole.

Let us now undertake a simple activity.

ACTIVITY 18.3

Suspend a bar magnet by a thread between pole pieces of a horse shoe magnate, as shown in Fig 18.22.

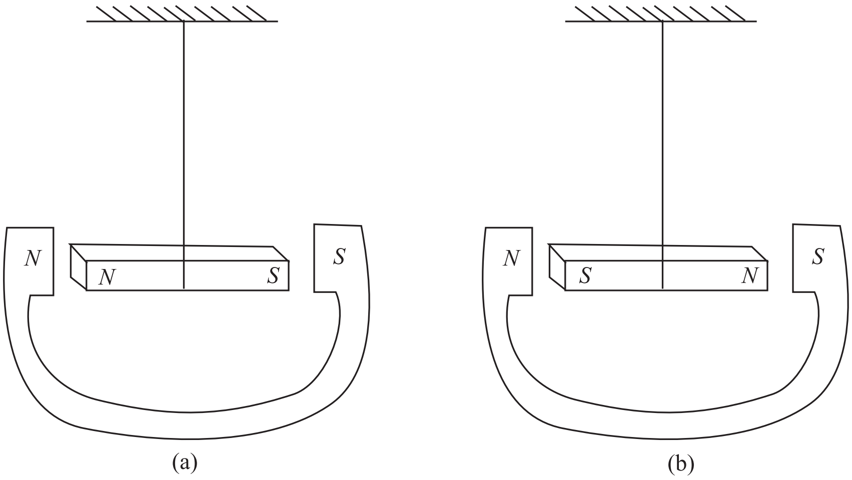


Fig. 18.22 : A bar magnet suspended between a horse shoe magnet

What will happen when the bar magnet shown in Fig. 18.24(a) is displaced slightly sideways? Since like poles repel, the bar magnet experiences a torque and tends to turn through 180º and get aligned, as shown in Fig. 18.22 (b). Since a current loop behaves as a magnet, it will align in an external field in the same way.

You have already studied the following equations in the lesson on electrostatics. The electric field of a dipole at a far point on its axis is given by

$$\mathbf{E} = \frac{1}{4\pi\mu_0}\frac{2\mathbf{P}}{x^3} \qquad (18.15\text{ b})$$

The magnetic field due to a current carrying coil is given by

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{2\,\mathrm{NIA}}{x^3} = \frac{\mu_0}{4\pi}\frac{2\mathbf{M}}{x^3} \qquad (18.15\text{ c})$$

where **M** is the magnetic dipole moment.

A comparison between these expressions leads us to the following analogies :

- A current loop behaves as a magnetic dipole with magnetic moment

$$\mathbf{M} = NI\mathbf{A} \qquad (18.15\text{ d})$$

- Like the poles of a magnetic dipole, the two faces of a current loop are inseparable.
- A magnetic dipole in a uniform magnetic field behaves the same way as an electric dipole in a uniform electric field.
- A magnetic dipole also has a magnetic field around it similar to the electric field around an electric dipole.

Thus magnetic field due to a magnetic dipole at an axial point is given by

$$\mathbf{B} = \frac{\mu_0}{4\pi}\frac{2\mathbf{M}}{x^3} \qquad (18.16)$$

whereas the field at an equatorial point is given by

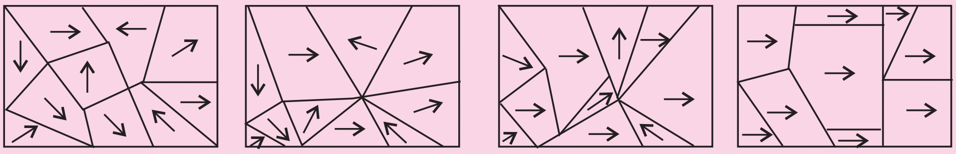
$$\mathbf{B} = -\frac{\mu_0}{4\pi}\frac{\mathbf{M}}{x^3} \qquad (18.17)$$

Magnetism in Matter

Based on the behaviour of materials in magnetic field, we can divide them broadly into three categories : (i) **Diamagnetic** materials are feebly repelled by a magnet. (ii) **Paramagnetic** materials are feebly attracted by a magnet. (iii) **Ferromagnetic** materials are very strongly attracted by a magnet. Substances like iron, nickel and cobalt are ferromagnetic. Let us study ferromagnetic behaviour of materials in some details.

Ferromagnetic materials, when placed even in a weak magnetic field, become magnets, because their atoms act as permanent magnetic dipoles. The atomic dipoles tend to align parallel to each other in an external field. These dipoles are not independent of each other. Any dipole strongly feels the presence of a neighboring dipole. A correct explanation of this interaction can be given only on the basis of quantum mechanics. However, we can qualitatively understand the ferromagnetic character along the following lines.

A ferromagnetic substance contains small regions called **domains**. All magnetic dipoles in a domain are fully aligned. The magnetization of domains is maximum. But the domains are randomly oriented. As a result, the total magnetic moment of the sample is zero. When we apply an external magnetic field, the domains slightly rotate and align themselves in the direction of the field giving rise to resultant magnetic moment. The process can be easily understood with the help of a simple diagram shown in Fig.18.23.



B = 0 (a) | **B** > 0 (b) | **B** >> 0 (c) | **B** >>> 0 (d)

Fig. 18.23 : Domains in a ferromagnetic substance

Fig. 18.23 (a) shows ten domains. For simplicity we take a two dimensional example. All the domains are so directed that the total magnetization of the sample is zero. Fig. 18.23 (b) shows the state after the application of an external magnetic field. The boundaries of the domains (Domain Walls) reorganise in such a way that the size of the domain having magnetic moment in the direction of the field becomes larger at the cost of others. On increasing the strength of external field, the size of favorable domains increases, and the orientation of the domain changes slightly resulting in greater magnetization (Fig. 18.23 (c)). Under the action of very strong applied field, almost the entire volume behaves like a single domain giving rise to saturated magnetization. When the external field is removed, the sample retains net magnetization. The domain in ferromagnetic samples can be easily seen with the help of high power microscope.

When the temperature of a ferromagnetic substance is raised beyond a certain critical value, the substance becomes paramagnetic. This critical temperature is known as Curie temperature T_c.

Table 18.1: Ferromagnetic substances and their curie temperatures

Substances	Curie temperature T_c (K)
Iron	1043
Nickel	631
Cobalt	1394
Gadolinium	317
Fe_2O_3	893

Example 18.5 : The smallest value of magnetic moment is called the Bohr Magneton $\mu_B = \dfrac{eh}{4\pi m}$. It is a fundamental constant. Calculate its value.

Solution :

$$\mu_B = \frac{eh}{4\pi m} = \frac{(1.6\times10^{-19}\,\text{C})\times(6.6\times10^{-34}\,\text{Js})}{4\times3.14\times(9\times10^{-31}\,\text{kg})}$$

$$= 9.34 \times 10^{-24}\ \text{J T}^{-1}$$

18.6.1 Torque on a Current Loop

A loop of current carrying wire placed in a uniform magnetic field (**B**) experiences no net force but a torque acts on it. This torque tends to rotate the loop to bring its plane perpendicular to the field direction. This is the principle that underlines the operation of all electric motors, meters etc.

Let us examine the force on each side of a rectangular current carrying loop where plane is parallel to a uniform magnetic field **B**. (Fig. 18.24 (a).)

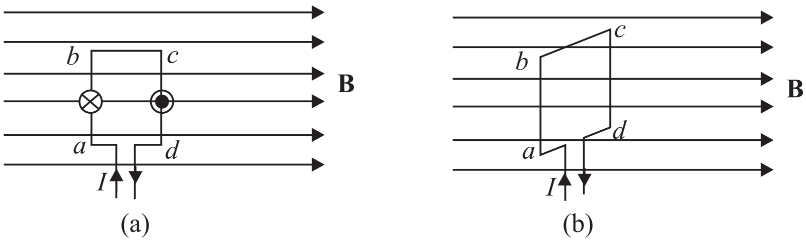


Fig. 18.24: Force on the sides of a rectangular loop when **(a)** the loop is parallel to the field, and **(b)** the coil is perpendicular to the field.

The sides *ad* and *bc* of the loop are parallel to **B.** So no force will act on them. Sides *ab* and *cd* are however, perpendicular to **B**, and these experience maximum force. We can easily find the direction of the force on *ab* and *cd*.

In fact, $|\mathbf{F}_{ab}| = |\mathbf{F}_{cd}|$ and these act in opposite directions. Therefore, there is no net force on the loop. Since $\mathbf{F}_{ab}$ and $\mathbf{F}_{cd}$ do not act along the same line, they exert a torque on the loop that tends to turn it. This holds good for a current loop of any shape in a magnetic field.

In case the plane of the loop were perpendicular to the magnetic field, there would neither be a net force nor a net torque on it (see Fig 18.26 (b)).

$$\text{Torque} = \text{force} \times \text{perpendicular distance between the force}$$

$$= B\,IL.\,b \sin\theta$$

Refer to Fig. 18.25 which shows a loop *PQRS* carrying current *I*. θ is the angle between the magnetic field **B** and the normal to the plane of the coil *n*. The torque is then

$$\tau = NBIL\,b \sin\theta$$

where *N* is the number of turns of the coil. We can rewrite it as

$$|\tau| = NBI\,A \sin\theta \qquad (18.18)$$

where *A* is area of the coil = *L x b*

$$|\tau| = |\mathbf{B}|\,|\mathbf{M}| \sin\theta \qquad (18.19)$$

where **M** = *NI***A** is known as the magnetic moment of the current carrying coil.

Thus, we see that the torque depends on *B, A, I, N* and θ

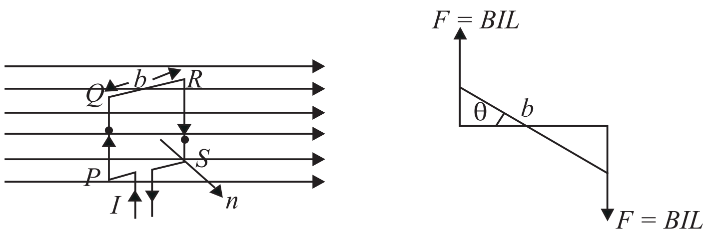


Fig. 18.25 : Torque on the current carrying loop

If a uniform rotation of the loop is desired in a magnetic field, we need to have a constant torque. The couple would be approximately constant if the plane of the coil were always along or parallel to the magnetic field. This is achieved by making the pole pieces of the magnet curved and placing a soft iron core at the centre so as to give a radial field.

The soft iron core placed inside the loop would also make the magnetic field stronger and uniform resulting in greater torque (Fig. 18.26).

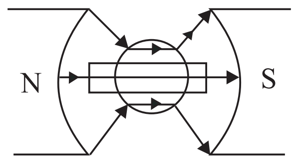


Fig. 18.26: Constant torque on a coil in a radial field

18.6.1 (a) Magnetic Dipole

The term magnetic dipole includes

(i) a current-carrying circular coil of wire, and

(ii) a small bar magnet

The magnetic field due to a magnetic dipole at a point

(i) situated at a distance r on the axis of the dipole is given by :

$$|\mathbf{B}| = \frac{\mu_0}{4\pi}\frac{2\mathbf{M}}{r^3}$$

(ii) situated at a distance r on the equatorial line is given by :

$$|\mathbf{B}| = \frac{\mu_0}{4\pi}\frac{\mathbf{M}}{r^3}$$

This implies that the field has a cylindrical symmetry about the dipole axis.

18.6.1(b) The Torque on a Magnetic Dipole Placed in a uniform magnetic field

We have seen in section18.6 that a current loop behaves as a magnetic dipole.

In 18.6.1 we have also seen that a current loop placed in a uniform magnetic field experiences a torque

$$\tau = \mathbf{M} \times \mathbf{B}$$

$$\Rightarrow \qquad |\tau| = |\mathbf{M}| \times |\mathbf{B}| \sin\theta$$

The direction of τ is normal to the plane containing **M** and **B** and is determined by the right hand cork screw rule.

Note that in all these expressions $\mathbf{M} = NI\mathbf{A}$.
where the direction of **A** is determined by the right hand rule.

18.6.2 Galvanometer

From what you have learnt so far, you can think of an instrument to detect current in any circuit. A device doing precisely this is called a galvanometer, which works on the principle that a current carrying coil, when placed in a magnetic field, experiences a torque.

A galvanometer consists of a coil wound on a non-magnetic frame. A soft iron cylinder is placed inside the coil. The assembly is supported on two pivots attached to springs with a pointer. This is placed between the pole pieces of a horse shoe magnet providing radial field (see Fig. 18.27).

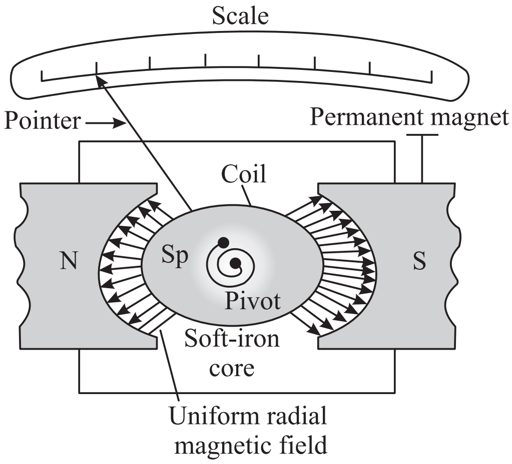


Fig. 18.27 : A moving coil galvanometer

To understand the working of a moving coil galvanometer, we recall that when a current is passed through the coil, it will rotate due to the torque acting on it. The spring sets up a restoring force and hence, a restoring torque. If α is the angle of twist and k is the restoring torque per unit twist or torsional contant, we can write $NBIA \sin\theta = k\,\alpha$. For $\theta = 90^\circ$, $\sin\theta = 1$. So, in the instant case, we can write

$$\therefore \qquad NBIA = k\alpha$$

$$\text{or} \qquad \frac{INBA}{k} = \alpha$$

$$\text{That is,} \qquad I = \frac{k\alpha}{NBA} \qquad (18.20)$$

where $\dfrac{k}{NBA}$ is called galvanometer constant. From this we conclude that

$$\alpha \propto I$$

That is, deflection produced in a galvanometer is proportional to the current flowing through it provided N, B A and k are constant. The ratio α/I is known as current sensitivity of the galvanometer. It is defined as the deflection of the coil per unit current. The more the current stronger the torque and the coil turns more. Galvanometer can be constructed to respond to very small currents (of the order of 0.1μA).

Sensitivity of a galvanometer : In order to have a more sensitive galvanometer,

- *N* should be large;
- *B* should be large, uniform and radial;
- area of the coil should be large; and
- tortional constant should be small.

The values of *N* and *A* cannot be increased beyond a certain limit. Large values of *N* and *A* will increase the electrical and inertial resistance and the size of the galvanometer. **B** can be increased using a strong horse shoe magnet and by mounting the coil on a soft iron core. The value of *k* can be decreased by the use of materials such as quartz or phospher bronze.

18.6.3 An Ammeter and a Voltmeter

(a) Ammeter : An Ammeter is a suitably shunted galvanometer. Its scale is calibrated to give the value of current in the circuit. To convert a galvanometer into an ammeter, a low resistance wire is connected in parallel with the galvanometer. The resistance of the shunt depends on the range of the ammeter and can be calculated as follows :

Let *G* be resistance of the galvanometer and *N* be the number of scale divisions in the galvanometer. Let *k* denote figure of merit or current for one scale deflection in the galvanometer. Then current which produces full scale deflection in the galvanometer is $I_g = Nk$

Let *I* be the maximum current to be measured by the galvanometer.

Refer to Fig. 18.28. The voltage between points *A* and *B* is given by

$$V_{AB} = I_g\, G = (I - I_g)\, S$$

so that

$$S = \frac{I_g\, G}{I - I_g} \qquad (18.21)$$

where *S* is the shunt resistance.

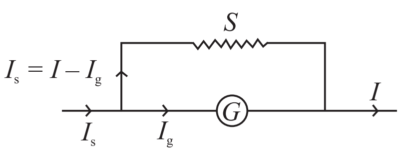


Fig. 18.28 : A shunted galvanometer acts as an ammeter

As G and S are in parallel, the effective resistance R of the ammeter is given by

$$R = \frac{GS}{G+S}.$$

As the shunt resistance is small, the combined resistance of the galvanometer and the shunt is very low and hence, ammeter resistance is lower than that of the galvanometer. An ideal ammeter has almost negligible resistance. That is why when it is connected in series in a circuit, all the current passes through it without any observable drop.

(b) Voltmeter : A voltmeter is used to measure the potential difference between two points in a circuit. We can convert a galvanometer into a voltmeter by connecting a high resistance in series with the galvanometer coil, as shown in Fig 18.29. The value of the resistance depends on the range of voltmeter and can be calculated as follows :

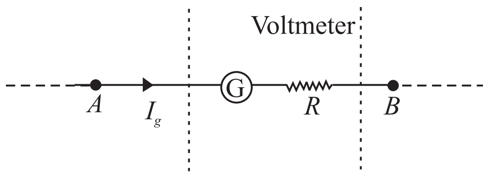


Fig. 18.29 : Galvanometer as a voltmeter

A high resistance, say R is connected in series with the galvanometer coil. If the potential difference across AB is V volt, then total resistance of the voltmeter will be $G + R$. From Ohm's law, we can write

$$I_g (G + R) = V$$

or $$G + R = \frac{V}{I_g}$$

$\Rightarrow$ $$R = \frac{V}{I_g} - G \qquad (18.22)$$

This means that if a resistance R is connected in series with the coil of the galvanometer, it works as a voltmeter of range 0-V volts.

Now the same scale of the galvanometer which was recording the maximum potential $I_g \times G$ before conversion will record the potential V after conversion into voltmeter. The scale can be calibrated accordingly. The resistance of the voltmeter is higher than the resistance of galvanometer. Effective resistance of the voltmeter, is given by

$$R_V = R + G$$

The resistance of an ideal voltmeter is infinite. It is connected in parallel to the points across which potential drop is to be measured in a circuit. It will not draw

Notes

any current. But the galvanometer coil deflects. Seems impossible! Think about it.

Example 18.6 : A circular coil of 30 turns and radius 8.0 cm, carrying a current of 6.0 A is suspended vertically in a uniform horizontal magnetic field of magnitude 1.0 T. The field lines make an angle of 90º with the normal to the coil. Calculate the magnitude of the counter torque that must be applied to prevent the coil from turning.

Solution : Here, $N = 30$, $I = 6.0$ A, $B = 1.0$ T, $\theta = 90º$, $r = 8.0$ cm $= 8 \times 10^{-2}$ m.

Area (A) of the coil $= \pi r^2 = \dfrac{22}{7} \times (8 \times 10^{-2})^2 = 2.01 \times 10^{-2}$ m^2

$$\therefore \quad \text{Torque} = N\,I\,B\,A \sin\theta$$

$$= 30 \times 6 \times 1{:}\,0 \times (2.01 \times 10\mathbf{-2}) \times \sin 90º$$

$$= 30 \times 6 \times (2.01 \times 10^{-2})$$

$$= 3.61 \text{ Nm}$$

Example 18.7 : A galvanometer with a coil of resistance 12.0 Ω shows a full scale deflection for a current of 2.5 mA. How will you convert it into (a) an ammeter of range 0 – 2A, and (b) voltmeter of range 0 – 10 volt ?

Solution : (a) *Here,* $G = 12.0\ \Omega$, $I_g = 2.5$ mA $= 2.5 \times 10^{-3}$ A, and $I = 2$A.
From Eqn. (18.21), we have

$$S = \frac{I_g G}{I - I_g}$$

$$= \frac{2.5 \times 10^{-3} \times 12}{2 - 2.5 \times 10^{-3}}$$

$$= 15 \times 10^{-3}\ \Omega$$

So, for converting the galvanometer into an ammeter for reading 0 – 2V, a shunt of $15 \times 10^{-3}\ \Omega$ resistance should be connected parallel to the coil.

(b)For conversion into voltmeter, let R be the resistance to be connected in series.

$$R = \frac{V}{I_g} - G$$

$$= \frac{10}{2.5 \times 10^{-3}} - 12 = 4000 - 12$$

$$= 3988\ \Omega$$

Thus, a resistance of 3988 Ω should be connected in series to convert the galvanometer into voltmeter.

INTEXT QUESTIONS 18.5

1. What is radial magnetic field ?
2. What is the main function of a soft iron core in a moving coil galvanometer?
3. Which one has the lowest resistance - ammeter, voltmeter or galvanometer? Explain.
4. A galvanometer having a coil of resistance 20 Ω needs 20 mA current for full scale deflection. In order to pass a maximum current of 3A through the galvanometer, what resistance should be added and how ?

WHAT YOU HAVE LEARNT

- Every magnet has two poles. These are inseparable.
- The term magnetic dipole may imply (i) a magnet with dipole moment $\mathbf{M} = m\ell$ (ii) a current carrying coil with dipole moment $\mathbf{M} = NIA$
- Magnetic field at the axis of a magnetic dipole is given by $\mathbf{B} = \dfrac{\mu_0}{4\pi}\dfrac{2\mathbf{M}}{x^3}$ and

  on the equatorial line by $\mathbf{B} = \dfrac{\mu_0}{4\pi}\dfrac{2\mathbf{M}}{x^3}$.
- A magnetic dipole behaves the same way in a uniform magnetic field as an electric dipole does in a uniform electric field, i.e., it experience no net force but a torque

  $\tau = \mathbf{M} \times \mathbf{B}$.
- Earth has a magnetic field which can be completely described in terms of three basic quantities called elements of earth's magnetic field :

  – angle of inclination,
  – angle of declination,and
  – horizontal component of earth's field.
- Every current carrying conductor develops a magnetic field around it. The magnetic field is given by Biot-Savart's Law :

  $$|dB| = \frac{\mu_0 I}{4\pi}\frac{dl\sin\theta}{r^2}$$

- Unit of magnetic field is tesla.
- Field at the centre of a flat coil carrying current is given by $|\mathbf{B}| = \dfrac{\mu_0 I}{2r}$.

  Ampere's circuital law gives the magnitude of the magnetic field around a conductor $\oint \mathbf{B} \cdot \mathbf{dl} = \mu_0 I$
- The Lorentz force on a moving charge q is $\mathbf{F} = q\,(\boldsymbol{v} \times \mathbf{B})$ and its direction is given by Fleming's left hand rule.
- The mechanical force on a wire of length L and carrying a current of I in a magnetic field $\mathbf{B}$ is $\mathbf{F} = \mathbf{B}\, I\, L$.
- Mutual force per unit length between parallel straight conductors carrying currents I_1 and I_2 is given by $\dfrac{F}{L} = \dfrac{\mu_0\, I_1 I_2}{2\pi r}$.
- Magnetic field due to a toroid, $B = \dfrac{\mu_0\, Ni}{2\pi r}$
- A charged particle traces a circular path of radius $R = \dfrac{m\upsilon}{Bq}$.
- Cyclotron is a device used to accelerate charged particles to high velocities.
- Cyclotron frequency $\nu_c = \dfrac{Bq}{2m\pi}$
- A current loop behaves like a magnetic dipole.
- A current carrying coil placed in a magnetic field experiences a torque given by

  $\tau = N\,B\;\; I\,A \sin\theta$

  $\quad = N\,B\,I\,A$, (if $\theta = 90°$)
- Galvanometer is used to detect electric current in a circuit.
- An ammeter is a shunted galvanometer and voltmeter is a galvanometer with a high resistance in series. Current is measured by an ammeter and potential difference by a voltmeter.

TERMINAL EXERCISES

1. A small piece of the material is brought near a magnet. Complete the following by filling up the blanks by writing Yes or No.

Material	Repulsion		Attraction	
	weak	strong	weak	strong
Diamagnetic				
Paramagnetic				
Ferromagnetic				

2. You have to keep two identical bar magnets packed together in a box. How will you pack and why?

   N S   OR   N   S

   N S   S   N

3. The magnetic force between two poles is 80 units. The separation between the poles is doubled. What is the force betweem them?

4. The length of a bar magnet is 10 cm and the area of cross-section is 1.0 cm^2. The magnetization I = 10^2 A/m. Calculate the pole strength.

5. Two identical bar magnets are placed on the same line end to end with north pole facing north pole. Draw the lines of force, if no other field is present.

6. The points, where the magnetic field of a magnet is equal and opposite to the horizontal component of magnetic field of the earth, are called neutral points

   (a) Locate the neutral points when the bar magnet is placed in magnetic meridian with north pole pointing north.

   (b) Locate the neutral points when a bar magnet is placed in magnetic meridian with north pole pointing south.

7. If a bar magnet of length 10 cm is cut into two equal pieces each of length 5 cm then what is the pole strength of the new bar magnet compare to that of the old one.

8. A 10 cm long bar magnet has a pole strength 10 A.m. Calculate the magnetic field at a point on the axis at a distance of 30 cm from the centre of the bar magnet.

9. How will you show that a current carrying conductor has a magnetic field arround it? How will you find its magnitude and direction at a particular place ?

10. A force acts upon a charged particle moving in a magnetic field, but this force does not change the speed of the particle, Why ?

11. At any instant a charged particle is moving parallel to a long, straight current carrying wire. Does it experience any force ?

12. A current of 10 ampere is flowing through a wire. It is kept perpendicular to a magnetic field of 5T. Calculate the force on its 1/10 m length.

13. A long straight wire carries a current of 12 amperes. Calculate the intensity of the magnetic field at a distance of 48 cm from it.

14. Two parallel wire, each 3m long, are situated at a distance of 0.05 m from each other. A current of 5A flows in each of the wires in the same direction. Calculate the force acting on the wires. Comment on its nature ?

15. The magnetic field at the centre of a 50cm long solenoid is 4.0×10^{-2} NA^{-1} m^{-1} when a current of 8.0A flows through it, calculate the number of turns in the solenoid.

16. Of the two identical galvanometer one is to be converted into an ammeter and the other into a milliammeter. Which of the shunts will be of a larger resistance ?

17. The resistance of a galvanometer is 20 ohms and gives a full scale deflection for 0.005A. Calculate the value of shunt required to change it into an ammeter to measure 1A. What is the resistance of the ammeter ?

18. An electron is moving in a circular orbit of radius 5×10^{-11} m at the rate of 7.0×10^{15} revolutions per second. Calculate the magnetic field **B** at the centre of the orbit.

19. Calculate the magnetic field at the centre of a flat circular coil containing 200 turns, of radius 0.16m and carrying a current of 4.8 ampere.

20. Refer to Fig. 18.30 and calculate the magnetic field at *A, B* and *C*.

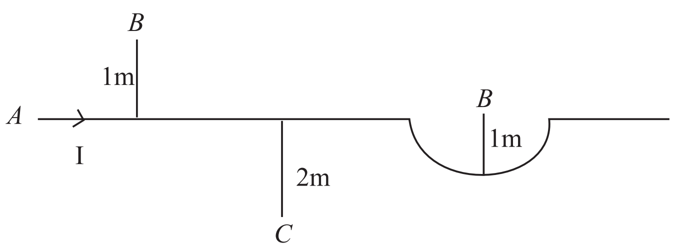


Fig. 18.30

ANSWERS TO INTEXT QUESTIONS

18.1

1. Suspend the magnet with a thread at its centre of mass. Let it come to equilibrium. The end of the magnet which points towards geographical north is its north pole.

2. Bring the ends of any two bars closer together. If there is attraction between them, one of the bars is a magnet and the other is an iron bar. Now lay down one of these bars on the table and strock along its length with the other. If uniform force is experienced, the bar in hand is a magnet and that on the table is iron piece. If non-uniform force is experienced, reverse is the case.

3. Suspending one of the bar magnets with thread, we can find its south pole. Then the end of the second magnet, which is repelled by the first, is its south pole.

18.2

1. (i)electrical (ii) magnetic as well as electrical.

2. A conductor in equilibrium is neutral i.e. it has no net electrical current. Due to their random motion, thermal electrons cancel the magnetic fields produced by them.

3. In first case length of wire $l_1 = 2\pi r$ In second case length of wire $l_2 = (2\pi r_2)2$.

   But $l_1 = l_2$

   $$\therefore \quad 2\pi r = 4\pi r_2 \quad \Rightarrow \quad r_2 = \frac{r}{2}$$

   Using $|\mathbf{B}| = \dfrac{\mu_0 nI}{2r}$

   $$|\mathbf{B}_1| = \frac{\mu_0 I}{2r}, \qquad |\mathbf{B}_2| = \frac{\mu_0 .2.I}{2\times\frac{r}{2}} = \frac{2\mu_0 I}{r} = 4\,\mathbf{B}$$

   That is, the magnetic **B** at the centre of a coil with two turns is four times stronger than the field in first case.

18.3

1. *c*

2. Both laws specify magnetic field due to current carrying conductors.

3. (i) *B*, (ii) *A*, (iii) *C*.

4. $B = \mu_0 \dfrac{n}{\ell} I \Rightarrow 4\pi \times \dfrac{10^{-7}\times n}{0.1m} \times 3\text{A} = 0.002$ or $n = \dfrac{.0002\times 10^{7}}{12\pi} = 50$ turns

18.4

1. The nature of the force will be attractive because the stream of protons is equivalent to electrons in the opposite direction.

2. The force exerted by a magnetic field on a moving charge is perpendicular to the motion of the charge and the work done by the force on the charge is zero. So the KE of the charge does not change. In an electric field, the deflection is in the direction of the field. Hence the field accelerates it in the direction of field lines.

3. The direction of current in each turn of the spring is the same. Since parallel currents in the same direction exert force of attraction, the turns will come closer and the body shall be lifted upward, whatever be the direction of the current in the spring.

18.5

1. Radial magnetic field is one in which plane of the coil remains parallel to it.

2. This increases the strength of magnetic field due to the crowding of magnetic lines of force through the soft iron core, which in turn increases the sensitivity of the galvanometer.

3. Ammeter has the lowest resistance whereas voltmeter has the highest resistance. In an ammeter a low resistance is connected in parallel to the galvanometer coil whereas in a voltmeter, a high resistance is connected in series with it.

4. A low resistance R_s should be connected in parallel to the coil :

$$R_s = \frac{G\,I_g}{I - I_g} = \frac{20 \times 20 \times 10^{-3}}{3 - 20 \times 10^{-3}} = 0.13\ \Omega$$

Answers To Problems in Terminal Exercise

1. 10^{-2} T m^{-1}

7. same.

8. 2.3×10^{-6}T

12. 5 N

13. 5 μN

14. attractive force of 10^{-4} N m^{-1}

15. $\frac{625}{\pi}$ turns.

17. 0.1Ω.

18. 4.48 πT

19. 1.2π mT

20. $B_A = 2 \times 10^{-7}$ T, $B_B = \pi \times 10^{-7}$ T and $B_C = 10^{-7}$ T.

312en19

19

ELECTROMAGNETIC INDUCTION AND ALTERNATING CURRENT

Electricity is the most convenient form of energy available to us. It lights our houses, runs trains, operates communication devices and makes our lives comfortable. The list of electrical appliances that we use in our homes is very long. Have you ever thought as to how is electricity produced?

Hydro-electricity is produced by a generator which is run by a turbine using the energy of water. In a coal, gas or nuclear fuel power station, the turbine uses steam to run the generator. Electricity reaches our homes through cables from the town substation. Have you ever visited an electric sub-station? What are the big machines installed there? These machines are called transformers. Generators and transformers are the devices, which basically make electricity easily available to us. These devices are based on the principle of electromagnetic induction.

In this lesson you will study electromagnetic induction, laws governing it and the devices based on it. You will also study the construction and working of electric generators, transformers and their role in providing electric power to us. A brief idea of eddy current and its application will also be undertaken in this chapter.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain the phenomenon of electromagnetic induction with simple experiments;*
- *explain Faraday's and Lenz's laws of electromagnetic induction;*
- *explain eddy currents and its applicaitons;*

- *describe the phenomena of self-induction and mutual induction;*
- *describe the working of ac and dc generators;*
- *derive relationship between voltage and current in ac circuits containing a (i) resistor, (ii) inductor, and or (iii) capacitor;*
- *analyse series LCR circuits; and*
- *explain the working of transformers and ways to improve their efficiency.*

19.1 ELECTROMAGNETIC INDUCTION

In the previous lesson you have learnt that a steady current in a wire produces a steady magnetic field. Faraday initially (and mistakenly) thought that a steady magnetic field could produce electric current . Some of his investigations on magnetically induced currents used an arrangement similar to the one shown in Fig.19.1. A current in the coil on the left produces a magnetic field concentrated in the iron ring. The coil on the right is connected to a galvanometer *G*, which can indicate the presence of an induced current in that circuit. It is observed that there is no deflection in *G* for a steady current flow but when the switch S in the left circuit is closed, the galvanometer shows deflection for a moment. Similarly, when switch *S* is opened, momentary deflection is recorded but in opposite direction. It means that current is induced only when the magnetic field due to the current in the circuit on the left ***changes***.

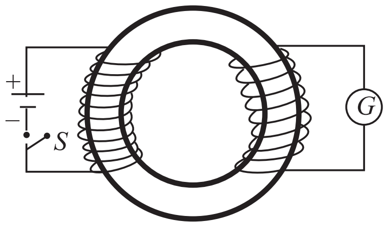


Fig. 19.1: Two coils are wrapped around an iron ring. The galvanometer *G* deflects for a moment when the switch is opened or closed.

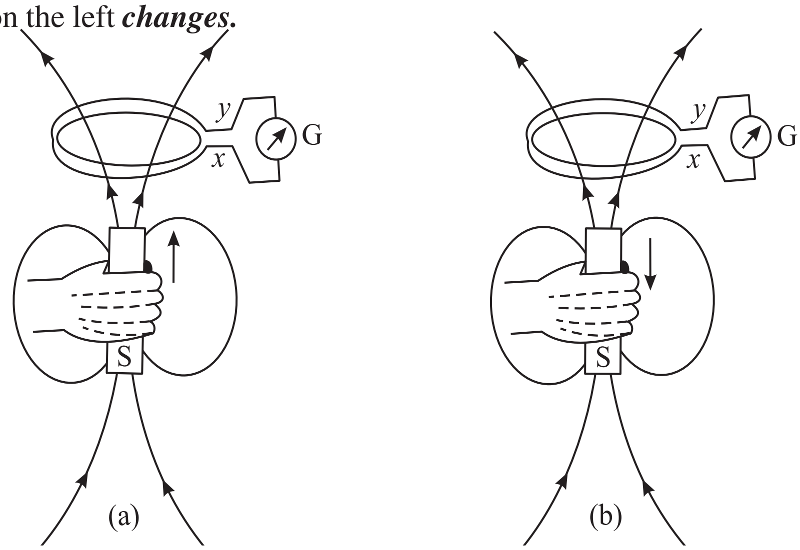
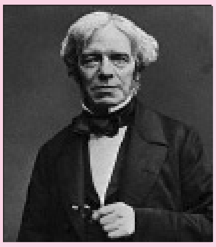


Fig. 19.2 : **a)** A current is induced in the coil if the magnet moves towards the coil, and **b)** the induced current has opposite direction if the magnet moves away from the coil.

The importance of a change can also be demonstrated by the arrangement shown in Fig.19.2. If the magnet is at rest relative to the coil, no current is induced in the coil. But when the magnet is moved towards the coil, current is induced in the direction indicated in Fig. 19.2a. Similarly, if the magnet is moved away from the coil, the a current is induced in the opposite direction, as shown in Fig.19.2b. Note that in both cases, the magnetic field changes in the neighbourhood of the coil. An induced current is also observed to flow through the coil, if this is moved relative to the magnet. The presence of such currents in a circuit implies the existence of an ***induced electromotive force (emf)*** across the free ends of the coil, i.e., x and y.

This phenomenon in which a magnetic field induces an emf is termed as ***electromagnetic induction***. Faraday's genius recognised the significance of this work and followed it up. The quantitative description of this phenomenon is known as Faraday's law of electromagnetic induction. We will discuss it now.



Michael Faraday (1791-1867)

British experimental scientist Michael Faraday is a classical example of a person who became great by shear hardwork, perseverance, love for science and humanity. He started his carrier as an apprentice with a book binder, but utilized the opportunity to read science books that he received for binding. He sent his notes to Sir Humphry Davy, who immediately recognised the talent in the young man and appointed him his permanent assistant in the Royal Institute.

Sir Humphry Davy once admitted that the greatest discovery of his life was Michael Faraday. And he was right because Faraday made basic discoveries which led to the electrical age. It is because of his discoveries that electrical generators, transformers, electrical motors, and electolysis became possible.

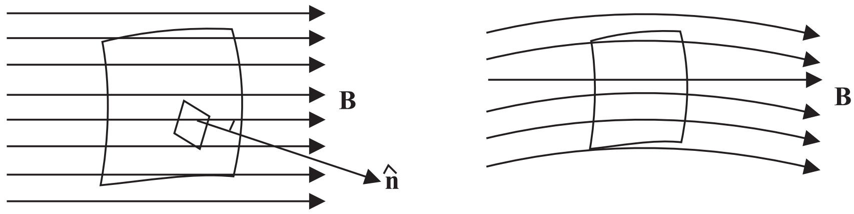
19.1.1 Faraday's Law of Electromagnetic Induction

The relationship between the changing magnetic field and the induced emf is expressed in terms of magnetic flux ϕ_B linked with the surface of the coil. You will now ask: What is magnetic flux? To define ***magnetic flux*** ϕ_B refer to Fig. 19.3a, which shows a typical infinitesimal element of area **ds**, into which the given surface can be considered to be divided. The direction of **ds** is normal to the surface at that point. By analogy with electrostatics, we can define the magnetic flux $d\phi_B$ for the area element **ds** as

$$d\phi_B = \mathbf{B}.\mathbf{ds} \qquad (19.1a)$$

The magnetic flux for the entire surface is obtained by summing such contributions over the surface. Thus,

$$d\phi_B = \sum \mathbf{B} \cdot \mathbf{ds} \qquad (19.1b)$$



(a) (b)

Fig. 19.3: a) The magnetic flux for an infinitesimal area ds is given by $d\phi_B$ = B.ds, and **b)** The magnetic flux for a surface is proportional to the number of lines intersecting the surface.

The SI unit of magnetic flux is **weber** (Wb), where 1 Wb = 1 Tm^2.

In analogy with electric lines and as shown in Fig.19.3b, the number of magnetic lines intersecting a surface is proportional to the magnetic flux through the surface.

Faraday's law states that *an emf is induced across a loop of wire when the magnetic flux linked with the surface bound by the loop changes with time. The magnitude of induced emf is proportional to the rate of change of magnetic flux*. Mathematically, we can write

$$|\varepsilon| = \frac{d\phi_B}{dt} \qquad (19.3)$$

From this we note that weber (Wb), the unit of magnetic flux and volt (V), the unit of emf are related as 1V = 1Wb s^{-1}.

Now consider that an emf is induced in a closely wound coil. Each turn in such a coil behaves approximately as a single loop, and we can apply Faraday's law to determine the emf induced in each turn. Since the turns are in series, the total induced emf ε_r in a coil will be equal to the sum of the emfs induced in each turn. We suppose that the coil is so closely wound that the magnetic flux linking each turn of the coil has the same value at a given instant. Then the same emf ε is induced in each turn, and the total induced emf for a coil with *N* turns is given by

$$|\varepsilon_r| = N|\varepsilon| = N\left(\frac{d\phi_B}{dt}\right) \qquad (19.4)$$

where ϕ_B is the magnetic flux linked with a single turn of the coil.

Let us now apply Faraday's law to some concrete situations.

Example 19.1 : The axis of a 75 turn circular coil of radius 35mm is parallel to a uniform magnetic field. The magnitude of the field changes at a constant rate from 25mT to 50 mT in 250 millisecond. Determine the magnitude of induced emf in the coil in this time interval.

Solution : Since the magnetic field is uniform and parallel to the axis of the coil, the flux linking each turn is given by

$$\phi_B = B\pi R^2$$

where R is radius of a turn. Using Eq. (19.4), we note that the induced emf in the coil is given by

$$|\varepsilon_r| = N\frac{d\phi_B}{dt} = N\frac{d(B\pi R^2)}{dt} = N\,\pi\,R^2\frac{dB}{dt} = N\,\pi\,R^2\left(\frac{B_2 - B_1}{t}\right)$$

Hence, the magnitude of the emf induced in the coil is

$$|\varepsilon_r| = 75\pi\ (0.035\text{m})^2\ (0.10\text{Ts}^{-1}) = 0.030\text{V} = 30\text{mV}$$

This example explains the concept of emf induced by a time changing magnetic field.

Example 19.2 : Consider a long solenoid with a cross-sectional area 8cm^2 (Fig. 19.4a and 19.4b). A time dependent current in its windings creates a magnetic field $B(t) = B_0 \sin 2\pi\nu t$. Here B_0 is constant, equal to 1.2 T. and ν, the frequency of the magnetic field, is 50 Hz. If the ring resistance $R = 1.0\Omega$, calculate the emf and the current induced in a ring of radius r concentric with the axis of the solenoid.

Solution : We are told that magnetic flux

$$\phi_B = B_0 \sin 2\pi\nu t A$$

since normal to the cross sectional area of the solenoid is in the direction of magnetic field.

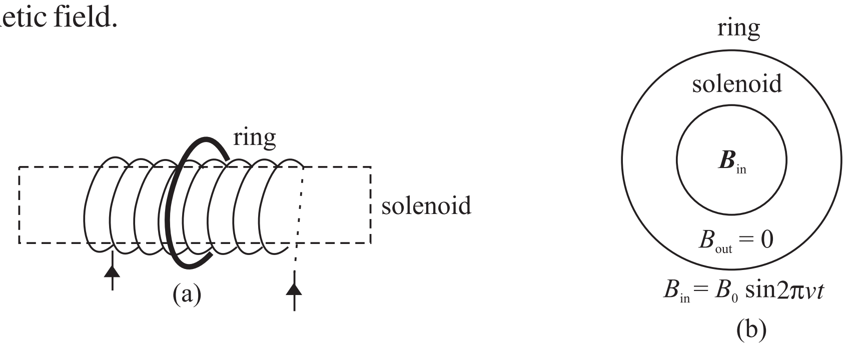


Fig.19.4 : **a)** A long solenoid and a concentric ring outside it, and **b)** cross-sectional view of the solenoid and concentric ring.

$$\text{Hence } |\varepsilon| = \frac{d\phi_B}{dt} = 2\pi\nu A B_0 \cos 2\pi\nu t.$$

$$= 2\pi\ .\ (50\text{s}^{-1})\ (8 \times 10^{-4}\text{m}^2)\ (1.2\ \text{T})\cos 2\pi\nu t$$

$$= 0.3 \cos 2\pi\nu t \text{ volts}$$

$$= 0.3 \cos 100\pi t \text{ V}$$

The current in the ring is $I = \varepsilon/R$. Therefore

$$I = \frac{(0.3\cos 100\pi t)\ \text{V}}{(1.0\Omega)}$$

$$= +0.3 \cos 100\pi t\ \ \text{A}$$

INTEXT QUESTIONS 19.1

1. A 1000 turn coil has a radius of 5 cm. Calculate the emf developed across the coil if the magnetic field through the coil is reduced from 10 T to 0 in (a) 1s (b) 1ms.

2. The magnetic flux linking each loop of a 250-turn coil is given by $\phi_B(t) = A + Dt^2$, where A = 3 Wb and D = 15 Wbs^{-2} are constants. Show that a) the magnitude of the induced emf in the coil is given by $\varepsilon = (2ND)t$, and b) evaluate the emf induced in the coil at t = 0s and t = 3.0s.

3. The perpendicular to the plane of a conducting loop makes a fixed angle θ with a spatially uniform magnetic field. If the loop has area S and the magnitude of the field changes at a rate dB/dt, show that the magnitude of the induced emf in the loop is given by $\varepsilon = (dB/dt)\, S \cos\theta$. For what orientation(s) of the loop will ε be a) maximum and b) minimum?

19.1.2 Lenz's Law

Consider a bar magnet approaching a conducting ring (Fig.19.5a). To apply Faraday's law to this system, we first choose a positive direction with respect to the ring. Let us take the direction from O to Z as positive. (Any other choice is fine, as long as we are consistent.) For this configuration, the positive normal for the area of the ring is in the z-direction and the magnetic flux is negative. As the distance between the conducting ring and the N-pole of the bar magnet decreases, more and more field lines go though the ring, making the flux more and more negative. Thus $d\phi_B/dt$ is negative. By Faraday's law, ε is positive relative to our chosen direction. The current I is directed as shown.

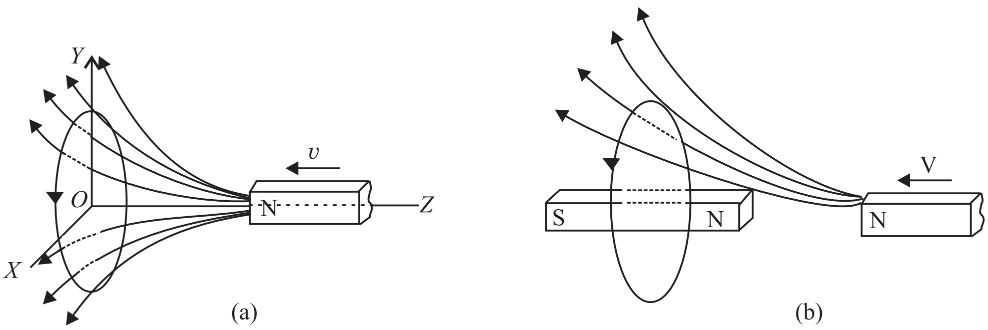


Fig.19.5: a) A bar magnet approaching a metal ring, and **b)** the magnetic field of the induced current opposes the approaching bar magnet.

The current induced in the ring creates a secondary magnetic field in it. This induced magnetic field can be taken as produced by a bar magnet, as shown in Fig.19.5 (b). Recall that induced magnetic field repels or opposes the original magnetic field. This opposition is a consequence of the law of conservation of energy, and is formalized as Lenz's law. **When a current is induced in a conductor, the direction of the current will be such that its magnetic effect opposes the change that induced it.**

The key word in the statement is 'oppose'-it tells us that we are not going to get something for nothing. When the bar magnet is pushed towards the ring, the current induced in the ring creates a magnetic field that opposes the change in flux. The magnetic field produced by the induced current repels the incoming magnet. If we wish to push the magnet towards the ring, we will have to do work on the magnet. This work shows up as electrical energy in the ring. Lenz's law thus follows from the law of conservation of energy. We can express the combined form of Faraday's and Lenz's laws as

$$\varepsilon = -\frac{d\phi}{dt} \qquad (19.5)$$

The negative sign signifies opposition to the cause.

As an application of Lenz's law, let us reconsider the coil shown in Example 19.2. Suppose that its axis is chosen in vertical direction and the magnetic field is directed along it in upward direction. To an observer located directly above the coil, what would be the sense of the induced emf? It will be clockwise because only then the magnetic field due to it (directed downward by the right-hand rule) will oppose the changing magnetic flux. You should learn to apply Lenz's law before proceeding further. Try the following exercise.

19.1.3 Eddy currents

We know that the induced currents are produced in closed loops of conducting wires when the magnetic flux associated with them changes. However, induced currents are also produced when a solid conductor, usually in the form of a sheet or plate, is placed in a changing magnetic field. Actually, induced closed loops of currents are set up in the body of the conductor due to the change of flux linked with it. These currents flow in closed paths and in a direction perpendicular to the magnetic flux. These currents are called eddy currents as they look like eddies or whirlpools and also sometimes called Foucault currents as they were first discovered by Foucault.

The direction of these currents is given by Lenz's law according to which the direction will be such as to oppose the flux to which the induced currents are due. Fig. 19.1.3 shows some of the eddy currents in a metal sheet placed in an increasing magnetic field pointing into the plane of the paper. The eddy currents are circular and point in the anticlockwise direction.

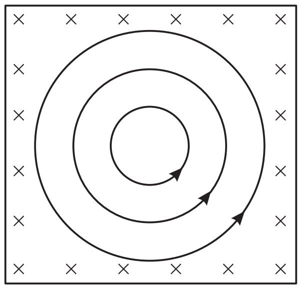


Fig. 19.1.3

The eddy currents produced in metallic bodies encounter little resistance and, therefore, have large magnitude. Obviously, eddy currents are considered undesirable in many electrical appliances and machines as they cause appreciable energy loss by way of heating. Hence, to reduce these currents, the metallic bodies are not taken in one solid piece but are rather made in parts or strips, called lamination, which are insulated from one another.

Eddy currents have also been put to some applications. For example, they are used in induction furnaces for making alloys of different metals in vacuum. They are also used in electric brakes for stopping electric trains.

INTEXT QUESTIONS 19.2

1. The bar magnet in Fig.19.6 moves to the right. What is the sense of the induced current in the stationary loop *A*? In loop *B*?

2. A cross-section of an ideal solenoid is shown in Fig.19.7. The magnitude of a uniform magnetic field is increasing inside the solenoid and **B** = 0 outside the solenoid. In which conducting loops is there an induced current? What is the sense of the current in each case?

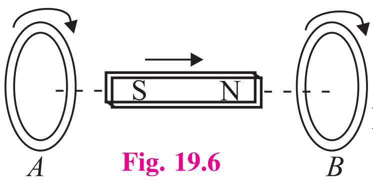


Fig. 19.6

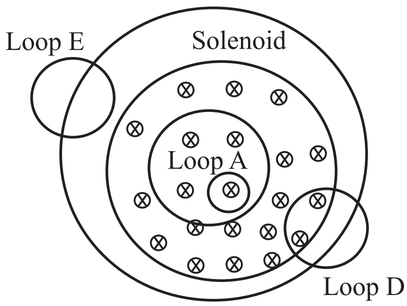


Fig. 19.7

3. A bar magnet, with its axis aligned along the axis of a copper ring, is moved along its length toward the ring. Is there an induced current in the ring? Is there an induced electric field in the ring? Is there a magnetic force on the bar magnet? Explain.

4. Why do we use laminated iron core in a transformer.

19.2 INDUCTANCE

When current in a circuit changes, a changing magnetic field is produced around it. If a part of this field passes through the circuit itself, current is induced in it. Now suppose that another circuit is brought in the neighbourhood of this circuit. Then the magnetic field through that circuit also changes, inducing an emf across it. Thus, induced emfs can appear in these circuits in two ways:

- By changing current in a coil, the magnetic flux linked with each turn of the coil changes and hence an induced emf appears across that coil. This property is called ***self-induction***.
- for a pair of coils situated close to each other such that the flux associated with one coil is linked through the other, a changing current in one coil induces an emf in the other. In this case, we speak of ***mutual induction*** of the pair of coils.

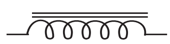
19.2.1 Self-Inductance

Let us consider a loop of a conducting material carrying electric current. The current produces a magnetic field **B**. The magnetic field gives rise to magnetic flux. The total magnetic flux linking the loop is

$$d\phi = \mathbf{B}\cdot d\mathbf{s}$$

In the absence of any external source of magnetic flux (for example, an adjacent coil carrying a current), the Biot-Savart's law tells us that the magnetic field and hence flux will be proportional to the current (I) in the loop, i.e.

$$\phi \propto I \quad \text{or} \quad \phi = LI \qquad (19.6)$$

where L is called self-inductance of the coil. The circuit elements which oppose change in current are called ***inductors.*** These are in general, in the form of coils of varied shapes and sizes. The symbol for an ***inductor*** is . If the coil is wrapped around an iron core so as to enhance its magnetic effect, it is symbolised by putting two lines above it, as shown here . The inductance of an indicator depends on its geometry.

(a) Faraday's Law in terms of Self-Inductance: So far you have learnt that if current in a loop changes, the magnetic flux linked through it also changes and gives rise to self–induced emf between the ends. In accordance with Lenz's law, the self-induced emf opposes the change that produces it.

To express the combined form of Faraday's and Lenz's Laws of induction in terms of L, we combine Eqns. (19.5) and (19.6) to obtain

$$\varepsilon = -\frac{d\phi}{dt} = -L\frac{dI}{dt} \qquad (19.7a)$$

$$= -L\left(\frac{I_2 - I_1}{t}\right) \qquad (19.7b)$$

where I_1 and I_2 respectively denote the initial and final values of current at $t = 0$ and $t = \tau$. Using Eqn. (19.7b), we can define the unit of self-inductance:

$$\text{units of } L = \frac{\text{unit of emf}}{\text{units of } dI/dt}$$

$$= \frac{\text{volt}}{\text{ampere / second}}$$

$$= \text{ohm-second}$$

An ohm-second is called a *henry*, (abbreviated H). For most applications, henry is a rather large unit, and we often use millihenry, mH (10^{-3} H) and microhenry μH (10^{-6}H) as more convenient measures.

The self-induced emf is also called the ***back emf***. Eqn.(19.7a) tells us that the ***back emf in an inductor*** depends on the rate of change of current in it and ***opposes the change in current.*** Moreover, since an infinite emf is not possible, from Eq.(19.7b) we can say that an instantaneous change in the inductor current cannot occur. Thus, we conclude that ***current through an inductor cannot change instantaneously.***

The inductance of an inductor depends on its geometry. In principle, we can calculate the self-inductance of any circuit, but in practice it is difficult except for devices with simple geometry. A solenoid is one such device used widely in electrical circuits as inductor. Let us calculate the self-inductance of a solenoid.

(b) Self-inductance of a solenoid : Consider a long solenoid of cross-sectional area A and length ℓ, which consists of N turns of wire. To find its inductance, we must relate the current in the solenoid to the magnetic flux through it. In the preceding lesson, you used Ampere's law to determine magnetic field of a long solenoid:

$$|\mathbf{B}| = \mu_0 n I$$

where $n = N/\ell$ denotes is the number of turns per unit length and I is the current through the solenoid.

The total flux through N turns of the solenoid is

$$\phi = N\,|\mathbf{B}|\,A = \frac{\mu_0 N^2 A I}{\ell} \qquad (19.8)$$

and self-inductance of the solenoid is

$$L = \frac{\phi}{I} = \frac{\mu_0 N^2 A}{\ell} \qquad (19.9)$$

Using this expression, you can calculate self-inductance and back emf for a typical solenoid to get an idea of their magnitudes.

INTEXT QUESTIONS 19.3

1. A solenoid 1m long and 20cm in diameter contains 10,000 turns of wire. A current of 2.5A flowing in it is reduced steadily to zero in 1.0ms. Calculate the magnitude of back emf of the inductor while the current is being reduced.

2. A certain length (ℓ) of wire, folded into two parallel, adjacent strands of length $\ell/2$, is wound on to a cylindrical insulator to form a type of wire-wound non-inductive resistor (Fig.19.8). Why is this configuration called non-inductive?

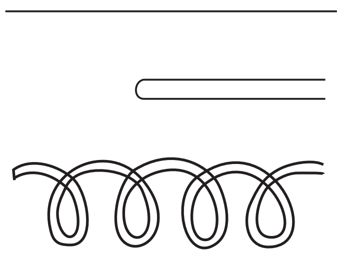


Fig.19.8: Wire wound on a cylindrical insulator

3. What rate of change of current in a 9.7 mH solenoid will produce a self-induced emf of 35mV?

19.2.2 LR Circuits

Suppose that a solenoid is connected to a battery through a switch (Fig.19.9). Beginning at $t = 0$, when the switch is closed, the battery causes charges to move in the circuit. *A* solenoid has inductance (L) and resistance (R), and each of these influence the current in the circuit. The inductive and resistive effects of a solenoid are shown schematically in Fig.19.10. The inductance (L) is shown in series with the resistance (R). For simplicity, we assume that total resistance in the circuit, including the internal resistance of the battery, is represented by R. Similarly, L includes the self-inductance of the connecting wires. A circuit such as that shown in Fig.19.9, containing resistance and inductance in series, is called an *LR* circuit.

The role of the inductance in any circuit can be understood qualitatively. As the current $i(t)$ in the circuit increases (from $i = 0$ at $t = 0$), a self-induced emf $\varepsilon = -L\, di/dt$ is produced in the inductance whose sense is opposite to the sense of the increasing current. This opposition to the increase in current prevents the current from rising abruptly.

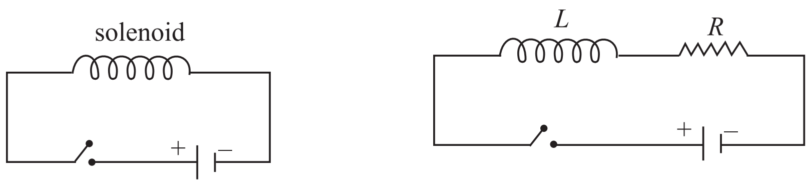


Fig. 19.9: LR Circuit

If there been no inductance in the circuit, the current would have jumped immediately to the maximum value defined by ε_0/R. But due to an inductance coil in the circuit, the current rises gradually and reaches a steady state value of ε_0/R as $t \rightarrow \tau$. The time taken by the current to reach about two-third of its steady state value is equal to by L/R, which is called the ***inductive time constant*** of the circuit. Significant changes in current in an LR circuit cannot occur on time scales much shorter that L/R. The plot of the current with time is shown in Fig. 19.10.

You can see that greater the value of L, the larger is the back emf, and longer it takes the current to build up. (This role of an inductance in an electrical circuit is somewhat similar to that of mass in mechanical systems.) That is why while switching off circuits cortaining large inductors, you should the mindful of back emf. The spark seen while turning off a switch connected to an electrical appliance such as a fan, computer, geyser or an iron, essentially arises due to back emf.

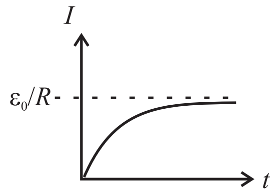


Fig.19.10 : Variation of current with time in a LR–ciruci t.

INTEXT QUESTIONS 19.4

1. A light bulb connected to a battery and a switch comes to full brightness almost instantaneously when the switch is closed. However, if a large inductance is in series with the bulb, several seconds may pass before the bulb achieves full brightness. Explain why.

2. In an LR circuit, the current reaches 48mA in 2.2 ms after the switch is closed. After sometime the current reaches it steady state value of 72mA. If the resistance in the circuit is 68Ω, calculate the value of the inductance.

19.2.3 Mutual Inductance

When current changes in a coil, a changing magnetic flux develops around it, which may induce emf across an adjoining coil. As we see in Fig. (19.11), the magnetic flux linking each turn of coil B is due to the magnetic field of the current in coil A.

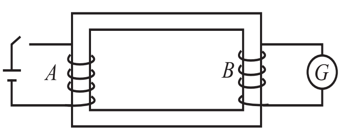


Fig. 19.11 : Mulual inductance of a pair of coils

Therefore, a changing current in each coil induces an emf in the other coil, i.e.

i.e., $$\phi_2 \propto \phi_1 \propto I_1 \Rightarrow \phi_2 = MI_1 \qquad (19.10)$$

where M is called the mutual inductance of the pair of coils. Also back emf induced across the second coil

$$e_2 = -\frac{d\phi}{dt}$$

$$= -M\frac{dI}{dt} = -M\left(\frac{I_2 - I_1}{t}\right) \quad (19.11)$$

where the curent in coil A changes from I_1 to I_2 in t seconds.

The mutual inductance depends only on the geometry of the two coils, if no magnetic materials are nearby. The SI unit of mutual inductance is also henry (H), the same as the unit of self-inductance.

Example 19.3 : A coil in one circuit is close to another coil in a separate circuit. The mutual inductance of the combination is 340 mH. During a 15 ms time interval, the current in coil 1 changes steadily from 28mA to 57 mA and the current in coil 2 changes steadily from 36 mA to 16 mA. Determine the emf induced in each coil by the changing current in the other coil.

Solution : During the 15ms time interval, the currents in the coils change at the constant rates of

$$\frac{di_1}{dt} = \frac{57\text{mA} - 23\text{mA}}{15\text{ms}} = 2.3 \text{ As}^{-1}$$

$$\frac{di_2}{dt} = \frac{16\text{mA} - 36\text{mA}}{15\text{ms}} = -1.3 \text{ As}^{-1}$$

From Eq. (19.11), we note that the magnitudes of the induced emfs are

$$\varepsilon_1 = -(340\text{mH})\,(2.3\text{As}^{-1}) = -0.78 \text{ V}$$

$$\varepsilon_2 = (340\text{mH})\,(1.3\text{As}^{-1}) = 0.44 \text{ V}$$

Remember that the minus signs in Eq. (19.11) refer to the sense of each induced emf.

One of the most important applicances based on the phenomenon of mutual inductance is transformer. You will learn about it later in this lesson. Some commonly used devices based on self-inductance are the choke coil and the ignition coil. We will discuss about these devices briefly. Later, you will also learn that a combination of inductor and capitator acts as a basic oscillator. Once the capacitor is charged, the charge in this arrangement oscillates between its two plates through the inductor.

INTEXT QUESTIONS 19.5

1. Consider the sense of the mutually induced emf's in Fig.19.11, according to an observer located to the right of the coils. (a) At an instant when the current i_1 is increasing, what is the sense of emf across the second coil? (b) At an instant when i_2 is decreasing, what is the sense of emf across the first coil?
2. Suppose that one of the coils in Fig.19.11 is rotated so that the axes of the coils are perpendicular to each other. Would the mutual inductance remain the same, increase or decrease? Explain.

19.3 ALTERNATING CURRENTS AND VOLTAGES

When a battery is connected to a resistor, charge flows through the resistor in one direction only. If we want to reverse the direction of the current, we have to interchange the battery connections. However, the magnitude of the current will remain constant. Such a current is called *direct current*. But a current whose magnitude changes continuously and direction changes periodically, is said to be an *alternating current* (Fig. 19.12).

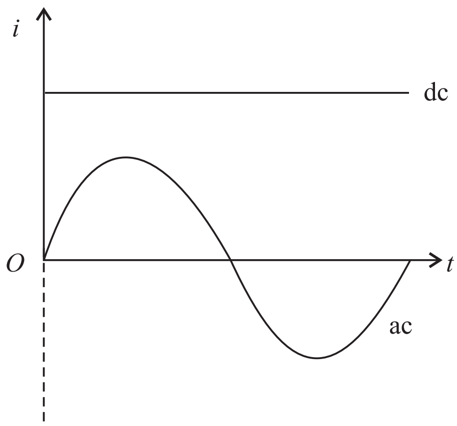


Fig. 19.12 : dc and ac current waveforms

In general, alternating voltage and currents are mathematically expressed as

$$V = V_m \cos \omega t \qquad (19.12a)$$

and

$$I = I_m \cos \omega t \qquad (19.12b)$$

V_m and I_m are known as the ***peak values*** of the alternating voltage and current respectively. In addition, we also define the root mean square (*rms*) values of V and I as

$$V_{rms} = \frac{V_m}{\sqrt{2}} = 0.707\ V_m \qquad (19.13a)$$

$$I_{rms} = \frac{I_m}{\sqrt{2}} = 0.707\, I_m \qquad (19.13b)$$

The relation between V and I depends on the circuit elements present in the circuit. Let us now study a.c. circuits containing (i) a resistor (ii) a capacitor, and (iii) an inductor only

George Westinghouse

(1846-1914)

If ac prevails over dc all over the world today, it is due to the vision and efforts of George Westinghouse. He was an American inventor and enterpreneur having about 400 patents to his credit. His first invention was made when he was only fifteen year old. He invented air brakes and automatic railway signals, which made railway traffic safe.

When Yugoslav inventor Nicole Tesla (1856-1943) presented the idea of rotating magnetic field, George Westinghouse immediately grasped the importance of his discovery. He invited Tesla to join him on very lucrative terms and started his electric company. The company shot into fame when he used the energy of Niagra falls to produce electricity and used it to light up a town situated at a distance of 20km.

19.3.1 AC Source Connected to a Resistor

Refer to Fig. 19. 13 which shows a resistor in an ac circuit. The instantaneous value of the current is given by the instantaneous value of the potential difference across the resistor divided by the resistance.

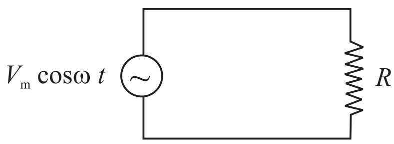


Fig. 19.13 : An ac circuit containing a resistor

$$I = \frac{V}{R}$$

$$= \frac{V_m \cos \omega t}{R} \qquad (19.14a)$$

The quantity V_m/R has units of volt per ohm,(i.e., ampere). It represents the maximum value of the current in the circuit. The current changes direction with time, and so we use positive and negative values of the current to represent the two possible current directions. Substituting I_m, the maximum current in the circuit, for V_m/R in Eq. (19.14a), we get

$$I = I_m \cos \omega t \qquad (19.14b)$$

Fig.19.14 shows the time variation of the potential difference between the ends of a resistor and the current in the resistor. Note that the potential difference and current are in phase i.e., the peaks and valleys occur at the same time.

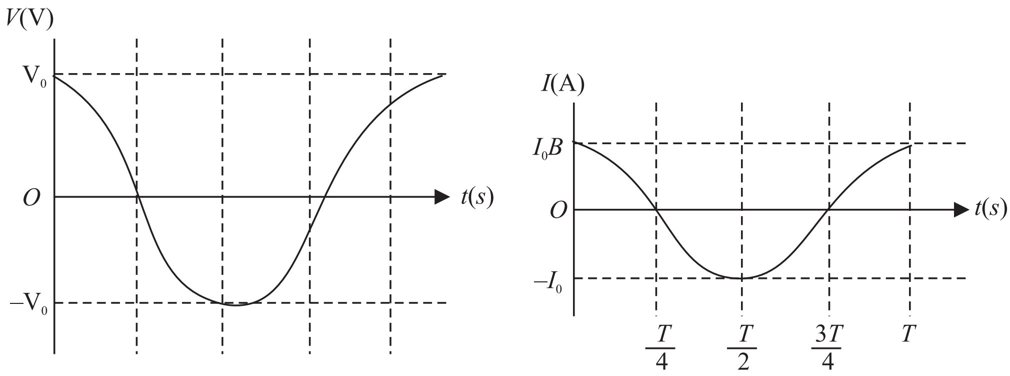


Fig. 19.14 : Time variation of current and voltage in a purely resistive circuit

In India, we have V_m = 310V and v = 50 Hz. Therefore for R = 10 Ω, we get

$$V = 310 \cos (2\pi\, 50t)$$

and
$$I = \frac{310}{10} \cos (100\pi t)$$
$$= 31 \cos (100\pi t)\text{A}$$

Since V and I are proportional to cos (100πt), the average current is zero over an integral number of cycles.

The ***average power*** $P = I^2R$ developed in the resistor is not zero, because square of instantaneous value of current is always positive. As I^2, varies periodically between zero and I^2, we can determine the average power, P_{av}, for single cycle:

$$P_{av} = (I^2R)_{av} = R(I^2)_{av} = R\left(\frac{I_m^2 + 0}{2}\right)$$

$$P_{av} = R\left(\frac{I_m^2}{2}\right) = R\, I^2_{rms} \qquad (19.15)$$

Note that the same power would be produced by a constant *dc* current of value ($I_m/\sqrt{2}$) in the resistor. It would also result if we were to connect the resistor to a potential difference having a constant value of $V_m/\sqrt{2}$ volt. The quantities $I_m/\sqrt{2}$ and $V_m/\sqrt{2}$ are called the rms values of the current and potential difference. The term rms is short for root-mean-square, which means "the square root of the mean value of the square of the quantity of interest." For an electric outlet in an Indian home where V_m = 310V, the rms value of the potential difference is

$$V_{rms} = V_m/\sqrt{2} \simeq 220\text{V}$$

This is the value generally quoted for the potential difference. Note that when potential difference is 220 V, the peak value of a.c voltage is 310V and that is why it is so fatal.

INTEXT QUESTIONS 19.6

1. In a light bulb connected to an ac source the instantaneous current is zero two times in each cycle of the current. Why does the bulb not go off during these times of zero current?
2. An electric iron having a resistance 25Ω is connected to a 220V, 50 Hz household outlet. Determine the average current over the whole cycle, peak current, instantaneous current and the rms current in it.
3. Why is it necessary to calculate root mean square values of ac current and voltage.

19.3.2 AC Source Connected to a Capacitor

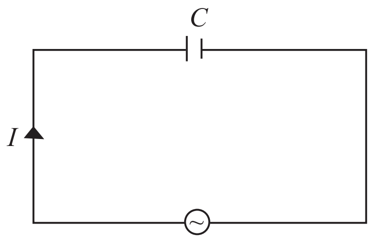


Fig.19.15 : Capacitor in an ac circuit

Fig.19.15 shows a capacitor connected to an ac source. From the definition of capacitance, it follows that the instantaneous charge on the capacitor equals the instantaneous potential difference across it multiplied by the capacitance ($q = CV$). Thus, we can write

$$q = CV_m \cos \omega t \qquad (19.16)$$

Since $I = dq/dt$, we can write

$$I = -\omega C V_m \sin \omega t \qquad (19.17)$$

Time variation of V and I in a capacitive circuit is shown in Fig.19.16.

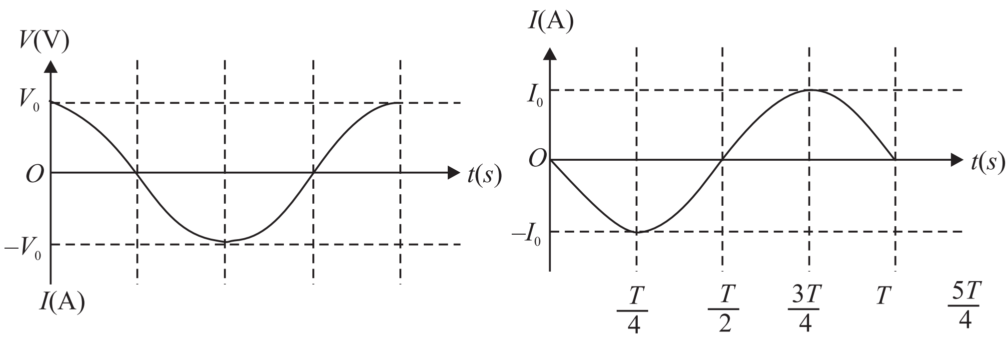


Fig.19.16: Variation in V and I with time in a capacitive circuit

Unlike a resistor, the current I and potential difference V for a capacitor are not in phase. The first peak of the current-time plot occurs one quarter of a cycle before the first peak in the potential difference-time plot. *Hence we say that the capacitor current leads capacitor potential difference by one quarter of a period.* One quarter of a period corresponds to a phase difference of π/2 or 90°. Accordingly, we also say that the potential difference lags the current by 90°.

Rewriting Eq. (19.17) as

$$I = -\frac{V_m}{1/(\omega C)} \sin \omega t \qquad (19.18)$$

and comparing Eqs. (19.14a) and (19.18), we note that $(1/\omega C)$ must have units of resistance. The quantity $1/\omega C$ is called the capacitive reactance, and is denoted by the symbol X_C :

$$X_C = \frac{1}{\omega C}$$

$$= \frac{1}{2\pi\nu C} \qquad (19.19)$$

Capacitive reactance is a measure of the extent to which the capacitor limits the ac current in the circuit. It depends on capacitance and the frequency of the generator. The capacitive reactance decreases with increase in frequency and capacitance. Resistance and capacitive reactance are similar in the sense that both measure limitations to ac current. But unlike resistance, capacitive reactance depends on the frequency of the ac (Fig.19.17). The concept of capacitive reactance allows us to introduce an equation analogous to the equation $I = V/R$:

$$I_{rms} = \frac{V_{rms}}{X_C} \qquad (19.20)$$

The instantaneous power delivered to the capacitor is the product of the instantaneous capacitor current and the potential difference :

$$P = VI$$

$$= -\omega C V^2 \sin \omega t \cos \omega t$$

$$= -\frac{1}{2}\omega C V^2 \sin 2\omega t \qquad (19.21)$$

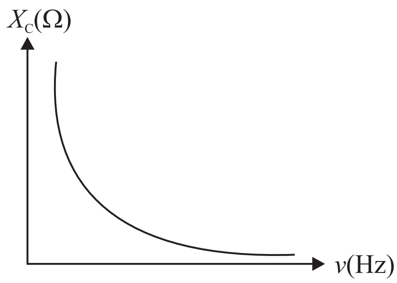


Fig.19.17 : Frequency variation of capacitive reactance

The sign of P determines the direction of energy flow with time. When P is positive, energy is stored in the capacitor. When P is negative, energy is released by the capacitor. Graphical representations of V, I, and P are shown in Fig.19.18. Note that whereas both the current and the potential difference vary with angular frequency w, the power varies with angular frequency 2w. The average power is zero. The electric energy stored in the capacitor during a charging cycle is completely recovered when the capacitor is discharged. On an average, there is no energy stored or lost in the capacitor in a cycle.

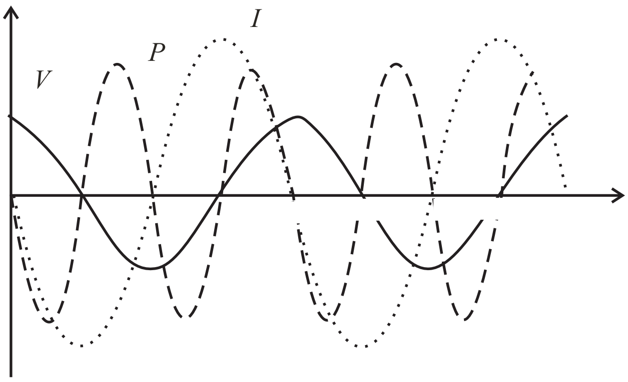


Fig.19.18 : Time variation of *V*, *I* and *P*

Example 19.5 : A 100 μF capacitor is connected to a 50Hz ac generator having a peak amplitude of 220V. Calculate the current that will be recorded by an rms ac ammeter connected in series with the capacitor.

Solution : The capacitive reactance of a capacitor is given by

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi(50\,\text{rad}\,\text{s}^{-1})(100\times10^{-6}\,\text{F})} = 31.8\Omega$$

Assuming that ammeter does not influence the value of current because of its low resistance, the instantaneous current in the capacitor is given by

$$I = \frac{V}{X_C}\cos\omega t = \frac{220}{31.8}\cos\omega t$$

$$= (-6.92\cos\omega t)\text{ A}$$

The rms value of current is

$$I_{rms} = \frac{I_m}{\sqrt{2}}$$

$$= \frac{6.92}{\sqrt{2}}$$

$$= 4.91\text{A}$$

Now answer the following questions.

INTEXT QUESTIONS 19.7

1. Explain why current in a capacitor connected to an ac generator increases with capacitance.
2. A capacitor is connected to an ac generator having a fixed peak value (V_m) but variable frequency. Will you expect the current to increase as the frequency decreases?
3. Will average power delivered to a capacitor by an ac generator to be zero? Justify your answer.
4. Why do capacitive reactances become small in high frequency circuits, such as those in a TV set?

19.3.3 AC Source Connected to an Inductor

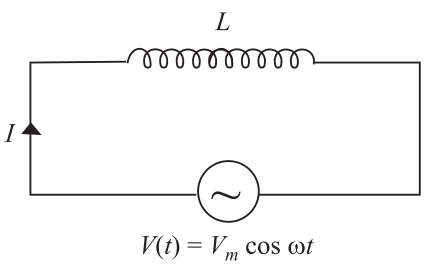


Fig.19.19 : An ac generator connected to an inductor

We now consider an ideal (zero-resistance) inductor connected to an ac source. (Fig. 19.19). If V is the potential difference across the inductor, we can write

$$V(t) = L\frac{dI(t)}{dt} = V_m \cos \omega t \qquad (19.22)$$

To integrate Eqn. (19.22) with time, we rewrite it as

$$\int dI = \frac{V_m}{L}\int \cos\omega t \, dt \; .$$

Since integral of cos x is sin x, we get

$$I(t) = \frac{V_m}{\omega L} \sin \omega t + \text{constant} \qquad (19.23a)$$

When $t = 0$, $I = 0$. Hence constant of integration becomes zero. Thus

$$I(t) = \frac{V_m}{\omega L} \sin \omega t \qquad (19.23b)$$

To compare $V(t)$ and $I(t)$ let us take $V_m = 220$V, $\omega = 2\pi(50)$ rads^{-1}, and $L = 1H$. Then

$$V(t) = 220 \cos (2\pi\, 50t) \text{ volt}$$

$$I(t) = \frac{220}{2\pi.50} \sin (2\pi\, 50t) = 0.701 \sin (2\pi\, 50t) \text{ ampere}$$

Fig.19.20. Shows time variation of V and I The inductor current and potential difference across it are not in phase. In fact the potential difference peaks one-quarter cycle before the current. We say that incase of an inductor current lags

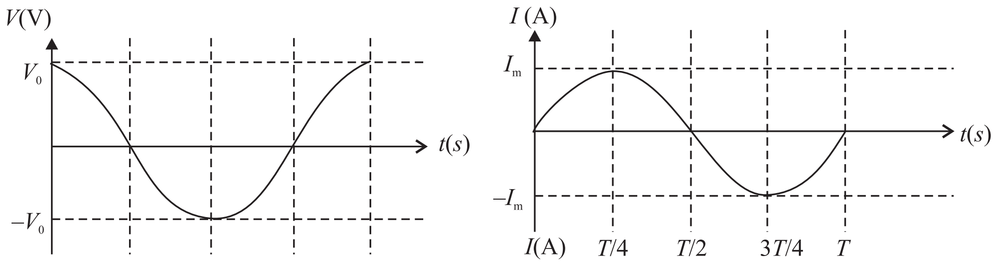


Fig. 19.20 : Time variation of the potential differeence across an inductor and the currentflowing through it. These are not in phase

the potential difference by $\pi/2$ rad (or 90°). This is what we would expect from Lenz's law. Another way of seeing this is to rewrite Eq. (19.23b) as

$$I = \frac{V_m}{\omega L}\cos\left(\omega t - \frac{\pi}{2}\right)$$

Because $V = V_m \cos \omega t$, the phase difference ($-\pi/2$) for I means that current lags village by $\pi/2$. This is in contrast to the current in a capacitor, which leads the potential difference. For an inductor, the current lags the potential difference.

The quantity ωL in Eq.(19.23b) has units of resistance and is called ***inductive reactance.*** It is denoted by symbol X_L :

$$X_L = \omega L = 2\pi \nu L \tag{19.24}$$

Like capacitive reactance, the inductive reactance, X_L, is expressed in ohm. ***Inductive reactance*** is a measure of the extent to which the inductor limits ac current in the circuit. It depends on the inductance and the frequency of the generator. Inductive reactance increases, if either frequency or inductance increases. (This is just the opposite of capacitive reactance.) In the limit frequency goes to zero, the inductive reactance goes to zero. But recall that as $\omega \rightarrow 0$, capcative reactance tends to infinity (see Table 19.1). Because inductive effects vanish for a dc source, such as a battery, zero inductive reactance for zero frequency is consistent with the behaviour of an inductor connected to a dc source. The frequency variation of X_L is shown in Fig. 19.21.

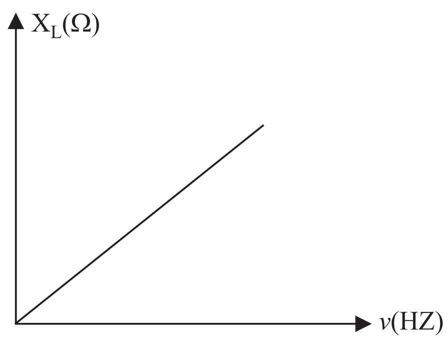


Fig.19.21 : The reactance of an inductor ($X_L = 2\pi\nu L$) as a function of frequency. The inductive reactance increases as the frequency increases.

Table 19.1: Frequency response of passive circuit elements

Circuit element	Opposition to flow of current	Value at low-frequency	Value at high-frequency
Resistor	R	R	R
Capacitor	$X_C = \frac{1}{\omega C}$	∞	0
Inductor	$X_L = \omega L$	0	∞

The concept of inductive reactance allows us to introduce an inductor analog in the equation $I = V/R$ involving resistance R :

$$I_{rms} = \frac{V_{rms}}{X_L} \tag{19.25}$$

The instantaneous power delivered to the inductor is given by

$$P = VI$$

$$= \frac{V_m^2}{\omega L} \sin \omega t \cos \omega t = \frac{V_m^2}{2\omega L} \sin 2\,\omega t \tag{19.26}$$

Graphical representations of V, I and P for an inductor are shown in Fig. 19.21. Although both the current and the potential difference vary with angular frequency, the power varies with twice the angular frequency. The average power delivered over a whole cycle is zero. Energy is alternately stored and released as the magnetic field alternately grows and windles. decays

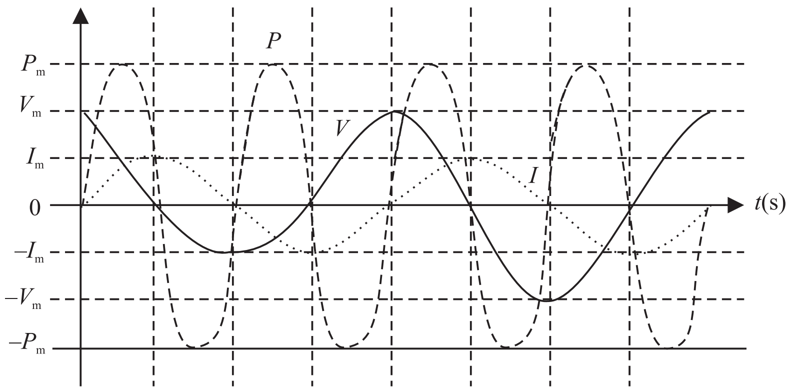


Fig. 19.21: Time variation of potential difference, current and power in an inductive circuit

Example 19.6 : An air cored solenoid has a length of 25cm and diameter of 2.5cm, and contains 1000 closely wound turns. The resistance of the coil is measured to be 1.00Ω. Compare the inductive reactance at 100Hz with the resistance of the coil.

Solution : The inductance of a solenoid, whose length is large compared to its diameter, is given by

$$L = \frac{\mu_0 N^2 \pi a^2}{\ell}$$

where N denotes number of turns, a is radius, and ℓ is length of the solenoid. On substituting the given values, we get

$$L = \frac{(4\pi \times 10^{-7})\ \text{Hm}^{-1}\ (1000)^2\ \pi(0.0125)^2\,\text{m}^2}{0.25\text{m}}$$

$$= 2.47 \times 10^{-3}\ \text{H}$$

The inductive reactance at a frequency of 100Hz is

$$X_L = \omega L \quad = 2\pi \left(100 \frac{\text{rad}}{\text{s}}\right) (2.47 \times 10^{-3})\ \text{H}$$

$$= 1.55\Omega$$

Thus, inductive reactance of this solenoid at 100Hz is comparable to the intrinsic (ohmic) resistance R. In a circuit diagram, it would be shown as

$L = 2.47$ H and $R = 1.00\ \Omega$

You may now like to test your understanding of these ideas.

INTEXT QUESTIONS 19.8

1. Describe the role of Lenz's law when an ideal inductor is connected to an ac generator.

2. In section 19.3.1, self-inductance was characterised as electrical inertia. Using this as a guide, why would you expect current in an inductor connected to an ac generator to decrease as the self-inductance increases?

19.3.4 Series LCR Circuit

Refer to Fig. 19.22. It shows a circuit having an inductor L, a capacitor C and a resistor R in series with an ac source, providing instantaneous emf $E = E_m \sin \omega t$. The current through all the three circuit elements is the same in amplitude and phase but potential differences across each of them, as discussed earlier, are not in the same phase. Note that

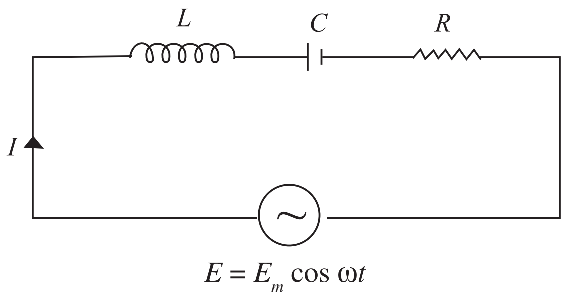


Fig. 19.22: A series *LCR* circuit

(i) The potential difference across the resistor $V_R = I_0 R$ and it will be in-phase with current.

(ii) Amplitude of P.D. across the capacitor $V_C = I_0 X_C$ and it lags behind the current by an angle $\pi/2$ and (iii) amplitude of P.D. across the inductor $V_L = I_0 X_L$ and it leads the current by an angle $\pi/2$.

