ELECTROMAGNETISM, MODERN PHYSICS, WAVES AND OPTICS

(For First Year University Students of Science & Engineering)

Physics Writers Series Creation

TABLE OF CONTENTS

CHAPTER 22: WAVE THEORY OF LIGHT REFLECTION, REFRACTION, INTERFERENCE, DIFFRACTION AND POLARIZATION

CHAPTER I
ELECTRIC CHARGE, POTENTIAL AND FIELD

1.0 Introduction

About a century ago, there were no electric lights, motors, radios, and television sets. Today, electricity is in such common use that we tend to give it little thought. Although the practical applications of electricity have been developed mostly in the twentieth century, the study of electricity has a long history. Observations of electrical attraction can be traced back to the time of the ancient Greeks. It was Thales (640-540 BC) who observed for the first time that a body could be made to attract other bodies by rubbing it with another bodies. Thales and others noticed that after amber has been rubbed, it attracts other subjects such as feathers. The Greek word for amber is elektron. Thus, a body made attractive by rubbing is said to be electrified. This branch of electricity was earliest to be discovered and it is known as Electrostatics. *Electrostatics is the study of electrical phenomena that are associated with charges and charge systems at rest.*

In this chapter, we will begin our study of electricity with a discussion of the concept of electric charges and how electric charge is associated with material bodies such as conductors, semi-conductors and insulators. We will then study Coulombs Law which describes the force exerted by one electric charge on another, after which we will introduce the concept of electric field and show how it can be described by the electric field lines which indicate the magnitude and direction of the field at any point. Lastly, we will discuss the behaviour of point charges including their motion in an electric field as well as electric dipoles in electric fields.

1.1 Electric Charge

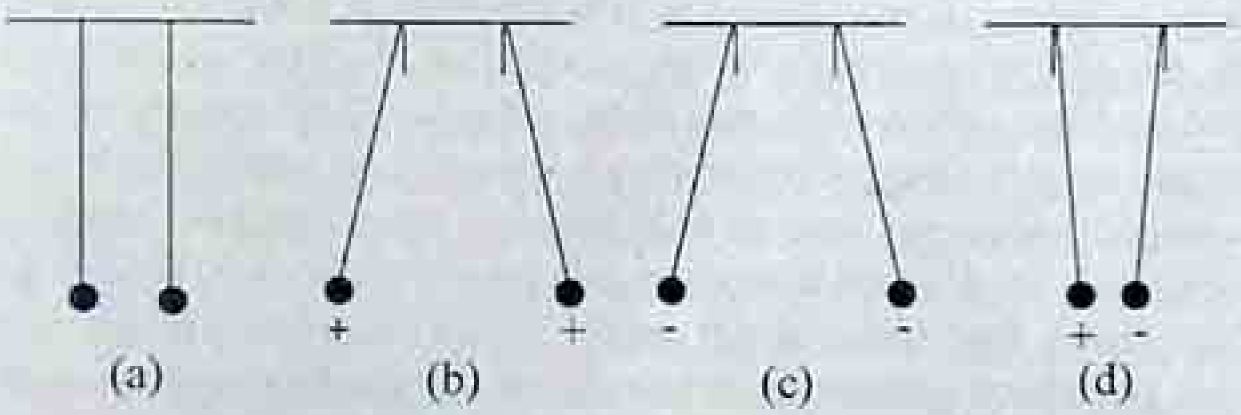


Fig.1.1: Charge of pith balls

As mentioned earlier, the early Greeks knew that amber when rubbed with a woolen cloth attracts small objects placed near it. The rubbing of the amber with a woolen cloth has enabled the amber to acquire electric charges which makes it possible to attract small objects. Many substances can be made to acquire electric charges in the same way. If, for example, we rub a hard rubber rod with fur or a glass rod with silk, we find that we can readily pick up small pieces of paper with them. In each case, we say that the rod is charged and it is said to possess electric charges. In our studies of mechanics, we introduced three fundamental physical quantities– these are length, mass and time. Here, we introduce the electric charges as one more fundamental physical quantity. Its unit is the coulomb.

Let us perform some experiments that will reveal the type of charges acquired by these rods when rubbed vigorously with either silk cloth or fur. Let us suspend two pith balls with silk thread as shown in Figure 1.1(a). Let us touch each ball with a glass rod that has been vigorously rubbed with silk. It is found as shown in Figure 1.1(b) that the balls repel each other and also repel the charged glass rod. Now if the uncharged pith balls are touched with a hard rubber rod which has been rubbed with fur, they again repel each other and repel the charged rubber rod. This is shown in Figure 1.1(c). However if we touch one pith ball with a charged rubber rod and the other with a charged glass rod, we find that the pith balls attract each other as shown in Figure 1.1(d).

We also find that the charged rubber rod attracts a charged glass rod. The electric charge on the glass rod must therefore be different from that on the rubber rod. It is obvious that there are two kinds of charges; one that appears on the hard rubber rod when rubbed with fur and the other which appears on a glass rod when rubbed with silk. The two kinds of charges attract each other. It was Benjamin Franklin an American Physicist (1706-1790) who named the electric charge on the glass rod as positive charge while that which appears on the charged rubber rod as negative. We can sum up from the foregoing experiment that like charges (positive or negative) repel and unlike charges (positive and negative) attract.

1.2 Types of Charge

There are two types of electric charge called positive and negative charge. Materials such as polythene and cellulose acetate can be charged by friction. When a strip of polythene is rubbed with a dry woolen material, it becomes negatively charged whereas a strip of cellulose acetate rubbed in the same manner with a cotton material would be positively charged.

The formation of the charges on the polythene (or cellulose acetate) stripped by friction can be explained in terms of movement of the electrons. When a polythene strip is rubbed with wool, some of the loosely-held electrons in wool are transferred to the polythene. The polythene now has an excess of electrons and hence becomes negatively charged. On the other hand, the woolen material having lost some electrons ends up with more positive than negative charges. It becomes positively charged. The positive charge on the woolen material is of the same magnitude as the negative charge on the polythene. Electric charge is always *conserved* in a *closed* system. The negative charge lost by the woolen material is equal to the negative charges gained by the polythene strip.

Similarly, when a cellulose acetate strip is rubbed with a piece of cotton cloth, electrons are transferred from the cellulose acetate to the cotton cloth. The cellulose acetate strip becomes positively charged and the cotton cloth negatively charged. *Like charges repel and unlike charges attract.* A polythene strip is charged by rubbing with a woolen cloth and hung from a nylon thread. When another charged polythene strip is held close to the suspended strip, the suspended strip is repelled. A similar result is obtained using two cellulose acetate strips which had been rubbed with a cotton cloth. This shows that like charges repel. When two different strips are used, example, a negatively charged polythene strip and a positively charged cellulose acetate strip, with one of them suspended, the suspended strip is attracted. This shows that unlike charges attract.

1.3 Electrical Conductors and Insulators

Materials such as copper, aluminium and silver permit electric charge to move easily from one part of a material to another. Such materials are known as good *conductors*. Other materials such as glass and plastic have no *free electrons* and charges bound to the atom and hence cannot move freely through the material. These are called *insulators*. A third group of material called *semi-conductors* are intermediate in their properties between good conductors and insulators with resistivity between that of a conductor and an insulator. Examples of such materials are silicon, carbon and germanium. Resistivity of good conductors is $10^{-8}\,\Omega m$. That of insulators is $10^{4}\,\Omega m$ and that of semi-conductors is $10^{-1}\,\Omega m$.

A material can be tested for electrical conduction by holding it in the hand and touching the metal cap of a charged electroscope known as Gold-leaf electroscope. When the metal cap is touched with a good conductor, the leaf collapses instantly. A poor conductor will cause the leaf to go down slowly. For an insulator, the divergence of the gold leaf goes down slightly but the gold leaf diverges again when the insulator is removed. Table 1.1 lists some examples of electrical conductors and insulators.

1.4 Electrostatic Hazards

When a person wearing rubber soled shoes moves about in a carpeted room, friction between the rubber soles and the synthetic fabric of the carpet causes charge to build up in the person's body. If the person reaches a metal door knob, he/she can feel a slight tingle. This occurs as a result of the movement of electric charges between the fingers and the door knob just before he/she touches the knob.

Table1.1: Electrical conductors and insulators

Conductors	Semiconductors	Insulators
Silver	Carbon	Plastic
Copper	Germanium	Glass
Aluminium	Silicon	Paper

In operating theaters, the gases used to anaesthetize patients are easily flammable. Static charges produced by the movements of blankets or clothing can cause sparks and consequently fire. As a precaution, anesthetic machines, trolleys and patients are linked to the ground by metal chains. In addition, doctors and nurses wear antistatic conducting shoes.
The same principle to prevent the buildup of static charges in operating theaters is used in petrol tankers. As the lorry moves through the air, friction between the lorry and the air causes electrostatic charges to build up on the lorry. A spark from this charge can cause the petrol vapor to explode. To prevent any explosion, a metal chain hangs from the lorry and touching the ground provides a path to the earth for the charges produced on the lorry.
Similarly, charges buildup on an aeroplane as it flies through the air. The charges have to be earthed when the plane lands. This is done by using special conducting rubber for the tires or by having a metal chain to conduct charges quickly to the ground.

1.5 Electric Charge and Structure of Matter

Let us review briefly the atomic structure of matter. This will make it clear how the electric charge (either positive or negative) is associated with material bodies. Ordinarily, matter is made up of three kinds of elementary particles: these are the electron which is negatively charged, the positively charged proton and the neutral neutron. These are the building blocks that form the atoms. The atoms combine with each other in different ways to form the molecules. The atom, from scientific studies has two parts to its structure– a heavy, positively charged nucleus surrounded by one or more negatively charged electrons. The nucleus, on the other hand is a closely packed group of protons and neutrons. Ordinarily, the positive and negative charges within the atom are equal, so the atom is said to be electrically neutral. Some of the electrons surrounding the nucleus may not be tightly bound to the nucleus, hence the atom is no longer electrically neutral but it will have a net positive charge and it is called an *ion*. If instead of losing an electron, the atom gains an electron or more, it will then have a net negative charge. It is therefore clear that atoms of some elements release electrons with relative ease while others accept electrons with ease.
The protons of the nucleus are not free to leave the nucleus just like the electrons. Hence, bodies become negatively charged only when they loose electrons. Thus, a glass rod rubbed with silk loses electrons and acquires a positive charge, while the silk gains the lost electrons and becomes negatively charged. It is clear that since the transfer of a certain number of electrons is responsible for the charging of both bodies, the positive charge on the glass rod and the negative charge on the silk must be equal in magnitude. It is important for us to realize that electric charges are not generated or created during the process of rubbing, but are merely transferred between the two bodies involved. However, the net charge produced is zero. This is the well- known *law of conservation of electric charges which states that the total quantity of charge in any process does not change*.

We have limited our discussions so far to amber, hard rubber and glass. Suppose we have two metal spheres mounted on a glass pedestal. One of the spheres is charged and the other is neutral. If we connect the two spheres with a bare copper wire, we will find that uncharged sphere has become charged and having the same type of charge as the charged sphere. If on the other hand, the two spheres are connected using a piece of dry wood, the uncharged sphere remains uncharged. Substances like the copper wire that conduct electricity are known as conductors while those like the dry wood and glass are non- conductors or insulators. Metals are generally good conductors while non-metals are non-conductors or insulators. Almost all naturally occurring materials fall into one or the other of these categories – conductors or insulators. However, there are few materials such as silicon and germanium that do not fall into either of the two distinct categories. They are called *semi-conductors*. We shall discuss this class of materials in greater detail in the later chapters. But for now, let us note that semi–conductors have very few electrons at low temperature, but the number of free electrons increases rapidly with increase in temperature.

1.6 Induced Charges

We saw in section 1.1 that an uncharged body can be charged by bringing it in contact with a charged body. Let us try to understand how this works. Suppose we have two metal plates (conductors), one is positively charged and the other is neutral. If we bring the two plates together and they touch each other, the electrons (negatively charged) in the neutral plates are attracted to the positively charged plate and some of these electrons will pass over it. Since the neutral plate now has a shortage of electrons, it will have a net positive charge. The former neutral plate is now charged by the flow of its electrons and this process of charging is known as *charging by conduction*.

Similarly, if the original plate is negatively charged and it is brought in contact with the neutral plate, some of the excess electrons in the charged plate will flow to the uncharged plate. It will then have a net charge and will be negatively charged.
However, we can charge a conductor without touching it with a charged body as outlined above. This process is known as *charging by induction*. Let us see how this method works. The required series of steps is illustrated in Figure 1.2. Figure 1.2 shows a metal conductor mounted on an insulating metal globe and a hard rubber rod that has been charged by rubbing fur. If the charged rod is brought near the conductor without actually touching it as shown in Figure 1.2(a), the electrons in the conductor are repelled to the end of the conductor farthest from the charged rod. This makes the end nearest the rod positively charged.

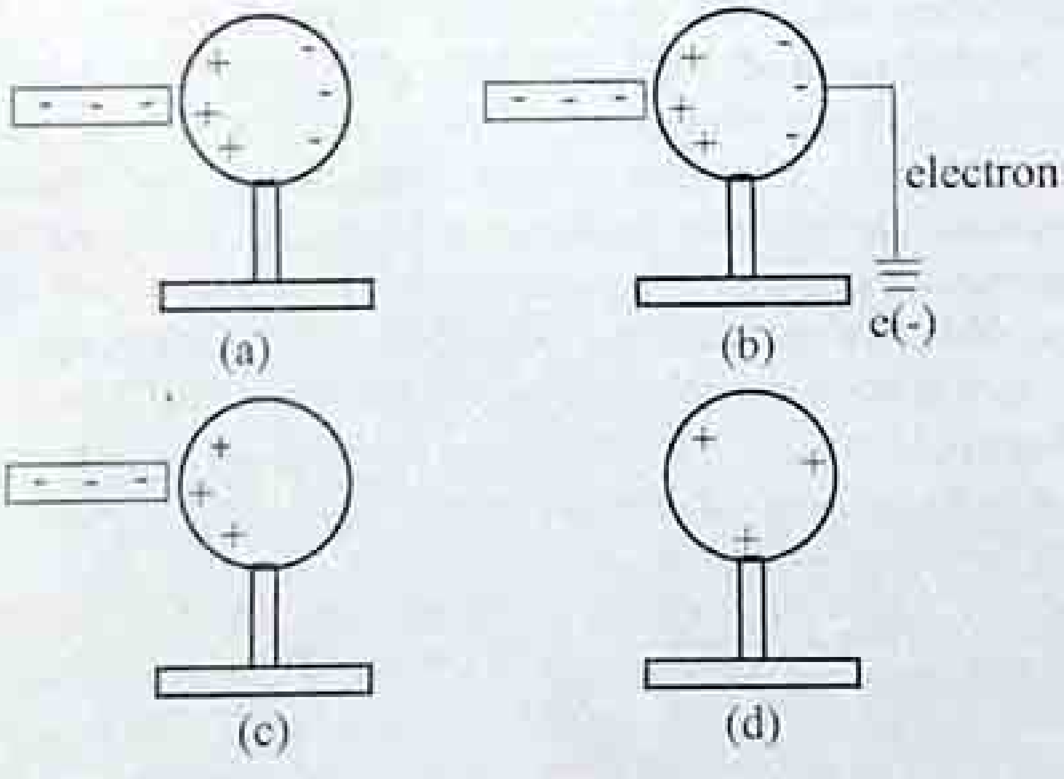


Fig. 1.2: Charging metal conductor by induction

Suppose we now connect the conductor to the ground with a conducting wire as shown in Figure 1.2(b). We now say that the conductor is grounded or earthed. The symbol ⏚ means grounded. Since the earth is so large, it can easily accept electrons or give up electrons without affecting its charge neutrality. It behaves like a reservoir. The repelled electrons to one end of the conductor will then move down the wire to the earth. This leaves the conductor positively charged. If the connecting wire is now removed, as shown in Figure 1.2(c), the conductor is left positively charged. Now when the charging rod is removed, the positive charges in the conductor will distribute themselves in a normal fashion, but the conductor remains positively charged. If the charging rod is removed before the connecting wire is cut, electrons will flow from the earth into the conductor, thereby making it neutral. It is therefore necessary to follow the sequence of steps.

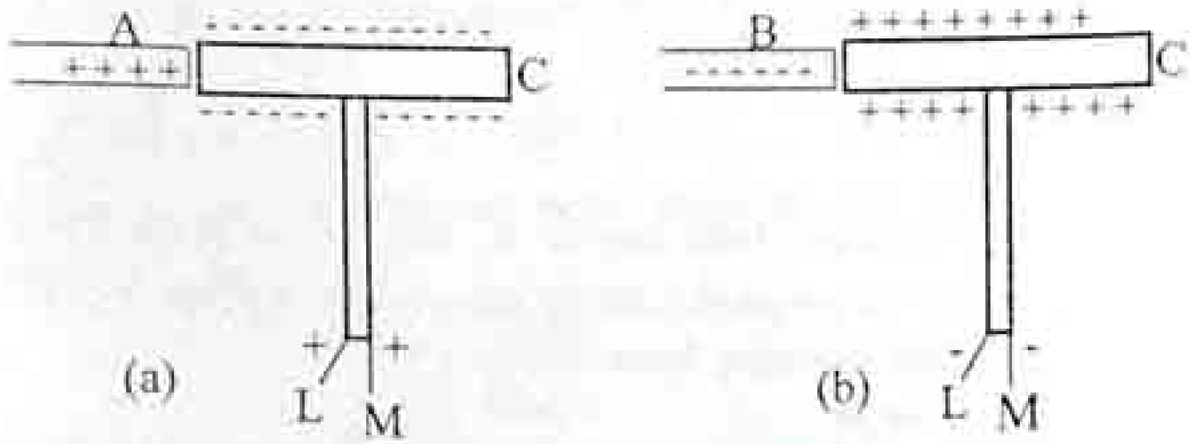


Fig. 1.3: (a) The gold leaf of an electroscope diverges when it is touched with a charged rod. (b) The gold leaf of an electroscope diverges when a charged object isbrought near the electroscope.

If we had started with a positively charged rod (glass rod rubber with silk), the conductor will be negatively charged. In this case, electrons will flow from the earth into the conductor and give the body a net negative charge.

Let us now use the electroscope to demonstrate charging by induction. An electroscope is a device that can be used for detecting the presence, and measures the quantity of charge. A simple example is the goldleaf electroscope which consists of two strips of gold leaves which are moveable. The leaves are connected by a metal rod to a metal knob outside the case. The rod is insulated from the case. When a charged rod is brought near the metal knob, a separation of charge takes place as shown in Figure 1.3(a); the two leaves become positively charged and repel each other.

When the charging glass is removed the charges in the conductor no longer remain separated and the leaves collapse. If we now touch the metal knob with the charging rod, the whole apparatus becomes charged by conduction and the leaves diverge. This is shown in Figure 1.3(b). The angular divergence of the leaves is proportional to the quantity of charge on the leaves. Therefore, a calibrated electroscope can be used to measure the quantity of charge.

Before we continue, it is necessary to have an idea of the magnitude of charges on bodies. It has been shown that the proton is about 2000 times more massive than the electron. Yet, their charges are exactly equal but opposite in sign. The charge of proton is $+e$ and that of the electron $-e$, where e is called the *fundamental unit of charge*. All charges occur in integral amounts of the fundamental unit of charge e. That is, charge is quantized. Any charge, Q occurring in nature can be written as $Q = \pm Ne$ where N is an integer. The quantization of electric charge is usually not noticed because N can be very large. For example, charging a plastic rod by rubbing it with a piece of fur typically transfers about 10^{10} electrons to the rod. The S.I unit of charge is coulomb, which will be defined later in terms of the unit of electric current, the ampere. The *coulomb (C) is the amount of charge flowing through a cross- sectional area of a wire in one second when the current in the wire is one ampere.* The fundamental unit of the electric charge, i.e. e is related to coulomb as shown below:

$$e = 1.60 \times 10^{-19} C$$

Charges from about $10nC(1nC = 10^{-9} C)$ to about $0.1\mu C(1\mu C = 10^{-6} C)$ can be produced in the laboratory by putting certain objects in intimate contact, often simply by rubbing their surfaces together. Such a procedure involves the transfer of many electrons.

Example 1.1

A small copper plate has a mass of 3g. The atomic number of copper is $Z = 29$ and the atomic mass is $63.5\, g/mol$. What is the total charge of all electrons in the copper plate?

Solution

We must first of all find the number N, of atoms in 3g of copper. Since 1 molecule of copper contains Avogadro's number of atoms and has a mass of 63.5g, the number of atoms in 3g of copper is (where Avogadro's constant $= 6.02 \times 10^{23}$).

$$N = \frac{3 \times 6.02 \times 10^{23}}{63.5} = 2.84 \times 10^{22}\, atoms$$

Each atom contains $Z = 29$ electrons, so the total charge,

$$Q = 2.84\times10^{22}\times29\times(1.6\times10^{-19}) = -1.32\times10^{5}\,C$$

1.7 Coulomb's Law

In gravitation, the force of attraction between two masses m_1 and m_2 is given by Newton's law of gravitation which states *that the force is proportional to the product of the masses and inversely proportional to the square of the distance between them*,

i.e. Gravitational force, $F = \dfrac{Gm_1m_2}{r^2}$

where G is the universal gravitational constant.

The force between two charged objects can either be attractive or repulsive, depending whether the charges are like or unlike. The magnitude of the force between two point charges obeys a similar inverse- square relationship with distance as in gravitational force.

This relationship is given by *Coulomb's Law* which states that *the force between two point charges is directly proportional to the product of the charges and inversely proportional to the square of the distance x between them*,

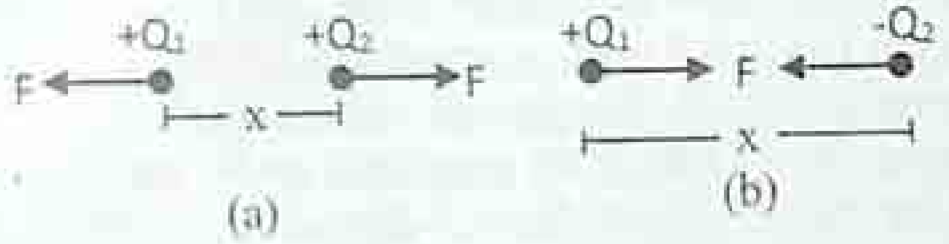


Fig1.4: Coulomb's Law

i.e. Force, $$F \propto \frac{Q_1Q_2}{x^2} \tag{1.1}$$

Electrostatic forces between two charges exist in pairs and satisfy Newton's third law of motion. In Figure 1.4, the force on the charge Q_1 is equal to the force on the charge Q_2 but the forces are acting in opposite directions.

From equation (1.1), the force between two charges Q_1 and Q_2 at a distance x apart may be written as

$$F = k\frac{Q_1Q_2}{x^2} \tag{1.2}$$

where k = the constant of proportionality.

If the changes are in a vacuum, the constant k in S.I. units is written as

$$k = \frac{1}{4\pi\varepsilon_0} = 9.0\times10^{9}\,Nm^2C^{-2} \tag{1.3}$$

where ε_0 is another constant known as the permittivity of free space and the value of ε_0 is $8.85\times10^{-12}\,Fm^{-1}$. Hence in free space or a vacuum,

$$F = \frac{Q_1Q_2}{4\pi\varepsilon_0x^2} \tag{1.4}$$

If the charges are in an insulating medium, the force, F is given by

$$F = \frac{Q_1Q_2}{4\pi\varepsilon x^2} \tag{1.5}$$

where ε is known as the permittivity of the medium.

The relation between ε and ε_0 is

$$\varepsilon = \varepsilon_0\varepsilon_r, \tag{1.6}$$

where ε_r is a non-dimensional constant known as the relative permittivity or dielectric constant of the medium.

The dielectric constant of air ε_r is 1.005. Since ε_r for air is approximately 1.0, the value of the permittivity ε for air maybe assumed to be equal to the value of ε_0. Hence within experimental errors, the equation

$$F = \frac{Q_1Q_2}{4\pi\varepsilon_0 x^2}$$

can be used for charges in air.

Table 1.2 shows the values of the dielectric constant ε_r of some materials.

Table1.2: Dielectric constant ε_r at 20°C

Material	ε_r
Vacuum	1,000 00
Air	1,005
Water	81
Paraffin	2.1
methyl alcohol	38
acetone	27
Metals	∞

Example 1.2

By assuming that a hydrogen atom consists of an electron of charge e^- orbiting a proton of charge e^+, calculate the ratio of the electric force between the electron and proton to the gravitational force between them. [mass of electron, $m_e = 9.1\times10^{-31}kg$, mass of proton, $m_p = 1.67\times10^{-27}kg$ and $G = 6.67\times10^{-11}Nm^2kg^{-2}$]

Solution

Using Coulomb's law, the electric force, $F_e = \dfrac{e^2}{4\pi\varepsilon_0 x^2}$

where x = radius of the orbit.

The gravitational force, $F_g = \dfrac{Gm_e m_p}{x^2}$

$$\therefore \quad \frac{F_e}{F_g} = \frac{e^2}{4\pi\varepsilon_0 x^2}\left(\frac{x^2}{Gm_e m_p}\right) = \frac{(1.6\times10^{-19})^2}{4\pi(8.85\times10^{-12})(6.67\times10^{-11})(9.1\times10^{-31})(1.67\times10^{-27})}$$

$$= 2.27\times10^{39}$$

The above example shows that the force between the orbiting electron and the nucleus is electric in nature. The gravitational force is negligible.

Equation (1.4) gives the magnitude of the force each charged body exerts on one another. The direction of the force is along the line joining the two bodies. We already know that if the two charges have the same sign, the force on each body is repulsive and therefore, directed away from the other. However, if the two bodies have opposite charges, the force on each body is attractive, and it is thus directed toward the other.

Coulombs law applies to particles e.g. electrons and protons as well as small charged bodies provided that their sizes are much less than their distances of separation (point charges). Equation (1.4) resembles Newton's law of gravitation in which the gravitational constant G is replaced by $1/4\pi\varepsilon_0$ and the electric charges are analogous to the gravitating masses. If the charges are situated in a medium other than air, Equation 1.4 takes a more general form:

$$F = \frac{Q_1Q_2}{4\pi\varepsilon r^2} \quad \text{as in equation 1.5.}$$

where ε is the permittivity of the medium. It thus follows that the force between any two given charges depends on the permittivity of the medium. For example, the force between any two charges is eighty times less in water than in free space.

Example 1.3

(a) Compare the electric and gravitational forces of attraction between the electron and the proton in a hydrogen atom assuming that their distance of separation is $5.3\times10^{-11}m$.

(b) Hence or otherwise, show that this ratio is independent of the distance of separation between the particles ($m_e = 9.1\times10^{-31}kg$, $m_p = 1.7\times10^{-27}kg$, $e = 1.6\times10^{-19}C$ and $G = 6.69\times10^{-11}Nm^2kg^{-2}$).

Solution

The electric force between the particles is given by coulombs law:

$$F_E = \frac{1}{4\pi\varepsilon_0}\frac{Q_1Q_2}{r^2} = 9.0\times10^9\,Nm^2C^{-2}\times\frac{(1.6\times10^{-19}C)^2}{(5.3\times10^{-11}m)^2} = 8.2\times10^{-8}N$$

The gravitational force is given by the Newton's law of gravitation:

$$F_G = \frac{GM_1M_2}{r^2} = \frac{(6.69\times10^{-11}Nm^2kg^{-2}\times9.1\times10^{-31}kg\times1.7\times10^{-27}kg}{(5.3\times10^{-11}m)^2} = 3.7\times10^{-27}N$$

The ratio of the two forces is thus, $\dfrac{F_E}{F_G} = \dfrac{8.2\times10^{-8}}{3.7\times10^{-27}} = 2.2\times10^{-19}$

Thus, the electrical force is greater than the gravitational force by a factor of about 10^{19}. The reason why electrical forces are not as apparent as gravitational forces in nature is that, matter in nature appears electrically neutral. The ratio of the forces is

$$\frac{F_E}{F_G} = \frac{\dfrac{1}{4\pi\varepsilon_0}\dfrac{Q_1Q_2}{r^2}}{\dfrac{GM_1M_2}{r^2}} = \frac{Q_1Q_2}{4\pi\varepsilon_0 GM_1M_2}$$

Hence, the ratio of the two forces is independent of the distance between the two particles.

Example 1.4

What is the magnitude of the electrostatic force of attraction between an α-particle and an electron $10^{-13}m$ apart?

Solution

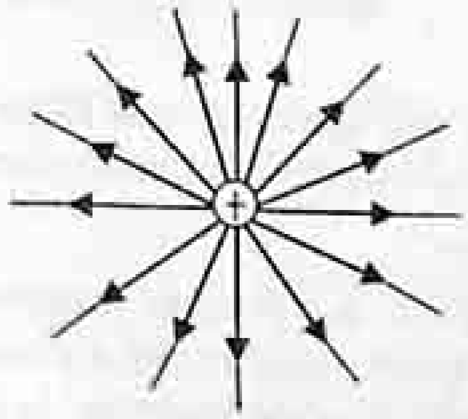
The charge on the α-particle is $-2e$ where e is the charge on one electron. The electrostatic force is given by:

$$F = \frac{1}{4\pi\varepsilon_0}\times\frac{2e^2}{r^2} = \frac{8.99\times10^9Nm^2C^{-2}\times2\times(1.6\times10^{-19})^2C^2}{(10^{-13}m)^2} = 4.6\times10^{-2}N$$

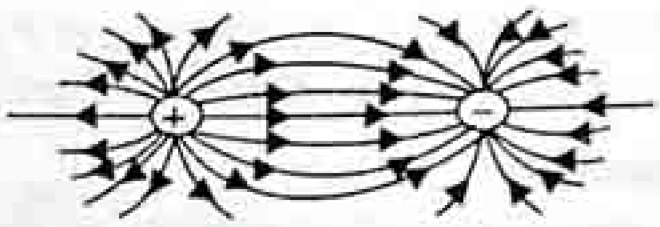
As we have seen from the two worked examples, if only the magnitude of coulomb's force is required, the arithmetic is very simple.