Due to different phases, we can not add voltages algebraically to obtain the resultant peak voltage across the circuit. To add up these voltages, we draw a phasor diagram showing proper phase relationship of the three voltages (Fig.19.23). The diagram clearly shows that voltages across the inductor and capacitor are in opposite phase and hence net voltage across the reactive components is $(V_L - V_C)$. The resultant peak voltage across the circuit is therefore given by

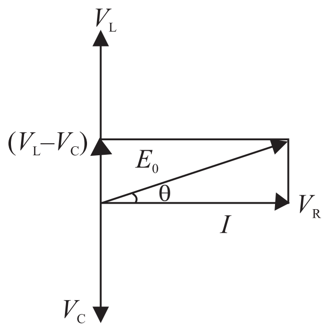


Fig. 19.23: Phasor diagram of voltages across *LCR*.

$$E_0 = \sqrt{(V_L - V_C)^2 + V_R^2}$$

$$= \sqrt{I_0^2\left\{(X_L - X_C)^2 + R^2\right\}}$$

or

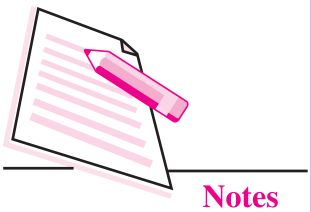
$$\frac{E_0}{I_0} = \sqrt{(X_L - X_C)^2 + R^2}$$

The opposition to flow of current offered by a *LCR* circuit is called its *impedance*. The impedance of the circuit is given by

$$Z = \frac{E_{rms}}{I_{rms}} = \frac{E_0}{I_0} = \sqrt{(X_L - X_C)^2 + R^2} = \sqrt{\left(2\pi\nu L - \frac{1}{2\pi\nu C}\right)^2 + R^2} \qquad (19.27)$$

Hence, the rms current across an *LCR* circuit is given by

$$I_{rms} = \frac{E_{rms}}{Z}$$



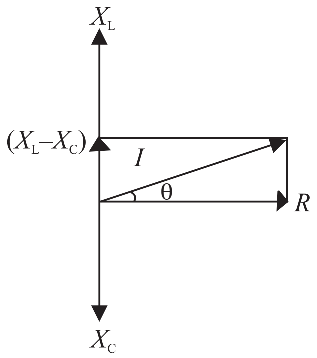


Fig. 19.24 : Phasor diagram for *Z*

Also from Fig.19.23 it is clear that in *LCR* circuit, the emf leads (or lags) the current by an angle ϕ, given by

$$\tan\phi = \frac{V_L - V_C}{V_R} = \frac{X_L I_0 - X_C I_0}{RI_0} = \frac{X_L - X_C}{R} \quad (19.28)$$

This means that R, X_L, X_C and Z can also be represented on a phasor diagram similar to voltage (Fig.19.24).

Resonance

You now know that inductive reactance (X_L) increases and capacitive reactance (X_C) decreases with increase in frequency of the applied ac source. Moreover, these are out of phase. Therefore, there may be a certain frequency v_r for which $X_L = X_C$:

i.e. $$2\pi v_r L = \frac{1}{2\pi v_r C}$$

$$\Rightarrow \quad v_r = \frac{1}{2\pi\sqrt{LC}} \quad (19.29)$$

This frequency is called *resonance frequency* and at this frequency, impedance has minimum value : $Z_{min} = R$. The circuit now becomes purely resistive. Voltage across the capacitor and the inductor, being equal in magnitude, annul each other. Since a resonant circuit is purely resistive, the net voltage is in phase with current ($\phi = 0$) and maximum current flows through the circuit. The circuit is said to be in resonance with applied ac. The graphs given in Fig.19.25 show the variation of peak value of current in an *LCR* circuit with the variation of the frequency of the applied source. The resonance frequency of a given *LCR* circuit is independent of resistance. But as shown in Fig.19.25, the peak value of current increases as resistance decreases.

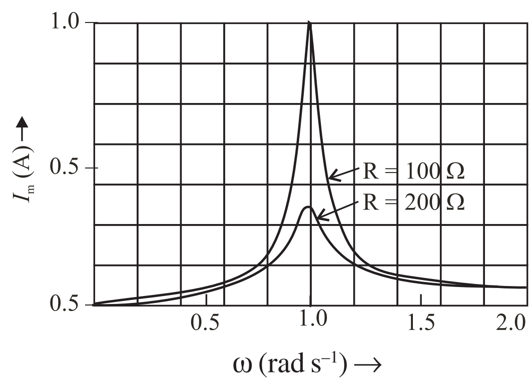


Fig.19.25 : Variation of peak current in a *LCR* circuit with frequency for (i) R = 100 Ω, and (ii) R = 200 Ω

The phenomenon of resonance in *LCR* circuits is utilised to tune our radio/TV receivers to the frequencies transmitted by different stations. The tuner has an inductor and a variable capacitor. We can change the natural frequency of the *L*-*C* circuit by changing the capacitance of the capacitor. When natural frequency of the tuner circuit matches the frequency of the transmitter, the intercepting radio waves induce maximum current in our receiving antenna and we say that particular radio/TV station is tuned to it.

Power in a LCR Circuit

You know that a capacitor connected to an ac source reversibly stores and releases electric energy. There is no net energy delivered by the source. Similarly, an inductor connected to an ac source reversibly stores and releases magnetic energy. There is no net energy delivered by the source. However, an ac generator delivers a net amount of energy when connected to a resistor. Hence, when a resistor, an inductor and a capacitor are connected in series with an ac source, it is still only the resistor that causes net energy transfer. We can confirm this by calculating the power delivered by the source, which could be a generator.

The instantaneous power is the product of the voltage and the current drawn from the source. Therefore, we can write

$$P = VI$$

On substituting for *V* and *I*, we get

$$P = V_m \cos \omega t \left[\frac{V_m}{Z} \cos(\omega t + \phi)\right]$$

$$= \frac{V_m^2}{Z} \cdot \frac{2\cos \omega t \cos(\omega t + \phi)}{2}$$

$$= \frac{V_m^2}{2Z} [\cos\phi + \cos(\omega t + \frac{\phi}{2})] \qquad (19.30)$$

The phase angle ϕ and angular frequency ω play important role in the power delivered by the source. If the impedance Z is large at a particular angular frequency, the power will be small at all times. This result is consistent with the idea that impedance measures how the combination of elements impedes (or limits) ac current. Since the average value of the second term over one cycle is zero, the average power delivered by the source to the circuit is given by

$$\text{Average Power} = \frac{V_m^2}{2Z} \cos\phi \qquad (19.31)$$

$$= \frac{V_m}{\sqrt{2}} \cdot \frac{V_m}{\sqrt{2}Z} \cos\phi = V_{rms} I_{rms} \cos\phi \qquad (19.32)$$

cosϕ is called *power factor* and is given by

$$\cos\phi = \frac{R}{Z} = \frac{R}{\sqrt{R^2 + (X_L - X_C)^2}} \qquad (19.33)$$

The power factor delimits the maximum average power per cycle provided by the generator. In a purely resistive circuit (or in a resonating circuit where $X_L = X_C$), $Z = R$, so that $\cos\phi = \frac{R}{R} = 1$. That is, when $\phi = 0$, the average power dissipated per cycle is maximum: $P_m = V_{rms} I_{rms}$.

On the other hand, in a purely reactive circuit, i.e., when $R = 0$, $\cos\phi = 0$ or $\phi = 90^0$ and the average power dissipated per cycle $P = 0$. That is, the current in a pure inductor or pure capacitor is maintained without any loss of power. Such a current, therefore, is called *wattless current.*

19.4 POWER GENERATOR

One of the most important sources of electrical power is called ***generator***. *A generator is a device that converts mechanical energy into electrical energy with the help of magnetic field.* No other source of electric power can produce as large amounts of electric power as the generator. A conductor or a set of conductors is rotated in a magnetic field and voltage is developed across the rotating conductor due to electromagnetic induction. The energy for the rotation of the conductors can be supplied by water, coal, diesel or gas or even nuclear fuel. Accordingly, we have hydro-generators, thermal generators, and nuclear reactors, respectively.

There are two types of generators

- alternating current generator or A.C. generator also called alternators.
- direct current generator or D.C. generator or dynamo.

Both these generators work on the principle of electromagnetic induction.

19.4.1 A.C. Generator or Alternator

A generator basically consists of a loop of wire rotating in a magnetic field. Refer to Fig.19.26. It shows a rectangular loop of wire placed in a uniform magnetic field. As the loop is rotated along a horizontal axis, the magnetic flux through the loop changes. To see this, recall that the magnetic flux through the loop, as shown in Fig. 19.26, is given by

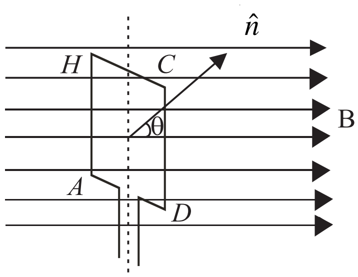


Fig. 19.26 : A loop of wire rotating in a magnetic field.

$$\phi(t) = \mathbf{B}.\hat{\mathbf{n}}A$$

where **B** is the field, $\hat{\mathbf{n}}$ is a unit vector normal to the plane of the loop of area A. If the angle between the field direction and the loop at any instant is denoted by θ, $\phi(t)$ can be written as

$$\phi(t) = AB\cos\theta$$

When we rotate the loop with a constant angular velocity ω, the angle θ changes as

$$\theta = \omega t \qquad (19.34)$$

$$\therefore \qquad \phi(t) = AB\cos\omega t$$

Now, using Faraday's law of electromagnetic induction, we can calculate the emf induced in the loop :

$$\varepsilon(t) = -\frac{d\phi}{dt} = \omega AB\sin\omega t \qquad (19.35)$$

The emf induced across a coil with N number of turns is given by

$$\varepsilon(t) = N\omega AB\sin\omega t \qquad (19.35a)$$

$$= \varepsilon_0 \sin\omega t$$

That is, when a rectangular coil rotates in a uniform magnetic field, the induced emf is sinusoidal.

An A.C. generator consists of four main parts (see in Fig.19.27 : (i) Armature, (ii) Field magnet, (iii) Slip-rings, (iv) Brushes.

An armature is a coil of large number of turns of insulated copper wire wound on a cylindrical soft iron drum. It is capable of rotation at right angles to the magnetic field on a rotor shaft passing through it along the axis of the drum. This drum of soft iron serves two purpose : it supports the coil, and increases magnetic induction through the coil. A field magnet is provides to produce a uniform and permanent radial magnetic field between its pole pieces.

Slip Rings provide alternating current generated in armature to flow in the device connected across them through brushs. These are two metal rings to which the two ends of the armatures are connected. These rings are fixed to the shaft. They are insulated from the shaft as well as from each other. Brushes are two flexible metal or carbon rods [B_1 and B_2 (Fig. 19.27)], which are fixed and constantly in touch with revolving rings. It is with the help of these brushes that the current is passed on from the armature and rings to the main wires which supply the current to the outer circuit.

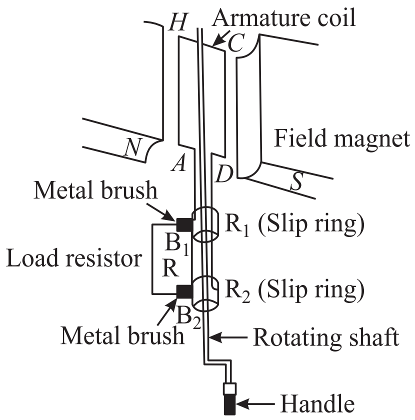


Fig.19.27 : Schematics of an ac generator

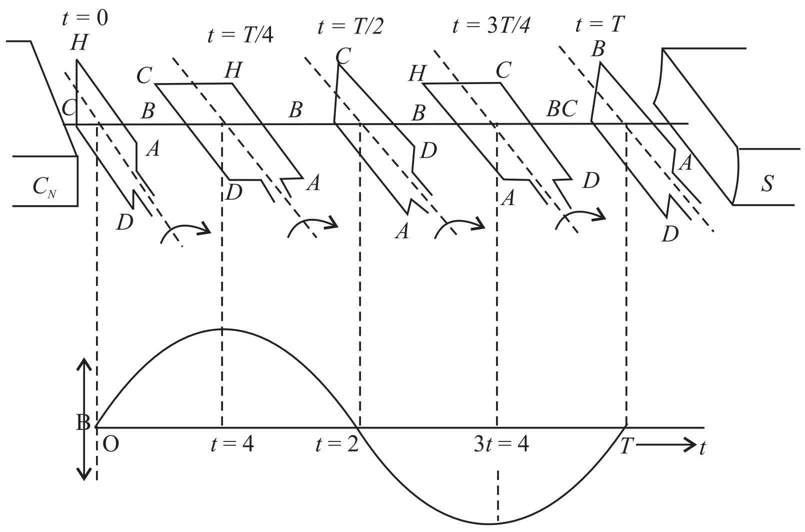


Fig. 19.28 : Working principle of an ac generator

The principle of working of an ac generator is illustrated in Fig.19.28.

Suppose the armature coil *AHCD* rotates in the anticlockwise direction. As it rotates, the magnetic flux linked with it changes and the current is induced in the coil. The direction of the induced current is given by Fleming's right hand rule. Considering the armature to be in the vertical position and its rotation in anticlockwise direction, the wire *AH* moves downward and *DC* moves upwards, the direction of induced emf is from *H* to *A* and *D* to *C* i.e., in the coil it flows along *DCHA*. In the external circuit the current flows along $B_1\ R\ B_2$ as shown in Fig.19.28(a). This direction of current remains the same during the first half turn of the armature. However, during the second half revolution (Fig.19.28(b)), the wire *AH* moves upwards while the wires *CD* moves downwards. The current flows in the direction *AHCD* in the armature coil i.e., the direction of induced current in the coil is reversed. In the external circuit direction is $B_2\ RB_1$. Therefore, the direction of the induced emf and the current changes after every half revolution in the external circuit also. Hence, the current thus produced alternates in each cycle (Fig. 19.28(c)).

The arrangement of slip rings and brushes creates problems of insulation and sparking when large output powers are involved. Therefore, in most practical generators, the field is rotated and the armature (coil) is kept stationary. In such a generator, armature coils are fixed permanently around the inner circumference of the housing of the generator while the field coil pole pieces are rotated on a shaft within the stationary armature.

19.4.2 Dynamo (DC Generator)

A dynamo is a machine in which mechanical energy is changed into electrical energy in the form of direct current. You must have seen a dynamo attached to a bicycle for lighting purpose. In automobiles, dynamo has a dual function for lighting and charging the battery. The essential parts of dynamo are (i) field magnet, (ii) armature, (iii) commutator split rings and (iv) brushes.

Armatures and field magnets differ in dynamo and alternator. In the dynamo, the field magnets are stationary and the armature rotates while in an alternator, armature is stationary (stator) and the field magnet (rotor) rotates.

In a dynamo, ac waveform or the sine wave produced by an a.c. generator is converted into d.c. form by the split ring commutator. Each half of the commuter is connected permanently to one end of the loop and the commutator rotates with the loop. Each brush presses against one segment of the commutator. The brushes remain stationary while the commutator rotates. The brushes press against opposite segments of the commutator and every time the voltage reverses polarity, the split rings change position. This means that one brush always remains positive while the other becomes negative, and a d.c. fluctuating voltage is obtained across the brushes.

A dynamo has almost the same parts as an ac dynamo but it differs from the latter in one respect: In place of slip ring , we put two split rings R_1 and R_2 which are the two half of the same ring, as shown in Fig.19.29(a). The ends of the armature coil are connected to these rings and the ring rotates with the armature and changes the contact with the brushes B_1 and B_2. This part of the dynamo is known as ***commutator***.

When the coil is rotated in the clockwise direction, the current produced in the armature is a.c. but the commutator changes it into d.c. in the outer circuit. In the first half cycle, Fig.19.29(a), current flows along *DCHA*. The current in the external circuit flows along $B_1 L B_2$. In the second half, Fig.19.29(b), current in the armature is reversed and flows along *AHCD* and as the

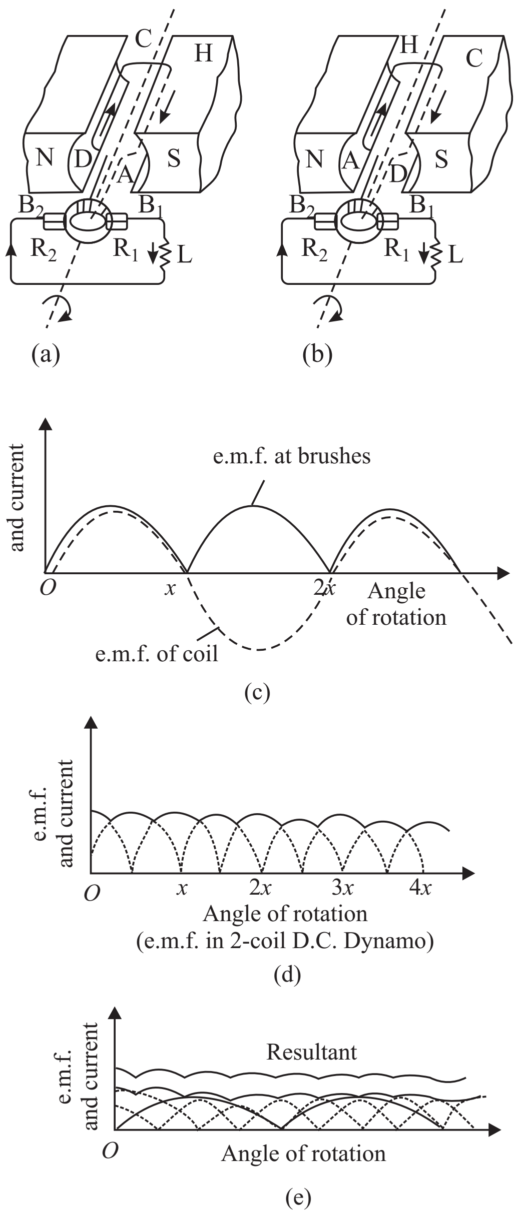


Fig. 19.29 : A dc generator

ring R_1 comes into contact with B_1 to B_2. Thus, current in the external circuit always flows in the same direction. The current produced in the outer circuit is graphically represented in Fig.19.29(c) as the coil is rotated from the vertical position, perpendicular to the magnetic lines of force. The current generated by such a simple d.c. dynamo is unidirectional but its value varies considerably and even falls to zero twice during each rotation of the coil.

One way of overcoming this variation would be to use two coils, mutually at right angles, and to divide the commutator ring into four sections, connected to the ends of the coils. In such a case, both these coils produce emf of the same type but they differ in phase by $\pi/2$. The resultant current or emf is obtained by superposition of the two, as shown in Fig. 19.29(d). In this way, the fluctuations are cosiderally reduced. Similarly, in order to get a steady current, we use a large number of coils, each consisting of good many turns. The commutator ring is divided into as many segments as the number of ends of coils, so that the coils work independently and send current into the outer circuit. The resultant current obtained is shown in Fig.19.29(e) which is practically parallel to the time axes.

INTEXT QUESTIONS 19.9

1. Distinguish between an ac and dc generator.
2. Name the essential parts of a generator?
3. Why do we use a commutator in a dc generator?
4. Where do you find the use of dynamo in daily life?

Low Voltage and Load Shedding

For normal operation of any electrical device, proper voltage is essential. If the voltage supplied by the electric supply company is less than the desired value, we face the problem of low voltage. In fact, low voltage is not as harmful to the appliance as the high voltage. However, due to **low voltage**, most of the appliances do not work properly. To overcome this, use voltage stabilizers. If the low voltage is within the range of the stabilizer, you will get constant voltage. You can use CVT (constant voltage transformers) also to get constant voltage.

As you know, the electricity generated at a power station is transmitted at high voltage to city sub–station. At the sub–station, voltage is reduced using a step down transformer. In order to avoid the danger of burning off the transformers, the supply undertakings try to keep the load on the transformer within the specified rating. If the transformer through which you receive the voltage is heavily loaded (more than the specified value), the supplier will either shed the load by cutting the supply from the power source, or request the consumers to decrease the load by switching off the (heating or cooling) applicances of higher wattages. This process is known as **load shedding.**

In case of load shedding, you can use inverters. Inverters are low frequency oscillator circuits which convert direct current from battery to alternating current of desired value and frequency (230V and 50Hz).

19.5 TRANSFORMER

Transformer is a device that changes (increases or decreases) the magnitude of alternating voltage or current based on the phenomenon of electromagnetic induction. A transformer has at least two windings of insulated copper wire linked by a common magnetic flux but the windings are electrically insulated from one another. The transformer windings connected to a supply source, which may be an ac main or the output of a generator, is called ***primary winding***. The transformer winding connected to the load R_L is called the secondary winding. In the secondary winding, emf is induced when a.c. is applied to the primary. The primary and secondary windings, though electrically isolated from each other, are magnetically coupled with each other.

Basically, a transformer is a device which transfers electric energy (or power) from primary windings to secondary windings. The primary converts the changing electrical energy into magnetic energy. The secondary converts the magnetic energy back into electric energy.

An ideal transformer is one in which

- the resistance of the primary and secondary coils is zero;
- there is no flux leakage so that the same magnetic flux is linked with each turn of the primary and secondary coils; and
- there is no energy loss in the core.

Fig. 19.30 illustrates the configuration of a typical transformer. It consists of two coils, called primary and secondary, wound on a core (transformer). The coils, made of insulated copper wire, are wound around a ring of iron made of isolated laminated sheets instead of a solid core. The laminations minimize eddy currents in iron. Energy loss in a transformer can be reduced by using the laminations of "soft" iron for the core and thick high conductivity wires for the primary and secondary windings.

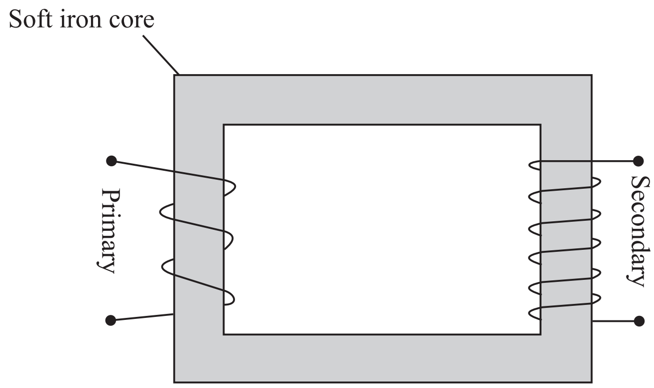


Fig.19.30 : A schematic representation of a transformer

We now discuss the working of a transformer in the following two cases:

(a) **Secondary an open circuit :** Suppose the current in the primary changes the flux through the core at the rate $d\phi/dt$. Then the induced (back) emf in the primary with N_p turns is given by

$$E_p = -N_p \frac{d\phi}{dt}$$

and the induced emf in the secondary widings of N_S turns is

or $$E_s = -N_s \frac{d\phi}{dt}$$

$$\frac{E_p}{E_s} = \frac{N_p}{N_s} \qquad (19.36)$$

(b) **Secondary not an open circuit :** Suppose a load resistance R_L is connected across the secondary, so that the secondary current is I_S and primary current is I_P. If there is no energy loss from the system, we can write

Power input = Power output

or $$E_p I_p = E_s I_s$$

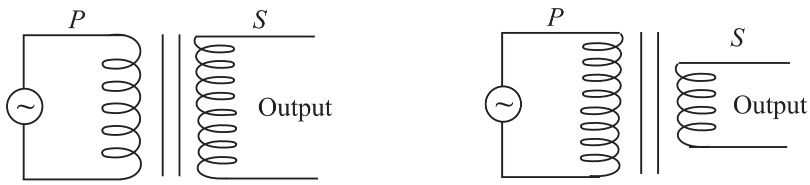
so that $$\frac{I_p}{I_s} = \frac{E_s}{E_p} = \frac{N_p}{N_s} = k. \qquad (19.37)$$

Thus when the induced emf becomes k times the applied emf, the induced current is $\frac{1}{k}$ times the original current. In other words, what is gained in voltage is lost in current.

19.5.1 Types of transformers

There are basically two types of transformers.

(i) ***A step-up transformer*** increases the voltage (decreases the current) in secondary windings. In such transformers (Fig.19.31a) the number of turns in secondary is more than the number of turns in primary.



(a) (b)

Fig. 19.31 : Iron cored **a)** step-up, and **b)** stepdown transformers

(ii) A step-down transformer decreases the voltage (increases the current) in the secondary windings. In such transformers (Fig 19.31b), the number of turns in secondary is less than the number of turns in the primary.

19.5.2 Efficiency of Transformers

While discussing the theory of the transformers we considered an ideal transformer in which there is no power loss. But in practice, some energy is always converted into heat in the core and the windings of the transformer. As a result, the electrical energy output across the secondary is less than the electrical energy input. The efficiency of a transformer is given by

$$\eta = \frac{\text{Energy output}}{\text{Energy input}} \times 100\%$$

$$= \frac{\text{Power output}}{\text{Power input}} \times 100\%$$

The efficiency of a transformer is less than 100%.

In a transformer the energy losses result from

(a) Resistive heating in copper coils - *cooper loss*,

(b) Eddy current losses in form of heating of iron core - *Eddy current loss.*

(c) Magnetization heating of the core during repeated reversal of magnetization - *hysteresis loss*.

(d) Flux leakage from the *core.*

Electrical Power Transmission

You have learnt how electricity is generated using ac or dc generators. You must have come across small units of generating sets in shops, offices and cinema halls. When power goes off, the mains is switched over to generator. In commercial use, generators which produce power of million of watts at about 15kV (kilo volt) is common. These generating plants may be hundreds of kilometers away from your town. Very large mechanical power (kinetic energy) is, therefore, necessary to rotate the rotor which produces magnetic field inside enormously large coils. The rotors are rotated by the turbines. These turbines are driven by different sources of energy.

To minimise loss of energy, power is transmitted at low current in the transmission lines. For this power companies step up voltage using transformers. At a power plant, potential difference is raised to about 330kV. This is accompanied by small current. At the consumer end of the transmission lines, the potential difference is lowered using step down transformers.

You may now like to know how high potential difference used to transmit electrical power over long distances minimises current. We explain this with an example. Suppose electrical power P has to be delivered at a potential difference V by supply lines of total resistance R. The current $I = P/V$ and the loss in the lines is $I^2R = P^2R/V^2$. It means that greater V ensures smaller loss. In fact, doubling V quarters the loss.

Electrical power is, thus, transmitted more economically at high potential difference. But this creates insulation problems and raises installation cost. In a 400kV supergrid, currents of 2500 A are typical and the power loss is about 200kW per kilometer of cable, i.e., 0.02% (percent) loss per kilometer. The ease and efficiency with which alternating potential differences are stepped-up and stepped-down in a transformer and the fact that alternators produces much higher potential difference than d.c. generators (25kV compared with several thousand volts), are the main considerations influencing the use of high alternating rather than direct potential in most situations. However, due to poor efficiency and power thefts, as a nation, we lose about} Rs. 50,000 crore annually.

Example 19.7 : What is the efficiency of a transformer in which the 1880 W of primary power provides for 1730 W of secondary power?

Solution : Given P_{pri}= 1880W and P_{sec} = 1730W. Hence

$$\text{Efficiency} = \frac{P_{sec}}{P_{pri}} \times 100$$

$$\therefore \quad = \frac{1730\,W}{1880\,W} \times 100 = 92\%$$

Thus, the transformer is 92% efficient.

Example 19.8 : A transformer has 100 turns in its primary winding and 500 turns in its secondary windings. If the primary voltage and current are respectively 120V and 3A, what are the secondary voltage and current?

Solution : Given $N_1 = 100$, $N_2 = 500$, $V_1 = 120V$ and $I_1 = 3A$

$$V_2 = \frac{N_2}{N_1} \times V_1 = \frac{500\text{turns}}{100\text{turns}} \times 120\ V = 600V$$

$$I_2 = \frac{N_1}{N_2} \times I_1 = \frac{100\text{turns}}{500\text{turns}} \times 3A = 0.6A$$

INTEXT QUESTIONS 19.10

1. Can a transformer work on dc? Justify your answer.
2. Why does step-up transformer have more turns in the secondary than in the primary?
3. Is the secondary to primary current ratio same as the secondary to primary voltage ratio in a transformer?
4. Toy trains often use a transformer to supply power for the trains and controls. Is this transformer step-up or step-down?

WHAT YOU HAVE LEARNT

- A current is induced in a coil of wire if magnetic flux linking the surface of the coil changes. This is known as the phenomenon of ***electromagnetic induction.***
- The induced emf ε in a single loop is given by ***Faraday's law:***

$$e = \frac{d\phi_B}{dt}$$

  where ϕ_B is the magnetic flux linking the loop.
- According to ***Lenz's Law***, the induced emf opposes the cause which produces it.

- Induced closed loops of current are set up on the body of the conductor (usually a sheet) when it is placed in a changing magnetic field. These currents are called eddy currents.
- If the current changes in a coil, a self-induced emf exists across it.
- For a long, tightly wound solenoid of length ℓ, cross - sectional area A, having N number of turns, the self–inductance is given by

$$L = \frac{\mu_0 N^2 A}{\ell}$$

- Current in an LR circuit takes some time to attain maximum value.
- The changing currents in two nearby coils induce emf mutually.
- In an LC circuit, the charge on the capacitor and the current in the circuit oscillate sinusoidally with the angular frequency ω_0 given by

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

- In an ac circuit, the voltage across the source is given by $V = V_m \cos \omega t$ and current $I = I_m \cos (\omega t + \phi)$
- In a purely resistive ac circuit, the voltage and current are in phase.

  The average power in such a circuit is $P_{av} = \dfrac{I_m^2 R}{2}$

- In a purely capacitive ac circuit, the current leads the voltage by 90^0. The average power in such a circuit is zero.
- In a purely inductive ac circuit, the current lags the voltage by 90^0. The average power in such a circuit is zero.
- In a series LCR circuit, $I_m = \dfrac{V_m}{Z} = \dfrac{V_m}{[R^2 + (X_L - X_C)^2]^{1/2}}$,

  where Z is the impedance of circuit : $Z = [R^2 + (X_L - X_C)^2]^{1/2}$

- For $X_L - X_C = 0$, an ac circuit is purely resistive and the maximum current $I_m = V_m/R$. The circuit is said to be in resonance at $\omega_o = 1/\sqrt{LC}$.
- The average power $P_{av} = V_{rms} . I_{rms} = I_{rms}^2 R$.
- A generator converts mechanical energy into electrical energy. It works on the principle of electromagnetic induction.

- A transformer is a static electrical device which converts an alternating high voltage to low alternating voltage and vice versa.
- The transformers are of two types: Step-up to increase the voltage, and Step-down : to decrease the voltage.
- The secondary to primary voltage ratio is in the same proportion as the secondary to primary turns ratio i.e.

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

- Main sources of power losses in a transformer are heating up of the windings and eddy current
- For transmission of power from a power station to our homes, transformers and transmission lines are used.

TERMINAL EXERCISES

1. Each loop in a 250-turn coil has face area $S = 9.0 \times 10^{-2}$ m^2. (a) What is the rate of change of the flux linking each turn of the coil if the induced emf in the coil is 7.5V? (b) If the flux is due to a uniform magnetic field at 45^0 from the axis of the coil, calculate the rate of change of the field to induce that emf.

2. (a) In Fig.19.32 what is the direction of the induced current in the loop when the area of the loop is decreased by pulling on it with the forces labelled F? B is directed into the page and perpendicular to it.

   (b) What is the direction of the induced current in the smaller loop of Fig.19.31b when a clockwise current as seen from the left is suddenly established in the larger loop, by a battery not shown?

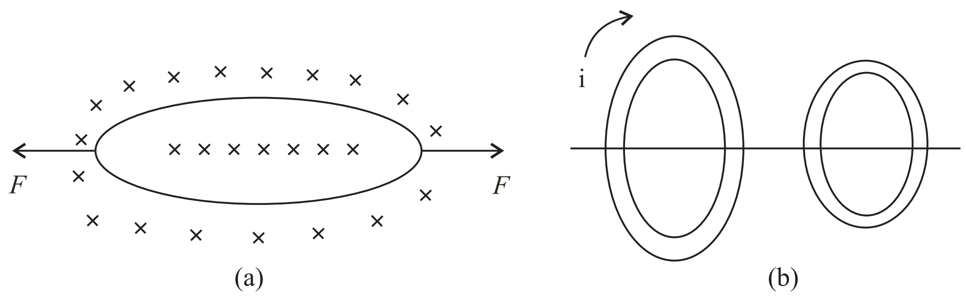


Fig. 19.32

3. (a) If the number of turns in a solenoid is doubled, by what amount will its self-inductance change?

   (b) Patrol in a vehicle's engine is ignited when a high voltage applied to a spark plug causes a spark to jump between two conductors of the plug. This high voltage is provided by an ignition coil, which is an arrangement of two coils wound tightly one on top of the other. Current from the vehicle's battery flows through the coil with fewer turns. This current is interrupted periodically by a switch. The sudden change in current induces a large emf in the coil with more turns, and this emf drives the spark. A typical ignition coil draws a current of 3.0 A and supplies an emf of 24kV to the spark plugs. If the current in the coil is interrupted every 0.10ms, what is the mutual inductance of the ignition coil?

4. (a) Why is the rms value of an ac current always less than its peak value?

   (b) The current in a 2.5μF capacitor connector to an ac source is given by $I = -4.71 \sin 377t$ μA

   Calculate the maximum voltage across the capacitor.

5. (a) Calculate the capacitive reactance (for C = 2 μF) and the inductive reactance (for L = 2 mH) at (i) 25Hz and (ii) 50Hz.

   (b) Calculate the maximum and rms currents in a 22 μH inductor connected to a 5V (rms) 100MHz generator.

6. A series *LCR* circuit with R = 580Ω, L = 31mH, and C =47 nF is driven by an ac source. The amplitude and angular frequency of the source are 65 *V* and 33 krad/s. Determine (a) the reactance of the capacitor, (b) the reactance of the inductor, (c) the impedance of the circuit, (d) the phase difference between the voltage across the source and the current, and (e) the current amplitude. Does current lead behind or lag the voltage across the source?

7. What is electromagnetic induction? Explain Faraday's laws of electromagnetic induction.

8. State Lenz's law. Show that Lenz's law is a consequence of law of conservation of energy.

9. What is self-induction? Explain the physical significance of self-inductance.

10. Distinguish between the self-inductance and mutual-inductance. On what factors do they depend?

11. How much e.m.f. will be induced in a 10H inductor in which the current changes from 10A to 7A in 9×10^{-2}s?

12. Explain why the reactance of a capacitor decreases with increasing frequency, whereas the reactance of an inductor increases with increasing frequency?

13. What is impedance of an *LCR* series circuit? Derive an expression for power dissipated in a.c. *LCR* circuit.

14. Suppose the frequency of a generator is increased from 60Hz to 120Hz. What effect would this have on output voltage?
15. A motor and a generator basically perform opposite functions. Yet some one makes a statement that a motor really acts as a motor and a generator at the same time? Is this really true?
16. A light bulb in series with an A.C. generator and the primary winding of a transformer glows dimly when the secondary leads are connected to a load, such as a resistor, the bulb in the primary winding will brighten, why?
17. If the terminals of a battery are connected to the primary winding of transformer, why will a steady potential differences not appear across the secondary windings.
18. The power supply for a picture tube in a colour television (TV) set typically requires 15,000V A.C. How can this potential difference be provided if only 230V are available at a household electric outlet?
19. Would two coils acts as transformer without an iron core? If so, why not omit the core to save money?
20. An ac source has a 10-volt out-put. A particular circuit requires only a 2V A.C. input. How would you accomplish this? Explain.
21. A person has a single transformer with 50 turns on one part of the core and 500 turns on the other. Is this a step-up or a step-down transformer? Explain.
22. Some transformers have various terminals or “taps” on the secondary so that connecting to different tap puts different functions of the total number of secondary windings into a circuit? What is the advantage of this?
23. A transformer in an electric welding machine draws 3A from a 240V A.C. power line and delivers 400A. What is the potential difference across the secondary of the transformer?
24. A 240-V, 400W electric mixer is connected to a 120-V power line through a transformer. What is the ratio of turns in the transformer? and How much current is drawn from the power line?
25. The primary of a step-up transformer having 125 turns is connected to a house lighting circuit of 220 V_{ac}. If the secondary is to deliver 15,000 volts, how many turns must it have?
26. The secondary of a step-down transformer has 25 turns of wire and primary is connected to a 220V ac. line. If the secondary is to deliver 2.5 volt at the out-put terminals, how many turns should the primary have?
27. The primary of a step-down transformer has 600 turns and is connected to a 120V ac line. If the secondary is to supply 5 volts at its terminal and electron current of 3.5A, find the number of turns in the secondary and the electron current in the primary?

28. A step-up transformer with 352 turns in the primary is connected to a 220V ac line. The secondary delivers 10,000 volts at its terminal and a current of 40 milliampere.

(a) How many turns are in the secondary?

(b) What is the current in the primary?

(c) What power is drawn from the line?

ANSWERS TO INTEXT QUESTIONS

19.1

1. $N = 1000$, $r = 5 \times 10^{-2}$m and $B_1 = 10$T $\quad B_2 = 0$T

a) For $t = 1$s,

$$|\mathbf{e}| = N\frac{(B_2 - B_1)}{t}\pi r^2$$

$$= 10^3 \times \frac{10 \times \pi \times 25 \times 10^{-4}}{1}$$

$$= 25\pi \text{V}$$

$$= 25 \times 3.14 = 78.50\text{V}$$

b) For $t = 1$ms

$$|e| = \frac{10^3 \times 10\pi \times 25 \times 10^{-4}}{10^{-3}}$$

$$= 78.5 \times 10^3 \text{V}$$

2. Since $\phi = A + Dt^2$, $e_1 = \frac{d\phi}{dt} = 2Dt$

$\therefore \quad e = Ne_1 = 2N\,Dt$

$$= 2 \times 250 \times 15t = 7500t$$

For $t = 0$, $e_1 = 0$ and hence $e = 0$V

For $t = 3$s, $e = 22500$V

3. $\phi = \mathbf{B}.\mathbf{S} = BS\cos\theta$

$$|e| = N\frac{d\phi}{dt}$$

$$|e| = \left|NS\frac{dB}{dt}\cos\theta\right| \because \theta \text{ is constt}$$

(a) $|e|$ is max.

when $\cos\theta = 1$, $\theta = 0$, i.e., The coil is normal to the field.

(b) $|e|$ is min.

when $\theta = 90$, i.e. coil surface is parallel to the field.

19.2

1. As we look on the coil from magnet side Anticlockwise for both *A* and *B*.
2. In all the loops except loop *E* there is a change in magnetic flux. For each of them the induced current will be anticlockwise
3. Yes, there is an induced current in the ring. The bar magnet is acted upon by a repulsive force due to the induced current in the ring.
4. To minimise loss of energy due to eddy currents.

19.3

1. $$e = L\frac{dI}{dt} = \omega\frac{N^2A}{\ell}\frac{(I_2 - I_1)}{t}$$

$$= \frac{4\pi\times10^{-7}\times\pi\times10^{-2}\times(2.5-0)}{1\times10^{-3}}$$

$$= 10^{-6}\,\text{V}$$

2. Because, current in the two parallel strands flow in opposite directions and oppose the self induced currents and thus minimize the induction effects.
3. $$3.5\times10^{-3} = 9.7\times10^{-3}\times\frac{dI}{dt}$$

$$= \frac{dI}{dt} = \frac{3.5}{9.7} = 0.36\text{A } s^{-1}$$

19.4

1. Because, the inductor creates an inertia to the growth of current by inducing a back emf

2. $2.2 \times 10^{-3} = \dfrac{L}{R}$

$\Rightarrow L = 2.2 \times 68 \times 10^{-3}\text{H}$

$= 150\text{mH}$

19.5

1. (a)If i_1 is increasing, the flux emerging out of the first coil is also increasing. Therefore, the induced current in the second coil will oppose this flux by a current flowing in clockwise sense as seen by O. Therefore B will be positive and A negative.

   (b)If i_2 is decreasing, flux emerging out of the first coil is decreasing. To increase it the induced current should flow in out anticlockwise sense leaving C at positive potential and D at negative.

2. No, the mutual inductance will decrease. Because, when the two coils are at right angles coupling of flux from one coil to another coil will be the least.

19.6

1. It actually does but we can not detect it, because the frequency of our domestic ac is 50Hz. Our eye can not detect changes that take place faster than 15 times a second.

2. (i) $I_{rms} = \dfrac{E_{rms}}{R} = \dfrac{220}{25}\dfrac{V}{\Omega} = 8.8\text{A}.$

   (ii) Peak value of current $I_m = \sqrt{2}\, I_{rms} = 1.4 \times 8.8 = 12.32$ A.

   Instantaneous current $= I_0 \sin 2\pi\nu t$

   $= 12.32 \sin 100\pi t$

   (iii) Average value of current over integral number of cycles will be zero.

3. Since an ac current varies sinusoidally, its average value over a complete cycle is zero but rms value is finite.

19.7

1. Capacitive reactance $X_C = \dfrac{1}{2\pi\nu C}$. As C increases X_C decreases and I increases.

2. A charged capacitor takes some time in getting discharged. As frequency of source increases it starts charging the capacitor before it is completely

discharged. Thus the maximum charge on capacitor and hence maximum current flowing through the capacitor increases though V_m is constant.

3. Because the energy stored in the capacitor during a charging half cycle is completely recovered during discharging half cycle. As a result energy stored in the capacitor per cycle is zero.

4. Capacitative reactance $X_C = \frac{1}{2\pi v C}$ as v increases X_C decreases. This is so because on capacitor plates now more charge accumulates.

19.8

1. In accordance with Lenz's law a back emf is induced across the inductor when ac is passed through it. The back emf $e = -L\frac{dI}{dt}$.

2. $I_{rms} = \frac{V_{rms}}{X_L}$ frequency increases, X_L $(= 2\pi\, vL)$ increases, hence I_{rms} decreases.

19.9

1. (i) The a.c. generator has slip rings whereas the d.c. generator has a split rings commutator.

   (ii) a.c. generator produces current voltage in sinusoidal form but d.c. generator produces current flowing in one direction all through.
2. Four essential parts of a generator are armature, field magnet, slip rings and brushes.
3. The commutator converts a.c. wave form to d.c. wave form.
4. Attached to the bicycle for lighting purpose.

19.10

1. No, because the working of a transformer depends on the principle of electromagnetic induction, which requires time varying current.
2. Because the ratio of the voltage in primary and secondary coils is proportional to the ratio of number of their turns.
3. No, they are reciprocal to each other.
4. Step-down transformer.

Answers To Problems in Terminal Exercise

1. (a) 3×10^{-2} W_b s^{-1} (b) 0.47 T s^{-1}

4. (b) 5×10^{-2} V

5. (a) (*i*) $\frac{1}{\pi} \times 10^4\ \Omega$ (*ii*) $\frac{1}{2\pi} \times 10^4\ \Omega$

(b) (*i*) 0.1 $\pi\Omega$ (*ii*) 0.2 $\pi\Omega$

6. (a) $6.7 \times 10^2\ \Omega$ (b) 99 Ω (c) 813.9 Ω (d) □ 4 rad

(e) 0.16 A (f) Current lags

11. 333.3V

23. 1.8A.

24. 1 : 2 , $\frac{10}{3}$A .

25. 8522 turns

26. 2200 turns

27. 25 turns, $\frac{1}{7}$A .

28. (a) 16000 lurns, (b) $\frac{20}{11}$A (c) 400W

312en20

20

REFLECTION AND REFRACTION OF LIGHT

Light makes us to see things and is responsible for our visual contact with our immediate environment. It enables us to admire and adore various beautiful manifestations of mother nature in flowers, plants, birds, animals, and other forms of life. Can you imagine how much shall we be deprived if we were visually impaired? Could we appreciate the brilliance of a diamond or the majesty of a rainbow? Have you ever thought how light makes us see? How does it travel from the sun and stars to the earth and what is it made of? Such questions have engaged human intelligence since the very beginning. You will learn about some phenomena which provide answers to such questions.

Look at light entering a room through a small opening in a wall. You will note the motion of dust particles, which essentially provide simple evidence that light travels in a straight line. An arrow headed straight line represents the direction of propagation of light and is called a ray; a collection of rays is called a **beam**. The ray treatment of light constitutes **geometrical optics**. In lesson 22, you will learn that light behaves as a wave. But a wave of short wavelength can be well opproximated by the ray treatment. When a ray of light falls on a mirror, its direction changes. This process is called *reflection*. But when a ray of light falls at the boundary of two dissimilar surfaces, it bends. This process is known as *refraction*. You will learn about reflection from mirrors and refraction from lenses in this lesson. You will also learn about *total internal reflection.* These phenomena find a number of useful applications in daily life from automobiles and health care to communication.

OBJECTIVES

After studying this lesson, you should be able to:

- *explain reflection at curved surfaces and establish the relationship between the focal length and radius of curvature of spherical mirrors;*

- *state sign convention for spherical surfaces;*
- *derive the relation between the object distance, the image distance and the focal length of a mirror as well as a spherical refractive surface;*
- *state the laws of refraction;*
- *explain total internal reflection and its applications in everyday life;*
- *derive Newton's formula for measuring the focal length of a lens;*
- *describe displacement method to find the focal length of a lens; and*
- *derive an expression for the focal length of a combination of lenses in contact.*

20.1 REFLECTION OF LIGHT FROM SPHERICAL SURFACES

In your earlier classes, you have learnt the laws of reflection at a plane surface. Let us recall these laws :

Law 1 –*The incident ray, the reflected ray and the normal to the reflecting surface at the point of incidence always lie in the same plane.*

Law 2 –*The angle of incidence is equal to the angle of reflection* :

$$\angle i = \angle r$$

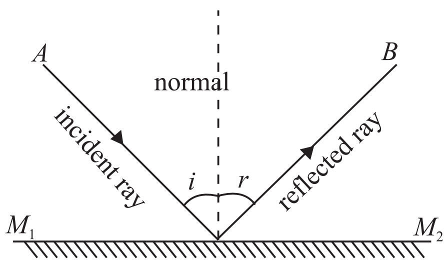
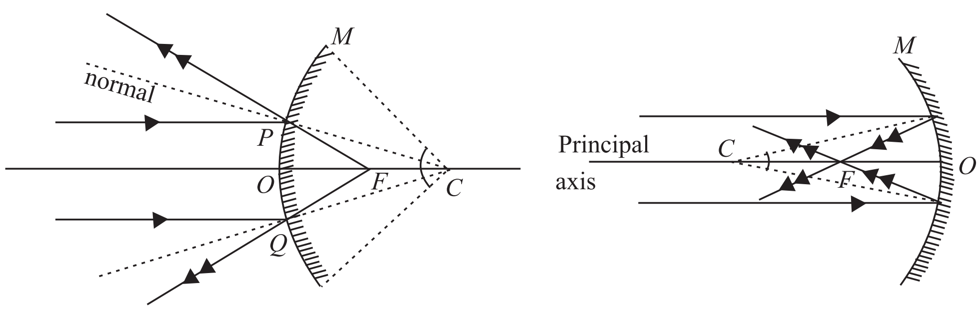


Fig. 20.1 : Reflection of light from a plane surface

These are illustrated in Fig. 20.1. Though initially stated for plane surfaces, these laws are also true for spherical mirrors. This is because a spherical mirror can be regarded as made up of a large number of extremely small plane mirrors. A well-polished spoon is a familiar example of a spherical mirror. Have you seen the image of your face in it? Fig. 20.2(a) and 20.2 (b) show two main types of spherical mirrors.



(a) (b)

Fig. 20.2 : Spherical mirrors : a) a convex mirror, and b) a concave mirror

Note that the reflecting surface of a convex mirror curves outwards while that of a concave mirror curves inwards. We now define a few important terms used for spherical mirrors.

The centre of the sphere, of which the mirror is a part, is called the ***centre of curvature*** of the mirror and the radius of this sphere defines its ***radius of curvature***. The middle point O of the reflecting surface of the mirror is called its ***pole***. The straight line passing through C and O is said to be the ***principal axis*** of the mirror. The circular outline (or periphery) of the mirror is called its ***aperture*** and the angle ($\angle MCM'$) which the aperture subtends at C is called the ***angular aperture*** of the mirror. Aperture is a measure of the size of the mirror.

A beam of light incident on a spherical mirror parallel to the principal axis converges to or appears to diverge from a common point after reflection. This point is known as ***principal focus*** of the mirror. The distance between the pole and the principal focus gives the ***focal length*** of the mirror. A plane passing through the focus perpendicular to the principal axis is called the ***focal plane***.

We will consider only small aperture mirrors and rays close to the principal axis, called **paraxial rays.** (The rays away from the principal axis are called **marginal** or **parapheral rays**.)

INTEXT QUESTIONS 20.1

1. Answer the following questions :

   (a) Which mirror has the largest radius of curvature : plane, concave or convex?

   (b) Will the focal length of a spherical mirror change when immersed in water?

   (c) What is the nature of the image formed by a plane or a convex mirror?

   (d) Why does a spherical mirror have only one focal point?

2. Draw diagrams for concave mirrors of radii 5cm, 7cm and 10cm with common centre of curvature. Calculate the focal length for each mirror. Draw a ray parallel to the common principal axis and draw reflected rays for each mirror.

3. The radius of curvature of a spherical mirror is 30cm. What will be its focal length if (i) the inside surface is silvered? (ii) outside surface is silvered?

4. Why are dish antennas curved?

20.1.1 Ray Diagrams for Image Formation

Let us again refer to Fig. 20.2(a) and 20.2(b). You will note that

- *the ray of light through centre of curvature retraces its path.*

- *the ray of light parallel to the principal axis, on reflection, passes through the focus; and*
- *the ray of light through F is reflected parallel to the principal axis.*

To locate an image, any two of these three rays can be chosen. The images are of two types : real and virtual.

Real image of an object is formed when reflected rays actually intersect. These images are inverted and can be projected on a screen. They are formed on the same side as the object in front of the mirror (Fig. 20.3(a)).

Virtual image of an object is formed by reflected rays that appear to diverge from the mirror. Such images are always erect and virual; these cannot be projected on a screen. They are formed behind the mirror (Fig. 20.3(b)).

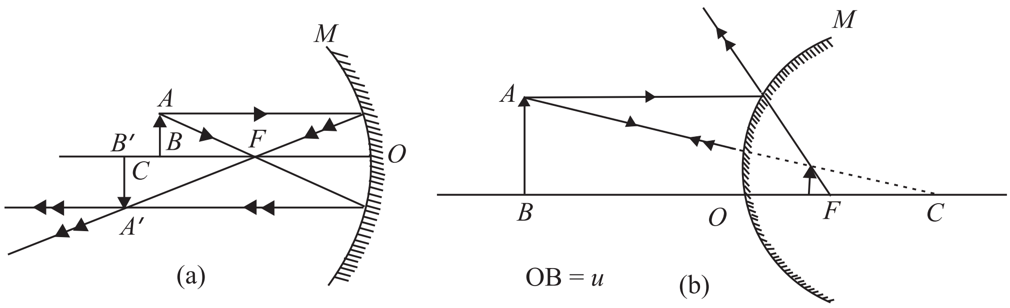


Fig. 20.3 : Image formed by a) concave mirror, and b) convex mirror

20.1.2 Sign Convention

We follow the sign convention based on the cartesian coordinate system. While using this convention, the following points should be kept in mind:

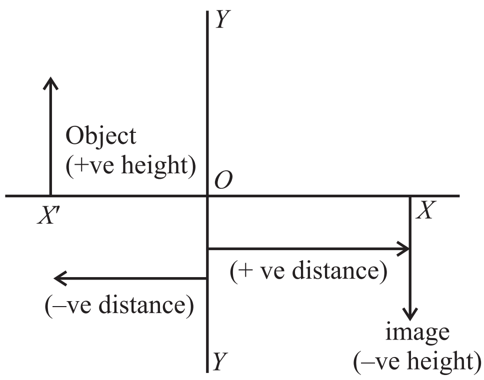


Fig. 20.4 : Sign convention

1. All distances are measured from the pole (*O*) of the mirror. The object is always placed on the left so that the incident ray is always taken as travelling from left to right.
2. All the distances on the left of *O* are taken as negative and those on the right of *O* as positive.
3. The distances measured above and normal to the principal axis are taken as positive and the downward distances as negative.

The radius of curvature and the focal length of a concave mirror are negative and those for a convex mirror are positive.

20.2 DERIVATION OF MIRROR FORMULA

We now look for a relation between the object distance (u), the image distance (v) and the local length f of a spherical mirror. We make use of simple geometry to arrive at a relation, which surprisingly is applicable in all situations. Refer to Fig. 20.5, which shows an object AB placed in front of a concave mirror. The mirror produces an image $A'B'$.

AX and AY are two rays from the point A on the object AB, M is the concave mirror while XA' and YA' are the reflected rays.

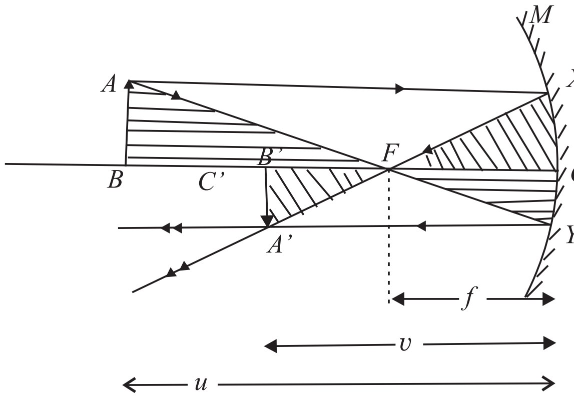


Fig. 20.5 : Image formation by a concave mirror: mirror formula

Using sign conventions, we can write

> In optics it is customary to denote object distance by v. You should not confuse it with velocity.

object distance, $OB = -u,$

focal length, $OF = -f,$

image distance, $OB' = -v,$

and radius of curvature $OC = -2f$

Consider ΔABF and ΔFOY. These are similar triangles. We can, therefore, write

$$\frac{AB}{OY} = \frac{FB}{OF} \tag{20.1}$$

Similarly, from similar triangles ΔXOF and $\Delta B'A'F$, we have

$$\frac{XO}{A'B'} = \frac{OF}{FB'} \tag{20.2}$$

But $AB = XO$, as AX is parallel to the principal axis. Also $A'B' = OY$. Since left hand sides of Eqns. (20.1) and (20.2) are equal, we equate their right hand sides. Hence, we have

$$\frac{FB}{OF} = \frac{OF}{FB'} \tag{20.3}$$

Putting the values in terms of u, v and f in Eqn. (20.3), we can write

$$\frac{-u-(-f)}{-f} = \frac{-f}{-v-(-f)}$$

$$\frac{-u+f}{-f} = \frac{-f}{-v+f}$$

On cross multiplication, we get

$$uv - uf - vf + f^2 = f^2$$

or

$$uv = uf + vf$$

Dividing throughout by uvf, we get the desired relation between the focal length and the object and image distances :

$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u} \qquad (20.4)$$

We next introduce another important term ***magnification***. This indicates the ratio of the size of image to that of the object :

$$m = \frac{\text{size of the image}}{\text{size of the object}} = \frac{h_2}{h_1}$$

But

$$\frac{A'B'}{AB} = \frac{-v}{-u}$$

$$\therefore \qquad m = -\frac{h_2}{h_1} = \frac{v}{u} \qquad (20.5)$$

Since a real image is inverted, we can write

$$m = \frac{A'B'}{AB} = -\frac{v}{u} \qquad (20.5b)$$

To *solve numerical problems, remember the following steps :.*

1. For any spherical mirror, use the mirror formula:

$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u}$$

2. Substitute the numerical values of the given quantities with proper signs.
3. Do not give any sign to the quantity to be determined; it will automatically be obtained with the relevant sign.
4. Remember that the linear magnification is negative for a real image and positive for a virtual image.
5. It is always better to draw a figure before starting the (numerical) work.

INTEXT QUESTIONS 20.2

1. A person standing near a mirror finds his head look smaller and his hips larger. How is this mirror made?
2. Why are the shaving mirrors concave while the rear view mirrors convex? Draw ray diagrams to support your answer.

3. As the position of an object in front of a concave mirror of focal length 25cm is varied, the position of the image also changes. Plot the image distance as a function of the object distance letting the latter change from $-x$ to $+x$. When is the image real? Where is it virtual? Draw a rough sketch in each case.

4. Give two situations in which a concave mirror can form a magnified image of an object placed in front of it. Illustrate your answer by a ray diagram.

5. An object 2.6cm high is 24cm from a concave mirror whose radius of curvature is 16cm. Determine (i) the location of its image, and (ii) size and nature of the image.

6. A concave mirror forms a real image four times as tall as the object placed 15cm from it. Find the position of the image and the radius of curvature of the mirror.

7. A convex mirror of radius of curvature 20cm forms an image which is half the size of the object. Locate the position of the object and its image.

8. A monkey gazes in a polished spherical ball of 10cm radius. If his eye is 20cm from the surface, where will the image of his eye form?

20.3 REFRACTION OF LIGHT

When light passes obliquely from a rarer medium (air) to a denser medium (water, glass), there is a change in its direction of propagation. ***This bending of light at the boundary of two dissimilar media is called refraction***.

When a ray of light is refracted at an interface, it obeys the following two laws :

Law I : The incident ray, the refracted ray and the normal to the surface at the point of incidence always lie in the same plane.

Law II : The ratio of the sine of the angle of incidence to the sine of the angle of refraction is constant for a given pair of media. It is independent of the angle of incidence when light propagates from a rarer to a denser medium. Moreover, for a light of given colour, the ratio depends only on the pair of media.

This law was pronounced by the Dutch scientist Willebrord van Roijen Snell and in his honour is often referred to as ***Snell's law***. According to Snell's law

$$\frac{\sin i}{\sin r} = \mu_{12}$$

where μ_{12} is a constant, called the *refractive index* of second medium with respect to the first medium, and determines how much bending would take place at the interface separating the two media. It may also be expressed as the ratio of the

Notes

velocity of light in the first medium to the velocity of light in the second medium

$$\mu_{12} = \frac{c_1}{c_2}$$

Refractive indices of a few typical substances are given in Table 20.1.Note that these values are with respect to air or vacuum. The medium having larger refractive index is optically denser medium while the one having smaller refractive index is called rarer medium. So water is denser than air but rarer than glass. Similarly, crown glass is denser than ordinary glass but rarer than flint glass.

Table 20.1 : Refractive indices of some common materials

Medium	μ
Vacuum/air	1
Water	1.33
Ordinary glass	1.50
Crown glass	1.52
Dense flint glass	1.65
Diamond	2.42

If we consider refraction from air to a medium like glass, which is optically denser than air[Fig. 20.6 (a)], then $\angle r$ is less than $\angle i$. On the other hand, if the ray passes from water to air, $\angle r$ is greater than $\angle i$ [Fig. 20.6(b)]. That is, the refracted ray bends towards the normal on the air–glass interface and bends away from the normal on water–air interface.

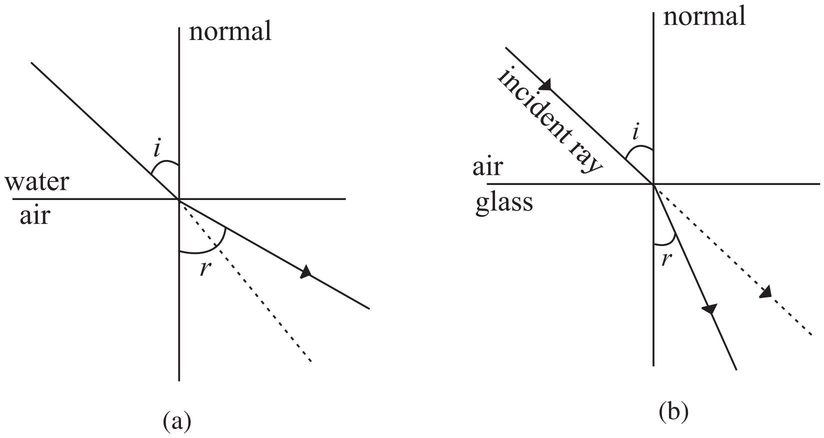


Fig. 20.6 : a) Refraction on air–glass interface, and **b)** refraction on water–air interface

Willebrord Van Roijen Snell (1580 – 1626)

Willebrord Snell was born in 1580 in Lieden. He began to study mathematics at a very young age. He entered the University of Leiden and initially studied law. But, soon he turned his attention towards mathematics and started teaching at the university by the time he was 20. In 1613, Snell succeded his father as Professor of Mathematics.

He did some important work in mathematics, including the method of calculating the approximate value of π by polygon. His method of using 96 sided polygon gives the correct value of π up to seven places while the classical method only gave this value upto two correct places. Snell also published some books including his work on comets. However, his biggest contribution to science was his discovery of the laws of refraction. However, he did not publish his work on refraction. It became known only in 1703, seventy seven years after his death, when Huygens published his results in "Dioptrics".

20.3.1 Reversibility of light

Refer to Fig. 20.6(b) again. It illustrates the principle of reversibility. It appears as if the ray of light is retracing its path. It is not always necessary that the light travels from air to a denser medium. In fact, there can be any combination of transparent media. Suppose, light is incident at a water-glass interface. Then, by applying Snell's law, we have

$$\frac{\sin i_w}{\sin i_g} = \mu_{wg} \tag{20.6}$$

Now, let us consider separate air-glass and air-water interfaces. By Snell's law, we can write

$$\frac{\sin i_a}{\sin i_g} = \mu_{ag}$$

and

$$\frac{\sin i_a}{\sin i_w} = \mu_{aw}$$

On combining these results, we get

$$\mu_{ag} \sin i_g = \mu_{aw} \sin i_w \tag{20.7}$$

This can be rewritten as

$$\frac{\sin i_w}{\sin i_g} = \frac{\mu_{ag}}{\mu_{aw}} \tag{20.8}$$

On comparing Eqns. (20.6) and (20.8), we get

$$\mu_{wg} = \frac{\mu_{ag}}{\mu_{aw}} \tag{20.9}$$

This result shows than when light travels from water to glass, the refractive index of glass with respect to water can be expressed in terms of the refractive indics of glass and water with respect to air.

Example 20.1 : A ray of light is incident at an angle of 30° at a water-glass interface. Calculate the angle of refraction. Given $\mu_{ag} = 1.5$, $\mu_{aw} = 1.3$.

Solution : From Eqn. (20.8), we have

$$\frac{\sin i_w}{\sin i_g} = \frac{\mu_{ag}}{\mu_{aw}}$$

$$\frac{\sin 30^0}{\sin i_g} = \frac{1.5}{1.3}$$

or $$\sin i_g = \left(\frac{1.3}{1.5}\right) \times \frac{1}{2}$$

$$= 0.4446$$

or $$i_g = 25°41'$$

Example 20.2 : Calculate the speed of light in water if its refractive index with respect to air is 4/3. Speed of light in vacuum = 3×10^8 ms^{-1}.

Solution : We know that

$$\mu = \frac{c}{\upsilon}$$

or $$\upsilon = \frac{c}{\mu}$$

$$= \frac{(3\times10^8 \text{ ms}^{-1})}{4/3}$$

$$= \frac{3\times10^8\times3}{4}$$

$$= 2.25 \times 10^8 \text{ ms}^{-1}$$

Example 20.3 : The refractive indices of glass and water are 1.52 and 1.33 respectively. Calculate the refractive index of glass with respect to water.

Solution : Using Eqn. (20.9), we can write

$$\mu_{wg} = \frac{\mu_{ag}}{\mu_{aw}} = \frac{1.52}{1.33} = 1.14$$

INTEXT QUESTIONS 20.3

1. What would be the lateral displacement when a light beam is incident normally on a glass slab?
2. Trace the path of light if it is incident on a semicircular glass slab towards its centre when $\angle i < \angle i_c$ and $\angle i > \angle i_c$.
3. How and why does the Earth's atmosphere alter the apparent shape of the Sun and Moon when they are near the horizon?
4. Why do stars twinkle?
5. Why does a vessel filled with water appear to be shallower (less deep) than when without water? Draw a neat ray diagram for it.
6. Calculate the angle of refraction of light incident on water surface at an angle of 52°. Take $\mu = 4/3$.

20.4 TOTAL INTERNAL REFLECTION

ACTIVITY 20.1

Take a stick, cover it with cycle grease and then dip it in water or take a narrow glass bottle, like that used for keeping Homeopathic medicines, and dip it in water. You will observe that the stick or the bottle shine almost like silver. Do you know the reason? This strange effect is due to a special case of refraction. We know that when a ray of light travels from an optically denser to an optically rarer medium, say from glass to air or from water to air, the refracted ray bends away from the normal. This means that the angle of refraction is greater than the angle of incidence. What happens to the refracted ray when the angle of incidence is increased? The bending of refracted ray also increases. However, the maximum value of the angle of refraction can be 90°. ***The angle of incidence in the denser medium for which the angle of refraction in rarer medium, air in this case, equals 90° is called the critical angle, i_C.*** The refracted ray then moves along the boundary separating the two media. If the angle of incidence is greater than the critical angle, the incident ray will be reflected back in the same medium, as shown in Fig. 20.7(c). Such a reflection is called ***Total Internal Reflection*** and the incident ray is said to be totally internally reflected. For total internal reflection to take place, the following two conditions must be satisfied.

- Light must travel from an optically denser to an optically rarer medium.
- The angle of incidence in the denser medium must be greater than the critical angle for the two media.

The glass tube in water in Activity 20.1 appeared silvery as total internal reflection took place from its surface.

An expression for the critical angle in terms of the refractive index may be obtained readily, using Snell's law. For refraction at the glass-air interface, we can write

$$\frac{\sin i}{\sin r} = \mu_{ga}$$

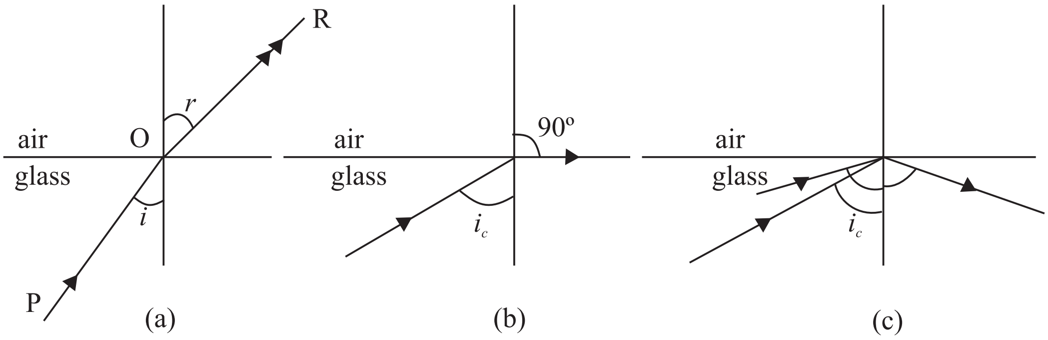


Fig. 20.7 : Refraction of light as it travels from glass to air for **a)** $i < i_c$, **b)** $i = i_c$ and **c)** $i > i_c$

Table 20.2 : Critical angles for a few substances

Substance	*μ*	*Critical angle*
Water	1.33	48.75°
Crown glass	1.52	41.14°
Diamond	2.42	24.41°
Dense flint glass	1.65	37.31°

Putting $r = 90°$ for $i = i_c$, we have

$$\frac{\sin i_c}{\sin 90°} = \mu_{ga}$$

or

$$\sin i_c = \mu_{ga}$$

Hence

$$\mu_{ag} = \frac{1}{\mu_{ga}} = \frac{1}{\sin i_c}$$

The critical angles for a few substances are given in Table 20.2

Example 20.4 : Refractive index of glass is 1.52. Calculate the critical angle for glass air interface.

Solution : We know that

$$\mu = 1/\sin i_c$$

$$\sin i_c = 1/\mu = \frac{1}{1.52}$$

$$\therefore \quad i_c = 42°$$

Much of the shine in transparent substances is due to total internal reflection. Can you now explain why diamonds sparkle so much? This is because the critical angle is quite small and most of the light entering the crystal undergoes multiple internal reflections before it finally emerges out of it.

In ordinary reflection, the reflected beam is always weaker than the incident beam, even if the reflecting surface is highly polished. This is due to the fact that some light is always transmitted or absorbed. But in the case of total internal reflection, cent percent (100%) light is reflected at a transparent boundary.

20.4.1 Applications of Refraction and Total Internal Reflection

There are many manifestations of these phenomena in real life situations. We will consider only a few of them.

(a) Mirage : Mirage is an optical illusion which is observed in deserts or on tarred roads in hot summer days. This, you might have observed, creates an illusion of water, which actually is not there.

Due to excessive heat, the road gets very hot and the air in contact with it also gets heated up. The densities and the refractive indices of the layers immediately above the road are lower than those of the cooler higher layers. Since there is no abrupt change in medium (see Fig. 20.9), a ray of light from a distant object, such as a tree, bends more and more as it passes through these layers. And when it falls on a layer at an angle greater than the critical angle for the two consecutive layers, total internal reflection occurs. This produces an inverted image of the tree giving an illusion of reflection from a pool of water.

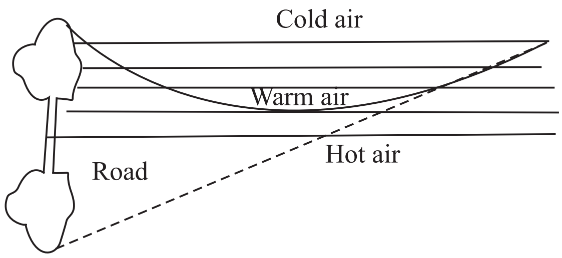


Fig. 20.8 : Formation of mirage

Totally Reflecting Prisms : A prism with right angled isosceles trianglur base or a totally reflecting prism with angles of 90°, 45° and 45° is a very useful device for reflecting light.

Refer to Fig. 20.9(a). The symmetry of the prism allows light to be incident on *O* at an angle of 45°, which is greater than the critical angle for glass i.e. 42°. As a result, light suffers total internal reflection and is deviated by 90°.

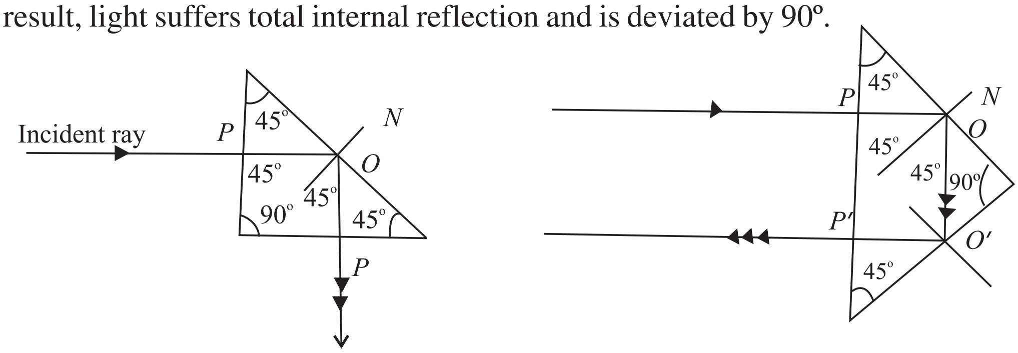


Fig. 20.9 : Totally reflecting prisms

Choosing another face for the incident rays, it will be seen (Fig. 20.9(b)) that the ray gets deviated through 180^0 by two successive total internal reflections taking place at O and O'.

Optical Fibres

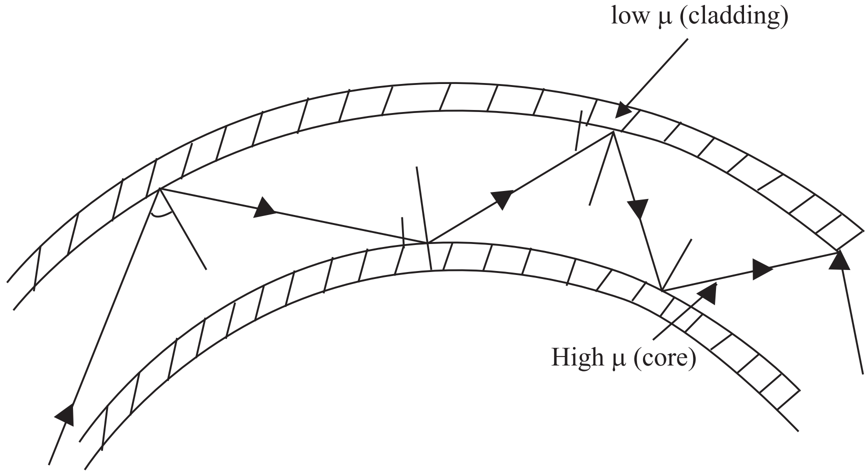


Fig. 20.10 : Multiple reflection in an optical fibre

An optical fibre is a hair-thin structure of glass or quartz. It has an inner core which is covered by a thin layer (called **cladding**) of a material of slightly lower refractive index. For example, the refractive index of the core is about 1.7 and that of the cladding is 1.5. This arrangement ensures total internal reflection. You can easily understand it, if you recall the conditions for total internal reflection.

When light is incident on one end of the fibre at a small angle, it undergoes multiple total internal reflections along the fibre (Fig. 20.10). The light finally emerges with undiminished intensity at the other end. Even if the fibre is bent, this process is not affected. Today optical fibres are used in a big way. A flexible light pipe using optical fibres may be used in the examination of inaccessible parts of the body e.g. laproscopic examination of stomach, urinary bladder etc. Other medical applications of optical fibres are in neurosurgery and study of bronchi. Besides medical applications, optical fibres have brought revolutionary changes in the way we communicate now. Each fibre can carry as many as 10,000 telephone messages without much loss of intensity, to far off places. That is why millions of people across continents can interact simultaneously on a fibre optic network.

INTEXT QUESTIONS 20.4

1. Why can't total internal reflection take place if the ray is travelling from a rarer to a denser medium?

2. Critical angle for glass is 42°. Would this value change if a piece of glass is immersed in water? Give reason for your answer.

3. Show, with the help a ray diagram how, a ray of light may be deviated through 90° using a i) plane mirror, and ii) totally reflecting prism. Why is the intensity of light greater in the second case?

4. A liquid in a container is 25cm deep. What is its apparant depth when viewed from above, if the refractive index of the liquid is 1.25? What is the critical angle for the liquid?

20.5 REFRACTION AT A SPHERICAL SURFACE

We can study formation of images of objects placed around spherical surfaces such as glass marbles (Kanchas), water drops, glass bottle, etc. For measuring distances from spherical refracing surfaces, we use the same sign convention as applicable for spherical mirrors. Refer to Fig. 20.11.

SPS′ is a convex refracting surface separating two media, air and glass. *C* is its centre of curvature. *P* is a point on *SPS′* almost symmetrically placed. You may call it the *pole*. *CP* is then the *principal axis*. *O* is a point object. *OA* is an incident ray and *AB* is the refracted ray. Another ray *OP* falls on the surface normally and goes undeviated after refraction. *PC* and *AB* appear to come from *I*. Hence *I* is the virtual image of *O*.

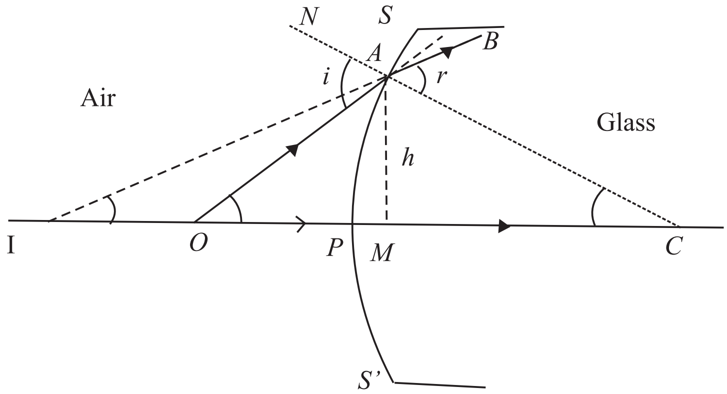


Fig. 20.11 : Refraction at a spherical surface

Let $\angle OAN = i$, the angle of incidence and $\angle CAB = r$, the angle of refraction. Using the proper sign convention, we can write

$$PO = -u \; ; PI = -v \; ; PC = +R$$

Let α, β, and γ be the angles subtended by *OA*, *IA* and *CA*, respectively with the principal axis and *h* the height of the normal dropped from *A* on the principal axis. In Δ*OCA* and Δ*ICA*, we have

$$i = \alpha + \gamma \quad (i \text{ is exterior angle}) \qquad (20.10)$$

and

$$r = \beta + \gamma \quad (r \text{ is exterior angle}) \qquad (20.11)$$

From Snell's law, we recall that

$$\frac{\sin i}{\sin r} = \mu$$

where μ is the refractive index of the glass surface with respect to air. For a surface of small aperture, *P* will be close to *A* and so *i* and *r* will be very small (sin $i \simeq i$, sin $r \simeq r$). The above equation, therefore, gives

$$i = \mu r \tag{20.12}$$

Substituting the values of *i* and *r* in Eqn. (20.12) from Eqns. (20.10) and (20.11) respectively, we get

$$\alpha + \gamma = \mu\,(\beta + \gamma)$$

or
$$\alpha - \mu\beta = \gamma\,(\mu - 1) \tag{20.13}$$

As α, β and γ are very small, we can take tan $\alpha \simeq \alpha$, and tan $\beta \simeq \beta$, and tan $\gamma \simeq \gamma$. Now referring to Δ*OAM* in Fig. 20.11, we can write

$$\alpha \approx \tan \alpha = \frac{AM}{MO} = \frac{AP}{PO} = \frac{h}{-u} \quad \text{(if } M \text{ is very near to } P\text{)}$$

$$\beta \approx \tan \beta = \frac{AM}{MI} = \frac{AM}{PI} = \frac{h}{-v}$$

and
$$\gamma \simeq \tan \gamma = \frac{AM}{MC} = \frac{AM}{PC} = \frac{h}{R}$$

Substituting for α, β and γ in Eqn. (20.13), we get

$$\frac{h}{-u} - \frac{\mu h}{v} = (\mu - 1)\,\frac{h}{R}$$

or
$$\frac{\mu}{v} - \frac{1}{u} = \frac{\mu - 1}{R} \tag{20.14}$$

This important relationship correlates the object and image distances to the refractive index μ and the radius of curvature of the refracting surface.

20.5.1 Reflection through lenses

A lens is a thin piece of transparent material (usually glass) having two surfaces, one or both of which are curved (mostly spherical). You have read in your earlier classes that lenses are mainly of two types, namely, convex lens and concave lens. Each of them is sub-divided into three types as shown in Fig. 20.12. Thus, you can have plano-convex and plano-concave lenses too.

Basic Nomenclature

Thin lens : If the thickness of a lens is negligible in comparison to the radii of curvature of its curved surfaces, the lens is referred to as a thin lens. We will deal with thin lenses only.

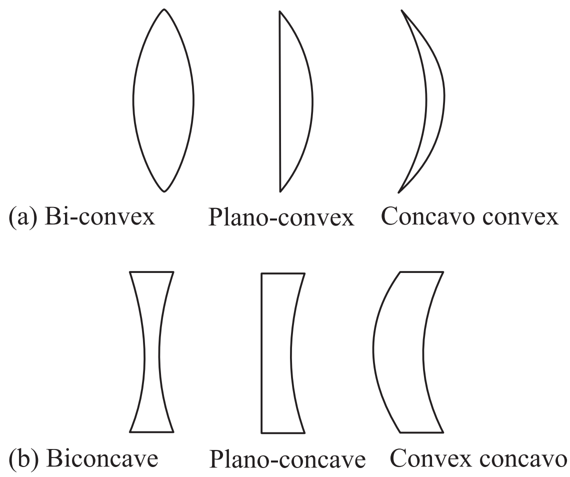


Fig. 20.12 : Types of lenses

Principal axis is the line joining the centres of curvature of two surfaces of the lens.

Optical centre is the point at the centere of the lens situated on the principal axis. The rays passing through the optical centre do not deviate.

Principal focus is the point at which rays parallel and close to the principal axis converge to or appear to diverge. It is denoted by *F* (Fig. 20.13) Rays of light can pass through a lens in either direction. So every lens has two principal focii, one on its either side.

Focal length is the distance between the optical centre and the principal focus. In Fig. 20.13, *OF* is focal length (*f*). As per the sign convention, *OF* is positive for a convex lens and negative for a concave lens.

Focal plane is the plane passing through the focus of a lens perpendicular to its principal axis.

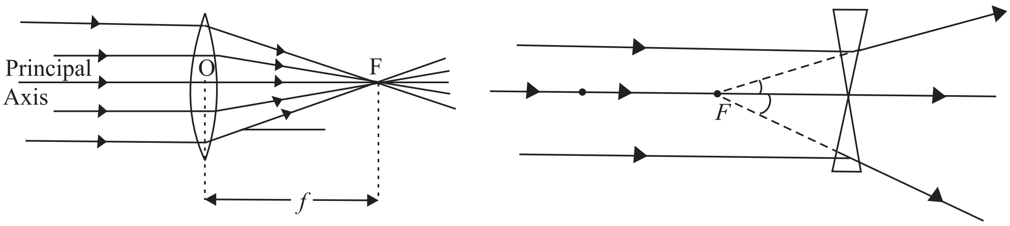


Fig. 20.13 : Foci of **a)** convex, and **b)** concave lenses

20.5.2 Lens Maker's Formula and Magnification

You can now guess that the focal length must be related to the radius of curvature and the refractive index of the material of the lens. Suppose that a thin convex lens *L* is held on an optical bench (Fig. 20.14). Let the refractive index of the material of the lens with respect to air be μ and the radii of curvatures of its two surfaces be R_1 and R_2, respectively. Let a point object be situated on the principal axis at P. C_1 and C_2 are the centres of curvature of the curved surfaces 1 and 2, respectively.

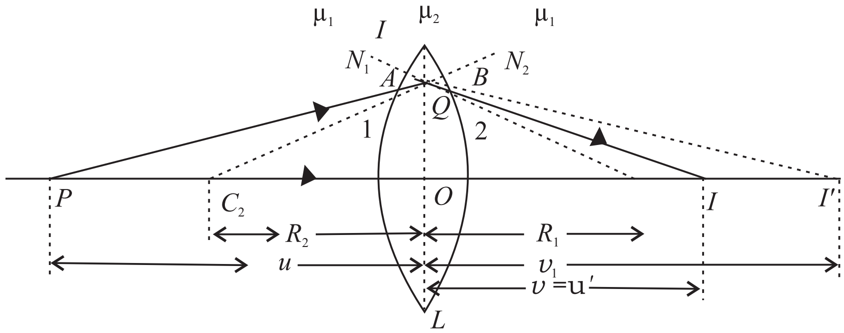


Fig. 20.14 : Point image of a point object for by a thin double convex lens

A ray from P strikes surface 1 at A. $C_1 N_1$ is normal to surface 1 at the point A. The ray PA travels from the rarer medium (air) to the denser medium (glass), and bends towards the normal to proceed in the direction AB. The ray AB would meet the principal axis C_2C_1 at the point I' *in the absence of the surface 2*. Similarly, another ray from P passing through the optical centre O passes through the Point I'. I' is thus the virtual image of the object P.

Then object distance $OP = u$ and image distance $OI' = v_1$ (say). Using Eqn. (20.14) we can write

$$\frac{\mu}{v_1} - \frac{1}{u} = \frac{\mu - 1}{R_1} \qquad (20.15)$$

Due to the presence of surface 2, the ray AB strikes it at B. C_2N_2 is the normal to it at point B. As the ray AB is travelling from a denser medium (glass) to a rarer medium (air), it bends away from the normal C_2N_2 and proceeds in the direction BI and meets another ray from P at I. Thus I is image of the object P formed by the lens. It means that image distance $OI = v$.

Considering point object O, its virtual image is I' (due to surface 1) and the final image is I. I' is the virtual object for surface 2 and I is the final image. Then for the virtual object I' and the final image I, we have, object distance $OI' = u' = v_1$ and image distance $OI = v$.

On applying Eqn. (20.12) and cosidering that the ray AB is passing from *glass to air,* we have

$$\frac{(1/\mu)}{v} + \frac{1}{v_1} = \frac{(1/\mu) - 1}{R_2}$$

or,
$$\frac{1}{\mu v} - \frac{1}{v_1} = \frac{1 - \mu}{\mu R_2}$$

Since the lens used is actually thin, points A and B may be considered very close to point a and hence C_1A is taken equal to C_1Q and C_2B as C_2Q.

Multiping both sides by μ, we get

$$\frac{1}{v} - \frac{\mu}{v_1} = \frac{\mu - 1}{R_2} \qquad (20.16)$$

Adding Eqns. (20.15) and (20.16), we have

$$\frac{1}{v} - \frac{1}{u} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \qquad (20.17)$$

It $u = \infty$, that is the object is at infinity, the incoming rays are parallel and after refraction will converge at the focus ($v = f$). Then Eqn. (20.17) reduces to

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \qquad (20.18)$$

This is called lens maker's formula.

From Eqns. (20.17) and (20.18), we can conclude that

- *The focal length of a lens depends on the radii of curvature of spherical surfaces. Focal length of a lens of larger radii of curvature will be more.*
- *Focal length of a lens is smaller if the refractive index of its material is high.*

In case a lens is dipped in water or any other transparent medium, the value of μ changes and you can actually work out that focal length will increase. However, if the density of the medium is more than that of the material of the lens, say carbon disulphide, the rays may even diverge.

20.5.3 Newton's Formula

Fig. 20.5.3 shows the image of object AB formed at $A'B'$ by a convex lens F_1 and F_2 are the first and second principal focii respectively.

Let us measure the distances of the object and image from the first focus and second focus respectively. Let x_1 be the distance of object from the first focus and x_2 be the distance of image from the second focus and f_1 and f_2 the first and second focal lengths, respectively as shown in Fig. 20.5.3.

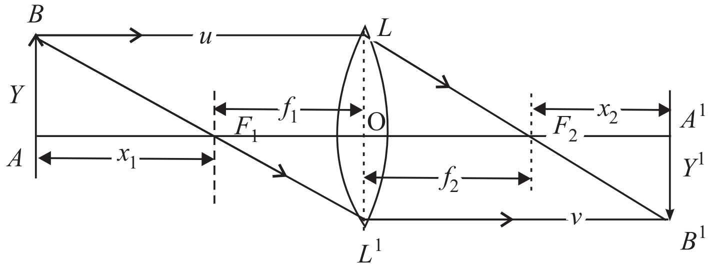


Fig. 20.5.3

Now, in similar “s, ABF_1 and $OL'F_1$

$$\frac{-y'}{y} = \frac{-f_1}{-x_1}$$

Also from similar Δs, OLF_2 and $A'B'F_2$

$$\frac{-y'}{y} = \frac{x_2}{f_2}$$

Comparing these two equations we get

$$x_1 x_2 = f_1 f_2$$

for $\quad f_1 \equiv f_2 \approx f$ (say), then $x_1 x_2 = +f^2$

or $\quad f = \sqrt{x_1 x_2}$

This relation is called Newton’s formula and can be conveniently used to measure the focal length.

20.5.4 Displacement Method to find the Position of Images (Conjugates points)

In the figure 20.5.4, $A'B'$ is the image of the object AB as formed by a lens L. $OA = u$ and $OA' = v$.

The principle of reversibility of light rays tells us that if we move the lens towards the right such that $AO = v$, then again the image will be formed at the same place.

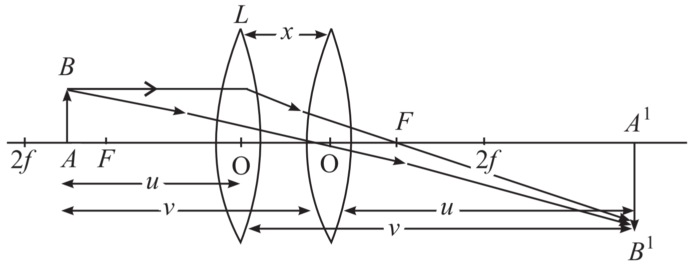


Fig. 20.5.4

Thus $\quad AA' = D = u + v \quad$ ...(i)

and the separation between the two positions of the lens:

$$OO' = x = (v - u) \quad \text{...(ii)}$$

Adding (i) and (ii) we get

$$v = \frac{x+D}{2}$$

and, subtracting (ii) from (i) we get

$$u = \frac{D-x}{2}$$

Substituting these values in the lens formula, we get.

$$\frac{1}{f} = \frac{1}{v} - \frac{1}{(-u)}$$

$$\frac{1}{f} = \frac{2}{x+D} + \frac{2}{D-x} = \frac{2}{D+x} + \frac{2}{D-x}$$

$$\frac{1}{f} = \frac{2(D-x+D+x)}{D^2-x^2}$$

$$\frac{1}{f} = \frac{4D}{D^2-x^2}$$

or $$f = \frac{D^2-x^2}{4D}$$

Thus, keeping the positions of the object and screen fixed we can obtain equally clear, bright and sharp images of the object on the screen corresponding to the two positions of the lens. This again is a very convenient way of finding f of a lens.

20.6 FORMATION OF IMAGES BY LENSES

The following properties of the rays are used in the formation of images by lenses:

- A ray of light through the optical centre of the lens passes undeviated.
- A parallel ray, after refraction, passes through the principal focus.
- A ray of light through F or F' is rendered parallel to the principal axis after refraction.

  Any two of these rays can be chosen for drawing ray diagrams.

The lens formula $\frac{1}{f} = \frac{1}{v} - \frac{1}{u}$ suggests the dependence of the image distance (v) on the object distance (u) and the focal length (f) of the lens.

The magnification of a lens is defined as the ratio of the height of the image formed by the lens to the height of the object and is denoted by *m* :

$$m = \frac{I}{O} = \frac{v}{u}$$

where *I* is height of the image and *O* the height of the object.

Example 20.5 : The radii of curvature of a double convex lens are 15cm and 30cm, respectively. Calculate its focal length. Also, calculate the focal length when it is immersed in a liquid of refractive index 1.65. Take μ of glass = 1.5.

Solution : From Eqn. (20.18) we recall that

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

Here R_1= + 15cm, and R_2 = –30cm. On substituting the given data, we get

$$\frac{1}{f} = (1.5 - 1)\left(\frac{1}{15} - \frac{1}{-30}\right)$$

$$\Rightarrow \qquad f = 20 \text{ cm}$$

When the lens is immersed in a liquid, μ will be replaced by $\mu_{\ell g}$:

$$\mu_{\ell g} = \frac{\mu_{ag}}{\mu_{al}}$$

$$= \frac{1.5}{1.65} = \frac{10}{11}$$

Therefore

$$\frac{1}{f_l} = (\mu_{\ell g} - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$= \left(\frac{10}{11} - 1\right)\left(\frac{1}{15} - \frac{1}{-30}\right)$$

$$= -\frac{1}{110}$$

$$\therefore \qquad f = -110\text{cm}$$

As *f* is negative, the lens indeed behaves like a concave lens.

20.7 POWER OF A LENS

A practical application of lenses is in the correction of the defects of vision. You may be using spectacles or seen other learners, parents and persons using spectacles. However, when asked about the power of their lens, they simply quote a positive or negative number. What does this number signify? This number is the power of a lens in dioptre. The power of a lens is defined as the reciprocal of its focal length in metre:

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

The S1 unit of power of a lens is m^{-1}. Dioptre is only a commercial unit generally used by opticians. The power of a convex lens is positive and that of a concave lens is negative. Note that greater power implies smaller focal length. Using lens maker's formula, we can relate power of a lens to its radii of curvature:

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

or

$$P = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

Example 20.6 : Calculate the radius of curvature of a biconvex lens with both surfaces of equal radii, to be made from glass ($\mu = 1.54$), in order to get a power of +2.75 dioptre.

Solution :

$$P = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$P = +2.75 \text{ dioptre}$$

$$\mu = 1.54$$

$$R_1 = R$$

and

$$R_2 = -R$$

Substuting the given values in lens makers formula, we get

$$2.75 = (0.54)\left(\frac{2}{R}\right)$$

$$R = \frac{0.54 \times 2}{2.75}$$

$$= 0.39 \text{ m}$$

$$= 39 \text{ cm}$$

20.8 COMBINATION OF LENSES

Refer to Fig. 20.15. Two thin convex lenses A and B having focal lengths f_1 and f_2, respectively have been kept in contact with each other. O is a point object placed on the common principal axis of the lenses.

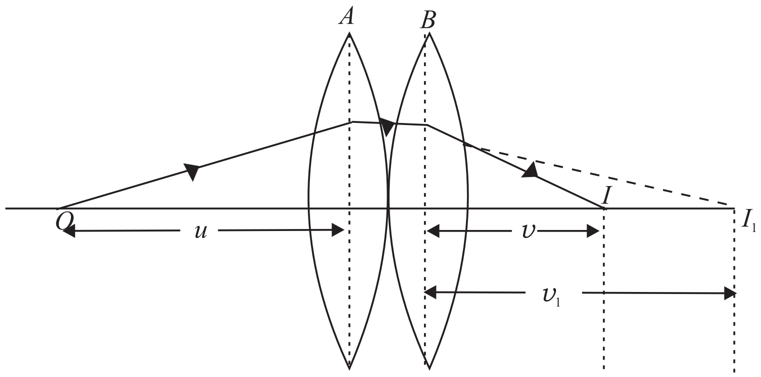


Fig. 20.15 : Two thin convex lenses in contact

Note that lens A forms the image of object O at I_1. This image serves as the ***virtual*** object for lens B and the final image is thus formed at I. If v be the object distance and v_1 the image distance for the lens A, then using the lens formula, we can write

$$\frac{1}{v_1} - \frac{1}{u} = \frac{1}{f_1} \tag{20.19}$$

If v is the final image distance for the lens B, we have

$$\frac{1}{v} - \frac{1}{v_1} = \frac{1}{f_2} \tag{20.20}$$

Note that in writing the above expression, we have taken v_1 as the object distance for the thin lens B.

Adding Eqns. (20.19) and (20.20), we get

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f_1} + \frac{1}{f_2} \tag{20.21}$$

If the combination of lenses is replaced by a single lens of focal length F such that it forms the image of O at the same position I, then this lens is said to be equivalent to both the lenses. It is also called the ***equivalent lens*** for the combination. For the equivalent lens, we can write

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{F} \tag{20.22}$$

where

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}. \tag{20.23}$$

If P is power of the equivalent lens and P_1 and P_2 are respectively the powers of individual lenses, then

$$P = P_1 + P_2 \qquad (20.24)$$

Note that Eqns.(20.23) and (20.24) derived by assuming two thin convex lenses in contact also hold good for any combination of two thin lenses in contact (the two lenses may both be convex, or concave or one may be concave and the other convex).

Example 20.7 : Two thin convex lenses of focal lengths 20cm and 40cm are in contact with each other. Calculate the focal length and the power of the equivalent lens.

Solution : The formula for the focal length of the combination $\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}$ gives

$$\frac{1}{F} = \frac{1}{20} + \frac{1}{40}$$

$$= \frac{3}{40}$$

or $$F = \frac{40}{3} = 13.3\text{cm} = 0.133\text{cm}$$

Power of the equivalent lens is

$$P = \frac{1}{F} = \frac{1}{0.133} = +7.5 \text{ dioptre.}$$

INTEXT QUESTIONS 20.5

1. On what factors does the focal length of a lens depend?
2. A lens, whose radii of curvature are different, is used to form the image of an object placed on its axis. If the face of the lens facing the object is reversed, will the position of the image change?
3. The refractive index of the material of an equi-double convex lens is 1.5. Prove that the focal length is equal to the radius of curvature.
4. What type of a lens is formed by an air bubble inside water?
5. A lens when immersed in a transparent liquid becomes invisible. Under what condition does this happen?
6. Calculate the focal length and the power of a lens if the radii of curvature of its two surfaces are +20cm and –25cm (μ = 1.5).

7. Is it possible for two lenses in contact to produce zero power?

8. A convex lens of focal length 40cm is kept in contact with a concave lens of focal length 20cm. Calculate the focal length and the power of the combination.

Defects in image formation

Lenses and mirrors are widely used in our daily life. It has been observed that they do not produce a point image of a point object. This can be seen by holding a lens against the Sun and observing its image on a paper. You will note that it is not exactly circular. Mirrors too do not produce a perfect image. The defects in the image formation are known as ***aberrations***. The aberrations depend on (i) the quality of lens or mirror and (ii) the type of light used.

Two major aberrations observed in lenses and mirrors, are (a) ***spherical aberration*** and (b) ***chromatic aberration***. These aberration produce serious defects in the images formed by the cameras, telescopes and miscroscopes etc.

Spherical Aberration

This is a monochromatic defect in image formation which arises due to the sphericity and aperture of the refracting or reflecting surfaces. The paraxial rays and the marginal rays form images at different points I_p and I_m respectively (Fig. 20.16)

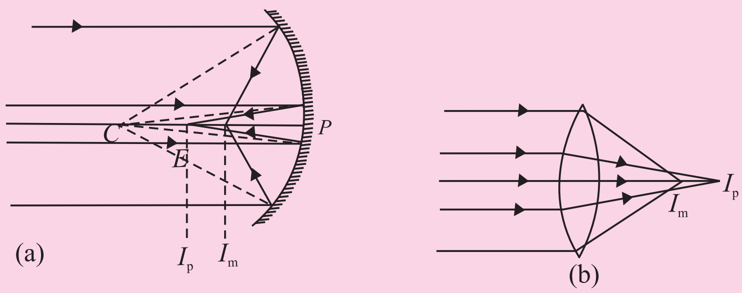


Fig. 20.16 : Spherical aberration in **a)** spherical mirror, and **b)** lens. I_p is image formed by the paraxial rays and I_m that formed by the marginal rays.

The ***spherical aberration*** in both mirrors and lenses can be reduced by allowing only the paraxial rays to be incident on the surface. It is done by using stops. Alternatively, the paraxial rays may be cut-off by covering the central portion, thus allowing only the marginal or parapheral rays to form the image. However, the use of stops reduces the brightness of the image.

A much appreciated method is the use of elliptical or parabolic mirrors.

The other methods to minimize spherical aberration in lenses are : use of plano convex lenses or using a suitable combination of a convex and a concave lens.

Chromatic Aberration in Lenses

A convex lens may be taken as equivalent to two small-angled prisms placed base to base and the concave lens as equivalent to such prisms placed vertex to vertex. Thus, a polychromatic beam incident on a lens will get dispersed. The parallel beam will be focused at different coloured focii. This defect of the image formed by spherical lenses is called ***chromatic aberration***. It occurs due to the dispersion of a polychromatic incident beam (Fig. 20.17. Obviously the red colour is focused farther from the lens while the blue colour is focused nearer the lens (in a concave lens the focusing of the red and blue colours takes place in the same manner but on the opposite side of it).

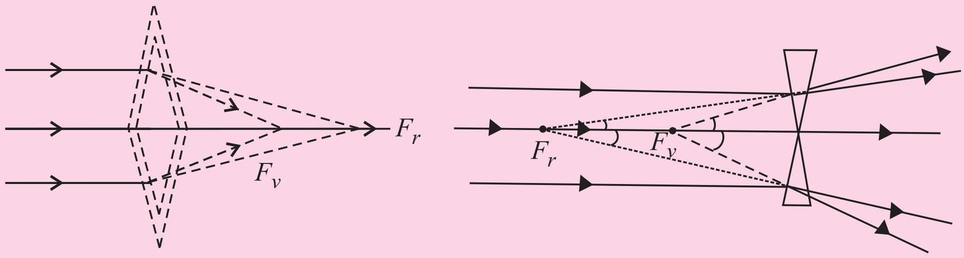


Fig. 20.17: Chromatic aberration

To remove this defect we combine a convergent lens of suitable material and focal length when combined with a divergent lens of suitable focal length and material. Such a lens combination is called an ***achromatic doublet*** . The focal length of the concave lens can be found from the necessary condition for achromatism given by

$$\frac{\omega_1}{f_1} + \frac{\omega_2}{f_2} = 0$$

WHAT YOU HAVE LEARNT

- Real image is formed when reflected rays actually intersect after reflection. It can be projected on a screen.
- The focal length is half of the radius of curvature.

$$f = \frac{R}{2}$$

The object and image distances are related to magnification as

$$m = \frac{v}{u}$$

- Refraction of light results in change in the speed of light when it travels from one medium to another. This causes the rays of light to bend towards or away from the normal.
- The refractive index μ determines the extent of bending of light at the interface of two media.
- Snell's law is mathematically expressed as

$$\frac{\sin i}{\sin r} = \mu_{12}$$

where i is the angle of incidence in media 1 and r is the angle of refraction in media 2.

- Total internal reflection is a special case of refraction wherein light travelling from a denser to a rarer media is incident at an angle greater than the critical angle:

$$\mu = \frac{1}{\sin i_C}$$

- Any transparent media bounded by two spherical surfaces or one spherical and one plane surface forms a lens.
- Images by lenses depend on the local length and the distance of the object from it.
- Convex lenses are converging while concave lenses are diverging.
- $$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$m = \frac{v}{u}$$

and $$\frac{1}{f} = \frac{1}{v} - \frac{1}{u}$$

are simple relationships between the focal length (f), the refractive index, the radii of curvatures (R_1, R_2), the object distance (u) and the image distance (v).

- Newton's formula can be used to measure the focal length of a lens.
- Displacement method is a very convenient way of finding focal length of a lens.

- Power of a lens indicates how diverging or converging it is:

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

Power is expressed in dioptre. (or m^{-1} in SI units)

- The focal length F of an equivalent lens when two their lenses of focal lengths f_1 and f_2 one kept in contact is given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2}$$

TERMINAL EXERCISES

1. List the uses of concave and convex mirrors.

2. What is the nature and position of image formed when the object is at (i) infinity (ii) $2f$ (iii) f in case of concave mirror and convex mirror.

3. List the factors on which lateral displacement of an incident ray depends as it suffers refraction through a parallel-sided glass slab? Why is the lateral displacement larger if angle of incidence is greater. Show this with the help of a ray diagram.

4. State conditions for total internal reflection of light to take place.

5. How is +1.5 dioptre different from –1.5 dioptre? Define dioptre.

6. Why does the intensity of light become less due to refraction?

7. A lamp is 4m from a wall. Calculate the focal length of a concave mirror which forms a five times magnified image of the lamp on the wall. How far from the wall must the mirror be placed?

8. A dentist's concave mirror has a radius of curvature of 30cm. How far must it be placed from a cavity in order to give a virtual image magnified five times?

9. A needle placed 45cm from a lens forms an image on a screen placed 90cm on the other side of the lens. Identify the type of the lens and determine its focal length. What is the size of the image, if the size of the needle is 5.0cm?

10. An object of size 3.0cm is placed 14cm in front of a concave lens of focal length 21 cm. Describe the nature of the image by the lens. What happens if the object is moved farther from the lens?

11. An object is placed at a distance of 100cm from a double convex lens which forms a real image at a distance of 20cm. The radii of curvature of the surfaces

of a lens are 25cm and 12.5 cm respectively. Calculate the refractive index of the material of the lens.

12. A ray of light is travelling from diamond to glass. Calculate the value of the critical angle for the ray, if the refractive index of glass is 1.51 and that of diamond 2.47.

13. A small object is placed at a distance of 15cm from two coaxial thin convex lenses in contact. If the focal length of each lens is 20cm. Calculate the focal length and the power of the combination and the distance between the object and its image.

14. While finding the focal length of a convex lens, an object was kept at a distance of 65.0 cm from the screen. Tow positions of the lens for which clear image of the object was formed on the screen were obtained. The distance between these two positions was found to be 15 cm. Calculate the focal length of the given lens.

ANSEWERS TO INTEXT QUESTIONS

20.1

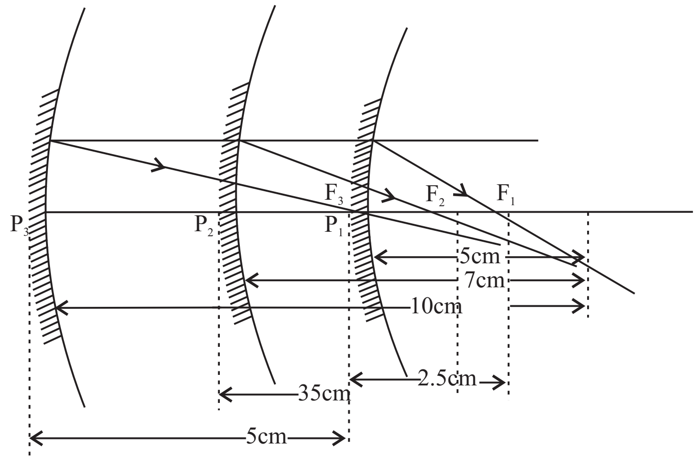
1. (a) plane mirror (its radius of curvature is infinitely large).

   (b) No. The focal length of a spherical mirror is half of its radius of curvature ($f \simeq R/2$) and has nothing to do with the medium in which it is immersed.

   (c) Virtual

   (d) This is because the rays parallel to the principal axis converge at the focal point *F*; and the rays starting from *F*, after reflection from the mirror, become parallel to the principal axis. Thus, *F* serves both as the first and the second focal point.

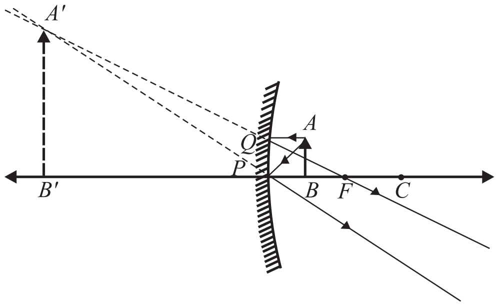
2. Focal lengths : 2.5cm, 3.5cm, 5cm.



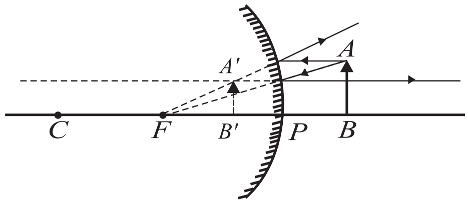
3. $f = -15$cm; $f = +15$cm.
4. The dish antennas are curved so that the incident parallel rays can be focussed on the receiver.

20.2

1. The upper part of the mirror must be convex and its lower part concave.
2. Objects placed close to a concave mirror give an enlarged image. Convex mirrors give a diminished erect image and have a larger field of view.

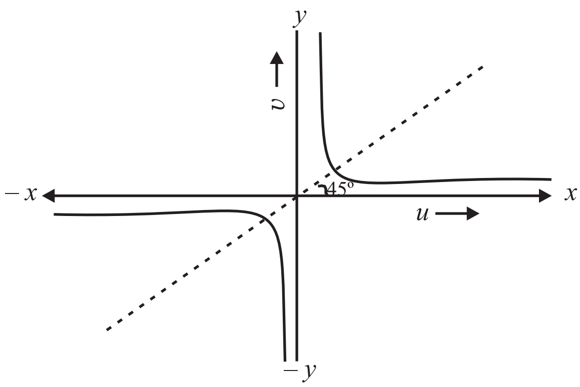


(a) Image formed by concave mirror

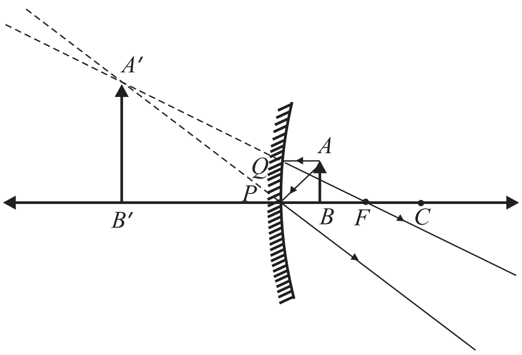


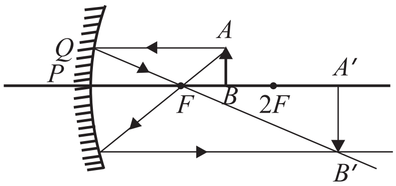
(b) Image formed by convex mirror

3. for $|u| > f$, we get real image; $u = -2f$ is a special case when an object kept as the centre of curvature of the mirror forms a real image at this point itself ($v = -2f$). For $u < f$, we get virtual image.



4. When (i) $u < f$, and (ii) $f < u < 2f$.





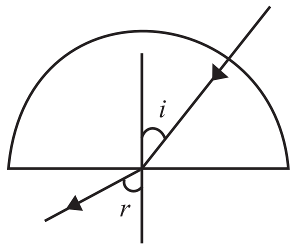
5. (i) 12cm in front of mirror, real and inverted, (ii) 0.8cm

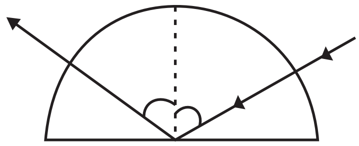
6. $v = -60$cm, $R = -24$cm 7. $u = -10$cm, $v = +5$cm

8. $v = 4$cm

20.3

1. No lateral displacement.



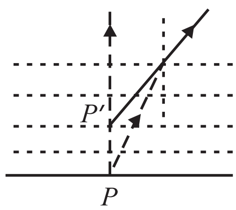


2. $\angle r > \angle i$ when $\angle i < \angle i$ Total internal reflection where $\angle i > \angle i_c$

3. The density of air and hence its refractive index decrease as we go higher in altitude. As a result, the light rays from the Sun, when it is below the horizon, pass from the rarer to the denser medium and bend towards the normal, till they are received by the eye of the observer. This causes the shape to appear elongated.

4. Due to the change in density of the different layers of air in the atmosphere, μ changes continuously. Therefore, the refractive index of air varies at different levels of atmosphere. This along with air currents causes twinkling of stars.

5. Due to refraction point P appears at P'.



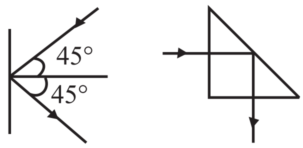
6. 36.2^0

20.4

1. Total internal reflection cannot take place if the ray travels from a rarer to a denser medium as the angle of refraction will always be less than the angle of incidence.

2. Yes the critical angle will change as

$$\mu_{ag} = \frac{1}{\sin i_c} \qquad \mu_{\omega g} = \frac{\mu_{ag}}{\mu_{aw}}$$

3. 

The intensity in the second case is more due to total internal reflection.

4. 20cm, $i_c = \sin^{-1} 0.8$

20.5

2. No. Changing the position of R_1 and R_2 in the lens maker's formula does not affect the value of *f*. So the image will be formed in the same position.

3. Substitute $R_1 = R$; $R_2 = -R$ and $\mu = 1.5$ in the lens maker's formula. You will get $f = R$.

4. Concave lens. But it is shaped like a convex lens.

5. This happens when the refractive index of the material of the lens is the same as that of the liquid.

6. $f = 22.2$ cm and $P = 4.5$ dioptre

7. Yes, by placing a convex and a concave lens of equal focal length in contact.

8. – 40cm, – 2.5 dioptre

Answers to Problems in Terminal Exercise

7. $f = -0.83$, 5m. 8. 12cm

9. $f = 30$cm, size of image = 10cm, converging lens

10. The image is erect, virtual and diminished in size, and located at 8.4cm from the lens on the same side as the object. As the object is moved away from the lens, the virtual image moves towards the focus of the lens but never beyond and progressively diminshes in size.

11. $\mu = 1.5$ 12. 37.7º

13. 10cm, 10 dioptre, 45 cm.

14. $f = 15.38$ cm

312en21

21

DISPERSION AND SCATTERING OF LIGHT

In the previous lesson you have learnt about reflection, refraction and total internal reflection of light. You have also learnt about image formation by mirrors and lenses and their uses in daily life. When a narrow beam of ordinary light is refracted by a prism, we see colour bands. This phenomenon has to be other than reflection or refraction. *The splitting of white light into its constituent colours or wavelengths by a medium is called* ***dispersion***. In this lesson, you will study about this phenomenon. A beautiful manifestation of this phenomenon in nature is in the form of rainbow. You will also learn in this lesson about the phenomenon of scattering of light, which gives sky its blue colour and the sun red colour at sunrise and sunset. Elementary idea of Raman effect will also be discussed in this lesson.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain dispersion of light;*
- *derive relation between the angle of deviation* (δ), *angle of prism (A) and refractive index of the material of the prism* (μ);
- *relate the refractive index with wavelength and explain dispersion through a prism;*
- *explain formation of primary and secondary rainbows;*
- *explain scattering of light and list its applications.and; and*
- *explain Raman effect.*

21.1 DISPERSION OF LIGHT

Natural phenomena like rings around planets (halos) and formation of rainbow etc. cannot be explained by the rectilinear propagation of light. To understand such events, light is considered as having wave nature. (You will learn about it in the next lesson.) As you know, light waves are transverse electromagnetic waves which propagate with speed 3×10^8 ms^{-1} in vacuum. Of the wide range of electromagnetic spectrum, the visible light forms only a small part. Sunlight consists of seven different wavelengths corresponding to seven colours. Thus, colours may be identified with their wavelengths. You have already learnt *that the speed and wavelength of waves change when they travel from one medium to another.* The speed of light waves and their corresponding wavelengths also change with the change in the medium. The speed of a wave having a certain wavelength becomes less than its speed in free space when it enters an optically denser medium.

The refractive index μ has been defined as the ratio of the speed of light in vacuum to the speed of light in the medium. It means that the refractive index of a given medium will be different for waves having wavelengths 3.8×10^{-7} m and 5.8×10^{-7} m because these waves travel with different speeds in the same medium. This ***variation of the refractive index of a material with wavelength is known as dispersion.*** This phenomenon is different from refraction. In free space and even in air, the speeds of all waves of the visible light are the same. So, they are not separated. (Such a medium is called a non-dispersive medium.) But in an optically denser medium, the component wavelengths (colours) travel with different speeds and therefore get separated. Such a medium is called *dispersive medium*. Does this suggest that light will exhibit dispersion whenever it passes through an optically denser medium. Let us learn about it now.

21.1.1 Dispersion through a Prism

The separation of colours by a medium is not a sufficient condition to observe dispersion of light. These colours must be widely separated and should not mix up again after emerging from the dispersing medium. A glass slab (Fig. 21.1) is not suitable for observing dispersion as the rays of the emergent beam are very close and parallel to the incident beam

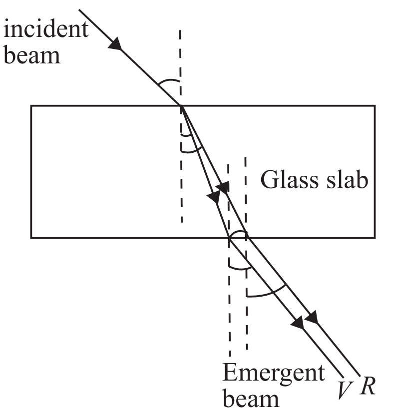


Fig. 21.1 : Passage of light through a glass slab

Newton used a prism to demonstrate dispersion of light. Refer to Fig. 21.2. White light from a slit falls on the face *AB* of the prism and light emerging from face *AC* is seen to split into different colours. Coloured patches can be seen on a screen. The face *AC* increases the separation between the rays refracted at the face *AB*. The incident white light *PQ* thus splits up into its component seven colours : Violet, indigo, blue, green, yellow, orange and red (VIBGYOR). The wavelengths travelling with different speeds are refracted through different angles and are thus separated. This *splitting of white light into component colours is known as dispersion*. *MR* and *MV* correspond to the red and violet light respectively. These colours on the screen produce the *spectrum.*

The bending of the original beam *PQN* along *MR* and *MV* etc. is known as ***deviation***. The angle between the emergent ray and the incident ray is known as the **angle of deviation.** Thus δ_v and δ_r represent the angles of deviation for violet light and red light, respectively.

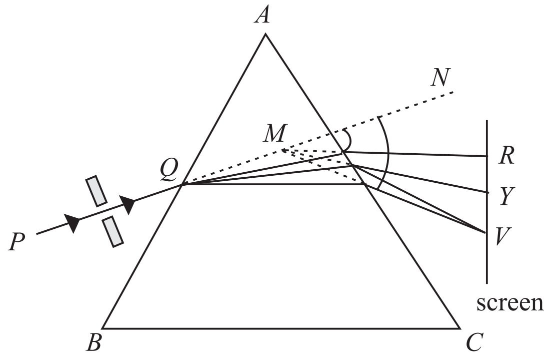


Fig. 21.2 : Dispersion of light by a prism

Read the following example carefully to fix the ideas on variation of the refractive index with the wavelength of light.

Example 21.1: A beam of light of average wavelength 600nm, on entering a glass prism, splits into three coloured beams of wavelengths 384 nm, 589 nm and 760 nm respectively. Determine the refractive indices of the material of the prism for these wavelengths.

Solution : The refractive index of the material of the prism is given by

$$\mu = \frac{c}{\upsilon}$$

where *c is* speed of light in vacuum, and υ is speed of light in the medium (prism).

Since velocity of a wave is product of frequency and wavelength, we can write

$$c = \nu\lambda_a \qquad \text{and} \qquad \upsilon = \nu\lambda_m$$

where λ_a and λ_m are the wavelengths in air and medium respectively and υ is the frequency of light waves. Thus

$$\mu = \frac{\nu\lambda_a}{\nu\lambda_m} = \frac{\lambda_a}{\lambda_m}$$

For 384 nm wavelength, the refractive index is

$$\mu_1 = \frac{600\times10^{-9}\,\text{m}}{384\times10^{-9}\,\text{m}} = 1.56$$

For wave length of 589 nm :

$$\mu_2 = \frac{600\times10^{-9}\,\text{m}}{58.9\times10^{-9}\,\text{m}} = 1.02$$

and for 760nm wavelength :

$$\mu_3 = \frac{600\times10^{-9}\,\text{m}}{760\times10^{-9}\,\text{m}} = 0.8$$

We have seen that the refractive index of a material depends on

- the nature of the material, and
- the wavelength of light.

An interesting outcome of the above example is that the variation in wavelength ($\Delta\lambda = \lambda_2 - \lambda_1$) produces variation in the refractive index ($\Delta\mu = \mu_2 - \mu_1$). The ratio $\frac{\Delta\mu}{\Delta\lambda}$ is known as the spectral *dispersive power of the material of prism* .

21.1.2 The Angle of Deviation

We would now establish the relation between the angle of incidence i, the angle of deviation δ and the angle of prism A. Let us consider that a monochromatic beam of light PQ is incident on the face AB of the principal section of the prism ABC [Fig.21.3]. On refraction, it goes along QR inside the prism and emerges along RS from face AC. Let $\angle A \equiv \angle BAC$ be the refracting angle of the prism. We draw normals NQ and MR on the faces AB and AC, respectively and produce them backward to meet at O. Then you can easily convince yourself that $\angle NQP = \angle i$, $\angle MRS = \angle e$, $\angle RQO = \angle r_1$, and $\angle QRO = \angle r_2$ are the angle of incidence, the angle of emergence and the angle of refraction at the faces AB and AC, respectively. The angle between the emergent ray RS and the incident ray PQ at D is known as the angle of deviation (δ).

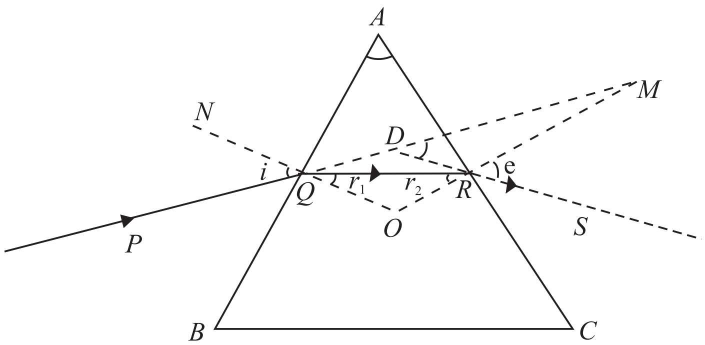


Fig. 21.3 : Refraction through a prism

Since $\angle MDR = \angle\delta$, As it is the external angle of the triangle QDR, we can write

$$\angle\delta = \angle DQR + \angle DRQ$$

$$= (\angle i - \angle r_1) + (\angle e - \angle r_2)$$

or $$\angle\delta = (\angle i + \angle e) - (\angle r_1 + \angle r_2) \qquad (21.1)$$

You may recall that the sum of the internal angles of a quadrilateral is equal to 360º. In the quadrilateral $AQOR$, $\angle AQO = \angle ARO = 90^0$, since NQ and MR are normals on faces AB and AC, respectively. Therefore

$$\angle QAR + \angle QOR = 180^0$$

or $$\angle A + \angle QOR = 180^0 \qquad (21.2)$$

But in ΔQOR

$$\angle OQR + \angle QRO + \angle QOR = 180^0$$

or $$\angle r_1 + \angle r_2 + \angle QOR = 180^0 \qquad (21.3)$$

On comparing Eqns. (21.2) and (21.3), we have

$$\angle r_1 + \angle r_2 = \angle A \qquad (21.4)$$

Combining this result with Eqn. (21.1), we have

$$\angle\delta = (\angle i + \angle e) - \angle A$$

or $$\angle i + \angle e = \angle A + \angle\delta \qquad (21.5)$$

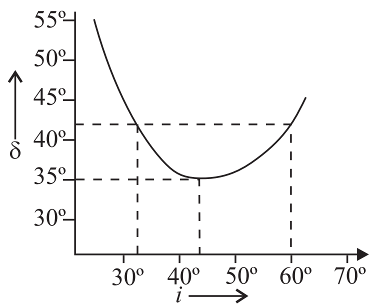


Fig. 21.4 : Plot between angle of incidence *i* and angle of deviation δ

Angle of Minimum Deviation

If we vary the angle of incidence *i,* the angle of deviation δ also changes; it becomes minimum for a certain value of *i* and again starts increasing as *i* increases further (Fig. 21.4). The minimum value of the angle of deviation is called *angle of minimum deviation* (δ_m). It depends on the material of the prism and the wavelength of light used. In fact, one angle of deviation may be obtained corresponding to two values of the angles of incidence. Using the principle of reversibility of light, we find that the second value of angle of incidence corresponds to the angle of emergence (*e*). In the minimum deviation position, there is only one value of the angle of incidence. So we have

$$\angle e = \angle i$$

Using this fact in Eqn.(21.5) and replacing δ by δ_m, we have

$$\angle i = \frac{\angle A + \angle\delta_m}{2} \qquad (21.6)$$

Applying the principle of reversibility of light rays and under the condition $\angle e = \angle i$, we can write $\angle r_1 = \angle r_2 = \angle r$, say

On substituting this result in Eqn. (21.4), we get

$$\angle r = \frac{\angle A}{2} \tag{21.7}$$

The light beam inside the prism, under the condition of minimum deviation, passes symmetrically through the prism and is parallel to its base. The refractive index of the material of the prism is therefore given by

$$\mu = \frac{\sin i}{\sin r} = \frac{\sin\left(\frac{A+\delta_m}{2}\right)}{\sin\frac{A}{2}} \tag{21.8}$$

The refractive index μ can be calculated using Eqn.(21.8) for a monochromatic or a polychromatic beam of light. The value of δ_m is different for different colours. It gives a unique value of the angle of incidence and the emergent beam is brightest for this incidence.

For a prism of small angle *A*, keeping *i* and *r* small, we can write

$$\sin i = i, \sin r = r, \text{ and } \sin e = e$$

Hence

$$\mu = \frac{\sin i}{\sin r_1} = \frac{i}{r_1} \quad \text{or} \quad i = \mu r_1$$

Also

$$\mu = \frac{\sin e}{\sin r_2} = \frac{e}{r_2} \quad \text{or} \quad e = \mu r_2$$

Therefore,

$$\angle i + \angle e = \mu (\angle r_1 + \angle r_2)$$

Using this result in Eqns. (26.4) and (26.5), we get

$$\mu \angle A = \angle A + \angle\delta$$

or

$$\angle\delta = (\mu - 1)\angle A \tag{21.9}$$

We know that μ depends on the wavelength of light. So deviation will also depend on the wavelength of light. That is why δ_V is different from δ_R. Since the velocity of the red light is more than that of the violet light in glass, the deviation of the red light would be less as compared to that of the violet light.

$$\delta_V > \delta_R.$$

This implies that $\mu_V > \mu_R$. This change in the refractive index of the material with the wavelength of light is responsible for dispersion phenomenon.

21.1.3 Angular Dispersion and Dispersive Power

The difference between the angles of deviation for any two wavelengths (colours) is known as the *angular dispersion* for those wavelengths. The angular dispersion between the red and violet wavelengths is $\delta_V - \delta_R$. In the visible part of the spectrum, the wavelength of the yellow colour is nearly the average wavelength of the spectrum. The deviation for this colour δ_Y may, therefore, be taken as the average of all deviations.

The ***ratio of the angular dispersion to the mean deviation is taken as the dispersive power*** (ω) of the material of the prism :

$$\omega = \frac{\delta_V - \delta_R}{\delta_Y}$$

We can express this result in terms of the refractive indices using Eqn. (21.9) :

$$\omega = \frac{(\mu_V - 1)\angle A - (\mu_R - 1)\angle A}{(\mu_V - 1)\angle A}$$

$$= \frac{\mu_V - \mu_R}{\mu_Y - 1} = \frac{\Delta\mu}{\mu - 1} \qquad (21.10)$$

Example 21.2 : The refracting angle of a prism is 30′ and its refractive index is 1.6. Calculate the deviation caused by the prism.

Solution : We know that $\delta = (\mu - 1)\angle A$

On substituting the given data, we get

$$\delta = (1.6 - 1) \times \frac{1^\circ}{2} = \frac{0.6}{2} = 0.3^\circ = 18'$$

Example 21.3 : For a prism of angle A, the angle of minimum deviation is $A/2$. Calculate its refractive index, when a monochromatic light is used. Given $A = 60^\circ$

Solution : The refractive index is given by

$$\mu = \frac{\sin\left(\frac{A + \delta_m}{2}\right)}{\sin(A/2)}$$

Now $\delta_m = A/2$ so that

$$\mu = \frac{\sin\left(\frac{A + A/2}{2}\right)}{\sin(A/2)} = \frac{\sin\left(\frac{3}{4}A\right)}{\sin\left(\frac{A}{2}\right)} = \frac{\sin\left(\frac{3}{4}A\right)}{\sin\left(\frac{A}{2}\right)} = \sqrt{2} = 1.4$$

Notes

INTEXT QUESTIONS 21.1

1. Most ordinary gases do not show dispersion with visible light. Why?
2. With your knowledge about the relative values of μ for the component colours of white light, state which colour is deviated more from its original direction?
3. Does dispersion depend on the size and angle of the prism?
4. Calculate the refractive index of an equilateral prism if the angle of minimum deviation is equal to the angle of the prism.

Rainbow formation

Dispersion of sunlight through suspanded water drops in air produces a spectracular effect in nature in the form of rainbow on a rainy day. With Sun at our back, we can see a brighter and another fainter rainbow. The brighter one is called the ***primary*** rainbow and the other one is said to be ***secondary rainbow***. Sometimes we see only one ***rainbow***. The bows are in the form of coloured arcs whose common centre lies at the line joining the Sun and our eye. Rainbow can also be seen in a fountain of water in the evening or morning when the sun rays are incident on the water drops at a definite angle.

Primary Rainbow

The primary rainbow is formed by two refractions and a single internal reflection of sunlight in a water drop. (See Fig. 21.5(a)). Descartes explained that rainbow is seen through the rays which have suffered minimum deviation. Parallel rays from the Sun suffering deviation of 137º.29′ or making an angle of 42º.31′ at the eye with the incident ray, after emerging from the water drop, produce bright shining colours in the bow. Dispersion by water causes different colours (red to violet) to make their own arcs which lie within a cone of 43º for red and 41º for violet rays on the outer and inner sides of the bow (Fig. 21.5 (b)).

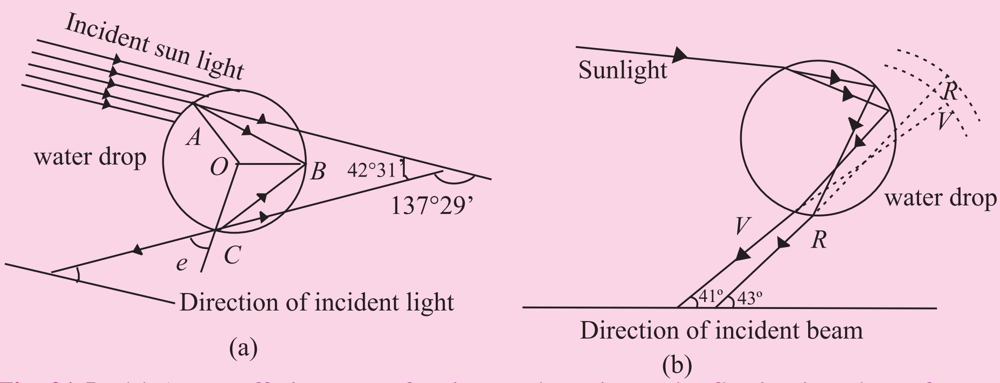


Fig. 21.5 : (a) A ray suffering two refractions and one internal reflection in a drop of water. Mean angle of minimum deviation is 137º29′, and **(b)** dispersion by a water drop.

Notes

Secondary Rainbow

The secondary rainbow is formed by two refractions and two internal reflections of light on the water drop. The angles of minimum deviations for red and violet colours are 231°. and 234°. respectively, so they subtend a cone of 51°. for the red and 54°. for the violet colour. From Fig.21.6 it is clear that the red colour will be on the inner and the violet colour on the outer side of the bow.

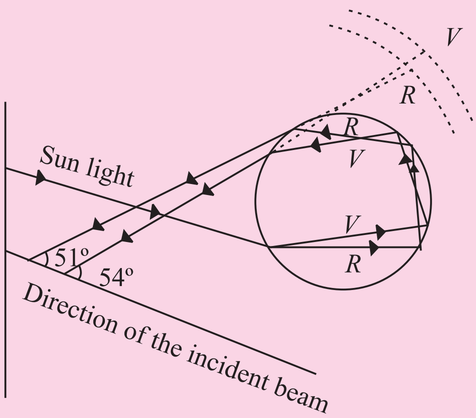


Fig. 21.6 : Formation of the secondary rainbow

The simultaneous appearance of the primary and secondary rainbows is shown in Fig.21.7. The space between the two bows is relatively dark. Note that the secondary rainbow lies above the primary bow.

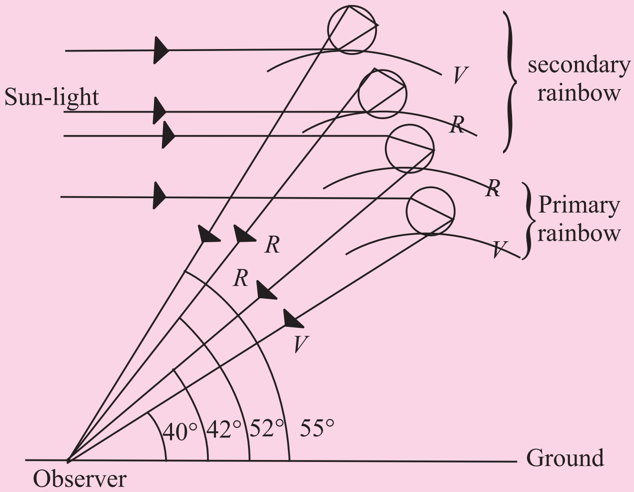


Fig. 21.7 : Simultaneous formation of the primary and secondary rainbow.

21.2 SCATTERING OF LIGHT IN ATMOSPHERE

On a clear day when we look at the sky, it appears blue. But the clouds appear white. Similarly, production of brilliant colours when sunlight passes through jewels and crystals also attracts our attention. You may like to know : How and why does it happen? These phenomena can be explained in terms of *scattering of light*. A solution of dust or particle-free benzene exposed to sunlight gives brilliant blue colour when looked sideways.

21.2.1 Scattering of Light

This phenomenon involves interaction of radiation with matter. Tiny dust particles are present in Earth's atmosphere. When sunlight falls on them, it gets diffused in all directions. That is why light reaches even those nooks and corners where it normally is not able to reach straight from the source.

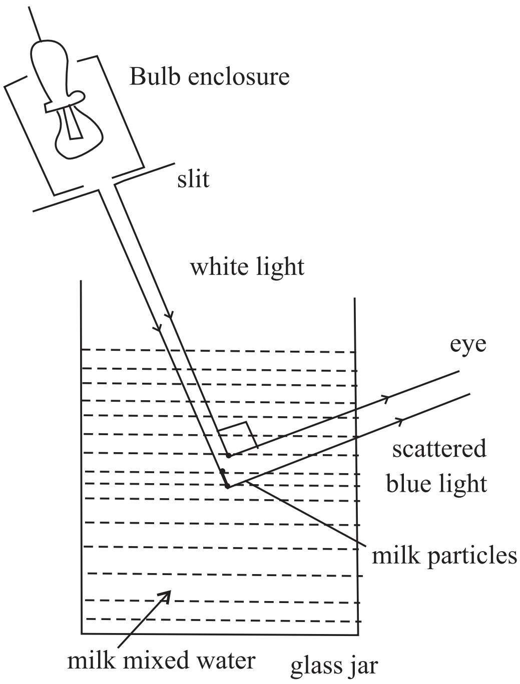


Fig. 21.8 : The scattering of light from milk particles

Let us perform a simple activity.

ACTIVITY 21.1

Take a glass jar or a trough, fill it with water and add a little milk to it. Now allow a narrow beam of light from a white bulb to fall on it. Observe the light at 90^o. You will see a bluish beam through water. This experiment shows that after scattering, the wavelenghts of light become a peculiarly different in a given direction (Fig. 21.14).

The phenomenon of scattering is a two step process : absorption of light by the scattering particle and then instant re-emission by it in all possible directions. Thus, this phenomenon is different from reflection. The scattered light does not obey the laws of reflection. It is important to note that the size of the particle must be less than the wavelength of light incident on it. A bigger sized particle will scatter all the wavelengths equally. The intensity of scattered light is given by *Rayleigh's law* of scattering. According to this law, ***the intensity of scattered light is inversely proportional to the fourth power of its wavelength***:

$$I \alpha \frac{1}{\lambda^4}$$

Here I is intensity and λ is wavelength of the scattered light. Thus, when white light is incident on the scattering particle, the blue light is scattered the most and the red light is scattered the least.

Example 21.4 : Waves of wavelength 3934Å, 5890Å and 6867Å are found in the scattered beam when sunlight is incident on a thin layer of chimeny smoke. Which of these is scattered more intensely?

Solution : The intensity of scattered light is given by

$$I \,\alpha\, \frac{1}{\lambda^4}$$

Since 3934Å is the smallest wavelength, it will be scattered most intensely.

On the basis of scattering of light, we can explain why sky appears blue, clouds appear white and the sun appears red at sunrise as well as at sunset.

C.V. Raman
(1888 – 1970)

Chandra Shekhar Venkat Raman is the only Indian national to receive Nobel prize (1930) in physics till date. His love for physics was so intense that he resigned his job of an officer in Indian finance department and accepted the post of Palit Professor of Physics at the Department of Physics, Calcutta University. His main contributions are : Raman effect on scattering of light, molecular diffraction of light, mechanical theory of bowed strings, diffraction of X-rays, theory of musical instruments and physics of crystals.

As Director of Indian Institute of Science, Bangalore and later as the founder Director of Raman Research Institute, he did yeoman's to Indian science and put it on firm footings in pre-independence period.

(A) Blue Colour of the Sky

We know that scattering of light by air molecules, water droplets or dust particles present in the atmosphere can be explained in accordance with Rayleigh's law. The shorter wavelengths are scattered more than the longer wavelengths. Thus, the blue light is scattered almost six times more intensely than the red light as the wavelength of the blue light is roughly 0.7 times that of the red. The scattered light becomes rich in the shorter wavelengths of violet, blue and green colours. On further scattering, the violet light does not reach observe's eye as the eye is comparatively less sensitive to violet than blue and other wavelengths in its neighbourhood. So, when we look at the sky far away from the sun, it appears blue.

Example 21.5 : What will be the colour of the sky for an astronant in a spaceship flying at a high attitude.

Solution : At a high attitude, in the absence of dust particle and air molecules, the sunlight is not scattered. So, the sky will appear black.

(B) White colour of the clouds

The clouds are formed by the assembly of small water drops whose size becomes more than the average wavelength of the visible light (5000Å). These droplets scatter all the wavelengths with almost equal intensity. The resultant scattered light is therefore white. So, a thin layer of clouds appears white. What about dense clouds?

(C) Red colour of the Sun at Sunrise and Sunset

We are now able to understand the red colour of the Sun at sunrise and sunset. In the morning and evening when the Sun is near the horizon, light has to travel a greater distance through the atmosphere. The violet and blue wavelengths are scattered by dust particles and air molecules at an angle of about 90°. The sunlight thus becomes devoid of shorter wavelengths and the longer wavelength of red colour reaches the observer (Fig. 21.9). So the Sun appears to us as red.

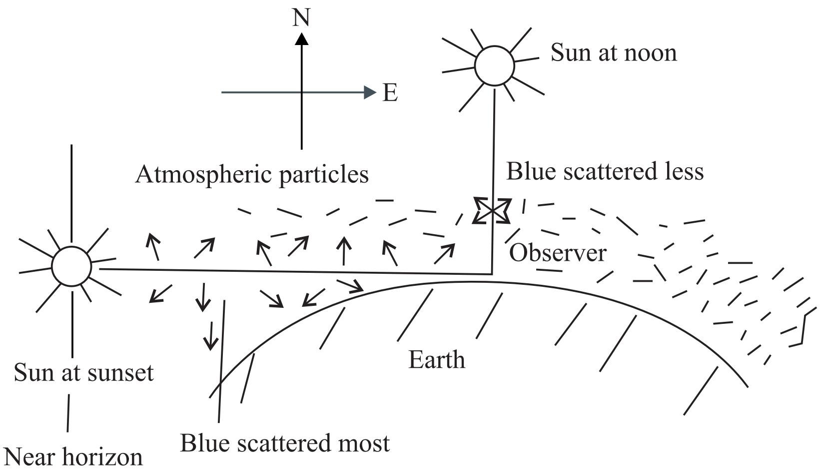


Fig. 21.9 : Red colour of the sun at sunset and sunrise (blue is scattered away).

At noon, the Sun is overhead and its distance from the observer is comparatively less. The blue colour is also scattered less. This results in the Sun appearing white, as a matter of fact, crimson.

21.2.2 Raman effect

When light radiation undergoes scattering from a transparent substance (solid, liquid or gas) then the frequency of the scattered radiation may be greater or less than the frequency of the incident radiation. This phenomenon is known as Raman effect as it was first observed by C. V. Raman in 1926. An analogue of this optical phenomenon was observed earlier by A. H. Compton in connection with the scattering of X-rays. The spectrum of the scattered radiation is known as Raman spectrum. This has lines having frequency greater than the frequency of the incident radiation (known as anti-Stokes' lines) as also lines having frequency less than the frequency of the incident radiation (called Stokes' lines).

A simple explanation of Raman effect can be given as follows. When light radiation interacts with a substance three possibilities may arise. In the first possibility, the light radiation interacting with the substance does not undergo any change of energy. Hence, its frequency remains unchanged. In the second possibility, the light radiation may impart some of its energy to the substance. As a result, the energy of the light radiation decreases. This leads to a decrease in the frequency of the scattered radiation (corresponding to Stokes' lines). In the third possibility, the incident radiation may interact with the substance which is already in the excited state. In the process, the radiation gains energy resulting into increase in its frequency (corresponding to anti-Stokes' lines).

Raman effect has lot of applications in various fields. C. V. Raman was awarded Nobel prize in physics for this discovery in 1930.

INTEXT QUESTIONS 21.2

1. Why dense clouds appear black?
2. Why does the sky appear deep blue after rains on a clear day?
3. Can you suggest an experiment to demonstrate the red colour of the Sun at sunrise and sunset?
4. The photographs taken from a satellite show the sky dark. Why?
5. What are anti stokes' lines?

WHAT YOU HAVE LEARNT

- Light of single wavelength or colour is said to be monochromatic but sunlight, which has several colours or wavelengths, is polychromatic.
- The splitting of light into its constituent wavelengths on entering an optically denser medium is called dispersion.
- A prism is used to produce dispersed light, which when taken on the screen, forms the spectrum.
- The angle of deviation is minimum if the angles of incidence and emergence become equal. In this situation, the beam is most intense for that colour.

- The angle of deviation and refractive index for a small-angled prism are connected by the ralation $\delta = (\mu - 1)A$.
- The rainbow is formed by dispersion of sunlight by raindrops at definite angles for each colour so that the condition of minimum deviation is satisfied.
- Rainbows are of two types : primary and secondary. The outer side of the primary rainbow is red but the inner side is violet. The remaining colours lie in between to follow the order (VIBGYOR). The scheme of colours gets reversed in the secondary rainbow.
- The blue colour of the sky, the white colour of clouds and the reddish colour of the Sun at sunrise and sunset are due to scattering of light. The intensity of scattered light is inversely proportional to the fourth power of the wavelength $\left(I \propto \frac{1}{\lambda^4}\right)$. This is called Rayleigh's law. So the blue colour is scattered more than the red.
- When light radiation undergoes scattering from a transparent substance, then frequency of scatered radiation may be greater or less than frequency of incident ratiation. This phenomenon is known as Raman effect.

TERMINAL EXERCISE

1. For a prism, show that $i + e = A + \delta$.
2. Would you prefer small-angled or a large-angled prism to produce dispersion. Why?
3. Under what condition is the deviation caused by a prism directly proportional to its refractive index?
4. Explain why the sea water appears blue at high seas.
5. The angle of minimum deviation for a 60^0 glass prism is 39^0. Calculate the refractive index of glass.
6. The deviation produced for red, yellow and violet colours by a crown glass are 2.84^o, 3.28^o and 3.72^o respectively. Calculate the dispersive power of the glass material.
7. Calculate the dispersive power for flint glass for the following data : μ_C = 1.6444, μ_D= 1.6520 and μ_F= 1.6637, where C, D & F are the Fraunhofer nomenclatures.
8. A lens can be viewed as a combination of two prisms placed with their bases together. Can we observe dispersion using a lens. Justify your answer.
9. Human eye has a convex lens. Do we observe dispersion with unaided eye?

ANSEWERS TO INTEXT QUESTIONS

21.1

1. The velocity of propagation of waves of different wavelengths of visible light is almost the same in most ordinary gases. Hence, they do not disperse visible light. Their refractive index is also very close to 1.

2. Violet, because $\lambda_r > \lambda_v$ and the velocity of the red light is more than that of the violet light inside an optically denser medium.

3. No

4. $\mu = \dfrac{\sin 60^\circ}{\sin 30^\circ} = \sqrt{3} = 1.732$

21.2

1. It absorbs sunlight

2. It becomes clear of dust particles and bigger water molecules. The scattering now takes place strictly according to Rayleigh's law.

3. We can take sodium thiosulphate solution in a round bottom flask and add a small quality of sulphuric acid. On illuminating this solution with a high power bulb, we can see a scenario similar to the colour of the sun at sunrise and sunset.

4. At very high altitudes no centres (particles) of scattering of sunlight are present. So the sky appears dark.

5. The spectral lines having frequency greater than the frequency of incident radiation are known as anti stokes' lines.

Answers to Problems in Terminal Exercise

5. 1.5 6. 0.27

7. 0.03

312en22

22

WAVE PHENOMENA AND LIGHT

In the preceding two lessons of this module, you studied about reflection, refraction, dispersion and scattering of light. To understand these, we used the fact that light travels in a straight line. However, this concept failed to explain redistribution of energy when two light waves were superposed or their bending around corners. These observed phenomena could be explained only on the basis of wave nature of light. Christian Huygens, who was a contemporary of Newton, postulated that light is a wave and the wave theory of light was established beyond doubt through experimental observations on interference and diffraction. In this lesson, you will also learn about polarisation, which conclusively proved that light is a wave and transverse in nature.

OBJECTIVES

After studying this lesson, you should be able to :

- *state Huygens' principle and apply it to explain wave propagation*;
- *explain the phenomena of interference and diffraction of light*;
- *explain diffraction of light by a single-slit; and*
- *show that polarisation of light established its wave nature; and*
- *derive Brewster's law.*

22.1 HUYGENS' PRINCIPLE

Huygens' postulated that light is a wave, which travels through a hypothetical medium called ether. This hypothetical medium has the strange property of occupying all space, including vacuum! The vibrations from the source of light propagate in the form of waves and the energy carried by them is distributed equally in all directions.

The concept of wavefront is central ***Huygens'principle.*** Let us first understand what a wavefront is with the help of a simple activity.

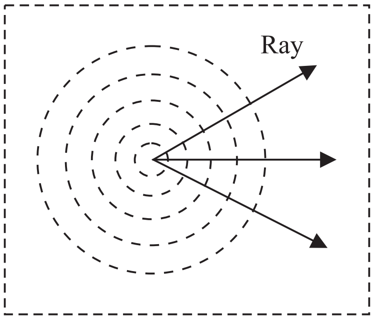


Fig. 22.1: Circular wavefronts on the surface of water

ACTIVITY 22.1

Take a wide based trough full of water and drop a small piece of stone in it. What do you observe? You will see that circular ripples due to the up and down motion of water molecules spread out from the point where the stone touched the water surface. If you look carefully at these ripples, you will notice that each point on the circumference of any of these ripples is in the same state of motion i.e., each point on the circumference of a ripple oscillates with the same amplitude and in the same phase. In other words, we can say that *the circumference of a ripple is the locus of the points vibrating in the same phase at a given instant and is known as* ***the wavefront.*** Therefore, the circular water ripples spreading out from the point of disturbance on the water surface represent a **circular wavefront.** Obviously, the distance of every point on a wavefront is the same from the point of disturbance, i.e., the source of waves.

For a point source emitting light in an isotropic medium, the locus of the points where all waves are in the same phase, will be a sphere. Thus, a point source of light ***emits spherical wavefronts.*** (In two dimensions, as on the water surface, the wavefronts appear circular.) Similarly, a line source of light emits ***cylindrical wavefronts.*** *The line perpendicular to the wavefront at a point represents the direction of motion of the wavefront at that point. This line is called the ray of light and a collection of such rays is called a beam of light.* When the source of light is at a large distance, any small portion of the wavefront can be considered to be a ***plane wavefront.***

The ***Huygens'principle*** states that

- Each point on a wavefront becomes a source of secondary disturbance which spreads out in the medium.
- The position of wavefront at any later instant may be obtained by drawing a forward common envelop to all these secondary wavelets at that instant.
- In an isotropic medium, the energy carried by waves is transmitted equally in all directions.

- If the initial shape, position, the direction of motion and the speed of the wavefront is known, its position at a later instant can be ascertained by geometrical construction. Note that the wavefront does not travel in the backward direction.

To visualise Huygens' construction, you may imagine a point source at the centre of a hollow sphere. The outer surface of this sphere acts as a primary wavefront. If this sphere is enclosed by another hollow sphere of larger radius, the outer surface of the second hollow sphere will act as a secondary wavefront. (The nearest mechanical analogue of such an arrangement is a football.) If the second sphere is further enclosed by another sphere of still bigger radius, the surface of the outermost (third) sphere becomes secondary wavefront and the middle (second) sphere acts as the primary wavefront. In two dimensions, the primary and secondary wavefronts appear as concentric circles.

22.1.1 Propagation of Waves

Now let us use Huygens' principle to describe the propagation of light waves in the form of propagation of wavefronts. Fig. 22.2 shows the shape and location of a plane wavefront *AB* at the time $t = 0$. You should note that the line *AB* lies in a plane perpendicular to the plane of the paper. Dots represented by *a*, *b*, *c*, on the wavefront *AB* are the sources of secondary wavelets. All these sources emit secondary wavelets at the same time and they all travel with the same speed along the direction of motion of the wavefront *AB*. In Fig. 22.2, the circular arcs represent the wavelets emitted from *a*, *b*, *c*, … taking each point as center. *These wavelets have been obtained by drawing arcs of radius*, $r = \upsilon t$, *where* υ *is the velocity of the wavefront and* t *is the time at which we wish to obtain the wavefront*, The tangent, *CD*, to all these wavelets represents the new wavefront at time $t = T$.

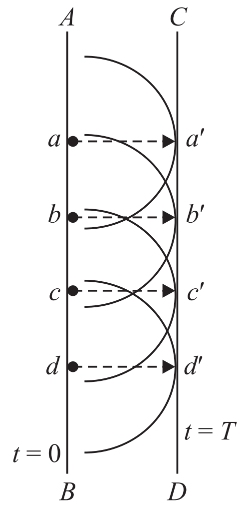


Fig. 22.2: Construction of a plane wave front

Let us take another example of Huygens' construction for an expanding circular wavefront. Refer to Fig. 22.3, which indicates a circular wavefront, centred at *O*, at time $t = 0$. Position *A*, *B*, *C* … represent point sources on this wavefront. Now to draw the wavefront at a later time $t = T$, what would you do? You should draw arcs from the points *A*, *B*, *C* …, of radius equal to the speed of the expanding wavefront multiplied by *T*. These arcs will

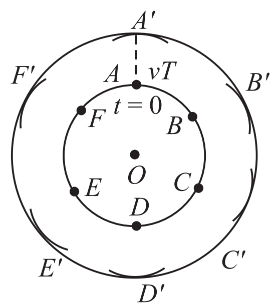


Fig. 22.3: Construction of circular wavefront using Huygens' principle

represent secondary wavelets. The tangents drawn to these arcs will determine the shape and location of the expanding circular wavefront at time *T*.

We hope you have now understood the technique of Huygens' construction. Now, you may like to know the ***physical significance of Huygens' construction.*** By determining the shape and location of a wavefront at a subsequent instant of time with the help of its shape and location at an earlier instant, we are essentially describing the propagation of the wavefront. Therefore, Huygens' construction enables us to describe wave motion.

INTEXT QUESTIONS 22.1

1. What is the relative orientation of a wavefront and the direction of propagation of the wave?

2. A source of secondary disturbance is emitting wavelets at an instant $t = 0$ s. Calculate the ratio of the radii of wavelets at $t = 3$s and $t = 6$s.

22.2 INTERFERENCE OF LIGHT

Let us first perform a simple activity:

ACTIVITY 22.2

Prepare a soap solution by adding some detergent powder to water. Dip a wire loop into the soap solution and shake it. When you take out the wire loop, you will find a thin film on it. Bring this soap film near a light bulb and position yourself along the direction of the reflected light from the film. You will observe beautiful colours. Do you know the reason? To answer this question, we have to understand the phenomenon of ***interference of light.*** In simple terms, ***interference of light refers to redistribution of energy due to superposition of light waves from two coherent sources.*** The phenomenon of interference of light was first observed experimentally by Thomas Young in 1802 in his famous two-slit experiment. This experimental observation played a significant role in establishing the wave theory of light. The basic theoretical principle involved in the phenomenon of interference as well as diffraction of light is the *superposition principle*.

22.2.1 Young's Double Slit Experiment

Young's experimental set up is shown schematically in Fig. 22.4. In his experiment, sunlight was allowed to pass through a pin hole S and then, at some distance away, through two pin holes S_1 and S_2 equidistant from S and close to each other.

According to Huygens' wave theory of light, spherical wavefronts would spread out from the pin hole S which get divided into two wavefronts by S_1 and S_2. If S is illuminated by a monocromatic source of light, such as sodium, these act as coherent sources and in-phase waves of equal amplitude from these sources superpose as they move beyond S_1S_2. As a consequence of superposition (of the two sets of identical waves from S_1 and S_2), redistribution of energy takes place and a pattern consisting of alternate bright and dark fringes is produced on the screen such as placed at C. Let us now learn the explanation of the observed fringe pattern in the Young's interference experiment.

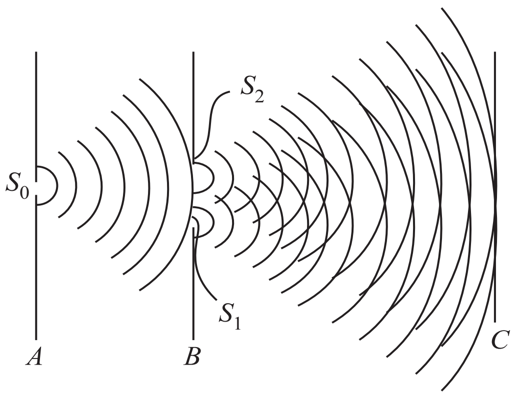


Fig. 22.4: Schematic arrangement of Young's double-slit experiment

Euygene Thomas Young (1773-1829)

Born on 16 June, 1773, Euygene Thomas Young will always be known for his study on the human ear, the human eye, how it focuses and on astigmatism. His research on colour blindness led him to the three component theory of colour vision. Working on human ears and eyes, he dedicated much time to the speed of sound and light. He knew that if two sound waves of equal intensity reached the ear 180º out of phase, they cancelled out each other's effect and no sound was heard. It occurred to him that a similar interference effect should be observed with two light beams, if light consisted of waves. This led Young to devise an experiment, now commonly referred to as the Young's double-slit experiment.

In his later years, Young devoted most of his time deciphering the Egyptian hieroglyphics found on the Rosetta stone discovered in the Nile Delta in 1799.

(a) Constructive Interference: You may recall from the superposition principle that some points on the screen *C* will have maximum displacement (or amplitude) because the crests due to one set of waves coincide with the crests due to another set of waves. In other words, at this point, the waves arrive in-phase and hence the total amplitude is much higher than the amplitude of individual waves. The same holds true for the points where the troughs due to one set of waves coincide with the troughs due to another set. Such points will appear bright because the intensity of light wave is proportional to the square of the amplitude. Superposition of waves at these points leads to what is known as ***constructive interference.***

(b) Destructive Interference: The points where the crests due to one set of waves coincide with the troughs due to the other set and vice-versa, the total amplitude is zero. It is so because the waves reach these points completely out of phase. Such points appear dark on the screen. These points correspond to ***destructive interference.***

(c) Intensity of fringes: To analyse the interference pattern, we calculate the intensity of the bright and dark fringes in the interference pattern for harmonic waves. Refer to Fig. 22.5, which is schematic representation of the geometry of Young's experiment. The phenomenon of interference arises due to superposition of two harmonic waves of same frequency and amplitude but differing in phase. Let the phase difference between these two waves be δ. We can write y_1 and y_2, the displacements at a fixed point *P* due to the two waves, as

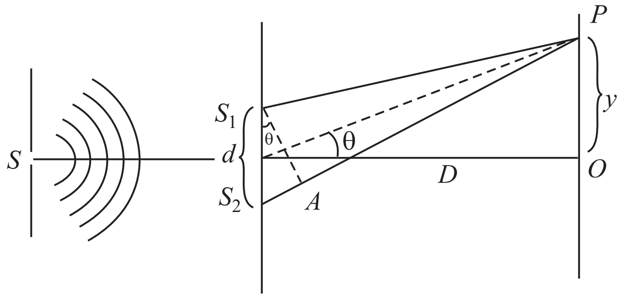


Fig. 22.5: Geometry of Young's double slit experiment

$$y_1 = a \sin \omega t$$

and $$y_2 = a \sin (\omega t + \delta)$$

where δ signifies the phase difference between these waves. Note that we have not included the spatial term because we are considering a fixed point in space.

According to the principle of superposition of waves, the resultant displacement is given by

$$y = y_1 + y_2$$

$$= a \sin \omega t + a \sin (\omega t + \delta)$$

$$= a [\sin \omega t + \sin (\omega t + \delta)]$$

$$= 2a \sin (\omega t + \frac{\delta}{2}) \cos\left(-\frac{\delta}{2}\right)$$

$$= A \sin (\omega t + \frac{\delta}{2})$$

where amplitude of the resultant wave is given by

$$A = 2a \cos (\delta/2).$$

The intensity of the resultant wave at point P can be expressed as

$$I \propto A^2$$

$$\propto 4a^2 \cos^2 (\delta/2) \qquad (22.1)$$

To see the dependence of intensity on the phase difference between the two waves, let us consider the following two cases.

Case 1: When the phase difference, $\delta = 0, 2\pi, 4\pi, \ldots, 2n\pi$

$$I = 4a^2 \cos^2 0$$

$$= 4a^2$$

Case 2: When, $\delta = \pi, 3\pi, 5\pi, \ldots\ldots, (2n+1)\pi$

$$I = 4a^2 \cos^2 (\delta/2)$$

$$= 0$$

From these results we can conclude that when phase difference between superposing waves is an integral multiple of 2π, the two waves arrive at the screen 'in-phase' and the resultant intensity (or the brightness) at those points is more than that due to individual waves (which is equal to $4a^2$). On the other hand, when phase difference between the two superposing waves is an odd multiple of π, the two superposing waves arrive at the screen 'out of phase'. Such points have zero intensity and appear to be dark on the screen.

(d) Phase Difference and Path Difference

It is obvious from the above discussion that to know whether a point on the screen will be bright or dark, we need to know the phase difference between the waves arriving at that point. The phase difference can be expressed in terms of the path difference between the waves during their journey from the sources to a point on the observation screen. You may recall that waves starting from S_1 and S_2 are in phase. Thus, whatever phase difference arises between them at the point P is because of the different paths travelled by them upto observation point from S_1 and S_2. From Fig. 22.5, we can write the path difference as

$$\Delta = S_2P - S_1P$$

We know that path difference of one wavelength is equivalent to a phase difference of 2π. Thus, the relation between the phase difference δ and the path difference Δ is

$$\Delta = \left(\frac{\lambda}{2\pi}\right)\delta \qquad (22.2)$$

From Eqn. (22.1) we note that bright fringes (corresponding to constructive interference) are observed when the phase difference is $2n\pi$. Using this in Eqn. (22.2) we find that the path difference for observing bright fringes is

$$(\Delta)_{\text{bright}} = \left(\frac{\lambda}{2\pi}\right) 2n\pi = n\lambda;\ n = 0, 1, 2, \ldots \qquad (22.3)$$

Similarly, for dark fringes, we get

$$(\Delta)_{\text{dark}} = (\lambda/2\pi)\,(2n+1)\,\pi$$

$$= (2n + 1)\frac{\pi}{2};\ n = 0, 1, 2, \ldots \qquad (22.4)$$

Having obtained expressions for the bright and dark fringes in terms of the path difference and the wavelength of the light used, let us now relate path difference with the geometry of the experiment, i.e., relate Δ with the distance D between the source and the screen, separation between the pin holes (d) and the location of the point P on the screen. From Fig. 22.5 we note that

$$\Delta = S_2P - S_1P = S_2A = d \sin\theta$$

Assuming θ to be small, we can write

$$\sin\theta \approx \tan\theta \approx \theta$$

and $$\sin\theta = x / D$$

Therefore, the expression for path difference can be rewritten as

$$\Delta = d \sin\theta = x\frac{d}{D} \qquad (22.5)$$

On substituting Eqn. (22.5) in Eqns. (22.2) and (22.3), we get

$$\frac{d}{D}(x_n)_{\text{bright}} = n\lambda$$

or $$(x_n)_{\text{bright}} = \frac{n\lambda D}{d};\ n = 0, 1, 2, \ldots \qquad (22.6)$$

and $$\frac{d}{D}(x_n)_{\text{dark}} = (n + \frac{1}{2})\lambda$$

or $$(x_n)_{\text{dark}} = (n + \frac{1}{2})\frac{\lambda D}{d};\ n = 0, 1, 2, \ldots \qquad (22.7)$$

Eqns. (22.6) and (22.7) specify the positions of the bright and dark fringes on the screen.

(e) Fringe width

You may now ask: How wide is a bright or a dark fringe? To answer this question, we first determine the location of two consecutive bright (or dark) fringes. Let us first do it for bright fringes. For third and second bright fringes, from Eqn. (22.6), we can write

$$(x_3)_{\text{bright}} = 3\,\frac{\lambda D}{d}$$

and
$$(x_2)_{\text{bright}} = 2\,\frac{\lambda D}{d}$$

Therefore, fringe width, β is given by

$$\beta = (x_3)_{\text{bright}} - (x_2)_{\text{bright}} = \frac{\lambda D}{d} \qquad (22.8)$$

You should convince yourself that the fringe width of an interference pattern remains the same for any two consecutive value of *n*. Note that fringe width is directly proportional to linear power of wavelength and distance between the source plane and screen and inversely proportional to the distance between the slits. In actual practice, fringes are so fine that we use a magnifying glass to see them.

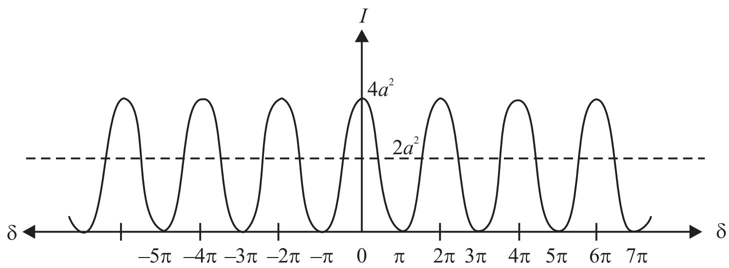


Fig. 22.6: Intensity distribution in an interference pattern

Next let us learn about the intensity of bright and dark fringes in the interference pattern. We know that when two light waves arrive at a point on the screen out of phase, we get dark fringes. You may ask : Does this phenomenon not violate the law of conservation of energy because energy carried by two light waves seem to be destroyed? It is not so; the energy conservation principle is not violated in the interference pattern. Actually, the energy which disappears at the dark fringes reappears at the bright fringes. You may note from Eqn. (22.1) that the intensity of the bright fringes is four times the intensity due to an individual wave. Therefore, in an interference fringe pattern, shown in Fig. (22.6), **the energy is redistributed and it varies between $4a^2$ and zero.** Each beam, acting independently, will contribute a^2 and hence, in the absence of interference, the screen will be uniformly illuminated with intensity $2a^2$ due to the light coming from two identical sources. This is the average intensity shown by the broken line in Fig. 22.6.

You have seen that the observed interference pattern in the Young's experiment can be understood qualitatively as well as quantitatively with the help of wave theory of light. To be sure that you have good understanding, answer the following intext questions.

INTEXT QUESTIONS 22.2

1. On what factors does the resultant displacement at any point in the region of superposition of two waves depend?
2. In Young's experiment, how is the constructive interference produced on the screen?
3. If we replace the pinholes S_1 and S_2 by two incandescent light bulbs, can we still observe the bright and dark fringes on the screen?
4. What are coherent sources? Can our eyes not act as coherent sources?

22.3 DIFFRACTION OF LIGHT

In earlier lessons, you were told that rectilinear propagation is one of the characteristics of light. The most obvious manifestation of the rectilinear propagation of light is in the formation of shadow. But, if you study formation of shadows carefully, you will find that, as such, these are not sharp at the edges. For example, the law of rectilinear propagation is violated when the light passes through a very narrow aperture or falls on an obstacle of very small dimensions. At the edges of the aperture or the obstacle, light bends into the shadow region and does not propagate along a straight line. ***This bending of light around the edges of an obstacle is known as diffraction.***

Before discussing the phenomenon of diffraction of light in detail, you may like to observe diffraction of light yourself. Here is a simple situation. Look at the street light at night and almost close your eyes. What do you see? The light will appear to streak out from the lamp/tube. This happens due to the diffraction (bending) of light round the corners of your eyelids.

Another way to observe diffraction is to use a handkerchief. Hold it close to your eyes and look at the Sun or a lamp. You will observe circular fringes, which form due to diffraction of light by small apertures formed by crissed-crossed threads.

In the above situations, the dimensions of the diffracting obstacle/aperture are very small. To observe diffraction, either of the following conditions must be satisfied:

a) *The size of the obstacle or the aperture should be of the order of the wavelength of the incident wave.*

b) *The separation between the obstacle or aperture and the screen should be considerably larger (a few thousand times) than the size of the obstacle or aperture.*

On the basis of the above observations, it is easy to understand why we normally do not observe diffraction of light and why light appears to travel in a straight line. You know that the wavelength of light is of the order of 10^{-6} m. Therefore, *to observe diffraction of light, we need to have obstacles or aperture having dimensions of this order!*

22.3.1 Diffraction at a Single Slit

Let us see how diffraction pattern appears for a simple opening like a single slit. Refer to Fig. 22.7. It shows the experimental arrangement for producing diffraction pattern. S is a monochromatic source of light. It is placed on the focal plane of a converging lens so that a plane wavefront is incident on a narrow slit. Another converging lens focusses light from different portions of the slit on the observation screen.

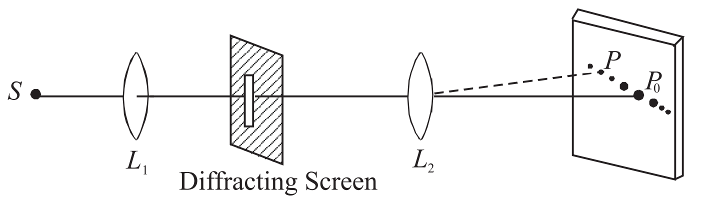


Fig. 22.7: Schematic representation of single slit dtiffraction

The salient features of the actual diffraction pattern produced by a single vertical slit from a point source as shown in Fig. 22.8 are :

- A horizontal streak of light along a line normal to the length of the slit.
- The horizontal pattern is a series of bright spots.

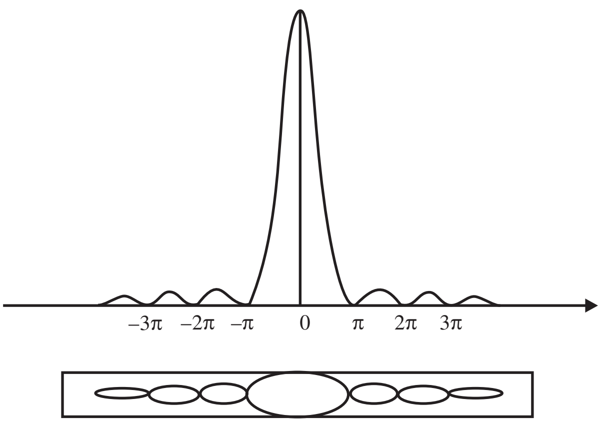


Fig. 22.8 : Observed differation pattern single of slit

- The spot at the centre is the brightest. On either side of this spot, we observe a few more symmetrically situated bright spots of diminishing intensity. The central spot is called *principal maxima* and other spots are called *secondary maxima.*
- The width of the central spot is twice the width of other spots.

To understand the theoretical basis of these results, we note that according to Huygens' wave theory, plane wavefronts are incident on the barrier containing the slit. As these wavefronts fall on the barrier, only that part of the wavefront passes through the slit which is incident on it. This part of the wavefront continues to propagate to the right of the barrier. However, the shape of the wavefront does not remain plane beyond the slit.

Refer to Fig. 22.9 which shows that each point of the aperture such as QPR ... Q′ form a series of coherent sources of secondary wavelets. In the central part of the wavefront to the right of the barrier, the wavelet emitted from the point *P*, say, spreads because of the presence of wavelets on its both sides emitted from the points such as *Q* and *R*. Since the shape of the wavefront is determined by the tangent to these wavelets, the central part of the wavefront remains plane as it propagates. But for the wavelets emitted from points *Q* and *Q′* near the edges of the slit, there are no wavelets beyond the edges with which these may superpose. Since the superposition helps to maintain the shape of the wavefront as plane, the absence of such superposing wavelets for the wavelets emitted from the points near the edges allows them to deviate from their plain shape. In other words, the wavelets at the edges tend to spread out. As a result, the plane wavefront incident on a thin aperture of finite size, after passing through it does not remain plane.

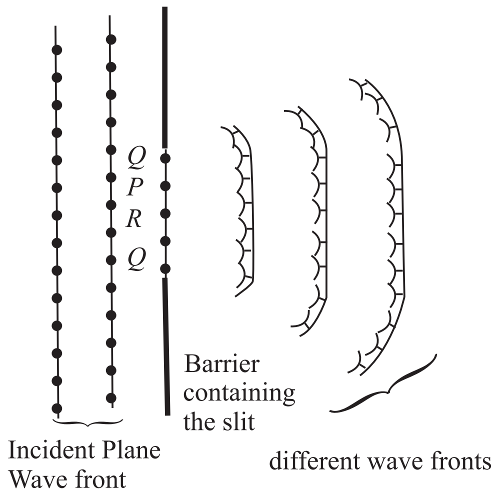


Fig. 22.9: Huygen's construction for diffraction of light from a narrow slit

To understa d the ***intensity distribution*** of the single-slit diffraction pattern, we determine the nature of the superposition of waves reaching the screen. In order to apply Huygens' principle, let us divide the width '*a*' of the slit into, say, 100 equal parts. Each of these can be considered as a sources of secondary wavelets. The wavelets emanating from these points spread out into the region to the right of the slit. Since the plane wavefront is incident on the slit, initially all points on it are in phase. Therefore, the wavelets emitted by these points are all in phase at the time of leaving the slit. Now let us consider the effect of the superposition of these wavelets at point O on the screen. The symmetry of the Fig. 22.10 suggests that the wavelets emitted from source of 1 and 100 will reach O in phase. It is so because both the wavelets travel equal path length. When they started their journey from the respective points on the slit, they were in phase. Hence they arrive at O in phase and superpose in such a manner as to give resultant amplitude much more than that due to the individual wavelets from the source 1 and 100. Similarly, for each wavelet from source 2 to 50, we have a corresponding wavelet from the source 99 to 51 which will produce constructive interference causing enhancement in intensity at the center O. Thus the point O will appear bright on the screen.