However, Coulomb's force is a vector quantity and therefore, a knowledge of the direction of the resultant force acting on a system of point charges is sometimes required. Such a problem is best carried out using vector algebra.

1.8 Electric Field



(a) Electric field pattern around a positive charge



(b) Electric field pattern between two unlike point charges

Fig. 1.5: Electric field pattern

An electric field is a region where a force acts on a charged body placed in the region. An electric field can be represented by lines of force. The lines of force in an electric field shows the paths along which a charged particle will move in the field. The direction shown on a line of force is the direction along which a free positive test charge would move. In an electric field, if the lines of force are close to each other, it denotes a strong field. Figure 1.5 shows a few electric field patterns.

The *electric field Intensity, E at a point in an electric field is the force per unit charge acting on a positive test charge placed at that point*. Each point in an electric field is associated a quantity known as the electric field intensity or electric field strength, E.
i.e.

$$\text{Electric field intensity, } E = \frac{\text{Force}}{\text{Charge}} = \frac{F}{q} \qquad (1.7)$$

The electric field intensity, E is a vector quantity, its direction is the direction of the force on a positive test charge placed at that point. The unit for electric intensity is Newton per coulomb (NC^{-1}) or volt per metre (Vm^{-1}).

Electric Field Intensity, E Due to a Point Charge
If there is a point charge $+Q$ at a point O, there exists an electric charge around the point O. The equation for the electric field intensity or electric field strength E at the point P, distance r from O, may be obtained by considering the force, F acting on a positive test charge $+q$ at P.
Using Coulomb's law, the force, F on the charge $+q$ at P is

$$F = \frac{Qq}{4\pi\varepsilon_0 r^2}$$

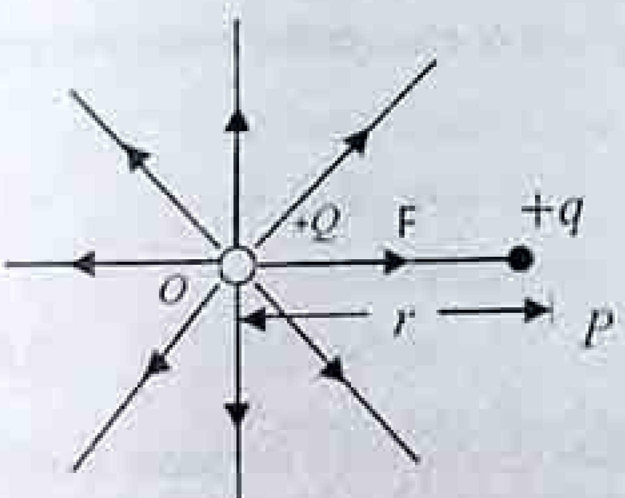


Fig 1.6: Electric field, E due $+Q$

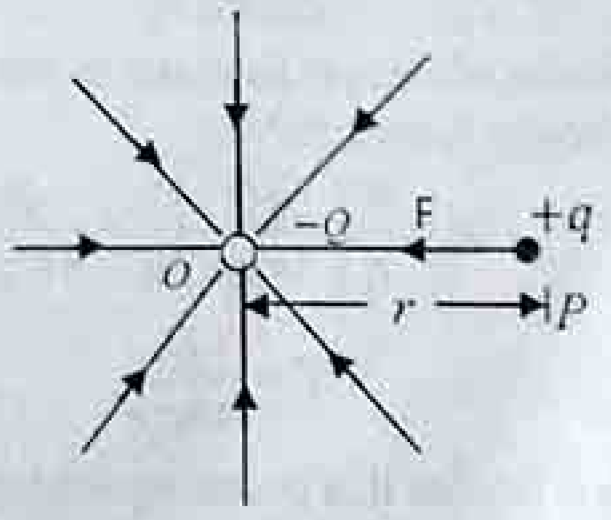


Fig 1.7: Electric field, E due to $-Q$

Hence the magnitude of the electric field at P, distance r from a point charge, $+Q$ or $-Q$ is:

$$E=\frac{F}{q}=\frac{Q}{4\pi\varepsilon_0 r^2} \qquad (1.8)$$

The direction of E is the same as the direction of F as shown in Figure 1.6 and Figure 1.7. In Figure 1.6 the direction is repulsion- away from O. In Figure 1.7, the direction is attraction-towards O.

1.9 Electric Potential

Another quantity associated with a point in an electric field is the electric potential. *The electric potential V at a point in an electric field is the work done to move a unit positive charge from infinity to point.* The electric potential at infinity is assumed to be zero. Electric potential is a scalar quantity and the S.I. unit is volt which is equivalent to joule per coulomb.

Electric Potential Due to a Point Charge

The electric potential at a point in an electric field due to a point charge can be derived using the principle of work done by a force.

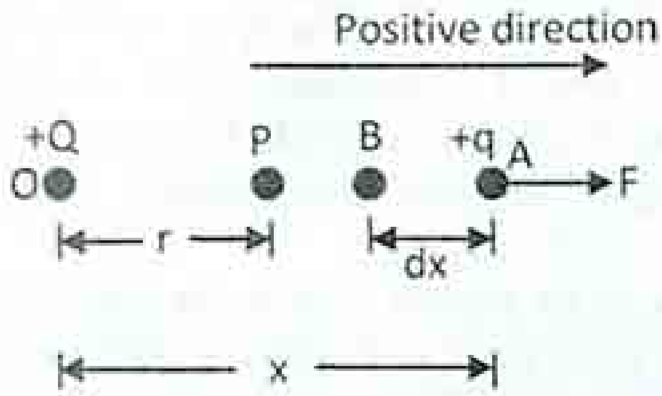


Fig: 1.8: Electric potential due to a point charge

Figure 1.8 shows a point charge $+Q$ at O. To derive the expression for the electric field potential V at the point P, distance r from charge $+Q$, consider a test charge $+q$ at that point A and distance x from O. Using Coulomb's law, the force on the test charge $+q$ is $F=\frac{Qq}{4\pi\varepsilon_0 x^2}$ in the positive direction which is the direction of increasing x.

The small work done, dW by the force F when the charge $+q$ is moved through a small displacement ($-dx$; negative because the displacement is in the opposite direction to F) is given by

$$dW=F(-dx)=\frac{-qQdx}{4\pi\varepsilon_0 x^2}$$

Hence the total work done to move the test charge $+q$ from infinity to the point P, distance r from a charge $+Q$, is

$$W=\int_{\infty} dW=\int_{\infty}\frac{-qQ}{4\pi\varepsilon_0 x^2}dx=\left[\frac{qQ}{4\pi\varepsilon_0 x}\right]_{\infty}^{r}=\frac{qQ}{4\pi\varepsilon_0 r}$$

Hence by the definition of electric potential as work per unit positive charge, the electrical potential at a point distance r from a charge $+Q$ is

$$V=\frac{W}{q}\left(\frac{qQ}{4\pi\varepsilon_0 r}\right)\frac{1}{q}=\frac{Q}{4\pi\varepsilon_0 r}$$

$$V=\frac{Q}{4\pi\varepsilon_0 r} \qquad (1.9)$$

The expression for V implies that points equidistant from a point charge have the same electric potential. These points in fact lie on the surface of a sphere whose center is at the point charge, and is known as an *equipotential surface.*

An equipotential surface is a surface where the points lying on it are all at the same potential. No work is done when a charge is moved from one point to another point of an equipotential surface.

Equipotentials cut the lines of force at right angles. Figure 1.9 shows the electric field and the equipotentials round a point charge.

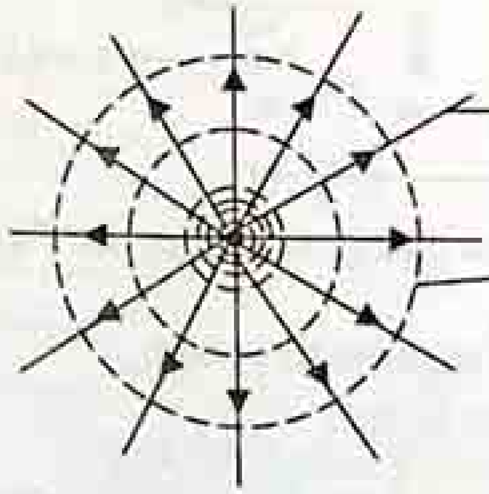


Fig. 1.9: Equipotential around a point charge

1.10 Electric Potential Energy

The *electric potential energy, U* of a charge at a point in an electric field is defined as *the work done to move the charge from infinity to that point.* A charge in an electric field has electric potential energy.

1.11 Potential Difference

The *potential difference* between two points in an electric field is *the work done per unit positive charge to move a positive charge from one point to another against the electric field.* Different points which are not on the same equipotential have different electric potentials. If V is the potential difference between two points in an electric field, the work done to move a test charge q against the electric field from one point to the other is

$$W = qV \qquad (1.10)$$

The work done on the charge q is changed into electric potential energy. The increase in electric potential energy is

$$\Delta U = qV$$

1.12 Relationship between E and V

The relationship between the electric field intensity E and the electric potential V can be established from the definition of E as force per unit positive charge and V as work done per unit positive charge. Figure 1.10 shows a test $+q$ at a point A distance x from the origin O.

If the electric field intensity at A is E, then the force acting on the charge $+q$ is

$$F = qE \qquad (1.11)$$

in the direction of E, the positive direction. To move the test charge $+q$ from **A** to **B**, through a small distance dx, the work done,

$$dW = force \times displacement = F(-dx) = -qEdx$$

Table 1.3: Magnitude of some electric fields, (E) in nature

Place	E (NC^{-1})
In household wires	10^{-2}
In radio waves	10^{-1}
In the atmosphere	10^{2}
In sunlight	10^{3}
Under a thundercloud	10^{4}
In a lightning bolt	10^{4}
In an X-ray tube	10^{6}
At the electron in a hydrogen atom	10^{11}

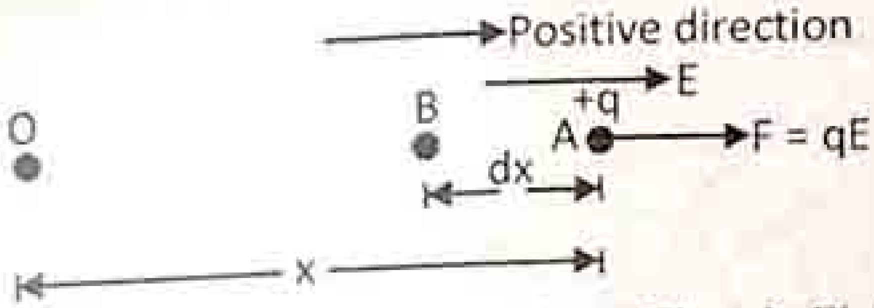


Fig. 1.10: Force on a charge in an Electric Field

The displacement is $(-dx)$ because it is in the negative direction.
If dV is the potential difference between A and B, then

$$dV = \frac{dW}{q} = -\frac{qEdx}{q}$$

$$\therefore \qquad E = -\frac{dV}{dx} = - \text{ (potential gradient)}$$

Or

$$dV = Edx$$

$$V = \int_{\infty}^{r} Edx \qquad (1.12)$$

1.13 Electric Field Lines

From equation 1.11 we have $F = qE$

and since E is a vector , the electric field is a vector field. We must then find a geometric way to enable us visualize the electric field. Let us determine the direction of the force exerted on a test charge q_0 by a positively charged particle Q_1. The situation is shown in Figure 1.11(a). We know that the positive charge Q_1 will exert a repulsive force on q_0. Therefore, the lines are drawn to indicate the direction of the force due to the given field on a positive test charge. In Figure 1.11(a), the lines point radially outward from the charge, while in Figure 1.11(b), the lines point radially inward toward the charge.

These lines which are imaginary indicate the directions of the force that would be exerted on a test charge in each case. These lines are known as lines of force or electric field lines. Since the electric field has the same direction as the force, the direction of the electric field is a tangent to the lines of the force.

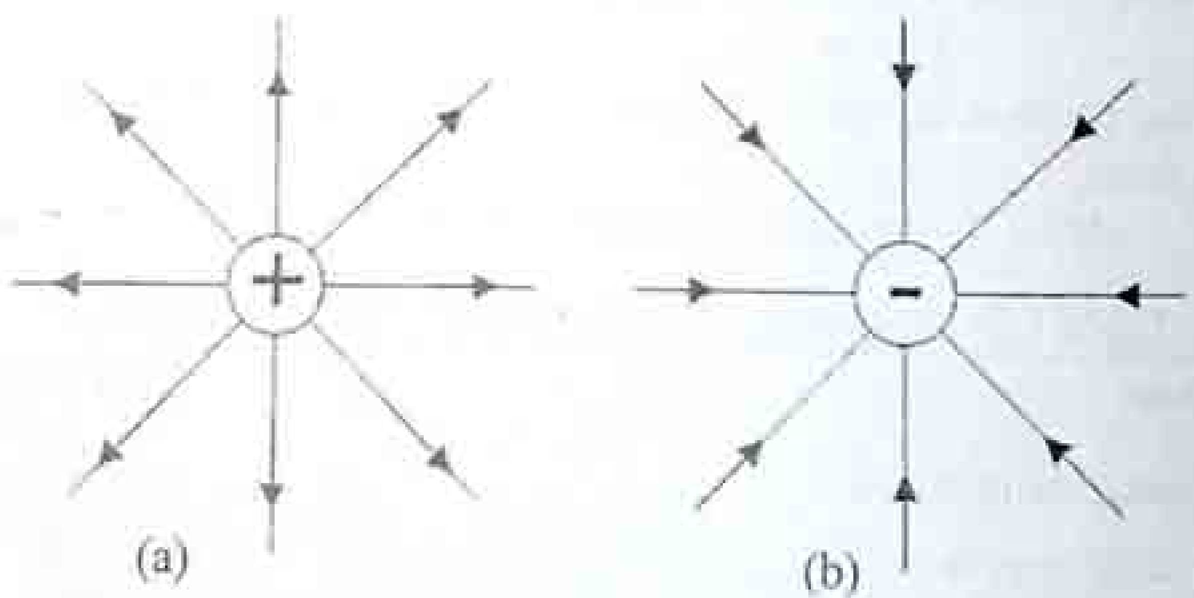


Fig. 1.11: Lines of force

Now, let us map out the lines of force between two charges of opposite signs. This is illustrated in Figure 1.12(a). The lines of force are directed from the positive charge to the negative charge. At point S_1 the least charge experience two forces; F_1 is a repulsive force due to the positive charge while F_2 is an attractive force due to the negative charge. The resultant of these forces is F and it is in the direction of the lines of force. Therefore, the field is directed tangentially to the lines of the force at S_1. For the case of two charges of the same sign (Figure 1.12b), the point N where the resultant intensity is zero is called neutral point. Since the force on the test charge has only one resultant direction, lines of force do not cross one another.

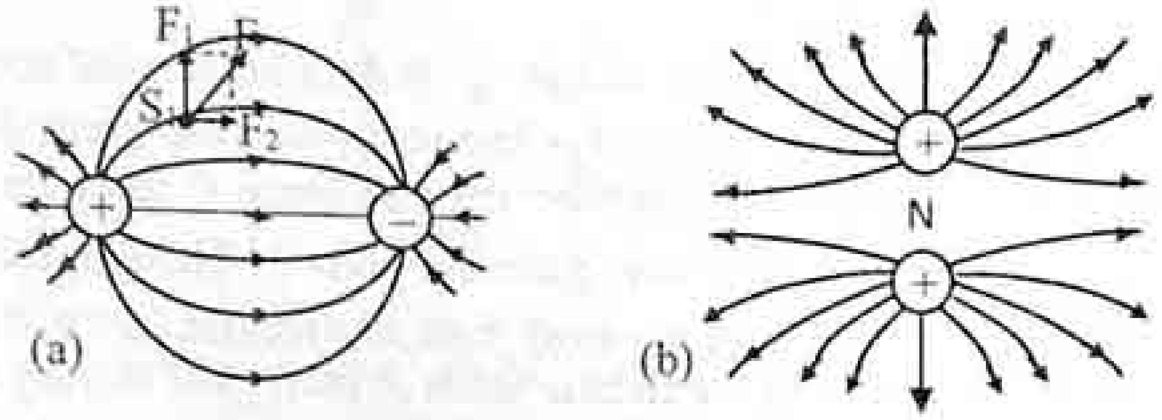


Fig. 1.12: Lines of force

It is also interesting to note that the lines of force are continuous, originating at one end on positive charge and terminating at the other end on a negative charge. For the radial lines in Figure (1.11a), the negative charges on which the lines of force terminate are considered to be located at infinity. We must note that the force in the electric field, is greater near the charges and the lines of force are much close together. In general, the closer the lines of force the stronger the electric field in that region. Actually, the lines can be drawn so that the total number crossing a unit area perpendicular to E is proportional to the magnitude of the electric field. Suppose we take two concentric spheres, the larger sphere has twice the radius of the other and also has four times the surface area of the smaller sphere. Since the number of lines passing through the two spheres is the same, then the number of lines per unit area is four times less for the larger sphere than for the smaller sphere. For a point charge, the electric field decreases as $1/r^2$. This means that electric field is four times less at the larger sphere than at the smaller sphere. Therefore, the magnitude of the electric field is proportional to the number of lines per unit area crossing a surface area normal to the direction of the lines of force.

It may be useful to summarize the properties of lines of force we have encountered in this section as follows:

1. The lines of force are drawn such that the magnitude of the electric field is proportional to the number of lines crossing a unit area perpendicular to the lines.
2. The lines of force are continuous and they start on positive and end only on the negative charges.
3. The tangent to the lines of force at every point gives the direction of the field.
4. Lines of force do not touch or intersect one another.

1.14 Calculation of Electric Field

We are now in a position to calculate the electric field, using Equation 1.11, for simple charge distributions. Since electric fields play a key role in physics, we shall devote this section to the determination of electric fields for various charge distribution.

Let us consider a point charge Q that is located at a distance r away from a test charge q_0 at point P. This is shown in Figure 1.13.

Fig.1.14: Charge Q separated from test charge q_0

The force exerted on the test charge q_0, according to Coulomb's law, is

$$F = \frac{1}{4\pi\varepsilon_0}\frac{Qq_0}{r^2}\hat{r}$$

where $\hat{r}$ is a unit vector pointing from Q to q_0

But $$E = \frac{F}{q_0}$$

$$E = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}$$

This is the electric field set up at point P by the charge Q. It does not depend upon the test charge. The point P at which the electric field is evaluated is known as the field point. So we see that the electric field E at a field point P located at a distance r from a charge Q is in the same direction as would be the force F on a test charge q_0 if it were placed at point P. The principle of superposition also applies to electric fields. That is, if there are more than one charge, say n charges, the electric field at the field point P is the vector sum of the electric fields of individual charges at the field point. Then the electric field is

$$E = E_1 + E_2 + E_3 + \ldots + E_n = \sum_{i=1}^{n} E_i$$

When E_i is the electric field at the field point P due to the i^{th} charge, Q_i.

Let us first of all perform some simple calculations without use of vectors.

Example 1.5
Find the electric field at a point $20cm$ from a charge of $20\mu C$. What force will the field exert on a charge of $5\mu C$, placed at that point?

Solution
The electric field due to the $20\mu C$ charge is

$$E = \frac{Q}{4\pi\varepsilon_0 r^2} = \frac{9.0\times10^9\,Nm^2C^{-2}\times20\times10^{-6}\,C}{(20\times10^{-2}\,m)^2} = 4.5\times10^6\,NC^{-1}$$

The force exerted on the $5\mu C$ is given by

$$F = \frac{Q_1Q_2}{4\pi\varepsilon_0 r^2} = \frac{9.0\times10^9\,Nm^2C^{-2}\times20\times10^{-6}\,C\times5\times10^{-6}\,C}{(20\times10^{-2}\,m)^2} = 22.5N$$

Example 1.6
When a $5nC$ test charge is placed at a point, it experiences a force of $2\times10^{-4}\,N$ in the x-direction. What is the electric field E at that point? If an electron is placed in this field, what force will be exerted on it?

Solution
Since the force on the positive test charge is in the x-direction, the electric field is also in the x-direction. From the definition, the electric field is

$$E = \frac{F}{q_0} = \frac{2\times10^{-4}\,N}{5\times10^{-9}} = 4\times10^4\,NC^{-1}.$$

The force on the electron will be given by $F = qE = 1.6\times10^{-19}\,C\times4\times10^4\,NC^{-1} = 6.4\times10^{-15}\,N$ in the negative x-direction.

Example 1.7
A positive charge $q_1 = 8nC$ is at the origin, and a second positive charge $q_2 = 12nC$ is on the x axis at x = $4m$. Find (i) the net electric field at a point P on the x axis at x = $7m$. (ii) the electric field at point Q on the y axis at y = $3m$ due to the charges.

Solution

(i) Net field at P is (Figure 1.14a). $E = \frac{1}{4\pi\varepsilon_0}\left[\frac{q_1}{7^2} + \frac{q_2}{3^2}\right]$ along positive x- axis

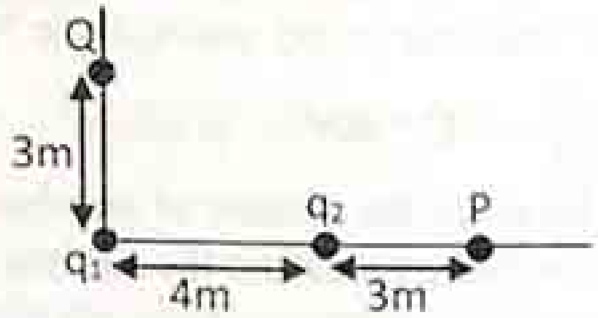


Fig 1.14a: Example 1.7

$$=9.0\times10^{9}C^{-2}\left[\frac{8\times10^{9}C}{(7m)^{2}}+\frac{12\times10^{9}C}{(3m)^{2}}\right]=1.47NC^{-1}+12.0NC^{-1}=13.5NC^{-1}$$ along positive x-axis.

(ii) First of all, from the geometry of the diagram (Fig. 1.14b), $\sin\theta=4/5=0.8$, $\cos\theta=3/5=0.6$.
The field due to q_1 at Q is in the positive direction of y axis and has magnitude.

$$E_1=\frac{1}{4\pi\varepsilon_0}\times\frac{q_1}{(3m)^2}=\frac{9.0\times10^{9}\times8\times10^{-9}}{3^2}=8NC^{-1}$$

Field E_2 at Q, due to q_2 has magnitude $E_2=\frac{1}{4\pi\varepsilon_0}\times\frac{q_2}{(5m)^2}=\frac{9.0\times10^{9}\times12\times10^{-9}}{5^2}=4.32NC^{-1}$

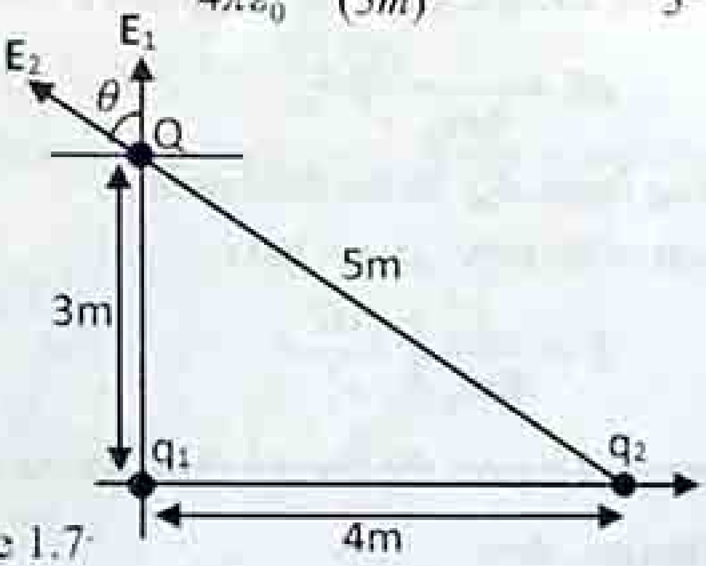


Fig. 1.14b: Example 1.7

Resultant field at Q in the x-direction $=E_2\sin\theta=-4.32\times0.8=-3.46NC^{-1}$.

Resultant field at Q in the $y-$direction $=E_1+E_2\cos\theta=7.99+-4.32\times0.6=10.6NC^{-1}$.

Total resultant field at Q is $E_R=\sqrt{E_x^2+E_y^2}=\sqrt{(-3.46)^2+(10.6)^2}=11.2NC^{-1}$.

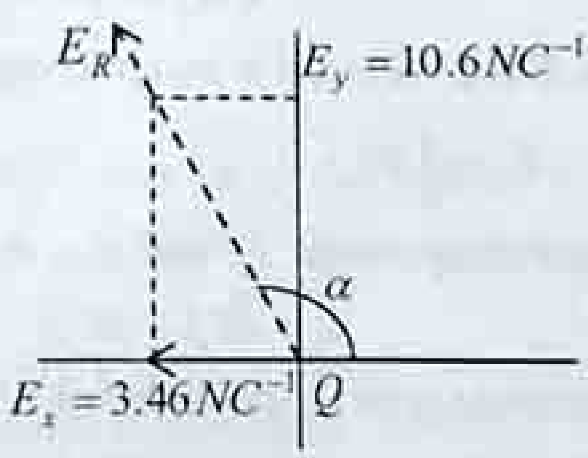


Fig. 1.14c: Example 1.7

This resultant E_R makes an angle α with x-axis shown in the diagram (Figure 1.14c) given by

$$\tan\alpha=\frac{E_y}{E_x}=\frac{10.6}{3.46}=3.06 \qquad \alpha=\tan^{-1}(3.06)=71.9^\circ$$

1.15 Continuous Charge Distribution

We have shown in the previous examples how electric field due to various discrete charge distributions can be calculated. Now we wish to calculate the electric field due to continuous charge distributions. The procedure is to divide the continuous charge distribution into infinitesimal charge dQ contained in a volume element ΔV that is large enough to contain many (billions of) individual charges or molecules and yet small enough that replacing ΔV by a differential dV and using calculus introduces negligible error. We describe the charge per unit volume by the volume density.

$$\rho = \frac{\Delta Q}{\Delta V} \therefore \Delta Q = \rho \Delta V$$

Often, charge is distributed in a thin layer on the surface of an object. In such a case, we define the other the surface of an object. In such a case, we define the surface charge density, σ as the charge per unit area.

$$\sigma = \frac{\Delta Q}{\Delta A} \qquad \therefore \Delta Q = \sigma \Delta A$$

Similarly, we sometimes encounter a charge distribution along a line in space. In such a case, we define the linear charge density, λ as the charge per unit length

$$\lambda = \frac{\Delta Q}{\Delta L} \therefore \qquad \Delta Q = \lambda \Delta L$$

The electric field produced by a given charge distribution can be calculated in a straight forward way using Coulomb's law.

If it is a volume charge distribution, we consider element $dQ = \rho dV$ that is small enough to be considered as a point charge. The electric field dE at a point P due to this charge is given by

$$dE = \frac{1}{4\pi\varepsilon_0}\frac{dQ}{r^2}\hat{r}$$

where $\hat{r}$ is unit vector pointing from dQ to the field point.

For volume charge distribution of density, ρ we have

$$E = \frac{1}{4\pi\varepsilon_0}\int_V \frac{\rho dV}{r^2}\hat{r}$$

$\int_V$ means integrating the expression over the charge distribution occupying volume, V.

If it is a surface charge distribution, then

$$E = \frac{1}{4\pi\varepsilon_0}\int_A \frac{\sigma dA}{r^2}\hat{r}$$

If it is a line charge distribution, then

$$E = \frac{1}{4\pi\varepsilon_0}\int_L \frac{\lambda dL}{r^2}\hat{r}$$

Remember that since E is a vector, it is easier to express the above equation $dE_x i + dE_y j$.

Then, $\qquad E_x = \int dE_x \qquad\qquad E_y = \int dE_y$

Having evaluated these components E_x and E_y, the electricity field at the point is

$$E = dE_x i + dE_y j$$

We illustrate this procedure with an example below.

1.16 Motion of a Charged Particle in an Electric Field

We now know that when a charged body is placed inside an electric field set up by other charges, it experiences a force due to this field. Recall from equation 1.11 that

$$F = qE$$

We spent the last section in determining E for some particular situations. Now suppose we know the electric field E, we wish to find out what force and torques that will act on a charged body placed in it. If the charged body has a mass m, we can also determine its acceleration and its subsequent motion in the field. According to Newton's second law, this force produces an acceleration a given by

$$a = \frac{F}{m} = \frac{qE}{m} \qquad (1.13)$$

Few examples will clarify these points.

Example 1.8

A proton of mass m and charge q is placed at rest inside a uniform electric field E and released. Describe its motion.

Solution

Figure 1.15 shows the proton inside the uniform electric field set up between two parallel plates which are S metres apart.

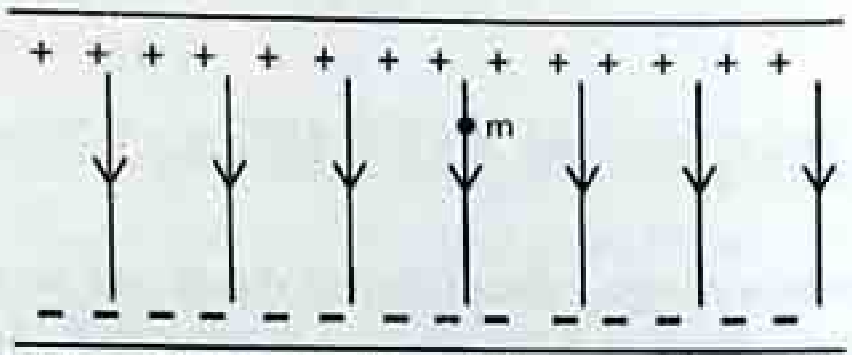


Fig. 1.15: A proton of mass m, released from rest in a uniform electric field.