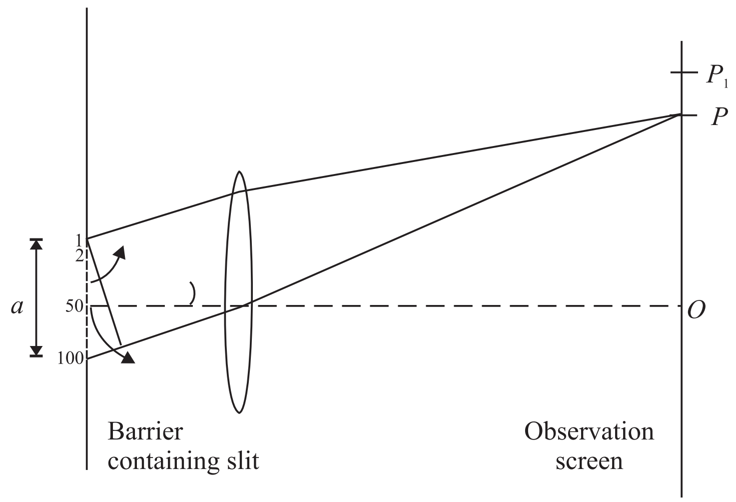


Fig. 22.10: Schematic representation of single slit diffraction

Now let us consider an off-axis point P on the observation screen. Suppose that point P is such that the *path difference* between the extreme points i.e. sources 1 and 100 is equal to λ. Thus the path difference between the wavelets from source 1 and 51 will be nearly equal to ($\lambda/2$).

You may recall from the interference of light that the waves coming from the sources 1 and 51 will arrive at P out of phase and give rise to destructive interference. Similarly, wavelets from the sources 2 and 52 and all such pair of wavelets will give rise to destructive interference at the point P. Therefore, we will have minimum intensity at point P. Similarly, we will get minimum intensity for other points for which the path difference between the source edges is equal to 2λ. We can imagine that the slit is divided into four equal points and we can, by similar pairing of 1 and 26, 2 and 27, ... show that first and second quarters have a path difference of $\lambda/2$ and cancel each other. Third and fourth quarters cancel each other by the same argument so that the resultant intensity will be minimum, and so on. *We can therefore conclude that when the path difference between the extreme waves diffracted by the extreme points in a particular direction is an integral multiple of* λ, *the resultant diffracted intensity in that direction will be zero.*

Let us now find intensity at a point P' which lies between the points P and P_1 and the *path difference between waves diffracted from extreme points is* $3\lambda/2$. We divide the wavefront at the slit into 3 equal parts. In such a situation, secondary wavelets from the corresponding sources of two parts will have a path difference of $\lambda/2$ when they reach the point P and cancel each other. However, wavelets from the third portion of the wavefront will all contribute constructively (presuming that practically the path difference for wavelets from this part is zero) and produce brightness at P'. Since only one third of the wavefront contributes towards the intensity at P' as compared to O, where the whole wavefront contributes, the intensity at P' is considerably less than that of the intensity at O. The point P' and all other similar points constitute secondary maxima.

However, you must note here that this is only a qualitative and simplified explanation of the diffraction at a single slit. You will study more rigorous analysis of this phenomena when you pursue higher studies in physics.

INTEXT QUESTIONS 22.3

1. Does the phenomenon of diffraction show that the light does not travel along a straight line path?
2. Distinguish between interference and diffraction of light.
3. Why are the intensity of the principal maximum and the secondary maxima of a single slit diffraction not the same?

22.4 POLARISATION OF LIGHT

In the previous two sections of this lesson, you learnt about the phenomena of interference and diffraction of light. While discussing these phenomena, we did not bother to know the nature of light waves; whether these were longitudinal or transverse. However, polarisation of light conclusively established that light is a transverse wave.

To understand the phenomenon of polarisation, you can perform a simple activity.

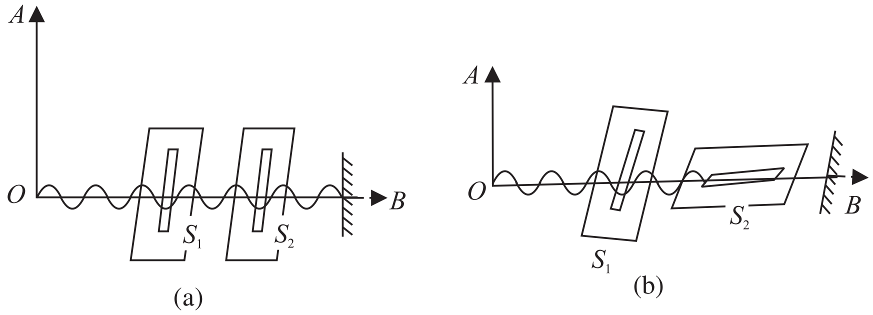


Fig. 22.11 : Transverse wave on a rope passing through a) two vertical slits, and b) one vertical and one horizontal slit

Take two card boards having narrow vertical slits S_1 and S_2 and hold them parallel to each other. Pass a length of a string through the two slits, fix its one end and hold the other in your hand. Now move your hand up and down and sideways to generate waves in all directions. You will see that the waves passing through the vertical slit S_1 will also pass through S_2, as shown in Fig. 22.11(a). Repeat the experiment by making the slit S_2 horizontal. You will see no waves beyond S_2. It means that waves passing through S_1 cannot pass through the horizontal slit S_2. This is because the vibrations in the wave are in a plane at right angles to the slits S_2, as shown in Fig. 22.11(b).

This activity can be repeated for light by placing a source of light at *O* and replacing the slits by two polaroids. You will see light in case(a) only. This shows that light has vibrations confined to a plane. It is said to be *linearly polarised* or *plane polarised* after passing through the first polaroid (Fig. 22.12).

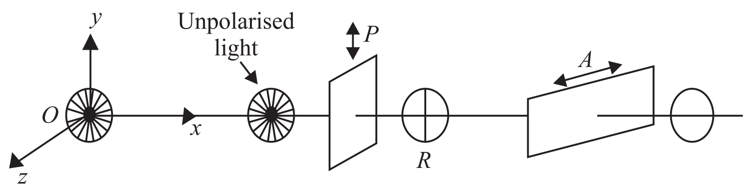
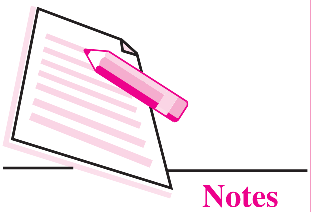


Fig. 22.12 : Schematics of the apparatus for observing polarisation of light

When an unpolarised light falls on glass, water or any other transparent material, the reflected light is, in general, partially plane polarised. Fig. 22.13 shows unpolarised light *AO* incident on a glass plate. The reflected light is shown by *OR* and the transmitted wave by *OT*. When the light is incident at polarising angle, the polarisation is complete. At this angle, the reflected and transmitted rays are at right angles to each other.

The polarising angle depends on the refractive index of the material of glass plate on which the (unpolarised) beam of light is incident. The relation between *r* and



i_p is obtained by using Snell's law (refer Fig. 22.13):

$$\mu = \frac{\sin i_p}{\sin r} = \frac{\sin i_p}{\sin(90 - i_p)}$$

$$= \frac{\sin i_p}{\cos i_p} == \tan i_p.$$

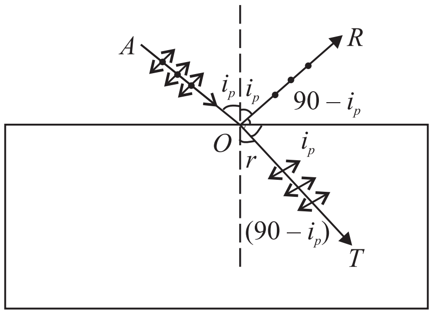


Fig. 22.13 : Polarisation of reflected and refracted light

This is known as ***Brewster's law.*** It implies that polarising angle i_p depends on the refractive index of the material. For air water interface, $i_p = 53°$. It means that when the sun is 37° above the horizontal, the light reflected from a calm pond or lake will be completely linearly polarised. Brewster's law has many applications in daily life. Glare caused by the light reflected from a smooth surface can be reduced by using polarising materials called *polaroids*, which are made from tiny crystals of quinine iodosulphate; all lined up in the same direction in a sheet of nitro cellulose. Such crystals (called dichoric) transmit light in one specific plane and absorb those in a perpendicular plan. Thus, *polaroid* coatings on sunglasses reduce glare by absorbing a component of the polarized light. Polaroid discs are used in photography as 'filters' in front of camera lens and facilitate details which would otherwise be hidden by glare. Polarimeters are used in sugar industry for quality control.

INTEXT QUESTIONS 22.4

1. Polarisation of light is the surest evidence that light is a transverse wave. Justify.
2. Is it correct to say that the direction of motion of a wave may not lie in the plane of polarisation?
3. Suppose a beam of unpolarised light is incident on a set of two *polaroids*. If you want to block light completely with the help of these polaroids, what should be the angle between the transmission axes of these polaroids?
4. Do sound waves in air exhibit polarization?

WHAT YOU HAVE LEARNT

- According to the Huygens' wave theory, light propagates in the form of wavefronts.
- The locus of all particles of the medium vibrating in the same phase at any instant of time is called the wavefront
- If two light sources emit light waves of the same frequency, same amplitude and move along the same path maintaining a constant phase difference between them, they are said to be coherent.
- When waves from two coherent sources superpose, a redistribution of energy takes place at different points. This is called the interference of light.
- For constructive interference, phase difference $\Delta = 2n\,\pi$ and for destructive interference, phase difference $\Delta = (2n + 1)\pi$.
- The bending of light near the corners of an obstacle or aperture is called diffraction of light.
- The phenomenon in which vibrations of light get confined in a particular plane containing the direction of propagation is called polarisation of light.

TERMINAL EXERCISE

1. Explain in brief the theories describing the nature of light.
2. What is a wavefront? What is the direction of a beam of light with respect to the associated wavefront? State the Huygens' principle and explain the propagation of light waves.
3. Obtain the laws of reflection on the basis of Huygens' wave theory.
4. What is the principle of superposition of waves? Explain the interference of light.
5. Describe Young's double slit experiment to produce interference. Deduce an expression for the width of the interference fringes.
6. What would happen to the interference pattern obtained in the Young's double slit experiment when
   (i) one of the slits is closed;
   (ii) the experiment is performed in water instead of air;
   (iii) the source of yellow light is used in place of the green light source;
   (iv) the separation between the two slits is gradually increased;

(v) white light is used in place of a monochromatic light;

(vi) the separation between the slits and the screen is increased;

(vii) two slits are slightly moved closer; and

(viii) each slit width is increased.

7. In Young's experimental set-up, the slit separation is 2 mm and the distance between the slits and the observation screen is 100 cm. Calculate the path difference between the waves arriving at a point 5 cm away from the point where the line dividing the slits touches the screen.

8. With the help of Huygens' construction, explain the phenomenon of diffraction.

9. How would you demonstrate that the light waves are transverse in nature?

10. Distinguish between the polarized and unpolarized lights.

11. State and explain Brewster's law.

12. The polarising angle for a medium is 60^o. Calculate the refractive index.

13. For a material of refractive index 1.42, calculate the polarising angle for a beam of unpolarised light incident on it.

ANSWERS TO INTEXT QUESTIONS

22.1

1. Perpendicular to each other ($\theta = \pi/2$) 2. ½

22.2

1. On the amplitude of the waves and the phase difference between them.

2. When the phase difference between the two superposing beams is an integral multiple of 2π, we obtain constructive interference.

3. No, it is so because two independent sources of light will emit light waves with different wavelengths, amplitudes and the two set of waves will not have constant phase relationship. Such sources of light are called incoherent sources. For observing interference of light, the sources of light must be coherent. When the light waves are coming from two incoherent sources, the points on the screen where two crests or two trough superpose at one instant to produce brightness may receive, at the other instant, the crest of the wave from one source and trough from the other and produce darkness. Thus, the whole screen will appear uniformly illuminated if the pinholes S_1 and S_2 are replaced by two incandescent light bulbs.

4. Coherent sources should emit waves

   (a) of same frequency and wavelength,

   (b) in phase or having constant phase difference, and

   (c) same amplitude and period.

   Moreover, these should be close. Our eyes may not meet this criterion.

22.3

1. Yes

2. Interference is the superposition of secondary waves emanating from two different secondary sources whereas diffraction is the superposition of secondary waves emanating from different portions of the same wavefronts.

3. Due to the increasing path difference between wavelets.

22.4

1. No. Because, in a longitudinal wave, the direction of vibrations is the same as the direction of motion of the wave.

2. No.   3. 90° or 270°   4. No.

Answers to Problems in Terminal Exercise

7. 0.1 mm   12.   1.73   13.   54°

312en23

23

OPTICAL INSTRUMENTS

We get most of the information about the surrounding world through our eyes. But as you know, an unaided eye has limitations; objects which are too far like stars, planets etc. appear so small that we are unable to see their details. Similarly, objects which are too small, e.g. pollen grains, bacteria, viruses etc. remain invisible to the unaided eyes. Moreover, our eyes do not keep a permanent record of what they see, except what is retained by our memory. You may therefore ask the question: How can we see very minute and very distant objects? The special devices meant for this purpose are called **optical instruments.**

In this lesson you will study about two important optical instruments, namely, a microscope and a telescope. As you must be knowing, a microscope magnifies small objects while a telescope is used to see distant objects. The design of these appliances depends on the requirement. (The knowledge of image formation by the mirrors and lenses, which you have acquired in Lesson 20, will help you understand the working of these optical instruments.) The utility of a microscope is determined by its magnifying power and resolving power. For a telescope, the keyword is *resolving power*. You must have read about Hubble's space telescope, which is being used by scientists to get details of far off galaxies and search for a life-sustaining planet beyond our solar system.

OBJECTIVES

After studying this lesson, you should be able to:

- *explain the working principle of simple and compound microscopes;*
- *derive an expression for the magnifying power of a microscope;*
- *distinguish between linear and angular magnifications;*
- *explain the working principle of refracting and reflecting telescopes; and*
- *calculate the resolving powers of an eye, a telescope and a microscope.*

23.1 MICROSCOPE

In Lesson 20 you have learnt about image formation by mirrors and lenses. If you take a convex lens and hold it above this page, you will see images of the alphabets/ words. If you move the lens and bring it closer and closer to the page, the alphabets printed on it will start looking enlarged. This is because their enlarged, virtual and errect image is being formed by the lens. That is, it is essentially acting as a magnifying glass or simple microscope. You may have seen a doctor, examining measels on the body of a child. Watch makers and jewellers use it to magnify small components of watches and fine jewellery work. You can take a convex lens and try to focus sunlight on a small piece of paper. You will see that after some time, the piece of paper start burning. A convex lens can, therefore start a fire. That is why it is dangerous to leave empty glass bottles in the woods. The sunlight falling on the glass bottles may get focused on dry leaves in the woods and set them on fire. Sometimes, these result in wild fires, which destroy large parts of a forest and/or habitation. Such fires are quite common in Australia, Indonesia and U.S.

As a simple microscope, a convex lens is satisfactory for magnifying small nearby objects upto about twenty times their original size. For large magnification, a compound microscope is used, which is a combination of basically two lenses. In a physics laboratory, a magnifying glass is used to read vernier scales attached to a travelling microscope and a spectrometer.

While studying simple and compound microscopes, we come across scientific terms like (i) near point, (ii) least distance of distinct vision, (iii) angular magnification or magnifying power, (iv) normal adjustment etc. Let us first define these.

(i) **Near point** is the distance from the eye for which the image of an object placed there is formed (by eye lens) on the retina. The near point varies from person to person and with the age of an individual. At a young age (say below 10 years), the near point may be as close as 7-8 cm. In the old age, the near point shifts to larger values, say 100-200 cm or even more. That is why young children tend to keep their books so close whereas the aged persons keep a book or newspaper far away from the eye.

(ii) **Least distance of distinct vision** is the distance upto which the human eye can see the object clearly without any strain on it. For a normal human eye, this distance is generally taken to be 25 cm.

(iii) **Angular magnification** is the ratio of the angle subtended by the image at the eye (when the microscope is used) to the angle subtended by the object at the unaided eye when the object is placed at the least distance of distinct vision. It is also called the magnifying power of the microscope.

(iv) **Normal Adjustment:** When the image is formed at infinity, least strain is exerted on the eye for getting it focused on the retina. This is known as normal adjustment.

(v) **Linear magnification** is the ratio of the size of the image to the size of the object.

(vi) **Visual angle** is the angle subtended by the object at human eye.

23.1.1 A Simple Microscope

When a convex lens of short focal length is used to see magnified image of a small object, it is called a simple microscope.

We know that when an object is placed between the optical center and the focus of a convex lens, its image is virtual, erect, and magnified and on the same side as the object. In practice, such a lens is held close to eye and the distance of the object is adjusted till a clear image is formed at the least distance of distinct vision. This is illustrated in Fig. 23.1, which shows an object AB placed between F and O. Its virtual image $A'B'$ is formed on the same side as the object. The position of the object is so adjusted that the image is formed at the least distance of distinct vision (D).

Magnifying power of a simple microscope

Magnifying power of an optical instrument is the ratio of the angle subtended by the image at the eye to the angle subtented by the object seen directly, when both lie at the least distance of distinct vision or the near point. It is also called angular magnification and is denoted by M. Referring to Fig. 23.1(a) and (b), the angular magnification of simple

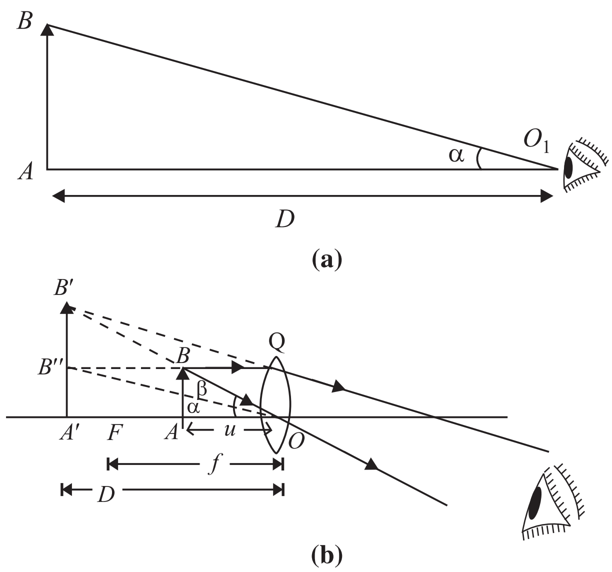


Fig.23.1 : Angular magnification of a magnifying glass

microscope is given by $M = \frac{\angle A'OB'}{AO'B} = \frac{\beta}{\alpha}$. In practice, the angles α and β are small. Therefore, you can replace these by their tangents, i.e. write

$$M = \frac{\tan\beta}{\tan\alpha} \qquad (23.1)$$

From $\Delta s\ A'OB'$ and AOB, we can write $\tan\beta = \frac{A'B'}{A'O} = \frac{A'B'}{D}$ and

$\tan\alpha = \frac{A'B''}{A'O} = \frac{AB}{D}$. On putting these values of $\tan\beta$ and $\tan\alpha$ in Eqn. (23.1), we get

$$M = \frac{A'B'}{D} \Big/ \frac{AB}{D} = \frac{A'B'}{AB}$$

Since $\Delta s\ AOB$ and $A'OB'$ in Fig 23.1(b) are similar, we can write

$$\frac{A'B'}{AB} = \frac{A'O}{AO} \qquad (23.2)$$

Following the standard sign convention, we note that

$$A'O = -D$$

and $$AO = -u$$

Hence, from Eqn. (23.2), we obtain

$$\frac{A'B'}{AB} = \frac{D}{u} \qquad (23.3)$$

If f is the focal length of the lens acting as a simple microscope, then using the lens formula $\left(\frac{1}{v} - \frac{1}{u} = \frac{1}{f}\right)$ and noting that $v = -D$, $u = -u$ and $f = f$, we get

$$\frac{1}{-D} - \frac{1}{-u} = \frac{1}{f}$$

or $$-\frac{1}{D} + \frac{1}{u} = \frac{1}{f}$$

Multiplying both the sides by D, and rearranging term, you can write

$$\frac{D}{u} = 1 + \frac{D}{f} \qquad (23.4)$$

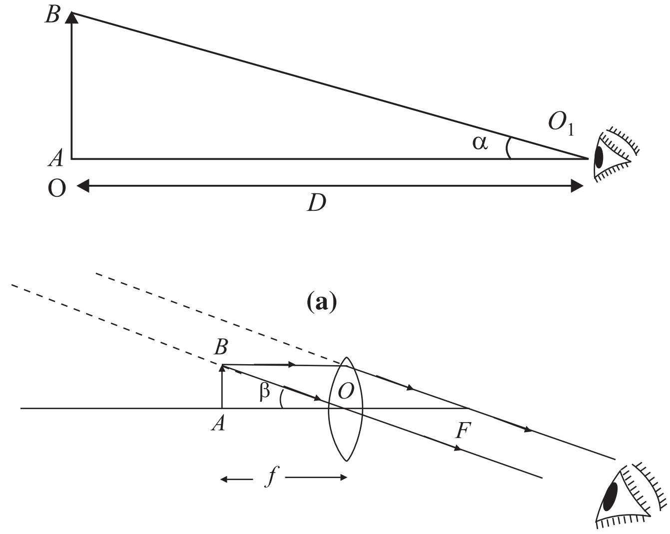
On combining Eqns. (23.3) and (23.4), we get

$$\frac{A'B'}{AB} = 1 + \frac{D}{f}$$

or
$$M = 1 + \frac{D}{f} \qquad (23.5)$$

From this result we note that lesser the focal length of the convex lens, greater is the value of the angular magnification or magnifying power of the simple microscope.

Normal Adjustment : In this case, the image is formed at infinity. The magnifying power of the microscope is defined as the ratio of the angle subtended by the image at the eye to the angle subtended by the object at the unaided eye when the object is placed at *D*. Fig 23.2(a) shows that the object is placed at the least distance of distinct vision *D*.



(b)

Fig.23.2 : Image formation for normal adjustment

The angles subtended by the object and the image at the unaided eye are α and β, respectively. The magnifying power is defined as

$$M = \frac{\beta}{\alpha}$$

In practice, the angles α and β are small, and, as before, replacing these by their tangents, we get

$$M = \frac{\tan\beta}{\tan\alpha}$$

i.e.
$$= \frac{AB}{AO} \Big/ \frac{AB}{AO_1}$$

$$= \frac{AO_1}{AO} = \frac{D}{f}$$

or
$$M = \frac{D}{f} \qquad (23.6)$$

You may note that in the normal adjustment, the viewing of the image is more comfortable. To help you fix your ideas, we now give a solved example. Read it carefully.

Example 23.1: Calculate the magnifying power of a simple microscope having a focal length of 2.5 cm.

Solution : For a simple microscope, the magnifying power is given by [Eqn. (23.5)] :

$$M = 1 + \frac{D}{f}$$

Putting D = 25 cm and f = 2.5cm, we get

$$M = 1 + \frac{25}{2.5} = 1 + 10 = 11$$

23.1.2 A Compound Microscope

A compound microscope consists of two convex lenses. A lens of short aperture and short focal length faces the object and is called the **objective**. Another lens of short focal length but large aperture facing the eye is called the **eye piece**. The objective and eye piece are placed coaxially at the two ends of a tube.

When the object is placed between F and $2F$ of the objective, its a real, inverted and magnified image is formed beyond $2F$ on the other side of the objective. This image acts as an object for the eye lens, which then acts as a simple microscope. The eye lens is so adjusted that the image lies between its focus and the optical center so as to form a magnified image at the least distance of distinct vision from the eye lens.

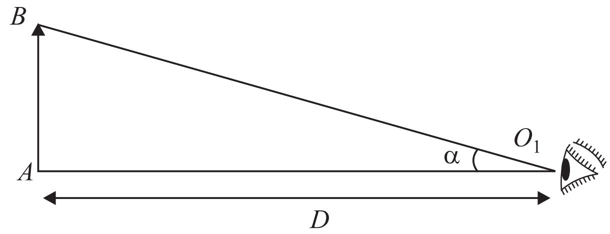
Magnifying Power of a compound microscope

Magnifying power of a compound microscope is defined as the ratio of the angle subtended by the final image at the eye to the angle subtended by the object at unaided eye, when both are placed at the least distance of distinct vision. It is denoted by M. By referring to Fig. 23.3, we can write

$$M = \frac{\beta}{\alpha}$$

Since the angles α and β are small, these can be replaced by their tangents, so that

$$M = \frac{\tan\beta}{\tan\alpha}$$



(a)

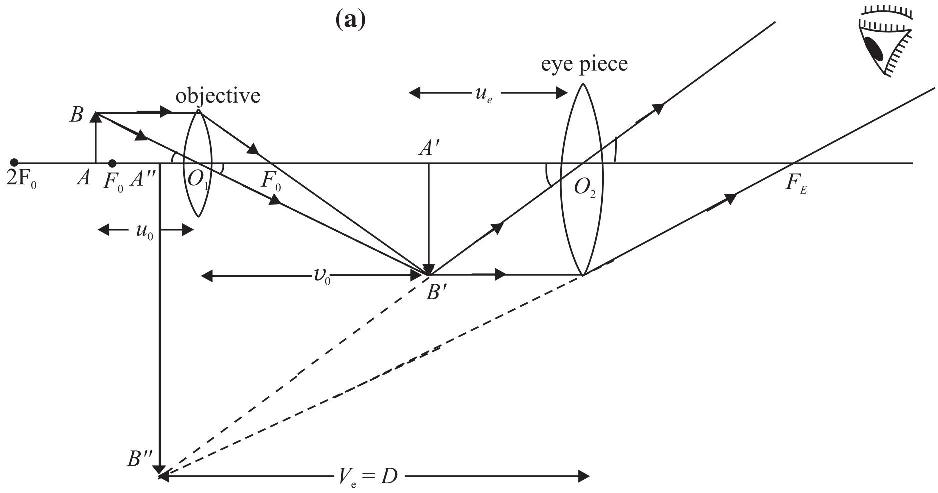


Fig.23.3 : Image formation by a compound microscope when the final image is formed at the least distance of distinct vision.

$$M = \frac{A''B''}{D} \Big/ \frac{AB}{D}$$

$$\Rightarrow \qquad M = \frac{A''B''}{AB} = \frac{A''B''}{A'B'} \cdot \frac{A'B'}{AB}$$

From similar Δs $A''B''O_2$ and $A'B'O_2$, we can write

$$\frac{A''B''}{A'B'} = \frac{A''O_2}{A'O_2} = \frac{D}{u_e}$$

Also from similar Δs$A'B'O_1$ and ABO, we have

$$\frac{A'B'}{AB} = \frac{v_o}{u_o}$$

Note that $m_e = \dfrac{A''B''}{A'B'}$ defines magnification produced by eye lens and $m_o = \dfrac{A'B'}{AB}$ denotes magnification produced by the objective lens. Hence

$$M = \frac{D}{u_e} . \frac{v_o}{u_o} = m_e \times m_o \qquad (23.7)$$

From Lesson 20, you may recall the lens formula. For eye lens, we can write

$$\frac{1}{\upsilon_e} - \frac{1}{u_e} = \frac{1}{f_e}$$

Multiply on both sides by υ_e to get

$$\frac{\upsilon_e}{\upsilon_e} - \frac{\upsilon_e}{u_e} = \frac{\upsilon_e}{f_e}$$

$$\Rightarrow \qquad \frac{\upsilon_e}{u_e} = 1 - \frac{\upsilon_e}{f_e}$$

Since f_e is positive and $\upsilon_e = -D$ as per sign convention, we can write

$$m_e = \frac{\upsilon_e}{u_e} = 1 + \frac{D}{f_e} \qquad (23.8)$$

On combining Eqns. (23.7) and (23.8), we get

$$M = \frac{\upsilon_o}{u_o} \times \left(1 + \frac{D}{f_e}\right)$$

In practice, the focal length of an objective of a microscope is very small and object AB is placed just outside the focus of objective. That is

$$\therefore \qquad u_o \approx f_0$$

Since the focal length of the eye lens is also small, the distance of the image $A'B'$ from the object lens is nearly equal to the length of the microscope tube i.e.

$$\upsilon_o \approx L$$

Hence, the relation for the magnifying power in terms of parameters related to the microscope may be written as

$$M = \frac{L}{f_0}\left(1 + \frac{D}{f_e}\right) \qquad (23.10)$$

Magnifying power in normal adjustment : In this case the image is formed at infinity. As discussed earlier, the magnifying power of the compound microscope may be written as

$$M = m_o \times m_e$$

$$= \frac{\upsilon_o}{u_o}\left(\frac{D}{f_e}\right)$$

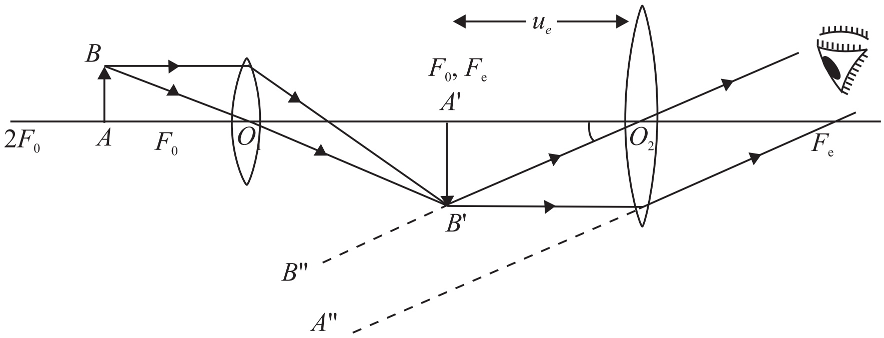


Fig. 23.4 : Compound microscope in normal adjustment

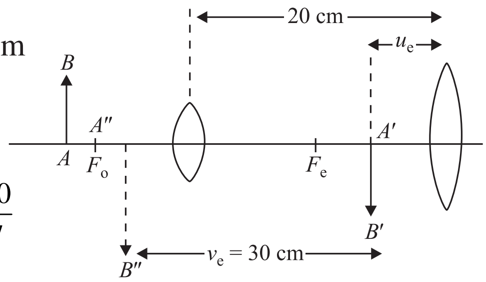
You may now like to go through a numerical example.

Example 23.2 : A microscope has an objective of focal length 2 cm, an eye piece of focal length 5 cm and the distance between the centers of two lens is 20 cm. If the image is formed 30 cm away from the eye piece, find the magnification of the microscope.

Solution : For the objective, f_o= 2 cm and f_e = 5 cm. For the eyepiece, υ_e = –30 cm and f_e = 5 cm. We can calculate υ_e using the relation

$$\frac{1}{\upsilon_e} - \frac{1}{u_e} = \frac{1}{f_e}$$

On solving, you will easily obtain $u_e = -\frac{30}{7}$ cm



For the objective lens

$$\upsilon_o = 20 - \frac{30}{7}$$

$$= \frac{110}{7} \text{ cm}$$

Using the formula

$$\frac{1}{\upsilon_o} - \frac{1}{u_o} = \frac{1}{f_o}$$

we have

$$\frac{1}{110/7} - \frac{1}{u_o} = \frac{1}{2}$$

or

$$u_o = -\frac{110}{48} \text{ cm}$$

The magnifying power of the objective

$$m_o = \frac{v_o}{u_o} = \frac{110/7}{-110/48} = -\frac{48}{7}$$

The magnification due to the eyepiece is

$$m_o = \frac{v_e}{u_e} = \frac{-30/1}{-30/7} = 7$$

Therefore, the magnification of the microscope is given by

$$M = (m_o)(m_e)$$

$$= \left(-\frac{48}{7}\right)(7) = -48$$

INTEXT QUESTIONS 23.1

1. What is the nature of images formed by a (i) simple microscope (ii) Compound microscope?
2. Differentiate between the magnifying power and magnification?
3. The magnifying power of a simple microscope is 11. What is its focal length?
4. Suppose you have two lenses of focal lengths 100 cm and 4 cm respectively. Which one would you choose as the eyepiece of your compound microscope and why?
5. Why should both the objective and the eyepiece of a compound microscope have short focal lengths?

23.2 TELESCOPES

Telescopes are used to see distant objects such as celestial and terrestrial bodies. Some of these objects may not be visible to the unaided eye. The visual angle subtended by the distant objects at the eye is so small that the object cannot be perceived. The use of a telescope increases the visual angle and brings the image nearer to the eye. Mainly two types of telescopes are in common use : refracting telescope and reflecting telescope. We now discuss these.

23.2.1 Refracting Telescope

The refracting telescopes are also of two types :

- **Astronomical telescopes** are used to observe heavenly or astronomical bodies.
- **Terrestrial telescopes** are used to see distant objects on the earth. So it is necessary to see an erect image. Even Galilean telescope is used to see objects distinctly on the surface of earth.

An astronomical telescope produces a virtual and erect image. As heavenly bodies are round, the inverted image does not affect the observation. This telescope consists of a two lens system. The lens facing the object has a large aperture and large focal length (f_o). It is called the *objective*. The other lens, which is towards the eye, is called the *eye lens*. It has a small aperture and short focal length (f_e). The objective and eye-piece are mounted coaxially in two metallic tubes.

The objective forms a real and inverted image of the distant object in its focal plane. The position of the lens is so adjusted that the final image is formed at infinity. (This adjustment is called normal adjustment.) The position of the eyepiece can also be adjusted so that the final image is formed at the least distance of distinct vision.

(a) When the final image is formed at infinity (Normal adjustment), the paraxial rays coming from a heavenly object are parallel to each other and they make an angle α with the principal axis. These rays after passing through the objective, form a real and inverted image in the focal plane of objective. In this case, the position of the eyepiece is so adjusted that the final image is formed at infinity.

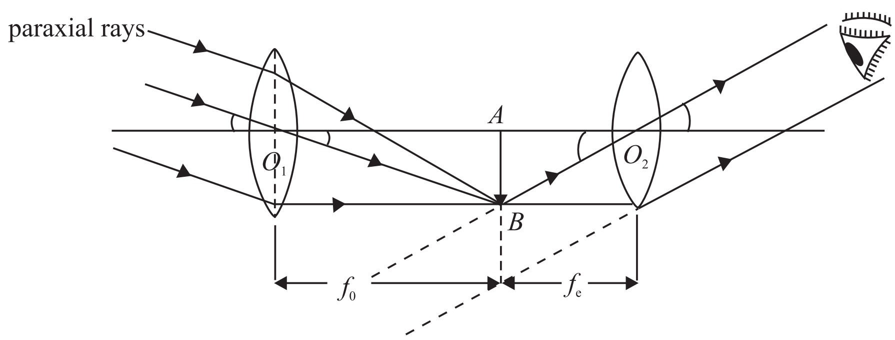


Fig 23.6 : Working principle of an astronomical telescope

Magnifying power of a telescope is defined as the ratio of the angle subtended by the image at the eye as seen through the telescope to the angle subtended by the object at objective when both the object and the image lie at infinity. It is also called **angular magnification** and is denoted by M. By definition,

$$M = \frac{\beta}{\alpha}$$

Since α and β are small, they can be replaced by their tangents. Therefore,

$$M = \frac{\tan \beta}{\tan \alpha}$$

$$= \frac{AB/AO_2}{AB/AO_1} = \frac{AO_1}{AO_2}$$

$$= \frac{f_o}{f_e} \qquad (23.11)$$

It follows that the magnifying power of a telescope in normal adjustment will be large if the objective is of large focal length and the eyepiece is of short focal length. The length of telescope in normal adjustment is $(f_o + f_e)$

(b) When the final image is formed at the least distance of distinct vision, the paraxial rays coming from a heavenly object make an angle α with the principal axis. After passing through the objective, they meet on the other side of it and form a real and inverted image *AB*. The position of the eyepiece is so adjusted that it finally forms the image at the least distance of distinct vision.

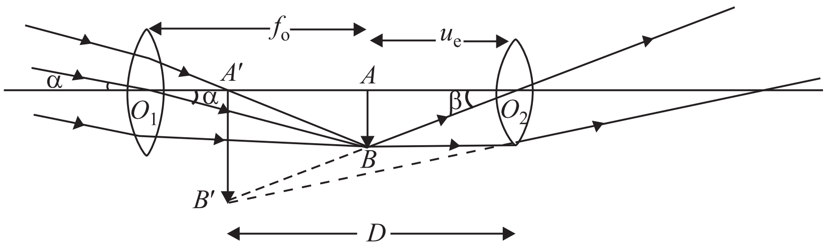


Fig 23.7 : Image formed by a telescope at *D*

Magnifying power: It is defined as the ratio of the angle subtended at the eye by the image formed at *D* to the angle subtended by the object lying at infinity:

$$M = \frac{\beta}{\alpha}$$

$$\approx \frac{\tan \beta}{\tan \alpha}$$

$$= \frac{AB/AO_2}{AB/AO_1} = \frac{AO_1}{AO_2}$$

$$= \frac{f_o}{u_e} \qquad (23.12)$$

Since $\frac{1}{v_e} - \frac{1}{u_e} = \frac{1}{f_e}$ for the eyepiece, we can write

$$\frac{1}{u_e} = \frac{1}{v_e} - \frac{1}{f_e}$$

$$= -\frac{1}{f_e}\left(1 - \frac{f_e}{v_e}\right)$$

or $$M = \frac{f_o}{u_e} = -\frac{f_o}{f_e}\left(1 - \frac{f_e}{v_e}\right) \qquad (23.13)$$

Applying the new cartesian sign convention $f_o = +f_o$, $v_e = -D$, $f_e = +f_e$, we can write

$$M = -\frac{f_o}{f_e}\left(1 + \frac{f_e}{D}\right) \qquad (23.14)$$

The negative sign of magnifying power of the telescope suggests that the final image is inverted and real. The above expression tells that the magnifying power of a telescope is larger when adjusted at the least distance of distinct vision to the telescope when focused for normal adjustment.

Example 23.3: The focal length of the objective of an astronomical telescope is 75 cm and that of the eyepiece is 5 cm. If the final image is formed at the least distance of distinct vision from the eye, calculate the magnifying power of the telescope.

Solution:

Here f_o = 75 cm, f_e = 5 cm, D = 25 cm

$$M = -\frac{f_o}{f_e}\left(1 + \frac{f_e}{D}\right) = \frac{-75}{5}\left(1 + \frac{5}{25}\right) = -18$$

23.2.2 Reflecting telescope

A reflecting telescope is used to see distant stars and possesses large light-gathering power in order to obtain a bright image of even a faint star deep in space. The objective is made of a concave mirror, having large aperture and large focal length. This concave mirror, being parabolic in shape, is free from spherical aberration.

Before the reflected rays of light meet to form a real, inverted and diminished image of a distant star at the focal plane of concave mirror, they are intercepted and reflected by a plane Mirror M_1M_2 inclined at an angle of 45 to the principal axis of the concave mirror. This plane mirror deviates the rays and the real image is formed in front of the eye piece, which is at right angle to the principal axis of concave mirror. The function of the eye- piece is to form a magnified, virtual image of the star enabling eye to see it distinctly.

If f_o is the focal length of the concave mirror and f_e is the focal length of eye piece, the magnifying power of the reflecting telescope is given by

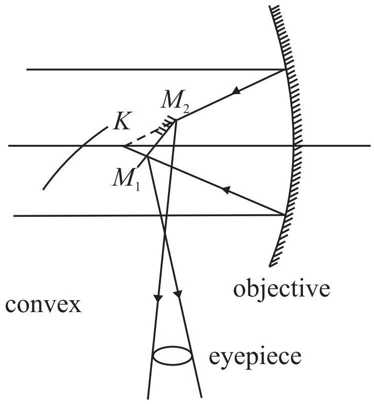


Fig 23.8 : Newtonion Reflector

$$M = \frac{f_o}{f_e}$$

Further, if D is the diameter of the objective and d is the diameter of the pupil of the eye, the brightness ratio is given by

$$B = D^2/d^2$$

The other form of the reflecting telescope is shown in Fig 23.9. It was designed by **Cassegrain**. In this case the objective has a small opening at its center. The rays from the distant star, after striking the concave mirror, are made to intercept at A_2 and the final image is viewed through the eyepiece.

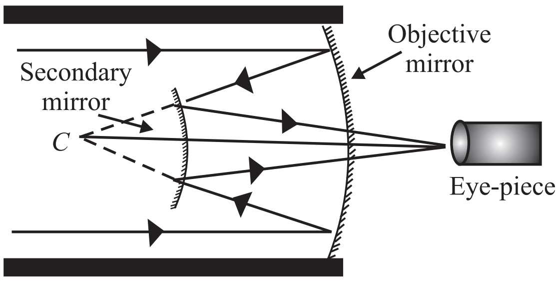


Fig 23.9 : Cassegrain reflector

There are several advantages of a reflecting telescope over a refracting telescope.

- Since the objective is not a lens, the reflecting telescopes are free from chromatic aberration. Thus rays of different colours reaching the objective from distant stars are focussed at the same point.
- Since the spherical mirrors are parabolic mirrors, free from spherical aberration, they produce a very sharp and distinct image.
- Even a very faint star can be seen through the reflecting telescope because they have large aperture and have large light-gathering power. The brightness of the image is directly proportional to the area of the objective :

$$B \ \alpha \ \frac{\pi D^2}{4}$$

where D is the diameter of the objective of the telescope. If d is the diameter of the pupil of the eye then brightness of the telescope B is defined as the ratio of light gathered by the telescope to that gathered by the unaided eye from the distant object

$$B = \frac{\pi D^2/4}{\pi d^2/4} = \frac{D^2}{d^2}$$

- In reflecting type of telescopes, there is negligible absorption of light.
- Large apertures of reflecting telescope enable us to see minute details of distant stars and explore deeper into space. That is why in recent years, astronomers have discovered new stars and stellar systems. You should look out for such details in science magazines and news dailies.

INTEXT QUESTIONS 23.2

1. How would the magnification of a telescope be affected by increasing the focal length of:

   (a) the objective ____

   ____

   (b) the eye piece ____

   ____

2. If the focal length of the objective of a telescope is 50 cm and that of the eyepiece is 2 cm. What is the magnification?
3. State one difference between the refracting and reflecting telescope.
4. What is normal adjustment?
5. If the telescope is inverted, will it serve as a microscope?

23.3 RESOLVING POWER : THE RAYLEIGH'S CRITERION

In earlier lessons, you have seen that the image of a point source is not a point, but has a definite size and is surrounded by a diffraction pattern. Similarly, if there are two point sources very close to each other, the two diffraction patterns formed by the two sources may overlap and hence it may be difficult to distinguish them as separate by the unaided eye. The resolving power of an optical instrument is its ability to resolve (or separate) the images of two point objects lying close to each other. Rayleigh suggested that two images can be seen as distinct when the first minimum of the diffraction pattern due to one object falls on the central maximum of the other. This is called *Rayleigh's criterion.*

If we assume that the pupil of our eye is about 2 mm in diameter, two points can be seen distinctly separate if they subtend an angle equal to about one minute of arc at the eye. **The reciprocal of this angle is known as the resolving power of the eye.**

Now let us calculate the resolving power of common optical instruments. We begin our discussion with a telescope.

23.3.1 Resolving Power of a Telescope

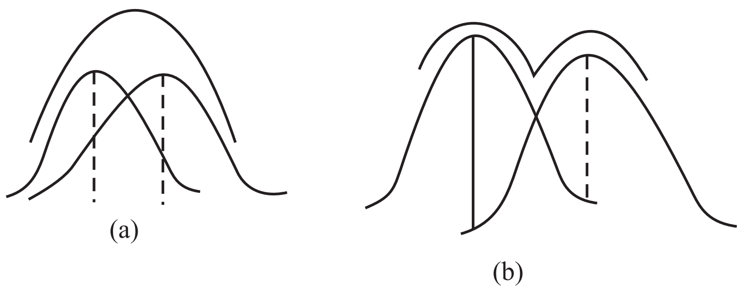


Fig. 23.10 : Rayleigh's criterion for resolution **a)** when the angular separation is less than θ, the two points are seen as one, and **b)** when the angular seperation is more than θ, the two points are distinctly visible.

The resolving power of a telescope is its ability to form separate images of two distant point objects situated close to each other. It is measured in terms of the angle subtended at its objective by two close but distinct objects whose images are just seen in the telescope as separate. This angle is called the **limit of resolution** of the telescope. If the angle subtended by two distinct objects is less than this angle, the images of the objects can not be resolved by the telescope. The smaller the value of this angle, higher will be the resolving power of the telescope. Thus, the reciprocal of the limit of resolution gives the resolving power of the telescope.

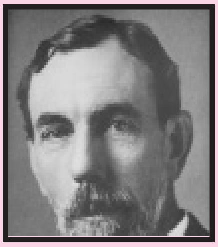
If λ is the wavelength of light, D the diameter of the telescope objective, and θ the angle subtended by the point object at the objective, the limit of resolution of the telescope is given by (Rayleigh's criterion)

$$\theta = \frac{1.22\lambda}{D}$$

Hence, the resolving power of the telescope.

$$(\text{R.P})_T = \frac{1}{\theta} = \frac{D}{1.22\lambda} \qquad (23.15)$$

From Eqn. (23.15) it is clear that to get a high resolving power, a telescope with large aperture objective or light of lower wavelength has to be used.



Lord Rayleigh
(1842 – 1919)

Born to the second Baron Rayleigh of Terling place, Witham in the country of Essex, England, John strutt had a very poor health in his childhood. Due to this he had a disrupted schooling. But he had the good luck of having Edward Rath and Stokes as his teachers. As a result, he passed his tripos examination in 1865 as senior Wrangler and become the first recipient of Smiths prize.

In addition to the discovery of Argon, for which he was awarded Nobel prize (1904), Rayleigh did extensive work in the fields of hydrodynamics, thermodynamics, optics and mathematics. His travelling wave theory, which suggested that elastic waves can be guided by a surface, paved way for researches in seismology and electronic signal processing. During the later years of his life, he also showed interest in psychiatry research. Lunar feature-crater Rayleigh and planetary feature crater Rayleigh on Mars are a tribute to his contributions.

Example 23.4: A telescope of aperture 3 cm is focussed on a window at 80 metre distance fitted with a wiremesh of spacing 2 mm. Will the telescope be able to observe the wire mesh? Mean wavelength of light $\lambda = 5.5 \times 10^{-7}$ m.

Solution: Given $\lambda = 5.5 \times 10^{-7}$ m and $D = 3$ cm $= 3 \times 10^{-2}$ m

Therefore, the limit of resolution

$$\theta = \frac{1.22\lambda}{D} = \frac{1.22 \times 5.5 \times 10^{-7}\,\text{m}}{3 \times 10^{-2}\,\text{m}} = 2.236 \times 10^{-5} \text{ rad}$$

The telescope will be able to resolve the wiremesh, if the angle subtended by it on the objective is equal to or greater than θ, the limit of resolution. The angle subtended by the wiremesh on the objective

$$\alpha = \frac{\text{spacing of wiremesh}}{\text{distance of the objective from the wiremesh}}$$

$$= \frac{2\text{ mm}}{80\text{ m}} = \frac{2 \times 10^{-3}}{80\text{ m}} = 2.5 \times 10^{-5} \text{ rad.}$$

As the angle 2.5×10^{-5} radian exceeds the limit of a resolution ($= 2.236 \times 10^{-5}$ radian), the telescope will be able to observe the wire mesh.

23.3.2 Resolving Power of a Microscope

The resolving power of a microscope represents its ability to form separate images of two objects situated very close to each other. The resolving power of a microscope is measured in terms of the smallest linear separation between the two objects which can just be seen through the microscope as separate. This smallest linear separation between two objects is called **the limit of resolution of the microscope.**

The smaller the value of linear separation, the higher will be the resolving power of the microscope. Thus, **the reciprocal of the limit of resolution gives the resolving power of the microscope.**

If λ is the wavelength of light used to illuminate the object, θ is the half angle of the cone of light from the point object at the eye and n is the refractive index of the medium between the object and the objective, the limit of resolution of the microscope is given by

$$d = \frac{\lambda}{2n\sin\theta} \qquad (23.16)$$

Thus the resolving power of microscope will be

$$(\text{R.P})_m = \frac{2n\sin\theta}{\lambda}$$

(23.17)The expression $2n\ \sin\theta$ is called numerical aperture (N.A) The highest value of N.A of the objective obtainable in practice is 1.6, and for the eye, N.A is 0.004.

It is clear from Eqn. (23.17) that the resolving power of a microscope can be increased by increasing the numerical aperture and decreasing the wavelength of the light used to illuminate the object. That is why ultraviolet microscopes and an electron microscope have a very high resolving power.

Applications in Astronomy

The astronomical (or optical) telescope can be used for observing stars, planets and other astronomical objects. For better resolving power, the optical telescopes are made of objectives having a large aperture (objective diameter). However, such big lenses are difficult to be made and support. Therefore, most astronomical telescopes use reflecting mirrors instead of lenses. These can be easily supported as a mirror weighs less as compared to a lens of equivalent optical quality.

The astronomical telescopes, which are ground-based, suffer from blurring of images. Also, ultraviolet, x-ray, gamma-ray etc. are absorbed by the earth's surface. They cannot be studied by ground-based telescopes. In order to study these rays coming from astronomical objects, telescopes are mounted in satellites above the Earth's atmosphere. NASA's Hubble space telescope is an example of such telescope. Chandra X-ray observation, Compton x-ray observation and Infrared telescopes have recently been set up in space.

INTEXT QUESTIONS 23.3

1. How can the resolving power of a telescope be improved?
2. What is the relationship between the limit of resolution and the resolving power of the eye?
3. If the wavelength of the light used to illuminate the object is increased, what will be the effect on the limit of resolution of the microscope?
4. If in a telescope objective is made of larger diameter and light of shorter wavelength is used, how would the resolving power change?

WHAT YOU HAVE LEARNT

- The angle subtended by an object at the human eye is called as the visual angle.
- The angular magnification or magnifying power of a microscope is the ratio of the angle subtended by the image at the eye to the angle subtended by the object when both are placed at the near point.
- Linear magnification is defined as the ratio of the size of the image to the size of the object.
- The magnifying power of a simple microscope is $M = 1 + \frac{D}{f}$, where D is least distance of distinct vision and f is focal length of the lens.
- In a compound microscope, unlike the simple microscope, magnification takes place at two stages. There is an eye piece and an objective both having short focal lengths. But the focal length of the objective is comparatively shorter than that of the eye piece.
- The magnifying power of a compound microscope is given as

$$M = m_0 \times m_e$$

But $m_e = 1 + \frac{D}{f}$. Therefore

$$M = \frac{\upsilon_o}{u_o}\left(1+\frac{D}{f_e}\right)$$

where υ_0 is distance between the image and the objective, u_0 is object distance from the objective, D is the least distance of distinct vision (= 25cm) and f_e is focal length of the eye-piece.

- Telescope is used to see the distant objects which subtend very small visual angle at the eye. The use of a telescope increases the visual angle at the eye. The far-off object appears to be closer to the eye and can be seen easily.
- Two types of telescopes are used (i) Refracting (ii) Reflecting.
- The objective of the refracting telescope is a converging lens. But the objective in a reflecting telescope is a spherical mirror of large focal length. There are several advantages of reflecting telescope over a refracting telescope.

  The magnifying power of a telescope is

$$M = f_0/f_e$$

  where f_0 is focal length of the objective and f_e is focal length of the eyepiece.

TERMINAL EXERCISE

1. What is the difference between simple and compound microscopes? Derive an expression for the magnification of a compound microscope.
2. Distinguish between the refracting and reflecting telescope. Draw a ray diagram for the Newton's telescope.
3. Derive an equation for the magnifying power of a refracting telescope.
4. What do you mean by the least distance of distinct vision? What is its value for a normal eye?
5. Can we photograph the image formed by a compound microscope? Explain your answer.
6. Define the resolving power of an optical instrument. What is the value of limit of resolution for a normal eye?
7. What are the main differences in the design of a compound microscope and a terrestrial telescope?
8. The eyepiece of a telescope has a focal length of 10 cm. The distance between the objective and eye piece is 2.1 m. What is the angular magnification of the telescope?

9. The image formed by a microscope objective of focal length 4 mm is 18 cm from its second focal point. The eyepiece has a focal length of 3.125 cm. What is the magnification of the microscope?

10. The objective of a telescope has a diameter three times that of a second telescope. How much more amount of light is gathered by the first telescope as compared to the second?

ANSWERS TO INTEXT QUESTIONS

23.1

1. Image formed by a simple microscope is virtual erect and magnified. whereas the image formed by a compound microscope is real, inverted and magnified.

2. Magnifying power is the ratio of the angle subtended by the image at eye piece to the angle subtended by the object placed at the near point. Magnification is the ratio of the size of image to the size of object.

3. $M = 11$, $m = 1+\frac{D}{f}$. Putting D =25 cm, we get $f = 2.5$ cm

4. If you choose the lens with 4 cm focal length, the magnifying power will be high because $m = \frac{f_o}{f_e}$

5. The magnifying power of a compound microscope is given by $M = \frac{-L}{f_o}\left(1+\frac{D}{f_e}\right)$

   Obviously, M will have a large value, if both f_o and f_e are small.

23.2

1. (a) Objective of large focal length increases the magnifying power of the telescope.

   (b) Magnification is reduced by increasing the focal length of eyepiece.

2. Magnification $m = \frac{f_o}{f_e} = \frac{50 \text{ cm}}{2 \text{ cm}} = 25$

3. The objectives of a telescope is a spherical mirror of large focal length instead of converging lens as in a refracting telescope.

4. A telescope is said to be in normal adjustment, if the final image is formed at infinity.

5. No

23.3

1. By taking a large aperture or by using a light of lower wavelength.
2. The limit of resolution of an eye is inversely proportional to its resolving power. Limit of resolution will also be increased.
3. Since resolving power of telescope is given by R.P = $\frac{D}{1.22\lambda}$, it would increase.

Answers To Problems in Terminal Exercise

8. 21 9. 400

10. 9 times.

312en24

24

STRUCTURE OF ATOM

So far you have studied about mechanical, thermal, electrical and magnetic properties of matter. Have you ever thought as to why do different materials have different properties? That is, why does chalk break so easily but a piece of aluminium flattens on impact? Why do some metals start conducting current when light falls on them? And so on. To understand such properties of materials, we recall that atoms are building blocks of all forms of matter. That is, despite its appearance being continuous, matter has definite structure on microscopic level which is beyond the reach of our sense of seeing. This suggests that to discover answers to above said questions, you need to know the structure of the atom.

Our understanding of the structure of atom has evolved over a period of time. In this lesson, we have discussed different atomic models. Starting with Rutherford's model based on his classic scattering experiment, in this lesson we have discussed Bohr's model of atom that explains the electronic structure. Bohr's theory also helps us to explain the atomic spectrum of hydrogen atom.

OBJECTIVES

After studying this lesson, you should be able to:

- *describe Rutherford's scattering experiment and its findings;*
- *explain Rutherford's atomic model and state its shortcomings.*
- *calculate the radius of Bohr's first orbit and velocity of an electron in it;*
- *derive an expression for the energy of an electron in a hydrogen atom; and*
- *draw the energy level diagram of a hydrogen atom and explain its spectrum.*
- *describe about the production,properties, types and uses of x-rays; and*
- *define Mosley law and Duane-Hunt law.*

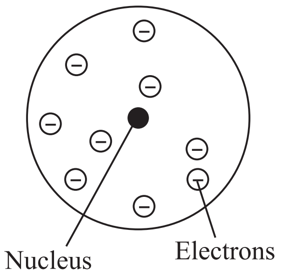


Fig. 24.1 : Plum-pudding model of atom

The Concept of Atom

The concept of atom is as old as human civilization. In ancient times Democritus in Greece and Kanad in India tried to explain the changes around us in terms of particles. But the exact theory of atom was presented by John Dalton, an English Chemist in 1808. He described atom as the smallest, indivisible particle endowed with all the properties of element and takes part in chemical reactions. Dalton's atom was an ultimate particle having no structure. This idea was accepted by the scientists in the nineteenth century as they knew nothing about the structure of atoms. The discovery of electrons by J.J. Thomson in 1897, while studying discharge of electricity through gases at low pressures, provided the first insight that atoms have a structure and negatively charged electrons are constituents of all atoms. Since the atom as a whole is neutral, it must also have equal amount of positive charge. Moreover, since electrons were thousands of times lighter than the atom, it was thought that the positively charged constituent of atoms carried the entire mass. On the basis of his experiments, Thomson suggested the *plum pudding model* of atom (Fig. 24.1). According to this, an atom is a tiny, uniformaly charged positive ball in which negatively charged electrons are suitably placed to make it neutral. It seemed perfectly reasonable then.

Our understanding of the structure of atom since the times of Thomson has improved considerably. Due to the pioneering works of Lord Rutherford, Niels Bohr, James Chadwick,Pauli, Schrodinger and others. In fact, our concept of new world of sub-atomic particles came into existence and has led to the invention of epoch making new technologies, like micro-electronics and nanotechnology.

24.1 RUTHERFORD'S EXPERIMENT ON SCATTERING OF A-PARTICLES

On the advice of Lord Rutherford, two of his students Geiger and Marsden performed an experiment in which a beam of α-particles was bombarded on a thin gold foil. The experimental arrangement used by them is shown in Fig. 24.2.

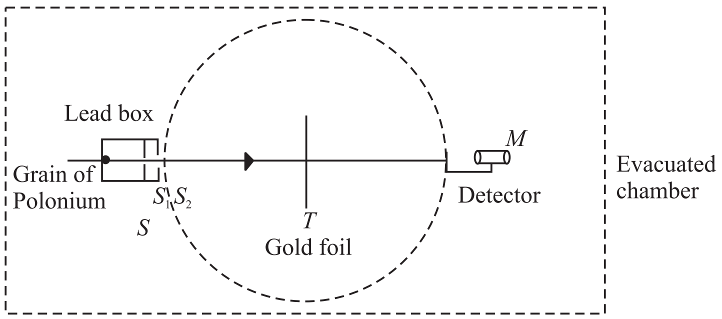


Fig. 24.2 : Schematics of experimental set up used for α-particle scattering

A well collimated fine pencil of α-particles from a source *S* was made to fall on a thin gold foil (*T*). The scattered α-particles were received on a ZnS fluorescent screen, which produced a visible flash of light when struck by an α-particles (and acted as detector), backed by a low power microscope (*M*). The detector was capable of rotation on a circular scale with *T* at the centre. The whole apparatus was enclosed in an evacuated chamber to avoid collisions of α-particles with air molecules. It was expected that if Thomson model was correct, most of the particles world go straight through the foil, with only minor deviation from the original path.

Geiger and Marsden observed that most of the α-particles suffered only small deflections, as expected. But a few got deflected at large angles (90º or more). Some of them (1 in 8000) even got deflected at 180º. Fig. 24.3 presents the experimental results. The large angle scattering of α-particles could not be explained on the basis of Thomson model of atom.

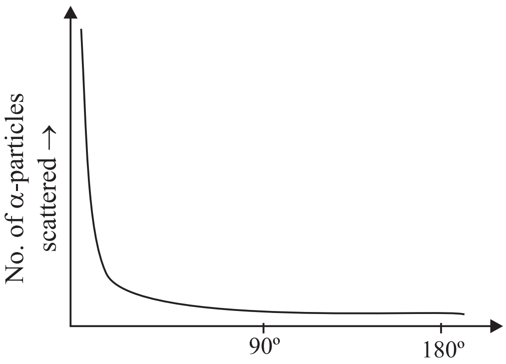


Fig. 24. 3 : Experimental result of Rutherford's experiment

To explain large angle scattering, Lord Rutherford suggested the nuclear model of atom. He argued that α-particles which pass at a large distance from the nucleus experience negligible coulombian repulsive force and hence, pass almost undeflected. However, closer to the nucleus a α-particle comes, greater force of repulsion it experiences and hence gets deflected at a greater angle. A few α-particles which proceed for a head-on collision towards the nucleus are scattered back by 180º along its direction of approach, as indicated by α-particle 4 in Fig. 24.4.

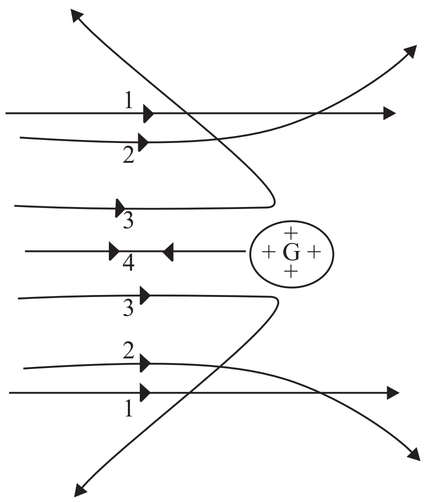
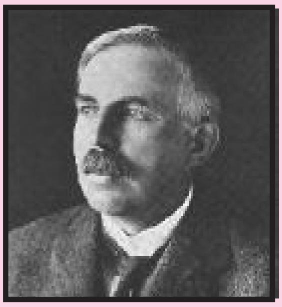


Fig. 24.4 : Paths traversed by α-particles scattered by a gold foil

According to Thomson model, α-particles should experience weak force due to electrons. However, since α-particles are about 7000 times heavier than electrons and travelled at high speed, large angle scattering strong repulsive force was required to be exerted.

Lord Rutherford

(1871–1937)



Born in New Zealand, Rutherford studied under J.J. Thomson at the Cavendish Laboratory in England. His poineering work on atom is a defining landmark. He developed Becquerel's discovery of radioactivity into an exact science and documented proof that the atoms of heavier elements, which had been thought to be immutable, actually disintegrate (decay) into various forms of radiation. In 1898, Rutherford discovered that two quite separate types of emissions came from radioactive atoms and he named them alpha and beta rays. Beta rays were soon shown to be high speed electrons. In 1907, he showed that the alpha particle was a helium atom stripped of its electrons. He and his assistant, Hans Geiger, developed Rutherford-Geiger detector to electrically detecting particles emitted by radioactive atoms. With this he could determine important physical constants such as Avogadro's number, the number of atoms or molecules in one gram-mole of material.

In 1911, Rutherford proposed the nuclear model of atom; that almost entire mass of an atom is concentrated in a nucleus 10^{-5} times the atom itself and electrons revolve around it. This second great work won him the **Nobel Prize** in chemistry in 1908.

Eminent Indian physicist, educationist and philosopher Dr. D.S. Kothari was one of his students and worked on pressure ionisation in stars.

24.1.1 Nuclear Model of Atom

Rutherford argued that large angle scattering of α-particles can be explained only by stipulating the presence of a hard, positively charged core of atom. Thus he proposed a new model of atom with following characteristics:

- The entire charge and most of the mass of the atom is confined in a very small ($\sim 10^{-15}$ m) central region, called the *nucleus.*
- The negatively charged electrons revolve at a distance around it such that the atom as a whole is electrically neutral and stable.

The nuclear model of atom proposed by Rutherford faced some difficulties. Some of the consequences of Rutherford's model contradicted experimental observations.

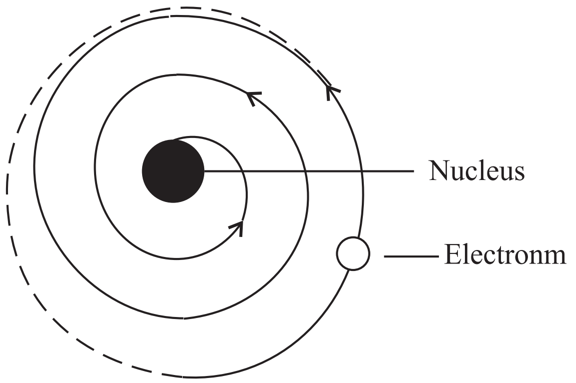


Fig. 24.5 : Motion of electrons in Rutherford's nuclear model of atom

(i) **Stability of the atom :** We know that electrons are negatively charged. These are attracted by the nucleus and get accelerated. An accelerated charged particle, according to clasical wave theory, emits elecromagnetic radiations. Hence, the revolving electrons should lose energy eventually and spiral into the nucleus (Fig. 24.5). This would have made the atom short-lived and contradicted the observed stability of matter.

(ii) **Frequency of electromagnetic radiation :** The electron spiralling towards the nucleus will emit electromagnetic radiations of all frequencies giving rise to a continuous spectrum. But experiments show that atoms emit radiations of certain well defined frequencies only (line spectra).

From the above discussion, you may be tempted to conclude that nuclear model of atom could not explain the experimental facts. Nevertheless, it contributed significantly to our understanding and was the first landmark in the right direction.

INTEXT QUESTIONS 24.1

1. Choose the correct answers :

   (a) In Rutherford's scattering experiment, target was bombarded with (i) β-rays (ii) γ -rays (iii) α-rays.

   (b) The nucleus is surrounded by :
   (i) electrons (ii) protons (iii) α-particles

   (c) The large angle scattering of α-particles indicated the presence of (i) some positively charged hard core inside the atom (ii) some porous core inside the atom (iii) negatively charged core.

2. Name two experimental observations that could not be explained by Rutherford's model.

> Bohr quantised energy as well as angular momentum to explain hydrogen spectrum; Planck quantised only energy to explain black body radiation.

24.2 BOHR'S MODEL OF HYDROGEN ATOM

To overcome the difficulties of Rutherford model, in 1913, Neils Bohr proposed a model of atomic structure, based on quantum ideas proposed by Max Planck. Bohr's model not only described the structure of atom but also accounted for its stability. It proved highly successful in explaining the observed spectrum of the hydrogen atom. Let us learn about it now.

Bohr's Postulates

Bohr started with the planetary model of atom. However, to overcome the problems that plagued Rutherford model, Bohr made several assumptions. These are known as Bohr's postulates. There are four postulates.

(i) *Electrons in an atom move in circular orbits around the nucleus with the centripetal force supplied by the Coulomb force of attraction between the electron and the nucleus. Mathematically, we can write*

$$\frac{mv^2}{r} = \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r^2} \tag{24.1}$$

where Z denotes the number of positive charges in the nucleus.

(ii) *Of the infinite number of possible circular orbits, only those orbits are allowed for which the value of orbital angular momentum of the electron is an integral multiple of* $h/2\pi$:

$$|\mathbf{L}| = mvr = \frac{nh}{2\pi} \tag{24.2}$$

where **L** is the orbital angular momentum, equal to *mvr* for a circular orbit. Here *h* is Planck's constant and *n* is an integer.

(iii) *An electron moving in an allowed orbit does not radiate any energy*. In these allowed orbits, the energy of the electron is constant. These orbits are called *stationary states.*

Note that an electron can move in a stationary state but its energy is constant.

(iv) *Energy is emitted by an atom only when its electron "falls" from an allowed higher energy level* E_f *to another allowed lower level* E_i. *The change in energy is the energy of the emitted photon*. Similarly, an electron only absorbs radiation when it "jumps" to a higher energy level from a lower energy level. The change in energy of an electron can be related to the frequency or wavelength of the emitted or absorbed photon:

For emission

$$\Delta E = E_i - E_f = h\nu \qquad (24.3a)$$

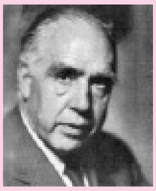
For absorption

$$\Delta E = E_f - E_i = h\nu \qquad (24.3b)$$

where ν is the frequency of the emitted photon.

Niels Henrik David Bohr

(1885-1962)



Niels Bohr was born in Copenhagen, Denmark. He grew up in an atmosphere most favourable to the development of his genius. His father was an eminent physiologist and was largely responsible for awakening his interest in physics while he was still at school. In the spring of 1912, he worked in Rutherford's laboratory in Manchester. He studied the structure of atoms on the basis of Rutherford's nuclear model of atom. He succeeded in working out and presenting a picture of atomic structure that explained atomic spectra of hydrogen atom.

In 1916, he was appointed Professor of Theoretical Physics at Copenhagen University, and in 1920 (until his death in 1962), he became head of the Institute for Theoretical Physics, established for him at that university.

Recognition of his work on the structure of atom came with the award of the **Nobel Prize** in Physics in 1922.

Note that these postulates beautifully combine classical and quantum ideas. For example, the first postulate is in accordance with classical physics while other postulates use quantum physics.

24.2.1 Energy Levels

To calculate the energy of an electron in *n*th orbit of radius r_n, we rewrite Eqn. (24.1) as

$$\frac{m\upsilon_n^2}{r_n} = \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_n^2}$$

where υ_n is speed of the electron in its orbit.

On multiplying both sides of this equation by mr_n^3 we get

$$m^2\upsilon_n^2 r_n^2 = \frac{1}{4\pi\varepsilon_0} m Z e^2 r_n$$

On combining this result with Eqn. (24.2), we get

$$m^2\upsilon_n^2 r_n^2 = n^2\frac{h^2}{4\pi^2} = \frac{m}{4\pi\varepsilon_0} Ze^2 r_n \qquad (24.4)$$

On re-arranging terms, we get an expression for the radius of the *n*th orbit :

$$r_n = 4\pi\varepsilon_0 \frac{n^2h^2}{4\pi^2 mZe^2}$$

$$= \frac{n^2h^2\varepsilon_0}{Ze^2 m\pi} \qquad n = 1, 2, 3, ... \qquad (24.5)$$

Note that radius of an orbit is directly proportional to second power of the number of orbit. It means that radius is more for higher orbits. Moreover, the relative values of the radii of permitted orbits are in the ratio $1^2, 2^2, 3^2, 4^2$, .. , ie. in the ratio 1 : 4 : 9 : 16 and so on. For hydrogen atom (Z = 1), the radius of its inner most orbit is called **Bohr radius**. It is denoted by a_0 and its magnitude is 5.3×10^{-11}m. In terms of a_0, the radii of other orbits are given by the relation

$$r_n = n^2 a_0$$

In shows that the spacing between consecutive orbits increases progressively. On inserting the value of r_n from Eqn. (24.5) in Eqn. (24.2), we get an expression for the speed of the electron in the *n*th orbit :

$$\upsilon_n = \frac{nh}{2\pi m r_n} = \frac{nh}{2\pi m} \cdot \frac{Ze^2 m\pi}{n^2h^2\varepsilon_0}$$

$$= \frac{1}{2}\frac{Ze^2}{\varepsilon_0 nh} \qquad (24.6)$$

From Lesson 16 you will recall that potential energy of a negative charge (electron in this case) in bringing it from infinity to a point at a distance *r* in a field of positive charge (nucleus in this case) is obtained by summing (integrating) the product of Coulomb force and distance :

$$U = -\frac{1}{4\pi\varepsilon_0}\int_{r_n}^{\infty} \frac{Ze^2}{r^2} dr$$

$$= \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r}\Bigg]_{r_n}^{\infty}$$

$$= -\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_n} \qquad (24.7)$$

since potential energy of the electron at infinity will be zero.

It readily follows from Eqn. (24.1) that

$$\frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r_n} = mv_n^2$$

Hence, potential energy of an electron in n^{th} orbit

$$U = -mv_n^2 \qquad (24.8)$$

Since kinetic energy

$$K.E = \frac{1}{2}mv_n^2 \qquad (24.9)$$

the total energy of the electron in nth orbit is given by

$$E = K.E + U$$

$$= \frac{1}{2}mv_n^2 - mv_n^2$$

$$= -\frac{1}{2}mv_n^2$$

Combining this result with Eqn. (24.6), we get

$$E = -\frac{m}{2}\left(\frac{2\pi Ze^2}{4\pi\varepsilon_0 nh}\right)^2$$

$$= -\frac{m}{8\varepsilon_0^2}\frac{Z^2e^4}{n^2h^2} \qquad (24.10)$$

$$= \frac{RZ^2}{n^2}\ ;\ n = 1, 2, 3... \qquad (24.11)$$

where

$$R = \frac{me^4}{8\varepsilon_0^2h^2} \qquad (24.12)$$

is called Rydberg constant, From Eqn. (24.11) we note that

- energy of an electron in various allowed orbits is inversely proportional to the square of the number of orbit; and
- the energy in an orbit is negative, which implies that the electron is bound to the nucleus.

Putting the standard values of $m = 9.11 \times 10^{-31}$kg, $e = 1.6 \times 10^{-19}$C, $\varepsilon_0 = 0.85 \times 10^{-11}\ C^2N^1m^{-2}$, and $h = 6.62 \times 10^{-34}$ Js in Eqn. (24.12), we obtain $R = 2.17 \times 10^{-18}$ J = 13.6eV, since 1eV = 1.6×10^{-19}J. On using this result in Eqn. (24.11), we find that the energy of an electron in n^{th} orbit of hydrogen atom (in eV) is given by

$$E_n = -\frac{13.6}{n^2} \qquad (24.13)$$

Thus every orbit can be specified with a definite energy; the energy of the first orbit being the lowest:

$$E_1 = -13.6 \text{ eV}$$

and the the highest energy state

$$E_\infty = 0$$

It means that different orbits represent different energy levels –13.6eV to 0. This is depicted in Fig. 24.6. $E = 0$ signifies that the electron is free.

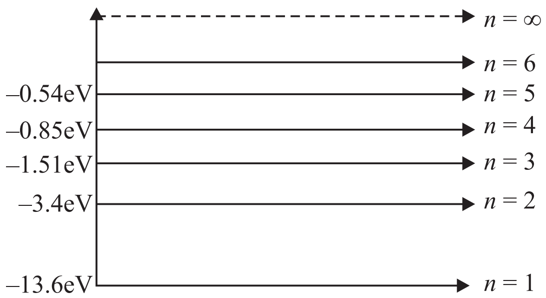


Fig. 24.6: Energy levels in hydrogen atom

According to Bohr's fourth postulate, the frequency υ_{mn} of the emitted (absorbed) radiation when the electron falls (jumps) from the *n*th state to the *m*th state is given by

$$v_{mn} = \frac{RZ^2}{h}\left(\frac{1}{m^2} - \frac{1}{n^2}\right) \quad (24.14)$$

Fraunhoffer Lines

The spectrum of sunlight, when examined carefully by a high power spectroscope, is found to be crossed by a large number of dark lines spread over the length of the continuous spectrum. Wollaston observed these lines in the year 1802. But their existence was studied by Fraunhoffer on the basis of Kirchoff's laws and named these as Fraunhoffer lines. The main body of the sun emits continuous spectrum but the atomosphere of comparatively much cooler vapours and gases in the Sun's atmosphere, called the chromosphere (~6000º C), absorb light corresponding to certain wavelengths. These appear as dark lines in the continuous spectrum of the sun.

Kirchhoff compared the absorbed wavelengths with the wavelengths emitted by various elements present on the earth and identified 60 terrestrial (existing on earth) elements present in the outer atomsphere of sun, e.g. oxygen, hydrogen, sodium, iron, calcium etc.

INTEXT QUESTIONS 24.2

1. Which of Bohr's postulates "fit" with classical physics and which support the ideas of quantum physics?

2. According to Bohr, why did an atom not collapse while its electrons revolved around the nucleus?

3. According to Bohr, what is happening in the atom when a photon of light is (i) emitted (ii) absorbed?

4. Write the energy of the first three orbits of hydrogen atom on the basis of Bohr's model.

5. An atom is excited to an energy level E_1 from its ground state energy level E_0. What will be the wavelength of the radiation emitted?

6. In case of hydrogen atom, the radius of the electron in its nth orbit is proportional to

   (i) $1/n$ (ii) $1/n^2$ (iii) n (iv) n^2

7. The total energy E_n of the electron in the nth orbit of hydrogen atom is proportional to

   (i) e^4 (ii) e^3 (iii) e^2 (iv) e

24.3 HYDROGEN SPECTRUM

Refer to Fig. 24.7. It shows frequency spectrum of hydrogen atom. As may be noted, the line spectrum of hydrogen consists of many lines in different regions of the spectrum. The various lines in a particular region of spectrum are found to have a pattern and may be represented by a common formula. So they are said to form a series. Let us about the series of hydrogen spectrum

Lyman series was discovered by in 1906. Accoding to Bohr, this series arises when an electron jumps to the first orbit (m = 1) from an higher orbit (n = 2, 3, 4...). The frequencies of various spectral lines of this series are given by

$$v_{1n} = \frac{R}{h}\left(\frac{1}{1^2} - \frac{1}{n^2}\right)$$

where n is natural number greater than one.

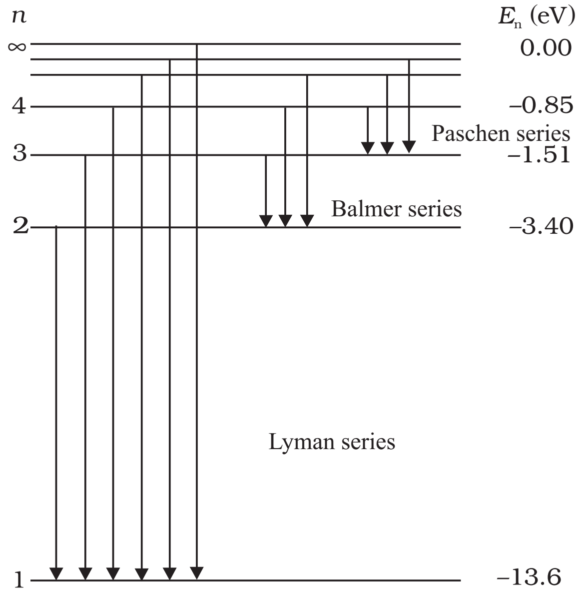


Fig. 24.7 : Energy Level diagram showing emission of various spectral series in hydrogen atom

Balmer series was discovered in 1885 in the visible region. According to Bohr, in this series, electron jumps to the second orbit ($m = 2$) from higher orbits ($n = 3, 4, 5...$). The frequencies of various spectral lines of the series are given by

$$v_{2n} = \frac{R}{h}\left(\frac{1}{2^2} - \frac{1}{n^2}\right);\ n > 2$$

Paschen series was discovered in 1908 in the near infra-red region. The existence of this series can be explained by assuming that electrons jump to third orbit ($m = 3$) from higher orbits ($n = 4, 5, 6...$). The frequencies of various spectral lines in the region are given by

$$v_{3n} = \frac{R}{h}\left(\frac{1}{3^2} - \frac{1}{n^2}\right);\ n_2 > 3$$

Brackett series was discovered in mid infra-red region. In this series, electrons jump to fourth orbit ($n = 4$) from higher orbits ($n = 5, 6...$). Therefore, the frequencies of various spectral lines in the region are given by

$$v_{4n} = \frac{R}{h}\left(\frac{1}{4^2} - \frac{1}{n^2}\right);\ n > 4$$

Pfund series was discovered in far infra-red region. According to Bohr, this series is obtained when electron jumps to fifth orbit ($n_1 = 5$) from any higher orbit ($n = 6, 7,...$). The frequencies of various spectral lines of the series are given by

$$v_{5n} = \frac{R}{h}\left(\frac{1}{5^2} - \frac{1}{n^2}\right);\ n > 5$$

The ingenuity of Bohr's model lies in the fact that it not only explained the already known spectrum but also predicted the existence of a number of series, which were observed later on. In fact, a new physics was born! Transition of the electrons from higher orbits to lower orbits showing emission of different series of spectral lines is shown in Fig. 24.8.

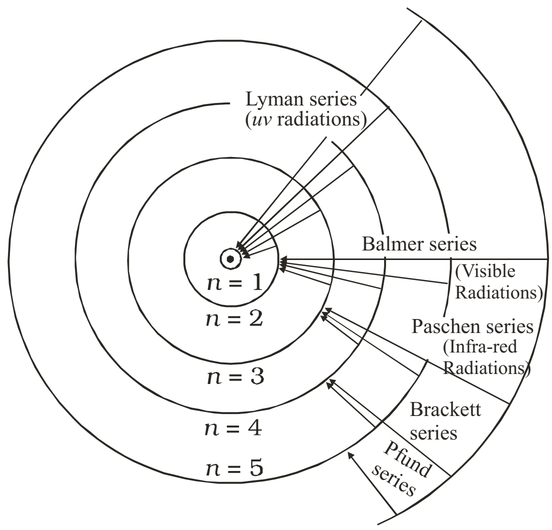


Fig. 24.8 : Permitted orbits in an atom of hydrogen and transitions leading to spectral lines of various series.

INTEXT QUESTIONS 24.3

1. The negative total energy of an orbital electron means that it a) has emitted a photon, b) is bound to the nucleus, c) is in stable equilibrium, d) satisfies Bohr's postulate

   $$L = \frac{nh}{2\pi}.$$

2. An electron jumps to the fourth orbit. When the electron jumps back to the lower energy level, the number of spectral lines emitted will be a) 6, b) 8, c) 5, d) 3.
3. Lymann series of spectral lines are emitted when electron jump from higher orbits to the .... orbit

   a) first, b) second, c) third, d) fourth.
4. Which physical property of electron was quantized by Bohr?
5. An electron jumps from third orbit to first orbit. Calculate the change in angular momentum of an electron?

24.4 X–RAYS

X–rays are produced when fast moving electrons are suddenly stopped by a heavy metal in a glass tube having extremely low pressure. The electrons emitted by the hot filament are focussed on the target which is made up of metal of high melting point and high atomic number as shown in Fig. 24.4.

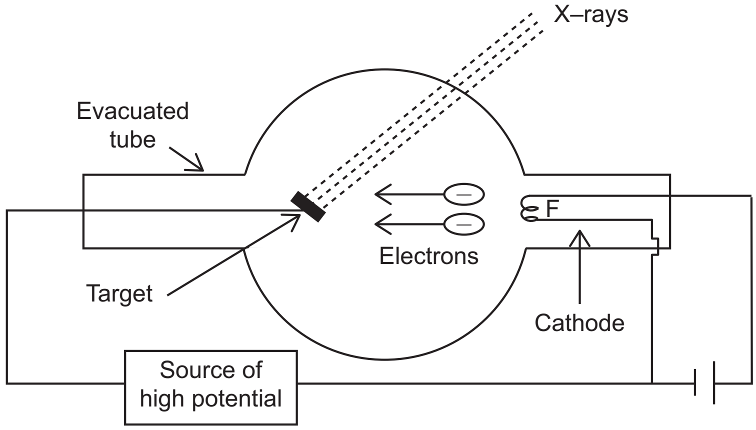


Fig. 24.4

When the electrons approach the target, 5% of their energy gets converted into X–rays and rest of the energy gets converted into heat, which is kept under control by the circulating cold water. The tube has extremely low pressure so that the electrons emitted from the hot filament (*F*) may directly hit the target without suffering collisions in between.

The intensity of the X–rays is controlled by adjusting the filament current and the quality is controlled by the accelerating voltage applied between the filament and target. This voltage usually ranges between 10 kV and 1 MV.

Properties of X–rays

X–rays show the following properties:

(i) They affect the photographic plate

(ii) They cause fluorescence in certain chemical compounds.

(iii) They ionize the gases.

(iv) They show no reflection in mirrors, no refraction in glass, no diffraction with the conventional gratings but when refined techniques are used with atomic layers of crystals, they show all these familiar phenomena of light.

(v) They do not get deviated by electric or magnetic field.

X– Rays Spectra: The element whose X–ray spectra is studied is placed at the place of target of the X–rays tube. The X–ray wavelengths are determined by the Bragg's spectrometer.

X–rays are of two types :

1. **Continuous X–rays:** All the X–ray tubes emit X–rays of all wavelengths beyond certain wavelength λ_{min}. Some of the important features of the spectrum of continuous X–rays are as follows:

   (i) The intensity of X–rays increases at all wavelengths as the voltage across the tube is increased.

   (ii) The shortest wavelength λ_{min} emitted is sharply defined and it depends on the voltage applied.

   (iii) As the voltage is increased, the wavelength at which the maximum emission occurs shifts towards the shortest wavelength side as shown in Fig. 24.5.

   Continuous X–rays are produced when the kinetic energy of the incident electrons is transformed into electromagnetic radiation upon collision with atoms. Before being stopped, electrons make several collisions and produce photons of all frequencies.

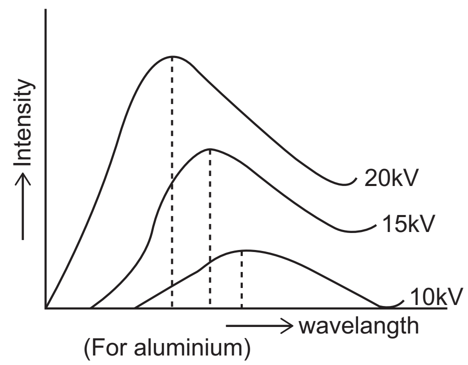


Fig. 24.5

   Photon of largest frequency is produced when an electron makes head -on collision with the atom and loses all its energy at once. For such a collision, the photon frequency ν_{max} is highest and corresponding wavelength λ_{min} is lowest.

   Thus, $$eV = h\nu_{max} = \frac{hc}{\lambda_{min}}$$

   This is known as Duane – Hunt Law.

2. **Characteristic X–rays:** In addition to continuous X–rays, X–rays tubes emit radiations which are characteristic of the target used. It is observed

that on the continuous spectrum, at certain wavelengths, large amounts of energy are radiated. The positions of these lines do not depend on the voltage applied but depend only on the nature of the target used.