Since the proton is released near the positive plate, it will experience a force, F in the direction of the field. We therefore have,

$$F = qE \text{ and } a = \frac{qE}{m}$$

The motion simply reduces to a problem of a body undergoing a uniformly accelerated motion. Then we have, $v_0 = 0$,

$$a = \frac{qE}{m}$$

Therefore, $v_f = at = \frac{qE}{m}t$; $\qquad S = \frac{1}{2}at^2 = \frac{qE}{2m}t^2$

And $\qquad v_f^2 = 2aS = 2\frac{qE}{m}S$

The kinetic energy after it has reached the lower plate is

$$K.E = \frac{1}{2}mv_f^2 = qES$$

We could have obtained this last expression from work- energy theorem since the work done by the force (qE) on the body is equal to change in its kinetic energy.

Example 1.9

A uniform electric field E, of magnitude $2000NC^{-1}$ acts along the negative direction of y-axis. An electron is projected into this field with a velocity $v_0 = 10^6 ms^{-1}$ perpendicular to the field i.e. along x-axis.

(a) By how much has the electron been deflected after it has travelled $1cm$ in the x direction?

(b) How far does the electron travel before it is brought momentarily to rest?

Solution

(a) The time it takes the electron to travel a distance 1cm in the x-axis is

$$t = \frac{x}{v_0} = \frac{10^{-2}}{10^6} = 10^{-8}s$$

In this time, the electron is deflected upward anti parallel to the field, a distance y given by

$$y = \frac{1}{2}at^2 = \frac{1}{2}\frac{eE}{m}t^2$$

Substitute $e = 1.6\times10^{-19}C, m_e = 9.11\times10^{-31}kg, E = 2000\,NC^{-1}$ and $t = 10^{-8}s$, we get

$y = 1.76 \times 10^{-2} m = 1.76 cm.$

(b) Since the charge of the electron is negative the force- eE acting on it, it is in the direction opposite the field. We then thus have a constant acceleration problem in which the acceleration of a particle is opposite to its initial velocity. To find the distance transverse before coming to rest, we can use $V^2 = V_0^2 + 2aX$

where $V_0 = 2 \times 10^6 ms^{-1}$,

$$X = -\frac{V_0^2}{2a} = \frac{mV_0^2}{2eE}$$

Making substitution from m, V_0, e and E, we obtain $X = 0.57 \times 10^{-2} m = 0.57 cm.$

1.17 Electric Dipole

In Example 1.7, we found the force on a charge q due to electric dipole and we found the electric field due to an electric dipole. *The electric dipole is a combination of two equal charges of opposite sign separated by a distance of* $2a$. The dipole moment P is defined as $2aQ$. Now suppose we place the electric dipole in a uniform electric field as shown in Figure 1.16, the dipole moment can be considered to be a vector pointing from $-Q$ to $+Q$ as shown in the figure and makes an angle θ with the electric field. Since the field is uniform, the force on the positive charge is QE and $-QE$ is on the negative charge. The net force on the dipole is therefore zero. However, there is a net torque, τ on the dipole. The magnitude of the torque, τ can be calculated about an axis through the centre O. It is given as,

$$\tau = aQE\sin\theta + aQE\sin\theta = 2aQE\sin\theta = PE\sin\theta \quad (1.14a)$$

This may be written as $\tau = E \times P$ (1.14b)

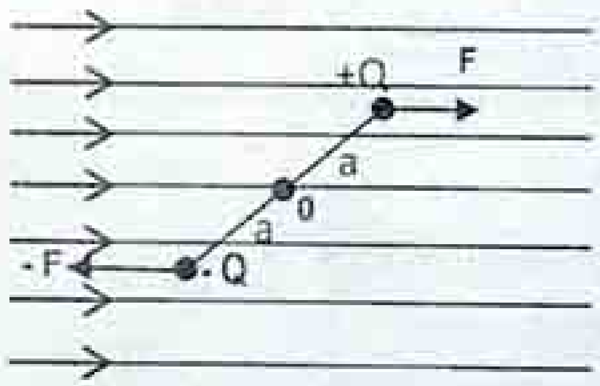


Fig 1.17: An electric dipole in a uniform external field E

This torque tends to align the electric dipole with the electric field. The electric field may cause the dipole to change the angle of alignment with the field say from θ_1 to θ. Then the work, W, done on the dipole by the electric field is given by

$$W = \int_{\theta_1}^{\theta} \tau d\theta \quad (1.15a)$$

$$= \int_{\theta_1}^{\theta} PE\sin\theta d\theta = PE\int_{\theta_1}^{\theta} \sin\theta d\theta$$

$$= PE(\cos\theta_1 - \cos\theta) \quad (1.15b)$$

The work done by the field decreases the potential energy U of the dipole in the field. Since we are interested only in changes in potential energy, let us take the potential energy U to be zero when P is perpendicular to the electric field i.e. $\theta_1 = 90°$ and $\theta = 180°$. Equation 1.15(b) becomes

$$W = -PE\cos\theta = P \cdot E \quad (1.16)$$

Example 1.10

The HCl molecule has a dipole moment of about $3.4\times10^{-30}\,C.m$. The two atoms are separated by about $1.0\times10^{-10}\,m$. What maximum torque would this dipole experience in a $2.5\times10^{4}\,NC^{-1}$ electric field and how much energy would be required to rotate the molecule 45° from its equilibrium position of minimum potential energy?

Solution

$$P=3.4\times10^{-30}\,C.m,\ E=2.5\times10^{4}\,NC^{-1}$$

From equation 1.14(a) we have

$$\tau = PE\sin\theta$$

We obtain the maximum torque when $\sin\theta=1$, substituting we have

$$\tau=(3.4\times10^{-30}\,C.m)(2.5\times10^{4}\,NC^{-1})=8.5\times10^{-26}\,Nm$$

From equation 1.15(b) we have $W = PE(\cos\theta_1-\cos\theta)$

The dipole has maximum potential energy when it is parallel to the electric field i.e. $\theta_1=0$. Then,

$$W = PE(1-\cos 45)=(3.4\times10^{-30}\,C.m)(2.5\times10^{4})(1-0.707)=2.5\times10^{-26}\,J.$$

Summary

1. There are two types of charges, negative and positive. Like charges repel each other while unlike charges attract.
2. Objects can be electrified or charged, either positively or negatively, by the removal or addition of electrons. The goldleaf electroscope is used to detect the presence of an electric charge.
3. Metals and other materials that allow free flow of electrons through them are called conductors. Materials that do not allow the flow of electrons are called insulators, then between the two types are materials called semiconductor.
4. Coulomb's law of force between two point charges is given by

$$F=\frac{1}{4\pi\varepsilon_0}\frac{Q_1Q_2}{r^2}$$

In S.I units, the permittivity of free space ε_0 is equal to $9\times10^{9}\,Nm^2C^{-2}$.

5. Electric field is a region where a stationary particle experiences a force. The electric field intensity, $E=\frac{F}{q_0}$.
6. Electric lines of force are continuous; they start from positive charges and end on negative charges. They do not touch or intersect one another.
7. A system consisting of two equal and opposite charges Q, separated by a small distance $2a$, is an electric dipole. The dipole moment is $P = 2aQ$. At a large distance from the dipole and along the perpendicular bisector, the electric field is

$$E=\frac{P}{4\pi\varepsilon_0 r^3}$$

8. Problems involving continuous charge distribution can be solved by integration as shown below:

$$\int dE=\frac{1}{4\pi\varepsilon_0}\int_V\frac{\rho dV}{r^2}\hat{r}$$

$\int_V$ means integrating the expression over the charge distribution occupying some volume, V. If it is a surface charge distribution, $E=\frac{1}{4\pi\varepsilon_0}\int_A\frac{\sigma dA}{r^2}\hat{r}$

If it is a line charge distribution, $E = \frac{1}{4\pi\varepsilon_0}\int_L \frac{\sigma dL}{r^2}\hat{r}$

Exercises 1

Which of the following statements is not correct in Questions 1.1 to 1.3

1.1 **A.** There are two types of electric charge positive and negative charge.
B. Materials such as polythene and cellulose acetate can be charged by friction.
C. The positive charge on the woolen material is not of the same magnitude as the negative charge on the polythene.
D. Electric charge is always conserved in a closed system.

1.2 **A.** Water is an insulator. **B.** Carbon is an insulator. **C.** Plastic is a semi-conductor.
D. Copper is a good conductor.

1.3 Dielectric constant of
A. Vacuum is 1 **B.** Air is $\frac{1}{2}$ **C.** Metal is infinity **D.** Water is 0

1.4 Which of the following statement is correct, water is?
A. Good conductor **B.** Poor conductor **C.** Semi–conductor **D.** Insulator

1.5 Silver is a:
A. Good conductor **B.** Poor conductor **C.** Semi–conductor **D.** Insulator

1.6 Carbon is a:
A. Good conductor **B.** Poor conductor **C.** Semi–conductor **D.** Insulator.

1.7 Charging by induction is
A. touching two positive charges **B.** touching two negative charges
C. Charging a body without touching it **D.** touching two irons

1.8 Which of the following is not correct
A. The charge on an electron is $e = 1.60\times10^{-19}C$ **B.** Mass of proton $= 1.67\times10^{-10}kg$
C. Mass of electron $= 9.1\times10^{-31}kg$ **D.** Dielectric constant of air is 1.005

1.9 Which of the following is incorrect:
A. Electric field intensity F/q **B.** The unit of electric field intensity is NC^{-1}
C. The unit of electric field intensity is Vm^{-1} **D.** Electric field intensity is a scalar quantity.

1.10 Which of the following is correct:
A. Equipotential surface is a surface where all points have equal potential. **B.** The surfaces have the same charge. **C.** There are no lines of force **D.** Point of equal electric intensity

1.11 Show that the ratio of electrical to gravitational forces is independent of distance between the particles.

1.12 What do you mean by the following: (i) Charging by induction. (ii) Earthing a conductor?

1.13 What do you understand by saying that electric charge is quantized?

1.14 Define the dielectric constant of a medium ε_r.

1.15 A plastic rod is rubbed against a wool shirt, thereby acquiring a charge of $-0.8\mu C$. How many electrons are transferred from the wool shirt to the plastic?

1.16 Compute the ratio of electric to gravitational force between two protons

1.17 Point charges $88\mu C, -55\mu C$ and $70\mu C$ are placed in a straight line. The central one is 0.75m from each of the others. Calculate the net force on each due to the other two.

1.18 Two charges $-Q_0$ $-Q_0$ and $-3Q_0$ are at l distance apart. These two charges are free to move but do not because there is a third charge nearby. What must be the charge and placement of the third charge for the first two to be in equilibrium?

1.19 Two equal charges $q_1 = q_2 = -6\mu C$ are on the y–axis at $y_1 = 3cm$ and $y_2 = -3cm$.

(i) What is the magnitude and direction of the electric field on the x-axis at $x = 4cm$. (ii) If a test charge $q_0 = 2\mu C$ is placed at $x = 4cm$, find the force the test charge experiences.

1.20 A quadruple consists of two dipoles that are close to each other as shown in Figure 1.18. The effective charge at the origin is $-2q$ and the other charges on the y-axis at $y = a$ and $y = -a$ are $+q$ each. (i) Find the electric field at a point on the x-axis far away so that $x >> a$. (ii) Find the electric field on the y-axis far so that $y >> a$.

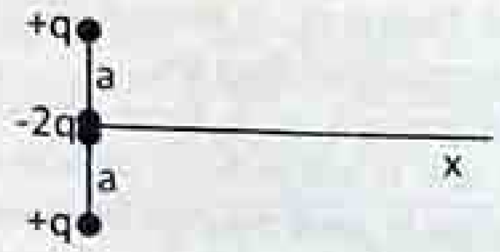


Fig. 1.18: Exercise 1.20

CHAPTER 2
GAUSS'S LAW AND ELECTRIC POTENTIAL

2.0 Introduction to Gauss's Law

The determination of electric field E, due to some charge distributions has occupied the attention of many scholars and Gauss's Law has been handy in this case. In principle, we can always determine the electric field due to a given charge distribution using Coulomb's law. This method always works and as we have seen it is tedious and complicated. However, in these days of fast computing, we can always find solution to problems no matter how complicated they are.

The Coulomb's law as we know it may be expressed in a form known as Gauss's law. The determination of the electric field using Gauss's law is not tedious or laborious. But the number of problems that can be solved using Gauss formulation is quite limited because the charge distributions must have special symmetry properties. Even the few that can be solved with Gauss's law are solved easily with elegance and beauty. However, the importance of Gauss's law does not lie on the few practical problems we can solve with it but in additional insight into the nature of electrostatic fields it gives us. We shall discuss the concept of flux of a vector field before discussing Gauss's law itself.

2.1 Electric Flux

Let us consider a closed surface S shown in Figure 2.1.

Suppose we take an elemental area ΔS and define a unit vector n which is always perpendicular to the area but pointing outwards. The elemental area ΔS and its direction is in the direction of the unit vector n.

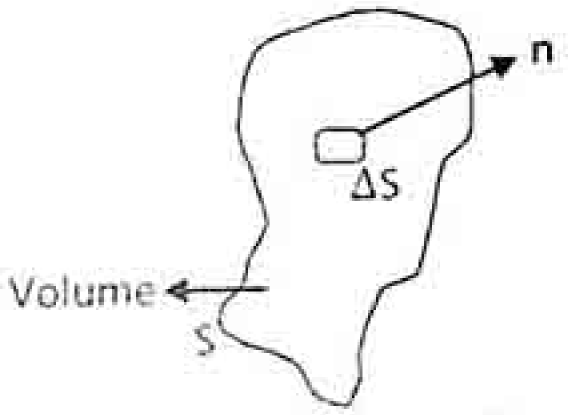


Fig. 2.1: A closed surface S on which asmall area ΔS is represented.

Hence,

$$\Delta S = \Delta S n$$

Now we are in a position to define the electric flux, Φ. Before we do so, let us consider a surface of area, A, through which a uniform electric field passes. This is shown in Figure 2.2. In Figure 2.2(a) the lines of force of the electric field are perpendicular to the surface of the area, the unit vector is parallel to the electric field. In this case we define the electric flux as

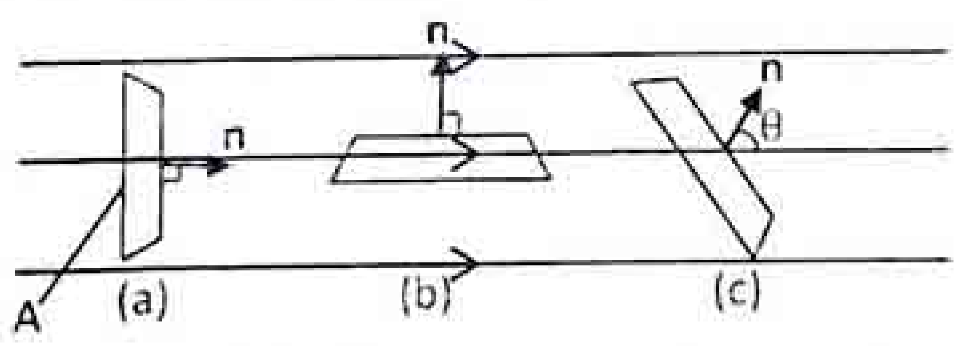


Fig. 2.2: A uniform electric field E passing through an area A at various inclinations with the field.

$$\Phi_{a\perp} = E_{\perp} \cdot A \qquad (2.1a)$$

We add the perpendicular sign ($\perp$) to the electric field E in Equation 2.1(a) to emphasize that the lines of force must be perpendicular to the area for the definition to be true. In the case of Figure 2.2(b), the lines of force are parallel to the area A, the unit vector is perpendicular to the field, hence the flux is zero.

$$\Phi_{(b)} = 0 \qquad (2.1b)$$

In Figure 2.2(c), the unit vector, n to the area makes an angle θ with the field, fewer lines of force will pass through the area. Then

$$\Phi_{(c)} = E_{\perp}A = EA\cos\theta \qquad (2.1c)$$

where $E\cos\theta$ is the component of the field E that is perpendicular to the surface A. In general, the electric flux, Φ is defined as

$$\Phi = E\cdot An \qquad (2.2)$$

We saw in the last chapter that the number of lines of force (N) through a unit area erected perpendicular to the lines is proportional to the magnitude of the electric field E. That is,

$$N/A \propto EA = \Phi$$

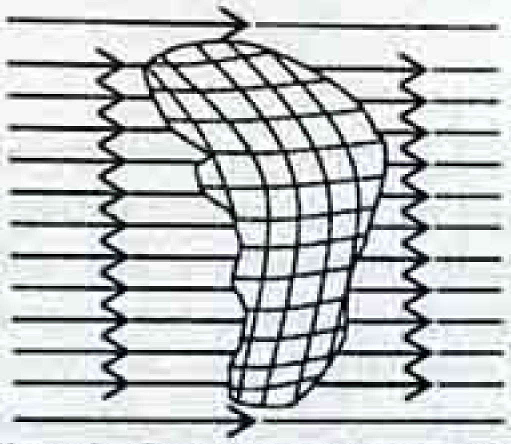


Fig. 2.3: Electric flux through an arbitrary surface

Therefore, the flux through an area is proportional to the number of lines of force passing through the area. Suppose our electric field is not uniform and the surface is not flat as shown in Figure 2.3, we divide the whole surface into small elemental areas ΔS_i in such a way that the electric field E_i at ΔS_i is uniform. Then, the electric flux through the entire surface is the sum of the fluxes through each of the elemental areas. In that case, we have

$$\Phi = \sum_{i=1}^{N} E_i\Delta S_i n$$

where E_i is the electric field passing through ΔS_i as ΔS_i gets smaller (i.e. $\Delta S \to 0$), the summation above is replaced by a integral sign and we have

$$\Phi = \int_i^n (E\cdot n)dS \qquad (2.3)$$

In most cases, we will be dealing with the flux through closed surfaces like the surface shown in Figure 2.3. A closed surface completely encloses a volume. In such a case, equation 2.3 becomes

$$\Phi = \oint (E\cdot n)dS \qquad (2.4)$$

where the integral sign written as $\oint$ indicates that the integral is over a closed surface.

Example 2.1

At each point on the flat rectangular surface shown in Figure 2.4, the electric field has a magnitude of $350\,NC^{-1}$ and makes an angle of $50°$ with n. Calculate the electric flux for this surface.

Solution

The area of the rectangle is given as $3\times10^{-2}\,m^2$. Since the field E is constant throughout the rectangle, using equation 2.2 we obtain

$$\Phi = EA\cos 50° = (350\,NC^{-1})(3\times10^{-2})(0.64) = 6.72\,Nm^2C^{-1}$$

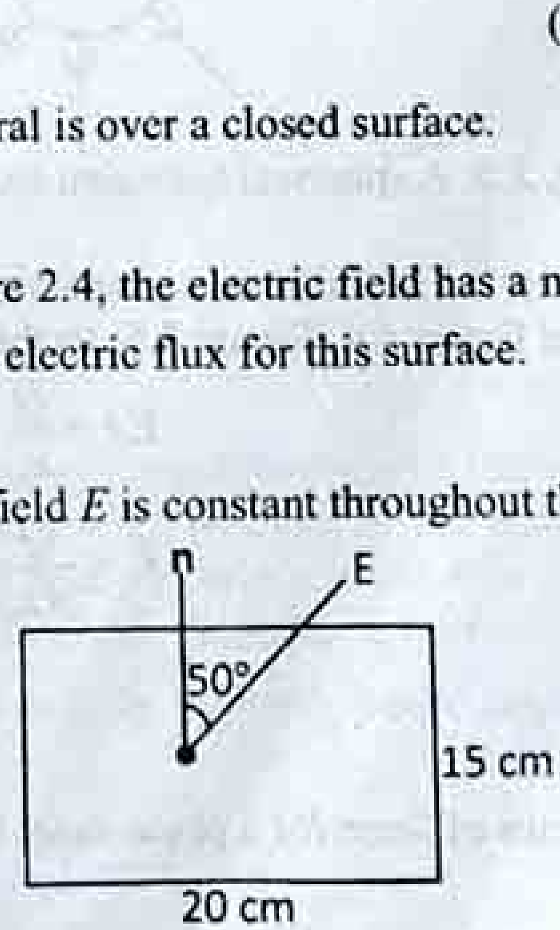


Fig. 2.4: Example 2.1

2.2 Gauss' Law

For continuous distribution of charges, the summation of electric field tends to integration as shown below.

$$E=\sum_{i=1}^{n}\frac{Q}{4\pi\varepsilon_0 r_i^2} \quad (2.5)$$

If the charge is continous over a surface S, then the electric flux Φ_E over the closed surface is defined as

$$\Phi_E=\oint EdS$$

(2.6)

Where dS is element of area. It thus follow from equation (2.5) and (2.6) that

$$\Phi_E=\frac{Q}{\varepsilon_0} \quad (2.7)$$

Equation 2.7 is known as Gauss' Law and it holds for any surface of arbitrary shape (called the Gaussian surface). Gauss' law offers an elegant method for calculating the electric field vector of a given charge distribution, provided the charge distribution has certain amount of symmetry. Conversely, Gauss' law can be used to calculate the charge inside a closed surface if the electric field vector at all points in the surface is known. In summary, Gauss' law states that *the surface integral of the normal component of an electric field through any closed surface containing a total charge Q is Q/ε_0.*

2.3 Applications of Gauss' Law

As mentioned above, Gauss' law can be applied in calculating the electric field due to asymmtric charge distributions such as a charged sphere or conducting sheet. Some of these applications are discussed below.

(a) **Electric Field due to a charged spherical conductor:** Consider a point charge Q at the centre of an imaginary sphere of radius r (Figure 2.5). The area of this sphere is $4\pi r^2$ and since by symmetry, the electric field is the same at every point on the sphere, the total flux is

$$\Phi_E=4\pi r^2E=4\mu r^2\frac{Q}{4\pi\varepsilon_0 r^2}=\frac{Q}{\varepsilon_0}$$

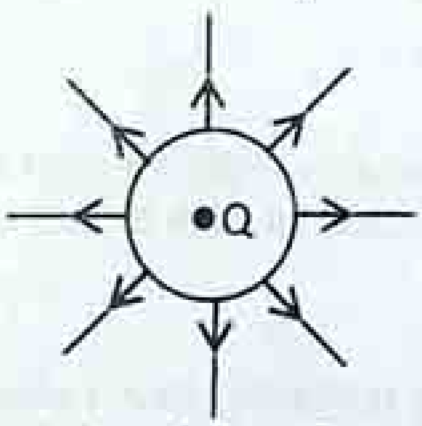


Fig. 2.5: A spherical Gaussian surface surrounds a positive point charge concentrically

(b) **Electric Field due to a plane sheet of charge:** For a charged conducting sheet of area A, the total flux is EA (Figure 2.6) so that from Gauss' law,

$$EA=\frac{Q}{\varepsilon_0}$$

Thus, $$E=\frac{Q}{A\varepsilon_0}$$

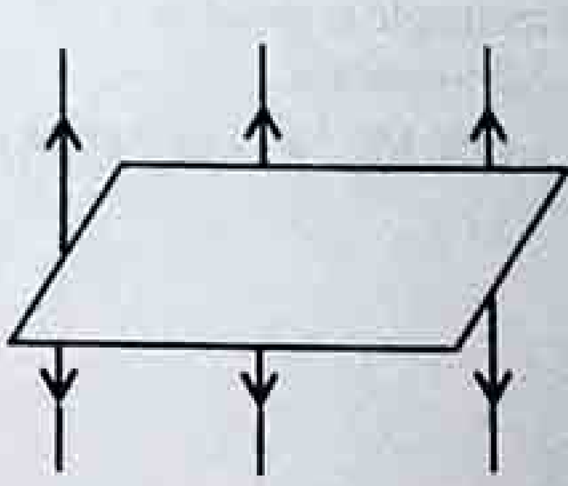


Fig. 2.6: Lines of force for a plane sheet of charge.

If the sheet is infinite, we can find the electric field near the conductor from the surface charge density $\sigma = Q/A$ which is constant for uniform charge distribution as

$$E = \frac{\sigma}{\varepsilon_0}$$

It therefore follows that once the charge density (charge per unit area) is known, the electric field can be computed.

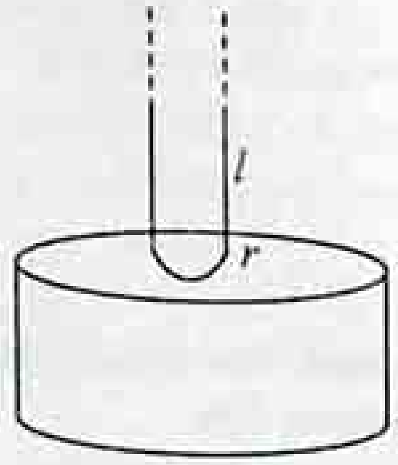


Fig. 2.7: A charged cylinder

(c) **Electric Field due to a charged cylinder:** For a cylinder of radius r and lenght l, the area is $2\pi rl$. The flux is therefore $2\pi rlE$. Thus, by Gauss' law

$$2\pi rlE = \frac{Q}{\varepsilon_0}$$

$$E = \frac{Q}{2\pi rl\varepsilon_0}$$

If the cylinder is of infinite length, we find the charge per unit length, $\lambda = Q/l$ such that from Gauss'law, we have

$$E = \frac{\lambda}{2\pi r\varepsilon_0}.$$

Example 2.2

A solid sphere of radius R is uniformly charged with a constant volume density, ρ. Find the electric field inside ($r < R$) and outside ($r > R$) the solid sphere.

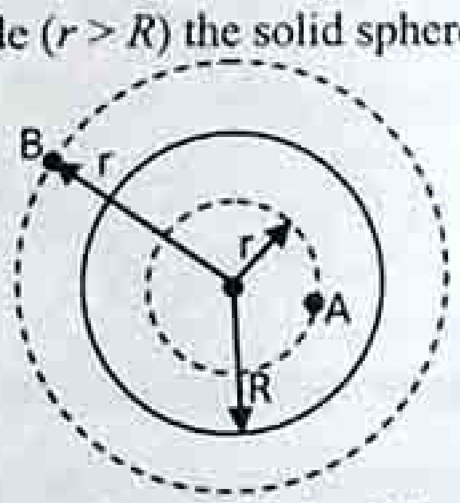


Fig. 2.8: A charged solid sphere

Solution

Figure 2.8 shows a solid sphere of radius R which is uniformly charged throughout the volume. Now we wish to find the electric field at point A (r < R) and B (r > R). Because of the spherical charge symmetry, we must choose spherical Gaussian surfaces. To find the field at point A (r < R) we must take a Gaussian surface passing through point A. The radius of the Gaussian sphere is r. The charge enclosed by our sphere is $\rho(4/3\pi r^3)$. The electric field is radially outward, hence, applying Gauss' law we have,

$$\oint E \cdot dS = \frac{\rho(4/3\pi r^3)}{\varepsilon_0}$$

$$E(4\pi r^2) = \frac{\rho}{\varepsilon_0}(4/3\pi r^3),\ E = \frac{\rho}{\varepsilon_0}\left(\frac{1}{3}r\right)$$

But $$\rho = \frac{Q}{V};\ E = \frac{\frac{Q}{(4/3R^3)}}{\varepsilon_0}\left(\frac{1}{3}r\right) = \frac{1}{4\pi\varepsilon_0}\frac{Qr}{R^3}$$

From the above equation, E is equal to zero if $r = 0$, and E increases linearly with r until $r = R$. To find the field at point B, we take a sphere passing through B (r > R). The charge enclosed by the Gaussian surface is just Q.

Applying Gauss' law we have

$$E(4\pi r^2) = \frac{Q}{\varepsilon_0};\qquad \therefore E = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}$$

The last equation above shows that for points outside the sphere, the charged sphere behaves as if it were a point charge Q located at the centre (see Figure 2.9).

The magnitude of the field as a function of r shows the plot of the magnitude of the electric field as a function of the distance r from the centre. The field increases linearly with r, until $r = R$ then it decreases as $1/r^2$.

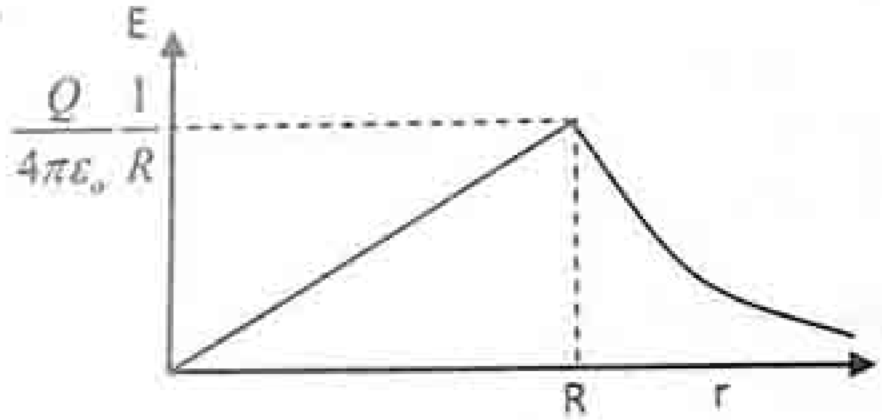


Fig. 2.9: The magnitude of the field as a function of r

2.4 Electric Field and Charge in Conductors

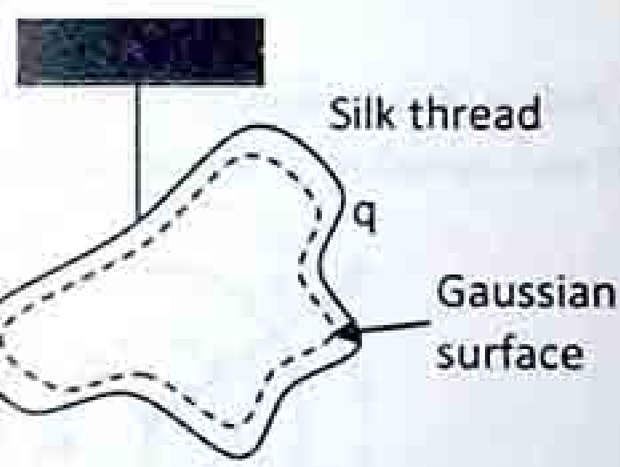


Fig. 2.10: An insulated conductor carrying an excess charge q.