Mosley's law

Mosley investigated the characteristic X–rays of a large number of elements. He found that some specific characteristic lines appeared in the spectra of all elements but at slightly differing wavelengths (Fig. 24.6). Each characteristic line obeyed a specific equation. For example, K_2 lines obey the following relation

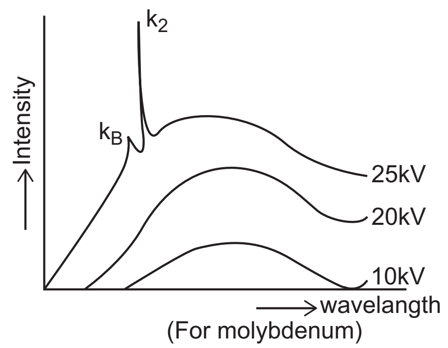


Fig. 24.6

$$\bar{\nu} = R\left[\frac{1}{1^2} - \frac{1}{2^2}\right](Z-1)^2$$

WHAT YOU HAVE LEARNT

- Rutherford's scattering experiment indicated the presence of small central region inside the atom where all the positive charge and most of the mass of the atom is concentrated. The region was named as the nucleus.
- Electrons revolve around the nucleus and total negative charge is equal to the total positive charge of the nucleus.
- Rutherford's model of atom could not explain satisfactorily the observed stability of the atom and the electromagnetic radiation emitted by the atoms.
- A satisfactory model of an atom was suggested by Niels Bohr based on four postulates.
- Permissible orbits for elecrons are those for which angular momentum $(I\omega) = nh/2\pi$
- Emission (absorption) of energy takes place when electron jumps from a higher orbit to a lower orbit (from a lower to a higher orbit).
- The radii of the permitted orbits in which the electron is free to revlove around the nucleus of the hydrogen atom are given by

$$a_n = \frac{n^2h^2}{4\pi^2 mke^2} = \frac{n^2h^2\varepsilon_0}{Ze^2m\pi}$$

For hydrogen atom, the radius of the first permitted orbit is $a = 0.53$Å.

- The energy of the electron in the n^{th} orbit of the hydrogen atom is given by

$$E_n = -\frac{e^4 m}{8h^2\varepsilon_0^2 n^2}$$

The negative sign of total energy indicates that the electron is bound to the nucleus.

- The frequency of the photon emitted when the electron moves from the energy level E_i to E_f is given by :

$$v_{mn} = \frac{R}{h}\left[\frac{1}{m^2} - \frac{1}{n^2}\right]$$

- x-rays are produced when fast moving electrons are suddenly stopped by a heavy metal.
- x-rays are of two types (i) continuous and (ii) characteristic.
- Duane-Hunt law $eV = h\nu_{max} = \dfrac{hc}{\lambda_{min}}$
- Mosley law $\bar{\nu} = R\left(\dfrac{1}{1^2} - \dfrac{1}{2^2}\right)(z-1)^2$

TERMINAL EXERCISE

1. In Rutherford's scattering experiment why do most of the α-particles pass straight through the target foil?
2. In Rutherford's α-particle scattering experiment, what observation led him to predict the existance of nucleus?
3. Why did Rutherford assume that electrons revolve in circular orbits around the nucleus?
4. What is the ratio of the energies of the hydrogen atom in its first excited state to that its second excited state?
5. What is the SI unit of Rydberg's constant?
6. The Rydberg constant for hydrogen is 1096700 m^{-1}. Calculate the short and long wavelength limits of Lyman series.
7. How many times does the electron of H-atom go round the first orbit in 1s?
8. Describe Rutherford's scattering experiment and discuss its findings and limitations.
9. State the postulates of Bohr's model of atom.

10. Derive an expression of the energy of the electron in the nth orbit of hydrogen atom.

11. Choose the correct answers

    (a) The total energy E_n of the electron in the nth orbit of hydrogen atom is proportional to.

    (i) $1/n^4$ (ii) $1/n^2$ (iii) $1/n^2$ (iv) $1/n$

    (b) The energy required to remove an electron from $n = 1$ to $n = \infty$ in the case of hydrogen atom is

    (i) 13.6 V (ii) 13.6 eV (iii) 13.6 MeV (iv) 13. 6 keV

    (c) For hydrogen atom when the electron jumps from higher energy level $n = 5, 6, 7,...$ etc. to the energy level $n = 4$, a set of spectral lines are obtained. These are called

    (i) Balmer Series (ii) Bracket Series

    (iii) Paschen Series (iv) Lyman Series

12. Calculate the radius of the third and fourth permitted orbits of electron in the hydrogen atom.

13. The energy transition in H-atom occurs from $n = 3$ to $n = 2$ energy level. Given $R = 1.097 \times 10^7 \text{ m}^{-1}$.

    (i) What is the wavelength of the emitted radiations?

    (ii) Will this radiation lie in the range of visible light?

    (iii) To which spectral series does this transition belong?

14. The ionisation potential of hydrogen is 13.6 volt. What is the energy of the atom in $n = 2$ state?

15. How are x-rays produced? Draw a labelled diagram to illustrate?

16. List the properties of x-rays and compare them with that of visible light.

17. How are continuous x-rays produced? What is the condition for the production of photons of highest frequency?

ANSWERS TO INTEXT QUESTIONS

24.1

1. a (iii), b (ii), c (i)

2. It could not explain the large angle scattering of particles from the gold foils as observed by Rutherford.

24.2

1. Bohr's first postulate is from classical physics; remaining three are from quantum physics.
2. Because the orbits are stationary.
3. (i) Electron falls from higher to lower energy state.

   (ii) Electron is excited to some higher energy state.
4. $E_1 = -13.6\text{eV}$, $E_2 = 3.4$ eV, $E_3 = -1.51\text{eV}$
5. $\lambda = \dfrac{hc}{E_i - E_0}$
6. (iv)

24.3

1. (b)
2. (a) Number of spectral lines emitted $= \frac{1}{2} n\,(n-1) = \frac{1}{2} \times 4\,(4-1) = 6$
3. (a)
4. Angular momentum of revolving electron.
5. From the *n*th states with principal quantum number *n* calculate the number of wavelengths observed in the spectrum from a hydrogen sample.

Answers to Problems in Terminal Exercise

4. 9 : 4

6. $\lambda_s = 911.4$ Å , $\lambda_l = 1215$ Å

7. 6.57×10^{15} Hz.

11. (a) (i), (b) (ii), (c) (iii), (d) (i), (e) (v)

13. (i) 6563 Å, (ii) visible (iii) Balmer Series. 14. –3.4 eV.

312en25

25

DUAL NATURE OF RADIATION AND MATTER

You must have seen films in cinema halls. The picture on the screen is produced by passing light through films which have the scenes shot on them. But have you ever thought as to how the sound is reproduced in the cinema? The sound is also recorded on the side of the film as a sound track. The light beam passing through this sound track falls on a photocell, which converts it into electrical pulses. These electrical pulses are converted into sound. In this lesson you will study the effect which governs the working of a photocell. It is called the *photoelectric effect*. It is also used in burglar alarm to detect intruders. Einstein's explanation of photoelectric effect led de Broglie (read as de Broy) to the wave-particle duality, i.e. matter exhibits wave as well as particle properties.

You now know that a particle is characterized by properties such as definite position, size, mass, velocity, momentum, etc. Its motion is described by Newton's laws of motion. On the other hand, a wave is characterized by properties such as periodicity in space-time, wavelength, amplitude, frequency, wavevelocity, etc. It transports energy, but no matter. That is, it extends in space unlike a particle, which is localised. The term wave-particle duality refers to the behaviour where both wave-like and particle-like properties are exhibited under different conditions by the same entity. His arguments were simple: Nature likes simplicity and loves symmetry. So if wave-particle duality can be exhibited by light, it should be exhibited by matter as well. You will learn about his explanation of matter waves in sec. 25.4.

OBJECTIVES

After studying this lesson, you should be able to:

- *explain photoelectric effect;*
- *describe the experimental arrangement to study photoelectric effect;*

- *state the laws of photoelectric emission;*
- *interpret the graphs between frequency of radiation and retarding potential;*
- *write deBroglie wavelength of matter waves associated with a particle of momentum* **p**; *and*
- *describe the experimental arrangement for the verification of matter waves.*

25.1 PHOTOELECTRIC EFFECT

In 1887, while working on propagation of electromagnetic waves, Hertz discovered that air in a spark gap became a better conductor when it was illuminated by ultraviolet rays. Further experiments by him showed that zinc became positively charged on irradiation by ultraviolet rays. In 1900, Leonard showed that electrons were emitted from a metal surface when light of sufficiently high frequency falls on it. This phenomenon is known as *photoelectric effect* and the electrons so emitted are called photoelectrons.

The emission of electrons from metals irradiated by light of a frequency greater than a certain characteristic frequency is called photoelectric effect.

The emission of electrons from metals can also take place when they are heated. This is known as **thermionic emission.** *Note that electrons gain energy from thermal energy in thermionic emission.*

25.1.1 Experimental Arrangement to Study Photoelectric Effect

Refer to Fig. 25.1. It shows a schematic diagram of the apparatus that can be used to study this phenomenon.

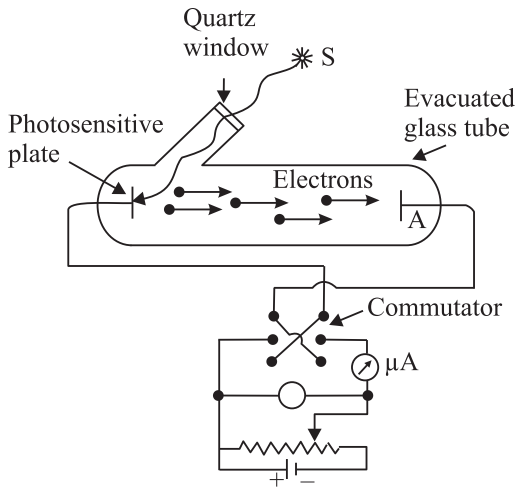


Fig. 25.1: Experimental arrangement for observing the photoelectric effect

A metallic cup *C* called photo cathode is sealed inside an evacuated tube along with another metal plate *A,* which is used to collect photoelectrons emitted by *C*.These electrodes are connected to a battery and microammeter circuit, as shown in Fig.25.1. The battery is so connected that the voltage on plate *A* is positive with respect to *C*. If the battery terminals are reversed, the voltage of the plate *A* will become negative relative to *C*.

To study the effect of intensity of incident light on the number of photoelectrons emitted by *C*, the collector plate *A* is kept at positive potential relative to *C*.

Keeping the frequency of incident light and the value of accelerating potential fixed, the photoelectrons emitted per unit area from the emitting surface vary linearly with the intensity of light, as shown in Fig. 25.2 (a).

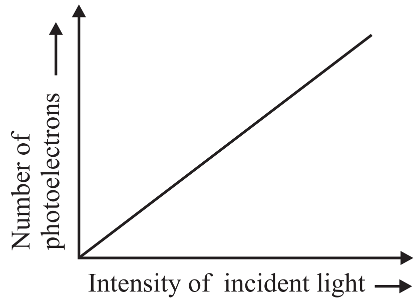


Fig. 25.2 : (a) Variation of number of photoelectrons with intensity of incident light

Case-I: *Plate A positive relative to C*

Let us first consider the case when the plate *A* is at a positive potential relative to *C*. When light of high frequency is incident on the emitter, it starts emitting electrons. Since *A* is at a higher potential relative to *C*, the emitted electrons experience an attractive force. When we increase the voltage on *A*, the kinetic energy of the photoelectrons increases. The current in the outer circuit shown by the microammeter depends on the number of electrons reaching the plate *A*. If we keep on increasing the voltage, a stage comes when all the emitted electrons are collected by the plate. The current is said to have *saturated* at this stage. If the voltage on the plate is increased further, the current remains constant in magnitude. This behaviour of current with respect to plate voltage is shown in Fig. 25.2(b). The voltage V_s is called *saturating voltage*.

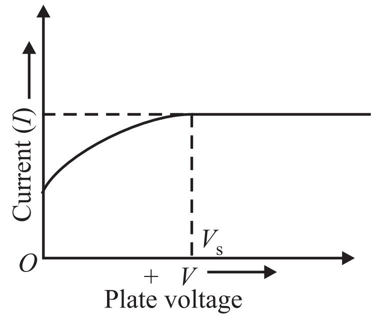


Fig. 25.2 : (b) Dependence of photoelectric current on plate voltage.

Case-II: *Plate A negative relative to C*

If *C* is at a positive potential relative to the plate *A* and light of a proper frequency is incident on the emitter, photoelectrons emitted by *C* will experience retarding potential, which impedes their motion towards *A*. Some of the electrons emitted from *C* may still reach the plate. This gives rise to current, which is registered by the microammeter. What does this mean? If the p.d between *A* and *C* only provides the force which makes the electrons move towards the plate, then none of the electrons should have reached the plate. Since such electrons have overcome the retarding potential while moving against it to reach the plate, they have some initial kinetic energy. This is also confirmed by observed results. For any incident light of particular frequency, if the retarding potential is gradually increased in magnitude, a stage is ultimately reached when none of the electrons reach the plate and the current becomes zero.

The minimum retarding potential for which the photoelectric current becomes zero for a particular frequency of incident light is called the stopping potential, V_0 *for that frequency.*

The work done by an electron W against the stopping potential V_0 is eV_0 where e is electronic charge. This work is done by the electron at the expense of its kinetic energy. So, we can write

$$eV_0 = \frac{1}{2} m v_{max}^2 \qquad (25.1)$$

The stopping potential V_0 was found by Millikan to depend on the frequency of the incident light. A plot of the stopping potential (V_0) versus the frequency of the incident light (ν) is shown in Fig. 25.3. You will note that there is a minimum cut-off frequency ν_0 below which ejection of electrons is not possible. It is called *threshold frequency.*

To study the effect of frequency of incident light on stopping potential, Millikan adjusted the intensity of light at a fixed value for various frequencies and studied the variation of photoelectric current with anode potential. He obtained different values of stopping potential for different frequencies of incident light. Moreover, the stopping potential is more negative for higher frequencies, as shown in Fig. 25.4. This implies that if the frequency of the incident light increases, the maximum kinetic energy of the photoelectrons also increases. Therefore, with increasing frequency, greater retarding potential is required to completely stop the movement of photoelectrons towards the anode.

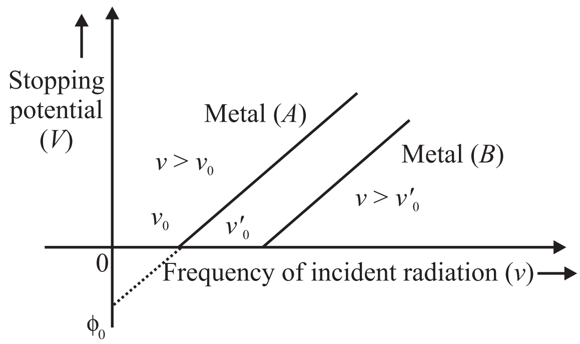


Fig. 25.3 : Stopping potential versus frequency of incident light.

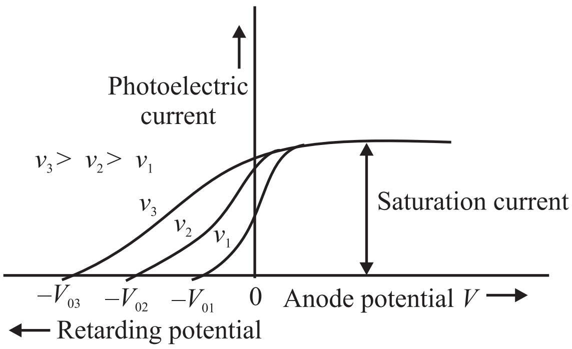


Fig. 25.4 : Photo electric current

This experiment also established that there exists a minimum cut-off frequency ν_0 for which stopping potential is zero. Moreover, photo emission begins as soon as light is incident on the material, i.e. photo emission is instantaneous, even if the incident light is dim. Now it is known that time lag between incident light and emission of photoelectrons from the emitter is of the order of 10^{-9}s.

Notes

These observations can be summarised as follows :

- *The maximum velocity of photoelectrons increases with frequency of incident light and depends on the nature of emitter material.*
- *The maximum velocity of photoelectrons does not depend on the intensity of incident light.*
- *For every material, there exists a thereshold frequency below which no photoelectrons are emitted.*
- *For a particular frequency, the number of photoelectrons emitted per unit area of the emitting surface is proportional to the intensity of the incident light.*
- *There is practically no time lapse* ($\sim 10^{-9}$s) *between the incidence of light on the metal and emission of electrons from it. In other words, photoelectric emission is an instantaneous process.*

INTEXT QUESTIONS 25.1

1. State whether the following statements are true or false :
   (a) In thermionic emission, electrons gain energy from photons.
   (b) The maximum velocity of photoelectron is independent of the frequency of incident radiation.
   (c) There exists a frequency v_0 below which no photoelectric effect takes place.
2. Refer to Fig. 25.3 and interpret the intercepts on x and y-axes and calculate the slope.
3. Draw a graph showing the variation of stopping potential ($-V_0$) with the intensity of incident light.

25.2 EINSTEIN'S THEORY OF PHOTOELECTRIC EMISSION

In 1905, Einstein proposed a simple but revolutionary explanation for the photoelectric effect. He assumed that light consists of bundles of energy, called photons and viewed photoelectric effect as a collision between a photon and a bound electron.

The energy E of a single photon is given by

$$E = hv \qquad (25.2)$$

Robert A. Millikan (1868-1953)

Robert Andrews Millikan was born on March 22, 1868 in U.S.A. During his undergraduate course, his favourite subjects were Greek and Mathematics. But after his graduation in 1891, he took, for two years, a teaching post in elementary physics. In this period, he developed interest in the subject. He received his Ph.D. (1895) for research on polarization of light emitted by incandescent surfaces.

Millikan spent a year (1895-1896) in Germany, at the Universities of Berlin and Göttingen. He returned at the invitation of A.A. Michelson to take appointment as his assistant at the newly established Ryerson Laboratory at the University of Chicago (1896). He became Professor at that University in 1910, a post which he retained till 1921. As a scientist, Millikan made numerous momentous discoveries in the fields of electricity, optics, and molecular physics. His earliest major success was the accurate determination of the charge carried by an electron, using the elegant "falling-drop method". He also proved that this quantity was a constant for all electrons demonstrating the quantised nature of charge.

He also verified experimentally Einstein's photoelectric equation, and made the first direct photoelectric determination of Planck's constant *h*. Throughout his life, Millikan remained a prolific author, making numerous contributions to scientific journals. He was awarded the Nobel Prize in Physics in 1923.

where v is the frequency of the incident light and h is Planck's constant. Let us now assume that a photon of energy hv is incident on the metal surface. Suppose ϕ_0 is the energy needed for an electron to come out of the metal surface. As you have studied earlier, this energy is also called the *work function* of the conductor. *The work function of a conductor is the minimum energy required by an electron to come out of the conductor surface.*

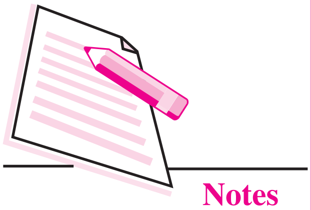
The typical values of work function for a few metals are given (in eV.) in Table 25.1, along with the corresponding threshold frequency (v_0).

Table 25.1: Work function and threshold frequencies of some typical metals

Metal	ϕ_0(eV)	v (Hz)
Sodium	2.5	6.07×10^{14}
Potassium	2.3	5.58×10^{14}
Zinc	3.4	8.25×10^{14}
Iron	4.8	11.65×10^{14}
Nickel	5.9	14.32×10^{14}

What do you think would happen when a photon of energy E (> ϕ_0) strikes the metal surface? We expect that out of the total energy E, an amount ϕ_0 would be used up by the electron to come out of the metal surface. The difference in energy, i.e. ($E - \phi_0$), would then be imparted to the emitted electron in the form of kinetic energy. (The electron may lose some energy in internal collisions before it escapes from the metal surface.) Mathematically, we can write

$$hv = \phi_0 + K_{max} \tag{25.3}$$



Albert Einstein
(1879-1955)

Albert Einstein was born in Wurttemberg, Germany, on March 14, 1879. In 1901, he acquired Swiss citizenship and, as he was unable to find a teaching post, he accepted a position as technical assistant in the Swiss Patent Office. During his stay at the Patent office, in his spare time, he produced much of his remarkable work, including the theory of photoelectric effect and the special theory of relativity. In 1909 he became Professor Extraordinary at Zurich. In 1911 he accepted the post of Professor of Theoretical Physics at Prague but returned to Zurich in the following year to fill a similar post. In 1914, he was appointed Director of the Kaiser Wilhelm Physical Institute and Professor in the University of Berlin. He became a German citizen in 1914. He was awarded the Nobel Prize in Physics in 1921 for his theory of photoelectric effect, though he is more famous for his theory of relativity. He remained in Berlin until 1933, when he renounced his citizenship for political reasons and immigrated to take the position of Professor of Theoretical Physics at Princeton in USA.

He became a US citizen in 1940 and retired from his post in 1945. He spent the later years of his life working on General Theory of Relativity and Unification of basic Forces. Einstein nurtured scientific humanism. He protested to President Roosevelt against the use of nuclear bombs for destruction of humanity. He is considered the greatest scientist to have ever walked on this planet and named scientist of the millenium.

Let us now see how observed results can be explained on the basis of this theory. Let us take

$$\phi_0 = h\nu_0.$$

Then Eqn. (25.3) takes the form

$$K_{max} = \frac{1}{2}mv^2 = h(\nu - \nu_o) \qquad (25.4)$$

This equation implies that

- For v_{max} to be positive, no emission can take place for $\nu < \nu_o$. That is, the incident light must have frequency above the threshold frequency.
- K_{max} is linearly proportional to $(\nu - \nu_o)$
- An increase in the intensity of incident light of frequency ν corresponds to an increase in the number of photons. Each and every photon has same

energy; there is no increase in the energy of photoelectrons. However, the no. of emitted electrons and hence photocurrent will increase with increase in intensity.

- Since photoelectric effect is produced by collisions between photons and electrons, the energy transfer from photons is instantaneous, i.e. there is almost no time lag.
- Since work function is a characteristic property of a material, v_o is independent of the intensity of incident light.

We see that Einstein's theory of the photoelectric effect successfully explains its physical origin.

To understand these concepts and get an idea about the values of physical parameters, go through the following examples carefully.

Example 25.1: Sodium has a work function of 2.3 eV. Calculate (i) its threshold frequency, (ii) the maximum velocity of photoelectrons produced when sodium is illuminated by light of wavelength 5×10^{-7} m, (iii) the stopping potential for light of this wavelength. Given $h = 6.6 \times 10^{-34}$ J s, $c = 3 \times 10^{8}$ m s^{-1}, 1eV = 1.6×10^{-19} J, and mass of electron $m = 9.1 \times 10^{-31}$ kg.

Solution: (i) The threshold frequency is given as $hv_0 = \phi_0$. Here, $h = 6.6 \times 10^{-34}$ Js and $\phi_0 = 2.3$ eV $= 2.3 \times 1.6 \times 10^{-19}$ J.

$$\therefore \qquad v_0 = \frac{\phi_0}{h}$$

$$= \frac{2.3 \times 1.6 \times 10^{-19}\ \text{J}}{6.6 \times 10^{-34}\ \text{J s}} = 5.6 \times 10^{14}\ \text{Hz}$$

(ii) From Einstein's photoelectric equation, we know that

$$hv = \phi_0 + K_{max} = \phi_0 + \frac{1}{2} m v_{max}^2 ,$$

Since $v = \dfrac{c}{\lambda}$, we can write

$$E = h \times \frac{c}{\lambda} = \phi_0 + \frac{1}{2}\left(m v_{max}^2\right)$$

where c is velocity of light and λ is its wavelength. On substituting the given values, we get

$$\therefore \qquad E = \frac{(6.6\times10^{-34}\,\text{J}s)(3\times10^{8}\,\text{ms}^{-1})}{5\times10^{-7}\,\text{m}}$$

$$= 3.96 \times 10^{-19}\ \text{J}$$

$$\Rightarrow \qquad 3.96 \times 10^{-19} = 2.3 \times 1.6 \times 10^{-19} + \frac{1}{2}\, m\, \upsilon_{\max}^2$$

$$= 3.68 \times 10^{-19} + \frac{1}{2}\, m\, \upsilon_{\max}^2$$

$$\therefore \qquad \upsilon_{\max}^2 = \frac{2\times0.28\times10^{-19}}{m} = \frac{2\times0.28\times10^{-19}}{9.1\times10^{-31}}$$

$$\therefore \qquad \upsilon_{\max} = \sqrt{\frac{0.56\times10^{-19}\,\text{J}}{9.1\times10^{-31}\,\text{kg}}} = 2.5 \times 10^{5}\ \text{m s}^{-1}$$

(iii) The stopping potential V_0 is given as

$$eV_0 = \frac{1}{2} m\,\upsilon_{\max}^2$$

$$\therefore \qquad V_0 = \frac{0.28\times10^{-19}\,\text{J}}{1.6\times10^{-19}\,\text{JV}^{-1}} = 0.18\ \text{V}$$

You may now like to answer some simple questions

INTEXT QUESTIONS 25.2

1. Calculate the momentum of a photon of frequency v.
2. If the wavelength of an electromagnetic radiation is doubled, how will be the energy of the photons change?
3. The intensity of incident radiation is doubled. How will it affect the kinetic energy of emitted photoelectrons.

25.3 PHOTOELECTRIC TUBE

You have studied the photoelectric effect in detail now. We know that when light of a frequency above v_0 is incident on a material, electrons are emitted. Their kinetic energies are different. We also know that flow of electrons constitutes current.

The photoelectric tube is based on photoelectric effect.

A photoelectric tube consists of an evacuated glass vessel which contains a semi-cylindrical cathode and an anode in the form of a straight wire. The cathode is coated with a metal of low work function to ensure emission of photoelectrons when light of a pre-decided frequency is incident on it. The threshold frequency above which a phototube responds determines the choice of this coating.

The anode is usually made of nickel or platinum. Electrical connections P_1 and P_2 are brought out on to the surface of the glass vessel. A battery and a microammeter are connected between the anode and the cathode to provide the accelerating voltage. The arrow on the battery indicates that the voltage applied by it can be varied. The microammeter placed in the circuit measures the current passing through it (Fig. 25.5a).

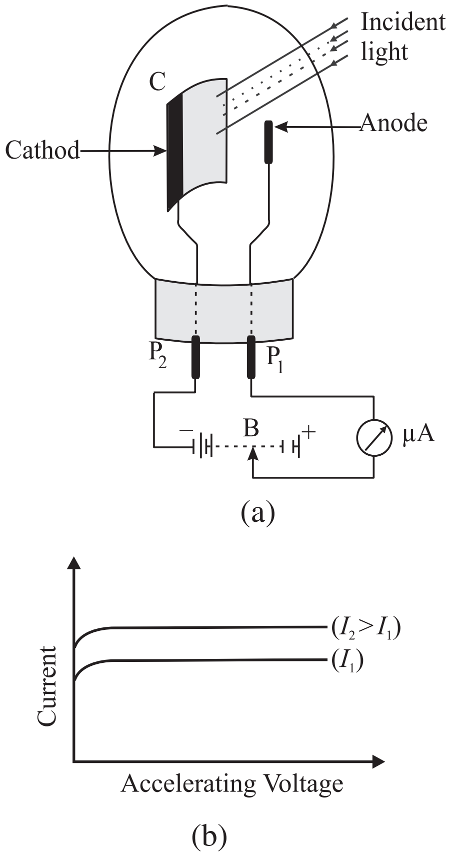


Fig. 25.5: Variation of current and accelerating voltage

To understand the working of a photoelectric tube, let us suppose that light of frequency higher than the threshold frequency is incident on the cathode. Some photoelectrons will be emitted even when accelerating potential between the cathode and the anode is zero. As you know, these electrons come out of the cathode with definite kinetic energy and reach the anode giving rise to a current, which is indicated by the microammeter. Let us now apply some accelerating voltage and see what happens. Obviously, more electrons will reach the anode and increase the current. This is shown in Fig. 25.5(b).

As we keep on increasing the voltage between the cathode and the anode of the photoelectric tube, current also increases. However, at high voltage, the current saturates to a fixed value as shown in Figure 25.5(b). ***The value of saturation current is determined by the intensity (I) of incident light.*** The magnitude of saturation current is of the order of nano-ampere ($\sim 10^{-9}$ A). It is seen that ***if the intensity of the light is increased, the saturation current also increases, as shown in the*** **Fig. 25.5. (b)**

25.3.1 Applications

Photoelectric cells find wide applications in processes wherever light energy has to be transformed into equivalent electric current.

(i) Reproduction of sound in cinematography (motion pictures): One of the most important applications of photoelectric cells is in reproduction of sound in films. A sound track is a track on the film of uniform width whose intensity varies in accordance with the audio frequency variations. Light is made to pass through this film and is then made to fall on the cathode of a photoelectric cell. The current developed in the circuit of the photoelectric cell is proportional to the audio frequency and the variations in current are also in accordance with the variations in the audio frequency. This current is then made to pass through a resistance. The voltage across the resistance is then suitably amplified and fed to a loudspeaker. The loudspeaker reproduces the sound as was originally recorded on the sound track. You will learn more about it on the optional module on photography and video-recording.

(ii) Transmitting pictures over large distances : Photo electric tubes are also used in systems that transmit pictures over large distances. The technique of transmission of pictures to large distances is called **photo-telegraphy.**

(iii) Other Uses : Many types of systems used for counting articles or living beings are based on photoelectric tubes. These are also used in burglar alarms, fire alarms, detectors used for detection of traffic law defaulters, in television camera for scanning, telecasting scenes and in industry for detecting minor flaws or holes in metal sheets.

INTEXT QUESTIONS 25.3

1. State whether the following statements are true or false:
   (a) The cathode in a phototube is biased positively with respect to anode.
   (b) The saturation current in a phototube depends on the frequency of incident radiation.
   (c) The saturation current in a photodiode increases with intensity of incident light.
2. State three applications of photoelectric tube.
3. A phototube is illuminated by a small bright source placed 100 cm away. When the same source of light is 50 cm away, what will be the effect on the number of electrons emitted by photo cathode?

In the previous section, you have studied Einstein's theory of photoelectric effect and learnt that light consists of photons. You have also learnt that the phenomena of interference and diffraction can be explained on the basis of wave theory of light. This duality in the nature of light came to be accepted by the physicists in the early 20th century. Thinking about the wave-particle duality of light, de Broglie asked himself the question : If light exhibits dual nature, will particles of matter also not act like waves? Successful resolution of this question led to de Broglie hypothesis.

25.4 THE DE BROGLIE HYPOTHESIS

As a young graduate student, de Broglie argued with a great amount of insight that since nature loves symmetry and simplicity in physical phenomena, ordinary "particles" such as electrons, and protons should also exhibit wave characteristics under certain circumstances. His argument runs as follows : Light is an electromagnetic radiation and exhibits wave-particle duality. Therefore, Einstein's mass-energy equivalence relation ($E = mc^2$), which essentially treats light as quantum of photon, a particle, can hold only if matter also exhibits wave character. He therefore proposed that the wavelength and frequency of matter waves should be determined by the momentum and energy of the particle in exactly the same way as for photons : $E = pc$ and the associated wavelength λ of a particle having momentum p is given by

$$\lambda = \frac{h}{p} \tag{25.5}$$

Since the momentum of such a particle is gives by $p = mv$, we can write

$$\lambda = \frac{h}{mv} \tag{25.6}$$

λ is called deBroglie wavelength. Eqn. (25.5) is a complete statement of wave-particle duality. It implies that a particle with a momentum p can exhibit wave-like properties and the wavelength of the associated matter waves is h/p. The converse is also true, i.e., a wave of wavelength λ can exhibit particle-like properties and the momentum of the wave-matter is h/λ

This hypothesis, submitted as Ph.D Thesis was initially rejected by the examiners. However, soon, experimental evidence proved de Broglie's argument. This has a very important inspirational lesson for us : We must keenly analyse every statement and try to seek experimental evidence.

The actual wavelength of anything macroscopic is incomprehensively small, as you can see by calculating it for a cricket ball. The case is quite different for elementary particles such as the electron. An electron has energy E when accelerated through potential difference V. Hence, we can write

$$\frac{1}{2}mv^2 = qV \qquad (25.7)$$

or

$$v = \sqrt{\frac{2qV}{m}}$$

so that

$$mv = p = \sqrt{2qmV} \qquad (25.8)$$

On combining this result with Eqn.(25.5), we find that de-Broglie wavelength is given by

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2qmV}} \qquad (25.9)$$

The constants appearing in Eqn. (25.9) have the values: $h = 6.625 \times 10^{-34}$ Js, $q = 1.602 \times 10^{-19}$ C and $m = 9.11 \times 10^{-31}$ kg. On substituting these values in Eqn.(25.9), we obtain

$$\lambda = \frac{6.625\times10^{-34}\text{Js}}{\sqrt{2\times(1.602\times10^{-19}\text{C})\times(9.11\times10^{-31}\text{kg})}\times\sqrt{V}}$$

$$= \frac{12.3}{\sqrt{V}}\times10^{-10}\ \text{m}$$

$$= \frac{12.3}{\sqrt{V}}\ \text{Å} \qquad (25.10)$$

It means that if an electron is accelerated through a potential difference of 100V, its wavelength will be given by

$$\lambda = \frac{12.3}{\sqrt{100}}\text{Å}$$

$$= 1.23\ \text{Å}$$

This is also the wavelength of an electron of energy 100eV. You can easily verify this using the relation

$$\lambda = \frac{h}{(2meE)^{½}}$$

The wavelength of matter waves associated with 100eV electrons lies in the X-ray region and is of the same order as the interatomic seperation in a solid. We therefore expect these to undergo diffraction by a crystal lattice.

The first experimental evidence of matter waves came from the work of Davisson and Germer, who were studying scattering of electrons by crystals. Let us learn about it now.

Louis Victor de Broglie

(1892-1987)

Louis de Broglie was born at Dieppe, France on 15th August, 1892. He first studied the arts and took his degree in history in 1910. Then, as his liking for science prevailed, he studied for a science degree, which he gained in 1913. In 1924 at the Faculty of Sciences at Paris University, he submitted a thesis *Recherches sur la Théorie des Quanta* (Researches on the quantum theory), which gained him his doctor's degree. This thesis contained a series of important findings, which he had obtained in the course of about two years. The ideas set out in that work served the basis for developing *wave mechanics,* a theory which has transformed our knowledge of physical phenomena on the atomic scale.

In 1929 he was awarded the Nobel Prize for Physics for his discovery of the wave nature of electrons.

25.4.1 Experimental Evidence for Existence of de Broglie Waves

The schematic diagram of Davisson-Germer experiment is shown in Fig.25.6. The set up consists of a filament *F*, which serves as a source of electrons. The electrons emitted from this filament are made to pass through a set of metal diaphragms having a number of slits. The electrons emitted by the filament come out in various directions. The metal diaphragms serve to

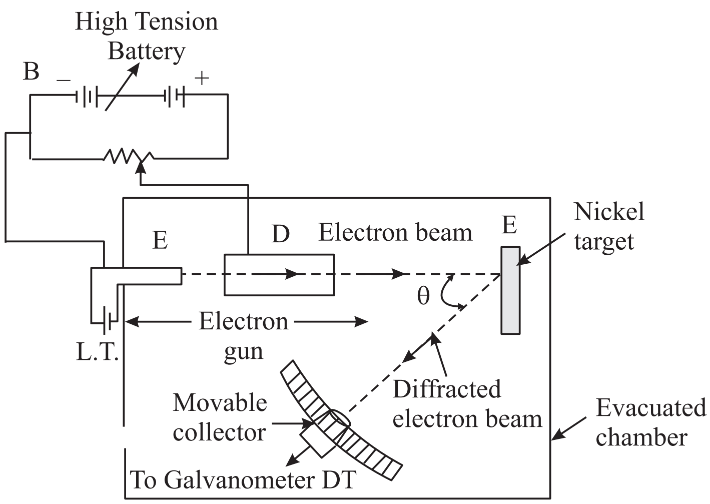


Fig. 25.6: Experimental set up to verify the existence of matter waves

collimate these electrons. Only those electrons which are able to pass through the slits in the various diaphragms are able to come out.

Note that the energy of the collimated stream of electrons is controlled by changing the magnitude of the accelerating voltage. The beam of electrons is made to fall perpendicularly on a single crystal of nickel. The set-up also contains a detector Dt which can be placed at any particular angle with respect to the normal to the target crystal. This detector determines the intensity of the reflected beam. Note that there is nothing special in the choice of nickel.

Fig. 25.7 shows a plot of detector current versus kinetic energy of incident electrons for $\theta = 50°$. As may be noted, the detector current shows a maxima for electrons of kinetic energy 54 eV. If you calculate the wavelength of these electrons, you will get

$$\lambda = \frac{6.62\times10^{-34}\,\text{Js}}{[2\times(9.1\times10^{-31}\,\text{kg})\times54\times1.6\times10^{-19}\,\text{J}]^{1/2}}$$

$$= 1.67\ \text{Å}$$

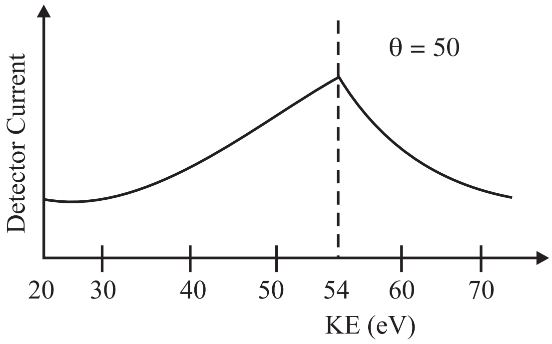


Fig. 25.7: Plot of detector current versus kinetic energy of electrons

25.4.2 Applications of de Broglie Waves

We now know that very small values of wavelength can be achieved by increasing the kinetic energy of electrons. From Lesson 23 you may recall that resolving power of an optical microscope depends on the wavelength of light used. In fact, the resolution increases with decreasing wavelength. Can you guess as to what would happen if a beam of very energetic electrons is used in a microscope instead of photons? Well, obviously you could obtain very high resolution and magnification by lowering the deBroglie wavelength associated with the electrons. This technique is used in electron microscopes. This is an extremely useful application of deBroglie waves. A comparison of structure and working of electron microscope with optical microscope is shown in Fig. 25.8.

Notes

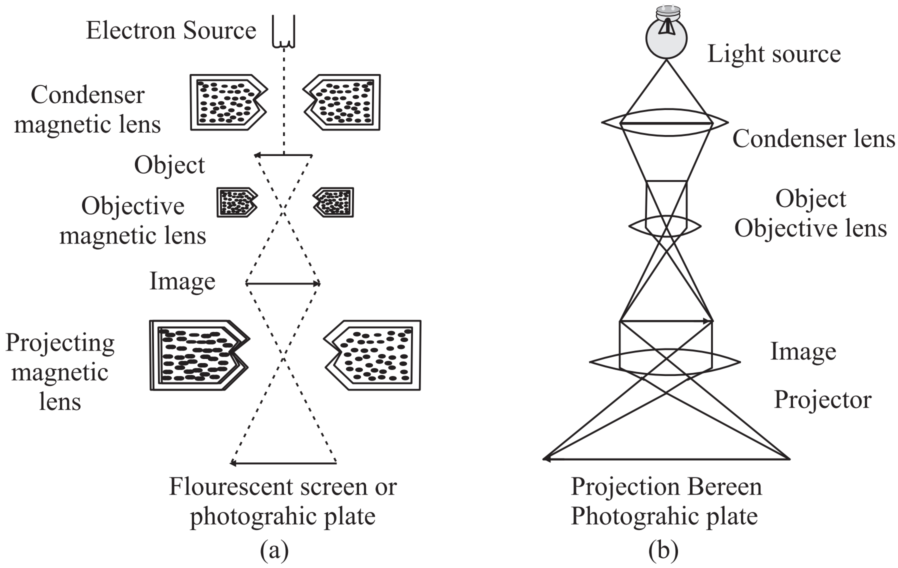
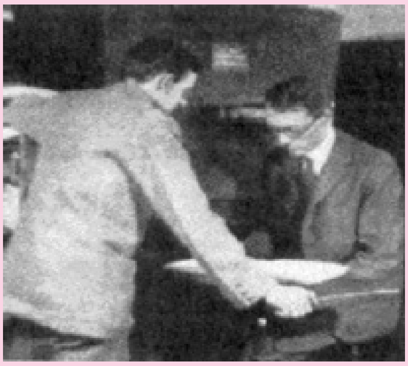


Fig. 25.8: a) Electron microscope, and **b)** Optical microscope

Story of Davisson and Germer's experiment



Germer had recorded in his notebook that he discovered a crack in the vacuum trap in the electron scattering apparatus on Feb. 5, 1925 when he was working with Clinton Davisson at Westren Electric, New York, U.S.A.This was not the first time their equipment had broken, and not the first time they had "resurrected" their precious nickel crystal by **heating it in vacuum and hydrogen.**

This particular break and the subsequent method of repair, however, had a crucial role to play in the later discovery of electron diffraction. By 6 April 1925, the repairs had been completed and the tube put back into operation. During the following weeks, as the tube was run through the usual series of tests, results very similar to those obtained four years earlier were obtained. Then suddenly, in the middle of May, unprecedented results began to appear. This puzzled Davisson and Germer so much that they halted the experiments a few days later, cut open the tube and examined the target (with microscopist F. F. Lucas) to see if they could detect the cause of new observations. What they found was this: the polycrystalline form of nickel target had been changed by the extreme heating until it had formed about ten crystal facets in the area from which the incident electron beam was scattered. Davisson and Germer surmised that the new scattering pattern must have been caused by the new crystal arrangement of the target. In other words, they concluded that it was the arrangement of the atoms in the crystals, not the structure of the atoms that was responsible for the new intensity pattern of scattered electrons.

During the summer of 1926, Davisson and his wife had planned a vacation trip to relax and visit relatives in England. Something was to happen on this particular trip. Theoretical physics was undergoing fundamental changes at this time. In the early months of 1926, Erwin Schrodinger's remarkable series of papers on wave mechanics appeared, following Louis de Broglie's papers of 1923-24 and Albert Einstein's quantum gas paper of 1925. These papers were the subject of lively discussions at the Oxford meeting of the British Association for the Advancement of Science.

Davisson, who generally kept abreast of recent developments in his field but appears to have been largely unaware of these developments in quantum mechanics, attended this meeting. He was surprised when he heard a lecture by Born in which his own and Kunsman's (platinum target) curves of 1923 were cited as confirmatory evidence for de Broglie's electron waves!

Davisson shared the 1937 Nobel Prize for Physics with G.P. Thomson (son of J.J. Thomson).

Electron Microscope

Electron microscopes are scientific instruments that use a beam of highly energetic electrons to examine objects on a very fine scale. This examination can yield the following information:

The surface features of an object or "how it looks", its texture; direct relation between these features and material properties (hardness, reflectivity, etc.), the shape and size of the particles making up the object; direct relation between these structures and materials properties (ductility, strength, reactivity, etc.), the elements and compounds that the object is composed of and the relative amounts of them; direct relationship between composition and material properties (melting point, reactivity, hardness, etc.). How are the atoms arranged in the object?

Electron microscopes were developed due to the limitations of optical microscopes, which are limited to 500× or 1000× magnification and a resolution of 0.2 micrometers. In the early 1930's, this theoretical limit had been reached and there was a scientific desire to see the finer details of the interior structures of organic cells (nucleus, *mitochondria,*etc.). This required 10,000× plus magnification which was just not possible using the microscopes available at that time.

The Transmission Electron Microscope (TEM) was the first Electron Microscope to be developed and is patterned exactly on the Light Transmission Microscope, except that a focused beam of electrons is used instead of light to image the specimen and gain information about its structure and composition. It was developed by Max Knoll and Ernst Ruska in Germany in 1931.

Transmission Electron Microscope (TEM)

A TEM works much like a slide projector. A projector throws a beam of light on the slide. As the light passes through the side, it is affected by the structures and objects on the slide. As a result, only certain parts of the light beam are transmitted through certain parts of the slide. This transmitted beam is then projected onto the viewing screen, forming an enlarged image of the slide.

TEMs works in the same way, except that they shine a beam of electrons through the specimen. Whatever part is transmitted is projected onto a phosphor screen for the user to see.

The electron gun, produces a stream of monochromatic electrons. This stream is focused to a small, thin, coherent beam by the use of condenser lenses 1 and 2. The first lens (usually controlled by the "spot size knob") largely determines the "spot size"; the general size range of the fm spot that strikes the sample. The second lens usually controlled by the "intensity or brightness knob" actually changes the size of the spot on the sample; changing it from a wide dispersed spot to a pinpoint beam. The beam is restricted by the condenser **aperture,** knocking out high angle electrons (those far from the optic axis. The beam strikes the **sample specimen** and parts of it are transmitted.

This transmitted portion is focused by the objective lens into an image Optional Objective and Selected Area metal **apertures** can restrict the beam; the Objective aperture enhances contrast by blocking out high-angle diffracted electrons, the Selected Area aperture enables the user to examine the periodic diffraction of electrons by ordered arrangements of atoms in the sample.

The image is passed down the column through the intermediate and projector lenses, being enlarged all the way.

The image strikes the phosphor image screen and light is generated, allowing the user to see the image. The darker areas of the image represent those areas of the sample that fewer electrons were transmitted through (they are thicker or denser). The lighter areas of the image represent those areas of the sample that more electrons were transmitted through (they are thinner or less dense).

Example 25.2: An electron is accelerated through a potential difference of 182 V. Calculate its associated wavelength.

Solutions: We know that deBroglie wavelength, $\lambda = \frac{h}{p} = \frac{12.3}{\sqrt{V}}$ Å. Here $V = 182$V.

$$\therefore \qquad \lambda = \frac{12.3}{\sqrt{182}}\text{Å} = \frac{12.3}{13.5} = 0.91\text{Å}$$

Example 25.3: Calculate the maximum kinetic energy of the emitted photoelectrons when light of frequency $v = 10^{15}$ Hz is incident on a zinc plate. The work function of zinc is 3.4 eV.

Solution: From Einstein's relation, we recall that

$$hv = \phi_0 + K_{max}$$

For this problem, $h = 6.625 \times 10^{-34}$ Js, $v = 10^{15}$ Hz, $E = hv = 6.625 \times 10^{-34} \times 10^{15}$ $= 6.625 \times 10^{-19}$ J and $\phi_0 = 3.4$ eV $= 3.4 \times 1.602 \times 10^{19}$J $= 5.4468 \times 10^{-19}$ J

$$\therefore \qquad K_{max} = E - \phi_0 = (6.625 - 5.447) \times 10^{-19}\text{J} = 1.178 \times 10^{-19}\text{ J}$$

INTEXT QUESTIONS 25.4

1. State whether the following statement are true or false:
   (a) According to deBroglie, stationary particles exhibit wave-like characteristics.
   (b) Matter waves are the same thing as deBroglie waves.
   (c) Very poor resolution can be obtained in a microscope using energetic electrons by lowering deBroglie wavelengths associated with electrons.
2. A 50 g ball rolls along a table with a speed of 20 cm s^{-1}. How large is its associated wavelength? Given $h = 6.625 \times 10^{-34}$ Js.
3. Why can we not observe de Broglie wavelength associated with a cricket ball?

WHAT YOU HAVE LEARNT

- Emission of electrons from a metal when light of proper frequency incident on its surface is called photoelectric emission.
- In photoelectric emission, electrons gain energy from light.
- The stopping potential increases with increase in frequency of incident light.
- There exists a frequency v_0 for every material below which no photoelectric effect takes place.
- The maximum velocity of photoelectrons increases with increasing frequency of incident light but is independent of the intensity of incident light.

- The number of photoelectrons emitted from each square centimeter of the emitting surface for any particular frequency is proportional to the intensity of incident light.
- Einstein assumed light to consist of photons, each having energy $h\nu$, where ν is frequency and h is Planck's constant.
- Photoemissive type of phototube is based on the photoelectric effect.
- The saturation current of a phototube increases with increasing intensity of the incident light.
- Particles in motion have waves associated with them. The wavelength is given by h/p, where, p is the momentum.

TERMINAL EXERCISE

1. In photoelectric emission, what happens to the incident photons?
2. What is the difference between a photon and a matter particle?
3. Why is the wave nature of matter not apparent in daily life?
4. How is velocity of photoelectrons affected if the wavelength of incident light is increased?
5. The threshold frequency of a metal is 5×10^{14} Hz. Can a photon of wavelength 6000Å emit an energetic photoelectron?
6. Does the threshold frequency for a metal depend on the incident radiations?
7. What are the various uses of photocell?
8. What was the aim of Davisson and Germer's experiment? On what principle does it depend?
9. Describe the experiment used for studying the photoelectric effect.
10. Explain the terms (a) Saturation voltage and (b)Stopping potential.
11. State the laws of photoelectric emission.
12. Describe the salient features of Einstein's theory of photoelectric effect.
13. Explain Einstein's relation: $h\nu = E_0 + K_{max}$
14. Calculate the wavelength associated with electrons moving with a velocity $\upsilon = 1 \times 10^8$ ms^{-1}. Take mass of electron = 9.1×10^{-31} kg and $h = 6.6 \times 10^{-34}$ J.s.
15. Describe an experiment which verifies the existence of deBroglie waves.

16. Show that the deBroglie wavelength associated with electrons accelerated through a potential V is given by the relation;

$$\lambda = \frac{12.3}{\sqrt{V}} \text{ Å}$$

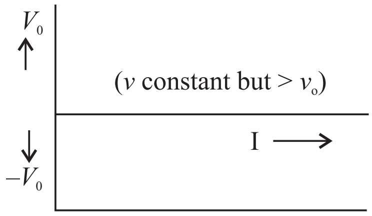
ANSWERS TO INTEXT QUESTIONS

25.1

1. (a) False (b) False (c) True
2. x – intercept gives the threshold frequency

   y – intercept gives $e \times$ work function (ϕ_0)

   $V_0 = \frac{h}{e} \nu - \frac{h}{e} \nu_0$. slope of graph gives $\frac{h}{e}$

3. 

25.2

1. $\lambda = \frac{h}{p} \Rightarrow p = \frac{h}{\lambda} = \frac{h}{c/\nu} = \frac{h\nu}{c}$
2. $E = hc/\lambda$

   If λ is doubled, E will become half
3. It is unchanged.

25.3

1. (a) False (b) False (c) True
2. (i) Reproduction of sound in films,

   (ii) Transmistting pictures over great distances.

   (iii) Thiefe detecting system.
3. Number of photo electrons will increase by a factor of 4.

Notes

25.4

1. (a) false, (b) True (c) True

2. $P = mv$ and $\lambda = \dfrac{h}{P}$

   Here $m = 50\text{g} = 0.05\text{kg}$ and $v = 20 \text{ cms}^{-1} = 0.02\text{ms}^{-1}$

   $\therefore \lambda = 6.6 \times 10^{-32}\text{m}$

3. From Eqn.(25.14) it is clear that if mass *m* is large, the value of λ will be small. Same is the case with cricket ball.

4. 7.25Å

Answers to Problems in Terminal Exercise

14. 7.25 Å

312en26

26

NUCLEI AND RADIOACTIVITY

So far you have learnt that atom is the smallest entity that acts as the building block of all matter. It consists of an extremely small *central core,* called the *nucleus,* around which electrons revolve in certain specified orbits. Though nucleus is very tiny, it is *amazingly complex and you may like to know more about it*. The march towards our understanding the physics of nuclei began towards the end of nineteenth century with the chance discovery of the natural phenomenon of radioactivity; disintegration of atomic nuclei to attain stability. This discovery provided us tools to probe the structure of nucleus : What is its size and mass? What does it contain? What forces make its costituent particles cling together and why?

In fact, the α-particles used by Geiger and Marsden to 'see' what was inside an atom were obtained from naturally occuring radioactive element ^{214}Bi. These inverstigations opened up very fertile and new avenues of research. A lot of good new physics of the atom began to emerge out and changed the course of developments in a short span of time. You will learn about these now.

OBJECTIVES

After studying this lesson, you should be able to :

- *determine the number of neutrons and protons in nuclei of different atoms;*
- *calculate the sizes of atomic nuclei;*
- *explain the nature of forces between nucleons;*
- *explain the terms 'mass defect' and 'binding energy';*
- *draw binding energy per nucleon curve and discuss the stability of atomic nuclei;*
- *discuss the phenomenon of radioactivity, and identify the three types of radioactive radiations;*

- *explain the growth and decay of radioactivity in a sample;*
- *calculate the half-life, and decay constant of a radioactive substance; and*
- *explain the uses of radioactivity in various fields.*

26.1 THE ATOMIC NUCLEUS

Soon after the discovery of nucleus in an atom by Rutherford in 1911, physicists tried to study as to what resides inside the nucleus. The discovery of neutron by James Chadwick in 1932 gave an impetus to these searches as it clearly suggested to the scientific world that the building blocks of the nucleus are the protons and the neutrons.

26.1.1 Charge and Mass

The atomic nucleus contains two types of particles, ***protons and neutrons***. While *protons* are *positively charged, neutrons* are *neutral*. The *electrons,* which revolve in certain specified orbits around the nucleus, are *negatively charged* particles. The magnitude of charge on a proton in a nucleus is exactly equal to the magnitude of charge on an electron. Further, the number of protons in a nucleus is also equal to the number of electrons so that the atom is as a whole is electrically neutral.

Neutrons and protons are collectively referred to as *nucleons.* Their combined number in a nucleus, that is the number of nucleons, is called the *mass number.* It is denoted by A. The number of protons in a nucleus (or the number of electrons in an atom) is called the *atomic number*. It is denoted by Z. The number of neutrons in a nucleus is usually denoted by $N = A - Z$. Usually $N \geq Z$. The difference $(N–Z)$ increases as A increases. Note that for a lithium nucleus containing 3 protons and 4 neutrons, the atomic number Z is 3, and the mass number A is 7.

Protons are slightly lighter than neutrons and almost the entire mass of an atom is concentrated in its nucleus. The mass of a nucleus is nearly equal to the product of A and the mass of a proton (or that of a neutron). Since mass of a proton is 1.67×10^{-27} kg, and A lies between 1 and 240 for most nuclei, the masses of nuclei vary roughly between 1.67×10^{-27}kg and 4.0×10^{-25}kg.

The charge of a nucleus is equal to Ze, where e is the fundamental unit of charge (that is the magnitude of charge on an electron). You may recall that it is equal to 1.6×10^{-19}C. For naturally occurring nuclei, Z varies from 1 to 92, while for transuranic elements (i.e. the artificially produced elements), Z varies from 93 to 105.

26.1.2 Size

The sizes of atomic nuclei are usually quoted in terms of their radii. Many nuclei are nearly spherical in shape and the radius R is given approximately by the formula

$$R = r_0 A^{1/3}$$

Here r_0 is the unit nuclear radius and its numerical value is taken as 1.2 fermi, a unit of length in honour of famous physicist Enrico Fermi. It is equal to 10^{-15}m. The radius of the lightest nucleus (hydrogen) is thus about 1.2 *f*, as *A* for hydrogen is one. The radius of the heaviest naturally occuring nucleus (uranium) is approximately 7.5*f*, as $A = 238$. You may note here that since the volume of any spherical object of radius *r* is equal to (4/3) πR^3, the volume of a nucleus is proportional to *A*, the mass number.

Can you now guess the volume of a nucleus relative to that of an atom? Knowing that the sizes of the nucleus and of the atom are approximately 10^{-15} m and 10^{-10}m, respectively, the volume of an atom is roughly 10^{+5} times the volume of a nucleus. To enable you to visualise these dimensions, the volume of a nucleus relative to atom is something like the volume of a bucket of water relative to the volume of water in Bhakra Dam.

You may now also like to know the order of magnitude of the density of nuclear matter. If we consider the lightest nucleus, hydrogen, whose mass is 1.673×10^{-27} kg and the radius is 1.2×10^{-15} m, and take it to be spherical, the density can be calculated using the relation

$$d_H = \frac{M_H}{\frac{4\pi}{3} R_H^3} = \frac{1.673 \times 10^{-27}\,\text{kg}}{\frac{4\pi}{3} \times \left(1.2 \times 10^{-15}\,\text{m}\right)^3} = 2.3 \times 10^{17}\,\text{kg m}^{-3}.$$

For oxygen, $R_0 = 3 \times 10^{-15}$ m and $M_0 = 2.7 \times 10^{26}$ kg, so that

$$d_0 = 2.39 \times 10^{17}\ \text{kg m}^{-3}$$

That is, the densities of hydrogen and oxygen are of the same order. You may recall that the density of water is 10^3 and density of mercury is 13.6×10^3 kgm^{-3}. It means that nuclear matter is extremely densely packed. To give you an idea of these magnitudes, if our earth were such a densely packed mass ($= 6 \times 10^{24}$ kg), it would be a sphere of radius 184m only. Similarly, the radius of nuclear sphere, whose mass will be equal to the mass of our sun will be 10 km!

26.1.3 Notation

The nucleus of an atom is represented by the chemical symbol of the element, with the *A* value as its superscript and *Z* value as its subscript; both on the left hand side of the chemical symbol. Thus if the chemical symbol of an element is, say, X, its nucleus is represented by ${}^{A}_{Z}\text{X}$. For example, for the nucleus of chlorine, which has 17 protons and 18 neutrons, we write ${}^{35}_{17}\text{Cl}$. Note that 35 here is mass number.

The atoms of different elements can have the same mass number, though they may have different number of protons. *Atoms having the same A value but different Z values are called Isobars.* Thus argon with $A = 40$ and $Z = 18$ is an isobar of calcium which has $A = 40$ and $Z = 20$. Note that isobars have different chemical properties since these are determined by Z. *Atoms of the same element having the same Z value but different A values are called isotopes.* Thus, chlorine with $Z = 17$ and $A = 35$, and chlorine with $Z = 17$ and A = 37, are isotopes of some element, chlorine. Since isotopes have same Z value, they show identical chemical properties. Note that isotopes differ in the number of neutrons in their nuclei. Atoms having the same number of neutrons in their nuclei are called the *isotones*. Thus, sodium with $A = 23$ and $Z = 11$ is an isotone of magnesium with $A = 24$ and $Z = 12$.

Example 26.1 : Calculate the number of electrons, protons, neutrons and nucleons in an atom of $^{238}_{92}\text{U}$.

Solution : $^{238}_{92}\text{U}$ symbolises uranium, which has 92 protons and 238 nucleus. Hence Atomic number $Z = 92$ = number of protons

Mass number $A = 238$ = number of (protons + neutrons) = Number of nucleons

$$\text{Number of neutrons } = A - Z$$
$$= 238 - 92$$
$$= 146.$$

Example 26.2 : Select the pairs of Isotopes, Isobars and Isotones in the following list.

$^{12}_{6}\text{C}$, $^{27}_{13}\text{Al}$, $^{39}_{19}\text{K}$, $^{22}_{14}\text{Si}$, $^{76}_{32}\text{Ge}$, $^{40}_{20}\text{Ca}$, $^{76}_{34}\text{Se}$, $^{14}_{6}\text{C}$

Solution : Isotopes – (Same Z - value) : $^{12}_{6}\text{C}$ and $^{14}_{6}\text{C}$

Isotones – [Same $A - Z$ values] : $\left[^{27}_{13}\text{Al and } ^{28}_{14}\text{Si}\right]$, $\left[^{39}_{19}\text{K and } ^{40}_{20}\text{Ca}\right]$

Isobars – (Same A values) : $^{76}_{32}\text{Ge}$ and $^{76}_{39}\text{Se}$

INTEXT QUESTIONS 26.1

1. Make groups of Isotopes, Isobars and Isotones from the following collection of different atoms:

Notes

${}^{16}_{8}O$, ${}^{207}_{82}Pb$, ${}^{12}_{6}C$, ${}^{40}_{18}A$, ${}^{3}_{1}H$, ${}^{7}_{3}Li$, ${}^{76}_{32}Ge$, ${}^{76}_{34}Se$, ${}^{3}_{2}He$, ${}^{40}_{20}Ca$, ${}^{7}_{4}Be$, ${}^{2}_{1}H$,

${}^{14}_{6}C$, ${}^{235}_{92}U$, ${}^{206}_{82}Pb$, ${}^{18}_{8}O$, ${}^{239}_{92}U$, ${}^{1}_{1}H$, ${}^{23}_{11}Na$, ${}^{27}_{13}Al$, ${}^{27}_{12}Mg$, ${}^{28}_{14}Si$, ${}^{37}_{17}Cl$, ${}^{35}_{17}Cl$

2. Fill in the blanks :
   (i) Neutron is .... than proton.
   (ii) The total number of protons and neutrons in an atom is called the .... number of that atom.
   (iii) The protons and neutrons together are called ....
   (iv) The number of neutrons in ${}^{27}_{13}Al$ = ...
   (v) The number of protons in ${}^{28}_{14}Si$ = ...
   (vi) Two atoms are said to belong to different elements if their .... numbers are different.

3. Which number cannot be different in two atoms of the same element – mass number, atomic number, neutron number?

26.1.4 Unified Atomic Mass

It has been experimentally determined that mass of proton (m_p) is 1836 times the mass of electron (m_e), and the mass of neutron (m_n) is 1840 m_e. Since the mass of an electron is negligibly small compared to the mass of a nucleon, the mass of an atom is effectively due to the mass of its nucleons. However, the neutron is slightly heavier than the proton. It is, therefore, desirable to choose a standard to express the masses of all the atoms (and also that of protons and neutrons). Now a days, atomic masses are expressed in terms of the actual mass of ${}^{12}_{6}C$ isotope of carbon. The unit of atomic mass, abbreviated as u, is defined as $(1/12)^{th}$ of the actual mass of ${}^{12}_{6}C$. We know that the value of the mass of a carbon atom is 1.99267×10^{-26}kg. Hence

$$1u = (1/12) \times \text{mass of one carbon atom with } A = 12$$

$$= (1/12) \times (1.99267 \times 10^{-26}\text{kg})$$

$$= 1.660565 \times 10^{-27}\text{kg}$$

$$= 1.66 \times 10^{-27}\text{ kg}$$

Since mass of a proton (m_p) is 1.6723×10^{-27}kg, and mass of a neutron (m_n) is 1.6747×10^{-27}kg, we can express these in terms of u :

$$m_p = \frac{1.6723 \times 10^{-27}}{1.6606 \times 10^{-27}}\text{ u} = 1.00727\text{ u}$$

and $$m_n = \frac{1.6747\times10^{-27}}{1.6606\times10^{-27}}\text{ u} = 1.00865\text{ u}$$

Can you now express the mass of an electron ($m_e = 9.1 \times 10^{-31}$ kg) in terms of u? Since we will use nuclear masses in u, it is quite useful to know its energy – equivalent. To do so, we use *Einstein's mass-energy equivalence* relation, viz

$$\text{Energy} = \text{mass} \times c^2$$

where c is velocity of light in vacuum. Thus

$$1\text{u} = (1.66 \times 10^{-27}\text{kg})\ (2.9979\times10^{8}\text{ms}^{-1})^2$$

$$= 14.92\times10^{-11}\text{J}$$

$$= \frac{14.92\times10^{-11}}{1.60\times10^{-13}}\text{MeV}$$

$$= 931.3\text{MeV}$$

Note that joule (J) is too big a unit for use in nuclear physics. That is why we have expressed u in MeV (million electron volts). 1MeV is the energy gained by an electron when accelerated through a potential difference of one million volts. It is equal to 1.6×10^{-13}J.

26.1.5 Mass Defect and Binding Energy

The mass of the nucleus of an atom of any element is always found to be less than the sum of the masses of its constitutent nucleons. This difference in mass is called ***mass-defect***. For example, the nucleus of deuterium isotope of hydrogen has one proton and one neutron. The measured masses of these particles are 1.6723×10^{-27}kg and 1.6747×10^{-27} kg, respectively. It means that total mass of a proton and a neutron is 3.34709×10^{-27} kg. But the mass of deuterium nucleus is 3.34313×10^{-27}kg. It means that the measured mass of deuterium nucleus is 3.96242×10^{-30}kg less than the measured masses of a proton plus a neutron. So we say that mass defect in the case of deuterium is 3.96242×10^{-30} kg. Let us denote it by Δm. Mathematically, for an atom denoted by ${}^{A}_{Z}\text{X}$, we can write

$$\text{Sum of the masses of the nucleons} = Zm_p + (A–Z)m_n$$

∴ $$\Delta m = [Z.m_p + (A–Z)m_n] – M \qquad (26.1)$$

where M is actual mass of nucleius.

Energy equivalent of mass defect is obtained by using mass-energy equivalence relation:

$$\text{BE} = \Delta m\, c^2 \text{ joules} \qquad (26.2)$$

For Deuterium

$$\text{BE} = (3.96242\times10–30\text{kg}) \times (2.998 \times 10^8\text{ms}^{-1})^2$$

$$= 35.164 \times10^{-14}\text{kg m}^2\text{ s}^{-2}$$

$$= 3.5164 \times10^{-13}\text{J}$$

$$= 2.223 \times 10^6\text{eV}$$

since 1eV = 1.602×10^{-19}J.

This means that we have to supply atleast 2.223MeV energy to free the constituent nucleons – proton and neutron – of deuterium nucleus. You can generalise this result to say that *mass defect appears as energy which binds the nucleons together*. This is essentially used up in doing work against the forces which make the nucleons to cling.

Binding Energy per nucleon, B= $\Delta m\, c^2/A$

or
$$\text{B} = \frac{[Zm_\text{p} + (A - Z)m_\text{n} - M\,]\,c^2}{A} \qquad (26.3)$$

For ${}^{12}_{6}\text{C}$, Z = 6 and A = 12. Therefore (A–Z) = 12 – 6 = 6. Also M = 12 u; (1u = 931.3MeV)

Therefore,

$$B = \frac{[6m_\text{p} + 6m_\text{n} - 12]\times 931.3}{12} \quad \text{MeV}$$

$$= 7.41\text{MeV}$$

where we have used m_p = 1.00727 u and m_n = 1.00865 u.

It suggests that on breaking the nucleus of carbon atom, nearly 90MeV energy will be released, which can be used for various purposes. This is obtained in nuclear fission of a heavy atom like ${}^{238}_{92}\text{U}$. You will learn about it in the next lesson. This is also the source of energy in an atom bomb.

The value of B is found to increase to about 8.8 MeV as we move from helium (A = 4) to iron (A = 56); thereafter it decreases gradually and drops to about 7.6 MeV for uranium (A = 238). Fig.26.2 shows the variation of binding energy per nucleons with mass number.

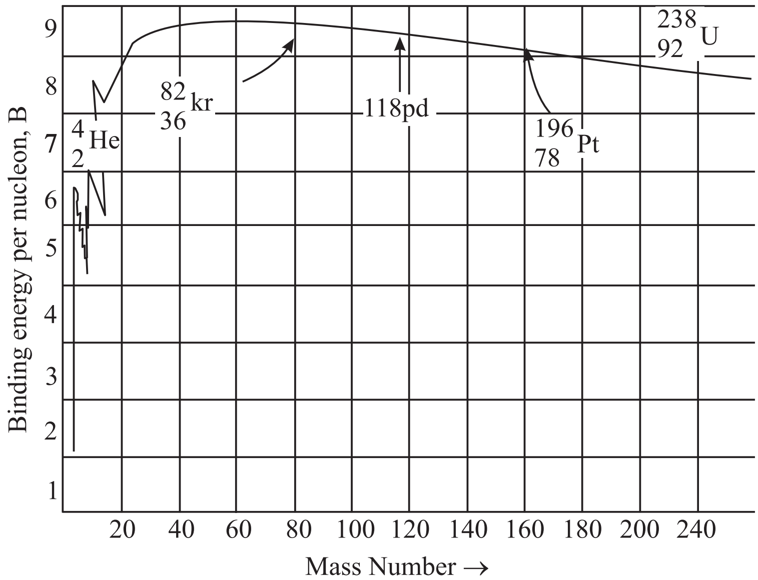


Fig. 26.2 : The variation of binding energy per nucleon with mass number

Note that binding energy curve shows sharp peaks for $^{4}_{2}$He, $^{12}_{6}$C, $^{16}_{8}$O and $^{20}_{9}$Ne. Moreover, *B* is small indicating that light nuclei with A < 20 are less stable. For example, the value of *B* for heavy hydrogen ($^{2}_{1}$He) is only 1.1 MeV per nucleon.

The subsidiary peaks occurring at $^{2}_{1}$He, $^{12}_{6}$C, $^{16}_{8}$O (even-even nuclei i.e. nuclei having even number of protons and even number of neutrons) indicate that these nuclei are more stable than their immediate neighbours.

The binding energy per nucleon curve is very useful in explaining the phenomena of nuclear fission and nucleon fusion.

Example 26.3 : Mass of a Boron ($^{10}_{5}$B) atom is 10.811 u. Calculate its mass in kg.

Solution : Since u $= 1.660565 \times 10^{-27}$kg,

$$10.811\text{u} = 10.811 \times 1.660565 \times 10^{-27}\text{kg}$$

$$= 17.952368 \times 10^{-27}\text{kg}$$

INTEXT QUESTIONS 26.2

1. The mass of the nucleus of $^{7}_{3}$Li atom is 6.01513 u. Calculate mass defect and binding energy per nucleon. Take, m_p = 1.00727 u.; m_n = 1.00865 u and 1 u = 931 MeV.

2. Calculate the radius of the nucleus of ${}^{8}_{4}\text{Be}$ atom.

   [Use $R = r_0 A^{1/3}$; $r_0 = 1.2 \times 10^{-15}$m]

26.2 HOW DO NUCLEONS CLING TOGETHER : NUCLEAR FORCE

Once physicists accepted the neutron-proton hypothesis of nucleus, an important question arose : How do nucleons cling together? In other words : What is the nature of force that binds nucleons? Since gravitation and electromagnetic interactions explain most of the observed facts, you may be tempted to identify one of these forces as the likely force. However, the extremely small size of the nucleus, where protons and neutrons are closely packed, suggests that forces should be strong, short range and attractive. These attractive forces can not have electrostatic origin because electrostatic forces between protons are repulsive. And if only these were operative, the nucleons would fly away, which is contrary to experience. Moreover, the forces between nucleons are responsible for the large binding energy per nucleon (nearly 8 MeV). Let us consider the gravitational force. No doubt, it is a force of attraction between every pair of nucleons. However, it is far too weak to account for the powerful attractive forces between nucleons. If the magnitude of nucleon-nucleon force is taken to be unity, the gravitational force would be of the order of 10^{-39}. We may, therefore, conclude that the purely attractive forces between nucleons are of a new type with no analogy whatsever with the forces known in the realm of classical physics. This new attractive force is called **nuclear force**.

26.2.1 Characteristic Properties

You may recall that the gravitational as well as electrostatic forces obey inverse square law. However, the nucleons are very densely packed and the nuclear force that holds the nucleons together in a nucleus must exist between the neighbouring nucleons. Therefore, nuclear force should be a short range force operating over very short distances (~10^{-15}m).

These nuclear forces must account for the attractive force between :

– a proton and a neutron;

– two protons; and

– two neutrons.

Since binding energy per nucleon, B is the same, irrespective of the mix of neutrons and protons in the nucleus, we are quite justified in considering the force between them as equivalent. That is, nuclear force is *charge independent*.

The nuclear force shows the property of *saturation,* which means that nucleons show only limited attraction. That is, each nucleon in a nucleus interacts with only neighbouring nucleons instead of all nucleons from one end of the nucleus to the other.

If nuclear forces had only attractive character, nucleons should have coalesced under their influence. But we all know that the average separation between nucleons is constant, resulting in a nuclear volume proportional to the total number of nucleons. The possible explanation is that nuclear forces exhibit attractive character only so long as nucleons are separated through a certain critical distance. For distances less than this critical value, the character of nuclear forces changes abruptly; attraction should change to repulsion. (You should not confuse this repulsion with electrostatic repulsion.) These qualitative aspects of nuclear forces are shown in Fig. 26.3

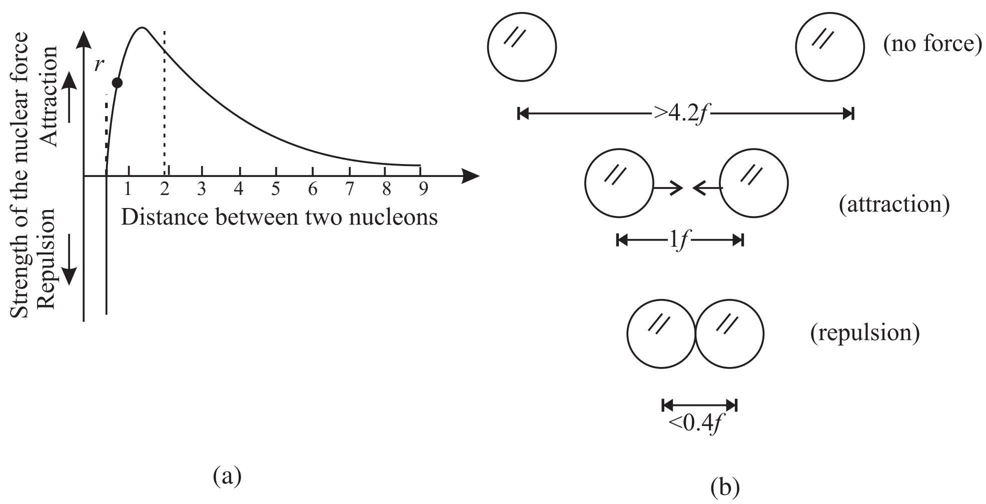


Fig. 26.3 : a) Typical variation of nuclear forces with distance, and **b)** effect of inter-nuclear distance on the force between nucleons.

26.3 RADIOACTIVITY

What is the age of our earth? How do geologists estimate the age of rocks and fossils found during excavations? What is radio-therapy which is used to treat malignant cells? The answers to all these interesting and useful questions are inherent in the study of radioactivity; a natural phenomenon in which atoms emit radiations to attain stability. Though it was discovered by chance, it opened flood gates for new physics. It finds wide use in industry, agriculture and medical care. Let us learn about it now.

26.3.1 Discovery

The story of discovery of radioactivity is very interesting. In 1896, French physicist A.H. Becquerel was working on the phenomenon of fluorescence (in which some

Notes

substances emit visible light when they are exposed to ultra-violet radiations). In one of the drawers of his desk, he had kept a collection of various minerals, besides several unopened boxes of photographic plates. Somehow, the collection of minerals remained untouched for a considerable period of time. One day Becquerel used one of the boxes of photographic plates to photograph something. When he developed the plates, he was disappointed to find that they were badly fogged as if previously exposed to light. He tried the other boxes of photographic plates and found them also in the same poor condition. He could not understand as to why plates were fogged because all the boxes were sealed and the plates inside were wrapped with thick black paper.

Becquerel was puzzled and investigated the situation further. He found that uranium placed in his drawer had done the damage and concluded that there must be some new type of penetrating radiation originating from the uranium salt. This radiation was named *Becquerel rays* and the phenomenon of emission of this radiation was named *radioactivity*. The elements exhibiting this phenomenon were called *radioactive elements*.

Soon after this discovery, and based on an exhaustive study, Madame Marie Curie alongwith her husband Pierre Curie, isolated an element from uranium ore by a painstaking method known as chemical fractionating. This new element, which was a million times richer in the mysterious rays than uranium, was given the name radium. Another radioactive element discovered by Madam Curie was named polonium in honour of her native country-Poland.

26.3.2 Nature of Radiations

In 1899, Lord Rutherford, a British physicist, analysed the Becquerel rays emitted by radioactive elements. He established the existence of two distinct components : α-particles and β-rays. The existence of third radiation – gamma rays – was established by P. Villard.

We know that nuclei of all atoms contain positively charged protons, which repel each other strongly due to electrostatic repulsion. To overcome this repulsion, neutrons in the nuclei act as glue. But in case of heavier nuclei, this electrostatic repulsion is so strong that even the addition of neutrons is not able to keep the nuclei stable. To achieve stability, such nuclei disintegrate spontaneously by emitting α and β particles along with γ-rays as shown in Fig 26.4. So, we can say that in natural radioactivity, α, β and γ-rays are emitted.

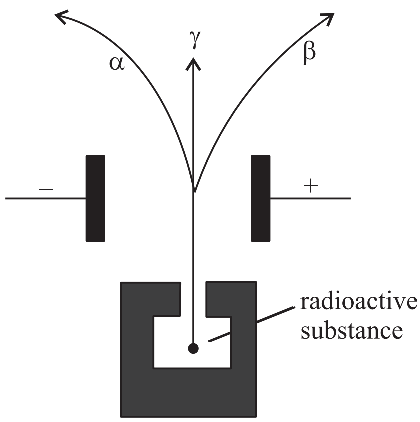


Fig. 26.4 : Emission of α, β and γ radiations

The emitted radiation is called the **radioactive radiation** and the process of disintegration (break-up) of atomic nuclei (by emitting α, β and γ-rays) is called **radioactive decay**. Sometimes, the break-up can be induced by bombarding stable nuclei with other light particles (like neutron and protons). It is then called ***artificial radio-activity.***

The characteristic features of this phenomenon are that it is spontaneous and in the case of α or β emission, a new nucleus belonging to a new element is formed. That is, one element gets converted into another element. This is thus a nuclear disintegration phenomenon and suggests the posibility of mutation of new nuclei. Let us first study the characteristic properties of α, β, and γ radiations.

(i) **α-particles**

Alpha particles are helium nuclei (^4_2He) and consist of two protons and two neutrons. Detailed studies of these particles revealed the following properties :

- Being charged particles, they get deflected in electric and magnetic fields.
- They produce fluorescence in substances like zinc sulphide and barium platino cyanide, affect a photographic plate, can induce radioactivity in certain elements and produce nuclear reactions.
- They have great ionizing power. A single particle in its journey through a gas can ionize thousands of gas atoms before being absorbed.
- They have little penetration power through solid substances, and get scattered by thin foils of metals. They can be stopped by 0.02 mm thick aluminum sheet.
- The energies of α particles emitted from a radioactive substance is a characteristic of the emitting nucleus. This corresponds to a variation in their velocity from $1.4 \times 10^7 \text{m s}^{-1}$ to $2.05 \times 10^7 \text{m s}^{-1}$.

(ii) **β-particles**

β-Particles can be both positively and negatively charged. They originate in the nucleus in the process of conversion of a neutron into a proton, and vice versa. Further studies of β-particles have revealed the following properties.

- Being charged particles, they get deflected by electric and magnetic fields.
- They produce fluorescence in materials like zinc-sulphide and barium plationcynide; and affect photographic plates.
- They can ionize gas atoms but to a much smaller extent than the α-particles.
- Negatively charged β-particles can pass through a few mm of aluminium sheets. They are about 100 times more penetrating than α-particles.

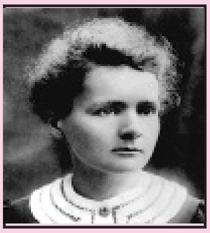
- Average energies of negative β-particles vary between 2 MeV and 3MeV. Due to their small mass, their velocities vary in range from 0.33*c* to 0.988*c*, where *c* is velocity of light.

(iii) **γ-rays**

γ-rays are electromagnetic waves of high frequency, and as such highly energetic. They are characterized with the following properties :

- They do not get deflected by electric or magnetic fields. They travel with velocity of light in free space.
- Their penetration power is more than that of α and β-particles; γ-rays can penetrate through several centimeters of iron and lead sheets.
- They have ionizing power that is smaller compared to that of α and β-particles.
- They can produce fluorescence in materials and affect a photographic plate.
- They knock out electrons from the metal surfaces on which they fall and heat up the surface. Hard γ-rays (i.e. high energy γ-rays) are used in radio therapy of malignant cells.

Marie Curie
(1867–1934)



Marie Curie shared the 1903 Nobel prize in physics with A. Henri Becquerel and her husband Pierre Curie for her studies in the field of radioactivity. She was the first person in the world to receive two Nobel prizes; the other Nobel prize she received was in chemistry in 1911. Later her daughter Joliot also won the Nobel prize in chemistry for her discovery of artificial radioactivity.

26.3.3 Radioactive Decay

In any radioactive decay, spontaneous emission consists of either a single α-particle or a β-particle. The emission of an α-particle from a radioactive nucleus (called *parent nucleus*) changes it into a new nucleus (new element is called *daughter nucleus*) with its atomic number decreased by two and its mass number decreased by four. Similarly, emission of a β particle changes the parent nucleus into a daughter nucleus with its atomic number increased by unity (if it is β^- emission) but its mass number remains unchanged. The emission of γ-rays does not change the atomic number or the mass number of the parent nucleus and hence no new nucleus is formed.

Note that in any nuclear disintegration, the charge number (*Z*) and the mass number (*A*) are always conserved. Thus for any radioactive nucleus, denoted by *X*, the nuclear transformations may be written as :

$$_Z^AX \xrightarrow{\alpha-\text{decay}} {}_2^4He + {}_{Z-2}^{A-4}Y$$
$$(\alpha\text{-particle})$$

$$_Z^AX \xrightarrow{\beta-\text{decay}} {}_{-1}^{0}e + {}_{Z-1}^{A}Y$$
$$(\beta\text{-particle})$$

$$({}_Z^AX)^* \xrightarrow{\gamma-\text{decay}} {}_Z^AX + \gamma$$

The asterisk over the symbol of element implies that it is in an excited state.

26.3.4 Law of Radioactive Decay

We now know that if we have a given amount of radioisotope, it will gradually decrease with time due to disintegrations. The law describing radioactive decay is very simple. The rate of radioactive disintegration is independent of external factors such temperature, pressure etc. and depends only on the law of chance. It states that ***the number of radioactive atoms disintegrating per second is proportional to the number of radioactive atoms present at that instant of time.*** This is called ***law of radioactive decay.***

Let N_0 be the number of radioactive atoms, at $t = 0$, and $N(t)$ be the number of radioactive atoms at time t. If dN denotes the number of atoms that decay in time dt, then $(N - dN)$ signifies the number of radioactive atoms at time $(t + dt)$. Hence, rate of decay

$$\frac{dN(t)}{dt} \propto N,$$

or

$$\frac{dN(t)}{dt} = -\lambda N(t) \qquad (26.4)$$

where λ denotes decay constant, which is characteristic of the radioactive substance undergoing decay. The negative sign signifies that the number of nuclei decreases with time. This relation can be rearranged as

$$\lambda = -\frac{1}{N(t)}\frac{dN(t)}{dt} \qquad (26.5)$$

Thus, decay constant (λ) may be defined as the ratio of the instantaneous rate of disintegration to the number of radioactive atoms present at that instant.

The law of decay is sometimes also expressed in exponential form and is also called the *law of exponential decay.* To obtain the exponential form, we integrate Eq. (26.4) with respect to time :

$$N(t) = N_0\exp(-\lambda t) \qquad (26.6)$$

The most important conclusion from this law is that N will become zero only when $t = \infty$.

We rewrite Eqn. (26.4) as

$$\frac{dN(t)}{N(t)} = -\lambda\, dt$$

On integration, we get

$\ln N(t) = -\lambda t + k$.

At $t = 0$, $N(t) = N_0$

$\therefore\ k = \ln N_0$

Hence

$\ln N(t) - \ln N_0 = -\lambda t$

or $\ln\left(\frac{N(t)}{N_0}\right) = -\lambda t$

On taking antilog, we obtain the required result:

$N(t) = N_0 \exp(-\lambda t)$

Notes

Thus, no radioactive element will disappear completely even after a very long time.

The radioactive decay law clearly shows that even if the number of atoms N_0 for different radioactive elements is same initially, at a later time they will have different values of $N(t)$ due to different values of their decay constants (λ). They will thus show different rates of disintegration. This is determined by their half-life ($T_{1/2}$) and average lives (T_a).

Units of Disintegration

The decay constant is measured in units of per second. The ***activity*** of a radioactive substance at any instant of time is measured by its rate of disintegration. Its SI unit has been named becquerel :

1 becquerel = 1 disintegration per second.

Another unit of the decay constant is ***curie.***

1 curie = 3.7×10^{10} disintegrations per second.

which is the rate of disintegration of radium (Ra) measured per second per gram.

Yet another unit is 'rutherford' (rd) :

1 rd = 10^6 disintegrations per second.

26.3.5 Half Life ($T_{1/2}$)

The half life ($T_{1/2}$) of any radioactive element is defined as the time in which the number of parent radioactive atoms decreases to half of the initial number.