A Gaussian surface is a closed surface in three dimensional space through which the flux of a vector field is calculated. Recall that a conductor is a material containing electrons and protons; the electrons are free to move about in the material. If a resultant electric field exists within a conductor, the electrons will experience a net force and will accelerate and a current will flow in the conductor. But under electrostatic conditions, there is no current flow in the conductor. When there is no current flow, the net force on the electrons in the conductor must be zero, and therefore the resultant electric field inside the conductor must be zero. We therefore conclude that under electrostatic conductions, the resultant electric field, E everywhere within the conductor is zero, i.e. $E = 0$.

Suppose we place an excess charge q on an insulated conductor. We wish to find where this excess charge resides. Figure 2.10 shows an insulated conductor carrying an excess charge q. Let us take a Gaussian surface immediately below the surface of the conductor. The Gaussian surface is shown in

Figure 2.10 by the dashed lines. Because the electric field inside the conductor is zero, the net flux for this Gaussian surface is zero. Applying Gauss's law to this surface we have

$$\oint E \cdot dS = \frac{Q_{net}}{\varepsilon_0}$$

But, since $E = 0$ inside the conductor $Q_{net} = 0$.

In other words, the sum of all the charges within the Gaussian surface is zero. Therefore, if there is an excess charge q on the conductor, it must reside outside the Gaussian surface. Since the Gaussian surface is just beneath the surface of the conductor, the excess charge must reside on the surface of the conductor.

Since the charge of a charged conductor resides on the surface, there must be an electric field on the surface. The electric field cannot have a component parallel to the surface; otherwise the electrons on the surface will experience a force and accelerate. This will definitely violate the electrostatic conditions. Therefore, the electric field at the surface of a charged conductor must be perpendicular to the surface at every point on the surface.

Example 2.3 Hydrogen Atom

In the classical model of the hydrogen atom, the electron revolves around the proton with a radius of $r = 0.5 \times 10^{-10} m$. The magnitude of charge of the electron and proton is $e = 1.6 \times 10^{-19} C$.

(a) What is the magnitude of the electric force between the proton and the electron?

(b) What is the magnitude of the electric field due to the proton at r?

(c) What is ratio of the magnitudes of the electrical and gravitational force between electron and proton? Does the result depend on the distance between the proton and the electron?

(d) In light of your calculation in (b), explain why electrical forces do not influence the motion of planets.

Solution

(a) The magnitude of the force is given by $F_e = \dfrac{e^2}{4\pi\varepsilon_0 r^2}$

We can substitute our numerical values and find that the magnitude of the force between the proton and the electron in the hydrogen atom is:

$$F_e = \frac{(9.0 \times 10^9\, Nm^2C^{-2})(1.6 \times 10^{-19})^2}{(5 \times 10^{-11})^2} = 9.2 \times 10^{-8}\, N$$

(b) The magnitude of the electric field due to the proton is given by

$$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{r^2} = \frac{(9.0 \times 10^9\, Nm^2C^{-2})(1.6 \times 10^{-19})}{(5 \times 10^{-11})^2} = 5.76 \times 10^{-11}\, NC^{-1}$$

(c) The mass of the electron is $m_e = 9.1 \times 10^{-31}\, kg$ and the mass of the proton is $m_e = 1.7 \times 10^{-27}\, kg$. Thus, the ratio of the magnitude of the electric and gravitational force is given by

$$y = \frac{\left(\dfrac{e^2}{4\pi\varepsilon_0 r^2}\right)}{\left(\dfrac{Gm_p m_e}{r^2}\right)} = \frac{e^2/4\pi\varepsilon_0}{Gm_p m_e} = \frac{(9.0 \times 10^9\, Nm^2C^{-2})(1.6 \times 10^{-19}\, C)^2}{(6.67 \times 10^{-11}\, Nm^2kg^{-2})(1.7 \times 10^{-27}\, kg)(9.1 \times 10^{-31}\, kg)} = 2.2 \times 10^{39}$$

which is independent of r, the distance between the proton and the electron.

(d) The electric force is 39 orders of magnitude stronger than the gravitational force between the electron and the proton. Then why are the large scale motions of planets determined by the gravitational force and not the electrical force. The answer is that the magnitudes of the charge of the electron and proton are equal. The best experiments show that the difference between these magnitudes is a number on the order of 10^{-24}. Since objects like planets have about the same number of protons as electrons, they are essentially electrically neutral. Therefore the force between planets is entirely determined by gravity.

Example 2.4: Millikan Oil –Drop Experiment

An oil drop of radius $r = 1.64 \times 10^{-6} m$ and mass density $\rho_{oil} = 8.51 \times 10^{2}\, kg/m^{3}$ is allowed to fall from rest and then enters into a region of constant external field E applied in the downward direction. The oil drop has an unknown electric charge q (due to irradiation by bursts of X-rays). The magnitude of the electric field is adjusted until the gravitational force $F_g = mg = mgj$ on the oil drop is exactly balanced by the electric force, $F_e = qE$. Suppose this balancing occurs when the electric field is $E = -E_y j = -(1.92 \times 10^{5}\, NC^{-1})j$, with $E_y = 1.92 \times 10^{5}\, NC^{-1}$.

(a) What is the mass of the oil drop?

(b) What is the charge on the oil drop in units of electric charge $e = 1.6 \times 10^{-19} C$?

Solution

(a) The mass density $\rho_{oil} V = \rho_{oil}(4\pi r^{3}/3)$

where the oil drop is assumed to be a sphere of radius r with volume $V = 4\pi r^{3}/3$.

We can substitute our numerical values into the symbolic expression for the mass,

$$M = \rho_{oil}(4\pi r^{3}/3) = (8.51 \times 10^{2}\, kg/m^{3})(4\pi/3)(1.64 \times 10^{-6} m)^{3} = 1.57 \times 10^{-14} kg$$

(b) The oil drop will be in static equilibrium when the gravitational force exactly balances the electrical force: $F_g + F_e = 0$. Since the gravitational force points downward, the electric force on the oil must be upward. Using our force laws, we have $0 = mg + qE$, $mg = -qE_y$.

With the electrical field pointing downward, we conclude that the charge on the oil drop must be negative. Notice that we have chosen the unit vector j to point upward. We can solve this equation for the charge on the oil drop: $q = -\dfrac{mg}{E_y} = \dfrac{(1.5 \times 10^{-14} kg)(9.8 ms^{-2})}{1.92 \times 10^{5}\, NC^{-1}} = -8.03 \times 10^{-19} C.$

Since the electron has charge $e = 1.6 \times 10^{-19} C$, the charge of the oil drop in units of e is

$$N = \frac{q}{e} = \frac{8.02 \times 10^{-19} C}{1.6 \times 10^{-19} C} = 5$$

You may at first be surprised that this number is an integer, but the Millikan oil drop experiment was the first direct experimental evidence that charge is quantized. Thus, from the given data we can assert that there are five electrons on the oil drop.

Summary

1. Having gone through this chapter, one should be able to define and explain Gauss's Law and its concepts. Starting from a simple point charge define, derive mathematically and explain Gauss's Law. Also attempt to extend the concepts to continuous charge distribution and applying it to solving problems.
2. Electric Flux is given by $\Phi = E \cdot A = EA\cos\theta$.
3. Gauss Law states that total flux through any closed surface is proportional to the net charge inside the surface. Mathematically it is given by

   Electric Flux, $\Phi = \oint E \cdot dS = \dfrac{Q}{\varepsilon_0}$
4. Electric Field at a distance R above infinite line charge of charge density, λ is given by $E = \dfrac{\lambda}{2\pi R\varepsilon_0}$

5. The field outside acharged sphere of radius R behaves like a pointcharge located at the centre of the sphere. $E=\frac{1}{4\pi\varepsilon_0}\frac{Q}{r^2}$.

Exercise 2

2.1 Which of the following statements is incorrect?
If a resultant electric field exist in a conductor,
A. The electrons will experience a net force **B.** Electron will be accelerated
C. Protons will flow **D.** Current will flow

2.2 The resultant electric field within a conductor is
A. Q/ε_0 **B.** zero **C.** maximum **D.** minimum value

2.3 Electrical forces do not influence planetary motion because
A. Electric field is zero **B.** Gravitational force is much greater than electric force
C. Electric force is greater **D.** Gravitational force is zero

2.4 Two charges $q_1=10\mu C$ and $q_2=-12\mu C$ are within a spherical surface of radius $10cm$. What is the total flux through the surface in Nm^2C^{-1} ?
A. -2.26×10^6 **B.** 5.56×10^6 **C.** 10.06×10^6 **D.** 20.76×10^6

2.5 A $60\mu C$ charge is at the centre of a cube of side $10cm$. (i) what is the total flux through the cube? **A.** $4.73\,Nm^2C^{-1}$ **B.** $10\,Nm^2C^{-1}$ **C.** $25\,Nm^2C^{-1}$ **D.** $40\,Nm^2C^{-1}$

2.6 What is the flux through one face of the cube above in Nm^2C^{-1}? **A.** 0.0 **B.** 6.0 **C.** 8.0 **D.** 10.0

2.7 Three concentric conducting spherical shells carry charges as follows: $4Q$ on the inner shell, -$2Q$ on the middle shell, and -$5Q$ on the outer shell. Analytically explain the charge distribution on the inner surface of the outer shell.

2.8 A circular plate has a radius of $7cm$. The plane of the plate is set at a $60°$ angle to a uniform electric field $E=355iNC^{-1}$. Calculate the flux through the plate.

2.9 A hemispherical surface of radius $5cm$ is placed in a uniform electric field, $E=250\,NC^{-1}$. What is the maximum electric flux that can pass through the surface?

2.10 A flat plate having an area of $5m^2$ is rotated in a uniform electric field whose magnitude is $7.4\times10^5\,NC^{-1}$. Calculate the electric flux through this area when the electric field is
(i) Perpendicular to the surface of the plate.
(ii) Parallel to the surface.
(iii) Makes an angle of $75°$ with the plane of the surface.

2.11 A uniform electric field $aj+bk$ intersects a surface of area A. What is the flux through this area if the surface lies (i) in the yz plane (ii) in the xz plane.

2.12 A $60cm$ circular loop is rotated in a uniform electric field until the position of maximum electric flux is found. The flux in this position is measured to be $8.5\times10^6\,Nm^2C^{-1}$. What is the magnitude of the electric field?

2.13 A pyramid with a $6m$ square base and height of 4m is placed in a vertical electric field of $52\,NC^{-1}$. Calculate the total electric flux through the pyramid's four slanted surfaces.

2.14 A cube of side L has one corner at the origin of coordinates, and extends a long x, y and z axes. Suppose the electric field in the region is given by $E=(a+by)j$. Determine the charge inside the cube.

2.15 The electric field everywhere on the surface of a hollow sphere of radius $11cm$ is measured to be equal to $3.8\times10^4\,NC^{-1}$ and points radially outward from the centre of the sphere. (i) What is the electric flux through this surface? (ii) How much charge is enclosed by this surface?

2.16 The electric field everywhere on the surface of a hollow sphere of radius $0.75m$ is measured to be equal to $8.9\times10^2\,NC^{-1}$ and points radially toward the centre of the sphere. What can you conclude about the nature and distribution of the charge inside the sphere?

2.17 A $100eV$ electron is fired directly toward a large metal plate that has a surface charge density of $-20\times10^{6}\,Cm^{-2}$. From what distance must the electron be fired if it is to just fail to strike the plate?

2.18 An electron remains stationary in an electric field directed downward in the earth's gravitational field. If the electric field is due to charge on two large parallel conducting plates, oppositely charged and separated by $2.3cm$, what is the surface density assumed to be uniform on the plates?

2.19 A positive charge of magnitude $2.5\mu C$ is at the centre of an uncharged spherical conducting shell of inner radius $60cm$ and an outer radius of $90cm$. (i) Find the charge densities on the inner and outer surfaces of the shell and the total charge on each surface. (ii) Find the electric field everywhere.

2.20 Two charged concentric spheres have radii of $10cm$ and $18cm$. The charge on the inner sphere is $6.0\times10^{-8}\,C$ and that on the outer sphere is $2.0\times10^{-8}\,C$. Find the electric field (i) at $r =$ $12cm$ and (ii) at $r = 25cm$.

CHAPTER 3
CAPACITANCE AND DIELECTRICS

3.0 Introduction

The capacitor is an important electronic device which is used to store charges for later use. We shall discuss how capacitors are used for energy storage and the effects of dielectrics on electric fields and potential differences. Electronic devices are classified as passive or active; capacitors are passive elements. They are classified as passive because they respond electrically to charges in a circuit in ways that depend only on their configuration and on the materials of which they are composed.

3.1 Capacitor

A capacitor consists of two isolated conductors that are close to one another but not touching. Figure 3.1 shows a typical capacitor in a vacuum.

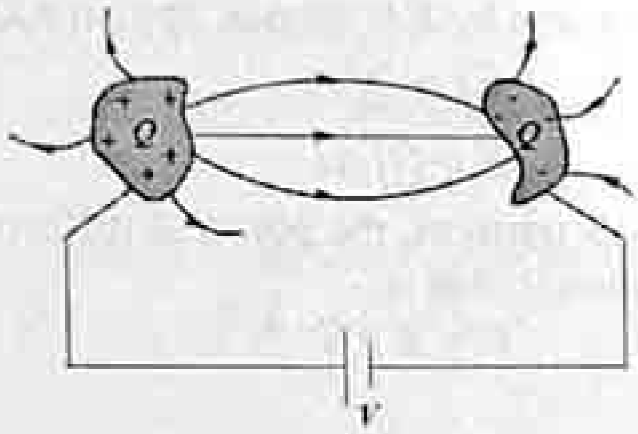


Fig. 3.1: Two conductors carrying equal and opposite charges

The two conductors are totally isolated from objects in their surroundings and they carry equal and opposite charges. The electric field around the capacitor is indicated by the lines of force in Figure 3.1; the lines originate from the positively charged conductor and terminate on the negatively charged conductor.

If the electrical potential difference between the two conductors is V, then the capacitor is characterised by Q, the magnitude of the charge in either conductor. It is found that the charge on each capacitor is proportional to the potential difference between the conductors, i.e.

$$Q \propto V$$

$$\therefore \quad Q = CV \tag{3.1}$$

where C, the constant of proportionality is the capacitance of the capacitor. The unit of capacitance is Coulomb per volt and this unit is called a farad (F). Most capacitors have capacitance considerably much smaller than a farad (F) and they are usually expressed in microfarad ($1\mu F = 10^{-6} F$) or in picofarad ($1pF = 10^{-9} F$).

3.2 Determination of Capacitance

Parallel plate capacitor

It is quite easy to put equal and opposite charges on conductors such as those shown in Figure 3.1. When battery terminals are connected to the conductors, one conductor acquires a negative ($-Q$) charge while the other acquires an equal amount of positive ($+Q$) charge. Figure 3.2 shows a parallel plate capacitor with plates of area, A separated by a distance d. Now, we wish to calculate the capacitance of the parallel plate capacitor. We assume that the plate separation d is smaller compared to the dimensions of the capacitor. With this assumption, the electric field between plates is uniform and perpendicular to each plate.

We can also ignore the fringing of the field lines at the edges since d is small. The space between the capacitor plates may be filled with air or any other insulator, however in this section, we shall assume that it is filled with air.

We want to relate E, the electric field between the plates by means of Gauss's law. The dashed line in the figure is a Gaussian surface of height h closed by plane caps of area A representing the cross-sectional area of the capacitor plates. The flux of E through the upper wall of the Gaussian surface is zero because that wall lies inside the capacitor plate (2) and we know that the electric field inside a conductor carrying a static charge is zero.

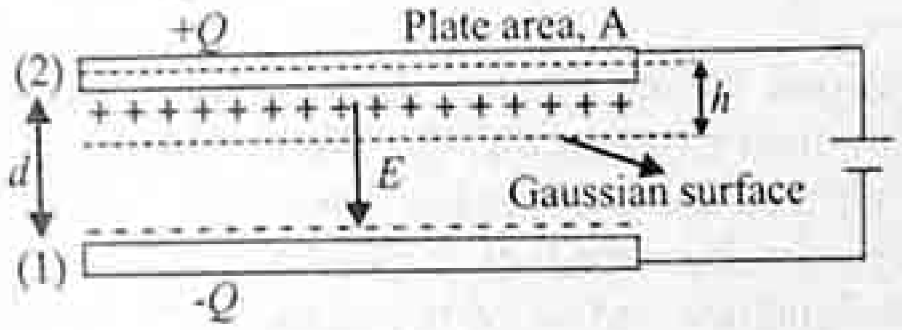


Fig. 3.2: A parallel plate capacitor whose plates have area A

The flux of E through the vertical wall of the Gaussian surface is also zero because E lies in the wall. Thus, the only flux through the Gaussian surface is through the lower wall which lies between the capacitor plates. The flux is simply given by EA. Hence, applying Gauss's law,

$$EA = Q/\varepsilon_0$$

$$E = Q/A\varepsilon_0 \tag{3.2}$$

Since the field between the plates is uniform, the potential difference between the plates equals the electric field times the plate separation d, that is

$$V = E \times d \tag{3.3a}$$

Thus, using Equation 3.2,

$$V = \frac{Qd}{A\varepsilon_0} \tag{3.3b}$$

Therefore,

$$C = Q/V = A\varepsilon_0/d \tag{3.3c}$$

Equation 3.3a which we used to arrive at Equation 3.3c is adequate for this treatment. However, an alternative and more general relationship which we shall use when we deal with cylindrical capacitors will now be introduced.

From Figure 3.2, the potential difference between the plates (1) and (2) is

$$V_2 - V_1 = -\int_1^2 E \cdot dl \tag{3.3d}$$

Substituting Equation 3.2 into Equation 3.3(d) and noting that E and dl point in opposite directions, we obtain

$$V_2 - V_1 = +\frac{Q}{A\varepsilon_0}\int_1^2 dl = \frac{Qd}{\varepsilon_0 A} \tag{3.3e}$$

Equation 3.3e relates the potential difference between the plates and the charge on each plate. Therefore, applying Equation 3.1, we can obtain the capacitance C as before.

$$C = \frac{Q}{V_2 - V_1} = \frac{\varepsilon_0 A}{d} \tag{3.3c}$$

Equation 3.3(c) holds only for parallel plate capacitors. It confirms the fact that capacitance of a fixed pair of conductors is dependent only on the geometry of the conductors. It shows that the capacitance of a parallel plate capacitor may be increased by increasing the area of the plates or decreasing the distance d between the plates. Since the capacitance depends upon the geometry of the capacitor, the expressions for the capacitances of capacitors of other geometry must necessarily be different from Equation 3.3(c).

Example 3.1

A $2.0F$ capacitor is desired. What should the area of the plates be if they are to be separated by a $4.5mm$ air gap?

Solution

This is a parallel plate capacitor, therefore, using Equation 3.3(c) we have

$$A = \frac{Cd}{\varepsilon_0} = \frac{(2F)(4.5\times10^{-3}m)}{(8.85\times10^{-12}C^2/Nm^2)} = 1\times10^9 m^2$$

3.3 Capacitors in Series and in Parallel

In a practical electronic circuit, capacitors are usually connected in a variety of ways. We show the two basic methods in which capacitors are connected in Figure 3.3. In drawing electronic circuits, a capacitor is represented by two equal parallel lines with horizontal connecting wires ⊣⊢.The battery is represented by the symbol, ⊣⊢; the shorter line represents the negative polarity of the battery. Now we wish to find a single capacitor that has equivalent capacitance as the combination in each case.

Let us consider the parallel connection in Figure 3.3(a); a battery of potential difference V is connected to points A and B. Each capacitor has the same potential difference V across it, because all the left hand plates of the capacitors (C_1, C_2 and C_3) are connected to point A and all the right-hand plates are connected to point B. Each capacitor therefore acquires a charge given by $Q_1 = C_1V$, $Q_2 = C_2V$ and $Q_3 = C_3V$.

The total charge that must leave the battery is the sum of charge stored in each capacitor. Thus,

$$Q = Q_1 + Q_2 + Q_3 = V(C_1 + C_2 + C_3) \tag{3.4a}$$

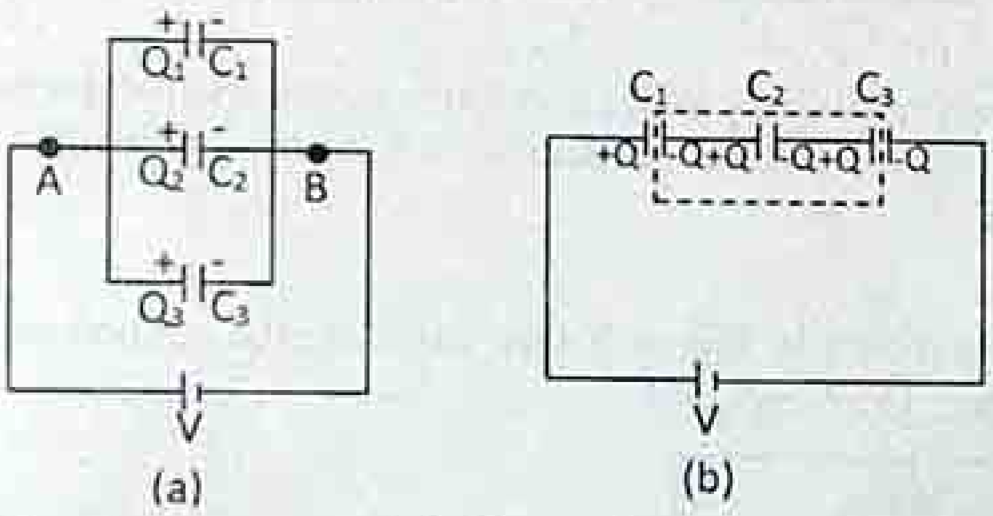


Fig. 3.3: Capacitors connected. (a) in parallel and (b) in series

A single capacitor that is equivalent to the combination must hold the same charge for the potential difference V. Therefore,

$$Q = C_{eq}V \tag{3.4b}$$

Comparing Equations 3.4(a) and 3.4(b) we see that

$$C_{eq} = C_1 + C_2 + C_3 \tag{3.5}$$

for capacitors connected in parallel.

Equation 3.5 shows that the net effect of combining capacitors in parallel is to increase the capacitance. This makes sense because by combining the capacitors in parallel, we are actually increasing area of the plates for charge to accumulate. If the area is increased, the capacitance must also increase accordingly. Equation 3.5 may be extended to any number n of capacitors connected in parallel:

$$C_{eq} = C_1 + C_2 + C_3 + C_4 + \ldots + C_n \tag{3.6}$$

For the series combination shown in Figure 3.3(b), the potential difference V across the combination is the potential difference between the left plate of C_1 and the right plate of C_3. The battery deposits $+Q$ charge on the left-hand plate of C_1 and $-Q$ charge on the right-hand plate of C_3. The net charge of the dotted area of the figure cannot change since no charge enters or leaves this region. The $+Q$ charge on the left-hand plate of C_1 induces $-Q$ on the right-hand plate of C_1. Similarly the $-Q$ charge on the right-hand plate of C_3 induces $+Q$ on the left-hand plate of C_3. Following this line of reasoning we can see how $+Q$ appears on the left-hand plate and $-Q$ appears on the right-hand plate

C_2. So we see that each capacitor in a series combination has the same charge Q. Therefore, the equivalent capacitance of the series combination must also have the same charge Q for the same potential difference V.

The total potential difference, V across the three capacitors in series must equal the sum of the potential difference across each capacitor:

$$V = V_1 + V_2 + V_3 \tag{3.7a}$$

But we know that

$$Q = C_1V_1,\ Q = C_2V_2,\ \text{and}\ Q = C_3V_3.$$

Solving for V_1, V_2 and V_3 and substituting for V_1, V_2 and V_3 in Equation 3.7(a) we obtain

$$\frac{Q}{C_{eq}} = \frac{Q}{C_1} + \frac{Q}{C_2} + \frac{Q}{C_3} \tag{3.7b}$$

where we have used the fact that $Q/C_{eq} = V$

or

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \tag{3.7c}$$

For n capacitors connected in series, Equation 3.7(c) may be extended to give:

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \ldots + \frac{1}{C_n} \quad \text{(capacitors in series)}$$

It is quite easy to get confused when solving problems involving network of capacitors when some are connected in series while others are connected in parallel. These examples illustrated below shows how such networks are analysed.

Example 3.2

For a combination of capacitors shown in Figure 3.4(a), calculate the equivalent capacitance of the combination and the charge on the $15\mu F$ capacitor.

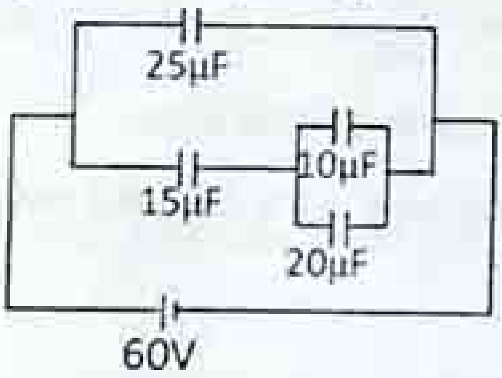


Fig. 3.4(a): Example 3.2

Solution

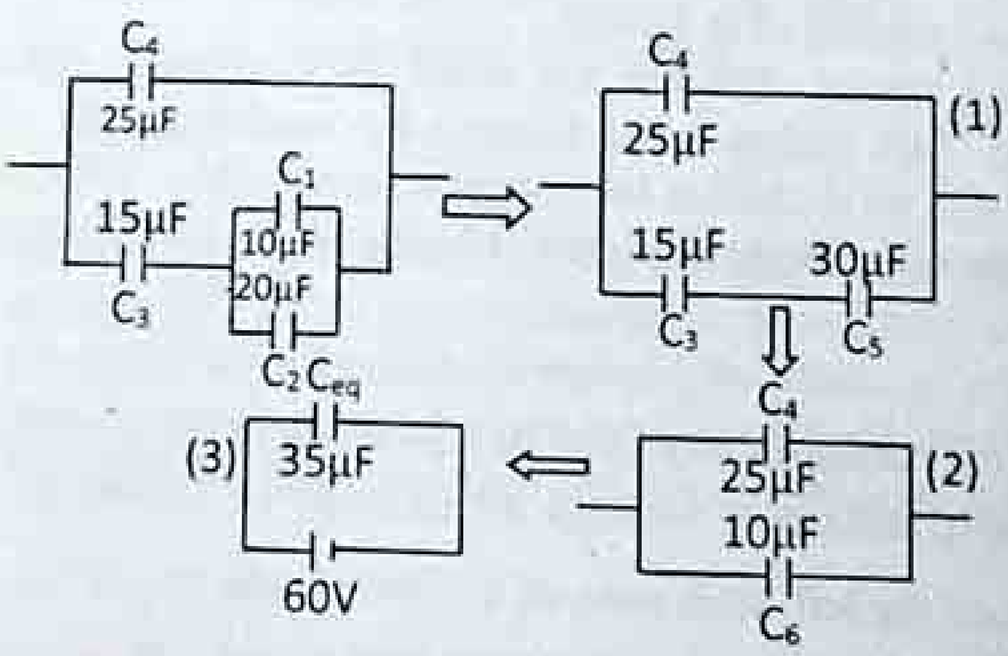


Fig. 3.4(b): Solution of Example 3.2

It is a good practice to show diagrammatically every successive stage in other to reduce the combination. We do this in Figure 3.4(b). To obtain stage (1), C_1 and C_2 are added in parallel. That is,

$$C_5 = C_1 + C_2 = 10 + 20 = 30\mu F$$

To obtain stage (2), C_3 are added in series. That is,

$$\frac{1}{C_6} = \frac{1}{C_3} + \frac{1}{C_5} = \frac{1}{15} + \frac{1}{30} = \frac{1}{10}$$

or $\quad C_6 = 10\mu F$

This leads to stage (3) as shown in the diagram. C_6 and C_4 are parallel and using the relation $C = 4\pi\varepsilon_0 A$, we have $C_{eq} = C_6 + C_4 = 10 + 25 = 35\mu F$

In order to find the charge on C_3, we have to work backwards following the stages. The charge on C_{eq} is: $\quad Q_{eq} = CV = (35 \times 10^{-6} F)(60V) = 2.1 \times 10^{-3} C$

The potential difference across C_6 is 60 Volts. Therefore, the charge on C_6 is:

$$Q_6 = (10 \times 10^{-6})(60) = 0.6 \times 10^{-3} C$$

But C_6 was obtained by combining C_3 and C_5 in series.

Therefore, C_3 and C_5 must accumulate the same charge as their equivalent which is C_6. Hence, $Q_3 = 0.6 \times 10^{-3} C$

Example 3.3

Find the charge on each of the capacitors shown in Figure 3.5 as well as the potential difference across each capacitor.