By definition, at $t = T_{1/2}$, $N = N_0/2$. Therefore, using Eqn. (26.6), we can write

$$N_0/2 = N_0 \exp(-\lambda T_{1/2})$$

or $\lambda T_{1/2} = \log_e 2$

or
$$T_{1/2} = \frac{\log_e 2}{\lambda}$$
$$= \frac{2.303 \times \log_{10} 2}{\lambda}$$
$$= \frac{2.303 \times 0.3010}{\lambda}$$
$$= \frac{0.693}{\lambda}$$

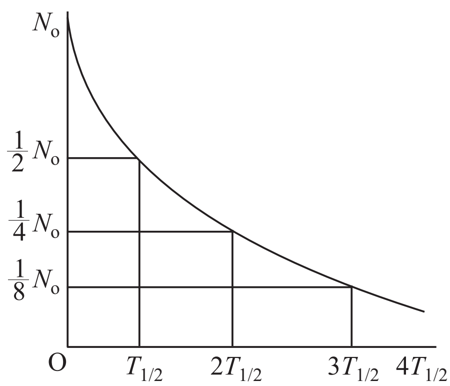


Fig. 26.5 : Radioactive decay curve

Thus, half-life of any radioactive substance is inversely proportional to its decay constant and is a characteristic property of the radioactive nucleus. The half-life of ${}^{14}_{6}C$ (radioactive carbon) is 5730 years. This means that one gram of ${}^{14}_{6}C$ will be reduced to 0.5 g in 5730 years. This number will be further reduced to $\frac{0.5}{2} = 0.25$ g in another 5730 years. i.e. in a total period of 11460 years. Refer to Fig. 26.5 to see how a radioactive sample decays with time.

Example 26.4 : An animal fossil obtained in the Mohanjodaro – excavation shows an activity of 9 decays per minute per gram of carbon. Estimate the age of the Indus Vally Civilisation. Given the activity of ^{14}C in a living specimen of similar animal is 15 decays per minute per gram, and half life of ^{14}C is 5730 years.

Solution : ^{14}C is radioactive isotope of carbon. It remains in fixed percentage in the living species. However, on death, the percentage of ^{14}C starts decreasing due to radioactive decay. Using radioactive decay law, we can write

$$N(t) = N_0 \exp(-\lambda t)$$

so that

$$N/N_0 = \exp(-\lambda t)$$

or

$$9/15 = \exp(-\lambda t)$$

or

$$\log_e (9/15) = -\lambda t$$

or

$$\text{loge} \left(\frac{15}{9}\right) = \lambda t$$

which gives

$$t = 1/\lambda \, [\log_e (15/9)]$$

Here $T_{1/2} = 0.693/\lambda = 5730$ years. Therefore,

$$t = 2.303 \times (5730/0.693) \, [\log_{10}15 - \log_{10}9]$$

Hence

$$t = 4224.47 \text{ years.}$$

Thus, the specimen containing carbon –14 existed 4224.47 years ago. Hence the estimated age of Indus valley civilsation is 4225 years.

INTEXT QUESTIONS 26.3

1. How can you say that radioactivity is a nuclear disintegration phenomenon?
2. Compare the ionizing and penetration powers of α, β and γ - radiations.

Notes

3. Apply the law of conservation of charge and mass numbers to determine the values of *a* and *b* in the following decay - equations :

   (i) ${}^{A}_{Z}X = {}^{4}_{2}He + {}^{b}_{a}Y + \gamma$

   (ii) ${}^{4}_{Z}X = {}^{0}_{-1}e + {}^{b}_{a}Y + \gamma$

4. The half-life of a radioactive substance is 5 years. In how much time, 10g of this substance will reduce to 2.5g?

Applications of Radioactivity

Radioactivity finds many applications in our every day life. Some of these are given below.

(i) ***In medicine :*** In the treatment of cancer (radiotherapy), a radio-active cobalt source which emits x-rays is used to destroy cancerous cells. The decay of a single radioactive atom can be registered by an instrument placed at a remote location outside a container wall. This high sensitivity is utilized in *tracer technique* as an important tool in medical diagnostics, like the detection of ulcer in any part of the body. A few radioactive atoms of some harmless element $\left({}^{24}_{11}Na\right)$ are injected into the body of a patient. Their movement can then be recorded. The affected part absorbs the radioactive atoms whose flow is, therefore, stopped and the diseased part of the body is easily located.

(ii) ***In agriculture :*** By exposing the seeds to controlled γ radiation, we are able to improve the quality and yield of crops, fruits and vegetables. Radiating these before their storage helps in saving from decay.

(iii) ***In geology :*** In estimating the age of old fossils. The normal activity of living carbon containing matter is found to be about 15 decays per minute for every gram of carbon. This activity arises from the small proportion of radioactive carbon –14 present in the atmosphere with the ordinary carbon –12. This isotope (${}^{14}C$) is taken by plants from the atmosphere and is present in animals that eat plants. Thus, about one part in 10^8 radioactive carbon is present in all living beings (all animals and plants). When the organism is dead, its interaction with the atmosphere (i.e. absorption, which maintains the above equilibrium) ceases and its activity begins to fall. From this, the age of the specimen can be approximately estimated. This is called ***carbon-dating*** and is the principle of determining the age of old fossils by archeologists.

The same technique has been used in estimating the age of earth from the measurements of relative amounts of ^{238}U and ^{206}Pb in geological specimens containing uranium ore. Assume that the specimen of ore contained only uranium and no lead at the time of birth of the earth. With the passage of time, uranium decayed into lead. The amount of lead present in any specimen will therefore indicate its age. The present age of the earth, using this method, has been estimated to be about 4 billion years.

(iv) ***In industry :*** γ-radiations are used to find the flaws (or imperfections) in the inner structure of heavy machinery. For example, if there is an air bubble inside, the penetration of γ-rays will be more at that point.

WHAT YOU HAVE LEARNT

- The nucleus in an atom contains positively charged protons and uncharged neutrons.
- The number of protons inside the nucleus of an atom of any element gives the atomic number of the element.
- The sum of the number of protons and neutrons in the nucleus of an atom is called its mass number.
- The atoms having same atomic number but different mass numbers are called isotopes.
- The atoms with same mass number but different atomic numbers are called isobars.
- The atoms with same number of neutrons are called isotones.
- The nucleons inside the nucleus of every atom are bound together by strong attractive nuclear forces which are short-range and charge-independent.
- The mass of a nucleus is found to be less than the sum of the masses of its nucleons. This difference in mass is called mass-defect. It is a measure of the binding energy.
- The size (volume) of the nucleus depends on its mass number.
- The spontaneous emission of α-particle or β-particle followed by γ-emission from any nucleus is called radioactivity.

- The α-particles have been identified as helium nuclei, while β-particles have been identified as fast moving electrons. The γ-rays are electromagnetic waves of extremely short wavelength.
- According to the law of radioactive decay, the number of radioactive atoms disintegrating per second is proportional to the number of radioactive atoms present at that instant.
- The half life of a radioactive substance is the time during which the number of radioactive atoms reduce to half of its original number.
- The law of exponential decay is $N(t) = N_0 \exp(-\lambda t)$.

TERMINAL EXERCISE

1. When does a radioactive sample disintegrate?
2. Differentiate between isotopes and isobars.
3. Explain the characteristics of binding energy per nucleon versus mass number curve.
4. What is the nature of nuclear force? Give its characteristics.
5. Explain how decay constant is related to half-life of a radioactive substance.
6. Define the following terms:

   (i) Atomic number; (ii) Mass number; (iii) Mass defect;

   (iv) Binding energy of nucleons; (v) Half-life; (vi) Average life;

   (vii) Decay constant.
7. State the law of radioactive decay.
8. What is carbon dating? What is its importance?
9. Calculate the number of neutrons, protons and electrons in the following atoms.

   (i) ${}^{23}_{11}Na$; (ii) ${}^{2}_{1}H$; (iii) ${}^{238}_{92}U$; (iv) ${}^{35}_{17}Cl$;
10. Calculate the mass defect and binding energy of nucleons for the following nuclei.

    (i) ${}^{4}_{2}He$; (ii) ${}^{7}_{3}Li$; (iii) ${}^{14}_{7}N$;

Given, 1 u = 1.660566×10^{-27}kg = 931 MeV, Mass of a proton = 1.007276 u. Mass of a neutron = 1.008665 u, Mass of $_2He^4$ atom = 4.00260 u, Mass of ^7_3Li atom = 7.01601 u, Mass of $^{14}_7N$ atom = 14.00307 u.

11. Using the present day abundance of the two main uranium isotopes and assuming that the abundance ratio could never have been greater than unity, estimate the maximum possible age of the earth's crust. Given that the present day ratio of ^{238}U and ^{235}U is 137.8 : 1; Half life of ^{238}U is = 4.5×10^9 year; and that of ^{235}U is 7.13×10^8 years.

12. If the activity of a redioactive sample drops to $\frac{1}{16}$th of its initial value in 1 hour and 20 minutes, Calculate the half-life.

ANSWERS TO INTEXT QUESTIONS

26.1

1.

Isotopes	Isobars	Isotones
$^{12}_6C$ and $^{14}_6C$	$^{76}_{32}Ge$ & $^{76}_{34}Se$	^2_1H & ^3_2He
^1_1H and ^2_1H & ^3_1H	$^{40}_{18}A$ & $^{40}_{20}Ca$	$^{14}_6C$ & $^{18}_8O$
$^{16}_8O$ & $^{18}_8O$	$^{76}_{32}Ge$ & $^{76}_{34}Se$	$^{23}_{11}Na$ & $^{24}_{12}Mg$
$^{35}_{17}Cl$ & $^{37}_{17}Cl$	^3_1H & ^3_2He	$^{27}_{13}Al$ & $^{28}_{14}Si$
$^{206}_{82}Pb$ & $^{207}_{82}Pb$	^7_3Li & ^7_4Be	$^{27}_{13}Al$ & $^{28}_{14}Si$
$^{238}_{92}U$ & $^{239}_{92}U$		

2. (i) heavier; (ii) mass; (iii) nucleons; (iv) 14; (v) 14 (vi) atomic.

3. Atomic number.

26.2

1. Δm = 1.041358 u; 969.5 MeV. 2. 2.4×10^{-15}m.

26.3

1. Nuclear disintegration usually involves α or β emission which results in change of atomic and mass numbers of the parent element. With the emission of α

and β particles, the heavier nuclei shed some of their mass resulting in comparatively lighter nuclei. Hence, it is a nuclear disintegration phenomenon.

2. Ionizing power of

   $\alpha > \beta > \gamma$

   Penetration power of

   $\alpha < \beta < \gamma$

3. i) a = Z –2 and b = A – 4

   ii) a = Z + 1 and b = A.

4. Two half life times are required – one for reduction from 10 to 5 grams and the other from 5 to 2.5 grams, i.e.. 10 years.

Answers to Problems in Terminal Exercise

9. (i) 12, 11, 11 (ii) 1, 1, 1 (iii) 146, 92, 921 (iv) 18, 17, 17

10. (i) 0.034, 28MeV (ii) 0.044, 37.86 MeV (iii) 0.10854, 101MeV

11. 6×10^9 years

12. 20 min

312en27

27

NUCLEAR FISSION AND FUSION

We all know that the sun supports life on the earth by continuously providing energy. It has been doing so for the last several billion years and will continue to do so for billions of years to come. What is the source of this huge amount of energy emitted by the sun? This question fascinated human mind always. But now we reliably know that the energy in the core of sun is produced by fusion of hydrogen nucli into helium at very high temperatures. This is also true of other stars. Imitation of these conditions in a fusion reactor is being highlighted as the ultimate source of all our energy requirements in coming years.

Similarly, you most have read about energy security and the role of nuclear energy to produce electricity in our nuclear reactors at Tarapore, Kota, Kaiga, Narora, Kalpakkam and Kakrapara. Similarly, you may have read in newspapers that on August 6, 1945, an atom bomb dropped over Hiroshima, a large city of Japan, destroyed the entire city almost completely in a span of a few seconds and lacs of lives were lost. It released an energy equivalent to that released by the explosion of a 20,000 ton TNT (tri-nitro toluene) bomb and was completely new in human history. Since then, more powerful (atomic, hydrogen and neutron) bombs have been made whose destructive power is equivalent to several Mega tons of TNT. The super powers are said to have stockpiled a large number of such bombs. The destructive power of their stock is so enormous that they can destroy the entire earth several times over. The physical process responsible for such colossal amount of energy is nuclear fission. You will now learn about these processes.

OBJECTIVES

After studying this lesson, you should be able to

- *state conservation laws for nuclear reactions;*

- *explain the terms nuclear chain reaction, controlled and uncontrolled fission chain reactions;*
- *describe working of a nuclear reactor; and*
- *explain the mechanism of production of energy in stars.*

27.1 CHEMICAL AND NUCLEAR REACTIONS

27.1.1 Chemical Reaction

We know that all substances are made up of atoms. In lesson 26, you learnt that electrons in the outermost orbit govern the chemical properties of an element. That is, atoms combine with other atoms or molecules (a group of atoms) and rearrange their valence electrons This is accompanied by reduction in their potential energy.

The formation of a new compound molecule due to rearrangement of valence electrons in interacting atoms and molecules with the release or absorption of energy is called a chemical reaction. In this process, the nucleus is not affected at all. Even the electrons in the inner orbits remain unaffected.

An example of a chemical reaction is the interaction of carbon atoms with oxygen molecules to produce carbon dioxide :

$$C + O_2 \rightarrow CO_2 + 4.08 \text{ eV} \qquad (27.1)$$

In this chemical reaction, 4.08* eV *energy is released for each reacting* carbon *atom. It is called the binding energy (B.E) of* CO_2 *molecule. Reactions which result in release of energy are said to be *exothermic*. Chemical reactions which require energy to be supplied to be initiated are *endothermic*. For example, if 4.08 eV of energy is given to a CO_2 molecule under suitable conditions, it will break up into its constituents:

$$CO_2 + 4.08 \text{ eV} \rightarrow C + O_2 \qquad (27.2)$$

As shown in Eq. (27.1), 4.08 eV energy leaves the system to form CO_2 gas. Therefore, the mass of CO_2 molecule will be less than the total mass of C and O_2 by a mass equivalent of 4.08 eV. The loss of mass Δm can be calculated using the relation $E = mc^2$:

$$\Delta m = \frac{4.08 \times 1.602 \times 10^{-19}}{9 \times 10^{16}} = 7.26 \times 10^{-36} \text{ kg} \qquad (27.3)$$

Such a small change in mass cannot be detected and we say that the mass is conserved in chemical reactions, though slight change of mass does occur.

Notes

The important points to be noted in chemical reactions are

- Energies of the order of 10 eV are involved.
- Change of mass is of the order of 10^{-35} kg, which is extremely small and we say that the mass is conserved.
- The total number of atoms of each type on the right hand side of the chemical equation is always equal to the total number of atoms of each type on the left hand side.

27.1.2 Nuclear Reactions

In nuclear reactions, the nuclei, not electrons, of the reactants interact with each other. They result in the formation of new elements. This process is also called transmutation of nuclei. From the previous lesson, you may recall that in nuclear reactions energies of the order of MeV are involved.

We know that the entire positive charge of an atom is concentrated in its nucleus, whose size is of the order of 10^{-15} m. The nucleus is surrounded by electrons revolving in certain specified orbits. These create a strong electrostatic potential barrier (also called the Coulomb barrier) as shown in Fig. 27.1. The Coulomb barrier is about 3 MeV for carbon nuclei and 20 MeV for lead nuclei. It means that a charged projectile aimed at a nucleus will experience strong repulsion by the Coulomb barrier of the target nucleus. If the kinetic energy of projectile is not large enough to penetrate the barrier, it will come back without producing any nuclear reaction. For a proton to enter a carbon nucleus and produce transmutation, its energy should be more than 3MeV or so. It is because of the large amounts of energy involved in nuclear reactions that we do not observe these reactions in everyday life at ordinary temperatures and pressures.

The phenomenon of nuclear ***transmutation*** or nuclear reaction was discovered by Lord Rutherford in the year 1919. He bombarded nitrogen gas with high energy α-particles of energy 7.7 M eV obtained from a polonium source. He observed

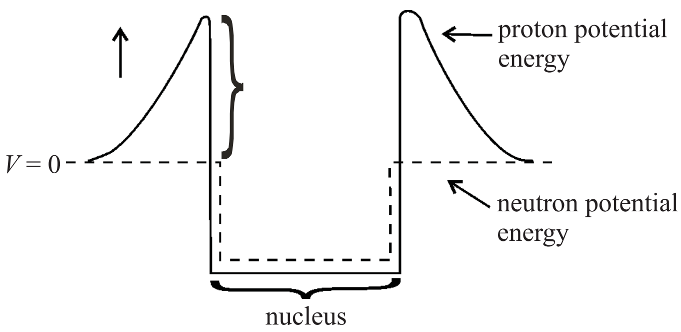


Fig. 27.1 : Proton and neutron potential energies near a nucleus

that nitrogen transformed into oxygen. This change was accomponied by high energy protons :

$$ {}^{4}_{2}\text{He} + {}^{14}_{17}\text{N} \rightarrow {}^{17}_{8}\text{O} + {}^{1}_{1}\text{H} \qquad (27.5)$$

The oxygen nuclei and protons carry away 6.5 MeV. Clearly this reaction can occur if 1.2 MeV energy is supplied from outside. Therefore, it is an endothermic nuclear reaction. When aluminium is bombarded by 7.7 M eV alpha particles from polonium, the following nuclear reaction takes place and 10.7 MeV energy is released:

$$ {}^{27}_{13}\text{Al} + {}^{4}_{2}\text{He} \rightarrow {}^{30}_{14}\text{Si} + {}^{1}_{1}\text{H} \qquad (27.6)$$

Here we see that more energy is released than the input energy; it is an exothermic reaction. Note that there is a gain of nearly 3 MeV energy per reaction, which is approximately 700,000 times the energy released in burning of one carbon atom. *But this reaction can't be used for production of energy because out of 125,000 incident alpha particles only one succeeds in producing the reaction.* Hence on the whole, there is much more energy spent than produced.

Nuclear reactions can also be produced by protons, deuterons, neutrons and other light nuclei. Of these, *neutrons are the best projectiles for producing nuclear reactions; being neutral particles, they do not experience Coulomb repulsion..* Thus even thermal neutrons (i.e. neutrons having energy 0.0253 eV) can penetrate the target nucleus and produce a nuclear reaction.

Some typical examples of nuclear reactions produced by protons, deutrons and neutrons are:

$$ {}^{6}_{3}\text{Li} + {}^{1}_{1}\text{H} \rightarrow {}^{3}_{2}\text{He} + {}^{4}_{2}\text{He} \qquad (27.7)$$

$$ {}^{10}_{5}\text{B} + {}^{2}_{1}\text{H} \rightarrow 3\,{}^{4}_{2}\text{He} \qquad (27.8)$$

$$ {}^{10}_{5}\text{B} + \text{n} \rightarrow {}^{7}_{3}\text{Li} + {}^{4}_{2}\text{He} \qquad (27.9)$$

Like chemical reactions, nuclear reactions also follow conservation laws.We state these now.

27.1.3 Conservation Laws for Nuclear Reactions

- *The sum of the mass numbers of the reactants is equal to the sum of mass numbers of the products. In Eqn.* (27.7), *mass number* 7 = 3 + 4 = 6 + 1 *is conserved.*

- *The sum of atomic numbers of the reactants is equal to the sum of atomic numbers of the products. In Eqn.* (27.7), *atomic number* 4 = 3 + 1 = 2 + 2 *is conserved.*
- *Nuclear reactions follow the law of conservation of energy. We know that mass is concentrated form of energy. Therefore the sum of input kinetic energy plus the mass of the reactants is equal to the output kinetic energy plus the mass of the products.*
- *Nuclear reactions follow the law of conservation of momentum, which results in distribution of kinetic energy among various product nuclei.*

Now, answer the following questions.

INTEXT QUESTIONS 27.1

1. Complete the following equations of nuclear reaction.

   (a) $^{19}_{9}F + ^{1}_{1}H \rightarrow ^{16}_{8}O + ?$

   (b) $^{27}_{13}Al + ^{1}_{0}n \rightarrow ? + ^{4}_{2}He$

   (c) $^{234}_{90}Th \rightarrow ^{234}_{91}Pa + ?$

   (d) $^{63}_{29}Cu + ^{2}_{1}D \rightarrow ^{64}_{30}Zn + ?$

2. Calculate the energy released in the nuclear reaction given below

   $^{10}_{5}B + ^{2}_{1}D \rightarrow 3^{4}_{2}He + Q$

   Given that $m(^{10}B)$ = 10.01294 u; $m\left(^{2}_{1}D\right)$ = 2.014103 u, and $m(^{4}_{2}He)$ = 4.002604 u.

3. $^{14}_{7}N$ nucleus, on bombarding with alpha particles, produces $^{17}_{8}O$. Write down the reaction equation and calculate the energy released.

   Given that: $m\left(^{14}_{7}N\right)$ = 14.003014 u; $m\left(^{17}_{8}O\right)$ = 16.999138 u; $m\left(^{4}_{2}He\right)$ = 4.002604 u; $m\left(^{1}_{1}H\right)$ = 1.007825 u and energy of α particle = 7.7MeV.

27.2 NUCLEAR FISSION

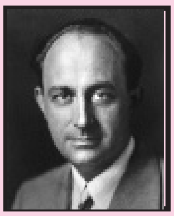
The story of discovery of fision is very fascinating. In the year 1938, Enrico Fermi, Otto Hahn and others irradiated uranium nuclei with slow neutrons to produce transuranic elements (having Z greater than 92), which do not occur in nature. When incident neutrons were captured by the uranium nuclei, the neutron-proton ratio increased. In reducing this ratio, it was expected that uranium would become β–active. That is a neutron would essentially behave as if it has changed into a proton resulting in the release of a β-particle and some energy according to the equation:

$$^{238}_{92}\text{U} + ^{1}_{0}\text{n} \rightarrow ^{239}_{93}\text{Np} + ^{0}_{-1}\text{e} + \text{Q} \qquad (27.10)$$

In this process, a new transuranic element having atomic number 93 was expected to be produced. In fact, Fermi and his co-researchers observed β–activities with half-lives different from any of the known values for heavy elements in the vicinity of uranium. From those observations, they concluded that transuranic elements had been produced. And to identify the element, they carried out chemical analysis but failed.

In the same year, Otto Hahn and Fritz Strausmann carried out a series of experiments and established that barium, an element of intermediate mass number, rather than a transuranic element, was one of the products of the reaction and it was accomponied by release of nearly 200 MeV of energy. This result – the product of slow neutron bombardment of uranium was barium – was completely unexpected and defied all knowledge of nuclear physics of that time. These findings were reported in *Nature* in Dec. 1938.

Initially, Lise Meitner and Otto Frisch explained these results on the basis of liquid drop model of nucleus and named this process *nuclear fission* using the analogy with biological cell division. Later on, Bohr and Wheeler calculated the amount of energy released in the process, confirming the physical basis of this model.



Enrico Fermi
(1901 – 1954)

Enrico Fermi, the Italy born physicist, was responsible for peaceful uses of nuclear energy for mankind. He demonstrated that nuclear transformations may occur in any element exposed to stream of neutrons. He achieved self-sustained nuclear fission chain reaction in 1942.

Fermi was only 25 years old when he formulated the Fermi–Dirac statistics, applicable to particles having half integral spin values (called fermions). At the time of his premature death, he was engrossed in theoretical studies of cosmic radiations.

27.2.1 Mechanism of Nuclear Fission

In the year 1939, Bohr and Wheeler developed the theory of fission using the analogy between nuclear forces and the forces which bind molecules in a liquid. They predicted that $^{235}_{92}U$ was more fissile than $^{235}_{92}U$. Refer to Fig. 27.2. If shows the schematics of nuclear fission of $^{235}_{92}U$ by thermal neutrons according to the equation.

$$^{235}_{92}U + ^{1}_{0}n \rightarrow ^{141}_{36}Ba + ^{92}_{36}Kr + 3^{1}_{0}n + Q \qquad (27.11)$$

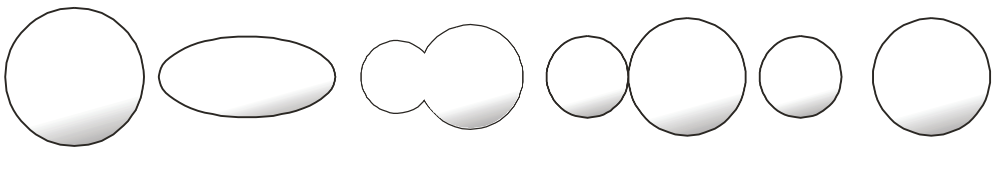


Fig. 27.2 : Nuclear-fission of a nucleus according to the liquid drop model

The emitted neutrons have energy of the order of a few MeV, and $Q \simeq 200MeV$.

Note that a fission event occurs within 10^{-17} s of neutron capture and fission neutrons are emitted within about 10^{-14} s of the event. Moreover, the fission fragments are of unequal mass; one being 1.5 to 2 times heavier than the other. Also, Eqn. (27.11) gives only one of the more than 40 different modes in which a $^{235}_{92}U$ nucleus can fission. It means that about 80 different nuclei of intermediate masses are produced in the fission of $^{235}_{92}U$. The heavier fragments lie in the mass range 125–150 with the a maximum around 140, whereas the lighter fragments lie in the range 80 – 110 with a maximum around 95. The number of neutrons emitted is either two or three and the average number of neutrons produced per fission of ^{235}U is 2.54

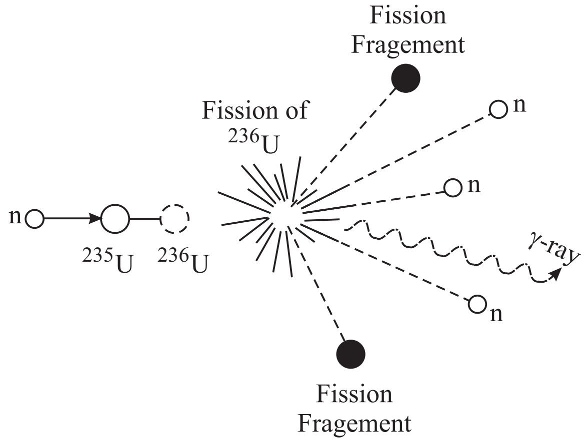


Fig. 27.3 : Nuclear fission

Bohr and Wheeler treated the nucleus as a charged spherically symmetric liquid drop in its equilibrium (lowest energy) state. According to them, when a nucleus captures a thermal neutron, the binding energy (BE) of this neutron, which is 6.8 MeV per atomic mass unit for ^{235}U, is released. This energy excites the nucleus and distorts its shape. While the force of surface tension tries to restore the original shape, the Coulomb force tends to distort it further. As a result, it oscillates between spherical and dumb bell shapes, as shown in Fig.27.2, depending on the energy of excitation. When the energy gained by the nucleus is large, the amplitude of these oscillatious pushes the nuclens into dumb bell shape. When the distance between the two charge centres exceeds a critical value, electrostatic repulsion between them overcomes nuclear surface tension and pushes the nucleus into two parts resulting in fission.

A substance like $^{235}_{92}U$ which undergoes fission by thermal neutrons is called a fissile material. Other fissile materials are $^{233}_{90}Th$, $^{233}_{92}U$ and $^{239}_{93}Pu$. You may note that all these nuclei have odd mass number and even atomic number.

We can estimate the amount of energy released in the fission of $^{235}_{92}U$ by calculating the mass defect as follows:

Table 27.1 Energy Generated in a Nuclear Reaction

Reactants	Mass	Products	Mass
^{235}U	235.0439 u	$^{141}_{56}Ba$	140.9139 u
^{1}n	1.008665 u	$^{92}_{36}Kr$	91.8973 u
		$3 \times Vn$	3.025995 u
Total mass	236.052565 u	**Total mass**	235.837195 u
Mass defect	0.21537u		
Energy released	$0.21537 \times 931 \simeq 200$ MeV		

27.2.2 Nuclear Chain Reaction

You have now learnt that when a neutron is captured by $^{235}_{92}U$, it splits into two fragments and 2-3 neutrons are emitted. These are capable of causing further fissions. This immediately presented the exciting possibility of maintaining a fission chain reaction in which each fission event removes one neutron and replaces that by more than two. When

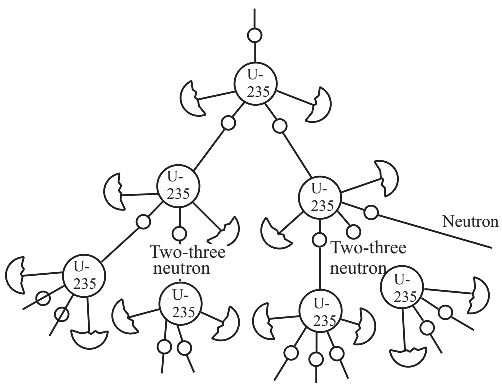


Fig. 27.4 : Nuclear Chain reaction

the rate of production of neutrons equals the rate of loss of neutrons, the reaction is said to be *self-sustained*. The device designed to maintain a self-sustained and controlled chain reaction is called a *nuclear reactor*.

Nuclear reactors are usually classified according to the purpose for which they are used. So a nuclear power reactor is used to produce electricity and a research reactor is used to produce radioisotopes for medical purposes, carrying out experiments for refinements or applied research. We also categorise nuclear reactors as fast and thermal, depending on the energy of neutrons causing fission. In India, we have thermal power reactors at Tarapore, Narora, Kota, Kaiga, etc. At Kalpakkam, we are developing a fast breeder research reactor.

You will now learn about a nuclear reactor in brief.

27.3 NUCLEAR REACTOR

Ever since the first nuclear reactor was constructed by Fermi and his co-workers at the university of Chicago USA, a large number of reactors have been built the world over primarily to meet demand for energy. Some countries generate as much as 70% of their total energy from nuclear reactors. In India, the contributions of nuclear energy is only about 2%, but efforts are on to increase this share. In absolute terms, we are generating about 20,000 MWe from nuclear reactors.

Nuclear reactors have huge complex structures and great care has to be exercised in designing them. The basic principle of a nuclear power plant is very simple and analogous to any power plant. The heat liberated in fission is used to produce steam at high pressure and high temperature by circulating a coolant, say water, around the fuel. (In a coal fired station, coal is burnt to produce steam. Since one fission event generates about 7×10^5 times more energy than that produced in burning one atom of carbon, we can cut down on emission of greenhouse gases substantially by switching over to nuclear energy. However, there are some complex social and political issues with global dimensions that will ultimately decide our ultimate nuclear energy options.)

The steam runs a turbine–generator system to produce electricity. (In research reactors, the heat is discharged into a river or sea. You many have heard about Bhahha Atomic Research Centre at Trombay, Mumbai or Indira Gandhi Atomic Research Centre at Kalpakkam. The heat generated by the research reactors at these centres is discharged into the Arabian sea and the Bay of Bengal, respectively.)

The general features of a reactor are illustrated in Fig. 27.5. All nuclear reactors consist of:

- A *reactor core,* where fission takes place resulting in release of energy. It has fuel rods (embedded in a modertor in a thermal reactors), and *control rods* to maintain the chain reaction at the desired level. *Coolant* is circulated to remove the heat generated in fission. Usually, heavy water or ordinary water are used as coolants and cadmium or boron are used for control rods.

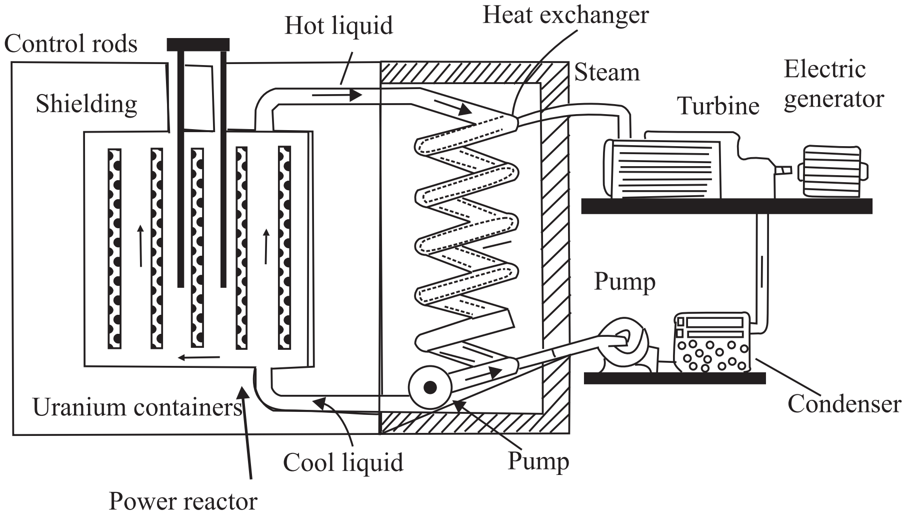


Fig. 27.5 : Schematic diagram of a nuclear reactor

- A *reflector is* put next to the core to stop neutron leakage from the core.
- The whole assembly is placed inside a vessel, called *pressure vessel*. Usually, a few inches thick stainless steel is used for this purpose.
- A thick *shield* is provided to protect the scientists and other personnel working around the reactor from radiations coming from the reactor core. It is usually in the form of a thick concrete wall.
- The entire structure is placed inside a *reactor building*. It is air tight and is maintained at a pressure slightly less than the atomospheric pressure so that no air leaks out of the building.

The heat generated inside the reactor core of a reactor due to fission is removed by circulating a *coolant*. The heated coolant is made to give up its heat to a secondary fluid, usually water in a heat exchanger. This generates steam, which is used to drive turbine-generator system to produce electricity in a power plant and discharged into a river/lake/sea in a research reactor.

INTEXT QUESTIONS 27.2

1. Why does a $^{238}_{92}U$ nucleus become β–active after absorbing a neutron?

2. Out of $^{238}_{92}U$, ^{141}Ba, ^{239}Pu, and $^{12}_{6}C$, which nucleus is fissile?

3. How much energy is released when $^{235}_{92}U$ undergoes nuclear fission?

27.4 NUCLEAR FUSION

You now know that uranium nucleus can be made to split into lighter nuclei resulting in release of huge amount of energy. You may now ask: Can we combine lighter nuclei to produce energy? To discover answer to this question, refer to the binding energy per nucleon (BE/A) curve (Fig.26.2). You will note that binding energy per nucleon increases as we go from hydrogen to helium . It means that helium is more stable than hydrogen. Consider the following reaction:

$$_1D^2 + {}_1D^2 \rightarrow {}_2He^4 + Q$$

You can easily calculate the B.E of reactants and products:

$$\text{Total B.E of reactants, } BE_1 = 2 \times 2.22 = 4.44 \text{ MeV}$$

$$\text{Total B.E of products, } BE_2 = 28.295 \text{ MeV}$$

$$Q = (BE_2 - BE_1) \simeq 24 \text{ MeV}$$

∴ Note that the energy released per nucleon in this reaction is 24/4=6 MeV, which is nearly seven times the energy released per nucleon (200/238 = 0.83 MeV) in a nuclear fission event.

The process in which two light nuclei combine to form a heavier nucleus is called nuclear fusion.

Table 27.2 : Binding Energy per nucleon (BE/A) of some light nuclei

Nucleon	*BE/A* (in MeV)
^{2}D	1.11
^{3}T	2.827
^{3}He	2.573
^{4}He	7.074
^{6}Li	5.332
^{7}Li	6.541

Fusion process presents itself as a more viable energy option. However, the process of fusion is more difficult to achieve than nuclear fission because both the deuterons are positively charged. When we try to bring them together to fuse into one nucleus, they repel each other very strongly and the reaction is ordinarily impossible.

To achieve this reaction, the deuterons have to be heated to nearly 10 million kelvin so that they acquire sufficient kinetic energy to overcome repulsion before they collide to fuse into helium nucleus. But the problems associated with maintaining such high temperatures continuously and containing the reactants together has not yet been solved fully. The controlled thermonuclear reaction necessary for harnessing this source of energy is however not far now.

Almost inexhaustible amount of deuterium (heavy hydrogen) is present in the ocean. Once we begin to harness this source, our energy problem should be solved for ever. We will get an endless supply of cheap electricity without any pollution. This is because one gram of deuterium (heavy hydrogen) yields about 100,000 kW h of energy.

27.4.1 Energy in the Sun and Stars

The stars like our sun are very massive objects. They have been continuously emitting tremendous amount of energy for the last billions of years.

Such a huge amount of energy cannot be obtained by burning conventional fuels like coal. Nuclear fission can also not be the source of this energy, because heavy elements do not exist in the sun in large quantity. The sun mainly consists of hydrogen and helium gases. Then you may like to know: What is the source of energy in the sun? This question has engaged human intellect for long. As a child, you must have gazed the sky when you learnt the rhyme: Twinkle twinkle litter star, How I wonder what you are!

You may know that the huge mass of the sun produces extremely strong gravitational field, which compresses its constituent gases by enormous pressure resulting in the rise of temperature to millions of kelvin at its centre. It has been estimated that the temperature at the centre of the sun is 20 million kelvin. At such high temperatures and pressures, gas molecules travel at high speeds and collide setting in thermonuclear reaction and resulting in the release of large amount of energy.

Bethe proposed that fusion of hydrogen into helium is responsible for the energy produced in stars:

$$4\,{}^{1}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + 2\,{}^{0}_{+1}\mathrm{e} + \mathrm{Q}$$

The overall result here is: four hydrogen nuclei fuse into a helium nucleus with the release of two positrons (electron-like microscope particles of the same mass but positive charge) and 26.8 MeV energy. ***The tremendous amount of energy released in a thermo-nuclear reaction is the source of energy in stars.*** The quantity of hydrogen in the sun is sufficient to keep it shining for nearly 8 billion years more.

INTEXT QUESTIONS 27.3

1. 200 MeV energy is released in fission of one ${}^{235}_{92}\mathrm{U}^{5}$ nucleus and 26.8 MeV energy is released in fusion of 4 protons. Which process releases more energy per unit mass?

2. Consider the following reactions:

   (*a*) ${}^{1}_{1}\mathrm{H} + {}^{7}_{3}\mathrm{Li} \rightarrow {}^{4}_{2}\mathrm{He} + Q$.

   (b) ${}^{2}_{1}\mathrm{H} + {}^{2}_{1}\mathrm{H} \rightarrow {}^{3}_{1}\mathrm{H} + {}^{1}_{1}\mathrm{H} + 4\mathrm{MeV}.$

   Calculate Q in the first reaction and mass of tritium in the second reaction.

   Given $m\left({}^{2}_{1}\mathrm{H}\right) = 2.014103\mathrm{u}$, $m\left({}^{4}_{2}\mathrm{H}\right) = 4.002604\mathrm{u}$, $m\left({}^{1}_{1}\mathrm{H}\right) = 1.007825\mathrm{u}$ and

   $m\left({}^{7}_{3}\mathrm{Li}\right) = 7.015982\mathrm{u}.$

27.5 NUCLEAR ENERGY

We need energy for all economic activities in life. The amount of energy consumed per capita is a measure of advancement of a nation. According to a recent UNESCO report (2007), we are consuming about 40% more than what mother earth can generate in the form of food, water and energy. In fact, the human society has been continuously striving for energy security and looking for newer sources of energy. Due to over use, conventional sources of energy are depleting very fast and may exhaust completely in the next one hundred years. The nuclear energy is perhaps an important option for meeting our future energy needs through peaceful applications. Let us discuss these now.

27.5.1 Peaceful Applications

The most important peaceful application of nuclear energy is in the *generation of electricity.* One of the main advantages of nuclear power plant is that the fuel is not required to be fed into it continuously like the gas or coal in a thermal power plant. Further, it does not *pollute the environment to the extent discharge of smoke or ash from fossil fuel/power plants do.* The fuel once loaded in a reactor runs for nearly 6 months at a stretch. Because of this nuclear power plants have been used to power huge ships and submarines.

However, spent fuel of a reactor is highly radioactive because a large number of radio-isotopes are present in it. India has developed its own facility to treat spent fuel and extract it from those *radio-isotopes which find uses in agriculture, medicine, industry and research.* To avoid the spread of radioactive radiations from the radioactive wastes, the radioactive wastes are generally embedded deep inside salt mines in heavy steel cases. Yet, it has evoked considerable controversy due to its destructive potential which was displayed on August 6, 1945, when an atom bomb was dropped on Hiroshima (Japan) killed hundred thousand people in a very short time. Subsequently, even more powerful hydrogen and nitrogen bombs have been developed. These can destroy this beautiful planet many times over.

Nuclear Power in India

The possibility of harnessing nuclear power for civil use was recognised by Dr H.J. Bhabha soon after India got independence. He outlined a three stage development plan for meeting country's nuclear power needs. These are :

- Employ pressurised Heavy Water Reaction (PHWR) fuelled by natural uranium to generate electricity and produce plutonium as a by-product.
- Set up fast breeder reactors burning the plutonium to breed U-233 from thorium.
- Develop the second stage and produce a surplus of fissile material.

Nuclear power has been produced in India through 14 small and one mid-sized nuclear power reactors in commercial operation, eight under construction and more planned. As of now, nuclear power contributes nearly 2×10^{10} kW h of electricity – 3% of total power capacity available.

Government policy is to have 20 GWe of nuclear capacity operating by 2020 and 25% nuclear contribution is foreseen by 2050.

INTEXT QUESTIONS 27.4

1. What type of reactors are used in India for power generation?
2. How much $^{235}_{92}U$ undergoes fission in an atomic bomb which releases energy equivalent to 20,000 tons of TNT. (Given that 1 g of TNT gives out 1000 calorie of heat).

27.5.2 Hazards of Nuclear Radiations and Safety Measures

The living and non-living things around us constitute our environment. In this environment, a delicate balance has existed for millions of years between the flora, fauna, acquatic and human life. This balance is now being threatened. One of the factors disturbing this balance is the ever increasing pollution in our environment. Out of the various types of pollutants present in our environment, the one which has very serious long term biological effects are the 'nuclear radiations'. Earlier these were present only because of natural sources like the radioactive minerals and cosmic rays, but now their presence is increasing day by day due to man-made sources. The major present day man-made sources of nuclear radiations are the nuclear tests, nuclear installations like the nuclear research facilities, nuclear reactors, and radio isotopes in treating diseases.

Nuclear radiations dissociate complex molecules of living tissues through ionisation and kill the cells. They induce cancerous growth, cause sterility, severe skin burns, and lower the body resistance against diseases. They disrupt the genetic process, mainly in the unborn child, and show their effects even upto five generations. Nuclear radiations affect us not only directly, but also indirectly by affecting the flora, fauna and the acquatic life around us. They kill vegetation, fishes and animals.

The damage caused by nuclear radiations depends on the exposed part of the body, as well as on the energy, intensity and the nature of the radiation. Different parts of human body show different sensitivities to radiation. The α-particles are, as a rule, quite harmful because of their high ionising power. The damaging effects of different radiations are generally compared in terms of their 'relative biological effectiveness', called the RBE factors. These factors for different particles/rays are given in Table 27.3.

Table 27.3: RBE factors of different radiations

Particles/rays	RBE factors
X-rays, γ -rays, β-particles	1
Thermal neutrons	2 to 5
Fast neutrons	10
α-particles, high energy ions of O, N, etc.	10 to 20

There is no control on natural sources of radiation. However, efforts can certainly be made to lower down radiation from man-made sources. Some of these are to:

- Avoid nuclear explosions.
- Minimise production of radio-isotopes.
- Extreme care should be exercised in the disposal of industrial wastes containing traces of radio-nuclides.
- Nuclear medicines and radiation therapy should be used only when absolutely necessary, and with well considered doses.

WHAT YOU HAVE LEARNT

- Valence electrons take part in chemical reactions and the energy involved in such reactions is of the order of 1eV.
- In a nuclear reaction, the atomic nuclei interact to form a new element.
- Energy involved in nuclear reaction is of the order of MeV.
- In a nuclear reaction, atomic number, mass number and charge are conserved.
- When a heavy nucleus like uranium is bombarded by slow neutrons, it splits into two fragment with release of 2-3 neutrons and 200 MeV energy. This process is known as nuclear fission.
- Substances that undergo fission are called fissile substances. ^{233}Th, ^{233}U, ^{235}U, and ^{239}Pu are fissile materials.
- Chain reaction occurs when more than one emitted neutron induce further fission for each primary fission.
- Nuclear reactor is a device to sustain controlled chain reaction.
- In nuclear fusion two light nuclei are fused into one.
- For producing nuclear fusion, the reacting nuclei must be heated to nearly 20 million kelvin to gain sufficient kinetic energy to overcome the Coulombian potential barrier.
- In stars energy is produced by nuclear fusion reaction.
- Amount of hydrogen consumed in the sun is nearly 400×10^6 ton per second.
- Radio-isotopes find diverse applications in agriculture, medicine and industry.

TERMINAL EXERCISE

1. How does a nuclear reaction differ from a chemical reaction?
2. What is the use of moderator and absorber in a fission reactor?

3. On the basis of B.E per nucleon versus mass number curve, explain nuclear fusion.

4. What is a nuclear reaction? State the conservation laws obeyed in nuclear reactions. Give threes examples of nuclear reactions.

5. What is nuclear fission? Give an example to illustrate your answer.

6. Calculate the mass of ^{235}U consumed to generate 100 mega watts of power for 30 days.

7. Heavy hydrogen undergoes the following fusion reaction

   $${}^{2}_{1}D + {}^{2}_{1}D \rightarrow {}^{4}_{2}He + 24\text{ MeV}$$

   Calculate the amount of heavy hydrogen used in producing the same energy as above. Compare the two results.

8. What is nuclear fusion? Write an equation of nuclear fusion to support your answer.

9. What is the source of energy in the sun? How is it generated? Illustrate with an example.

10. Describe the construction of an atomic reactor.

11. Calculate the energy released in a fusion reaction

    $$3\,({}^{4}_{2}He) \rightarrow {}^{12}_{6}C$$

    Given, the mass of on α-particle = 4.00263u.

ANSWERS TO INTEXT QUESTIONS

27.1

1. a. ${}^{19}_{9}F + {}^{1}_{1}H \rightarrow {}^{16}_{8}O + {}^{4}_{2}He;$

   b. ${}^{27}_{13}AI + {}^{1}_{0}n \rightarrow {}^{24}_{11}Na + {}^{4}_{2}He;$

   c. ${}^{234}_{90}Th \rightarrow {}^{234}_{90}Pa + {}^{0}_{-1}e;$

   d. ${}^{63}_{29}Cu + {}^{2}_{1}D \rightarrow {}^{64}_{30}Zn + {}^{1}_{0}n$

2. 17.9MeV

3. ${}^{14}_{7}N + {}^{4}_{2}He \rightarrow {}^{17}_{8}O + {}^{1}_{1}H + 6.5\text{MeV}.$

27.2

1. Due to increase of n/p ratio above the natural ratio, its stability decreases. To decrease the ratio to attain more stability, it emits a β-particle.

2. 239Pu

3. 200 MeV.

27.3

1. (1) In fission the energy released is 0.84 MeV/u where as in fusion. It is 6.7 MeV/u. Thus energy released per unit mass is more in the later case.

2. (*a*) 17.3 MeV, (b) 2.69 MeV.

27.4

1. Pressurized Heavy Water Reactor   2. nearly 1 kg.

Answers to Problems in Terminal Exercise

6. 30.6 kg   7. 146.6 g

11. 7.35 MeV

312en28

28

SEMICONDUCTORS AND SEMICONDUCTING DEVICES

Ever since man moved out of the cave and settled into a civil society, his quest for comfort has increased continuously. The invention of fire and wheel proved turning points in human history. Probably, the next big development was the grey revolution, which transformed the way of communication, transportation and living. Sitting in our living rooms, we can connect to our loved ones face-to-face across oceans and continents using computer mediated video-conferencing.

To make all this possible solid state electronic devices have played a significant role. Electronics is a branch of science and technology in which electrons are manipulated to do some specific tasks. Scientists have studied the electrical nature of materials and developed a concept of **energy bands** in terms of which solids can be classified as conductors, semiconductors and insulators. Semiconductors are mostly used for developing electronic devices. Silicon and germanium are the most familiar semiconductor materials. Normally, the conductivity of a semiconductor lies in–between the conductivities of metals and insulators. However, at absolute zero, the semiconductor also acts like a perfect insulator. The conductivity of a semiconductor is influenced by adding some impurity element called *dopant.*

In this lesson you will learn about various types of semiconductors, their behaviour and how they are combined to form useful devices such as Zener diode, solar cell, photodiode, light emitting diode and transistor, etc. You will also learn to draw the I-V characteristics of Zener diode, LED, photo diode and solar cell.

OBJECTIVES

After studying this lesson, you should be able to :

- *explain what energy bands are and how they are used to classify materials as conductors, insulators and semiconductors;*

- *differentiate between (i) intrinsic and extrinsic and (ii) n-type and p-type semiconductors;*
- *explain formation of depletion region and barrier potential in a p-n junction diode;*
- *describe I-V characteristics of a p-n junction diode in the forward and reverse biases;*
- *describe different type of diodes, viz. zener, LED, photo diode and solar cell and their I-V characteristics;*
- *explain the action of a transistor;*
- *describe the effect of doping, size and function of different regions in a transistor;*
- *list the differences between p-n-p and n-p-n transistors;*
- *list different configurations in which a transistor can be connected and describe their input and output characteristics; and*
- *compare different configurations of a transistor in terms of their input/output resistance, gain and applications.*

28.1 ENERGY BANDS IN SOLIDS

While studying the structure of atoms you have learnt that electrons in an isolated atom stay in certain descrete, well defined energy states. When two atoms come closer to form a stable structure, such that the separation between them tends to be lesser than their diameter (d), the energy states tend to overlap, which is forbidden by Pauli's exclusion principle. Hence, they get modified and corresponding to each of the interaction energy states, two energy states are created: one slightly lower than the normal state which is called the bonding state and the other slightly higher than the normal state called the antibonding state.

In solids, very large number of atoms (typically 10^{23} atoms per cm^3) come together to form crystals. If N atoms interact corresponding to each of the energy states, 2N energy states are created. All these energy states are so close to each other (typically $\Delta E \sim 10^{-23}$ eV) that we cannot practically discriminate between them. This quasi continuous distribution of energy states, which are though separate but practically indiscriminable, is called **energy band**.

The process of interaction of energy states (and thereby energy band formation) starts from outer unfilled energy states and then proceeds to valence level. The band formed of unfilled energy levels is called conduction band and the one formed of filled valence levels is called valence band. The relative position of these bands, at equilibrium separation, determines the conduction characteristics of a solid.

28.1.1 Classification of Solids as Conductors, Semiconductors and Insulators on the basis of Energy Bands

If in a solid at equilibrium separation, the conduction band (CB) and valence bands (VB) overlap as it happens in case of metals the material is conductor [Fig. 28.1(a)].

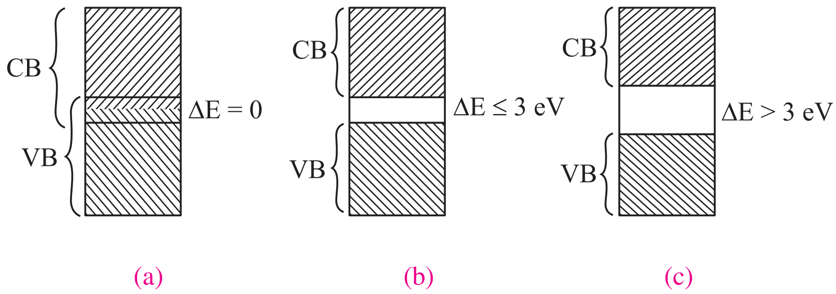


Fig. 28.1 Energy band in (a) Conductors (b) Semiconductors (c) Insulators

If at equilibrium separation the conduction band is completely empty, valence band is completely full and there is a small band gap ($\Delta E \leq 3$ eV) between the highest level of valence band and lowest level of conduction band, called a forbidden energy gap, the solid is a semiconductor. [Fig. 28.1(b)].

If at equilibrium separation, the CB is completely empty, VB is completely filled and there is a large band gap ($\Delta E > 3$ eV) the solid is an insulator.

28.2 INTRINSIC AND EXTRINSIC SEMICONDUCTORS

Semiconductors are classified on the basis of their purity as intrinsic (pure) and extrinsic (impure) semiconductors. Let us now learn about these.

28.2.1 An Intrinsic Semiconductor

Pure silicon and germanium are intrinsic semiconductors as the electrons in these elements are all tightly held in their crystalline structure, i.e., they do not have free electrons. When energy is added to pure silicon in the form of heat, say, it can cause a few electrons to break free of their bonds, leaving behind a hole in each case. (The absence of electrons is treated as positively charged particles having the same amount of positive charge as the negative charge on an electron.) These electrons move randomly in the crystal. These electrons and holes are called *free carriers,* and move to create electrical current. However, there are so few of them in pure silicon that they are not very useful.

Note that in an intrinsic semiconductor, electrons and holes are always generated in pairs and the negative charge of free electrons is exactly balanced by the positive charge of holes. However, a hole only shifts its position due to the motion of an electron from one place to another. *So we can say that when a free electron moves in a crystal because of thermal energy; its path deviates whenever it collides with a nucleus or other free electrons. This gives rise to a zig-zag or random motion, which is similar to that of a molecule in a gas.*

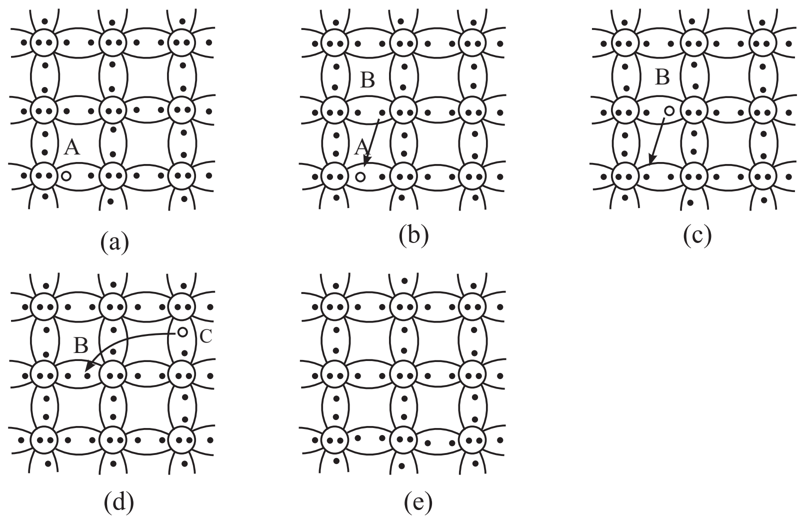


Fig. 28.2 : Movement of electrons and holes in a semiconductor

Now refer to Fig. 28.2(a) and consider the electron- hole pair generated at point *A*. The free electron drifts in the crystal leaving behind a hole. The broken bond now has only one electron and this unpaired electron has tendency to acquire an electron and complete its pair by forming a covalent bond. Due to thermal energy, the electron from neighbouring bond, say at point *B*, may get excited to break its own bond and jump into the hole at *A*. As a result, the hole at *A* vanishes and a new hole appears at *B* (Fig. 28.2(c)). Thus motion of electron from point *B* to point *A* causes the hole to move from *A* to *B*.

You may now like to ask: What will happen when hole at *B* attracts and captures a valence electron from neighbouring bond at *C*? The movement of electron from *C* to *B* causes movement of hole from *B* to *C* [see Fig. 28.2(d) and (e)]. Conventionally, the flow of electric current through the semiconductor is taken in the same direction in which holes move.

At absolute zero temperature, all valence electrons are tightly bound to their parent atoms and intrinsic semiconductor behaves as an insulator. At room temperature, the thermal energy makes a valence electron in an atom to move away from the influence of its nucleus. Therefore, a covalent bond is broken and electron becomes free to move in the crystal, resulting in the formation of a vacancy, called hole. ***Thus, due to thermal energy, some electron-hole pairs are generated and semiconductor exhibits small conductivity.*** For example, at room temperature (300 K), Ge has intrinsic carrier concentration of about $2.5\times 10^{19}\,m^{-3}$ electron-hole pair. As temperature increases, more electron-hole pairs are generated and conductivity increases. Alternatively, we can say that resistivity decreases as temperature increases. It means that semiconductors have negative temperature coefficient of resistance.

28.2.2 An Extrinsic Semiconductor

You now know that intrinsic semiconductors have high resistivity. Also their conductivity shows little flexibility. For these reasons, *intrinsic (pure) semiconductors are of little use; at best these can be used as a heat or light sensitive resistance.* These limitations are overcome by adding a small and measured quantity of another material to intrinsic (pure) semiconductor, which either increases the number of holes or electrons.

Note that the word impurity is being used here because we are adding atoms of some other element to a pure material.

The process in which some atoms of a pure or intrinsic semiconductor are replaced by impurity atoms from their lattice-sites is called **doping** and the impurity so added is called **dopant**. Such doped semiconductors are called **extrinsic** semiconductors.

The dopped semi-conductor normally belongs to group IV of periodic table and the dopants are generally taken from either **group III** (having three valence electrons) or **group V** (having five valence electrons) of the Periodic Table. Fig. 28.3 shows a small portion of the Periodic Table. Here groups III and V have been highlighted to indicate the types of materials generally used for doping.

	III	IV	V	VI
II	Al	Si	P	S
Zn	Ga	Ge	As	Se
Cd	In	Sn	Sb	Te
Hg				

Fig. 28.3 : A part of the Periodic Table. Group III and V elements are used for doping an intrinsic semiconductor of group IV.

Normally we add a very small amount of impurity atoms to the pure simiconductor. It is of the order of one atom per 10^8 atoms of intrinsic semiconductor. These atoms change the balance of charge carriers; either they add free electrons or create holes. Either of these additions makes the material more conducting. Thus, most of the charge carriers in extrinsic semiconductors originate from the impurity atoms.

28.2.3 *n*-and *p*-type Semiconductors

From the electronic configuration of Si ($1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^2$), you will recall that ten electrons are tightly bound to the nucleus and four electrons revolve around the nucleus in the outermost orbit. In an intrinsic silicon semiconductor, the Si atom attains stability by sharing one electron each with four neighbouring Si atoms. (This is called *covalent bonding)*. The same holds true for germanium; its electronic configuration is $1s^2$, $2s^2$, $2p^6$, $3s^2$, $3p^6$, $3d^{10}$, $4s^2$, $4p^2$. When silicon (or germanium) is doped with a pentavalent (five electrons in the outermost orbit) atom like phosphorus, arsenic or antiomony, four electrons form covalent bonds with the four neighbouring silicon atoms, but the fifth (valence) electron remains unbound and is available for conduction, as shown in Fig. 28.4. Thus, when a silicon (or germanium) crystal is doped with a pentavalent element, it develops excess free electrons and is said to be an *n-type* semiconductor. Such impurities are known as *donor impurities.*

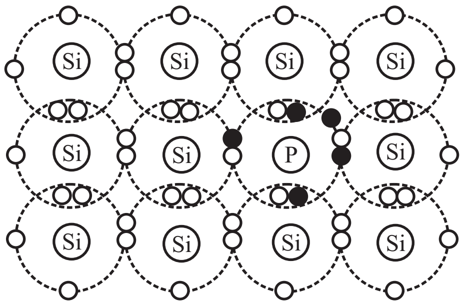


Fig. 28.4 : Covalent bonding in a *n*-type semiconductor

If silicon (or germanium) is doped with a trivalent (three electrons in the outermost shell) atom like boron, aluminium, gallium or indium, three valence electrons form covalent bonds with three silicon atoms and deficiency of one electron is created. This deficienty of electron is referred to as *hole.* It is shown in Fig. 28.5. Such a semiconductor is said to be a *p*-type semiconductor and the impurities are known as *acceptor impurities.*

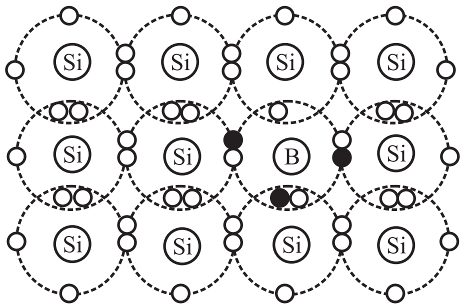


Fig. 28.5 : Colvalent bonding in a *p*-type semiconductor

You may now like to ask: Is a *n*-type semiconductor negatively charged or a p-type semiconductor positively charged? The answer to this question is not in affermative.

In fact, the number of free electrons is exactly equal to the total number of holes and positively charged ions and a semiconductor, whether intrinsic or doped, is electricially neutral.

Note that in a *p*-type semiconductor, more holes are created due to addition of acceptor impurity than by breaking covalent bonds due to thermal energy at room temperature. Hence, the net concentration of holes is significantly greater than that of electrons. That is, in a *p*-type semiconductor, the holes are the majority charge carriers. Similarly we can say that electrons are the majority charge carriers in n-type semiconductors.

INTEXT QUESTIONS 28.1

1. At 300 K, pure silicon has intrinsic carrier concentration of 1.5×10^{16} m^{-3}. What is the concentration of holes and electrons?

2. The *n*-type semiconductor is obtained by doping with
   (i) trivalent impurity
   (ii) pentavalent impurity
   (iii) tetravalent impurity
   (iv) trivalent as well as tetravalent

3. An intrinsic semiconductor can be converted into an extrinsic semiconductor by addition of .... This process in called ....

4. Electrons in *n*-type semiconductor and holes in *p*-type semiconductor are the .... carriers.

5. An extrinsic semiconductor has .... resistivity as compared to an intrinsic semiconductor.

28.3 A *p-n* JUNCTION

You now know that *n*-type and *p*-type semiconductors respectively have electrons and holes as majority charge carriers. What do you think will happen if a *n*-type material is placed in contact with a *p*-type material? Shall we obtain some useful device? If so, how? To answer such questions, let us study formation and working of a *p-n* junction.

28.3.1 Formation of *a p-n* Junction

To form a *p-n* junction, the most convenient way is to introduce donor impurities on one side and acceptor impurities into the other side of a single semiconducting crystal, as shown in Fig. 28.6.

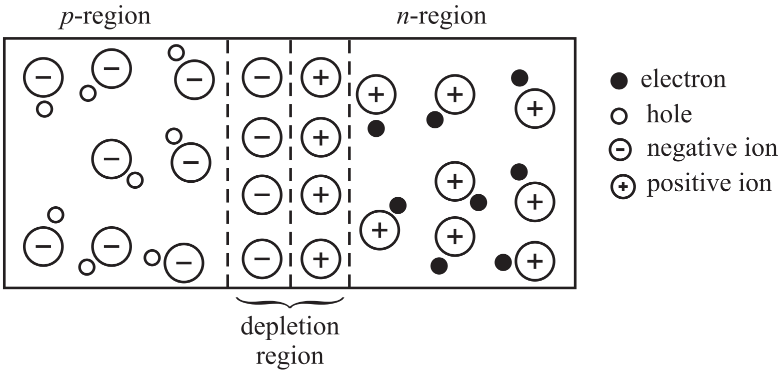


Fig. 28.6 : A p-n junction with depletion region

We now know that there is greater concentration of electrons in the *n*-region of the crystal and of holes in the *p*-region. Because of this, electrons tend to diffuse to the *p*-region and holes to the *n*-region and recombine. Each recombination eleminates a hole and a free electron. This results in creation of positively and negatively charged ions near the junction in *n* and *p* regions, respectively. As these charges accumulate, they tend to act as shield preventing further movement of electrons and holes across the junction. Thus, after a few recombinations, a narrow region near the junction is depleted in mobile charge carriers. It is about 0.5 μm thick and is called the *deplection region* or *space-charge region.*

Due to accumulation of charges near the junction, an electric field is established. This gives rise to electrostatic potential, known as *barrier potential.* This barrier has polarities, as shown in Fig. 28.7. When there is no external electric field, this barrier prevents diffusion of charge carriers across the junction.

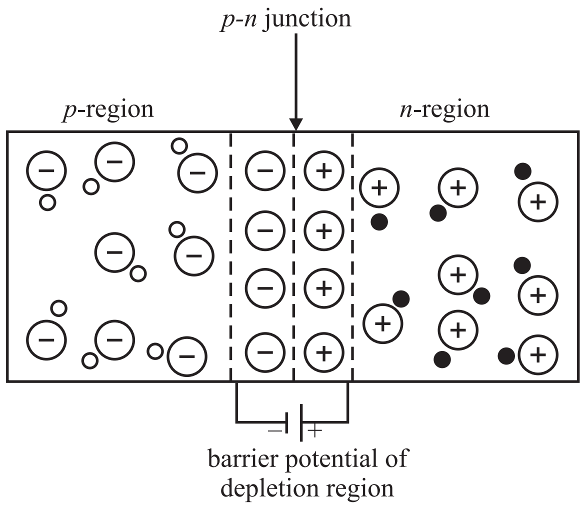
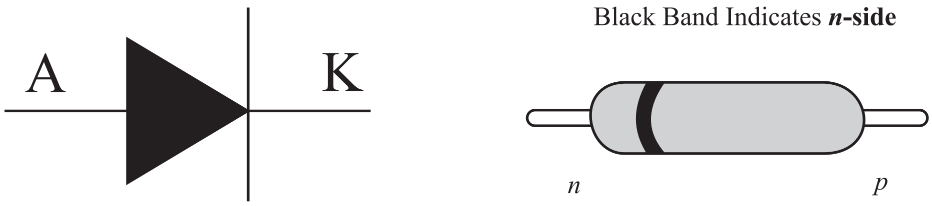


Fig.28.7 : Barrier potential due to depletion region

The barrier potential is characteristic of the semiconductor material. It is about 0.3 V for Ge and about 0.7 V for Si. The junction acts as a diode. It is symbolically represented as shown in Fig. 28.8(a). Here *A* corresponds to *p*-region and acts as an anode. Similarly, K indicates n-region and corresponds to a cathode. Fig 28.8 (b) shows a picture of p-n junction diode available in market.



(a) (b)

Fig. 28.8: (a) Symbol of a p-n junction (diode). The arrow gives the direction of conventional current. It is from p to n region **(b)** A p-n junction diode available in the market.

You may have noted that semiconductor diodes are designated by two letters followed by a serial number. The first letter indicates the material: A is used for material with a band gap of 0.6 eV to 1.0eV such as germanium. B is used for material with a band gap of 1.0eV to 1.3eV, such as silicon. The second letter indicates the main application: A signifies detection diode, B denotes a variable capacitance diode, E for tunnel diode, Y for rectifying diode and Z denotes Zener diode. The serial numbers specify power rating, peak reverse voltage, maximum current rating, etc. (We have to refer to manufacturer's catalogue to know exact details.) For example, *BY*127 denotes a silicon rectifier diode and *BZ*148 represents a silicon Zener diode.

To make visual identification of anode and cathode, the manufacturers employ one of the following ways :

- the symbol is painted on the body of the diode;
- red and blue marks are used on the body of the diode. Red mark denotes anode, whereas blue indicates the cathode;
- a small ring is printed at one end of the body of the diode that corresponds to the cathode. The band in Fig. 28.8(b) indicates the n-side of the p-n junction.

Note that we have to work within the specified ranges of diode ratings to avoid damage to the device.

INTEXT QUESTIONS 28.2

1. Fill in the blanks:

   (a) When a p-n junction is formed, the .... diffuse across the junction.

   (b) The region containing uncompensated acceptor and donor ions is called .... region.

(c) The barrier potential in silicon is .... V and in germanium, it is .... V.

(d) In a *p-n* junction with no applied electric field, the electrons diffuse from *n*-region to *p*-type region as there is .... concentration of .... in *n*-region as compared to *p*- region.

2. Choose the correct option:

(a) The potential barrier at the *p-n* junction is due to the charges on the either side of the junction. These charges are
(i) majority carriers
(ii) minority carriers
(iii) fixed donor and acceptor ions.
(iv) none of above

(b) In a *p-n* junction without any external voltage, the junction current at equillibrium is
(i) due to diffusion of minority carriers only
(ii) due to diffusion of majority carriers only
(iii) zero, as no charges are crossing the junction
(iv) zero, as equal and opposite charges are crossing the junction

(c) In a semiconductor diode, the barrier potential repels
(i) minority carriers in both the regions
(ii) majority carriers in both the regions
(iii) both the majority and the minority carriers
(iv) none of the above

3. Why is depletion region named so? What is depletion region made of?

28.4 FORWARD AND REVERSE BIASED *p-n* JUNCTION

Biasing means application of voltage. To make a *p-n* junction to conduct, we have to make electrons move from the *n*-type region to the *p*-type region and holes moving in the reverse direction. To do so, we have to overcome the potential barrier across the junction by connecting a battery to the two ends of the *p-n* junction diode. The battery can be connected to the *p-n* junction in two ways:

- Positive terminal of the battery connected to the *p*-side and negative terminal of the battery connected to the *n*-side. This is called **forward bias** [Fig. 28.9(a)].
- Positive terminal of the battery connected to the *n*-side and negative terminal of the battery connected to the *p*-side. This is called **reverse bias** [Fig. 28.9(b)].

When a junction is forward biased and the bias exceeds barrier potential, holes are compelled to move towards the junction and cross it from the *p*-region to the *n*-region. Similarly, electrons cross the junction in the reverse direction. This sets in **forward current** in the diode. The current increases with voltage and is of the order of a few milliampere. Under the forward bias condition, the junction offers low resistance to flow of current. Can you guess its magnitude? The value of junction resistance, called **forward resistance,** is in the range 10Ω to 30Ω.

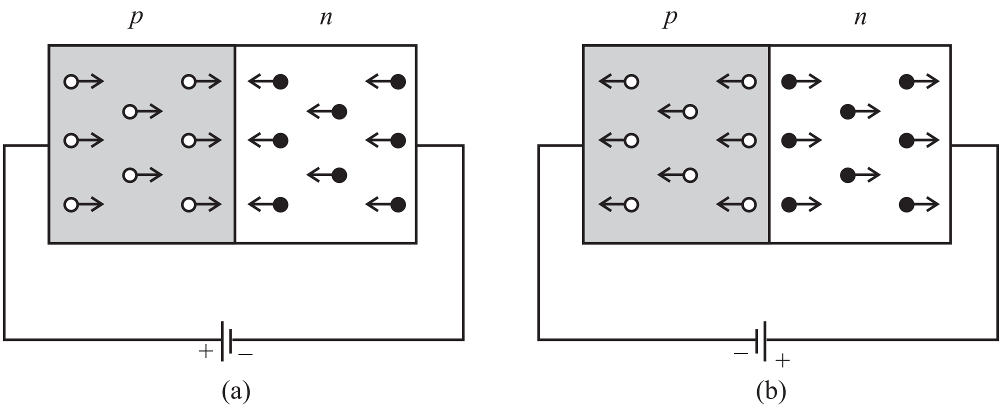


Fig. 28.9 : (a) Forward biased, and **(b)** reverse biased *p-n* junction

When the *p-n* junction is **reverse biased,** holes in the *p*-region and electrons in the *n*-region move away from the junction. Does it mean that no current shall flow in the circuit? No, a small current does flow even now because of the fewer number of electron-hole pairs generated due to thermal excitations. This small current caused by minority carriers is called **reverse saturation current** or **leakage current.** In most of the commercially available diodes, the reverse current is almost constant and independent of the applied reverse bias. Its magnitude is of the order of a few microamperes for Ge diodes and nanoamperes in Si diodes.

A *p-n* junction offers low resistance when forward biased, and high resistance when reverse biased. This property of *p-n* junction is used for ac rectification.

When the reverse bias voltage is of the order of a few hundred volt, the current through the *p-n* junction increases rapidly and damages it due to excessive power dissipation. The voltage at which a diode breaks down is termed as **breakdown voltage.** Physically, it can be explained as follows: When a reverse bias is applied, a large electric field is established across the junction. This field (i) accelerates the available minority carriers, which, in turn, collide with the atoms of the semiconductor material and eject more electrons through energy transfer (avalanche effect), and (ii) breaks covalent bonds by exerting large force on electrons bound by the bonds. This results in creation of additional electron-hole pairs in the junction region (Zener effect). Both these processes give rise to large reverse current even for a small increment in reverse bias voltage. This process is termed as *Zener breakdown*.

INTEXT QUESTION 28.3

1. Define forward bias.
2. Define reverse bias.
3. Fill in the blanks:

   (a) When forward bias is applied on a *p-n* junction diode, the width of the depletion region ....

   (b) When a *p-n* junction diode is reverse biased, the width of depletion region

   (c) When the reverse bias voltage is made too high, the current through the *p-n* junction .... abruptly. This voltage is called.

4. Choose the correct option:

   (a) In a forward biased junction

   (i) the holes in the *n*-region move towards the *p*-region

   (ii) there is movement of minority carriers

   (iii) charge carriers do not move

   (iv) majority carriers in both the regions (*n* and *p*-regions) move into other regions.

   (b) In a reverse biased junction

   (i) there is no of potential barrier
   (ii) there is movement of majority carriers only
   (iii) there is movement of minority carriers only
   (iv) none of the above

5. State two types of reverse breakdowns which can occur in a *p-n* junction diode and differentiate between them.

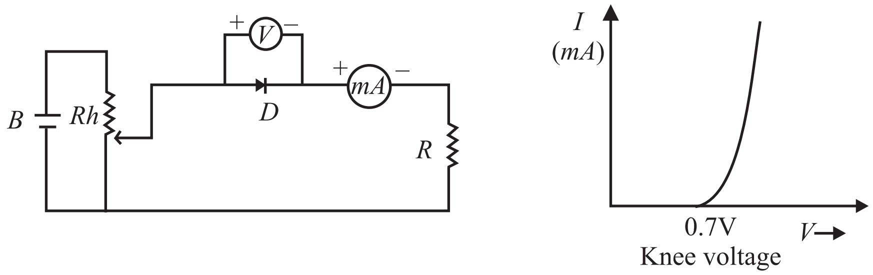
28.5 CHARACTERISTICS OF *p-n* JUNCTION DIODES

The practical application of a semiconductor device in electronic circuits depends on the current and voltage (*I-V*) relationship, as it gives vital information to a circuit designer as well as a technician. Therefore, with the help of *I–V* characteristics, we can know how much current flows through the junction diode at a particular voltage.

28.5.1 Forward Bias Characteristics

Refer to Fig. 28.10(a). You will note that to draw forward bias characterstic of a *p-n* junction diode, the positive terminal of a battery (*B*) is connected to *p*-side of the diode through the rheostat. (Alternative by we can use a variable battery.) The voltage applied to the diode can be varied with the help of the rheostat. The milliammeter (*mA*) measures the current in the circuit and voltmeter (*V*) measures the voltage across the diode. The direction of conventional current is the same as the direction of the diode arrow. Since current experiences little opposition to its flow through a forward biased diode and it increases rapidly as the voltage is increased, a resistance (*R*) is added in the circuit to limit the value of current. If this resistance is not included, the diode may get permanently damaged due to flow of excessive current through it.

The *I-V* characteristic curve of a *p-n* junction in forward bias is shown in Fig. 28.10(b).



(a) (b)

Fig. 28.10 : (a) Circuit diagram *I-V* characteristics of a *p-n* junction diode in forward bias, and **(b)** typical characterstics curve.

Note that the characteristic curve does not pass through origin; instead it meets the *V*-axis at around 0.7V. It means that the *p-n* junction does not conduct until a definite external voltage is applied to overcome the barrier potential. The forward voltage required to get the junction in conduction mode is called **knee voltage.** It is about 0.7 V for Si and 0.3 V for Ge *p-n* junction.

This voltage is needed to start the hole-electron combination process at the junction. As the applied voltage is increased beyond knee voltage, the current through the diode increases linearly. For voltage of around 1V, the current may attain a value of 30-80 mA.

28.5.2 Reverse Bias Characteristics

To draw reverse bias characteristics of a *p-n* junction, we use the circuit diagram shown in Fig. 28.11 (a). If you compare it with Fig. 28.10(a) for forward *I-V* characteristics, you will note two changes:

(i) The terminals of the junction are reversed.

(ii) Instead of milliammeter, microammeter (μ A) is used.

A typical *I-V* characteristic curve of a *p-n* junction in reverse bias is shown in Fig 28.11(b).

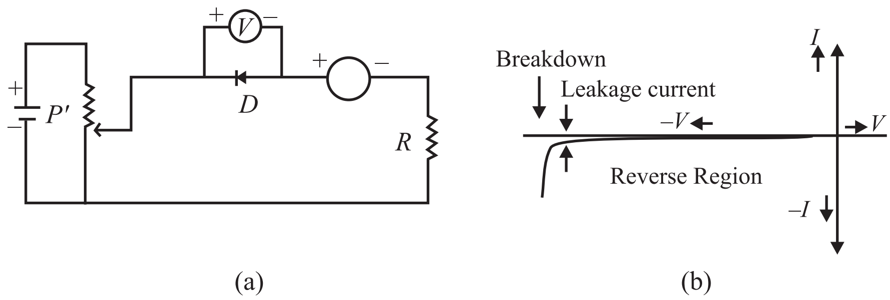


Fig.28.11 : a) Circuit diagram to obtain *I-V* characteristics of a *p-n* junction in reverse bias, and **b)** reverse bias characteristic curve

Note that the junction current is comparatively much less in reverse bias for all voltages below the breakdown voltage. And at breakdown voltage, the current increases rapidly for a small increase in voltage. Moreover, comparison of Fig. 28.10(b) and 28.11(b) reveals that a *p-n* junction diode offers low resistance when it is forward biased and high resistance when reverse biased. At the breakdown voltage in reverse biased *p-n* junction diode, the sharp increase in reverse current is due to sudden decrease in resistance offered by the junction.

From this we may conclude that a *p-n* junction diode conducts in only one direction, i.e. has unidirectional conduction of current, with electrons flowing from the *n*-type region to *p*-type end in forward bias.

You may have seen turnstiles at a metro subway station that let people go through in only one direction. A diode is a one-way turnstile for electrons.

p-n junction diodes find wide applictions. These include :

1. The unidirectional conducting property of a diode is used to convert ac voltage into dc voltage as a *rectifier*. Diodes are also used in adaptors to recharge batteries of cell phones, CDplayers, laptops, etc. You will study about it in detail in the next lesson.
2. A device that uses batteries often contains a diode as it simply blocks any current from leaving the battery, if it is reverse biased. This protects the sensitive electronics in the device.

INTEXT QUESTIONS 28.4

1. Explain the concept of knee voltage.
2. (a) The knee voltage in case of silicon diode is .... whereas in germanium diode it is ....

   (b) In a *p-n* junction diode, the current flows only in .... direction.

   (c) The reverse saturation current is of the order of .... for germanium diodes.

3. Choose the correct option :

   (a) The I-V characteristics of a p-n junction diode in forward bias show

   (i) a non-linear curve

   (ii) linear curve

   (iii) linear as well as non-linear portions

   (iv) none of above

   (b) When a p-n junction is forward biased and the voltage is increased, the rapid increase in current for relatively small increase in voltage occurs

   (i) almost immediately

   (ii) only when the forward bias exceeds the potential barrier

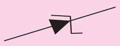
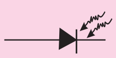
   (iii) when there is breakdown of the junction

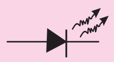
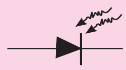
   (iv) none of the above

28.6 TYPES OF DIODES

By adjusting the levels of doping, doping material and the geometry (size, area etc.) of a p-n junction diode, we can modify its electrical and optical behaviour. In this section, we have listed diodes whose properties have been deliberately modified to obtain specific capabilities. Each of these diodes has its own schematic symbol and reflects its nature and functions.

You can use the following table to make a comparison between different diodes:

Name	Symbol	Construction	Principle mechanism	Main	Main use function
Zener diode		p-n junction diode with heavily doped p- & n- regions. Very narrow depletion layer (< 10 nm).	Zener breakdown mechanism	Provides continuous current in reverse breakdown voltage region without being damaged.	Voltage stabilization or regulation
Photo-diode		p-n junction diode. Uses light (or photo) emitting semicon-ductor materials, with very thin p-region, whose thickness is determined by wavelength of radi-ation to be detected	Photovoltaic effect into electrical current in	Converts an optical input controls in VCR & TV reverse bias.	Receivers for remote

LED		*p-n* junction diode with materials having band energies correspoding to near infrared region or visible light region (GaAsP or InP)	Electrolu minous	Changes an electrical inputto a light output in forward bias.	Used in multimeters, digital watches, instrument displays, calculators, switch boards, burglar alarm and remote control devices
Solar cell		*p-n* junction diode in which either *p* or *n* region is made very thin to avoid significant absorption of light before reaching the junction	Photovoltaic effect	Conversion of solar energy into electrical energy	1. In satellites to power systems. 2. To charge batteries. 3. Calculators

28.6.1 I–V Characteristics of Zener diode

Zener diode is fabricated by heavily doping both the *p*- and *n*-sides of the junction. Hence depletion layer formed is very thin $(< 10^{-6}\,\text{m})$. And the electric field across the depletion layer is extremely high $(\sim 5\times 10^{6}\ \text{N C}^{-1})$ even for a small reverse bias voltage of 5 V. The I-V characteristics of a Zener diode is shown in Fig. 28.12. It is seen that when the applied reverse voltage (V) reaches the breakdown voltage (V_z) of the Zener diode, there is a large change in the current. After the breakdown voltage v_z, a large change in the current can be produced by almost insignificant change in the reverse bias voltage. Zener voltage remains constant, even though the current through the Zener diode varies over a wide range.

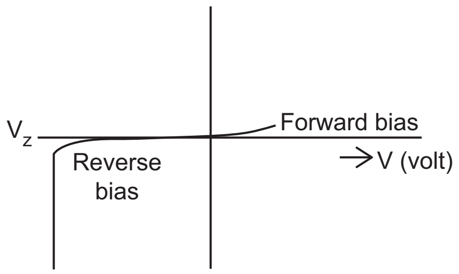


Fig. 28.12

28.6.2 I-V Characteristics of light-emitting diode

In light-emitting diode (LED) when the forward current of diode is small the intensity of light emitted is small. As the forward current increases, the intensity of the emitted light increases and reaches a maximum value. Further increase in the forward current results in a decrease of light intensity. LEDs are biased such that the light emitting efficiency is maximum.

The I-V characteristics of a LED is similar to that of a Si junction diode as shown in Fig. 28.13. But the threshold voltages are much higher and slightly different for each colour. The reverse breakdown voltage of LEDs are very low, typically around 5V.

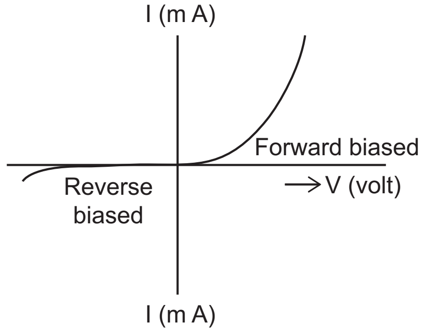


Fig. 28.13

28.6.3 I-V Characteristics of Photo diode

The photo diode is fabricated such that the generation of electron – hole pairs takes place in or near the depletion region in the diode. Due to the electric field of the junction, the electrons and holes are separated before they recombine. The direction of the electric field is such that the electrons reach the *n*-side and the holes reach the *p*-side. The electrons are collected on the n-side and the holes are collected on the *p*-side giving rise to an emf. When an external load is connected, the current flows. The magnitude of the photocurrent depends on the intensity of the incident light.

I-V Characteristics of a photo diode are shown in Fig. 28.14.

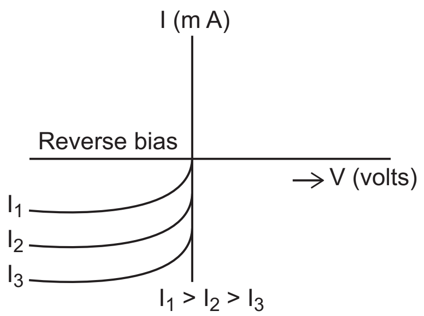


Fig. 28.14

28.6.4 I-V Characteristics of Solar Cell

The generation of emf, when the light falls on a solar cell is due to the following three basic processes: generation, separation and collection. Generation of electron – hole pairs is due to the light (with $h\nu > E_g$) close to the junction. Separation of electrons and holes is due to the electric field of the depletion layer. Electrons are swept to the *n*-side and the holes to the *p*-side.

The electrons reaching the *n*-side are collected by the front contact and the holes reaching the *n*-side are collected by the back contact. Thus the *p*-side becomes positive and the *n*-side negative giving rise to photo voltage.

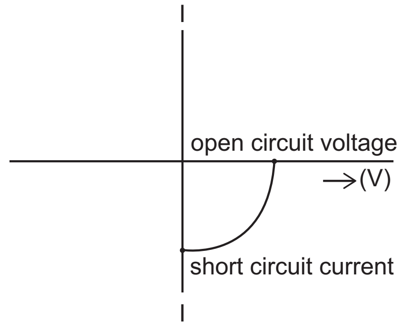


Fig. 28.15

The I-V characteristic curves of a solar cell are shown in Fig. 28.15. They are drawn in the fourth quadrant of the coordinate axis. This is because a solar cell does not draw current but supplies the same to the load.