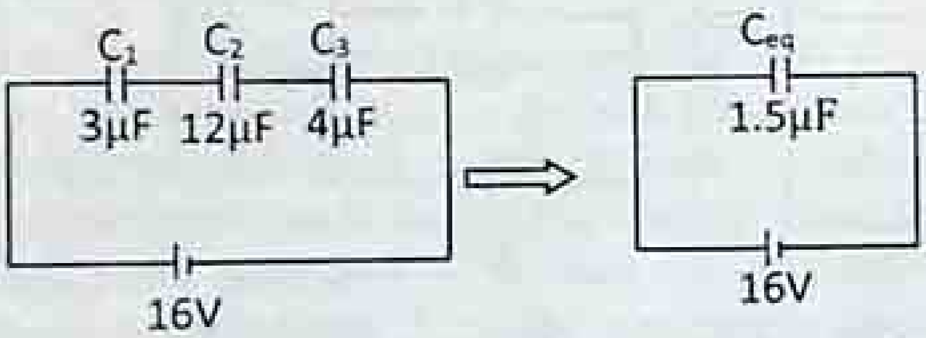


Fig 3.5: Example 3.3

Solution

In order to solve this problem, we need to find the equivalent capacitance for the capacitor:

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{1}{3\mu F} + \frac{1}{12\mu F} + \frac{1}{4\mu F}$$

$$C_{eq} = 1.5\mu F$$

The charge on the equivalent capacitor is $Q_{eq} = 24\mu C$

Since the three capacitors are in series, they store the same charge as the equivalent capacitance. The voltage on each capacitor is:

$$V_1 = \frac{Q}{C_1} = \frac{24 \times 10^{-6} C}{3 \times 10^{-6} F} = 8\,Volts \qquad V_2 = \frac{Q}{C_2} = \frac{24 \times 10^{-6} C}{12 \times 10^{-6} F} = 2\,Volts$$

and $\quad V_3 = \dfrac{Q}{C_3} = \dfrac{24 \times 10^{-6} C}{4 \times 10^{-6} F} = 6\,Volts$

We must note that $V_1 + V_2 + V_3 = 16\ Volts$, the voltage of the battery. This is an expected result.

Example 3.4

Find the equivalent capacitance for the circuit shown in Figure 3.6(a). Also find the charge and the potential difference across each capacitor.

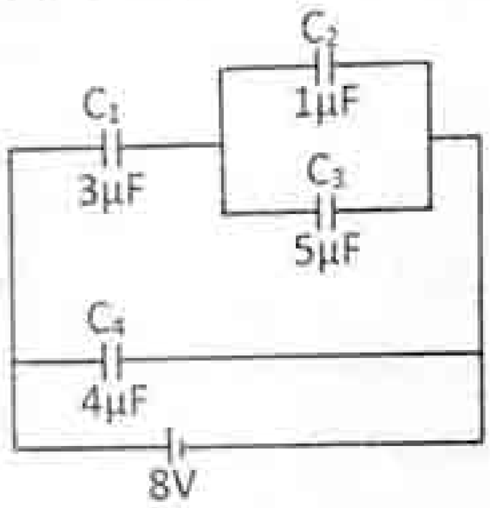


Fig 3.6(a): Example 3.4

Solution

As usual, we draw the diagram for every successive reduction, this is shown in Figure 3.6(b). We see that C_2 and C_3 are parallel, these two can be replaced with a capacitor having a capacitance $C_5 = 6\mu F$

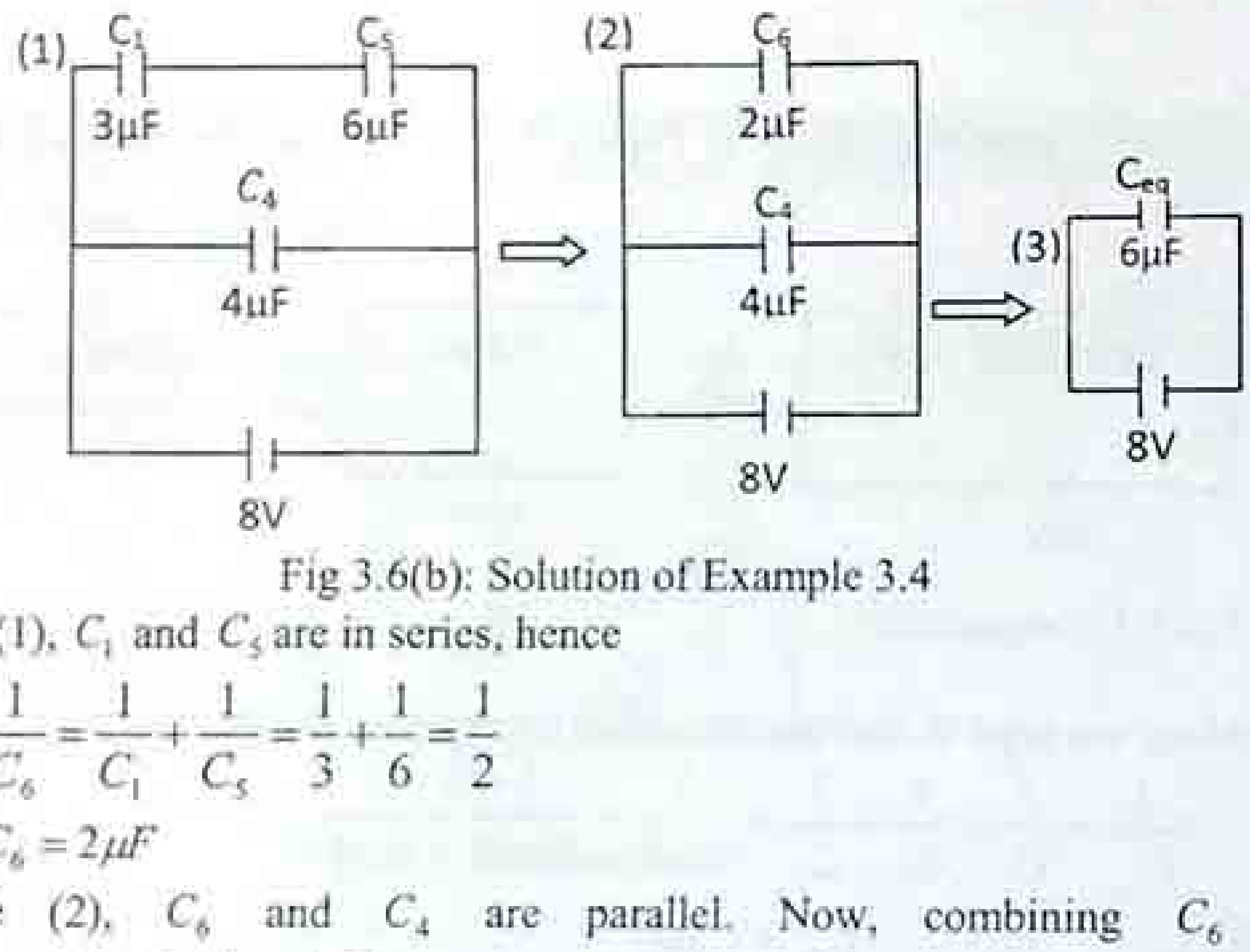


Fig 3.6(b): Solution of Example 3.4

At stage (1), C_1 and C_5 are in series, hence

$$\frac{1}{C_6} = \frac{1}{C_1} + \frac{1}{C_5} = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}$$

$$C_6 = 2\mu F$$

At stage (2), C_6 and C_4 are parallel. Now, combining C_6 and C_4 we obtain $C_{eq} = C_6 + C_4 = 2 + 4 = 6\mu F$.

The equivalent charge, $\qquad Q_{eq} = C_{eq}V = 48\times10^{-6}C.$

Now, to find the charge on each capacitor we must trace our steps backwards as we did in the last example. We added C_6 and C_4 in parallel in order to obtain C_{eq}; this means that the potential difference across C_6 and C_4 is the same as that across C_{eq}. Then we have $V_4 = 8\,volts$ and $V_6 = 8volts$.

$$Q_4 = C_4V_4 = (4\times10^{-6})(8) = 32\times10^{-6}C$$

Similarly, $Q_6 = 16\times10^{-6}C$.

However, C_6 is the equivalent of C_1 and C_5 in series.

Therefore, the charge on C_1 is $Q_1 = 16\times10^{-6}C$

$$V_1 = \frac{Q_1}{C_1} = 5.33\, volts$$

Similarly, $V_5 = \dfrac{16\times10^{-6}C}{6\times10^{-6}F} = 2.66\, volts$

and $\quad Q_5 = 15.96\,\mu C$

The sum of V_1 and V_2 is equal to 8 volts as expected. C_5 was obtained by combining C_2 and C_3 in parallel.Hence, C_2 and C_3 have the same voltage as C_5

That is, $\qquad V_2 = V_3 = 2.66\, volts$

$$Q_2 = C_2V_2 = (1\times10^{-6})(2.66) = 2.66\,\mu C$$

and $\quad Q_3 = (5\times10^{-6})(2.66)C = 13.30\,\mu C$

The charge on C_5 must be equal to Q_2 added to Q_3.

3.4 Energy Stored in Capacitors

We saw previously that all charge configurations have electrical potential *energy, U.* This energy is equal to the work done in assembling the individual charged particles originally at rest at infinity and far apart from one another. The energy is stored in the system.

The capacitor stores energy, this energy is equal to the work done in charging the capacitor. A capacitor is charged by transferring charge from one of the two conductors to the other. This process requires the expenditure of energy which is usually provided by the battery when it is connected to a capacitor. The energy expended by the battery in charging the capacitor is stored in the capacitor. This energy is however recoverable when the capacitor is allowed to discharge.

Let us consider an uncharged capacitor, it requires no work to carry the first bit of charge from one plate to the other. However, when some charge are on each plate, there is now a potential difference between the plates and it now requires work to transfer additional charges to the plates. The more charges already on a plate, the more work is required to transfer more. Suppose, at a certain stage of charging a capacitor, charge q has been transferred from one conductor to the other, there exists a potential difference V between the two plates. The voltage across the plates is

$$V = \frac{q}{C}$$

The work done ΔW necessary to transfer an additional charge Δq is $\Delta W = V\Delta q$

$$\Delta W = \frac{q}{C}\Delta q$$

or

$$dW = \frac{q}{C}\,dq \qquad (3.8)$$

The total work that must be done to increase the charge from $q=0$ to $q=Q$ equal to the energy stored in the capacitor. Therefore,

$$U = W = \int dW = \frac{1}{C}\int_0^Q q\,dq \qquad (3.9)$$

Integrating Equation 3.7 we have $\qquad U = \dfrac{Q^2}{2C} \qquad (3.10)$

When the capacitor C carries $+Q$ and $-Q$ on its two conductors. Using the definition $Q = CV$,

Equation 3.8 may be expressed as $\qquad U = \dfrac{1}{2}\dfrac{Q^2}{C} = \dfrac{1}{2}CV^2 \qquad (3.11a)$

$$= \frac{1}{2} QV \qquad (3.11b)$$

This result can also be obtained graphically by noting that since $q = VC$, we expect q to increase as V increases; the plot of V against q is a straight line as shown in Figure 3.7. The work done in transferring charge against potential difference V, which is given by $\Delta W = V\Delta q$ is equal to the area of the strip shown in the diagram. The total work done to increase the charge from 0 to q is the area under the straight line from $q = 0$ to $q = Q$. That gives as before $U = \frac{1}{2} QV$

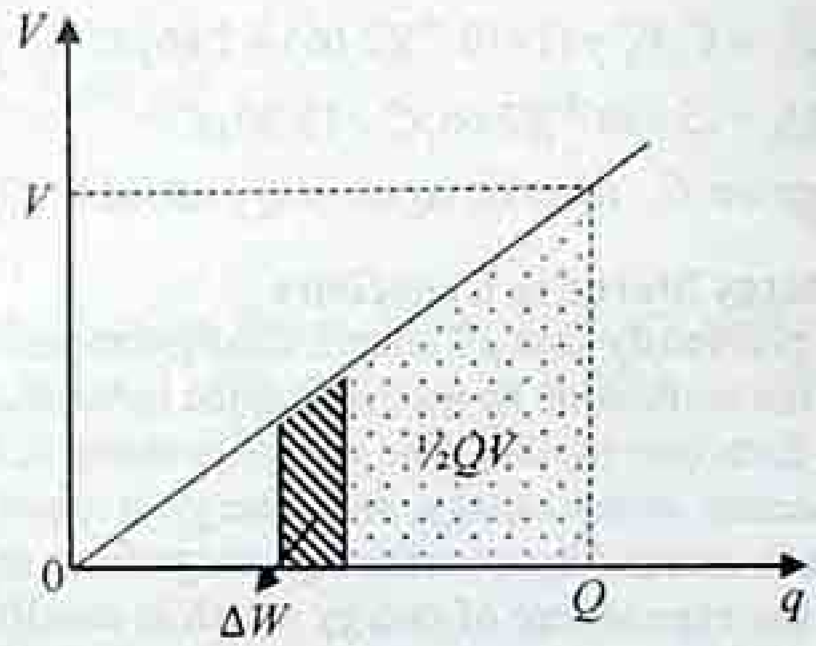


Fig. 3.7: The work done in charging a capacitor.

We have seen that the electric field in a parallel plate capacitor, if we ignore the fringe effects, is uniform and has the same value for all points between the plates. The magnitude of the electric field is related to the electric potential difference V between the plates as

$$V = Ed \qquad (3.12)$$

Where d is the separation between the plates. We had obtained the capacitance for a parallel plate capacitor as in equation 3.3(c):

$$C = \frac{\varepsilon_0 A}{d} \qquad (3.3c)$$

Now substituting Equation 3.13 and 3.3(c) into Equation 3.11(a) we obtain

$$U = \frac{1}{2} \varepsilon_0 E^2 Ad \qquad (3.13)$$

But Ad is the volume of the parallel plate capacitor. Dividing Equation 3.13 by the volume, we obtain the energy density of the parallel plate capacitor.

$$u = \frac{U}{Ad} = \frac{1}{2} \varepsilon_0 E^2 \qquad (3.14)$$

Equation 3.14 states that *the electric energy density per unit volume in any region of space is proportional to the square of the electric field in that region*. Even though we have derived Equation 3.14 for a special case of parallel plate capacitor, it is true in general. Thus, if electric field exists at any point in space, there is energy stored in that point in amount per unit volume equal to $\frac{1}{2}\varepsilon_0 E^2$.

Example 3.5

A $2.0\mu F$ capacitor is charged by a $12V$ battery. It is disconnected from the battery and then connected to an uncharged $5.0\mu F$ capacitor. Determine the total stored energy
(a) before the two capacitors are connected and (b) after they are connected.

Solution

The situation is illustrated in Figure 3.8(a) and Figure 3.8(b). The $12V$ battery charged the $2\mu F$ capacitor fully.

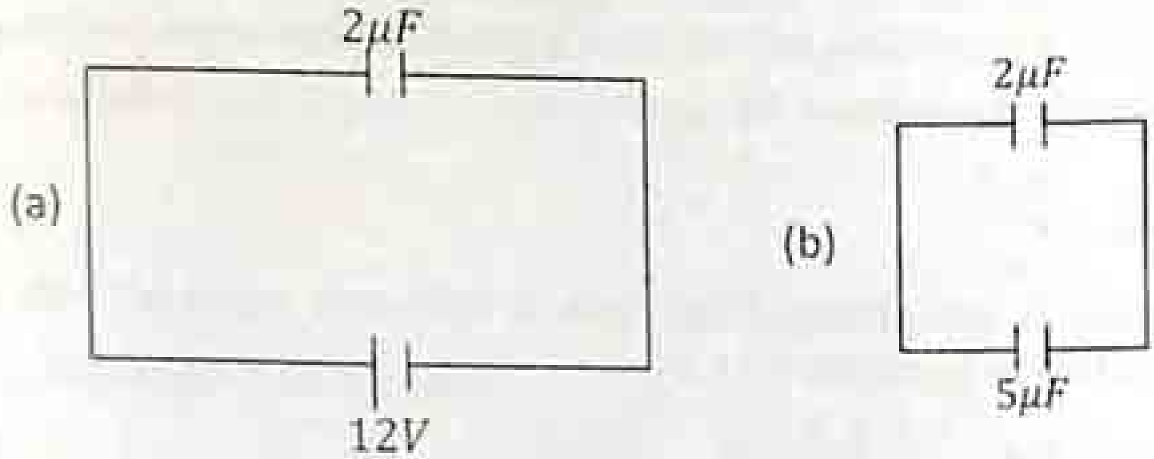


Fig. 3.8: Example 3.5

Therefore, the energy stored in the capacitor using Equation 3.11(a) is:

$$U = \frac{1}{2}CV^2 = \frac{1}{2}(2\times10^{-6})(12)^2 = 1.44\times10^{-4}\,J$$

After the battery has been removed, the $5\mu F$ capacitor is connected in parallel with the charged $2\mu F$ capacitor. The equivalent capacitance of the network is $7\mu F$ stored in $2\mu F$ capacitor is

$$Q = CV = 2\times12 = 24\mu C$$

Hence, the energy stored in this combination is $U_p = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}\frac{(24\times10^{-6})^2}{7\times10^{-6}} = 4.11\times10^{-5}\,J$

Comparing the two energies, we see that energy retained is less than the initial energy. This is because of losses to heat in redistributing the charge.

3.5 Dielectrics

So far, we have assumed in our discussions of capacitors that capacitors are in vacuum or in air. In other words, there are no other materials between the two conducing plates of the capacitor. However, there is a dielectric material (which is insulating) between the plates of practical capacitors. It is found experimentally that when a dielectric material is inserted between the conductors of a capacitor, the capacitance of the capacitor is increased by a certain factor K called the dielectric constant of the dielectric. Suppose the capacitance of a capacitor with a vacuum between the two conductors is C_0 and the capacitance is C when the space is filled with a material of dielectric constant K. Then

$$C = KC_0 \tag{3.15}$$

the dielectric constant, K has no unit and it is characteristic of a given material. The values of the dielectric constant K for some common materials used in capacitors are given in Table 3.1.

Table 3.1: Dielectric constant and strengths of common materials at 20°C

Materials	**Dielectric constant (K)**	**Dielectric strength (KV/mm)**
Vacuum	1.0	
Glass(pyrex)	4.5	14.0
Paper	3.5	16.0
Porcelain	6.0 – 8.0	5.7
Ruby Mica	5.0 – 7.5	10.0 – 100.0
Paraffin	2.3	10.0
Rubber	2.8	21.0
Water (distilled)	80.0	3.0
Ethyl alcohol	0	24.0
Air (at 1 atm)	1.0	3.0
Beeswax	2.9	
Titanium dioxide	100.0	6.0
Polstyrene	2.5	24.0
Neoprene	6.9	12.0

Michael Faraday was the first to find that the effect of dielectric material filling increases the capacitance of the capacitor by a factor of K. Suppose we have a parallel plate capacitor in a vacuum, we know, from Equation 3.3(c), that its capacitance C_0 is:

$$C_0 = \frac{\varepsilon_0 A}{d} \tag{3.3c}$$

When the space between the plates is completely filled with a dielectric material with dielectric constant K, its capacitance C according to Equation 3.15 is

$$C = KC_0 = \frac{K\varepsilon_0 A}{d} \tag{3.16}$$

Equation 3.16 reduces to Equation 3.3(c) if $K = 1$, which corresponds to a vacuum between the plates. Let us define K as the permittivity ε of the material. Thus,

$$\varepsilon = K\varepsilon_0 \tag{3.17}$$

Then the capacitance of a parallel plate capacitor (Equation 3.16) becomes

$$C = \frac{\varepsilon A}{d}$$

We must note that ε_0 is always greater than ε the permittivity of free space. However, $\varepsilon = \varepsilon_0$ if $K =$ 1; this corresponds to free space. Let us also recall that the energy density stored in the electric fields E is (from Equation 3.14)

$$u = \frac{1}{2}\varepsilon_0 E^2$$

With the introduction of the dielectric materials, the energy density becomes

$$u = \frac{1}{2}K\varepsilon_0 E^2 = \frac{1}{2}\varepsilon E^2 \tag{3.18}$$

Effect of Dielectric on Capacitance of Capacitor

Let us now examine other effects of inserting a dielectric material between the two conductors of a capacitor. Suppose a capacitor is connected to a battery of voltage V_0 and the capacitor is fully charged. It has acquired a charge Q_0. Therefore,

$$Q_0 = C_0 V_0$$

Then the battery is disconnected, before a dielectric material with a dielectric constant K is inserted in the space between the conductors. The charge on each plate is still Q_0. Here it is necessary to disconnect the battery before the dielectric material is inserted; this is to ensure the charge remains constant. Then according to Equation 3.15;

$$C = KC_0$$

Hence, $$\frac{Q_0}{V} = \frac{KQ_0}{V_0} \quad \text{(constant charge)}$$

Therefore, $$V = \frac{V_0}{K} \tag{3.19}$$

Equation 3.19 states that *the potential difference V between the plates decreases by a factor K*. Let us consider a parallel plate capacitor with a distance d between the two conductors. It is connected to a battery; the potential difference between the plates is V_0. Then as above $Q_0 = C_0 V_0$. When a dielectric material is inserted, the potential difference across the plates is still V_0 then we have

$$Q = CV_0 = KC_0 V_0 = KQ_0 \tag{3.20}$$

Thus, the charge on a capacitor with a fixed potential difference between its plates is increased by factor K because of the presence of a material with dielectric constant K. Let the electric field between the parallel plate capacitor be E_0 and it is E after the material with dielectric constant K is inserted into the space between the plates. If the separation between the plates is d, then,

$$E_0 = \frac{V_0}{d}$$

Similarly,

$$E = V/d$$

Then using Equation 3.19 we have,

$$E = \frac{E_0}{K} \tag{3.21}$$

From Equation 3.19, we see that the electric field within the dielectric is reduced by a factor equal to the dielectric constant. However, it is not reduced to zero as it would have been in a conductor.

3.6 Molecular View of the Effects of a Dielectric

We wish to explain the action of the dielectric inserted between capacitor plates from an atomic viewpoint. The molecules in a dielectric material have permanent dipole moments. In the absence of an electric field, such molecules have random orientation owing to thermal agitation (Figure 3.9).

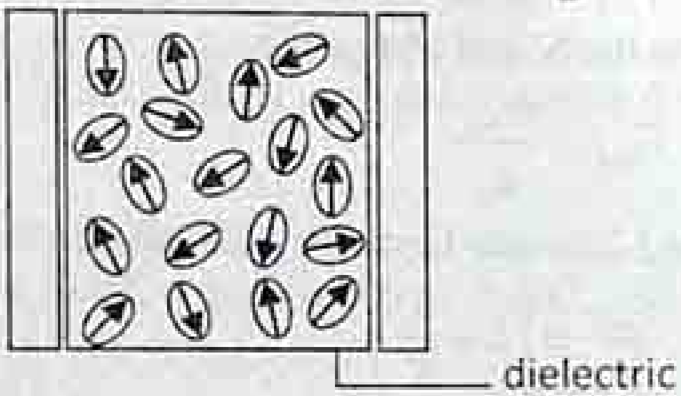


Fig. 3.9: Molecules with a permanent electric dipole moment showing their random orientation in the absence of an external electric field

When the plates of the capacitor are charged, the resulting electric field causes the dielectric molecules to align themselves with the field. The electrons are attracted to the positively charged plate and the nuclei to the negatively charged plate (Figure 3.10). The electrons in insulators not been free, are slightly displaced. This alignment called polarization, produces no net charge within the body of the dielectric. However, net negative charges appear on the edge of the dielectric facing the negative plate of the capacitor. The net charges on the dielectric are bound or induced charges as opposed to free charges in the plates. The induced charges create electric field E_{ind} that is opposite to the original field E (Figure 3.11). The process therefore weakens the electric field within the dielectric as well as outside. Since $E = V/d$, V will also be reduced and hence the capacitance$(= Q/V)$ is enhanced. For materials with non-polar molecules (where the centre of nuclei and centre of electronic cloud coincide), polarization also occurs in an electric field owing to temporary displacement of charge within the molecule, however when the capacitor is being discharged, depolarization or dis-alignment takes a very short time compared to the case of polar molecules which can take much longer time.

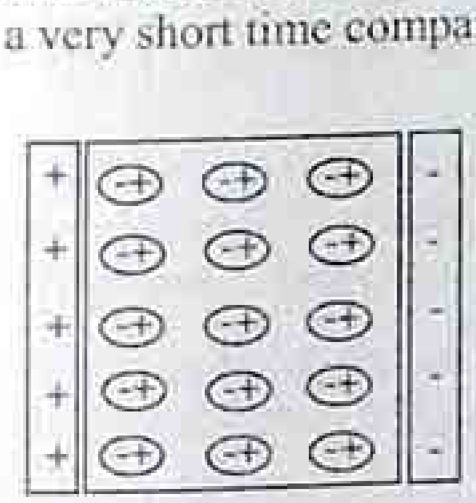


Fig.3.10: The applied external electric field shifts all the positive charges in the direction of the field

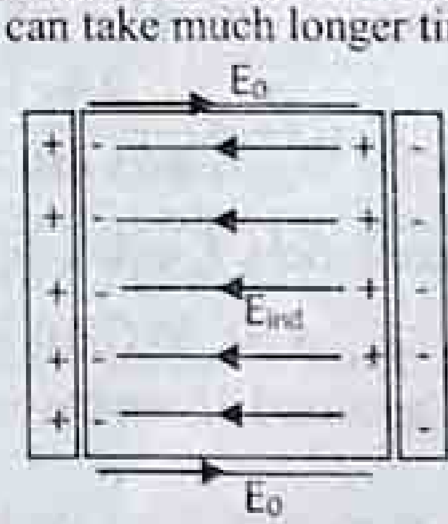


Fig. 3.11: Induced charges on a dielectric in an external field

From Figure 3.11, $\vec{E}_0$ and $\vec{E}_{ind}$ are in opposite directions, hence the net field within the dielectric is

$$E_0 - E_{ind} = E$$

But using Equation 3.21, E may be written as

$$E_0 - E_{ind} = \frac{E_0}{K}$$

Therefore,

$$E_{ind} = E_0\left(1 - \frac{1}{K}\right) \tag{3.22}$$

The electric field between two parallel plate capacitor is given as:

$$E_0 = \frac{\sigma}{\varepsilon_0} \tag{3.23a}$$

where $\sigma = Q/A$ is the surface charge density on the conductor. In the same way, we define the induced surface charge density $\sigma_{ind} = Q_{ind}/A$, as the bound charge on the surface of the dielectric. Therefore, the electric field due to the bound charges is

$$E_{ind} = \frac{\sigma_{ind}}{\varepsilon_0} \tag{3.23b}$$

Substituting Equations 3.23a and 3.23b into Equation 3.22 we obtain;

$$\frac{\sigma_{ind}}{\sigma} = \left(1 - \frac{1}{K}\right) \tag{3.24a}$$

And $$Q_{ind} = Q\left(1 - \frac{1}{K}\right) \tag{3.24b}$$

Since K is always greater than 1, from Equation 3.24b, the bound charge Q_{ind} is always less than the free charge on each plate of the capacitor.

Example 3.6

Two parallel plates of area $100cm^2$ are each given equal but opposite charges of $8.9\times10^{-7}C$. The electric field within the dielectric material filling the space between the plates is $1.4\times10^6 V/m$.
(a) Find the dielectric constant of the material and (b) determine the magnitude of the charge induced on each dielectric surface.

Solution

The area of each Plate A is $A = 10^{-2}m^2; Q = 8.9\times10^{-7}C$

Therefore, $\sigma = Q/A = 8.9\times10^{-5}C/m^2$

But, $E_0 = \frac{\sigma}{\varepsilon_0} = \frac{8.9\times10^{-5}C/m^2}{8.85\times10^{-12}(C^2/N.m^2)} \cong 10^7\ N/C$

The electric field, E within the dielectric is given by $E = E_0 - E_{ind}$

Thus, $E_{ind} = 10^7\ N/C - 0.14\times10^7\ N/C = 0.86\times10^7\ N/C$

From Equation 3.22, we have

$$E_{ind} = E_0\left(1 - \frac{1}{K}\right)$$

$$0.86\times10^7\ N/C = 10^7\left(1 - \frac{1}{K}\right)$$

Solving the above for K, we have $K = 7.14$

To find the magnitude of the induced charge on each plate, we make use of Equation 3.24(b). That is,

$$Q_{ind} = Q\left(1 - \frac{1}{K}\right) = (8.9\times10^{-7}C)(0.86) = 7.65\times10^{-7}C$$

The induced charge is less than the free charge on the capacitor plates. This is an expected result.

Example 3.7: A Paper-Filled Capacitor

A parallel-plate capacitor has plates of dimensions $2.0cm$ by $3.0cm$ separated by a $1.0mm$ thickness of paper. Find its capacitance?

Solution

Because $K = 3.7$ for paper, we have

$$C = K\frac{\varepsilon_0 A}{d} = 3.7\left(\frac{(8.85\times10^{-12}C^2/Nm^2)(6.0\times10^{-4}m^2)}{10\times10^{-3}m}\right) = 20\times10^{-12}F = 20pF$$

Example 3.8

Given a $7.4pF$ air-filled capacitor, you are asked to convert it to a capacitor that can store up to $7.4\mu J$ with a maximum voltage of $652V$. What dielectric constant should the material have that you insert to achieve these requirements?