INTEXT QUESTIONS 28.5

1. Choose the correct option

   (a) A zener diode is operated in

   (i) Forward bias

   (ii) Reverse bias

   (iii) Both of the above

   (iv) None of the above

   (b) Zener dide is

   (i) A highly doped *p-n* junction diode

   (ii) A lowly doped *p-n* junction diode

   (iii) A moderately doped *p-n* junction diode

   (iv) Another name of normal *p-n* junction diode

   (c) A zener diode is used as a

   (i) amplifier

   (ii) rectifier

   (iii) constant current device

   (iv) constant voltage device

2. Fill in the blanks

   a) The zener diode is based on the .... breakdown mechanism.

   b) A photodiode is operated in .... bias.

   c) In a photodiode, the *p-n* junction is made from .... semiconductor material.

   d) LED's are made up of the conductor material from .... of the periodic table.

   e) The light emitting diodes operate in .... bias.

   f) The .... arrow in the symbol of LED symbolizes .... of light.

   g) In an LED light is emitted due to .... of electrons and holes.

   h) LED is based on the principle of ....

   i) Solar cells are based on .... effect.

   j) When sunlight having energy .... than the band gap energy falls on the solar cell, it is .... and frees electron-hole pairs.

3. How does the separation of electrons and holes take place in a solar cell?

28.7 TRANSISTORS – *pnp* AND *npn*

In the preceeding sections, you have learnt about a *p-n* junction diode, which permits current to flow in only one direction. This limits its applications to rectification and detection. A more useful semiconductor device is a bipolar junction transistor.

The invention of transistor by John Bardeen, Walter Brattain and William Shockley in 1948 at Bell laboratory in USA revolutionised the electronic industry. The transistors find many any varied uses in our daily life ranging from gas lighter to toys to amplifiers, radio sets and television. In the form of switching device, these can be used to regulate vehicular traffic on the roads. They form key elements in computers, space vehicles, power systems in satellites and communication.

A transistor is basically a silicon or germanium crystal containing three alternate regions of *p* and *n*-type semiconductors as shown in Fig. 28.16. These three regions are called *emitter*(*E*), *base*(*B*) and *collector*(*C*). The middle region is the base and the outer two regions are emitter and collector. Note that the emitter and collector are of the same type (*p* or *n*) and collector is the largest of the three regions.

The base terminal controls the current flowing between the emitter and the collector. This control action gives the transistor an added advantage over the diode, which has no possibility of controlling the current flow. Depending on the type of doping, the transistors are classified as *n-p-n* or *p-n-p*. In general, the level of doping decreases from emitter to collector to base.

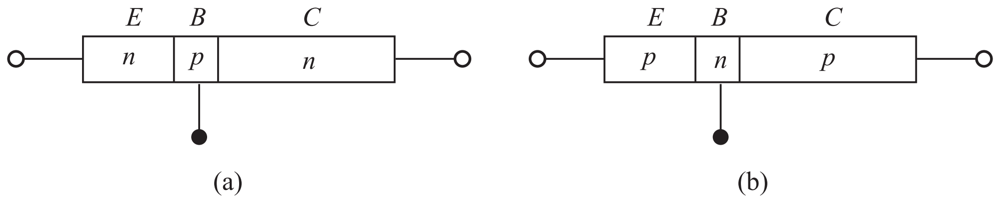
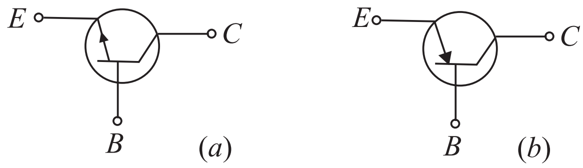


Fig. 28.16 : (a) *n-p-n,* and **(b)** *p-n-p* transistor

The names of the terminals of a transistor give clear indication of their functions. In case of a *n-p-n* transistor, the majority carriers (electrons) from the emitter are injected into base region. Since base is a very lightly doped thin layer, it allows most of the electrons injected by the emitter to pass into the collector. Being the largest of three regions, the collector dissipates more heat compared to the other two regions.



Figs 28.17 : Symbols of **a)** *n-p-n,* and **b)** *p-n-p* transistors

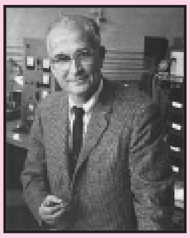
The symbolic representations of *n-p-n* and *p-n-p* transistors are shown in Fig. 28.17. The arrow head indicates the direction of flow of conventional current.

You may now like to ask : Why does the arrow head point outward in case of *n-p-n* transistor and inward in case of *p-n-p* transistor?

In a *n-p-n*, transistor, the emitter current is due to flow of electrons from emitter to base, and the conventional current flows from base to emitter and hence the arrow head points out from the base. In case of *p-n-p* transistor, the emitter current comprises flow of holes from emitter to base. Thus the conventional current flows from emitter to base.

Since transistors are bipolar devices, their operation depends on both the majority and minority carriers.

William Bradford Shockley
(1910 – 1989)



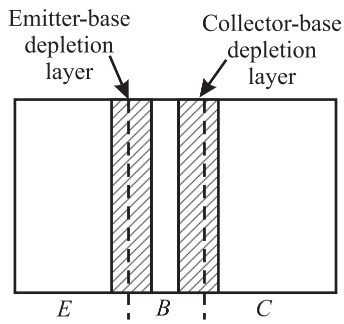
England born, American physicist W.B. Shockley was one of the three scientists who received 1956 Nobel Prize in physics for the discovery of transistor. Basically a solid state physicist, shockley contributed significantly to the development of theoretical understanding of bands in semiconductors, order and disorder in alloys; theory of vacuum tubes, theory of dislocations and theory of ferromagnetic domains. He is truely one of the pioneers of electronic revolution.

28.7.1 Working Principle

You are familiar with the working of a *p-n* junction. We now discuss the working principle of a transistor and consider an *n-p-n* transistor first because it is more commonly used.

When no voltage is applied across the transistor, diffusion of free electrons across the junctions produces two depletion layers, as shown in Fig. 28.18. For each depletion layer, the barrier potential is about 0.7V at 25°C for a silicon transistor and 0.3V for a germanium transistor. As you may be aware, silicon transistors are more widely used than germanium transistors because of higher voltage rating, greater current ratings, and low temperature sensitivity. For our discussion, we refer to silicon transistors, unless otherwise indicated.

Since the three regions in a transistor have different doping levels, the depletion layers have different widths. If a region is heavily doped, the concentration of ions near the junction will be more, resulting in thin depletion layer and vice versa. Since the base is lightly doped as compared to emitter and collector, the depletion layers extend well into it, whereas penetration in emitter/collector regions is to a lesser extent (Fig. 28.18). Moreover, the emitter depletion layer is narrower compared to collector depletion layer.



Figs 28.18: Depletion layers in a transistor when no voltage is applied

In order to made a transistor funciton properly, it is necessary to apply suitable voltages to its terminals. This is called **biasing** of the transistor.

A *n-p-n* Transistor

A typical biasing scheme of a *n-p-n* transistor is shown in Fig. 28.19(a). Note that *the emitter-base junction is forward biased while the collector-base junction is reverse biased.* We therefore expect a large emitter current and low collector current. But in practice, we observe that the collector current is almost as large as the emitter current. Let us understand the reason. When forward bias is applied to the emitter, free electrons in the emitter have to overcome the barrier potential to enter the base region [see Fig. 28.19(b)]. When V_{BE} exceeds barrier potential (0.6 to 0.7V for silicon transistor), these electrons enter the base region, as shown in Fig. 28.19(c). Once inside the base, these electrons can flow either through the thin base into the external base lead or across the collector junction into the collector region. The downward component of base current is called *recombination current*. It is small because the base is lightly doped and only a few holes are available. Since the base region is very thin and it receives a large number of electrons, for $V_{BE} > 0.7V$, most of these electrons diffuse into the collector depletion layer. The free electrons in this layer are pushed (by the depletion layer field) into the collector region [(Fig. 28.19(d)] and flow into the external collector lead. So, we can say that a steady stream of electrons leaves the negative source terminal and enters the emitter

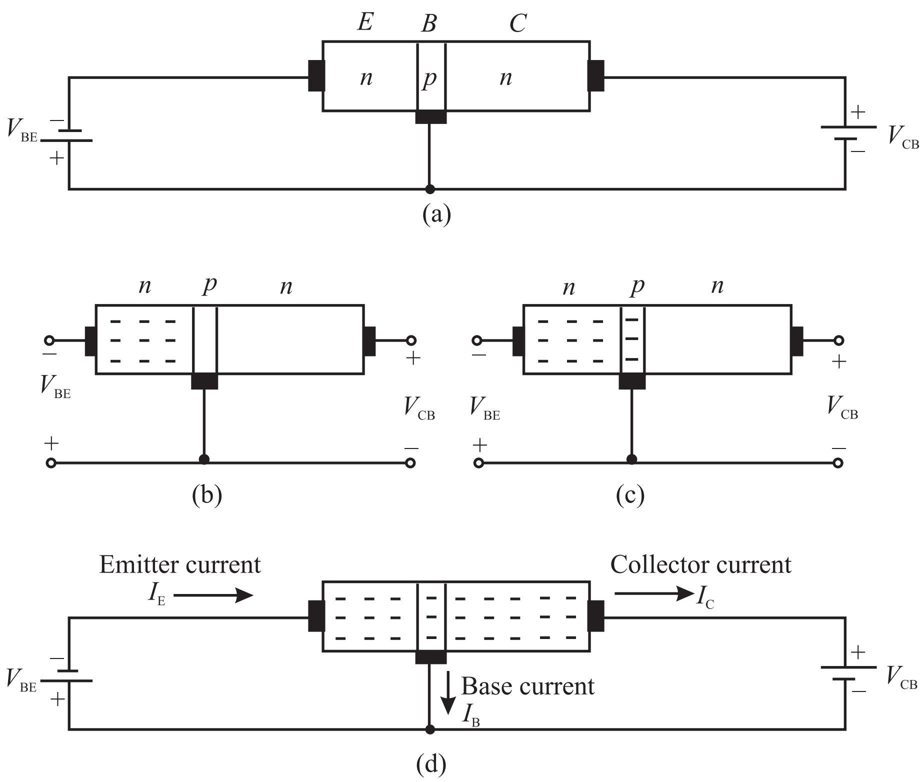


Fig. 28.19 : A *n-p-n* transistor when **(a)** emitter is forward-biased and collector is reverse-biased, **(b)** free electrons in an emitter, **(c)** free electrons injected into base; and **(d)** free electrons pass through the base to the collector.

region. The forward bias forces these electrons to enter the base region. Almost all these electrons diffuse into the collector depletion layer through the base. The depletion layer field then pushes a steady stream of electrons into the collector region. In most transistors, more than 95 percent emitter-injected electrons flow to the collector; less than 5 percent flow to the external base lead.

From this you should not conclude that you can connect two discrete diodes back to back to get a transistor. This is because in such a circuit, each diode has two doped regions and the overall circuit would have four doped regions and the base region would not be the same as in a transistor. *The key to transistor action, therefore, is the lightly doped thin base between the heavily doped emitter and the intermediately doped collector*. Free electrons passing through the base stay in base for a short time and reach the collector.

The relation between collector current (I_C) and emitter current (I_E) is expressed in terms of signal current gain, α, of a transistor. It is defined as

$$\alpha = \frac{I_C}{I_E}. \qquad (28.1)$$

You should note that the value of α is nearly equal to but always less than one.

Similarly, we can relate the collector current to the base current in a transistor. It is denoted by greek letter beta:

$$\beta = \frac{I_C}{I_B} \qquad (28.2)$$

Beta signifies the current gain of the transistor in common-emitter configuration. The value of β is significantly greater than one.

Since emitter current equals the sum of collector current and base current, we can write

$$I_E = I_C + I_B$$

On dividing throughout by I_C, we get

$$\frac{I_E}{I_C} = 1 + \frac{I_B}{I_C}. \qquad (28.3)$$

In terms of α and β, we can rewrite it as

$$\frac{1}{\alpha} = 1 + \frac{1}{\beta}$$

or

$$\beta = \frac{\alpha}{1-\alpha} \qquad (28.4)$$

Let us now consider how a *p-n-p* transistor differs from a *n-p-n* transistor in its details.

A *p-n-p* Transistor

A *p-n-p* transistor biased for operation in the active region is shown in Fig 28.20. Note that we reverse the battery terminals when *n-p-n* transistor is substituted by *p-n-p* transistor.

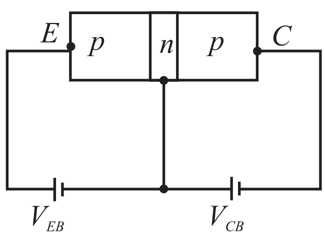


Fig. 28.20 : A p-n-p transistor biased for active operation

As before, the emitter - base junction is forward biased by battery of voltage V_{EB} and the collector base junction is reverse biased by a battery of voltage V_{CB}. The resistance of the emitter-base junction is very small due to its forward bias as compared to the collector-base junction (which is reverse biased). Therefore, we apply small forward bias voltage (0.6V) to the emitter-base junction, whereas the reverse bias voltage applied to the collector-base junction is of much higher value (9V).

The forward bias of emitter-base junction makes the majority carriers, that is the holes, in emitter (*p*-region), to diffuse to the base (*n*-region), on being repelled by the positive terminal of the battery. As width of the base is extremely thin and it is lightly doped, very few (two to five percent) of total holes that enter the base recombine with electrons and 95% to 98% reach the collector region. Due to reverse bias of the collector- base region, the holes reaching this region are attracted by the negative potential applied to the collector, thereby increasing the collector current (I_C). Therefore, increase in emitter current (I_E) increases collector current. And Eqns. (28.1) – (28.4) hold in this case as well.

INTEXT QUESTION 28.6

1. Choose the correct option:
   (a) The arrow head in the symbol of a transistor points in the direction of
      (i) hole flow in the emitter region
      (ii) electron flow in emitter region
      (iii) majority carriers flow in the above region
      (iv) none of the above
   b) The emitter current in a transistor in normal bias is
      (i) less than the collector current
      (ii) equal to sum of base current and collector current
      (iii) euqal to base current
      (iv) none of the above
2. Fill in the blanks
   (a) A ransistor has .... regions and .... junctions.
   (b) In a transistor, ... has the least thickness.
   (c) The emitter region is .... doped, whereas .... region has the least doping.
   (d) The collector of the transistor has .... size and .... doping.

(e) The transistor is said to be in active region when .... junction is forward biased and .... junction is reverse biased.

(f) The two types of transistors are .... and ....

You now know the working principle of a transistor. Let us learn the various ways in which a transistor is biased.

28.7.2 Transistor Configurations

A transistor is a two-port device; it can take an input and deliever an output. For both input and output, two terminals are needed. This can be done in a transistor by making one of the three terminals common. The configurations of a transistor in which one of the terminals is common to both input and output are shown in Fig. 28.21.

- When emitter is common to both input and output circuits, we obtain common emitter (*CE*) configuration (Fig. 28.21a);
- When base is common to both input and output circuits, we obtain common base (*CB*) configuration (Fig. 28.21b); and
- When collector is common to both input and output circuits, we have common collector (*CC*) configuration (Fig.28.21c).

In each of these configurations, the transistor characteristics are unique. The *CE* configuration is used most widely because it provides voltage, current and power gains. In the *CB* configuration, the transistor can be used as a constant current source while the *CC* configuration is usually used for impedance matching.

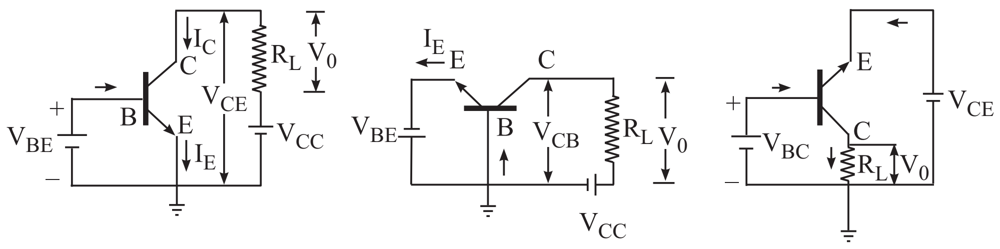


Fig. 28.21: Transistor configuration: **a)** *CE*, **b)** *CB*, and **c)** *CC*

For each configuration, we can plot three different characteristics: a) input characteristics, b) output characteristics, and c) transfer characteristics, depending on the nature of quantities involved.

Table 28.2 gives various quantities related to each of these characteristics in all the three configurations and the transistor constants of interest.

Table 28.2: Physical quantities of interest in different characteristics of a transistor

Configuration	Input Characteristic	Output characteristic	Transfer characteristic	Important transistor constant
CE	V_{BE} and I_B with V_{CE} as parameter	V_{CE} and I_C with I_B as parameter	I_B and I_C	Current gain, β
CB	V_{BE} and I_E with V_{CB} as parameter	V_{CB} and I_C with I_E as parameter	I_E and I_C	Large signal current gain, α
CC	V_{CB} and I_B with V_{CE} as parameter	V_{CE} and I_E with I_B as parameter	I_B and I_E	

To work with a transistor, you will be required to identify its base, emitter and collector leads. To do so, you can follow the following steps.

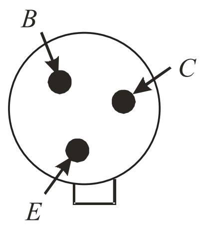


Fig. 28.22 : Identifying transistor leads.

Look for the a small notch provided on the metallic cap. The terminal close to the notch is emitter. To identify other two terminals, turn the transistor up-side-down. You can easily identify the base and the collector as shown in Fig. 28.22.

Like a *p-n* junction diode, transistors are also designated with two letters followed by a serial number. The first letter gives an indication of the material. *A* is for germanium and *B* is for silicon. The second letter indicates the main application: *C* is used for audio frequency transistors, *D* for power transistors and *F* for radio-frequency transistors. The serial number consists of digits assigned by the manufacturer for identification. For example, *AC* 125 represents germanium transistor for *AF* applications.

28.8 TRANSISTOR CHARACTERISTICS

As mentioned earlier, operation of a transistor can be studied with input and output *I-V* characteristics. The nature of these characteristics is unique and depends on the configuration used. Let us first study *CE* configuration.

28.8.1 Common Emitter (*CE*) Configuration of a *npn* Transistor

Common emitter characteristics of a transistor relate voltage and current when emitter is common to both input and output circuits. The circuit diagram for *CE* characteristics of a *n-p-n* transistor is shown in Fig. 28.23. V_{BB} is a variable *dc* supply of 0-3V and V_{CC} is a variable *dc* supply of 0-15V. R_1 and R_2 are potentiometers and R is a variable resistor. It is used to control base to emitter voltage, V_{BE}.

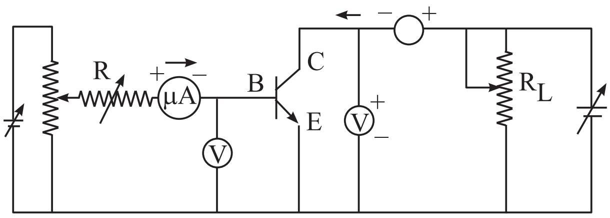


Fig. 28.23 : Circuit diagram for input and output characteristics of a transistor in *CE* configuration.

Input characteristics

In *CE* configuration, the input characteristics show the variation of I_B with V_{BE} when V_{CE} is held constant. To draw this characteristic, V_{CE} is kept at a suitable value with the help of R and R_1. Then V_{BE} is changed in steps and corresponding values of I_B are measured with the help of microammeter, connected to base. Fig. 28.24. shows typical input characteristics of a *n–p-n* transistor in *CE* configuration.

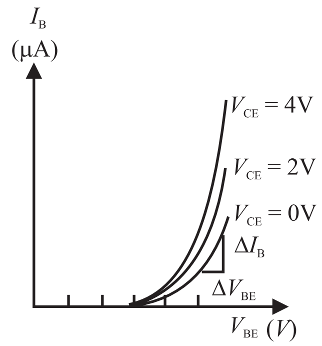


Fig. 28.24 : Input characteristics of a typical *npn* transistor in *CE* configuration

Note that for a given value of V_{CE}, the curve is as obtained for forward biased *p-n* junction diode. For $V_{BE} < 0.5$V, there is no measurable base current ($I_B = 0$). However, I_B rises steeply for $V_{BE} > 0.6$V.

From the reciprocal of the slope of input characteristic, we get input resistance of the transistor defined as the ratio of small change in base - emitter voltage to the small change produced in the base current at constant collector - emitter voltage:

$$R_{ie} = \left.\frac{\Delta V_{BE}}{\Delta I_B}\right|_{V_{CE}} \qquad (28.5)$$

Usually, the value of R_{ie} is in the range 20-100Ω. You should note that since the curve is not linear, the value of input resistance varies with the point of measurement. As V_{CE} increases, the curve tends to become more vertical and the value of R_{ie} decreases.

Output characteristics

The output characteristic curves depict the variation of collector current I_C with V_{CE}, when base current I_B is kept constant. To draw output characteristics, I_B is fixed, say at 10 μA, by adjusting R_1 and R. V_{CE} is then increased from 0 to 10 V in steps of 0.5V by varrying R_2 and the corresponding value of I_C is noted. Similarly, the output characteristics can be obtained at I_B = 40μA, 60μA, 80μA. However, in no case, the maximum base current rating of the transistor should be exceeded.

The output characteristics of this configuration are shown in Fig. 28.25.

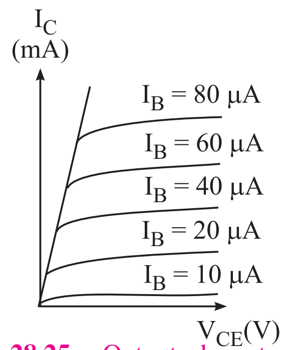


Fig. 28.25 : Output characteristics of a typical *npn* transistor in CE configuration

From the output characteristics, you will note that I_C changes with increase in V_{CE} for a given value of I_B and I_C increases with I_B for a given V_{CE}. *From these characteristics, we can calculate output admittance* (h_{oe}):

$$h_{oe} = \frac{\Delta I_C}{\Delta V_{CE}} \tag{28.6}$$

where Δ denotes a small change.

28.8.2 Common Emitter (*CE*) Configuration of a *pnp* Transistor

In the preceding section, you learnt to draw input and output characterstics of a *n-p-n* transistor in common emitter configuration. Now we will consider a *p-n-p* transistor. Fig. 28.26 shows the circuit diagram for *CE* characteristics of a *p-n-p* transistor. The transistor is biased to operate in the active region. The microammeter and voltmeter are used in the base- emitter circuit to measure the base current (I_B) and the voltage between base and emitter. Similarly, milliammeter and voltmeter are connected in collector-emitter circuit to measure the collector current (I_C) and voltage between collector and emitter (V_{CE}).

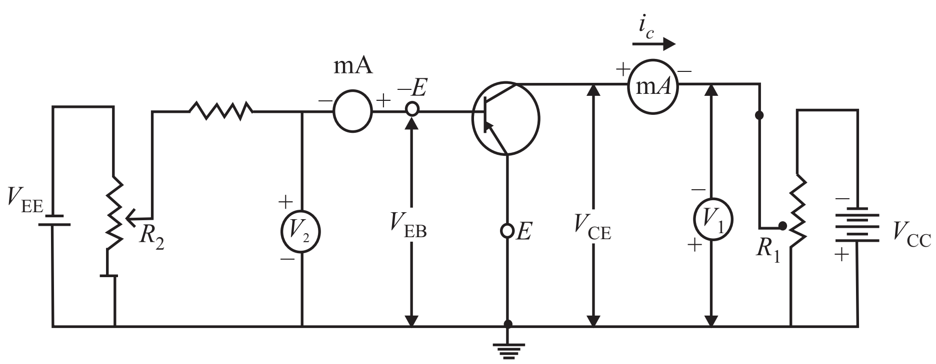


Fig. 28.26 : Circuit diagram for obtaining input and output characteristics of a *p-n-p* transistor in *CE* configuration

Input Characteristics

Input characteristics are graphs between V_{BE} and I_B at different constant values of V_{CE}.

To plot input characteristics, the potentiometer R_1 in the emitter- collector circuit is adjusted till the voltmeter shows constant value. Then potentiometer in the emitter-base circuit is adjusted in such a way that base-emitter voltage is zero. For this value, base current is also observed to be zero. Keeping the V_{CE} constant, V_{BE} is increased gradually and change in base current is noted with the help of microammeter. To plot input characteristics at $V_{CE} = -2$V, say, the potentiometer in emitter-collector circuit is adjusted till the voltmeter in the same circuit reads 2V. Then potentiometer in the emitter -base circuit is adjusted to make V_{BE} zero. Then V_{BE} is increased gradually, keeping V_{CE} constant. Similarly the input characteristics of the transistor in the CE configuration can be drawn for different values of V_{CE} = – 6V, 1V and so on. Fig. 28.27 shows typical input characteristics of *CE* configuration. As may be noted, the nature of *input characteristics is similar to the forward characteristics of p-n junction diode.* The base current remains zero as long as the base voltage is less than the barrier voltage (for silicon transistor, it is ~0.7V). As the base voltage exceeds barrier voltage, current begins to increase slowly and then rises abruptly.

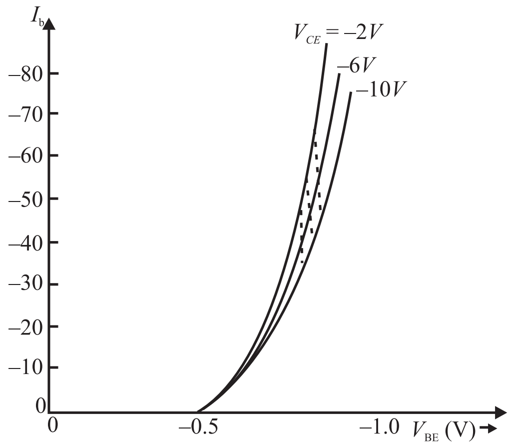


Fig. 28.27 : Input characteristics of a typical *p-n-p* transistor in *CE* configuration.

You may also recall that these curves are similar to the ones obtained for the *CE* configuration for *n-p-n* transistor.

From the reciprocal of the slope of the curve of input characteristic, the a.c input resistance of the transistor can be calculated.

- ***a.c* input resistance (R_{in})** of the transistor in *CE* configuration is expressed as:

$$R_{in} = \left.\frac{\Delta V_{BC}}{\Delta I_B}\right|_{V_{CE}} = \text{constant} \qquad (28.7)$$

In this configuration R_{in} is typically of the order of one $k\Omega$.

Output Characteristics

These are graphs between collector-emitter voltage (V_{CE}) and the collector current (I_C) at different constant values of base current (I_B).

To draw these characteristics, V_{CE} is made zero and V_{BE} is adjusted till the microammeter in the base-emitter circuit is set to read a constant value. Thus V_{CE} is adjusted to make I_B constant at a particular value. Now keeping I_B constant, V_{CE} is increased from zero in a number of steps and the corresponding collector current I_C is noted with the help of milliammeter connected in series with collector.

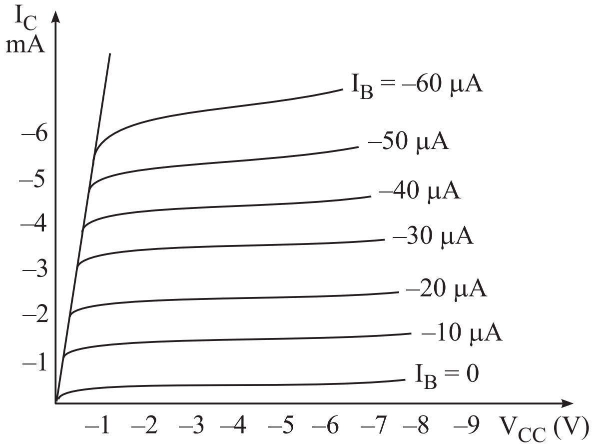


Fig 28.28 : Output characteristics of a typical *pnp* transistor in *CB* configuration

How can we plot the output characteristics at I_B = 50 μA? To do so, V_{BE} is adjusted till milliammeter reads 50 μA. Increase V_{CE} gradually and note correspoding values of I_C. The graph betweent V_{CE} and I_C gives the output characteristics at I_B = 50 μA. Similarly, the output characteristics can be obtained at I_B = 100 μA, 200 μA and so on. Fig. 28.28 shows output characteristics of *p-n-p* transistor for *CE* configuration.

Example 28.1 : Calculate the current gain β of a transistor if the current gain α = 0.98

Solution: $\beta = \dfrac{\alpha}{1-\alpha} = \dfrac{0.98}{1-0.98} = 49$

Example 28.2 : In a transistor, 1 mA change in emitter current changes collector current by 0.99 mA. Determine the a.c current gain.

Solution: Given $\Delta I_e = 1$ mA $= 1 \times 10^{-3}$ A and $\Delta I_c = 0.99$ mA $= 0.99 \times 10^{-3}$ A

Therefore, a.c current gain of the transistor $\alpha = \dfrac{\Delta I_c}{\Delta I_e} = \dfrac{0.99\times10^{-3}}{1\times10^{-3}\,\text{A}}$ A $= 0.99$

INTEXT QUESTIONS 28.7

1. Fill in the blanks
   (a) The .... curve relates the input current with input voltage, for a given output voltage.

(b) The .... curve relates the output current with the output voltage for a given input current.

(c) In common emitter configuration of a transistor, the .... and .... are the output terminals

(d) The .... and .... are the input terminals, whereas .... and .... are the output terminals of a transistor in common base configuration.

WHAT YOU HAVE LEARNT

- Semiconductors are materials like silicon (Si) and germanium (Ge), which have conductivities midway between insulators and conductors.
- Semiconductors are of two types : Intrinsic (pure) and extrinsic (dopped).
- Extrinsic semiconductors can be *p*-type (dopped with 3rd group impurities) or *n*-type (doped with 5^{th} group impurities).
- A *p-n* junction diode consists of a *n*-type region and a *p*-type region, with terminals on each end.
- When a *p-n* junction is formed, diffusion of holes and electrons across the junction results in a depletion region which has no mobile charges.
- The ions in the region adjacent to the depletion region generate a potential difference across the junction.
- A forward biased *p-n* junction offers low resistance to flow of electrons.
- A reverse biased *p-n* junction diode offers high resistance to flow of current.
- A *p-n* junction allows current to flow in only one direction.
- There are various types of diode e.g. photo diode light emitting diode, Zener diode and solar cell.
- A photo diode is always connected in reverse bias.
- A transistor consists of three separate regions (emitter, base and collector) and two junctions. Emitter is most heavily doped and base is the least doped. While collector has the largest size, base is the thinnest.
- Transistor can either be *n-p-n* type or *p-n-p* type.
- A transistor can be connected in any of the three configurations: common collector (*CE*), common base (*CB*) or common emitter (*CE*).
- The characteristics of a transistor vary according to the configuration of the transistor.
- *CE* configuration is preferred over other configurations as it provides high current gain and voltage gain.

TERMINAL EXERCISE

1. Describe the most important characteristic of a *p-n* junction diodes.
2. Explain the formation of depletion region in a *p-n* junction diode.
3. Which charge carriers conduct forward current in a *p-n* junction diode?
4. Differentiate between
   (i) Forward bias and reverse bias
   (ii) Avalancehe and zener breakdown
5. Explain the working of *p-n-p* and *n-p-n* transistors.
6. Define current gains α and β of a transistor.
7. For $\alpha = 0.998$, calculate change in I_C if change in I_E is 4 mA.
8. What are energy bands in solids? How are they formed? How do we classify solids as conductor semiconductors and insulators on the basis of energy bands?

ANSWERS TO INTEXT QUESTIONS

28.1

1. 1.5×10^{15} each 2. (ii)

3. impurity, doping 4. majority 5. lower

28.2

1. (a) majority carriers (b) depletion region
   (c) 0.7, 0.3 (d) higher, electrons
2. (a) (iii), (b) (iii), (c) (ii)

28.3

3. (a) decreases (b) increases (c) increases, breakdown voltage
4. (a) (iv); (b) (iii)

28.4

2. (a) 0.7 V, 0.3 V; (b) one (c) micro ampere
3. (a) (iii) ; (b) (ii)

28.5

1. (ii), (i), (iv)
2. (a) Zener (b) reverse (c) light sensitive
   (d) group III-V (e) forward (f) emission
   (g) recombination (h) electroluminiscence
   (i) photovoltaic (j) more, absorbed
3. Separation of electrons and holes takes palce due to the electric field of the depletion layer.

28.6

1. (a) (i); (b) (ii)
2. (a) Three, two; (b) Base (c) Most heavily, base
   (d) largest size, moderate (e) Emitter-base, collector-base
   (f) *npn, pnp*

28.7

1. (a) input characteristic (b) output characteristics
   (c) collector, emitter (d) base and emitter, base and collector

Answers to Problems in Terminal Exercise

7. 3.992 mA

312en29

29

APPLICATIONS OF SEMICONDUCTOR DEVICES

In the last lesson, you learnt the working principle of semiconductor devices like *p-n* junction diode, Zener diode, LED, solar cells and transistors. Due to their miniature size and special electrical properties, these devices find applications in almost every household appliances and gadgets like gas lighter, security alarm, radio, TV, telephone, tape recorder, CD player, computer, fan regulator, emergency lights etc. All control mechanisms in big industries and flight control equipments in an aeroplane and power systems in satellites use semiconductor devices. In a way, it is now difficult to imagine life without these.

In this lesson you will learn some simple applications of diodes and transistors. This discussion is followed by an introduction to elements of **digital electronics.** This branch of electronics handles special types of signals/waveforms, which can assume only two values, 0 and 1. Digital electronics is based on the concept of **logic gates.** These gates accept input in digital form and give output according to the logic operation it is supposed to perform. You will learn about logic gates, their symbols and circuit implementation in this lesson.

OBJECTIVES

After studying this lesson, you should be able to:

- *explain the use of diode as a half-wave and a full-wave rectifier;*
- *explain the use of Zener diode as voltage regulator;*
- *describe the uses of a transistor as an amplifier, a switch and an oscillator;*
- *explain the logic gates with their Truth Tables; and*
- *realize logic gates using simple circuit elements.*

29.1 APPLICATIONS OF *p-n* JUNCTION DIODES

You now know that a *p–n* junction exhibits asymmetric electrical conduction, i.e., its resistance in forward bias is different from that in reverse bias. This property of a diode is used in rectification, i.e., conversion of an ac signal into a dc signal (of constant magnitude). In every day life, we may need it to charge a cell phone, laptop etc. Let us now learn about it.

29.1.1 *p-n* Junction Diode as a Rectifier

You have learnt in Lessons of Module 5 that the electricity supply in our homes provides us ac voltage. It is a sinusoidal signal of frequency 50 Hz . It means that voltage (or current) becomes zero twice in one cycle, i.e., the waveform has one positive and other negative half cycle varying symmetrically around zero voltage level. The average voltage of such a wave is zero. Let us now learn the mechanism to convert an ac into dc.

(a) Half-Wave Rectification

Refer to Fig. 29.1. The signal from ac mains is fed into a step down transformer T which makes it available at the terminals X and Y. The load resistance R_L is connected to these terminals through a *p-n* junction diode D. You may now like to ask : Why have we used a step down transformer? This is done due to the fact that most devices require voltage levels lower than 220V. The stepped down ac signal is obtained at the output of stepdown transformer. The potential at terminal X with respect to Y will vary as a sine function with time, as shown in Fig. 29.2(a). In the positive half cycle, during the time interval 0 to $T/2$, diode D will be forward biased and conduct, i.e., current flows through R_L from A to B. However, during the negative half cycle, i.e., in the interval $T/2$ to T, D is reverse biased and the junction will not conduct, i.e. no current flows through R_L. This is shown in Fig. 29.2(b). Since the *p-n* junction conducts only in one-half cycle of the sine wave, it acts as a half-wave rectifier.

During the non-conducting half cycle, the maximum reverse voltage appearing across the diode is equal to the peak ac voltage V_m. The maximum reverse voltage that a diode can oppose without breakdown is called its **Peak Inverse Voltage(PIV)**. For rectification, we must choose a diode having PIV greater than the peak ac voltage to be rectified by it; otherwise it will get damaged. The dc voltage, V_{dc} across R_L, as measured by voltmeter in case of half-wave rectifier, is given by

$$V_{dc} = V_m/\pi \qquad (29.1)$$

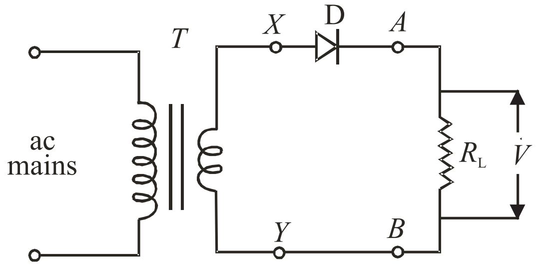


Fig. 29.1: Half wave rectifier circuit

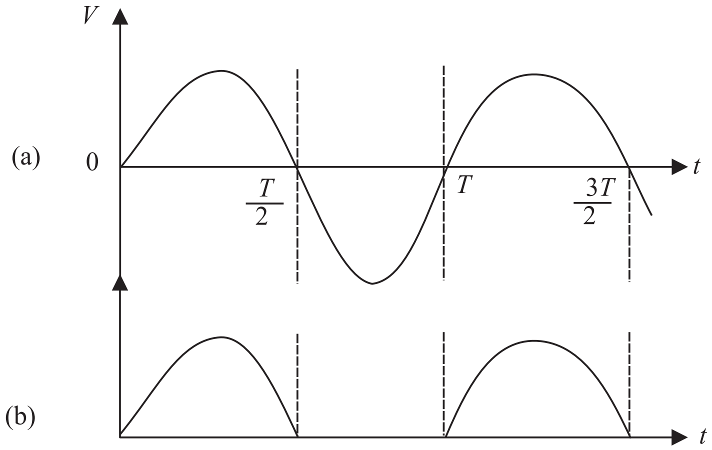


Fig. 29.2: (a) Input ac voltage, and **b)** half-wave rectified output

where V_m is the peak ac voltage. The dc current I_{dc} through the load resistance R_L is given by

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{V_m}{\pi R_L} \qquad (29.2)$$

Note that in this case, we are utilizing only half of the input power and obviously it is not an efficient way of obtaining dc. You may logically think that instead of one, we should use two diodes in such a way that they conduct in alternate cycles. This is known as full-wave rectification. Let us learn about it now.

(b) Full-Wave Rectification

For full-wave rectification, we feed the input signal in a centre tapped step down transformer. (It has two identical secondary windings connected in series.) D_1 and D_2 are two *p-n* junction diodes, as shown in Fig. 29.3. One end of the load resistance R_L is connected to the central point *Y* of the secondary windings and the other end is connected to the cathode terminals of the diodes D_1 and D_2. The anodes of these diodes are connected respectively to the ends *X* and *Z* of the secondary windings. The potentials at the ends *X* and *Z* are in opposite phase with respect to *Y*, i.e., when potential of *X* is positive, *Z* will be negative and vice versa. It is shown graphically in Fig. 29.4 (a) and (b).

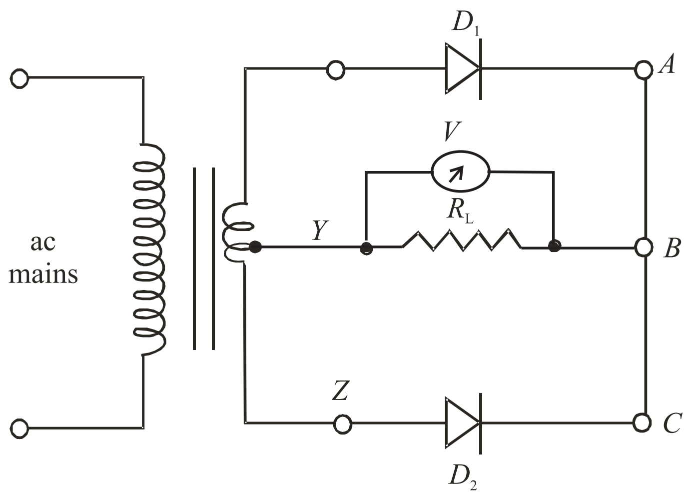


Fig. 29.3 : A full-wave rectifier circuit using two diodes

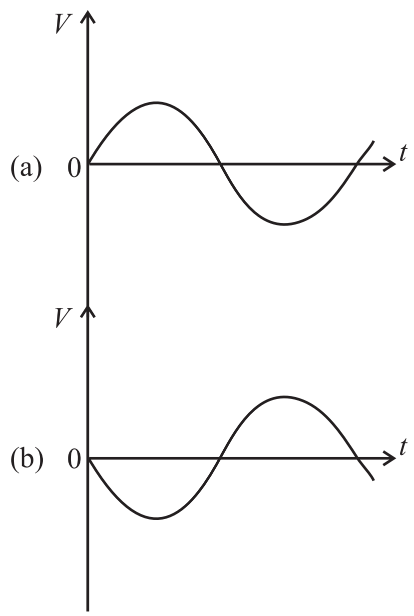


Fig. 29.4 : (a) Potential at point *X* is positive with respect to *Y*, and **(b)** potential of point *Z* is negative with resptect to *Y*

Suppose that to start with, terminal *X* is positive and *Z* is negative with respect to *Y*. In this condition, diode D_1 will conduct but D_2 will not conduct. The current will flow through the load from *B* to *Y* and the output voltage across R_L is as shown in Fig 29.5(a). During the next half cycle, terminal *X* will be negative and *Z* will be positive. Under this condition, diode D_2 conducts and current will again pass through the load resistance in the same direction, that is from *B* to *Y* . The corresponding waveform is shown in Fig. 29.5(b).And the net output across R_L is pulsating , as shown in Fig. 29.5(c).

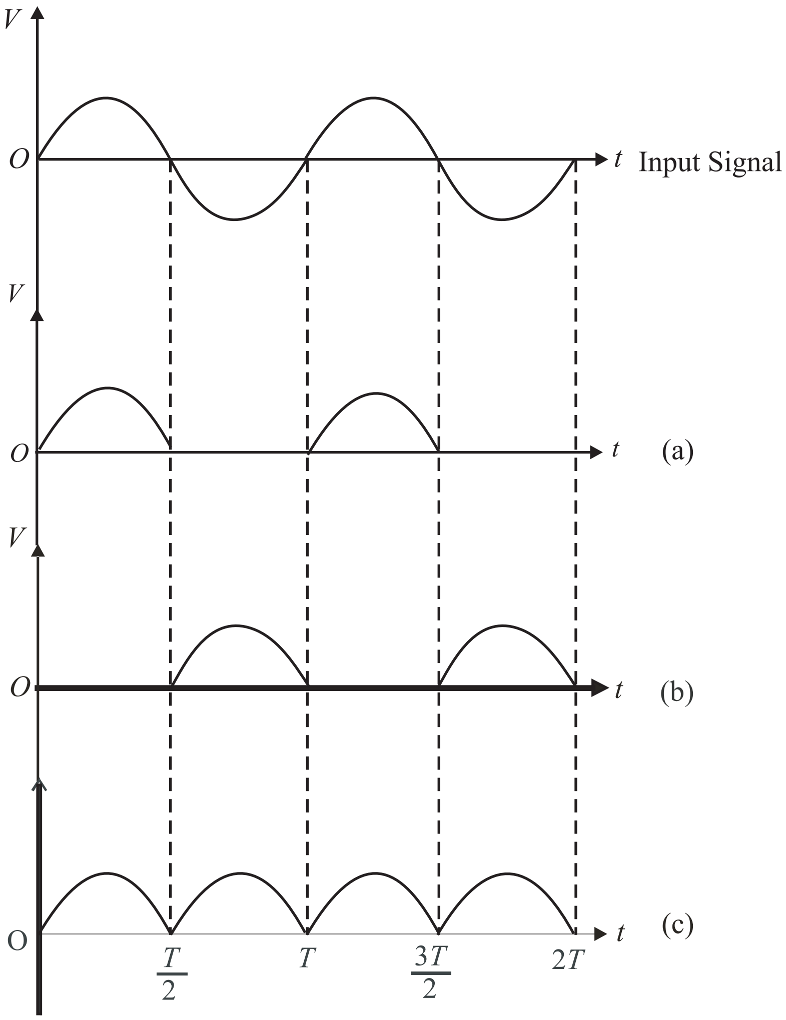


Fig. 29.5 : Voltage across R_L **a)** when D_1 conducts, **b)** D_2 conducts, **c)** net output of full wave rectifier

Since current through the load now flows over the entire cycle of the sine wave, this is called full-wave rectification. The dc voltage V_{dc} and dc current I_{dc} are given by

$$V_{dc} = 2 \times V_m/\pi \tag{29.3}$$

and

$$I_{dc} = \frac{V_{dc}}{R_L} = \frac{2V_m}{\pi R_L} \tag{29.4}$$

Note that the unidirectional current flowing through the load resistance after full-wave rectification pulsates from maximum to minimum (zero) and is not useful for any practical application. To reduce the fluctuating component and obtain more steady current, we filter the pulsating part. You may be eager to know as to how do we achieve this. Let us now discover answer to this important question.

Filtering

We recall that impedance offered by a capacitor to the flow of ac depends on its freqeuncy. Therefore, a capacitor C connected across the load resistance, as shown in Fig. 29.6, filters out high frequency component.

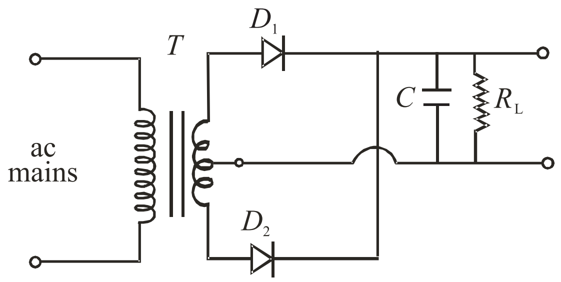


Fig. 29.6 : Circuit diagram for *capcitor*-filter in full-wave rectification

The capacitor gets charged to nearly maximum potential V_m when diode D_1 conducts for period $t = T/4$. When the current tends to decrease for $T/4 < t < T/2$, the capacitor discharges itself and tries to maintain current through the load, reducing fluctuations considerably, as shown in Fig. 29.7. The larger the value of capacitor and the load resistance, the lower will be the fluctuations in the rectified dc. The capacitor C connected across the load to reduce fluctuations is called a *filter capacitor*. In a power supply, we use LC and C-L-C (or π) filters to reduce the rippling effect. You will learn about these in detail in your higher classes.

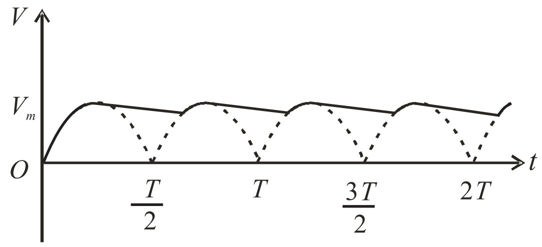


Fig 29.7: Output voltage when capacitor is used to filter ac

Special p-n junction, called Zener diode, acts as voltage regulator in reverse bias. You will now study about it.

In high quality power supplies combination of inductors and capacitor $L - C - L$ *or* $C - L - C$ is used. Depending on the way, these components are connected these filters are called '*T*' or 'Π'.

29.1.2 Zener Diode as a Voltage Regulator

The half-and full-wave rectifiers with filters are the simplest type of power supplies. These provide almost pure dc but have one deficiency. When load current is increased by decreasing resistance, the output voltage drops. This is because, when large current is drawn, the filter capacitor gets discharged more and its voltage across the load resistor reduces. Similarly, if the ac input changes, the dc output voltage also varies. Obviously, a supply with varying output voltage affects the performance of different devices being operated with it. For example, if we operate an amplifier, the quality of sound reproduced by it will get deteriorated.

To remove this deficiency, a Zener diode is used with simple power supplies which gives constant dc voltage. Such a circuit is called regulated power supply.

The Zener regulated voltage supply circuit is shown in Fig. 29.8. It consists of a Zener diode with breakdown voltage V_z. This will be equal to the stabilized output voltage V_O A suitable series resistance R_s is included to control circuit current and dissipate excess voltage. The anode of Zener diode is connected to the negative terminal of input supply, and the cathode is connected in series with R_s to positive terminal of input supply, that is, the Zener is connected in reverse bias condition. The load resistance is connected across the Zener diode. The Zener regulator will only operate if the input supply voltage to the regulator, V_i is greater than V_z . After breakdown, the voltage across it remains nearly constant and is independent of the current passing through it. The current I_s flowing passing through R_s is given by the equation

$$I_s = (V_i - V_z)/R_s \qquad (29.5)$$

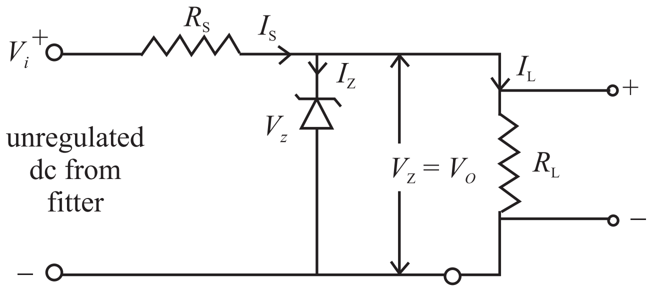


Fig. 29.8 : Zener diode as a stabilizer

This current divides in two parts: the Zener current I_z and load current I_L. Applying Kirchoff's law, we can write

$$I_s = I_z + I_L$$

or $\quad I_z = I_s - I_L \qquad (29.6)$

For Zener diode to operate, some current $I_{Z_{min}}$ should always flow through it. Therefore, the load current I_L should always be less than the main current I_s. Typical value of $I_{Z_{min}}$ may range from 5 mA to 20 mA.

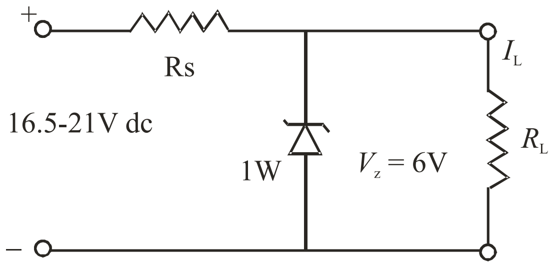
If load current is zero, the entire I_s will pass through Zener diode and output voltage V_O will be equal to V_z. When some load current is drawn, say I_L, the Zener current will decrease by the same amount but the output voltage will remain V_z. Similarly, if the ac main voltage increases or decreases, the input voltage, V_i will increase or decrease accordingly. It will result in change of I_s given by Eqn.(29.5). Due to change in I_s, the change in V_i will appear as a drop across the series resistance R_s. The Zener voltage V_z and hence V_O will remain unchanged. Thus we see that the output voltage has been stabilized against the variations in the current and the input voltage.

The power dissipation in Zener diode is given by the relation

$$P_d = V_z \times I_z \qquad (29.7)$$

This dissipation should not exceed the maximum power dissipation rating recommended by the manufacturer for Zener diode. Let us now understand the design of a Zener regulated power supply with one example.

Example 29.1: The load current varies from 0 to 100 mA and input supply voltage varies from 16.5 V to 21 V in a circuit. Design a circuit for stabilized dc supply of 6 V.



Solution: We choose a Zener diode of 6 V. Let $I_{Z_{min}}$ be 5 mA. The maximum current will flow through the Zener when there is no load current. Its magnitude will be (100+5) mA= 0.105A.

The value of R_s is determined by the minimum input voltage and maximum required current:

$$R_s = \frac{V_{Z_{min}} - V_z}{I_{max}} = \frac{16.5\text{V} - 6\text{V}}{105\text{mA}} = 100\ \Omega$$

The current through the Zener diode will be maximum when the input voltage is maximum, that is 21 V and $I_L = 0$. Therefore, the maximum Zener current I_{max} = (21V– 6V)/100 Ω $\simeq$ 0.15 A.

The maximum power dissipation in the diode is 6V × 0.15A= 0.9W.

It means that we should use a Zener diode of 6 V, 1 W and resistance Rs. of 100Ω. It should be connected in the circuit as shown above. It will give a stable output of 6 V for the specified ranges of load and input variation.

INTEXT QUESTIONS 29.1

1. Draw a circuit of full-wave rectifier with a filter capacitor.
2. What will be the output voltage, if you connect a Zener diode in forward bias instead of reverse bias in the regulator circuit of Example 29.1?

29.2 TRANSISTOR APPLICATIONS

You learnt the working principle of transistor in detail in the last lesson. Normally, the collector is reverse biased and no current flows in collector-emitter circuit. If we pass a very small current in the base circuit, a very large current starts flowing in the collector circuit. This property has made a transistor indispensable for vast electronic applications. But here we have discussed its applications as an amplifier, as a switch, and as an oscillator (frequency generator).

29.2.1 Transistor as an Amplifier

An electrical signal is voltage or current, which is coded with some useful information. For example, when we speak in front of a microphone, its diaphragm vibrates and induces a very small voltage in its coil, depending on the intensity of sound. This induced voltage appears as a weak signal and can not operate a loudspeaker to reproduce sound. To make it intelligible, it is fed into a device called amplifier. The amplifier increases the level of input signal and gives out magnified output. If V_i is the input signal voltage fed to the amplifier and V_O denotes the amplified output, their ratio is called *voltage gain.*

i.e., $$A_V = \frac{V_O}{V_i} \qquad (29.8)$$

Similarly, we can define the current gain and power gain as

$$A_I = \frac{i_O}{i_i} \qquad (29.10)$$

$$A_p = \frac{P_O}{P_i} \qquad (29.11)$$

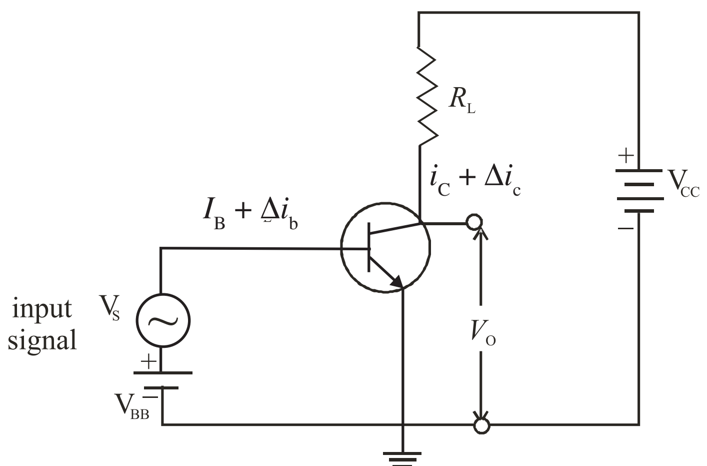


Fig. 29.9: Basic amplifier circuit using a *n-p-n* transistor in *CE* mode

The circuit for transistor as an amplifier is shown in Fig. 29.9. Here an *n-p-n* transistor is used in CE mode. Its collector is reverse biased through the load resistance R_L by the battery V_{CE}. When a base current I_B flows, some collector current I_C will start flowing. On decreasing I_B, a stage will be reached when I_C becomes almost zero. This is the lower limit of variation of I_B. Similarly, on increasing I_B again, a stage of saturation is reached and I_C stops increasing. This corresponds to the upper limit of variation of I_B. For faithful amplification of input signal, a base current equal to the mean of these two limiting values of I_B is passed through the base by forward biasing it with battery V_{BB}. We can choose the operating point in the centre of linear operating range of the transistor. This is called biasing of the base. A signal source providing an input signal υ_S is connected in series with υ_{BB}.

Due to addition of oscillating signal voltage υ_S to υ_{BB}, the base current changes by an amount Δi_b around the dc biasing current I_B. The signal voltage is kept low so that the signal current Δi_b if added and subtracted from I_B does not cross the upper and lower limits of the base current variation. Otherwise, the transistor will go into cut off or saturation region and the amplified output will be highly distorted and noisy. Note that signal current

$$\Delta i_b = \upsilon_S / r_i \qquad (29.12)$$

where r_i is the input impedance. This change in base current Δi_b results in a large change in collector current, say Δi_c given by

$$\Delta i_c = \beta \Delta i_b = \beta \upsilon_s / r_i \qquad (29.13)$$

where β is the ac current amplification factor, equal to $\Delta i_c / \Delta i_b$. From (Eqn. 29.13) we get

$$\upsilon_s = \Delta i_c \times r_i / \beta \qquad (29.14)$$

By applying Kirchhoff's law to the output circuit in Fig. 29.9, we have

$$V_{CC} = V_{CE} + I_C R_L \qquad (29.15)$$

On differentiating Eqn. (29.15), we get

$$dV_{CC} = dV_{CE} + dI_C \times R_L \qquad (29.16)$$

Since V_{CC} is constant, $dV_{CC} = 0$. Therefore, we get

$$dV_{CE} = -\,dI_C \times R_L$$

But dV_{CE} is the change in output $\Delta \upsilon_0$ and dI_C in i_c. Therefore,

$$\Delta \upsilon_0 = -\,\Delta i_c \times R_L$$

The voltage gain A_v of the amplifier is given by

$$A_V = \upsilon_0 / \upsilon_S = -\,(\Delta i_c \times R_L)/(\Delta i_c \times r_i / \beta)$$

$$= -\beta \times R_L / r_i \qquad (29.17)$$

The ratio β/r_i is called **transconductance** of transistor and is denoted by g_m. Hence Eqn. (29.17) can be written as

$$A_V = -g_m \times R_L \quad (29.18)$$

The negative sign indicates that input and output are in opposite phase, i.e. they differ in phase by 180°. The power gain is given by

$$A_P = A_I \times A_V = \beta \times A_V \quad (29.19)$$

Note that power gain does not mean that the law of conservation of energy is violated in an amplifier. The ac power output of the amplifier is more than the ac input signal power but this gain is achieved at the cost of dc power supplied by the voltage source.

John Bardeen
(1908 – 1991)

John Bardeen is the only researcher in history of science who received two Nobel Prizes in Physics. He was born in Madison, Wisconcin USA, in a highly educated family. He was so bright a kid that his parents moved him from third grade to Junior high school. He did his graduation in Electrical Engineering. But, he also had to struggle for his career. After spending three years as geophysicist with Gulf Oil Company, he went to Princeton for his Ph.D. in Mathematical Physics. After a brief stint at Harvard and Minnesota and in Naval Ordnance Labs, he joined William Shockley's research group at Bell Laboratories. With Walter Brattain, he devloped the first transistor for which Bardeen, Brattain and Shockley were conferred the 1956 Nobel Prize in Physics.

Bardeen shared his second Nobel in 1972 with Leon C Cooper and R Schieffer for their theoretical work on superconductivity.

INTEXT QUESTION 29.2

1. For a *CE* mode amplifier, υ_i is 20 mV and υ_o is one volt. Calculate voltage gain.

2. The P_0 of an amplifier is 200 times that P_i. Calculate the power gain.

3. For a CE amplifier, $R_L = 2000\ \Omega$, $r_i = 500\ \Omega$ and $\beta = 50$. Calculate voltage gain and power gain.

29.2.2 Transistor as a switch

In day-to-day life, we use electrical switches to put the gadgets like lamps, fans, machines on or off manually. Note that the switch has two distinct states, viz on and off. In electronics, we come across situations where we need to apply an input to some device in the form of two distinct voltage levels. This is as if we were operating a switch. When switch is on, one voltage level is applied but when switch is off, the other one is applied. Typically, such voltage levels are used in computers, where digital signals are employed. This is done by using a trasistor in the non-linear region of its operation. In the transister characteristics shown in Fig 29.10, we see two extreme regions: *cut-off region and saturation region*. The (jagged) region below the zero base (I_B = 0) signifies the *cut off* regions. The transistor does not conduct and entire supply voltage V_{CC} appears across the transistor between the collector and the emitter (V_{CE}). That is, the output voltage at the collector is V_{CC}.

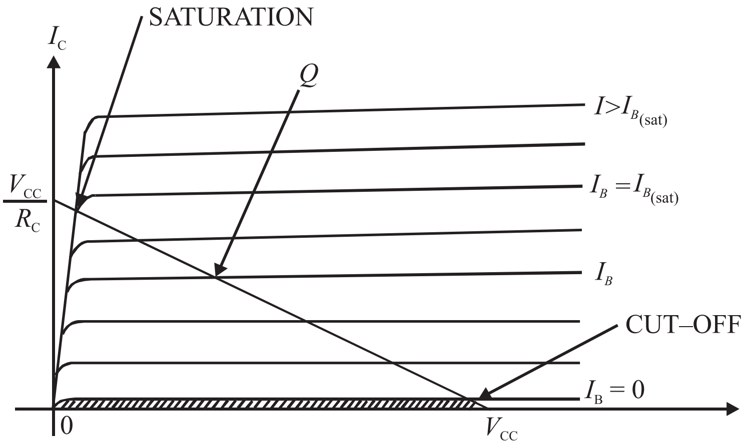


Fig. 29.10 : Transistor output characteristics

When the base current I_B is greater than its saturation value, the transistor conductor fully and collector-emitter voltage V_{CE} is almost zero. In such a case, the output voltage obtained between collector and ground is zero and entire voltage drop appears across R_L. That is , the collector current $I_C = \frac{V_{CC}}{R_L}$.

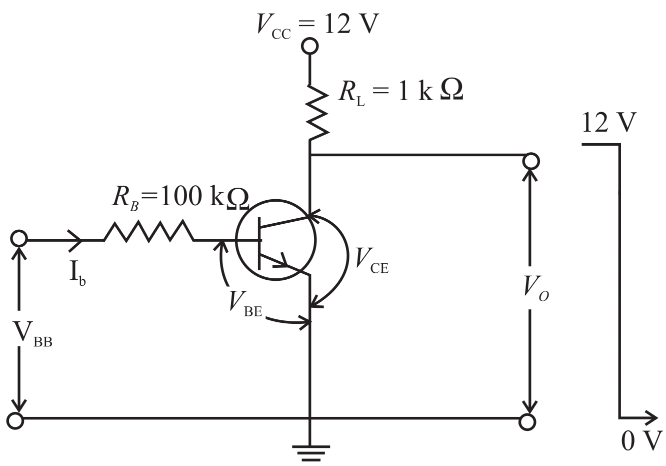


Fig. 29.11: Transistor as a switch

Fig 29.11 shows a typical circuit of transistor as a switch. The control signal for switching the transistor on or off is given in the form of V_{BB}. For the input loop, we can write

$$I_B R_B + V_{BE} - V_{BB} = 0$$

When $V_{BB} = 0$, we get

$$I_B = -\frac{V_{BE}}{R_B} \qquad (29.20)$$

Since I_B is less than zero, the transistor is cut off, and

$$V_0 = V_{CC} \qquad (29.21)$$

If $V_{BB} = 5V$, and $V_{BE} = 0.7$ V for the chosen transistor, from Eqn. (29.20) we get

$$I_B\ (100\ k\Omega) + 0.7V - 5\ V = 0.$$

$$\therefore \qquad I_B = \frac{5V - 0.7V}{100k\Omega} = 43\ \mu A$$

For normal transistors, this value of base current is enough to drive the transistor to full saturation. In this case, $V_O = V_{CE_{sat}} = 0$ and the collector current

$$I_C = \frac{V_{CC}}{R_L} = \frac{12V}{1K\Omega} = 12mA.$$

This kind of switch can also be used as an indicator in displays. For example, if we connect an LED is series with the collector resistor, as shown in Fig 29.12, the collector current drives the LED on for high (+5V) input, and it lights up. Whenever input is zero, the LED is off because no collector current flows through the circuit.

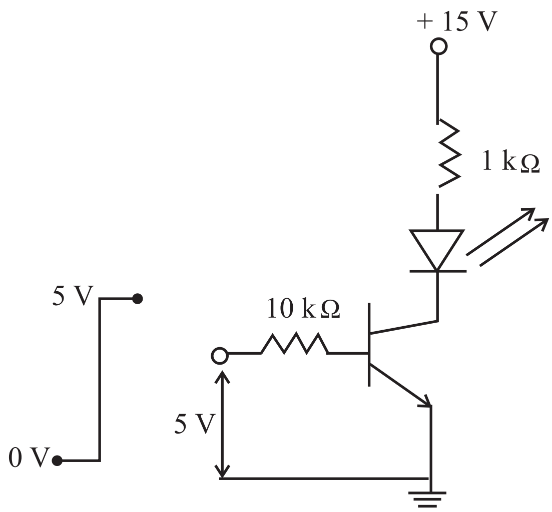


Fig. 29.12: LED indicator using transistor switch

Another major application of transistors is to generate an oscillating signal of desired frequency. This is done by a special circuit called an **oscillator**. The oscillators find many applications, particularly in radio transmitters to generate the carrier wave frequency. These are also used in clock generators, electronic watches and computers etc. There are various types of oscillators. We here discuss a typical oscillator circuit using a transistor.

29.2.3 Transistor as an Oscillator

An electronic oscillator is a device which generates continuous electrical oscillations. In a simple oscillator circuit, a parallel *LC* circuit is used as resonant circuit and an amplifier is used to feed energy to the resonant circuit. It can generate frequencies from audio to radio range depending on the choice of *L* and *C*.

We know that when a charged capacitor is connected across an inductor, the charge oscillates. But due to loss of energy by radiation and heating of wires, the energy is lost and the amplitude of oscillations decays with time. To build a sinusoidal oscillator, where the oscillations are sustained (i.e. they do not decay), we need an amplifier with positive feedback. The basic idea is to feed a part of output signal in input signal. By adjusting the gain of the circuit and the phase of the feedback signal, energy dissipated in each cycle is replenished to get sustained oscillations of desired frequency.

Schematically, we can depict an oscillator to be made up of two main blocks: an amplifier with gain *A*, and a feedback circuit with feedback factor β, as shown in Fig 29.13.

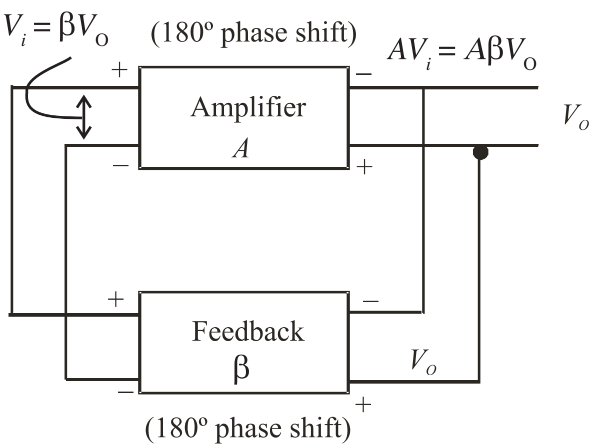


Fig. 29.13: Schematic diagram of an oscillator

In case, $A\beta < 1$, V_O decreases continuously. On the other hand, if $A\beta > 1$, V_0 increases gradually.But if $A\beta = 1$, we get constant value of V_0 leading to sustained oscillations.

Now, we consider a *CE* amplifier, like the one discussed in Sec. 29.2.1. It has 180° phase difference between the input and output, i.e. it has negative gain ($-A$). To keep the total feedback gain $A\beta = 1$, we require that β is also negative; equal to $-A^{-1}$. That is, it is necessary to introduce a phase shift of 180° in the feedback circuit as well.

In Fig. 29.14, we have shown a circuit diagram of an oscillator using *LC* tank circuit and a transistor amplifier in *CE* mode. This is called Colpitt's Oscillator.

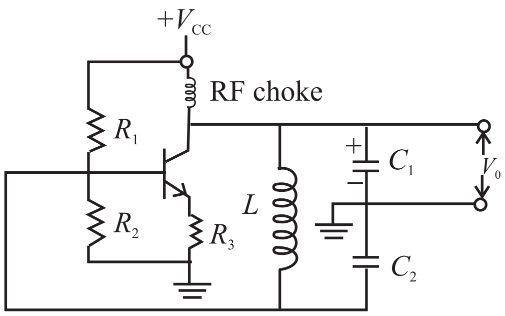


Fig. 29.14 : Colpitt's Oscillator

In this circuit C_1, C_2 and L form the tank circuit. The oscillating current is generated in this circuit, which is at its resonant frequency. The output is obtained across C_1, the feedback is provided across C_2 connected to the base of the transistor amplifier in *CE* mode. In this case 180° is introduced by the amplifier and another 180° phase shift is provided by the capacitor C_2 which is connected between ground and other end of the inductor coil. Hence, the total loop gain is positive. When the gain of transistor amplifier is sufficiently large at the resonant frequency, we obtain sustained oscillations at the output.

29.3 LOGIC GATES

In electronics, we come across mainly two types of waveforms. The information carried by these waveforms is called signal. When the signal takes any value within a range of amplitude at any instant of time, it is called a continuous signal. When the signal takes the value only at certain times, it is called a discrete signal. When the signal takes only particular finite number of amplitude values, it is called a **digital signal** (Fig. 29.15).

The digital signal varies in steps and typically has only two widely separated values '0' and '1'. These are called bits. Normally 0V corresponds to bit '0' and 5 V corresponds to bit '1'. Since the levels are so widely separated, any noise riding on the signal within the range of almost 2V, [(0V + 2 V) for level '0' and (5V – 2V) for level '1', does not affect the signal value, Hence these signals are immune to noise. The signals used in a computer are digital. The information is coded in the form of digital signals by a series of bits arranged in different order. Each bit is a pulse of fixed time duration.

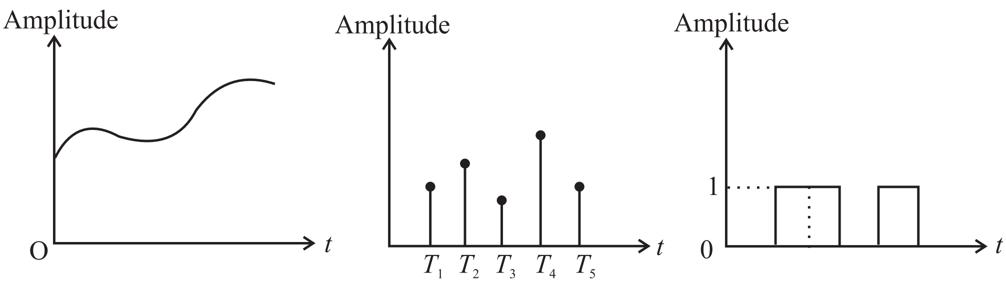


Fig. 29.15: a) continuous signal, **b)** discrete signal, and **c)** digital signal

Different mathematical operations can be performed on the digital signal. The mathematics governing these operations is called Boolean algebra.

In Boolean algebra, the basic operations are addition and multiplication. If it is a digital data that takes value 0 or 1, the following identities hold:

$$A \times 0 = 0 \qquad (29.22)$$

$$A + 1 = 1 \qquad (29.23)$$

The circuits which perform these operations are called *logic gates*. Let us now learn about basic logic gates.

29.3.1 Basic Logic Gates

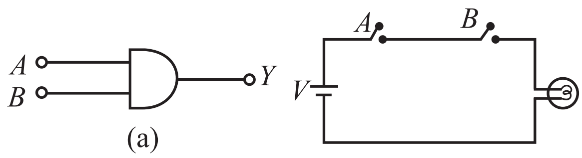
Logic gates are devices which have one or more inputs and one output. They give different output when the input bits differ in their arrangement. The output produced by these gates follows the laws of Boolean logic. There are three basic types of logic gates :

1. AND Gate, 2. OR Gate, 3. NOT Gate

These gates perform multiplication, addition and inversion (negation) operations, respectively. Let us now learn the working of these logic gates.

1. AND Gate

An AND gate can have two or more inputs but only one output. The logic symbol of a two input AND gate is given Fig 29.16(a). We can understand the behaviour of an AND gate by considering a number of electrical switches connected in series. For examples, switches A and B are two inputs of the gate and the bulb gives the output Y. The ON switch stands for logic input '1' and OFF switch stands for logic input '0'. In this case, the bulb will glow only if it is connected to the supply voltage. This will happen only if both A and B switches are simultaneously ON (or '1'). The behaviour of output Y at various values of A and B is shown in Table in Fig. 29.16(c). This table is called *Truth Table*.



A	B	Y
0	0	0
0	1	0
1	0	0
1	1	1

(c)

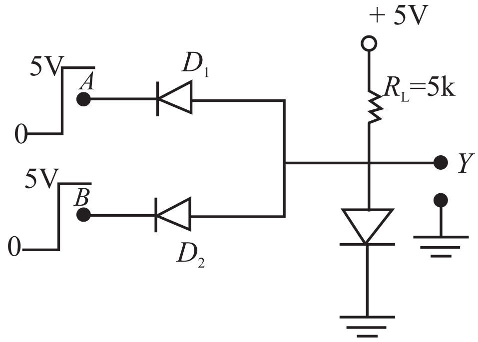


Fig 29.16: a) Symbol of AND gate, **b)** switch implementation of AND gate, **c)** Truth Table of AND Gate, and d) diode implementation of AND gate.

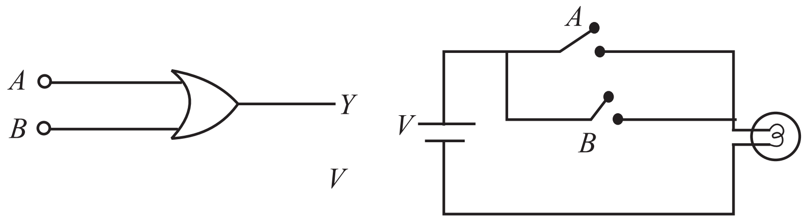
The Boolean expression for the AND operation is represented as

$Y = A.B = AB = A \times B$ and read as *A* AND *B*.

Realization of AND Gate : The logic gate realized by using diodes is called a DDL Gate (Diode–Diode Logic Gate). The diode implementation of a two-input AND gate is shown in Fig.29.16 (d). The anodes of two diodes D_1 and D_2 connected in parallel are forward biased by a 5 V battery through a 5 kΩ resistance. The output is taken from the anode. Cathode wires *A* and *B* serve as input terminals. When either *A* or *B* or both the terminals are grounded, the respective diode will conduct and a potential drop will develop across the resistance and output will be 0.7 V, i.e. logic ‘0’. When both the terminals are connected to 5V (i.e. for input 1, 1), neither of the diodes will conduct and output will be 5 V, i.e. logic ‘1’

2 OR Gate

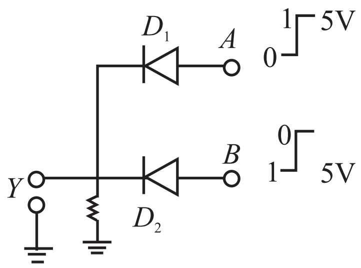
The OR gate can have two or more inputs and only one output. The logic symbol of a two input OR gate is given in Fig 29.17(a). We can explain the behaviour of an OR gate with the help of a number of electrical switches connected in parallel. For a two input OR gate, two switches are connected, as shown in Fig.29.17(b). The switch *A* and *B* are the two inputs of the gate and the bulb gives output *Y*. The ON switch stands for logic input ‘1’ and OFF switch stands for logic input ‘0’. The glowing bulb stands for logic output ‘1’ and the non-glowing bulb for logic output ‘0’. In this case, when either *A* OR *B* or both the switches are ON, the supply voltage reaches the output and the bulb glows. The input-output correlation for an OR gate is shown in the Truth Table given in Fig. 29.17(c).



(a) (b)

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	1

(c)



(d)

Fig 29.17: a) Symbol of OR gate, **b)** switch implementation of OR gate, **c)** Truth Table of OR gate, and **d)** diode implementation of OR gate

The Boolean expression for an OR operation is represented as

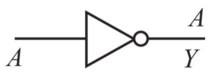
$Y = A + B$ and read as A or B.

Realization of OR Gate: The diode implementation of a two-input OR gate is shown in Fig. 29.17 (d). The cathodes of diodes D_1 and D_2 connected in parallel are grounded through a 5 kΩ resistance. The output is taken from the cathode and the two anode wires A and B serve as input terminals. When either A or B or both the terminals are connected to the positive terminal of the 5 V battery, the respective diode/diodes will conduct and potential at the output will be bout 5V i.e. logic '1'. When both the switches are open, output will be 0 V i.e. logic '0'.

3 NOT Gate

Another important gate used in digital signal handling is the **NOT** gate, which inverts the signal, i.e., if input is '1' then output of NOT gate is '0' and for '0' input, the output is '1'.

The symbol for NOT gate is shown in Fig. 29.18(a). The Truth Table of NOT gate is shown in fig. 29.18(b).



(a)

A	$Y = A$
0	1
1	0

(b)

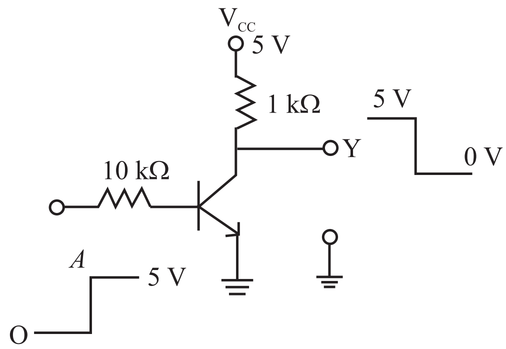


Fig. 29 18: (a) Symbol of NOT gate, (b) Truth Table of NOT gate, and (c) circuit implementation of NOT gate

The circuit to implement a NOT gate is identical to that used for a transistor as a switch. This is shown in Fig. 29.18(c). When input A is at ' 0' level, transistor is off and the entire V_{CC} voltage (5V) appears at the output Y. When input A is '1' (5V), the transistor conducts and output voltage Y is '0'.

The inversion operation is indicated by a bar on the top of the symbol of the input e.g. in the Truth Table we can write, Y = NOT (A) = $\overline{A}$

So far we have discussed basic logic gates. You may now ask: Can we combine these to develop other logic gates? You will discover answer to this question in the following section.

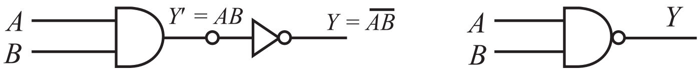
29.3.2 Combination Logic Gates

Two most important gates formed by combination of logic gates are (1) NAND [NOT+AND] and (2) NOR [NOT+OR] gates. In digital electronics, a NAND gate or a NOR gate serves as a building block because use of multiple number of either of these gates allows us to obtain OR, AND and NOT gates. For this reason, these are called universal gates. Let us now learn about combination logic gates.

1. NAND Gate

The NAND Gate is obtained by combining AND gate and NOT gate, as shown in Fig. 29.19 (a). Here the output Y of AND gate is inverted by the NOT gate to get the final output Y. The logic symbol of a NAND gate is shown in Fig. 29.19(b). The Truth Table of a NAND gate is given in Fig. 29.19(c). It can be obtained by inverting the output of an AND gate. The truth table of a NAND gate shows that it gives output '1' when at least one of the inputs is '0' The Boolean expression of a NAND operation is represented as

$$Y = A.B = A \times B = AB$$



(a) (b)

A	B	$Y' = AB$	$Y = \overline{AB}$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

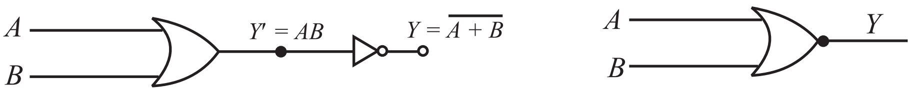
(c)

Fig. 29.19 : a) NAND as combination logic gate, **b)** symbol of NAND GATE, and **c)** Truth Table of a NAND gate

2. NOR Gate

The NOR gate, obtained by combining an OR gate and NOT gate, is shown in Fig. 29.20(a) Here the output of OR gate, Y', is inverted by the NOT gate to get the final output Y. The logic symbol of a NOR gate is given in Fig. 29.20(b). The Truth Table of a NOR gate given in Fig. 29.20(c), can be arrived at by inverting the output of an OR gate. The Truth Table of a NOR gate shows that it gives output ‘1’ only when both the inputs are ‘0’

The Boolean expression for a NOR operation is represented as $Y = \overline{A + B}$.



(a) (b)

A	B	$Y' = A+B$	$Y = \overline{A+B}$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

Fig. 29.20 : a) NOR as combination logic gate, **b)** symbol of NOR gate, and **c)** Truth Table of NOR gate

As mentioned earlier, the NAND and NOR gates are basic building blocks of all the logic gates. Let us now see, how we can obtain the three basic gates AND, OR and NOT by using NAND gates.

29.3.3 Realization of Basic Gates from NAND Gate

The NAND gate is considered to be the universal gate because all other gates can be realized by using this gate.

(a) Realization of a NOT gate : If two input leads of a NAND gate are shorted together, as shown in Fig. 29.21, the resulting gate is a NOT gate. You can convince yourself about this by writing its truth table.

Here we have $A = B$

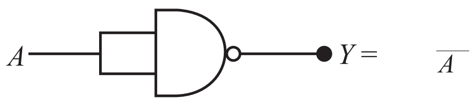
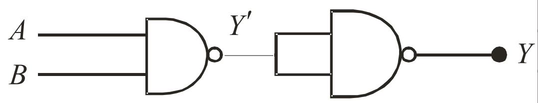


Fig. 29.21 : NAND gate as NOT gate

(b) **Realization of an AND gate** : The AND gate can be realized by using two NAND gates. The output of one NAND gate is inverted by the second NAND gate used as NOT gate as shown in Fig 29.22(a). The combination acts as an AND gate, as is clear from the Truth Table given in Fig. 29.22(b).



A	B	$Y' = \overline{AB}$	$Y = AB$
0	0	1	0
0	1	1	0
1	0	1	0
1	1	0	1

(b)

Fig. 29. 22: a) NAND gates connected to implement AND gate and **b)** Truth Table of AND gate using NAND gate

c) **Realization of an OR gate :** The OR gate can be realized by using three NAND gates. Two NAND gates are connected as inverters and their outputs are fed to the two inputs of a NAND gate, as shown in Fig. 29.23. The combination acts as an OR gate.

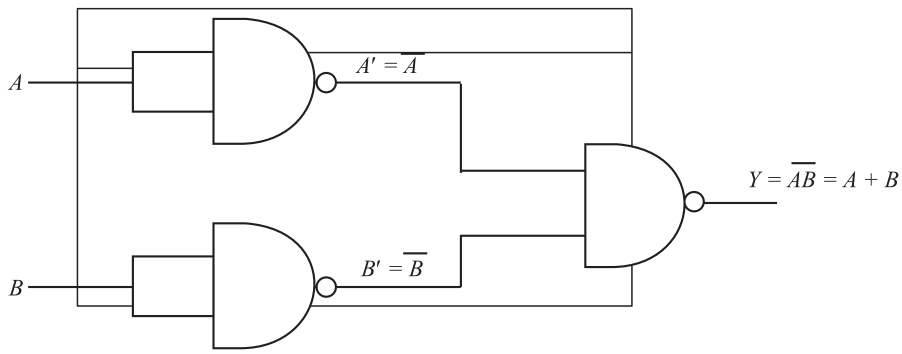


Fig. 29.23 : Three NAND gates connected as OR gate

INTEXT QUESTIONS 29.3

Complete the following table from Fig. 29.23 to prove that it is an OR gate.

A	*B*	*A′*	*B′*	*Y*
0	0	–	–	–
0	1	–	–	–
1	0	–	–	–
1	1	–	–	–

WHAT YOU HAVE LEARNT

- A *p-n* junction diode can be used as a rectifier to convert ac into dc.
- A half-wave rectified dc contains more ac component than the full-wave rectified dc.
- A Zener diode stablizes the output of a power supply.
- In a stabilizer, the Zener diode dissipates more power when the current taken by the load is less.
- For amplification, a transistor needs input current.
- Transistor can be used as a switch by biasing it into saturation and cut-off regions.
- There are three basic logic gates: AND,OR and NOT.
- NAND gate is a universal gate because it can be used to implement other gates easily.

TERMINAL QUESTION

1. Why the Peak Inverse Voltage (PIV) of a *p-n* junction diode in half-wave rectifier with filter capacitor is double of that without the capacitor?
2. Explain how a Zener diode helps to stabilize dc against load variation.
3. What should be the range of variation of amplitude of input signal for proper working of an amplifier?
4. Draw a circuit using diodes and transistors to implement a NOR gate.

ANSWERS TO INTEXT QUESTIONS

29.1

1. See Fig.29.6

2. In case of full wave rectifier, both diodes D_1 and D_2 charge C to maximum voltage of V_{max} in alternate half cycles. Hence, the PIV of the diodes should be $2 \times V_{max}$.