Solution

The capacitance with the dielectric in place is given by $C = KC_{air}$ and the energy stored is given by

$$U = \tfrac{1}{2}CV^2 = \frac{1}{2}KC_{air}V^2$$

so, $$K = \frac{2U}{C_{air}V^2} = \frac{2(7.4\times10^{-6}J)}{(7.4\times10^{-12}F)(652V)^2} = 4.7$$

3.7 Some Commercial Capacitors and Applications

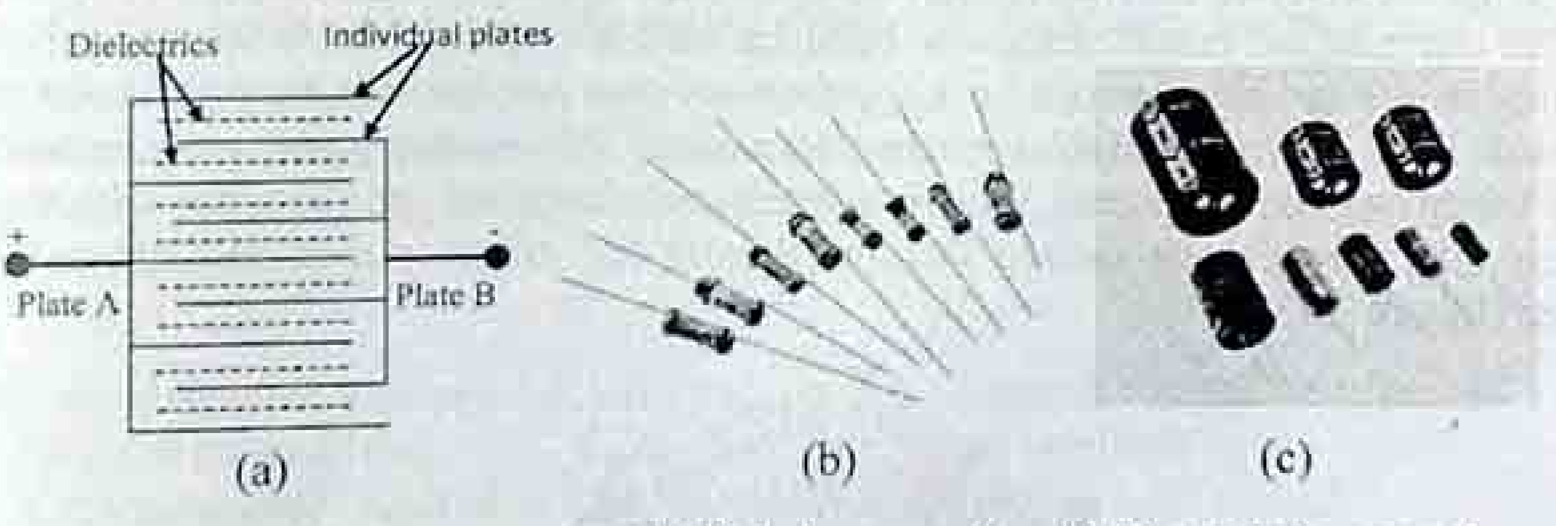


Fig. 3.12: (a) Multi-plate capacitor (b) Tubular capacitor (c) Electrolytic capacitor

Several factors are taken into account when constructing capacitors. For example, to increase the capacitance to be able to hold large quantities of electricity, use is made of a large number of plates each of large area. A diagram of such a multi-plate capacitor is as shown in Figure 3.12a.

The dielectric between the plates in a high grade capacitor can be mica. Note that when there are N plates, the capacitance is increased by factor N - 1 since there are effectively N - 1 individual capacitors.

Cheap commercial type capacitors of capacitance up to 10 microfarad are much more compact. The tubular capacitor consists of plates of thin metal foil (aluminium or tin) separated by a sheet of paper or plastic film and rolled up under pressure into a small package. The whole unit is then encapsulated in plastic film (Figure 3.12b). The dielectric in this case achieves three functions – (i) increases the capacitance because of its polarization, (ii) provides physical separation of the metal foil so that the sheet can be very close together without being in physical contact (small d leads to high C), and (iii) the dielectric strength of the paper is greater than that of air, therefore greater potential difference can be attained without dielectric breakdown. Recall that as more charge is placed on the plates of a capacitor, the potential difference builds up until finally the dielectric can no longer support the potential gradient. The term dielectric strength refers to that property of the insulator which determines the maximum potential gradient which can be applied to the material before its insulating properties are destroyed by a disruptive discharge of electricity through the insulator. Typical values of dielectric strength for some dielectric are also given in Table 3.1 in units of kV per mm of thickness of material.

Another form of compact capacitor is the electrolytic capacitor. This is made by using an insulating layer formed by chemical action directly on the metal plates of the capacitor. The small space between the layers is filled with an electrolyte in liquid or paste form which constitutes one of the plates (Figure 3.12c).

The thin film capacitors which are used in modern electronic industry are usually of the parallel plate configuration in which case, two conducting layers separated by a layer of a suitable dielectric are supported by an insulating substrate about 100 times thicker. An aluminium film is a typical conducing layer while SiO and Ta_2O_5 films form suitable dielectrics.

Capacitors find use in many devices, for example, the energy storage capability is used in (i) the photovoltaic system. This is a system where sunlight is converted into electricity when incident on suitable semiconductors. During the day the excess energy from such units is stored in capacitors for use at night when there is no sunlight; (ii) the flash attachment to a camera uses a capacitor to store energy needed to provide the sudden flash of light; (iii) Some emergency light units would have a bank of capacitors which will remain fully charged when electricity from the public power supply is on. When the public power supply is cut off, the energy stored in the capacitors is used to provide light for some hours until the capacitors are discharged.

Capacitors are also used in power supply units to remove ripples that arise when alternating current from the mains is converted into direct current. The direct current is required in many applications, for example, for charging car batteries and also batteries used in electricity experiments in the laboratory; for operating radios and calculators e.t.c. Capacitors are also employed in tuning circuits of radio and TV receivers. In this case, a variable air capacitor is used. Alternative plates are connected together, one group being capable of rotation. This is actually a multi-plate-type capacitor with the difference that one set of plates can be rotated thereby altering the effective area and hence the effective capacitance (Figure 3.13).

Fig. 3.13: Variable capacitor

Summary

1. A capacitor is a device for storing electrical charge and energy. It consists of two conductors, closely spaced but insulated from each other, carrying equal but opposite charges. The capacitance C of a capacitor is defined by $C = charge/potential = Q/V$
   The unit of capacitance is the Farad (F) = Coulomb/Volt.

2. Capacitance depends only on the geometrical arrangement of the conductors and not on the charge or the potential difference. A parallel plate capacitor has a capacitance given by $C = \varepsilon_0 A/d$, where A is the area of the plates and d the distance between the plates. The capacitance of a cylindrical capacitor is given by $C = 2\pi\varepsilon_0 L/\ln(b/a)$, where L is the length of the capacitor and a and b are the radii of the inner and outer conductors respectively. A spherical capacitor having plates of inner and outer radii a and b respectively has capacitance $C = 4\pi\varepsilon_0 ab/(b-a)$ as $b \to \infty$, we have $C = 4\pi\varepsilon_0 a$. Which is the capacitance of an isolated sphere of radius a

3. When capacitors of capacitance $C_1, C_2, C_3 \ldots$ are connected in parallel, the equivalent capacitance C is given by: $C = C_1 + C_2 + C_3 + \ldots$

   For capacitors in series, the equivalent capacitance if given by:
   $1/C = 1/C_1 + 1/C_2 + 1/C_3 + \ldots$

4. The electrostatic potential energy stored in a capacitor is given by $U = \frac{1}{2}Q^2\Big/C = \frac{1}{2}QV = \frac{1}{2}CV^2$. It is defined as the work required to charge it. This energy can be considered to be stored in the electric field between the plates. The energy per unit volume or energy density in the electric field E is given by $u = \frac{1}{2}\varepsilon_0 E^2$.

5. When a dielectric (insulator) is inserted between the plates of a capacitor, the molecules in the dielectric field within are polarized and the electric field within the dielectric is weakened. If E_0 is the field without the dielectric, the field with the dielectric is $E = E_0/K$

   where K is called dielectric constant. The decrease in the electric field leads to an increase in the capacitance by factor K. Thus, $C = KC_0$, where C_0 is the capacitance without the dielectric. The permittivity of a dielectric ε is defined as $\varepsilon = K\varepsilon_0$.

   The bound or induced charge on the surface of the dielectric is given by $Q_{ind} = Q(1 - 1/K)$

   Since K is always greater than 1, the bound charge Q_{ind} is always less than the free charge on each plate of the capacitor.

6. Capacitors are used in many devices because of the storage capability. They also find use when $a.c$ is being converted to $d.c$. Other applications include use in tuning circuits of radio and television receivers.

Exercise 3

3.1 Cells are connected in parallel in order to: **A.** Increase the charge available
B. Reduce cost of wiring **C.** Increase the current available
D. Reduce the time required to fully charge them effect use

3.2 The combined resistance of twoequalresistors connected in parallel is equal to
A. One half the resistance of one resistor **B.** Twice the resistance of one resistor
C. Four times the resistance of one resistor **D.** One fourth the resistance of one resistor

3.3 Materials which can store electrical energy are called
A. Magnetic materials **B.** Semiconductors
C. Dielectric materials **D.** Paper conductors

3.4 The dielectric constant of air is practically taken as
A. More than unity **B.** Unity **C.** Less than unity **D.** Zero

3.5 A $1\mu F$ capacitor is charged using a constant current of $10\mu A$ for $20s$. What is the energy finally stored by the capacitor **A.** $2\times10^3 J$ **B.** $2\times10^2 J$ **C.** $4\times10^3 J$ **D.** $4\times10^{-1} J$

3.6 A $1000\mu F$ capacitor, initially uncharged is charged by a steady currentof $50A$. How long will it take for the potential difference across the capacitors to reach $2.5V$.

A. $205s$ B. $505s$ C. $100s$ D. $400s$

3.7 What is the capacitance C of a capacitor charged at a potential difference of $60V$ whose quantity of charge is $3\times10^{-4}C$? A. $5\mu V$ B. $5\mu F$ C. $1.8\times10^{-2}V$ D. $1.8\times10^{-2}F$

3.8 Three capacitors of capacitances $3\mu F$, $4\mu F$ and $6\mu F$ are connected in series. What is their equivalent capacitance? A. $4/3\,\mu F$ B. $3/4\,\mu F$ C. $3\mu F$ D. $13\mu F$

3.9 The capacitor stores energy and this energy is equal to the work done in

A. charging the capacitor B. charging the battery

C. storing the charge D. lifting a coulomb of charge

3.10 What is the charge on the $1.5\mu F$ Capacitor?

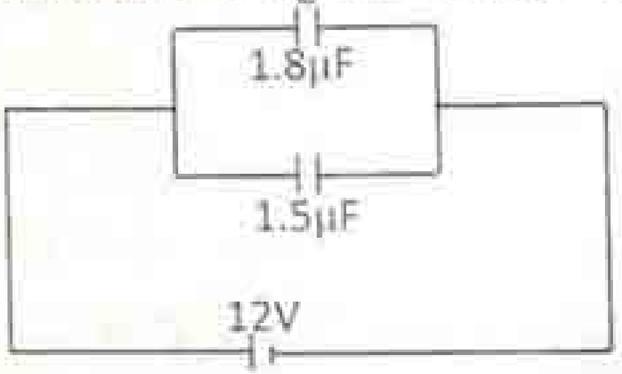


Fig. 3.14: Problem 3.10

A. $0.87\times10^{-5}C$ B. $2.16\times10^{-5}C$ C. $3.96\times10^{-5}C$ D. $1.80\times10^{-5}C$

3.11 The dimensions of the plates of a parallel-plate capacitor are $8cm$, and they are separated by a distance of $2mm$. Calculate the capacitance if (i) air is between the plates, and (ii) glass of dielectric constant 5 fills the space between the plates.

3.12 To achieve very large capacitance, what variables must be considered in designing the capacitor?

3.13 A capacitor having a capacitance of $2\mu F$ is charged to a difference of potential of $100V$. Find the charge on this capacitor?

3.14 Define capacitance and obtain an expression for the capacitance of a parallel plate capacitor?

3.15 The potential difference between the plates of a capacitor is $1000V$ and each plate carries of $250\,\mu C$. Calculate the capacitance and energy stored in the capacitor?

3.16 Deduce expressions for the combined capacitance of two capacitors (a) connected in series (b) connected in parallel.

3.17 Two capacitors of $6\mu F$ are connected in series and then connected to an external source of voltage $1000V$. Find (i) the total capacitance (ii) the potential drop across each capacitor (iii) the charge on each capacitor?

3.18 Describe what happens to a parallel-plate capacitor when (a) an insulating sheet and (b) a metal slab is inserted between the two plates.

3.19 Find the equivalent capacitance of the arrangement in Figure 3.15?

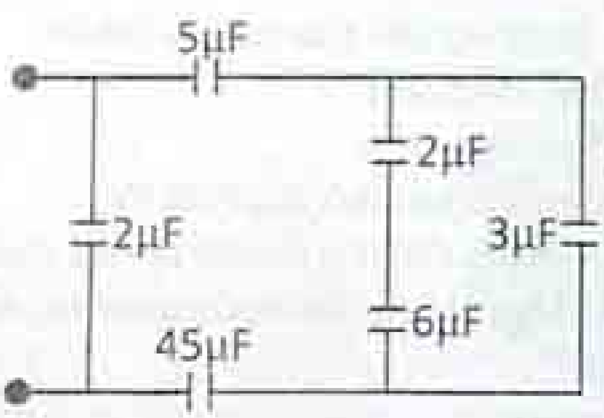


Fig. 3.15: Problem 3.19

3.20 Two parallel wires are suspended in a vacuum. When the potential difference between the two wires is $48V$, each wires has a charge of $72\times10^{-6}C$ (the two charges are of opposite sign). Calculate the capacitance of the parallel wires system?

CHAPTER 4
CURRENT ELECTRICITY

4.0 Introduction

We devoted the last three chapters to the study of stationary charges. In this chapter, we shall study charges in motion; the flow of charges is called the electric or just simply current. The flow of charges is normally within a fixed conducting wire such as copper wire. The conducting wire is composed of atoms sitting at lattice sites. The charges moving through the conducting wire continuously collide with these lattice atoms thereby impeding their motion.

Figure 4.1(a) shows charges in motion inside a conducting wire; because of the collisions, their motion is highly random and not directional. Therefore, to sustain the flow of charges from one end of the wire to the other, a force must be applied to the charges. This electric force that propels the charges through the conductor is as a result of the electric field inside the conductor. Let us recall here that in the chapter on electrostatics, we showed that electric field does not exist inside a conductor except when the charges are in motion. Figure 4.1(b) shows the motion of charges in a conductor under the influence of electric field. This means that there must be a potential difference between both ends of the wire.

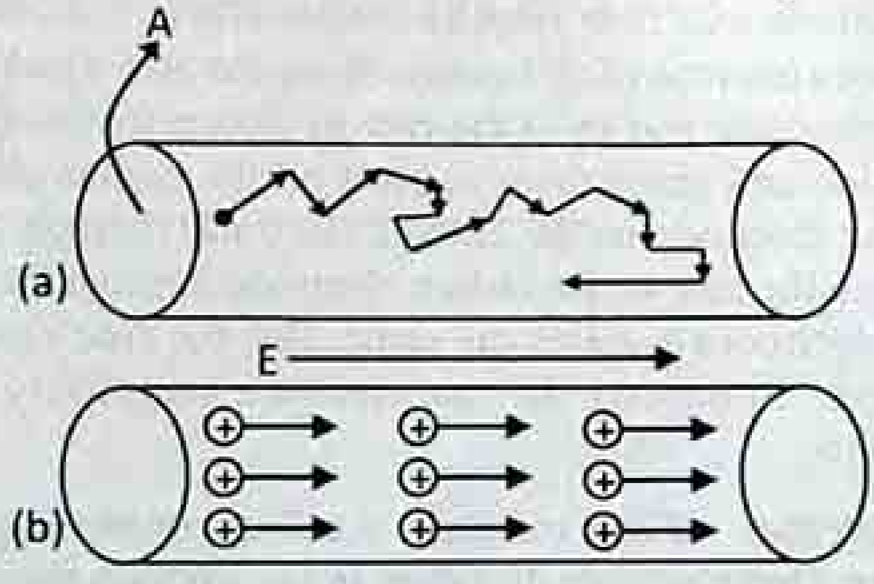


Fig. 4.1: Charged particles in motion within conducting wire

It is therefore clear that we need an energy source to set charges in motion and to maintain this motion. Batteries, dry cells, photocells, fuel cells and electric generators are few examples of sources of energy.

4.1 The Electric Battery

The electric battery, which was discovered by Volta in 1800, produces electricity by transforming chemical energy into electrical energy. Detailed description of the workings of a battery and other electric cells is beyond the scope of this book. But such descriptions can be found in Chemistry textbooks (any interested student may look them up). We shall just describe how a simple chemical battery works. Figure 4.2 shows a simple chemical battery.

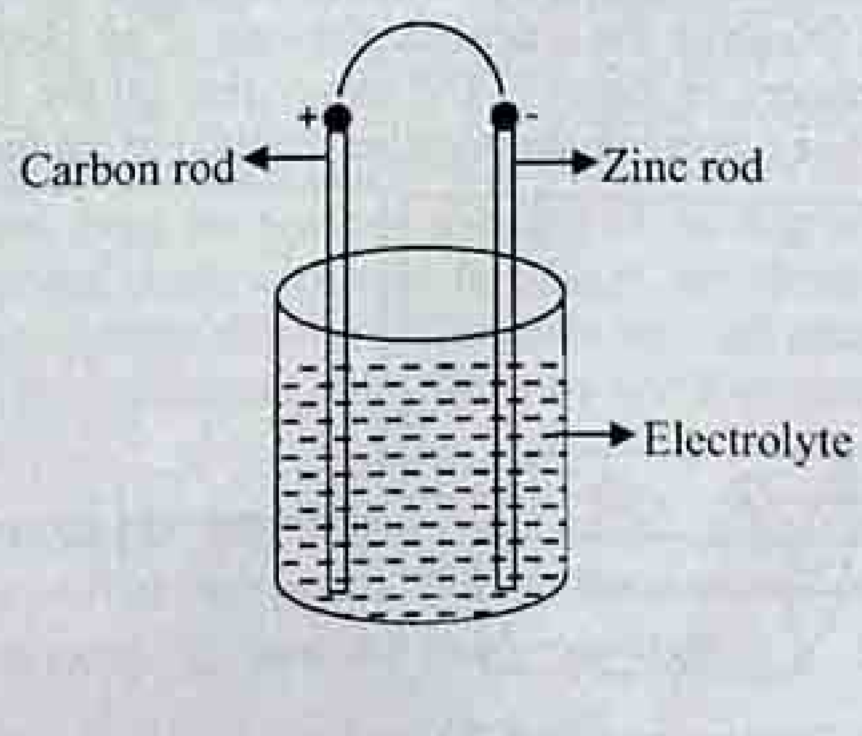


Fig. 4.2: Simple cell

The rods known as the electrodes are made of carbon and zinc. The electrodes are usually one of dissimilar metals. The metals are immersed in dilute sulphuric acid (H_2SO_4). The solution is known as the electrolyte and that part of the electrode remaining outside the solution is called the terminal.

Many substances such as salts, acids and bases ionise when they dissolve in water. This means that the molecules of the substance separate into pieces which do not share electrons evenly. Thus,

$$H_2SO_4 \rightarrow H^+ + H^+ + SO_4^{2-}$$

These ions remain in the electrolyte. The zinc electrode dissolves into the solution and for every zinc ion that enters the solution, two electrons are left behind. Thus,

$$Zn^{2+} \rightarrow Zn^{++} + e^- + e^-$$

To maintain the charge neutrality of the electrolyte, the hydrogen ions through some chemical reactions pull electrons from the carbon electrode to form neutral hydrogen atoms which are not soluble in the solution. The hydrogen atoms bubble off the solution as gas. Thus, the carbon electrode becomes positively charged and the zinc electrode is negatively charged. The positive electrode is called the anode while the negative electrode is the cathode. Since the electrodes are oppositely charged, there is a potential difference between the two terminals. If the two terminals are connected with a conducting wire, electrons will flow from zinc electrode to the carbon electrode. As electrons leave the zinc electrode, more zinc ions (Zn^{++}) escape from the zinc electrode into the solution. Since the electrons from the zinc electrode end up at the carbon electrode, this electrode tends to attract the positive ions from the solution and cause them to plate on the carbon electrode. This means that as electrons flow from the zinc electrode to the carbon electrode through the external wire, an equal positive charge is carried from the zinc to the carbon electrode through the solution by ions. However, the net charge on the two electrodes remains the same. But the zinc electrode loses ions which are gained by the carbon electrode. In the end, the zinc electrode completely dissolves; it is used up and the battery voltage will be zero.

It is clear from the foregoing that we can construct a cell by immersing two dissimilar conducting materials into any electrolyte. One electrode dissolves in the solution thereby losing chemical energy. The energy is converted into electrical energy by the external current flowing between the electrodes. Then more zinc can be dissolved and even the carbon electrode also suffers disintegration. After a time, one or the electrode is completely used up and no charge can flow anymore.

The potential difference between the terminals of a battery when no charge is allowed to flow in an external circuit is known as the electromotive force or the e.m.f of the battery. When charge flows in an external circuit, the potential difference between the terminals is lower than the e.m.f of the battery. This is due to the internal resistance of the battery; the older the battery, the greater the internal resistance. We shall discuss this later.

4.2 Electric Current

We say that we have an electric circuit when a conducting wire is connected to the terminals of a battery. Figure 4.3(a) and (b) show the schematic drawing of an electric circuit. We have used the symbol to represent the battery or the cell. The longer line on the symbol represents the positive terminal while the shorter line represents the negative terminal of the battery.

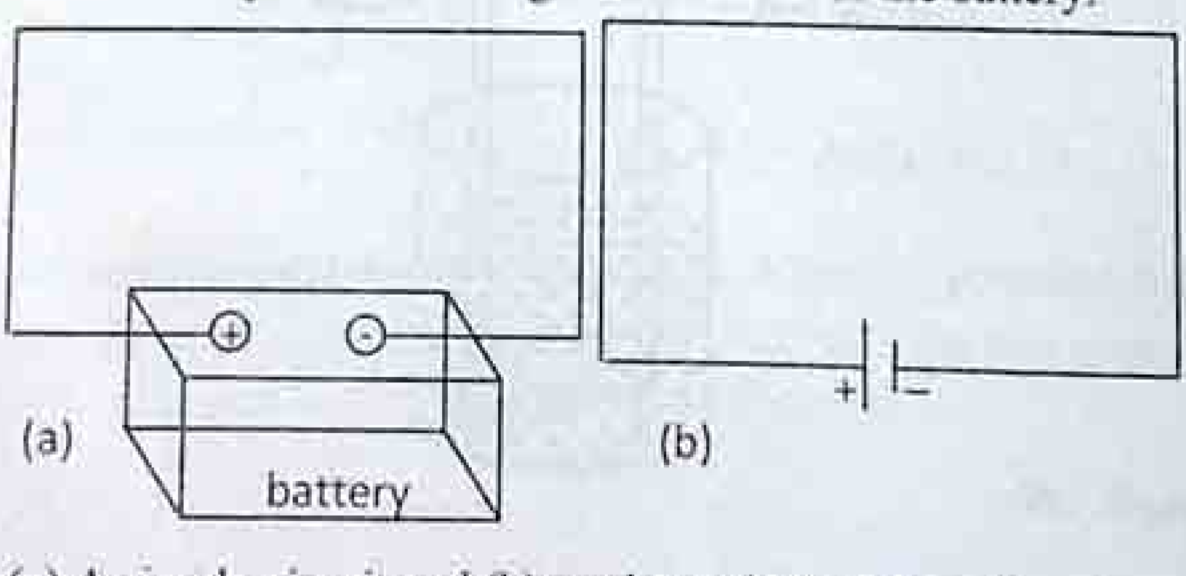


Fig. 4.3: (a) A simple circuit and (b) a schematic drawing of an electric circuit

For the purpose of defining the electric current, let us consider an enlarged segment of a conducting wire connecting the two terminals of the battery. This is shown in Figure 4.4; the wire has a cross-sectional area A.

The electric current is defined as the charge (ΔQ) passing through a given cross-sectional area, A of the wire per unit time. That is

$$I = \frac{\Delta Q}{\Delta t} \tag{4.1}$$

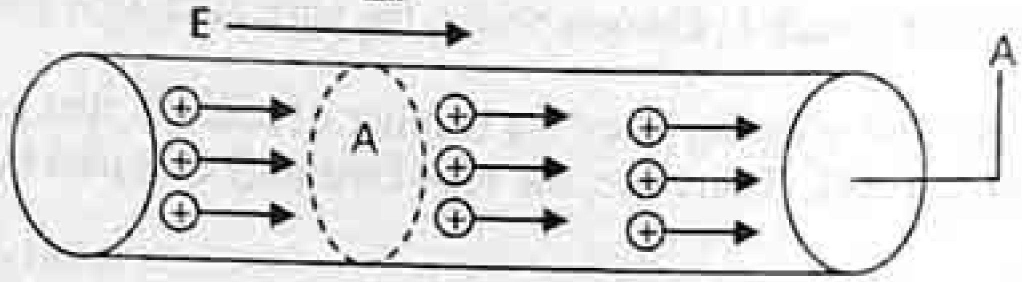


Fig. 4.4: An enlarged segment of a conducting wire with cross-sectional area A

where ΔQ is the quantity of charge passing through area A in time Δt. The symbol I used for current is Coulomb per second, which is called ampere (A). As $\Delta t \to 0$, Equation 4.1 may be written in a differential form as $I = \frac{dQ}{dt}$.

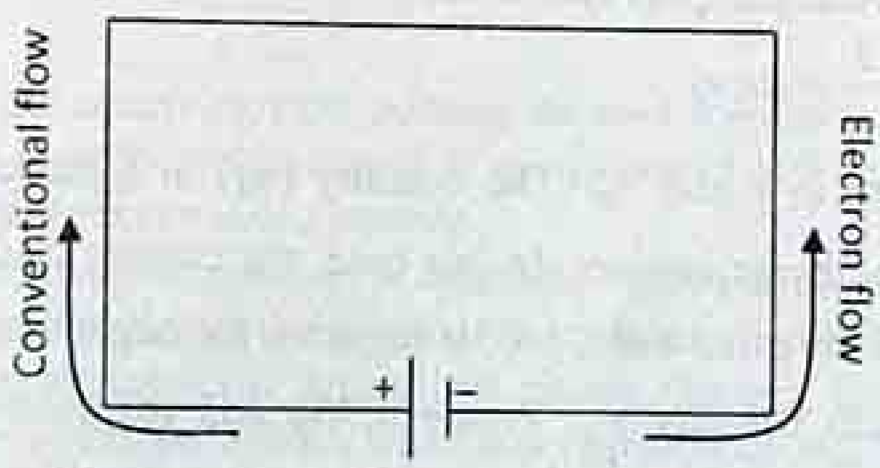


Fig. 4.5: Conventional current flow

The current, I is clearly a scalar quantity but it does have a sense (not strictly a direction, if it has direction then it would be a vector). As we have mentioned earlier, the conducting wire is made up of atoms. The atom is made up of a positively charged nucleus which is surrounded by a cloud of electrons. These electrons are free to move when subjected to an electric field. Thus, when a conducting wire is connected to the terminals of a battery as shown in Figure 4.5, the free electrons in the wire are attracted to the positive terminal of the battery. Electrons as we discussed above, also leave the negative terminal of the battery and enter the wire. Thus, there is a continuous flow of electrons through the wire to the positive terminal. The positively charged nucleus does not move within the conductor, therefore the current flow in a conductor is mainly due to the flow of negatively charged electrons. The electron flow is thus from the negative terminal through the wire to the positive terminal of the battery. However, by convention, the current in the wire is assumed to be due to the flow of positive charges from the positive terminal to the negative terminal of the battery. Actually, for all practical purposes, positive charges flowing in one direction is exactly equivalent to negative charges flowing in the opposite direction. Therefore, we shall adopt the historical convention of positive charges flowing in the conductor. Thus, when we speak of current flowing in a circuit we mean the conventional current.

Let us now find a relation between the current in a wire and the properties of the charges within the wire. The wire shown in Figure 4.6 has a length L and a cross-sectional area A. Suppose, the charges within the wire have a drift velocity v_d and the charge on each is q. The time, t it takes for the moving charges to go from one end of the wire to the other is just L/v_d. The wire of length L has a volume of AL.

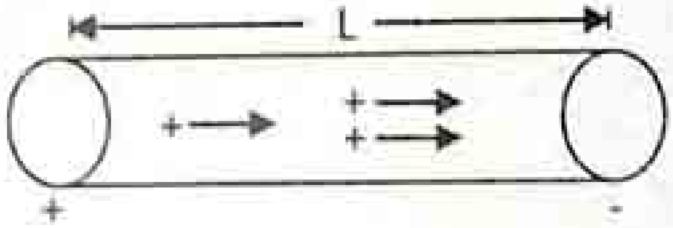


Fig. 4.6: A piece of wire of length L; charges within the wire drift with velocity, v_d.

Suppose there are n charges per unit volume. Then the number of charges that pass through a cross-section of the wire in time t is just nAL. Therefore, the total charge Q that passes through area A in a time is

$$Q = qnAL = qnAv_d t \tag{4.2}$$

From equation 4.1, we obtain the expression for current as

$$I = \frac{Q}{t} = qnAv_d \tag{4.3}$$

Equation 4.3 shows that the current in the wire depends upon the cross-section of the wire used. It is more convenient to talk of *current density, j*. It is defined as *the electric current per unit cross-sectional area at any point in space*. Assuming the current flow to be uniform, we have:

$$j = \frac{I}{A} = qnv_d \tag{4.4}$$

The current density is proportional to q, n and v_d. The quantity (nq) in Equation 4.4 has units of C/m^3 and it is positive for positive charge carriers. In this case, the current density is in the same direction as the drift velocity v_d. On the other hand, (nq) is negative for negative charge carriers, the current density, j is opposite the drift velocity v_d. Figure 4.7(a) and (b) shows the positive and negative charge carriers, the direction of drift velocity and the current densities. The sense of the conventional current is also shown.

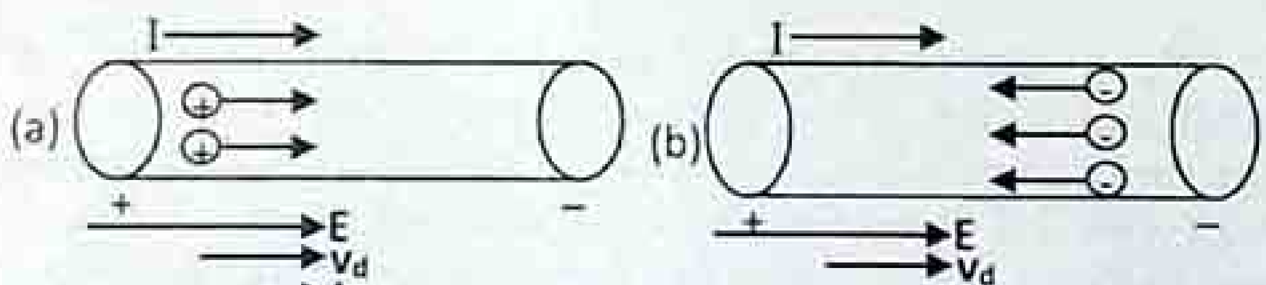


Fig. 4.7: (a) Positive charge carriers drifting in the direction of the applied electric field while (b) negative charge carriers drift opposite the applied force.

Example 4.1

A 0.40 mm diameter copper wire carries a current of 3 μA. Find (a) the current density and (b) the electron drift density.

Solution

(a) The cross-sectional area A of the wire is

$$A = \frac{\pi d^2}{4} = \frac{(3.14)(0.4\times10^{-3}m)^2}{4} = 12.56\times10^{-8}m^2$$

$$\therefore \quad j = \frac{I}{A} = \frac{3\times10^{-6}A}{12.56\times10^{-8}m^2} = 23.89Am^{-2}$$

(b) Since there is one free electron per atom, the density of free electrons, n, is the same as the density of Cu atoms. The atomic mass of Cu is 63.5 u, therefore, 63.5 gram of Cu contains 6.02×10^{23} free electrons. The density of copper from the Handbook of Physics and Chemistry is $\rho=8.98\times10^{3}\,\text{kg/m}^3$. Therefore,

$$n=\left(\frac{6.02\times10^{23}\,electrons}{63.5\times10^{-3}\,kg}\right)\left(8.89\times10^{3}\,kg/m^{3}\right)=8.43\times10^{28}\,electrons/m^{3}$$

From equation 4.4, the drift velocity is

$$v_d=\frac{j}{nq}=\frac{23.89\,Am^{-2}}{(8.43\times10^{28}\,m^{-3})(1.6\times10^{-19}\,C)}=1.77\times10^{9}\,ms^{-1}$$

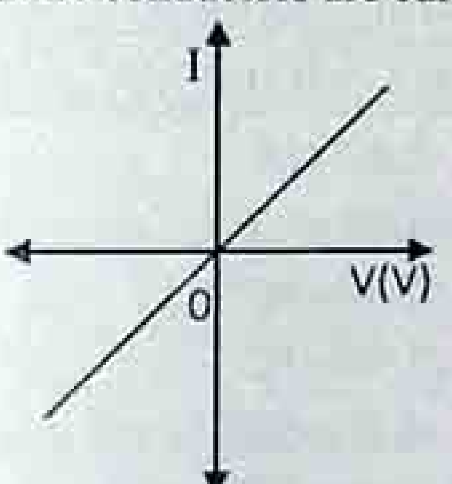
4.3 Resistance and Resistors

We have seen from the previous section that for electric current to flow in a circuit, there must be a potential difference between the two ends of the wire. This potential difference is usually provided by a battery. Suppose the same potential difference is maintained between the ends of geometrically identical rods of copper and iron, different current are obtained. The characteristics of the conductor that determines how much current obtained is known as the *resistance*, R. George Ohm was the first to establish experimentally the relationship between the current I, in a conducting wire and the potential difference, V applied.

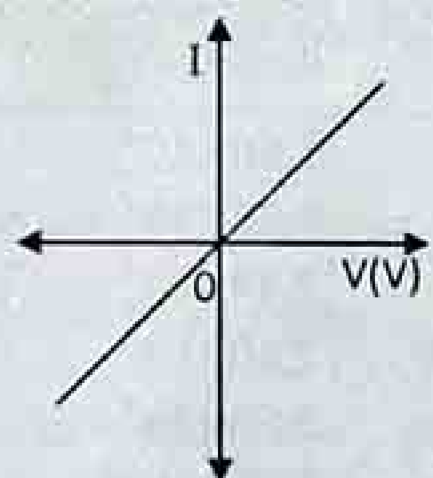
4.4 Ohm's Law

Ohm's law states that the steady current passing through a metallic conductor which is not a site of electromotive force, is directly proportional to the potential difference across it, if the temperature and other physical conditions remains constant.

That is, $I\propto V$ if temperature = constant. A conductor which obeys Ohm's law is known as an ohmic conductor. For an ohmic conductor, the graph of the current, I against the potential difference V, i.e its I-V characteristics, is a straight line graph that passes through the origin (see Figure 4.8). Some examples of ohmic conductors are eureka wires and carbon resistors.



(a) I-V characteristics of eureka wires



(b) I-V characteristics of carbon resistors

Fig. 4.8: Characteristic Plot of Ohmic conductors.

For non-ohmic conductors such as thermionic diodes, junction diodes, photo cells and dilute sulphuric acid, their I-V characteristics are not straight lines passing through the origin.

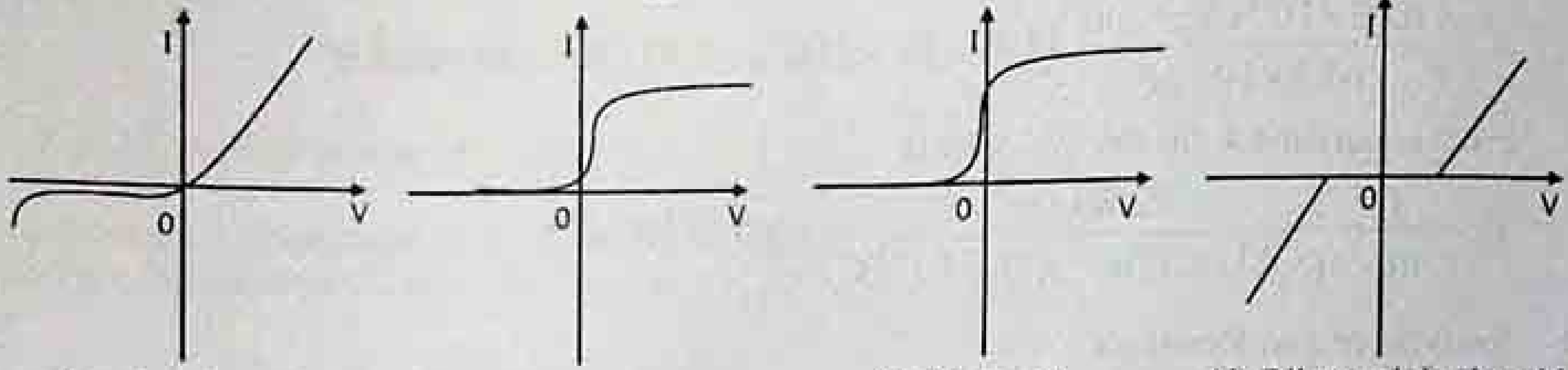
Note that the expression, $$\frac{V}{I}=R \qquad (4.5)$$

is not a representation of Ohm's law but a representation of the resistance R. A non-ohmic conductor also has a resistance which can be found using the equation 4.5 but its value is not a constant.

Ohms's law is represented by the expression $\frac{V}{I}=\text{constant}$ if the temperature is constant. This implies that at constant temperature, the resistance of an ohmic conductor is independent of the current I or the potential difference V. It is also found experimentally that the resistance, R of a wire is proportional to the length of the wire and inversely proportional to the cross-sectional area A.

That is $$R \propto \frac{L}{A}$$

or $$R = \rho \frac{L}{A} \qquad (4.6)$$



(a) Junction diode (b) Thermionic diode (c) Photo cell (d) Dilute sulphuric acid

Fig. 4.9: Characteristic Plot of non-Ohmic conductors

where ρ is the constant of proportionality, which is known as the *resistivity* of the material.

Equation 4.6 makes sense because if the length of the wire is increased, one would expect the resistance to increase. On the other hand, a thin wire (with small cross-sectional area) provides a greater resistance than a thicker wire (with large cross-sectional area) since a thicker wire provides more area for the charges to move through. The resistivity, ρ depends only on the material and the temperature. From equation 4.6, we obtain

$$\rho = R\frac{A}{L} \qquad (4.7)$$

The unit of resistivity ρ is ohm-metre (Ωm). The values of ρ for common materials are listed in Table 4.1.

Table 4.1: Resistivities and temperature coefficients of resistivity for commonly used materials

Material	Resistivity (Ωm)	Temperature Coefficient $[(C^{\circ})^{-1}]$
Silver	1.60×10^{-8}	3.80×10^{-3}
Copper	1.70×10^{-8}	3.90×10^{-3}
Gold	2.44×10^{-8}	3.40×10^{-3}
Aluminium	2.80×10^{-8}	3.90×10^{-3}
Tungsten	5.60×10^{-8}	4.50×10^{-3}
Iron	10.00×10^{-8}	5.00×10^{-3}
Platinum	11.00×10^{-8}	3.92×10^{-3}
Lead	22.00×10^{-8}	3.90×10^{-3}
Carbon	3.50×10^{-8}	-0.50×10^{-3}
Nichrome	150.00×10^{-8}	0.40×10^{-3}
Steel	40.00×10^{-8}	1.50×10^{-3}
Manganin	44.00×10^{-8}	0.01×10^{-3}
Mercury	96.00×10^{-8}	0.90×10^{-3}
Germanium	0.46	-48.00×10^{-3}
Silicon	640.00	-74.00×10^{-3}
Glass	$10^{10}-10^{14}$	
Hard Rubber	$10^{13}-10^{16}$	

It is clear from Table 4.1 that silver has the lowest resistivity ρ and is thus the best conductor. The reciprocal of the resistivity is called *conductivity*, σ.

$$\sigma = \frac{1}{\rho} \tag{4.8}$$

The conductivity σ has units of $(\Omega m)^{-1}$.

Let us consider a piece of wire of length L, with a cross-sectional area A and a potential difference V between the two ends. The wire is shown in Figure 4.10. The electric field inside the wire E is equal to V/L since the field is uniform. From equation 4.5, we can express the potential difference V in terms of the resistance of the wire. That is $V = IR$.

Using equation 4.6, the above expression becomes

$$V = I\rho\frac{L}{A} = \frac{I}{A}\rho L = j\rho L$$

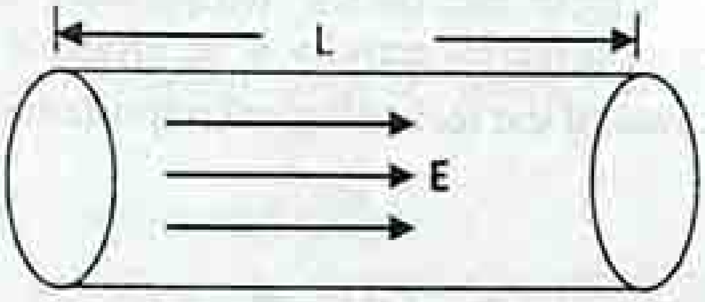


Fig. 4.10: Potential difference between the two ends is $V = EL$

Simplifying further, we obtain

$$j = \sigma E \tag{4.9}$$

Equation 4.9 is another form of Ohm's law. The resistivity's of the materials that obey Ohm's law are independent of the magnitude E of the applied electric field.

Example 4.2

A $20.0m$ length wire $1.50mm$ in diameter has a resistance of 2.5Ω. What is the resistance of a $35.0m$ length of wire $3.00mm$ in diameter made of the same material?

Solution

We need to find the resistivity of the material. From equation 4.7, we have

$$\rho = R_1\frac{A_1}{L_1} = \frac{(2.5\Omega)(3.14)(0.75\times10^{-3})^2 m^2}{20m} = 0.22\times10^{-6}\Omega m$$

Now the resistance of the wire of $3.0mm$ in diameter and $35m$ long is

$$R_L = \rho\frac{L_1}{A_2} = \frac{(0.22\times10^{-6}\Omega m)(35.0m)}{(3.14)(1.5\times10^{-3})^2} = 1.10\Omega$$

Example 4.3

A wire with a resistance of 6.0Ω is drawn out through a die so that its new length is three times its original length. Find the resistance of the longer wire assuming that the resistivity of the material are unchanged during the drawing process.

Solution

The resistivity of the material is

$$\rho = (6)\frac{A_1}{L_1}\Omega m$$

Therefore, $R_2 = (6)\left(\frac{A_1}{L_1}\right)\left(\frac{L_2}{A_2}\right)\Omega$

Since the length L_2 is stretched to $3L_1$, the new cross-sectional area is $1/3A_1$, then we have

$$R_2 = (6)\left(\frac{A_1}{L_1}\right)\left(\frac{L_2}{A_2}\right) = 6\left(\frac{A_1}{A_2}\right)\left(\frac{L_2}{L_1}\right) = 6(3)(3) = 54\Omega$$

4.5 Superconductivity

There are some materials for which the resistivity is zero below a certain temperature, called the critical temperature T_c. This phenomenon is called *superconductivity* and was discovered by the Dutch Physicists, H. Onnes in 1911 shortly after he successfully liquefied helium. Figure 4.11 shows a plot of the resistance of mercury against temperature, it gives a critical temperature of 4.2K for mercury. In the space of about 0.05K, the resistance drops abruptly to an immeasurably low value.
Many metallic compounds are also superconductors. For example, the superconducting alloy NB_3Ge discovered in 1973 has a critical temperature of 23.2K. This was the highest known critical temperature until 1986 when it was discovered that certain ceramic oxides become superconducting at much higher temperatures. For example the critical temperature for yttrium-barium-copper oxide ($YBa_2Cu_3O_2$) is about 92K. A lot of research activity is in progress towards obtaining easily usable materials with still higher T_c because of the technological importance of the superconductors.

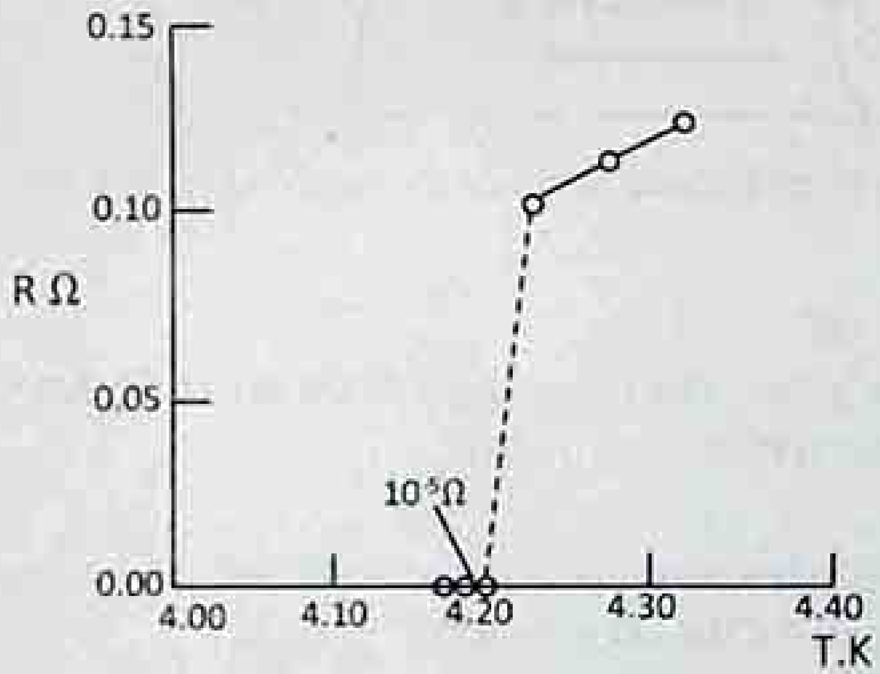


Fig. 4.11: Resistivity of mercury versus temperature showing sudden decrease at critical temperature T = 4.2K

This importance stems from the fact that materials in the superconducting state have zero resistance, therefore currents once established in closed superconducting circuits persist for weeks without decreasing, even though there is no power supply in such circuits. That means no energy loss in such circuits. Some applications of superconductors include:

(1) They are used in the manufacture of high field superconducting magnets which produce no heat. It is found cost effective despite the use of expensive liquid helium which boils at 4.2K for cooling the superconductor in the case of Nb_3Ge. With yttrium-barium-copper ceramic oxide, relatively inexpensive liquid nitrogen which boils at 77K can be used for the refrigeration. However, one disadvantage of the ceramics is that they are brittle and difficult to use.
(2) They are used in cables for distributing power without energy loss except energy required to refrigerate the line.
(3) There are also applications of superconductors in very sensitive measuring instruments and in computer technology.

The phenomenon of superconductivity cannot be understood in terms of classical physics. Instead, quantum mechanics is needed to explain the phenomenon, but that is beyond the scope of this book.

4.6 Variation of Resistivity, ρ with temperature (T)

The resistivity of all metallic conductors increases with increasing temperature (Figure 4.12(a)).

Over a narrow range of temperatures near the ice point, approximately the empirical law can represent the resistivity of pure metals.

$$\rho_T = \rho_0\left\{1 + \alpha(T - T_0) + \beta(T - T_0)^2\right\} \tag{4.10}$$

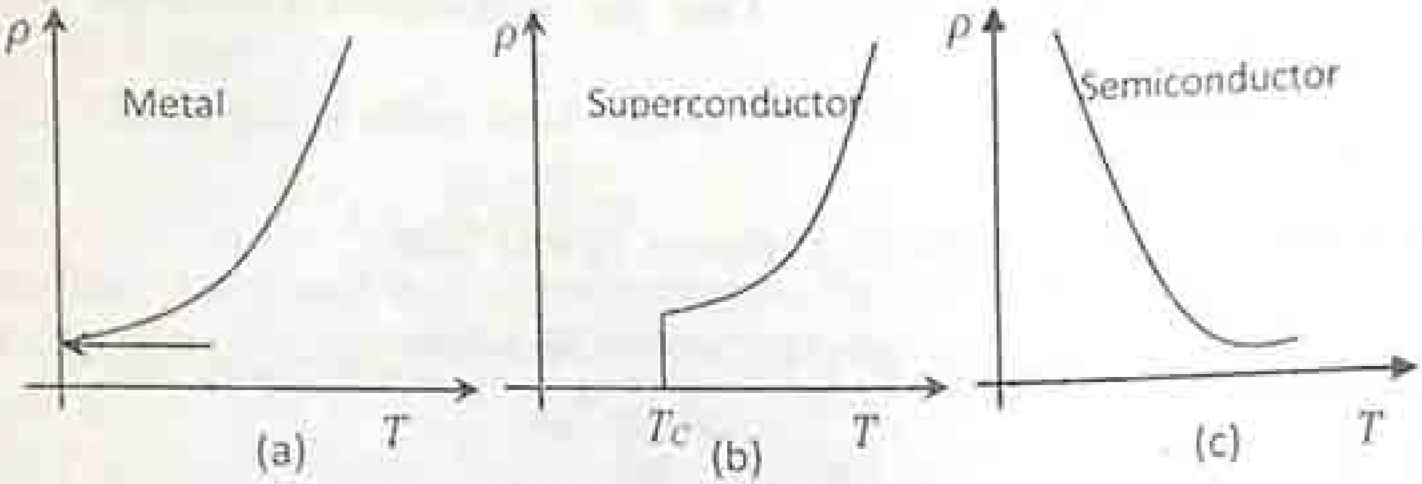


Fig. 4.12: Variation of ρ with T for various materials

where α and β are constants, and ρ_T and ρ_0 are the resistivity of the metal at temperatures T and T_0. The constant β is very small compared with α, and for moderate temperature ranges, the law of variation may be taken as

$$\rho_T = \rho_0\left\{1 + \alpha(T - T_0)\right\} \tag{4.11}$$

The quantity α is a constant known as the *temperature coefficient of resistivity* appropriate to the particular range of temperature in question.

The resistivity of carbon decreases with increasing temperature and its α is negative. For pure metal, the value of α is about $4.0\times10^{-3}K^{-1}$. While for some alloys such as manganin, α could be as low as $10^{-6}K^{-1}$. The low value of α for alloys makes them suitable for constructing standard resistors.

A number of materials (e.g. mercury) have been found to exhibit the property of superconductivity. As the temperature is decreased, the resistivity at first decreases regularly, like that of a metal. At the so-called critical temperature, usually in the range 0.1 to 20K, the resistivity suddenly drops to zero (Figure 4.12(b)). A current once established in a superconducting ring will continue of itself, apparently indefinitely, without the presence of any driving field.

The resistivity of a semiconductor decreases rapidly with increasing temperature (Figure 4.12(c)). This characteristics makes semiconductor suitable for making temperature sensitive resistors called thermistor, which are used as safeguards against current fluctuations in electric circuits

From the first equation above,

$$\alpha = \frac{\rho_T - \rho_0}{\rho_0(T - T_0)} \tag{4.12}$$

For a particular temperature T,

$$\alpha = \frac{1}{\rho}\frac{d\rho}{dT} \tag{4.13}$$

Thus, α may be defined as *the fractional increase in resistivity per degree increase in temperature.* The unit of α is K^{-1} or $°C^{-1}$. For a conductor of fixed length l and cross-sectional area A, we may use the relation $R = \rho l / A$ and rewrite equation 4.14 as

$$R = R_0\left(1 + \alpha(T - T_0)\right) \tag{4.14}$$

For a given specimen of wire, this change in ρ or R with temperature may be used to measure the temperature.

Example 4.4

A silver wire has a length of $l = 10m$ and cross-sectional area $A = 6mm^2$ at 20°C. If the conductivity of silver is $5.8\times10^7\,\Omega m^{-1}$, calculate

(a) The resistivity of the wire
(b) The resistance at 20°C
(c) If a p.d of 20V is maintained over the length of the wire, calculate the current flowing through the wire.
(d) Find the resistivity of the wire at 100°C. Take the temperature coefficient of resistance $\alpha = 3.9\times10^{-3}{}^{\circ}C^{-1}$.

Solution

(a) Resistivity, ρ is given by $\rho = \frac{1}{\sigma} = (5.8\times10^{7})^{-1}\Omega m = 1.72\times10^{-8}\Omega m$

(b) The resistance of the wire at 20°C is $R = \frac{\rho l}{A} = \frac{1.72\times10^{-8}\Omega m\times10m}{6\times10^{-6}m^{2}} = 2.87\times10^{-2}\Omega$

(c) The current flowing through the wire is $I = \frac{V}{R} = \frac{20V}{2.87\times10^{-2}\Omega} = 696.86A$

(d) The resistivity of the wire at 100°C is given by $\rho_{100} = \rho_{20}\{1+\alpha(T_{100}-T_{20})\}$

i.e. $\rho_{100} = 1.72\times10^{-8}\Omega m\{1+3.9\times10^{-3}{}^{\circ}C^{-1}(100-20)\}$

$= 1.72\times10^{-8}\{1+3.9\times10^{-3}(80)\} = 2.26\times10^{-8}\Omega m$

Example 4.5
The temperature coefficient of a tungsten wire at room temperature (20°C) is 1.51Ω. What will be its resistance at 0°C, 100°C and 1000°C?

Solution
From equation (4.14), $R = R_0\{1+\alpha(T-T_0)\}$,
The values of $R_0 = 1.50\Omega$, $T_0 = 20°C$, $T = 0°C$, we then get
$R = 1.50\Omega\{1+(0.0045\,{}^{\circ}C^{-1})(0-20°C)\} = 1.365\Omega$
At $T = 100°C$: $1.50\Omega\{1+(0.0045\,{}^{\circ}C^{-1})(100-20°C)\} = 2.04\Omega$.
At $T = 1000°C$: $1.50\Omega\{1+(0.0045\,{}^{\circ}C^{-1})(1000-20°C)\} = 8.12\Omega$.

4.7 Electric Power
Suppose the battery shown in Figure 4.13 is connected to an electrical device. The device may be any of the following: electrical motor, electric heater, stove, toaster, pure resistor, a storage battery. The battery sets up an electric current I in the circuit. Point A is at a higher potential than point B, because point A (in Figure 4.13) of the device is connected to the positive terminal of the battery and point B is connected to the negative terminal of the battery.

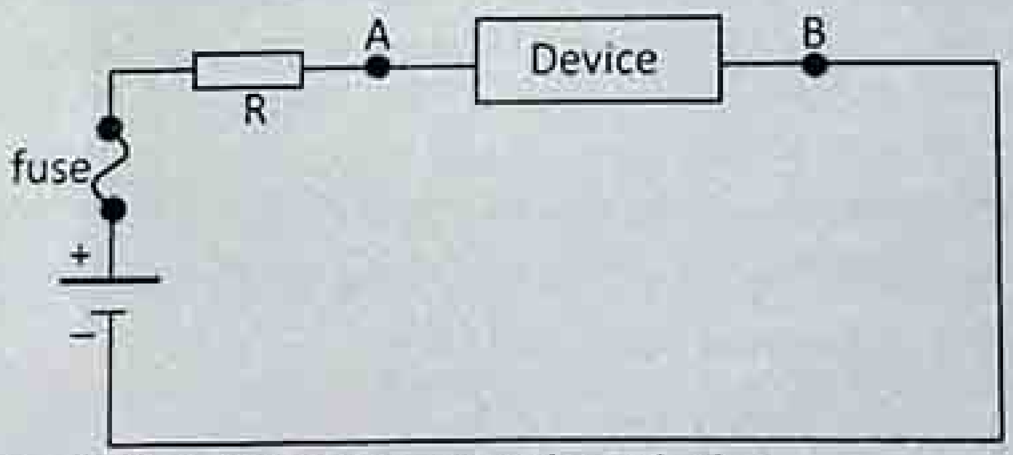


Fig. 4.13: A battery connected to adevice

The potential difference between A and B is just V. Now if a charge dq moves from A to B because it is moving from a higher potential to a lower potential, its potential energy decreases by Vdq. By conservation of energy principle, this reduction in potential energy of the charge will apply in some form of energy. The particular form of energy depends upon the type of device between A and B. In a time dt, the energy transferred to the device is then $dU = dqV$.

But, $I = \frac{dq}{dt}$ or $dq = Idt$

Therefore, $dU = IdtV$

Since power, P is the rate at which work is being done, we have

$$P = \frac{dU}{dt} = IV \tag{4.15}$$

If the device between A and B is a pure resistance, the energy appears as heat in the resistor. Using Ohm's law (equation 4.5), equation 4.15 can be expressed in the following ways:

$$P = IV = I^2R \text{ or } \frac{V^2}{R}$$

The unit of electric power is watt ($1W = 1J/s$)

Example 4.6

An X-ray tube takes a current of $7.0MA$ and operates at a potential difference of $80kV$. What power is dissipated?

Solution

Using equation 4.15 above, we have $P = IV = (7\times10^{-3}A)(80\times10^{3}V) = 560W$

If the device in Figure 4.12 is a pure resistance, the electrical energy is converted to thermal energy. This thermal energy heats up the device. The device and the current carrying wire may become hot enough to start fire. In order to avoid electrical fires or in factories, thick wires, heavy enough for expected load should be utilised in electrical wiring. When a wire carries more current than it can carry, it is said to be overloaded and the wire actually becomes hot. However, a thicker wire (large area) used under the same circumstances does not become overloaded. To prevent overloading, fuses are usually installed in circuits. A typical fuse in a circuit is shown in Figure 4.13. When the current, in the circuit exceeds some particular value, the wire in the fuse burns up and we have an open circuit and the current does not flow anymore, thus protecting the electrical device from the damage due to overheating.

A $10A$ fuse, for example, burns up when the current passing through it exceeds $10A$. If the fuse in a circuit repeatedly burns up, either there is a fault in the circuit such as a short or there are too many devices drawing current in the circuit. Since each household in Nigeria receives $220V$ from EEDC, the current requirements of each device can be determined.

Thus, a light bulb rated $60W$ at $200V$ draws a current of $I = P/V = 60/220 = 0.27A$. Therefore, one can determine the total current requirements of the appliances in a particular household. If the total current is higher than the fuse rating, it is then not advisable to turn on all the appliances at the same time.

EEDC charges its customers for the energy, not power, their devices consumed. Since power is the rate energy transformed, the total energy used by any appliance is simply its power consumption multiplied by the time it is on. If the power is in watts and the time is in seconds, then the energy is in Joules or watts-sec. However, EEDC usually specify energy with a much larger unit, the kilowatt-hour, $1kWh = (10^3W)(3600\text{ sec}) = 3.6\times10^6 J$.

Example 4.7

The wiring a house must be thick enough so it does not become so hot as to start a fire. What diameter must a copper wire be if it is to carry a maximum current of $40A$ and produce no more than $1.8W$ of heat per metre of length?

Solution

From Table 4.1, we find, for copper, that the resistivity $\rho = 1.67 \times 10^{-8}\,\Omega m$

Then, $P = I^2 R$ but $R = \rho \frac{L}{A} = \rho \frac{L}{\pi (d/2)^2}$

where d is the diameter of the wire. $\therefore$ $\quad P = I^2 \rho \frac{4L}{\pi d^2}$

$$\text{or } d = \left[\frac{4I^2 \rho L}{\pi P}\right]^{1/2} = \left[\frac{(4)(40.A)^2(1.67 \times 10^{-8}\,\Omega m)(1m)}{(3.14)(1.8W)}\right]^{1/2} = 4.4 \times 10^{-3}\,m$$

Example 4.8

An electric fan draws 2.0A of current from a 220V source. Find (a) the power rating of the fan, (b) electrical resistance R, and (c) the cost of operating the fan during the month of April, if it operates continuously and electric energy costs 8 kobo per kilowatt hour (kWh).

Solution

(a) $P = IV = (2.0A)(220V) = 440W = 0.44kW$

(b) $R = \frac{V}{I} = \frac{220V}{2.0A} = 110\Omega$

(c) Now let us determine the total energy consumed during the month of April- 30 days.

$$Energy = Pt = (440W)(30days)\left(\frac{24hours}{1day}\right) = 317kWh$$

Total cost of operation = ₦ (317) (0.08) = ₦ 25.36.