3. $R_z = 100\,\Omega$, $R_s = 100\,\Omega$ and $R = R_z + R_s = 200\,\Omega$
Hence,

$$I = \frac{21}{200} = 0.105\text{A}$$

and $V = IR = 0.105 \times 100$
$= 10.5\text{V}$

29.2

1. $|A_v| = \frac{V_0}{V_i} = \frac{1\text{V}}{20\text{ mV}} = 50$.

2. $A_p = \frac{P_0}{P_i} = 200$.

3. $$|A_v| = \frac{\beta \times R_L}{r_i} = \frac{50 \times 2000\Omega}{500\Omega} = 200$$

$$A_p = \beta A_v = 50 \times 200 = 10000.$$

29.3

A	B	A'	B'	Y
0	0	0	0	0
0	1	1	1	1
1	0	1	0	1
1	1	1	1	1

30

312en30

COMMUNICATION SYSTEMS

Communication is a basic characteristic of all living beings. Communication entails transmitting and receiving information from one individual/place to another. In the world of animals, communication is made by mechanical, audio and chemical signals. You may have observed how sparrows begin to chirp loudly on seeing an intruder, who can put their life in danger. However, human beings are blessed with very strong means of communication – speech. We can express what we see, think and feel about whatever is happening around us. That is to say, we use sound (an audible range, 20Hz - 20kHz) and light (in visible range, 4000 Å – 7000 Å), apart from mechanical (clapping, tapping) and opto-mechanical signals (nodding, gesturing), for communication. You must realise that language plays a very significant role in making sense out of spoken or written words.It comes naturally to us. Prior to the written alphabet, the mode of communication was oral. The second era of communication began with the invention of printing press. Invention of the telegraph in the early nineteenth century marked the beginning of the third stage. Revolutionary technological developments enabled as rapid, efficient and faithful transfer of information. Using tools and techniques such as telegraph, fax, telephone, radio, mobiles, satellites and computers, it is possible to communicate over long distances. The oceans and mountain ranges no longer pose any problem and the constraints of time and distance seem to be non-existent. On-line learning, (education), publishing (research), banking (business) which were topics in science fiction not too long ago, are now routine activities. In fact, combination of computers with electronic communication techniques has opened a very powerful and fertile field of information and communication technologies (ICT_s).

Have you ever thought about the technology that has made all this development possible? You will discover answers to this question in this lesson.

OBJECTIVES

After studying this lesson, you will be able to:

- list the elements of a long distance communication system;
- explain the terms analogue and digital signals;
- describe how electromagnetic waves act as carriers of information;
- Specify the bandwidth of signals (speech, TV and digital data);
- list variouis transmission media and state bandwidths specific to them;
- explain importance of ground, sky and space wave propagation;
- state need for modulation; and
- explain production and detection of amplitude modulation wave.

30.1 A MODEL COMMUNICATION SYSTEM

Communication systems endeavour to transmit information from

– one to one, i.e., point-to-point communication;
– one to many, i.e., broadcast communication; and
– many to many, i.e., telephone conference call or a chat room.

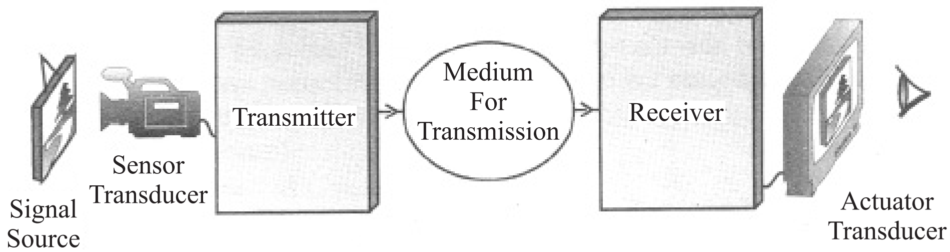


Fig. 30.1 : A schematic arrangement for the communication system.

In a typical modern day communication system, the information is in the form of electrical signals (voltage or current), spread over a range of frequencies called the signal **bandwidth**. (Some **noise** gets added to the signal and tries to obscure the desired information.) For scientific analysis of any system, we model the system into its basic components. You will now learnt about these.

30.1.1 Elements of a Communication System

Refer to Fig. 30.1. It shows building blocks of a typical communication system. As may be noted, the essential elements of a communication system are:

- a source of signal, a sensor transducer and a **transmitter,** which launches the signal carrying information,
- an intervening **medium/channel** to guide and carry the signal over long distances, and
- a signal **receiver** and an actuator transducer to intercept the signal and retrieve the information.

30.2 TYPES OF SIGNALS – ANALOGUE AND DIGITAL

You now know that communication of information involves use of signals, which are classified on the basis of their origin and nature. Accordingly we have

- continuous time (analog) end discrete time (digital) signals;
- coded and uncoded signals;
- periodic and aperiodic signals;
- energy and power signals; and
- deterministic and random signal.

Of these, we will consider only analog and digital systems. The sound produced by human being in conversation/interaction or photograph are converted into continuously varying electrical analog signal [Fig. 30.2(a)]. But in modern electronic communiction systems, these are converted into discrete form, which has finite values at different instances of time and zero otherwise [Fig. 30.2 (b), (c)] form Fig. 30.2, you will note that the waveforms used to represent correspond to a particular frequency and are periodic; while one of these is sinusoidal, the another is pulsed. In fact these may be viewed as a sub-class of sine and square waveforms.

Information can be packaged in both analog (or continuous) and digital (or discrete) forms. Speech, for example, is an analog signal which varies continuously with time. In contrast, computer files consist of a symbolic "discrete-time" digital signal.

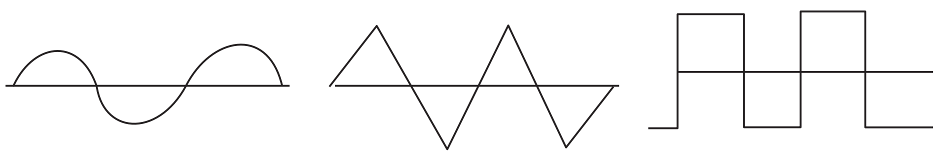


Fig. 30. 2 : *Examples of (a) continuous (sinusoidal) and (b) discrete signals.*

In the digital format, signals are in the form of a string of **bits** (abbreviated from **bi**nary digi**ts**), each bit being either 'ON' or 'OFF' (1 or 0). The **binary** system refers to a number system which uses only two digits, 1 and 0 (as compared to the **decimal** system which uses ten digits from 0 to 9). We can convert all information-bearing signals into discrete-time, amplitude-quantised digital signals. In a compact disc (CD), the audio is stored in the form of digital signals, just as a digital video disc (DVD) stores the video digitally.

Communication systems can be either fundamentally analog, such as the amplitude modulation (AM) radio, or digital, such as computer networks. Analog systems are in general, less expensive than digital systems for the same application, but digital systems are more efficient give better performance (less error and noise), and greater flexibility.

30.3 BAND WIDTH OF SIGNALS

The most crucial parameter in communication systems is the signal bandwidth, which refers to the frequency range in which the signal varies. However, it has different meaning in analog and digital signals. While analog bandwidth measures the range of spectrum each signal occupies, digital bandwidth gives the quantity of information contained in a digital signal. For this reason, analog bandwidth is expressed in terms of frequency, i.e. Hz, the digital bandwidth is expressed in terms of bits per second (bps). The frequency range of some audio signals and their bandwidths are given in Table 30.1. Note that human speech has bandwidth of nearly four kilo hertz. The bandwidth is about 10kHz in amplitude modulated (AM) radio transmission and 15kHz in frequency molulated (FM) transmission. However, the quality of signal received from FM broadcast is significantly better than that from AM. The bandwidth of a video signal is about 4.2MHz and television broadcast channel has bandwidth of 6MHz. The bandwidth of a typical modem, a device used for communication of digital signals over analog telephone lines, are 32kbps, 64 kbps or 128 kbps.

Table 30.1: Typical audiobandwidths

Source	Frequency range(H_E)	Bandwidth (kHz)
Guitar	82–880	... 0.8
Violin	196–2794	... 2.6
Vowels (a,e,i,o,u)	250–5000	... 4.0
consonants		
Telephone signal	200–3200	... 3.0

30.3.1 Electromagnetic Waves in Communication

In communication, we use different ways to transport the electrical signal from the transmitter to the receiver. From Modules on electricity and magnetics, you may recall that current passes through a metal conductor in the form of current signal or voltage drop, through air in the form of electromagnetic radiation or converted into light signal and sent through an optical fibre. Irrespective of the mode transmission of signal is governed by the classical theory of electromagnetic wave propagation, given by Maxwell.

As the name suggests, e.m. waves consist of electric and magnetic fields, which are inseparable. An electric field varrying in time produces a space-time varrying magnetic field, which, inturn, produces electric field. This mutually supporting role results in propagation of electromagnetic waves according to e.m. laws. The pictorial representation of a plane e.m. wave is shown in Fig. 30.3.

Mathematically we can express these as $E = E_0 \sin(kz - \omega t)$ and $H = H_o \sin(kz - \omega t)$. The direct experimental evidence for the existence of e.m. waves came in 1888 through a series of brilliant experiments by Hertz. He found that he could detect the effect of e.m. induction at considerable distances from his appparatus.

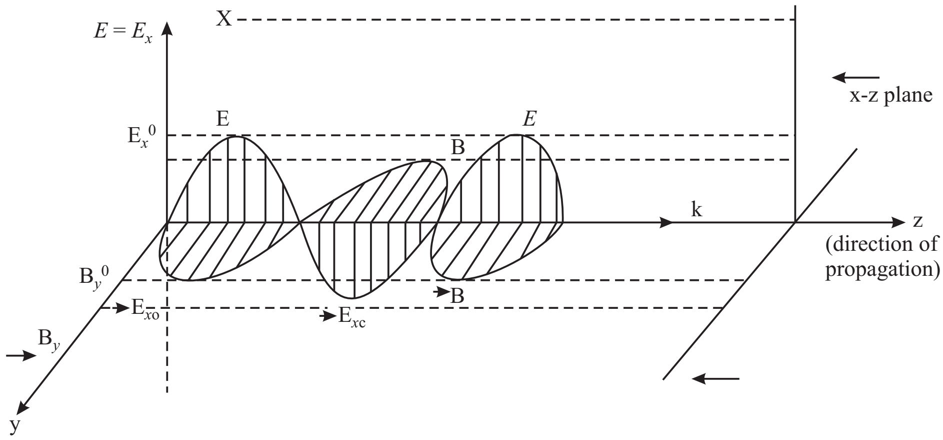
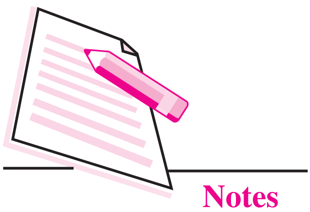


Fig. 30. 3: *Propagation of electromagnetic waves*

By measuring the wavelength and frequency of e.m. waves, he calculated their speed, which was equal to the speed of light. He also showed that e.m. waves exhibited phenomena similar to those of light. The range of wavelengths, as we now know is very wide from radio waves (λ is 1m to 10m) to visible light (400nm) as shown in Fig. 30.4. This generated a lot of interest and activity. In 1895 Indian physicist Jagadis Chandra Bose produced waves of wavelength in the range 25mm to 5m and demonstrated the possibility of radio transmission. This work was put to practical use by Guglielmo Marconi who, succeeded in transmiting e.m. waves across the Atlantic Ocean. This marked the beginning of the era of communication using e.m. waves. Marconi along with Carl Ferdinand Braun, received the 1909 Nobel Prize in physics for their work on wireless telegraphy.



Notes

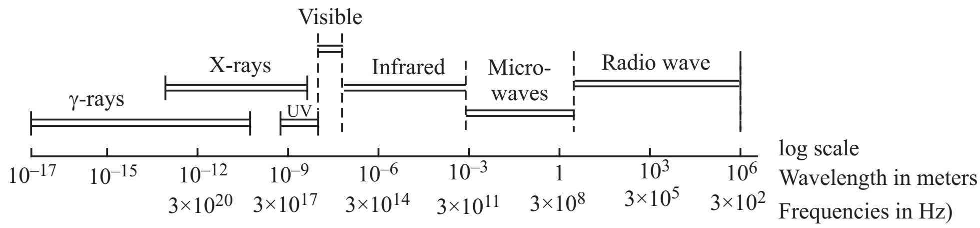


Fig. 30.4: *The electromagnetic spectrum: The wave values of length correspond to vacuum (or air) The boundaries between successive regions of the spectrum are not sharply defined.*

In a communication system, a transmitter radiates electromagnetic waves with the help of an antenna. These waves propagate in the space and captured by the receiver. At the receiver, another antenna extracts energy (information) from the electromagnetic waves. Now we use radio waves for different purposes television (TV) broadcasts, AM (amplitude modulated) and FM (frequency modulated) radio broadcasts, police and fire radios, satellite TV transmissions, cell phone conversations, and so on. Each such signal uses a different frequency, and that is how they are all separated.

You will learn the details of the mechanism of these transmissions and working of some common communication devices in the following two lessons. In Table 30.2, we have listed internationally accepted electromagnetic spectrum relevant for radio and TV broadcast, popular band names, and their application.

(Frequencies ν in Hz are related to wavelengths λ in m in vacuum through the relationship $c = \nu\lambda$, where $c = 3 \times 10^8$ m/s is the speed of electromagnetic waves in vacuum.)

Table 30.2: Radio frequency bands

Band	Frequency Range	Wavelength Range	Application
Extremely Low Frequency (ELF)	< 3 kHz	> 100 km	Mains electricity
Very Low Frequency (VLF)	3 - 30 kHz	100 – 10 km	SONAR
Low Frequency (LF)	30 - 300 kHz	10 – 1 km	Marine navigater
Medium Frequency (MF)	300 kHz - 3 MHz	1 km – 100 m	Medium wave radio
High Frequency (HF) Very High Frequency (VHF)	3 - 30 MHz 30 – 300 MHz	100 – 1 m 10 – 1 m	short wave radio FM radio
Ultra High Frequency (UHF)	300 MHz – 3 GHz	1 m – 10 cm	commercial, TV, Radio, Radar
Super High Frequency (SHF)	3 – 30 GHz	10 – 1 cm	Satellite communiction, cellular mobile, commercial TV

AM radio is broadcast on bands, popularly known as the Long wave*:* 144 - 351 kHz (in the LF), the Medium wave: 530 - 1,700 kHz (in the MF), and the Short wave*:* 3 – 30 MHz (HF). **Medium wave** has been most commonly used for commercial AM radio broadcasting. **Long wave** is used everywhere except in North and South Americas, where this band is reserved for aeronautical navigation. For long- and medium-wave bands, the wavelength is long enough that the wave diffracts around the curve of the earth by ground wave propagation, giving AM radio a long range, particularly at night. **Short wave** is used by radio services intended to be heard at great distances away from the transmitting station; the far range of short wave broadcasts comes at the expense of lower audio fidelity. The mode of propagation for short wave is ionospheric.

Table 30. 3 : Frequency ranges for commercial FM-radio and TV broadcast

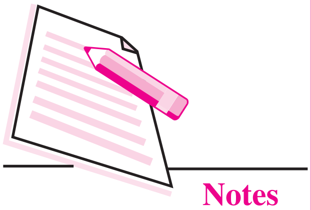
Frequency Band	Nature of Broadcast
41 – 68 MHz	VHF TV
88 – 104 MHz	FM Radio
104 – 174 MHz	S Band (Sond-erkanal meaning Special Channel) for cable TV networks
174 – 230 MHz	VHF TV
230 – 470 MHz	H (Hyper) Band for cable TV networks
470 – 960 MHz	UHF TV

Frequencies between the broadcast bands are used for other forms of radio communication, such as walkie talkies, cordless telephones, radio control, amateur radio, etc.

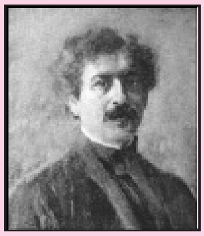
You must have read about Internet enabled mobile phones and Internet Protocol Television. Have you ever thought as to which technology is enabling such empowerment? Is it fibre optic communication? Does laser play any role? You will learn answers to all such questions in the next unit.

INTEXT QUESTIONS 30.1

1. What is an electromagnetic wave?
2. Calculate the wavelength of a radio wave of frequency of 30 MHz propagating in space.
3. What is the frequency range of (i) visible light, (ii) radio waves?



Jagadis Chandra Bose
(1858 – 1937)



Jagadis Chandra Bose, after completing his school education in India, went to England in 1880 to study medicine at the University of London. Within a year, he took up a scholarship in Cambridge to study Natural Science at Christ's College – one of his lecturers at Cambridge, Professor Rayleigh had a profound influence on him. In 1884 Bose was awarded B.A degree by Cambridge university and B.Sc degree by London University. Bose then returned to India and took teaching assignment as officiating professor of physics at the Presidency College in Calcutta (now Kolkata). Many of his students at the Presidency College were destined to become famous in their own right. Satyendra Nath Bose who became well known for his pioneering work on Bose-Einstein statistics and M.N. Saha who gave revolutionary theory of thermal ionisation, which enabled physists to classify the stars into a few groups.

In 1894, J.C. Bose converted a small enclosure adjoining a bathroom in the Presidency College into a laboratory. He carried out experiments involving refraction, diffraction and polarization. To receive the radiation, he used a variety of junctions connected to a highly sensitive galvanometer. He developed the use of *galena* crystals for making receivers, both for short wavelength radio waves and for white and ultraviolet light. In 1895, Bose gave his first public demonstration of radio transmission, using these electromagnetic waves to ring a bell remotely and to explode some gunpowder. He invited by Lord Rayleigh, to give a lecture in 1897. Bose reported on his microwave (2.5 cm to 5 mm) experiments to the Royal Institution and other societies in England. But Nobel prize alluded him probably for want of vivid practical appliction of this work by him. By the end of the 19th century, the interests of Bose turned to response phenomena in plants. He retired from the Presidency College in 1915, and was appointed Professor Emeritus. Two years later the Bose Institute was founded in Kolkata. Bose was elected a Fellow of the Royal Society in 1920.

30.4 COMMUNICATION MEDIA

There are two types of communication channels: wireline (using guided media) or wireless (using unguided media). *Wireline* channels physically connect the transmitter to the receiver with a "wire," which could be a twisted pair of transmission lines, a coaxial cable or an optical fibre. Consequently, wireline channels are more private and comparatively less prone to interference than wireless channels. Simple wireline channels connect a single transmitter to a single receiver, i.e., it is a point-to-point connection. This is most commonly observed in the telephone network, where a guided medium in the form of cable carry the signal from the telephone exchange to our telephone set. Some wireline channels operate in the broadcast mode, i.e., one or more transmitters are connected to several receivers, as in the cable television network.

Wireless channels are much more public, with a transmitter antenna radiating a signal that can be received by any antenna tuned close by. In radio transmission, the wireless or unguided propagation of radio waves from the transmitter to the receiver depends on the frequency of the electromagnetic waves. As you will learn in this lesson, the waves are transmitted as ground (or surface) waves, sky waves, or space waves by direct line-of-sight using tall towers, or by beaming to artificial satellites and broadcasting from there. Wireless transmission is flexible endowed with the advantage that a receiver can take in transmission from any source. As a result, desired signals can be selected by the tuner of the receiver electronics, and avoid unwanted signals. The only disadvantage is that the interference and noise are more prevalent in this case.

For transmitting em signals, we use microwave frequencies, you may recall that it varies from 1GHz to 300GHz. This frequency range is further divided into various bands. Indian satellite INSAT – 4C operates in the C band (4 – 8 GHz), whereas Edusat operates in Ku bond (12–18 GHz).

30.4.1 Transmission lines

For guided signal transmission, a transmission line – a material medium forms a path. As such, the construction of a transmission line determines the frequency range of the signal that can be passed through it. Fig. 30.5 shows some typical transmission lines. The simplest form of transmission line is a pair of parallel conductors separated by air or any dielectric medium. These are used in telephony. However, such lines tend to radiate, if the separation between the conductors is nearly half of the frequency corresponding to the operating frequency. This may lead to noise susceptibility, particularly at high frequencies, and limit their utility. To overcome this problem, we use a *twisted pair of* wires. These are used in computer networking.

At high signal frequencies $(\leq 3\text{GHz})$ we minimise radiation losses by using *coaxial cables,* where one conductor is hollow and the second conductor is placed inside it at its centre throughout the length of the cable. These conductors are separated by dielectric spacer layers of polythylene and the electric field is confined in the annular space in between the conductors. These cables are used for carrying cable TV signals. It is important to note that ideally the dielectrics should have infinite resistance. But in practice, their resistance is finite and that too decreases with frequency. As a result, even coaxial cables are useful in a limited range (upto a maximum of 40GHz when special dielectric materials are used). Beyond 40GHz, we use *waveguides*. However, for frequencies greater than 300GHz, their dimensions become too small (is 4mm or so) and it presents practical problems. Above this frequency, we use optical fibres for guided wave transmission.

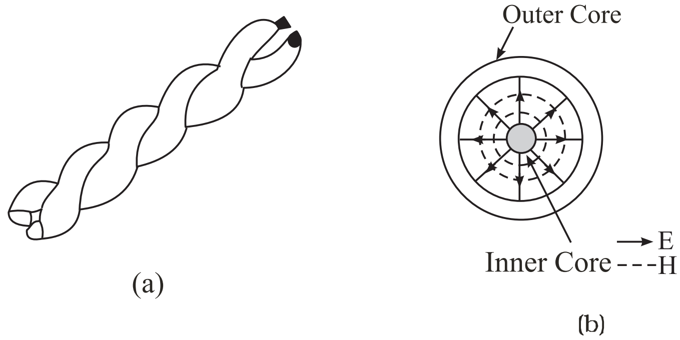


Fig. 30.5: (a) A twisted pair **(b)** A coaxial cable

30.4.2 Optical Fibre

The 1960 invention of the **laser** (acronym for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation) completely revolutionized communication technology. The laser, which is a highly coherent source of light waves, can be used as an enormously high capacity carrier wave for information carrying signals (voice, data or video) transmitted through an **optical waveguide**, such as an **optical fibre**. The basic principle involved in all long distance communication systems is **multiplexing**, i.e., simultaneous transmission of different messages over the same pathways. To illustrate it, let us consider transmission of an individual human voice. The frequency band required for transmitting human voice extends from v_1 = 200Hz to v_2 = 4000 Hz, i.e., the information contained in this frequency band can be transmitted in any band whose width is $v_1 - v_2$ = 3800 Hz, regardless of the region of the spectrum in which it is located. Higher frequency regions have far more room for communication channels, and hence, have a much greater potential capacity than the lower frequencies. The frequency corresponding to the visible optical region at 600 nm is 5×10^{14} Hz, while that at a wavelength of 6 cm is 5 x 10^9 Hz. Thus, the communication capacity of visible light in an optical fibre is about 100,000 times greater than that of a typical microwave in a metallic conductor.

The most extensively used optical waveguide is the step-index optical fibre that has a cylindrical central glass or plastic core (of refractive index n_1) and a cladding of the same material but slightly (about 1%) lower refractive index (n_2). There is usually an outer coating of a plastic material to protect the fibre from the physical environment (Fig. 30.6)

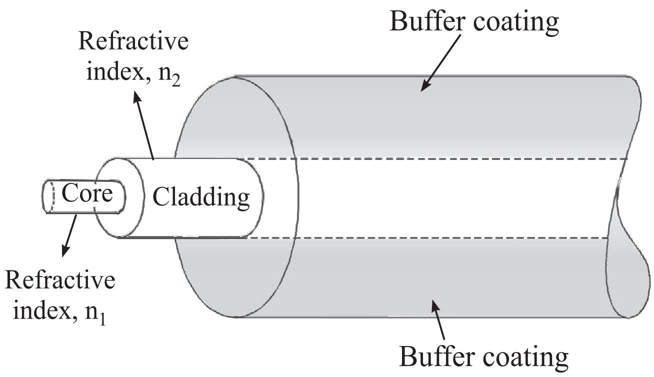


Fig. 30.6: A typical optical fibre with a doped silica core and a pure silica cladding.

When light from the core (n_1) is incident on the interface of the cladding ($n_2 < n_1$), the *critical angle* of incidence for *total internal reflection* is given by $\theta_c = \sin^{-1}(n_2/n_1)$. Thus in an optical fibre, the light ray is made to enter the core such that it hits the core-cladding interface at an angle $\theta_1 > \theta_c$. The ray then gets guided through the core by repeated total internal reflections at the upper and lower core-cladding interfaces. You may recall from wave optics that when a plane wave undergoes total internal reflection, a wave propagates in the cladding (rarer medium) along the interface, with its amplitude decreasing exponentially away from the interface. The entire energy of the wave in the core is reflected back, but there is a power flow along the interface in the cladding. Such a wave is called an *evanescent wave*, and is extensively used in integrated optics for the coupling the energy of a laser beam into a thin film waveguide (Fig. 30.7)

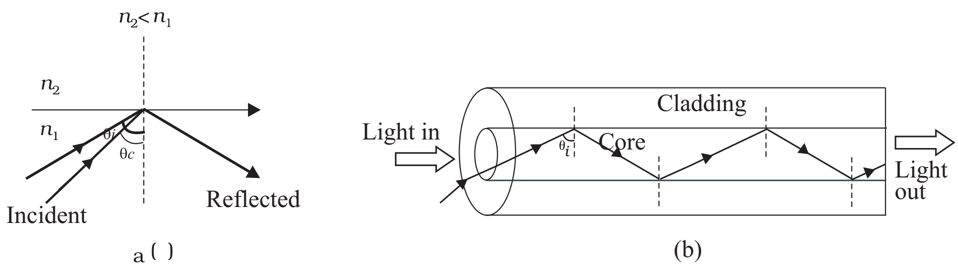


Fig. 30.7: (a) Total internal reflection **(b)** Ray confinement in actual optical fibre

INTEXT QUESTIONS 30.2

1. What is a coaxial cable? Write down its frequency range of operation.
2. State the basic principle used for guiding light in an optical fibre.

30.5 UNGUIDED MEDIA

The wireless communication between a transmitting and a receiving station utilising the space around the earth, i.e. atmosphere is called *space communication.* The earth's atmosphere plays a very interesting role the propagation of e.m. waves from one place to another due to change in air temperature, air density, electrical conductivity and absorption characteristics with height. For example, most of the radiations in infarred region are absorbed by the atmosphere. The ultaviolet radiations are absorbed by the ozone layer.

Five layers are considered to play main role in communicaiton:

- *C layer* at about 60km above the surface of earth reflects e.m. waves in the frequency range 3kHz – 300kHz. It is therefore used for direct long range communication.
- *D layer* at a height of about 80km reflects e.m. waves in the low frequency range (3kHz – 300kHz) but absorbs waves in the medium frequency range (300 kHz – 3MHz) and high frequency range (3 – 30MHz).
- *E layer* ata height of about 110km helps in propagation of waves in the medium frequency range but reflects waves in the high frequency range in the day time.
- F_1 *layer* at a height of about 180 km lets most of the high frequency waves to pass through.
- F_2 *layer* (at a height of 300 km in daytime and 350 km at night) reflects e.m. waves upto 30MHz and allows waves of higher frequencies to pass through.

You may recall from your easlier classes that, based on the variation of temperature, air density and electrical conductivity with altitude, the atmosphere is thought to be made up of several layers. The atmospheric layer close to the earth called the *troposphere* extends up to about 12 km above the sea level. The temperature in troposphere vary between 290K (at the equator) to 220K (at tropopause). The air density is maximum but electrical conductivity is the least compared to other layers. The next layer up to about 50 km is called the *stratosphere*. An ozone layer is in the lower stratosphere extends from about 15 km to about 30 km. The layer above the stratosphere and up to about 90 km is called the *mesosphere*. The minimum temperature in mesosphere in about 180K. Beyond mesosphere upto 350km, there is a zone of ionised molecules and electrons called the *ionosphere*. In ionosphere, temperature increases with height to about 1000k. The ionosphere affects the propagation of radio waves. It is divided into D, E, F and F_2 regions based on the number density of electrons, which increases with height from about $10^9 m^{-3}$ in Dregion to $10^{11} m^{-3}$ in region and $10^{12} m^{-3}$ in F_2 layer[1]. These variations in temperature, density and conductivity arise due to different absorption of solar radiations at different heights and changes in composition etc.

The essential feature of space communication is that a signal emitted from an antenna of the transmitter has to reach the antenna of the receives. Depending on the frequency fo radio wave, it can occur as *ground wave, space wave, sky wave* and via satellite communication. Let us now learn about these.

30.5.1 Ground Wave Propagation

In ground *wave* propagation, the electromagnetic waves travel along the surface of the earth. These can bend around the corners of the objects but are affectd by terrain. A vertical antenna is used to transmit electromagnetic waves. If electric field E is vertical, and the magnetic field B is horizontal, the direction of propagation k is horizontal but perpendicular to both E and B vectors. The material properties of the ground, such as its conductivity refractive index and dielectric constant, are seen to control propagation of such waves. That is why ground waves propagation is much better over sea than desert. In practice, ground waves are rapidly attenuated due to scattering by the curved surface of the the earth. A larger wavelength results in smaller attenuation. That is, ground waves are more useful as lower frequencies & constitute the only way to communicate into the ocean with submarines. Moreover, this mode of propagation is suitable for short range communication. For these reasons, ground wave propagation is used for radio wave (300kHz – 3MHz) transmission.

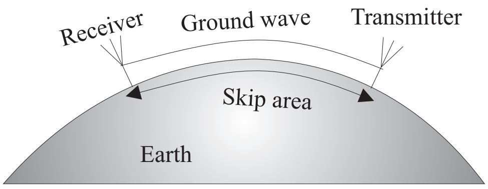


Fig. 30.8: Ground wave propagation

30.5.2 Sky Wave or Ionospheric Propagation

In **sky wave** or **ionospheric** propagation, the electromagnetic waves of frequencies between 3MHz – 30MHz launched by a transmitting antenna travel upwards, get reflected by the ionophere and return to distant locations. In this mode, the reflecting ability of the ionosphere controls the propagation characteristics of the sky wave. The ionosphere acts as an invisible electromagnetic "mirror" surrounding the earth – at optical frequencies it is transparent, but at radio frequencies it reflects the electromagnetic radiation back to earth.

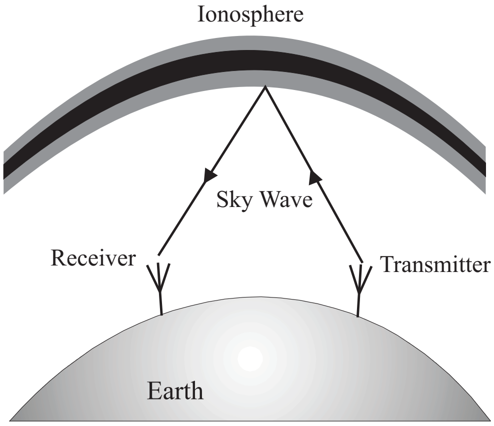


Fig. 30.9: Skywave propagation

The maximum distance along the surface of the earth that can be reached by a single ionospheric reflection ranges between 2010 and 3000 km depending on the altitude of the reflecting layer. The communication delay encountered with a single reflection ranges between 6.8 and 10 ms, a small time interval. This mode of propagation is used for long-distance (short wave) communication in the frequency range approximately between 5 and 10 MHz. Above 10 MHz, the waves pass through the ionosphere and do not reflect back to the earth. It is, however, subject to erratic daily and seasonal changes due to variations in the number density and height of the ionized layers in the ionosphere. The composition of the ionosphere at night is different than during the day because of the presence or absence of the sun. That is why international broadcast is done at night because the reflection characteristics of the ionosphere are better at that time.

30.5.3 Space Wave Propagation

You may have seen very high antennas at radio station. These are used for broadcasting. In space wave propagation, some of the VHF radio waves (30 MHz – 300MHz) radiated by an antenna can reach the receiver travelling either directly through space or after reflection by the curvature of the earth. (Note that earth reflected waves are different from ground waves.)

In practice, direct wave mode is more dominant. However, it is limited to the so-called *line-of-sight* transmission distances and curvature of earth as well as height of antenna restrict the extent of coverage.

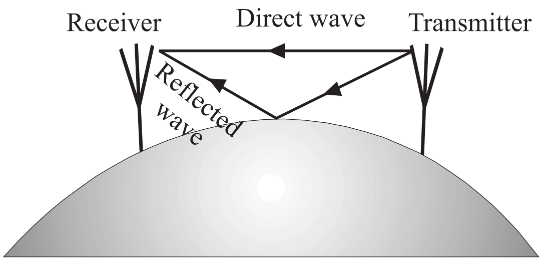


Fig. 30.10: Space-wave propagation

So far you have learnt that ground waves suffer conduction losses, space waves have limitations due to line of sight and sky waves penetrate the ionisalion beyond a certain frequency. Some of these difficulties were circumvented with the launch of communication satellities in the 1950s. Satellite communication has brought about revolutionary changes in the form and format of transmission and communication. We can now talk is real time at a distance. Let us now learn about it.

30.5.4 Satellite Communication

The basic principle of satellite communication is shown in Fig. 30.11. The modulated carrier waves are beamed by a transmitter directly towards the satellite.

The satellite receiver amplifies the received signal and retransmits it to earth at a different frequency to avoid interference.

These stages are called uplinking and down-linking.

As we have seen already in connection with communication with light waves, the capacity of a communication channel can be increased by increasing the frequency of communication. How high up can we go in frequency? You now know that the ionosphere does not reflect waves of frequencies above 10 MHz, and for such high frequencies we prefer space wave propagation with direct transmission from tall towers. But this line-of-sight transmission also has a limited range or reach. Hence for long-range wireless communication with frequencies above 30 MHz, such as for TV transmission in the range of 50-1000 MHz, communication through a satellite is used.

The gravitational force between the earth and the satellite serves as the centripetal force needed to make the satellite circle the earth in a f*reefall* motion at a height of about 36,000 km. An orbit in which the time of one revolution about the equator exactly matches the earth's rotation time of one day is called a *geostationary* orbit, i.e., the satellite appears to be stationary relative to the earth. Ground stations transmit to orbiting satellites that amplify the signal and retransmit it back to the earth. If the satellites were not in geostationary orbits, their motion across the sky would have required us to adjust receiver antenna continually. Two other orbits are also currently being used for communication satellites: (i) *polar circular orbit* at a height of about 1000 km almost passing over the poles (i.e., with an inclination of 90^0), and (ii) *highly elliptical inclined orbit* (with an inclination of 63°) for communications in regions of high altitudes.

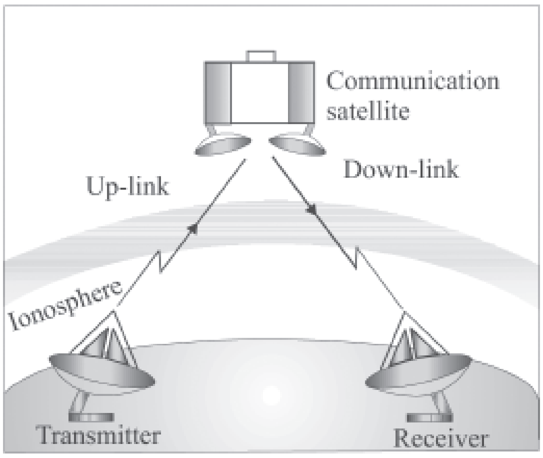


Fig. 30.11: Satellite communication.

INTEXT QUESTIONS 30.3

1. Why do you hear some radio stations better at night than in the day?
2. Choose the correct option in each case:
   (a) Frequencies in the UHF range normally propagate by means of
      (i) Ground Waves
      (ii) Sky Waves
      (iii) Surface Waves
      (iv) Space Waves.
   (b) Satellites are used for communication
      (i) With low (< 30 MHz) frequencies and for a small range
      (ii) With low (< 30 MHz) frequencies and for a long range
      (iii) With high (> 30 MHz) frequencies and for a small range
      (iv) With high (> 30 MHz) frequencies and for a long range.

EDUSAT

The Indian Space Research Organisation (ISRO), Department of Space, Government of India, launched an exclusive education satellite EDUSAT in Sept. 2004. The satellite has its footprints all over the country and operates in KU band. It is designed to provide services for seven years. This satellite has capability for radio and TV broadcast, Internet-based education, data broadcasting, talk-back option, audio-video interaction, voice chat on Internet and video conferencing. It has opened up numerous possibilities: a teacher of a leading educational institution in a city may video-conference with students of a remote school, or school drop-outs in villages may receive Internet-based education support and get back into mainstream education system. EDUSAT has the capability of telecasting 72 channels. A large number of networks have been created by state governments and national institutions including NIOS. Such networks are being successfully used to impart education even in regional languages.

30.6 MODULATION – ANALOGUE AM AND FM, DIGITAL (PCM)

The process of processing a signal to make it suitable for transmission is called *modulation.* Most of the information-bearing signals in day-to-day communication are audio signals of frequency less than 20 kHz. For small distances, we can form direct link. But it is not practical to transmit such signals to long distances. This is because of the following two reasons:

- The signal should have an antenna or aerial of size comparable to the wavelength of the signal so that the time variation of the signal is properly sensed by the antenna. It means that for low-frequency or long-wavelength signals, the antenna size has to be very large.
- The power carried by low frequency signals is small and can not go far. It is because of cantinuous docline or attemation due to absorption/radiation loss. It means that for long distance transmission high frequencies should be used. But these can not carry useful information. We are therefore confronted with a situation analogous to the following:

On a front port, Indian army spots advancing enemy forces. To minimise loss of life and save the post from falling to enemy, they need reinforcement from the base camp. But by the time an army jawan goes, conveys the message and the reinforcement reaches, the port would have fallen. Therefore, it wants a carrier, say a horse, which can run fast. But the horse can not deliver the message. The way out is: Put the jawan on the horseback; let the horse run and jawan convey the message.

For signal transmission, audio signal acts as jawan and high (radio) frequency acts as the horse (carrier). So we can say that by super imposing a low frequency signal on a high frequency carrier wave, we process a signal and make it suitable for transmission. We convert the original signal into an electrical signal, called the *base band signal* using a signal generator. Next we super impose the base band signal over carrier waves in the modulator. The change produced in the carrier wave is known as modulation of the carrier wave and the message signal used for modulation is known as *modulating signal*. The carrier wave can be continuous or pulsed. Since a sinusoidal wave, is characterised by amplitude, frequency and phase it is possible to modulate (i.e. modify) either of these physical parameter.

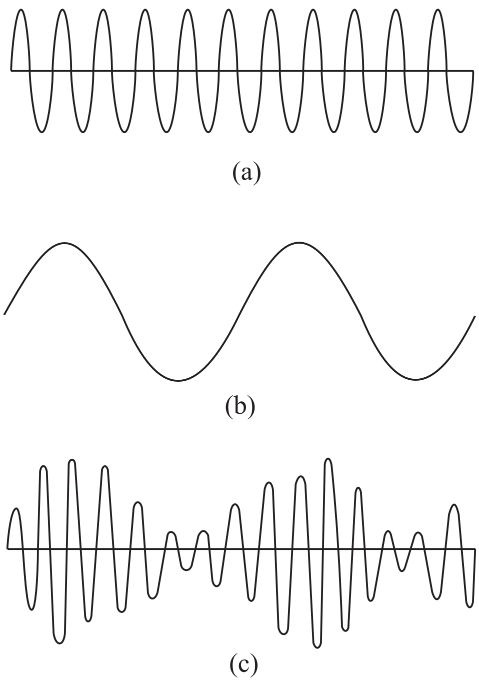


Fig. 30.12: Modulation of a carrier wave by a modulating signal: **(a)** a sinusoidal carrier wave of high frequency, **(b)** a modulating signal (message or information signal) of low frequency, **(c)** amplitude modulated carrier wave.

This is known as analog modulation. There are different types of analog modulation: **Amplitude Modulation** (**AM**); **Frequency Modulation** (**FM**); and **Phase Modulation (PM),** respectively For pulsed carrier waves, **Pulse Code Modulation** (**PCM**) is the preferred scheme.

In Amplitude modulation, the amplitude of a high-frequency carrier wave (Fig. 30.12a) is modified in accordance with the strength of a low-frequency audio or video modulating signal (Fig.30.12.b). When the amplitude of the modulating wave increases, the amplitude of the modulated carrier also increases and vice-versa — the envelope of the modulated wave takes the form depending on the amplitude and frequency of modulating signal (Fig. 30.12.c) .

To understand this, we write expressions for instantaneous amplitudes of audio signal and carrier wave:

$$\upsilon_a(t) = \upsilon_{ao} \sin\omega_a t \qquad (30.1a)$$

and

$$\upsilon_c(t) = \upsilon_{co} \sin\omega_c t \qquad (30.1b)$$

where w_a and w_c are the angular frequencies and $\boldsymbol{\upsilon}_{ao}$ and $\boldsymbol{\upsilon}_{co}$ denote of audio and carrier waves, respectively. denote the amplitudes. In amplitude modulation the modulationg (audio) signal is superimposed on the carrier wave, so that the amplitude of the resultant modulated wave can be expressed as

$$A(t) = \upsilon_{co} + \upsilon_a(t) = \upsilon_{co} + \upsilon_{ao} \sin\omega_a t$$

$$= \upsilon_{co}\left[1 + \frac{\upsilon_{ao}}{\upsilon_{co}} \sin \omega_a t\right] \qquad (30.2)$$

Hence the modulated wave can be expressed as

$$\upsilon_c^{\text{mod}}(t) = A \sin \omega_c t = \upsilon_{co}\left[1 + \frac{\upsilon_{ao}}{\upsilon_{co}} \sin \omega_a t\right] \sin w_c t \qquad (30.3)$$

From Eqn. (30.3) we note that the instantaneous amplitude of the modulated wave is determined by the amplitude and frequency of the analog audio signal. The ratio $\upsilon_{ao}/\upsilon_{co}$ gives us a measure of the extent to which carrier amplitude is varied by the analog modulating signal and is known as amplitude modulation index. We will denote it by m_a. In terms of modulation index, we can rewrite Eqn. (30.3) as

$$\upsilon_c^{\text{mod}} = \upsilon_{co}(1 + m_a \sin\omega_a t)\sin\omega_c t$$

$$= \upsilon_{co}\sin\omega_c t + \upsilon_{co} ma \sin\omega_a t \sin\omega_c t$$

$$= \upsilon_{co}\sin\omega_c t + \frac{\upsilon_{co} m_a}{2}\cos(\omega_c - \omega_a)t - \frac{\upsilon_{co} m_a}{2}\cos(\omega_c + \omega_a)t \qquad (30.4)$$

From Eqn. (30.4) we note that

- the modulated wave shown in Fig. 30.12(c) has three components. The first term represents carrier wave the second term whose frequency is lower than that of the carrier wave, constitutes the lower side band, and the third term with frequency higher than the carrier wave is the upper side band; and
- the frequency of the modulating signal is not directly contained in the amplitude modulated wave.

If the modulating signal in an AM system is given by

$\upsilon_a = 4\sin 6283t$ and frequency of the lower side band is 3.5×10^5Hz, the angular frequency of the carrier wave is given by

$$\omega_C = \omega_a + 2\pi \times (3.5) \times 10^5$$
$$= 6283 + 22 \times 10^5$$
$$= (2200 + 6.283) \times 10^3 \text{rad}$$
$$= 2.206 \times 10^6 \text{rad}$$

It is important to appreciate that the most efficient information transfer takes place when maximum power transmitted by the comunication system is contained in the side bands.

The block diagram of a basic analog AM transmitter is shown in Fig. 30.13 (a). The oscillator provides a fixed frequency and the power amplifier modulates the signal.

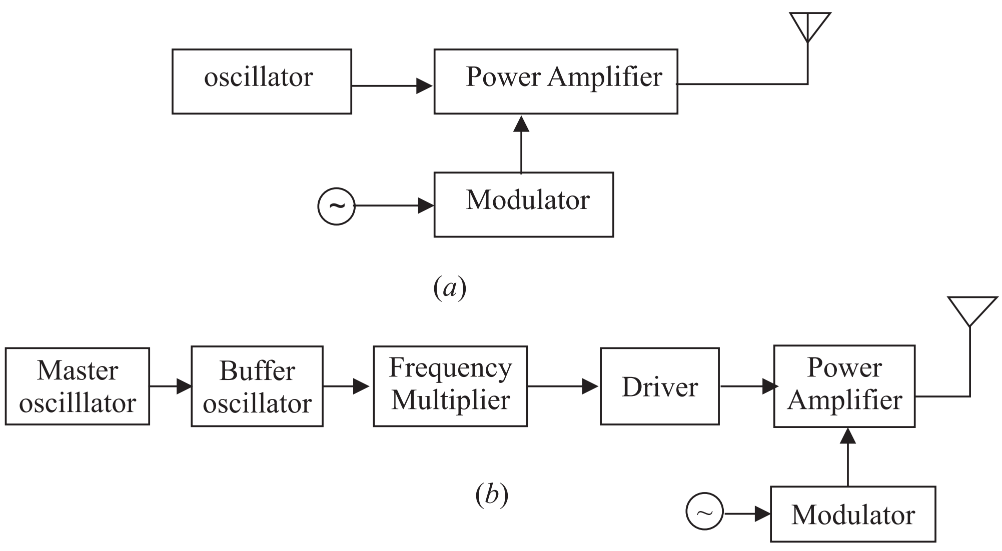


Fig. 30.13 Block diagram of **(a)** a basic and **(b)** practical AM transmitter

For any broadcast, the maximum power that can be radiated is controlled by the GOI. It is in the range 500W to 50kW for radio transmitters. Every broadcaster is allocated a definite frequency, which has to be observed strictly to avoid interfrence with other signals. To ensure this, undesirable frequencies are filtered out by using coupling circuits. We will not go into these details further.

The most popular form of radio communiction in India over the past 50 years had been medium wave (520 – 1700kHz) and short wave (4.39 – 5.18MHz; 5.72 – 6.33MHz) analog AM broadcast. It continues to have the widest spread, though analog FM broadcast is now being preferred because of better quantity. Moveover, radio waves are now comparatively free and private broadcasters are also entering the field in a big way. FM radio stations are also being created by educational institutions for education as well as empowerment of rural youth and homemakers. In TV transmission, audio is frequency modulated whereas the video (picture) is amplitude modulated.

In **frequency modulation**, the amplitude of the carrier wave remains constant, but its frequency is continuously varied in accordance with the instantaneous amplitude of the audio or video signal. When the amplitude of the modulating signal voltage is large, the carrier frequency goes up, and when the amplitude of the modulating signal is low, the carrier frequency goes down, i.e., the frequency of the FM wave will vary from a minimum to a maximum, corresponding to the minimum and maximum values of the modulating signal (Fig. 30.14).

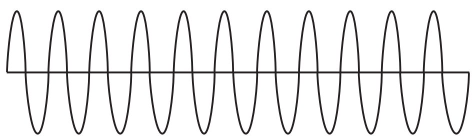


Fig. 30.14: Frequency modulated career wave

An FM Transmitter essentially contains an oscillator, whose frequency of the carrier is varied depending on the input audio signal. (It is usually accomplished by varying capacitance in an LC oscillator or by changing the charging current applied to a capacitor, for example, by the use of a reverse biased diode, since the capacitance of such a diode varies with applied voltage.) After enhancing the power of the modulated signal, it is fed to the transmission antenna. Low-frequency radio broadcast stations use amplitude modulation, since it is a simple, robust method.

Phase modulation involves changing the phase angle of the carrier signal is accordance with the modulating frequency. Analog pulse modulation is either amplitude modulated or time modulated. Similarly, digital pulse modulation is of two types: pulse code modulation and pulse delta modulation.

In **pulse code modulation**, the modulating signal is first sampled, and the magnitude (with respect to a fixed reference) of each sample is quantised. (It is a digital representation of an analog signal where the magnitude of the signal is sampled regularly at uniform intervals of duration T_s. The binary code is transmitted usually by modulating an analog current in a transmission medium such as a landline whereas pulse code modulation is used in digital telephone systems and for digital audio recording on compact discs.

30.7 DEMODULATION

The modulated signal carrying the information, once radiated by the antenna, travels in space. Since there are so many transmitting stations, thousands of signals reach our antenna.

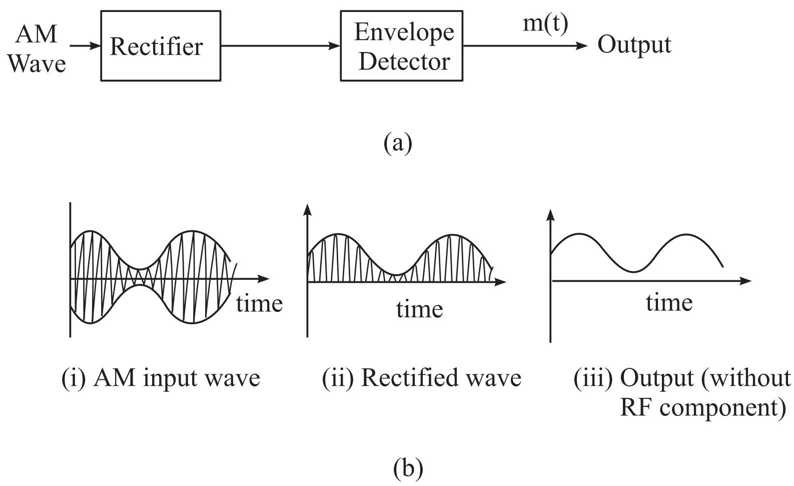


Fig. 30.15: (a) Block diagram of a detector for AM signal. (b) (i) AM modulated input wave (ii) Rectified wave (iii) output demodulated wave

We have to choose the desired signal and decouple the carrier wave and the modulating signal. This process is known as *demodulation*. The modulated signal of the form shown in Fig. is passed through 30.15(a) a rectifier to produce the output shown in Fig. 30.15(b). This envelop of the signal (b) is the message signal. In order to rereavel the message signal, the signal is passed through a envelop detectors (demodulator) which may consists of a simple RC circuit.

INTEXT QUESTIONS 30.4

1. Choose the correct option in each case:

   (a) Modulation is used to

      (i) reduce the bandwidth used

(ii) separate the transmissions of different users

(iii) ensure that information may be transmitted to long distances

(iv) allow the use of practical antennas.

(b) AM is used for broadcasting because

(i) it is more noise immune than other modulation systems

(ii) it requires less transmitting power compared to other systems

(iii) it avoids receiver complexity

(iv) no other modulation system can provide the necessary bandwidth for faithful transmission.

2. What is the optimum size of a radio antena.

COMMUNICATION APPLICATIONS

In recent years, the world of communication has advanced rapidly from printed texts to the telegraph, the telephone, the radio, the television, mobiles, Internet and computer conferencing (Audio and video). Countries all over the world are striving to achieve high standards of national and international communications. Radio and TV broadcasting through communication satellites is routinely achieved to reach out to the majority of the population even in remote corners of the globe. The domestic system of automatic telephone exchanges is usually connected by modern networks of fibre-optic cable, coaxial cable, microwave radio relay, and a satellite system.

Cellular or mobile telephone services are now widely available and include roaming service, even to many foreign countries. The cellular system works as a radio network of base stations and antennas. (The area of a city covered by one base station is called a cell, whose size ranges from 1 km to 50 km in radius.) A cell phone contains both a low-power transmitter and a receiver. It can use both of them simultaneously, understand different frequencies, and can automatically switch between frequencies. The base stations also transmit at low power. Each base station uses carefully chosen frequencies to reduce interference with neighbouring cells.

In a situation where multiple personal computers are used, as perhaps in your local study centre, it helps to get all the computers connected in a network so that they can "talk" to each other, and we can

- share a single printer between computers;
- share a single Internet connection among all the computers;
- access shared files and documents on any computer;
- play games that allow multiple users at different computers; and
- send the output of a device like a DVD player to other computer(s).

To install such a network of personal computers, there are three steps:

- Choose the technology for the network. The main technologies to choose between are standard Ethernet, phone-line-based, power-line-based and wireless.
- Buy and install the hardware.
- Configure the system and get everything talking together correctly.

The Internet is a vast network of computers throughout the world. It combines many different forms of communications. As the technology advances it could replace all other forms of communication by combining them into one. Magazines and newspapers are already being put online along with libraries, art, and research. Unlike most forms of communication, it facilitates access to vast store of information through the World Wide Web (WWW). The World Wide Web is the multimedia part of the Internet and combines text with sound, photos, drawings, charts, graphs, animation, and even video. New innovations such as Java, a web-based programming language, allow simple tasks to be performed inside the document. The more widespread the Internet becomes, the more important and powerful type of communication it will become. In India, several hundred thousand schools are being provided access to computers and Internet to impure the quality of education. The MHRD is developing a one stop portal –Sakshat– which can be accessed by you. The National Institute of open schooling is also contributing for it.

You may be aware of variouis complex communication systes in use around us; like: radio, TV, fax machine, internet etc.

A cathode-ray tube in a television set, has a “cathode” which emits a ray of electrons in a vacuum created inside a glass “tube”. The stream of electrons is focused and accelerated by anodes and hits a screen at the other end of the tube. The inside of the screen is coated with phosphor, which glows when struck by the beam. The cathods beam carries the vedio signals from the object and forms the image an the screen accordingly.

In a fax machine, the document to be transmitted is scanned by a photo sensor to generate a signal code before it is transmitted through a telephone line.

A modem (modulator/demodulator) can convert a digital bit stream into an analog signal (in the modulator) and vice-versa (in the demodulator). It is used as a transmitter to interface a digital source to an analogue communication channel, and also as a receiver to interface a communication channel to a digital receiver

WHAT YOU HAVE LEARNT

- In a typical modern-day long distance communication system, the information is in the form of electrical signals (voltage or current).
- The essential elements of a communication system are (i) a transmitter (ii) a medium or mechanism to carry the signal over long distances, and (iii) a receiver to intercept the signal and retrieve the information.
- An antenna or aerial is essentially a system of conductors, which is an effective radiator and absorber of electromagnetic waves in the desired radio frequency region.
- Analog signals are physical signals that vary continuously with time while digital signals have the form of discrete pulses.
- Digital communication systems are more efficienct, give better performance, and greater flexibility than their analog counterparts.
- AM radio is broadcast on three bands, the Long wave at 144 – 351 kHz (in the LF), the Medium wave at 530 – 1,700 kHz (in the MF), and the Short wave at 3 – 30 MHz (HF). FM radio is broadcast on carriers at 88 – 104 MHz (in the VHF). Commercial TV transmission is in the VHF-UHF range.
- Electrical communications channels are wireline (using guided media) or wireless (using unguided media).
- Multiplexing refers to the process of simultaneous transmission of different messages (each with some frequency bandwidth) over the same path way. The higher the frequency of the carrier, the higher is its message-carrying capacity.
- Comparing the different wireline channels, the communication capacity of visible light (of frequency of about 10^{14} Hz) in an optical fibre is thus much larger than that of typical microwave (of frequency of about 10^{9} Hz) in a metallic conductor.
- An optical fibre guides a light beam (from a laser) from its one end to the other by the process of total internal reflection at the interface of the inner core (of refractive index n_1) and the cladding (of refractive index $n_2 > n_1$).
- In the wireless radio transmission, a system of conductors called antenna or aerial launches the carrier radio waves in space and also detects them at the receiver location. The propagation of radio waves in the atmosphere depends on the frequency of the waves. Low and medium frequency radio waves up to about 1 MHz are used in ground (or surface) wave communication. Medium frequency (MF) waves of 300 kHz – 3 MHz are largely absorbed by the ionosphere. The high-frequency (HF) waves of 3 – 30 MHz are, however, reflected back by the ionosphere. VHF and UHF waves are transmitted either

by direct line-of-sight using tall towers (space wave or tropospheric propagation), or by beaming to artificial satellites and broadcasting from there.

- The cellular or mobile telephone system works as a radio network in which a city is divided into 'cells' of 1 km to 50 km in radius, and each cell is covered by one base station. A cellular phone contains a low-power transmitter and a low-power receiver.
- An analogue signal is completely described by its samples, taken at equal time intervals T_s, if and only if the sampling frequency $f_s = 1/T_s$ is at least twice the maximum frequency component of the analogue signal.
- Low frequencies can not be transmitted to long distances using aerials or antennas of practical dimensions. Low-frequency messages are loaded on a high frequency carrier signal by a process called modulation. In amplitude modulation (AM), the amplitude of a high-frequency carrier wave are modified in accordance with the strength of a low-frequency information signal. In frequency modulation (FM), the amplitude of the carrier wave remains constant, but its frequency is continuously varied in accordance with the instantaneous amplitude of the information signal, i.e., the frequency of the modulated carrier wave varies from a minimum to a maximum corresponding to the minimum and maximum values of the modulating signal.
- In the digital pulse code modulation (PCM) technique, the modulating signal is first sampled, the magnitude (with respect to a fixed reference) of each sample is quantised, and then the binary code is usually transmitted modulating an analogue current in a landline.

TERMINAL EXERCISE

1. What are the essential elements of a communication system?
2. What is an antenna?
3. What are the important characteristics of a receiver in a communication system?
4. Distinguish between the terms analogue and digital signals. Define a 'bit'.
5. The VHF band covers the radio frequency range of 30 – 300 MHz. Using the known relationship of speed to frequency and wavelength of an electromagnetic wave, determine the VHF wavelength range in vacuum. Take the speed of light in vacuum to be 3×10^8 ms^{-1}
6. Long distance radio broadcasts use shortwave bands. Explain.
7. Satellites are used for long distance TV transmission. Justify.
8. The core of an optical fibre is made of glass with a refractive index of 1.51 and the cladding has a refractive index of 1.49. Calculate the critical angle for total internal reflection.

9. List some advantages of creating a local network of personal computers.

10. What do you understand by modulation. Explain its need.

11. Explain the role of modulation and demodulation in communication system.

ANSWERS TO INTEXT QUESTIONS

30.1

1. E.M. waves are time varying electrical and magnetic field travelling in space with a speed of 3×10^8 ms^{-1} at right angles to each other.

2. $\lambda = \frac{C}{\nu} = \frac{3\times10^8 ms^{-1}}{30\times10^6 s^{-1}} = 10$ m

3. (i) frequency range of visible light is 10^{14} Hz–10^{15} Hz

   (ii) frequency range of radio waves is 30 kHz–300 MHz

30.2

1. Co-axial cable is a pair of point to point connector where in one conductor is in the form of hollow cylinder and the other a solid wire at the axis of the first conductor the two being separated by an insulator. The one used for frequency range 3.0 GHz-40 GHz.

2. The basic principle used in optical fibre is total internal reflecion due to which light beam may move along an optical fibre without any loss in energy.

30.3

1. Sky wave communication is normally better during night, because in absence of sun ionosphere's composition is settled so that it acts as a better reflector.

2. (a) iv, (b) (iii)

30.4

1. (a) (iii), (b) (iii)

2. The minimum size of a transmitting antenna is comparable to wavelength of signal to be transmitted. For maximum power transmission size of antenna should be at least $\lambda/4$.

Answer to Problems in Terminal Exercise

5. 10 m- 1 m

8. $\sin^{-1} n_2/n_1 = 80.66°$

Senior Secondary Course
Semiconductors And Semiconductor Devices
Student's Assignment – 8

Maximum Marks: 50 **Time : 1½ Hours**

Instructions

- Answer All the questions on a seperate sheet of paper
- Give the following information on your answer sheet:
    - Name
    - Enrolment Number
    - Subject
    - Assignment Number
    - Address
- Get your assignment checked by the subject teacher at your study centre so that you get positive feedback about your performance.

Do not send your assignment to NIOS

1. Name the majority charge carriers in n-type semiconductor? 1
2. Draw symbol for a n-p-n transistor. 1
3. Explain the meaning of the term doping in semiconductors. 1
4. What is the effect of forward biasing a p-n junction on the width of deplition region around it? 1
5. How do you identity collector and emitter in a transter. 1
6. Draw the logic symbol of a NOR gate. 1
7. Out of silicon and germanium which has more free charge carrier density at room temperature. Why? 1
8. In common base configuration current gain is less than 1 but still be can have a voltage gain. How? 1
9. Distinguish between a LED and a solar cell. Draw diagram of each. 2
10. Draw the characterstics of a pn junction diode in (*i*) forward bias (*ii*) reverse bias. 2
11. In a half wave reitifier input frequency is 50 H*z*. What is its output frequency? What is the out put frequency of a full wave rectifier for the same input frequency. 2

12. Two amplifiers are connected one after the other in series. The first amplifier has a voltage gain of 10 and the second has a voltage gain of 20. If the input signal is 0.01v, calculate the output ac signal. 2

13. How can you realize an AND gate with the help of *p-n* junction diodes? Draw the circuit and explain to truth table. 4

14. For a common emitter amplifier, the audio signal voltage across a 5 k Ω collector resistance is 5v. Suppose the current amplification factor of the transter is 100, find the input signal voltage and base current, if the base resistance is 1 kΩ . 4

15. Define current gain in common base configuration and common emitter configuration. Establish a relation between the two. 4

16. With the help fo suitable diagrams explain 4

    (*a*) how does a capacitor convert functuating *ac* steady *dc*.

    (*b*) how a zener diode stabilizes dc output against load variations.

17. Explain : 4

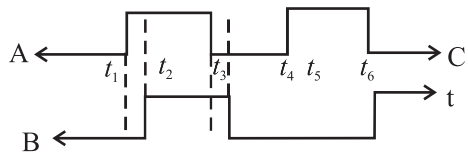
    (*i*) Why a transistorhas to be biased for using it as an amplifer,

    (*ii*) How the range of variation of amplitude of input signal is decided for the proper working of a transistor,

    (*iii*) Why the voltage gain of an amplifier can not be increased beyond a limit by increasing load resistance.

18. Identity the logic gates indicates by circuits given below.

Corresponding to the input signal at A and B as shown below draw output waveform for each ats.



19. With the help of a circuit diagram explain how a transistor can be used as an amplifier? 5

20. Draw a circuit diagram for studying the charactertics. Draw the input and output charactertics and explain the current gain obtained. 5

QUESTIONS PAPER DESIGN

Subject: **Physics** Class: **Senior Secondary**
Paper Marks: **80** Duration: **03 Hrs.**

1. Weightage by Objectives

Objective	Marks	% of the Total Marks
Knowledge	20	25
Understanding	40	50
Application and Skill	20	25
Total	**80**	**100**

2. Weightage by Types of Question

Type of Questions	Marks × No. of Questions	Marks Allotted
Essay (E)	6 × 4	24
Short Answers I (SA1)	4 × 7	28
Short Answers II (SA2)	2 × 9	18
Multiple Choice Questions (MCQ)	1 × 10	10
Total	**30 Questions**	**80 Marks**

3. Weightage as per the Content

Sr. No.	Module	Marks
1	Motion, Force and Energy	14
2	Mechanics of Solids and Fluids	06
3	Thermal Physics	06
4	Oscillations and Waves	06
5	Electricity and Magnetism	16
6	Optics and Optical Instruments	14
7	Atoms and Nuclei	08
8	Semiconductor Devices and Communication	10
	Total Marks	**80**

4. Difficulty Level

	Easy	Average	Difficult	Total
Percent Weightage	25 %	50 %	25 %	100 %
Marks Allotted	20	40	20	
No. of Questions	08	15	07	

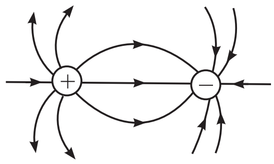
Note : Some internal choices is given in application questions.

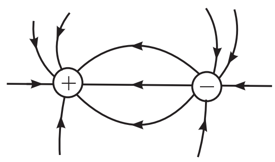
5. Time Management

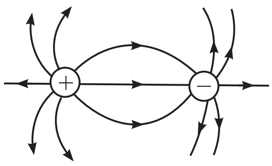
Type of Questions	Total Time 180 minutes
Essay (E)	60
Short Answers I (SA1)	50
Short Answers II (SA2)	35
Multiple Choice Questions (MCQ)	15
Reading and Revision	20

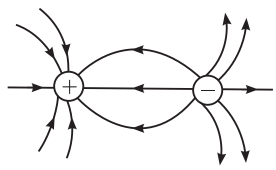
SAMPLE QUESTIONS PAPER

1. Which of the following diagrams best represents electric field lines between two equal and opposite charges?

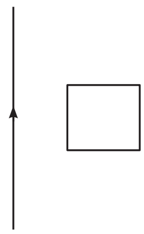
A. 

B. 

C. 

D. 
 1

2. A rectangular current carrying loop of wire is placed near a long current carrying straight conductor with one of its sides parallel to the counductor as shown in figure. The loops will be



   A. atracted towards the conductor whatever be the direction of current in the loop.
   B. repelled by the conductor whatever be the direction of current in the loop.
   C. attracted towards to conductor when the current in the loop flows anticlockwise.
   D. repelled by the conductor when the current in the loop flows anticlockwise. 1

3. The energy required to cause a transition of an electron from n = 1 to n = 2 orbit of hydrogen atom is ....

   A. 10.2 eV
   B. 13.6 eV
   C. 12.09 eV
   D. 1.51 eV

4. The number of electrons emitted by cathods in a photo tube is 3.0×10^4 when it is illuminated by a small bright source placed 50 cm away. If the source is carried to a distance of 100 cm, the number of electrons emitted by photocathode would be

   A. 3.0×10^4
   B. 7.5×10^3
   C. 1.5×10^4
   D. 6.0×10^4 1

5. In a forward biased p-n junciton diode

   A. charge carriers do not move at all.
   B. only the holes in the p-region move towards the n-region

C. only the minority charge carriers move due to thermal energies

D. majority charge carrier deffuse from either side towards the junction 1

6. A n-type extrinsic semiconducting material is

A. negatively charged

B. positively charged

C. electrically neutral

D. negatively or positively charged depending on the dopant 1

7. For a CE mode amplifier, vi is 10 mV and vo is one volt. The voltage gain of the amplifier would be

A. 50 B. 20

C. 100 D. 10 1

8. In a n-type semiconductor

A. electrons are majority charge carriers and dopants are trivalent atom.

B. electrons are minority carriers and dopants are pentavalent atoms

C. holes are the minority carrier and dopants are pentavalent atoms

D. holes are the majority charge carriers and dopants are trivalent atoms 1

9. A secondary rainbow occurs when a ray of light undergoes

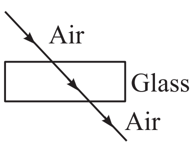
A. one refraction and one internal refleciton in a rain drop

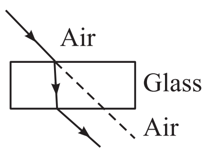
B. two refreactions and two internal reflections is a rain drop

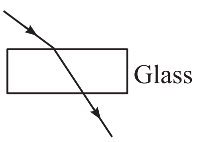
C. two refreactions and one internal refelction in a rain drop

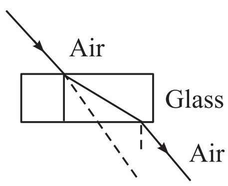
D. one refraction and two internal reflections in a rain drop 1

10. Which of the following represents the refraction of a ray of light through a glass slab correctly

A. 

B. 

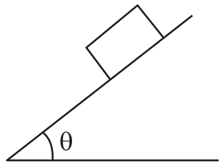
C. 

D. 

1

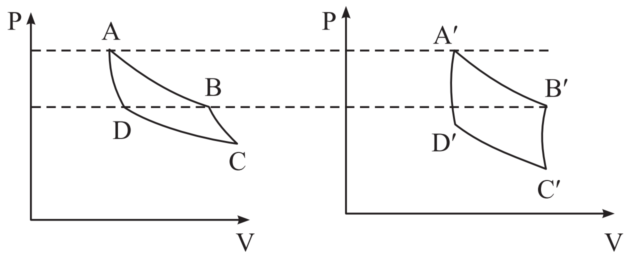
11. State second law of motion. Show that for an object of constant mass, acceleraiton of the object is directly proportional to the force applied on it. 2

12. A body of mass m is placed on a rough plane inclined at an angle θ. Draw a diagram showing forces acting on the body and write expression for frictional force if the block stays stationary on the inclined plane. 2



13. State stoke's law for the viscous force acting on a sphere of radius r, moving in a fluid of viscosity η with velocity v. Obtain the unit of coefficient of viscosity in terms of kg, metre and second. 2

14. Figures below show the p-v diagrams of two carnot engines. Which of the two engines is more efficient. (It is given that both engines draw equal heat from the source). 2



15. A spectral line in thelight from a star shows a shift towards the red end of the spectrum. If the shift is 0.03%, calculate the velocity of recession of the star. 2

16. How will you convert a galvanometer into (i) an ammeter (ii) a voltmeter? 2

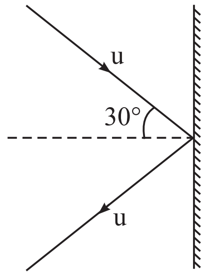
17. Draw a diagram showing refraction of a ray of light through a glass prism. Mark angle of deviation and angle of emergence inthe diagram. 2

18. In fission of a ${}^{236}_{92}U$ nucleus 200 MeV energy is released whereas in fusion of 4 protons 26.8 Mev energy is released. Whichof these processes gives more energy per unit atomic mass? Explain. 2

19. The axis of a 100 turn circular coil (area of cross-seciton 3.85×10^{-3} m^2) is parallel to a uniform magnetic field. The magnitude of the field changes at a constant range from 25 mT to 50 mT. in 250 ms. Calculate the magnitude of induced emf across the coil. 2

20. Explain (i) Isothermal (ii) Adiabatic (iii) isobaric and (iv) isochoric processes. 4

21. A ball of mass 50g strikes a rigid wall at an angle 30° with a speeed of 10 ms^{-1} and gets reflected without any change in speed as shown in figure. Find the magnitude and direciton of the impulse imprted to the ball by the wall. 4



22. Explain why is a small drop of liquid always spherical. Obtain an expression for pressure difference between inside and outside a spherical liquid drop. 4

23. A transverse harmonic wave on a string is given by

$$y(x,t) = 3.0\sin(36t + 0.018x + \pi/4)$$

where x and y are in cm and t in seconds. The positive direection of x is from left to right.

A. What is the speed and direction of propagation of the wave?

B. What are its amplitude and initial phase?

C. What is the least distance between its two successive crests? 4

24. An alternating voltage $E = E_0 \sin wt$ is applied to a circuit comprising of an inductor (L), a capacitor (C) and a resistor (R) in series. Obtain an expression for (i) impedence of the circuit and (ii) phase angle between voltage and current. 4

25. Distinguish between angular dispersion and dispersive power. for a prism of angle A = 60°, the angle of minimum deviation is also A. Calculate its refractive index. 4

26. What is mass defect? The mass of the nucleus of ${}^{14}_{7}N$ atom is 14.00307 u. Calculate mass defect and binding energy per nucleon. Take mp = 1.00727 u, mn = 1.00865 u and 1u = 931 meV. 4

or

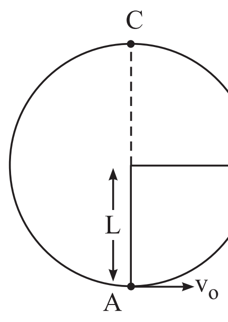
An electron, a protons and an alpha particle have the same kinetic energy, which of these particles has the shortest de Broglie wave and which have longest wavelength.

27. Define resistance. State the factors on which the resistance of a conductor depends. How are resistances connected in (i) series (ii) parallel. Find the expressions for equivalent resistance of two different resistors in these two arrangements. 6

28. What is meant by rectification? State the principle of a rectifier. With the help of a circuit diagram, explain the working of a half wave rectifier. Draw the input and output waveforms for a half wave rectifier. 6

29. What is meant by interference of light? Describe Young's double slit experiment to obtain interference pattern on a screen. Show that the intensity of the resultant wave in Young's experiment is proportional to the square of the amplitude of the incident wave. Discuss the conditions for constructive and destructive interference. 6

30. A bob of mass m, tied to a string of length L is rotated in a vertical circle in such a manner that the string has zero tension at point C, as shwon in figure. Find the ratio of the kinetic energies at points B and C. 6



or

A body of mass 20 kg is initially moving with a speed of 5 ms^{-1}. A force of 40 N is applied on the body for 3 seconds.

(i) find the final speed of the body after 3 seonds.

(ii) What is the distance covered during this period?

(iii) How much work has been done during this period?

(iv) Find the initial K.E. of the body.

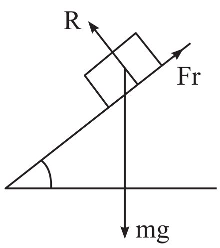
(v) Find the final K.E. of the body.

(vi) Show that the work done is equal to the change in K.E. for the body.

MARKING SCHEME

1. A
2. D
3. A
4. B
5. D
6. C
7. C
8. C
9. B
10. B
11. Statement of the law (module 1 p 65) 1

    Direction according to the law (module 1 p 65) 1

12. 

1

    $F_r = mg \sin\theta$

    or $F_r \leq \mu mg \cos\theta$ 1

13. According to stoke's law 1

    $F = 6\pi\eta rv$

    $\eta = \frac{F}{6\pi rv} = \frac{kgms^{-2}}{m \cdot ms^{-1}} = kg\ m^{-1}s^{-1}$ 1

14. Engine (ii) is more efficient $\frac{1}{2}$

    $\because \eta = \frac{W}{Q}$ $\frac{1}{2}$

Q is same for both $\frac{1}{2}$

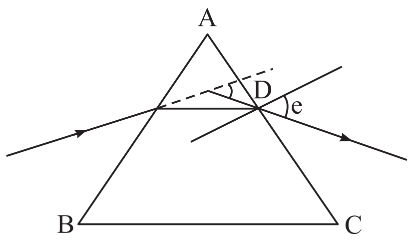
and w = area enclosed in p-v diagram which is more in case of II $\frac{1}{2}$

15. $\frac{\Delta\lambda}{\lambda} = \frac{v}{c}$ 1

$v = c\frac{\Delta\lambda}{\lambda} = 3\times10^8 \times \frac{0.03}{100} = 9\times10^4 \text{ ms}^{-1}$ 1

16. (a) By connecting a suitable low resistance in parallel to the galvanometer coil 1

(b) By connecting a suitable high resistance in serices with the galvanometer coil. 1

17. 

1+1

18. More energy per unit amu is released in fusion. 1

In fission it is less than 1 MeV/amu $\frac{1}{2}$

In fusion it is more than 6.7 MeV/amu $\frac{1}{2}$

19. N = 100, A = 3.85×10^{-3} m^2, B = (50 – 25) = 25 mT

$|e| = NA\left(\frac{B_2 - B_1}{t}\right)$ 1

$= 100\times3.85\times10^{-3}\times\frac{20\times10^{-1}}{250\times10^{-3}}$ $\frac{1}{2}$

= 3.85 × 10–2 V

= 38.5 mV $\frac{1}{2}$

20. Explanation of

(i) Isothermal process (p302) 1

(ii) Adiabatic process (p302) 1

(iii) Isobaric process (p303) 1

(iv) Isochoric process (p303) 1

Marking Scheme

21. $p_i = 0.05 \times 10 = 0.5 \text{ kg ms}^{-1} = p_f$

$p_i^x = 0.5 \cos 30 \qquad p_i^y = 0.5 \sin 30$ 1

$p_f^x = 0.5 \cos 30 \qquad p_f^y = 0.5 \sin 30$

Impulse $= p_f - p_i = \left(p_x^x i + p_f^y\right) - \left(p_i^x + p_i^y j\right)$ 1

$= \left(p_f^x - p_i^x\right) i$

$= -\ 0.866 \text{ kg ms}^{-1}$ 1

Impulse $= 0.866 \text{ kg ms}^{-1}$ $\frac{1}{2}$

22. Explanation 1

Derivation 3

23. (a) $v = \frac{w}{k} = \frac{36}{0.018} = 2000 \text{ cm s}^{-1} = 20 \text{ ms}^{-1}$ 1

The wave is travelling towards negative 1

x direction

(b) $a = 3.0$ cm ϕ_o = TA $\frac{1}{2} + \frac{1}{2}$

(c) Least distance between two succesion crests $\frac{1}{2}$

$= \lambda = \frac{2\pi}{0.018} = 3.5 \text{ m}$ $\frac{1}{2}$

24. Connection diagram $\frac{1}{2}$

Phase diagram $\frac{1}{2}$

Derivation of expression for z 2

Derivation of expression for tan θ 1

25. The diference between the angles of deviation for any two wavelengths (colours) is known as the angular dispersion for those wavelength 1

The ratio of the angular dispersion to the mean deviation is known as the dispersive power (w) of the material of the prism $\frac{1}{2}$

$$w = \frac{s_v - \delta_r}{s_y} \qquad \frac{1}{2}$$

$$= \frac{\sin\left(\frac{A + \sin}{2}\right)}{\sin\frac{A}{2}} \qquad 1$$

$$= \frac{\sin 60°}{\sin 30°} \qquad \frac{1}{2}$$

$$= \sqrt{3} \qquad \frac{1}{2}$$

26. The mass of the nucleus of an atom of any element is always found to be less than the sum of the masses of its constituent nucleon. The difference between the sum of masses of nucleon and the mass of the nucleus is called mass defeact. 1

$$\Delta m = 7mp + 7mn - m\left({}^{14}_{7}N\right) \qquad \frac{1}{2}$$

$= 7 \times 1.00727 + 7 \times 1.00865 - 14.00307$

$= 7.05089 + 7.06055 - 14,00307$

$= 14.11144 - 14.00307$

$= 0.10837$ u 1

BE $= \Delta m \times 931$ $\frac{1}{2}$

$= 100.89247$ Mev $\frac{1}{2}$

$$\text{BE/amu} = \frac{\Delta m \times 931}{A}$$

$$= \frac{100.89247}{14} = 7.206 \text{ MeV} \qquad \frac{1}{2}$$

or

26. $\lambda = \frac{h}{P}$ $\frac{1}{2}$

$$\text{K.E.; } E = \frac{1}{2}mv^2 = \frac{m^2v^2}{2m} = \frac{p^2}{2m} \qquad \frac{1}{2}$$

Marking Scheme

$\Rightarrow \quad \lambda = \dfrac{h}{\sqrt{2mE}}$ $\frac{1}{2}$

$\Rightarrow \quad \lambda \propto \dfrac{1}{\sqrt{m}}$ $\frac{1}{2}$

of the 3 particles given α-particle is the heaviest and electron is the lightest particle. 1

$\therefore$ λe will be greatest and $\lambda\alpha$ will be shortest. $\frac{1}{2}+\frac{1}{2}$

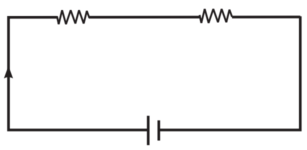
27. Resistance of a conductor is the ratio of P.D applied across it and the current flouring through it

i.e. $R = \dfrac{V}{I}$ $\frac{1}{2}$

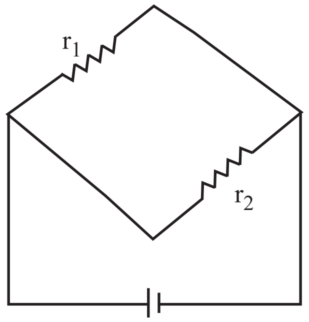
Resistance of a conductor depends on

Nature of material, length, area of cross section and temperature of the counductor 1

(i) resistors connected end to end such that same current flows through each of them $\frac{1}{2}$



(ii) resistors connected in such a way that one end of all the conductors is connected to positive terminal of the battery and the other end of negative terminal. In this case same P.D. is applied accross all the resistors. 1



In series combination $V = V_1 + V_2$

$= Ir_1 + Ir_2$

$$= I(r_1 + r_2)$$

$$\frac{V}{I} = R = r_1 + r_2 \qquad 1$$

In parallel combination $\quad I = I_1 + I_2$

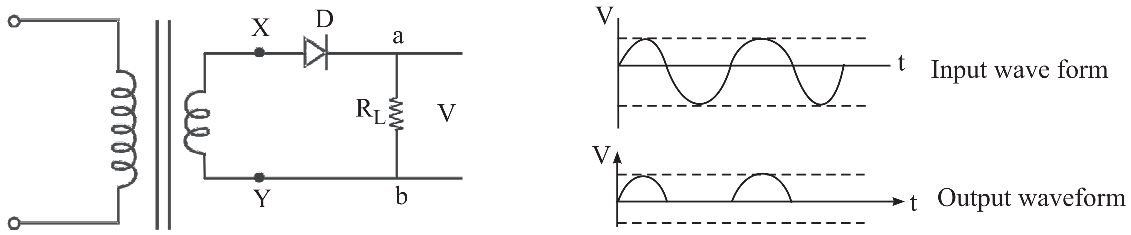
$$= \frac{V}{r_1} + \frac{V}{r_2} = V\left(\frac{r_1 + r_2}{r_1 r_2}\right) \qquad 2$$

$$R = \frac{V}{I} = \frac{r_1 r_2}{r_1 + r_2}$$

28. The process of conversion of ac into dc is called rectification 1

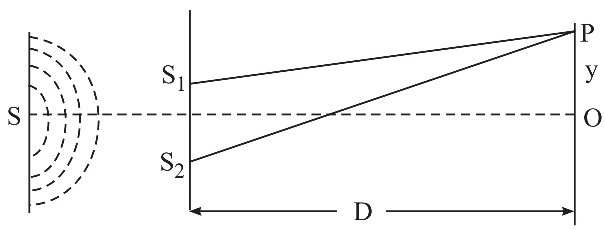
A pn junction diode conducts when it is forward biased and does not conduct in reverse bias 1

for the input wave form shown the diode is forward biased during the first half cycle $(0 - T/2)$ and hence conducts and carrent flows through R_C from a to b. During the next half cycle $(T_2 - T)$D is reverse biased and hence one current flows through R_C during $(T - 3T/2)$ again current from through R_C from



2 + 2

29. The phenomenon of redistribution of energy in space due to super position of light waves from two coherent sources. 1



Description of experimental set up of Young's double slit experiment

Derivation of A = $2a\cos\left(\frac{\delta}{2}\right)$ $1\frac{1}{2}$

Marking Scheme

$$I \propto A^2$$

$$\propto 4A^2 \cos^2\left(\frac{\delta}{2}\right)$$ $\frac{1}{2}$

Constructive interference

$$I_{max} = 4a^2$$

when $\cos^2\delta/2 = 1$

$$\cos \delta/2 = 1$$

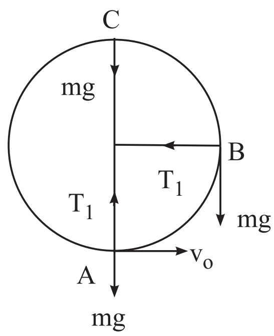
$$\delta = 0, 2\pi, 4\pi ... 2n\pi$$ $\frac{1}{2}$

Destructive interference

$$I_{min} = 0$$

when $\cos^2 \delta/2 = 0$

$$\delta = \pi, 3\pi, 5\pi ... (2n + 1)\pi$$ $\frac{1}{2}$



30. $T - mg = \frac{mv_o^2}{L}$ $\frac{1}{2}$

$$0 - mg = \frac{mv_c^2}{L} \Rightarrow v_c = \sqrt{gL}$$ 1

$$T_1 - 0 = \frac{mv_B^2}{L}$$ $\frac{1}{2}$

$$\frac{1}{2}mv_o^2 - \frac{1}{2}mv_B^2 = mgL$$ 1

$$\Rightarrow \quad v_B^2 = 3gL$$ $\frac{1}{2}$

$$\frac{k\varepsilon \text{ at B}}{k\varepsilon \text{ at c}} = \frac{\frac{1}{2}mv_B^2}{\frac{1}{2}mv_c^2} = \frac{v_B^2}{v_c^2} = \frac{3gL}{gL} = \frac{3}{1}$$ $1\frac{1}{2}$

$= 3 : 1$

or

(i) $a = \frac{F}{m} = \frac{40}{20} = 2ms^{-2}$ $\frac{1}{2}$

$v = u + at$

$= 5 + 2 \times 3 = 11\ ms^{-1}$ $\frac{1}{2}$

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

$= 5\times3+\frac{1}{\not{2}}\times\not{2}\times(3)^2$

$= 15 + 9 = 24$ 1

(iii) $W = FS$

$= 40 \times 24$

$= 960$ J 1

(iv) Initial K.E. $= \frac{1}{2}mu^2$

$=\frac{1}{\not{2}}\times \overset{10}{\not{20}}\times(5)^2$

$= 250$ J 1

(v) Final K.E. $= \frac{1}{2}mv^2$

$=\frac{1}{\not{2}}\times \overset{10}{\not{20}}\times(11)^2$

$=1210$ J 1

(vi) Change in K.E. = 1210 – 250 = 960J

Also work done = 960 J

So work done = change in K.E. of the body 1

For more information please visit QR code

www.bseh.ac.in

contact@bseh.ac.in