Summary

1. Electric current in a conductor is the charge passing through a cross-sectional area per unit time, $I = dQ/dt$.
2. When an electric field E is established in a conductor, the charge carriers acquire a mean drift speed v_d in the direction of E and the current density j is given by $j = I/A = qnv_d$.
3. The resistance of a wire across which there is a potential drop V is given by $R = V/I$ where I is the current flowing along the wire.
4. Ohms law states that provided the temperature and other physical conditions of the conductor is constant the potential difference between the ends of the conductor is proportional to current passing through it.
5. A given material obeys Ohm's law if its resistance is independent of the applied voltage. Materials which obeys Ohm's law are said to be Ohmic (e.g. metals) while those that do not obey it are non-ohmic (e.g. transistors).
6. The resistance of a wire is proportional to the length and inversely proportional to the cross-sectional area. $R = \rho \frac{L}{A}$
7. The reciprocal of resistivity ρ is known as conductivity, $\sigma = \frac{1}{\rho}$.
8. The power supplied to a segment of a circuit equals the product of the current and the voltage drop across the segment $P = IV$. The unit of power is the watt or joule per second.
9. Superconductors are materials for which the resistivity is zero below the critical temperature T_c. The materials are of technical importance e.g. for use in high field superconducting magnets, cables in power transmission e.t.c. because they are lossless.

10. Conduction in nerve cells is different from conduction in metals. However, ideas in physics are required to understand conduction in nerve cells.

Exercises 4

4.1 Calculate the drift velocity of the free electrons in a copper wire of cross-sectional area $1.0mm^2$ when the current flowing through the wire is $2.0A$. (Number of free electrons in copper is $1\times10^{29}\,m^{-3}$).
A. $1.25\times10^{-4}\,ms^{-1}$ B. $1.50\times10^{-4}\,ms^{-1}$ C. $1.45\times10^{-4}\,ms^{-1}$ D. $1.35\times10^{-4}\,ms^{-1}$

4.2 An electric current $2A$ flows in a heating coil of resistance 50 for 3 minutes, 20 seconds. Determine the heat produced. A. $30.0kJ$ B. $25.0kJ$ C. $40.0\,kJ$ D. $35.0kJ$

4.3 A $0.30mm$ diameter copper wire carries a current of $2A$. Find the current density.
A. $3.60A/m^2$ B. $2.83A/m^2$ C. $4.00A/m^2$ D. $2.50A/m^2$

4.4 Which of the following are non-ohmic conductors?
I. Thermionic diodes II. Photocells III. Dilute sulphuric acid IV. Carbon resistors.
A. I & II only B. II & IV only C. I, II, III & IV only D. I, II & III only.

4.5 Which of the following equations is another form of Ohm's law? (Where the symbols have their usual meaning).
A. $j=\sigma E$ B. $\sigma=\frac{1}{\rho}$ C. $v_d=\frac{j}{nq}$ D. $j=\frac{I}{A}$

4.6 The resistivity depends on the
A. The resistance of Material & temperature B. Pressure & resistance
C. Conductivity & current D. Temperature & voltage

4.7 An X-ray tube takes a current of $5\,mA$ and operates at a potential difference of $50kV$. What power is dissipated?
A. $250W$ B. $500W$ C. $560W$ D. $350W$

4.8 Which of the following are applications of superconductors?
I. Manufacture of high field superconducting magnet.
II. Use in cable for distributing power.
III. Use in very sensitive measuring instruments and its computer technology.
A. I only B. III only C. II & III only D. I, II & III only.

4.9 A beam current in a TV tube is $1.9mA$. The beam cross-section is circular with radius $0.5mm$. (i) How many electrons strike the screen per second? (ii) What is the current density?

4.10 A 10 m length of aluminum wire has a diameter of $1.5mm$. It carries a current of $12A$. Find (i) the current density (ii) the drift velocity (iii) the electric field in the wire. Aluminium has approximately 10^{29} free electrons per m^3.

4.11 Near the earth, the density of protons in the solar wind is $8.7\times10^6\,m^{-3}$ and their speed is $4.7\times10^5\,m/s$. (i) Find the current density of these protons (ii) if the earth's magnetic field did not deflect them, the protons would strike the earth. What total current would the earth receive?

4.12 Suppose that the current through a conductor decreases exponentially with time according to $I(t)=I_o e^{-t/\tau}$ where I is the initial energy (at $t=0$), and τ is the time constant having the dimensions of time. Consider a fixed observation point within the conductor. (i) How much charge passes this point between $t=0$ and $t=\tau$. (ii) How much charge passes this point between $t=0$ and $t=30\tau$.

4.13 A charge $+q$ moves in a circle of radius r with speed v. (i) Express the frequency f with which the charge passes a particular point in terms of r and v. Show that the average current is q and express it in terms of v and r.

4.14 Calculate the resistance at 20°C of $60m$ length of aluminium wire having a cross-sectional area of $0.3mm^2$.

4.15 Suppose that you wish to fabricate a uniform wire out of 1 gram of copper. If the wire is to have a resistance of $R = 0.5\Omega$, and all of the copper is to be used. What will be (i) the length and (ii) the diameter of the wire?

4.16 A potential difference of $6V$ is found to produce a current of $0.14A$ in a $2.6m$ length of conductor having a uniform radius of $0.3cm$ (i) calculate the resistivity of the material (ii) what is the resistance of the conductor?

4.17 A cylindrical tungsten conductor has initial length L_1, and cross-sectional area A. The metal is drawn uniformly to a final $L_2 = 10L_1$ and annealed. If the resistance of the conductor at the new length is 75Ω, what is the initial value of R?

4.18 The conductance G of an object is defined as the reciprocal of the resistance R; $G = 1/R$. The unit of conductance is mho which is also called Siemens (S). What is the conductance (in Siemens) of an object that draws $800mA$ of current at $12.0V$?

4.19 A 40Ω resistor is made from a coil of copper wire whose total mass is $10.2g$. Show that the length and the cross-sectional area may be written as follows: $L = \sqrt{\frac{Rm}{\rho d}}$ $\quad A = \sqrt{\frac{m\rho}{Rd}}$

where ρ is its resistivity, d the density, m its mass and R is the resistance.

4.20 By combining a carbon and a nichrome resistor in series, one can devise an equivalent resistor whose resistance is independent of temperature. What percentage of the resistance should be contributed by carbon?

CHAPTER 5
DC CIRCUITS AND INSTRUMENTS

5.0 Introduction

In chapter 4, we discussed the basic principles of electricity. In this chapter, we will apply the methods of the proceeding chapter to analyse steady or direct current (dc) circuits. The current in a dc circuit does not change with time; it is steady. The behaviour of alternating current in an alternating current (ac) circuits will be studied in later chapters.

In drawing an electric circuit, symbols are used to represent the elements in the circuit. The symbol ⊣⊢ as discussed in section 4.3 is used to represent the battery while ⊣▭⊢ represents resistor and two equal parallel lines ⊣⊢ represent the capacitor. The conducting wires whose resistances are negligible compared to other resistances in the circuit are drawn simply as straight lines.

5.1 Resistors in Series and Parallel

Two or more resistors are said to be connected in series if they are connected end to end and the same current passes through each. This is shown in Figure 5.1 (a) where three resistors with resistances R_1, R_2 and R_3 are connected in series.

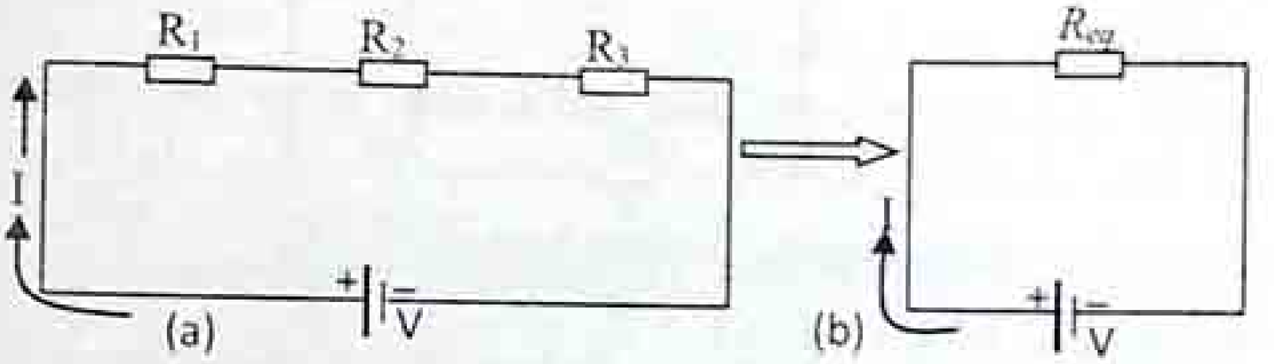


Fig. 5.1: Resistances connected (a) in series and (b) the equivalent circuit

The circuit can be simplified for analysis by finding the equivalent resistance (R_{eq}) of the resistances that are connected in series. The equivalent circuit is shown in Figure 5.1(b). This can be accomplished by noting that the same current I pass through each resistor and that the sum of the voltage drops across each resistor is equal to V. Now let V_1, V_2 and V_3 be the voltage drop across each of the resistors R_1, R_2 and R_3. By Ohm's law we have, $V_1 = IR_1$, $V_2 = IR_2$ and $V_3 = IR_3$.

Therefore, we have $\quad V = V_1 + V_2 + V_3 = IR_1 + IR_2 + IR_3$

For the equivalent single resistor R_{eq} (shown in Figure 5.1(b)) that will draw the current, we have

$$V = IR_{eq}$$

Equating these two expressions, we have $IR_{eq} = IR_1 + IR_2 + IR_3$

Therefore,

$$R_{eq} = R_1 + R_2 + R_3 \tag{5.1}$$

Equation 5.1 states that the equivalent resistance of three resistors in series is the sum of the three resistances. The above analysis, however, can be extended to include n resistors in series. Thus, equation 5.1 becomes

$$R_{eq} = R_1 + R_2 + R_3 + R_4 + \cdots + R_n \quad \text{(Resistors in series)}$$

It is clear that when more resistances are added in series to a circuit, the larger the equivalent resistance.

The parallel combination shown in Figure 5.2(a) is quite different. The resistors are connected so that the current from the battery or source splits into separate branches; the same potential drop exists across each resistor. As above, we wish to find the equivalent resistance of the three resistors R_1, R_2 and R_3 connected in parallel. Let I_1, I_2 and I_3 be the current passing through R_1, R_2 and R_3

respectively. Since charge is conserved, the current flowing into the junction A must be equal to the current flowing out of the junction.

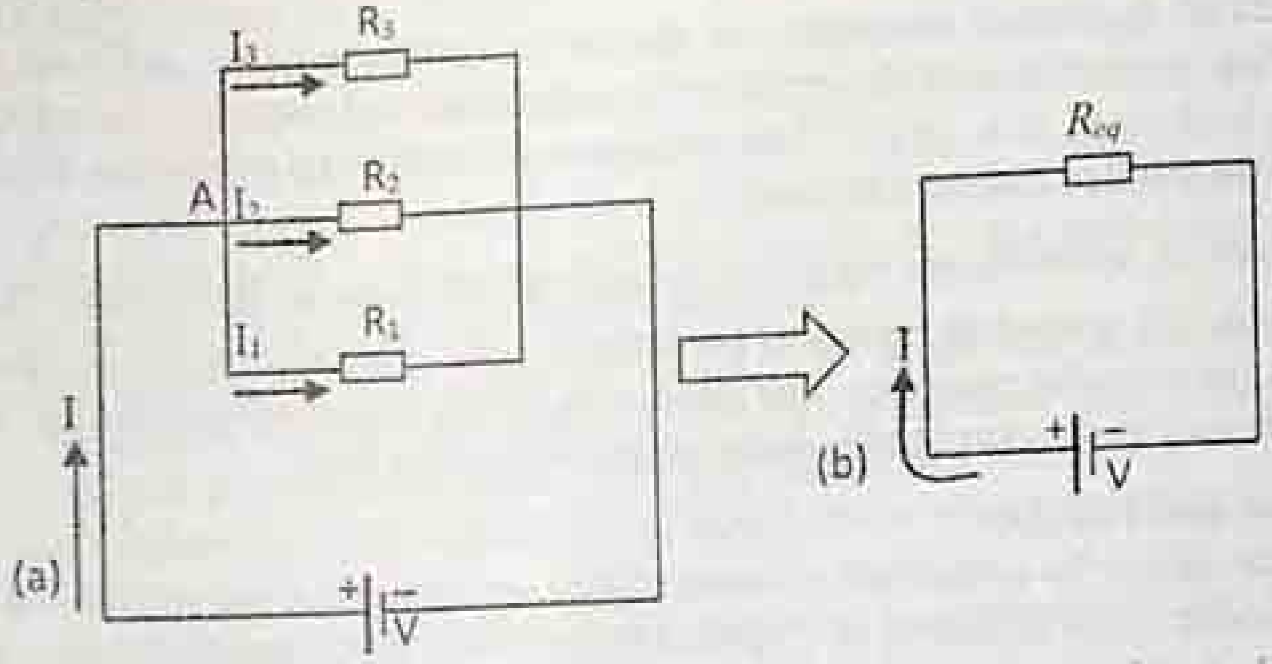


Fig. 5.2: Resistances connected in (a) parallel and (b) the equivalent circuit

Thus, $I = I_1 + I_2 + I_3$

The voltage drop across each resistor is the same and is equal to the full voltage V of the battery. Hence, $I_1 = V/R_1$, $I_2 = V/R_2$ and $I_3 = V/R_3$

From the equivalent circuit shown in Figure 5.2(b) we have $I = V/R_{eq}$.

Combining these equations, we have $I = I_1 + I_2 + I_3$

$$\frac{V}{R_{eq}} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3}$$

Therefore,
$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \qquad (5.2)$$

This represents the equation for the equivalent resistance for the Resistors in parallel connection. It is clear from equation 5.2 that the equivalent resistance R_{eq} is less than any of the resistor in parallel. Equation 5.2 can be extended for n resistors in parallel to give

$$\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots + \frac{1}{R_n}$$

Armed with equation 5.1 and 5.2, one can reduce a system of resistors to single equivalent resistor.

Example 5.1

Find the equivalent resistance of the circuit shown in Figure 5.3 and the current passing through the R_4 resistor.

Solution

It is very important to draw the diagram of the circuit at every stage of reduction. From Figure 5.3, it is clear that R_1 and R_2 are in parallel. They can be replaced using equation 5.2, thus,

$$\frac{1}{R_{1,2}} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{2}$$

Hence, $R_{1,2} = 2\Omega$. This is shown in stage (b) and from this stage we see that $R_{1,2}$ and R_3 are in series. They can be replaced using equation 5.1, thus:

$$R_7 = R_{1,2} + R_3 = 6\Omega$$

From stage (c) we see that R_4 and R_7 are in parallel and they can be replaced with R_8, thus:

$$\frac{1}{R_8}=\frac{1}{R_4}+\frac{1}{R_7}=\frac{1}{4}$$

Therefore, $R_8 = 4\Omega$

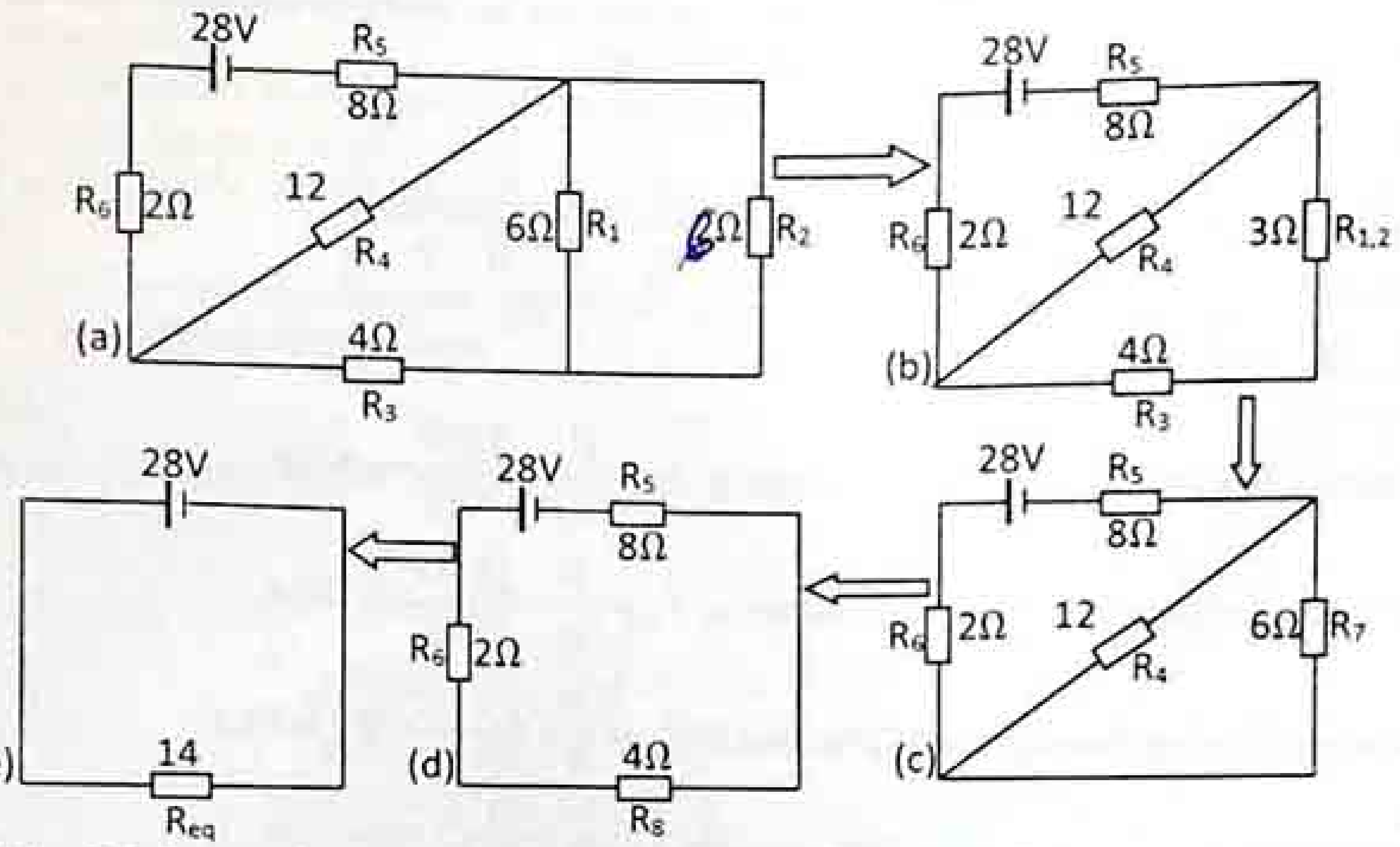


Fig. 5.3: Successive steps in reducing a system of series and parallel resistances

From stage (d) it is clear that R_5, R_6 and R_8 are in series and the equivalent resistance of the resistors in the circuits is 14 Ω as shown in (e). The current, I in the circuits is $I = V/R = 2A$.

In order to find the current through any of the resistors in the circuit, we have to work backwards, since R_{eq} is obtained by adding R_5, R_6 and R_8 in series. The same current I passes through each of them.

Therefore, the voltage drop across R_8 is just $V_8 = IR_8 = 8V$. In stage (c), we combine R_4 and R_7 in parallel to obtain R_8, hence the voltage drop across R_4 and R_7 is $V_8 = 8V$. Thus, I_4 is just $V_8/R_4 = 0.67A$. The current passing through other resistors can be similarly obtained.

Example 5.2

A student carrying out an experiment to investigate Ohm's law required to limit the current flowing through a 10 Ω resistor to 20 A when a power supply of 240 V was connected. How should an auxiliary resistor be connected in the circuit and what is the value of its resistance?

Solution

The equivalent single path circuit should require a resistance of $R=\frac{V}{I}=\frac{240V}{20A}=12\Omega$

Since the equivalent resistor is greater than 10 Ω, the connection should be in series. The auxiliary resistance is $R_a = 12\Omega - 10\Omega = 2\Omega$.

Example 5.3

Three resistors 2Ω, 6Ω and 9Ω are connected in series and a 120 V battery connected across the combination (a) calculate the equivalent resistance (b) the current flowing in each resistor.

Solution

(a) The equivalent resistance is $R_{eq} = R_1 + R_2 + R_3 = 2 + 6 + 9 = 17\Omega$

(b) The current flowing through the circuit is the same and is $I = \frac{V}{R_{eq}} = \frac{120V}{17\Omega} = 7.06A.$

Example 5.4
If the three resistors in Example 5.3 are now connected in parallel, calculate (a) the equivalent resistance (b) the current flowing through each resistor?

Solution

(a) The equivalent resistance is $\frac{1}{R_{eq}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{2} + \frac{1}{6} + \frac{1}{9} = \frac{7}{9}$

$\therefore \quad R_{eq} = \frac{9}{7} = 1.29\Omega$

(b) The current flowing through the 2Ω resistor is $I_1 = \frac{V}{R_1} = \frac{120V}{2\Omega} = 60A.$

The current flowing through the 6Ω resistor is $I_2 = \frac{V}{R_2} = \frac{120V}{6\Omega} = 20A.$

The current flowing through the 9Ω resistor is $I_3 = \frac{V}{R_3} = \frac{120V}{9\Omega} = 13.33A.$

5.2 Electrical Network Analysis - Kirchhoff's Laws

This is the principle guiding how electrical circuits with more than one path for the flow of current can be connected into an equivalent single-path circuit. In analyzing such circuits, Ohm's law is adequate and is readily applied. However, there are more complicated circuits which cannot be readily reduced to single-path equivalent circuits. For such complicated circuits, Kirchhoff's rules which are a generalized form of Ohm's law are required.

Kirchhoff's first rule is based on the law of conservation of charge and governs the flow of current at any junction (or Nodal Point) of an electric circuit. This law states that *in a network circuit carrying steady currents, the sum of the currents flowing into a junction is equal to the sum of the current flowing out of it.* Thus, a junction A in Figure 5.4 is analysed as follows:

i.e. $$I_1 = I_2 + I_3 + I_4$$

or $$\sum I = 0$$

where it is assumed that the current flowing into the junction are positive and those flowing out of it are negative.

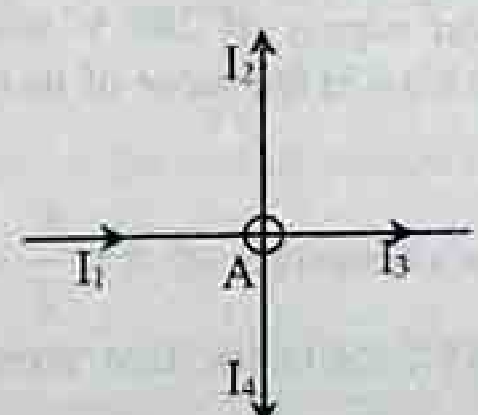


Fig. 5.4: Kirchhoff's first rule

The second rule on the other hand deals with the potential drop across the resistor in the circuit and is based on the law of conservation of energy. It states that *the sum of average drops in all resistances in a closed circuit (closed loop) is equal to the sum of the emfs in the circuit.* Thus, for any closed circuit (closed loop):

$$\sum IR = \sum E$$

The two rules allow for the calculation of currents flowing through each branch and/or potential drop across each resistance in a closed circuit. The procedure employed is as follows:

1. Assign a direction to the current in each branch of the circuit.
2. Identify each closed loop and design a direction for adding up the *emfs* and the potential drops around each closed loop. We must note that all currents in the chosen direction in any closed loop is positive (otherwise negative) and all *emfs* traversed from the negative to the positive terminal are positive (otherwise negative). Thus, the potential drop IR across each resistor is positive (otherwise negative).
3. Apply the Kirchhoff's current rule at each junction. If there are n junctions, $n-1$ equations can be generated using the current rule.
4. Apply the Kirchhoff's second rule to each closed loop. Thus, if there are n loops, n equation can be generated.
5. The solutions to the simultaneous equations generated can be obtained either by substitution method or by applying determinants

Example 5.5

Find the current I_1, I_2 and I_3 in the circuit shown in Figure 5.5.

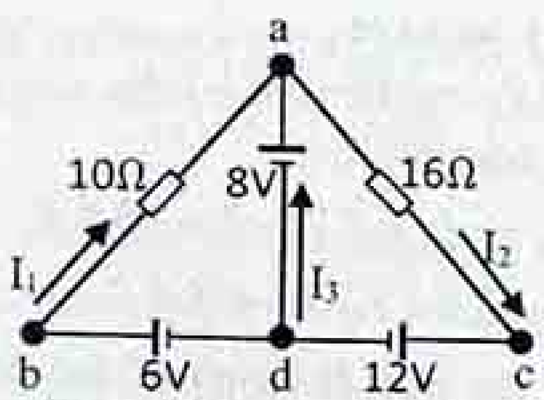


Fig. 5.5: Example 6.5

Solution

Let us assign directions to the currents, I_1 and I_3 are flowing towards junction a while I_2 is leaving the junction. Using Kirchhoff's first rule we have: $I_2 = I_1 + I_3$.

Now consider a closed loop adba and start tracing from a to d to b and back to a. the voltage drops encountered according to the loop rule is: $-8+6-10I_1=0$; $I_1=-0.2A$.

Now take the loop acda; this loop contains the other unknown current I_2. Tracing through this loop, we have: $-16I_2-12V+8V=0$; $I_2=-0.25A$.

Substituting these values I_1 and I_2 in the above expression, we obtain the value of I_3. Thus:

$$-0.25A=-0.2A+I_3; \qquad I_3=-0.5A.$$

The negative signs in the currents indicate that the true directions of the currents are opposite ones assumed. The magnitudes are however correct.

Example 5.6

Calculate the current flowing in each branch of the network circuit shown in the Figure 5.6 below.

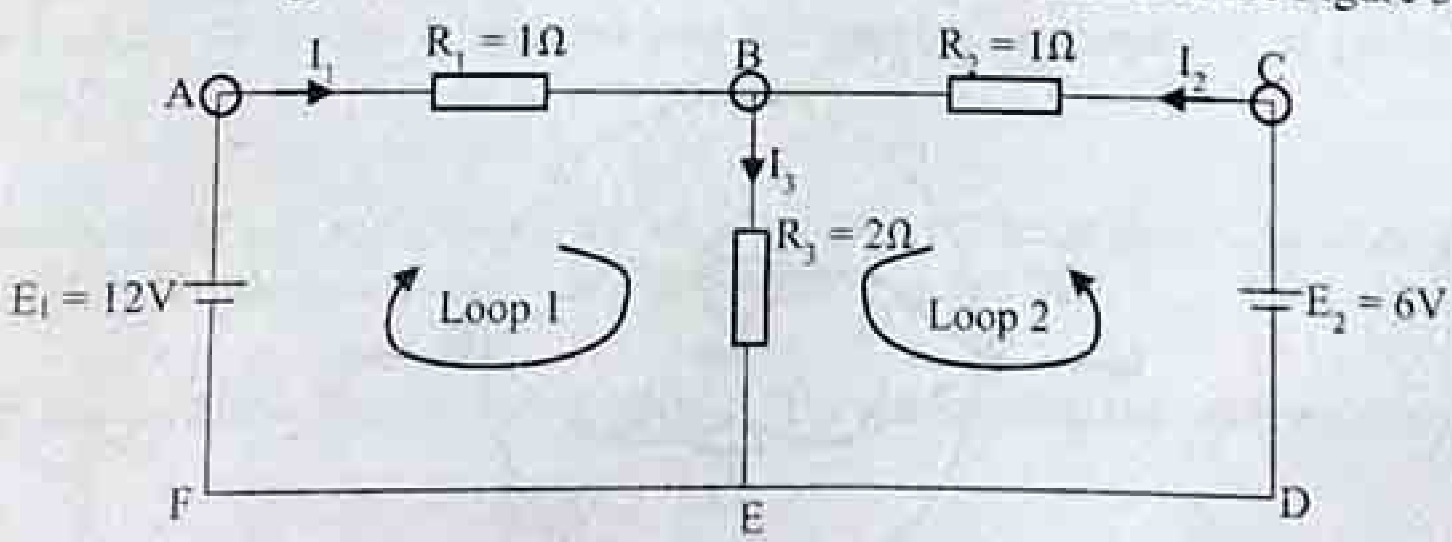


Fig. 5.6: Example 5.6

Solution

There are three possible closed loops in the entire circuit: ACDFA, ABEFA and BCDEB. Currents I_1, I_2 and I_3 are assigned to each resistor with directions indicated. Applying Kirchoff's first law at the junction B, we have: $I_1 + I_2 = I_3$ (1)

Also, applying Kirchoff's second law to the loop ACDFA gives the following equation for the potential drop across each resistor in the loop:

$$12 + I_1 - I_2 - 6 = 0 \text{ or } I_1 = I_2 - 6 \quad (2)$$

For loop ABEFA, $12 + I_1 + 2I_3 = 0$ or $I_1 = -12 - I_3$ (3)

For loop BCDEB, $6 + I_2 + 2I_3 = 0$ or $I_2 = -6 - 2I_3$ (4)

Using equation (1) in equation (2) and (3) to eliminates I_1 gives

$$I_3 - I_2 = I_2 - 6 \text{ or } I_3 - 2I_2 = -6 \quad (5)$$

And $I_3 - I_2 = -12 - 2I_3$ or $3I_3 - I_2 = -12$ (6)

Multiply equation (5) by 3 and subtract from (6), we have

$$-5I_2 = -6 \text{ or } I_2 = -6/-5 = 1.2A$$

Substituting the value of I_2 in equation (4), we obtain

$$1.2 = -6 - 2I_3 \text{ or } I_3 = 7.2/-2 = -3.6A$$

From equation (1), $I_1 = I_3 - I_2 = -3.6 - 1.2 = -4.8A$

The negative sign in I_1 and I_3 shows that the true direction of the current is opposite the one assumed. The magnitude is however correct.

5.3 Ammeters and Voltmeters in DC Circuits

A meter to measure currents is called ammeter (also called *milliammeter, microammeter* etc depending on the size of the current to be measured). While a meter to measure potential differences is called voltmeter (also called *millivoltmeter, microvoltmeter, etc*).

The most common type of ammeter or voltmeter is a modified moving coil galvanometer, in which a pivoted coil of fine wire carrying a current is deflected by the magnetic interaction between this current and the magnetic field of a permanent magnet. The resistance of the coil of a typical instrument is of the order of 10 to 100Ω, and a current of the order of a few milliamperes will produce full scale deflection (f.s.d). An ammeter must have a low resistance compared with the resistance in the rest of the circuit. On the contrary, a voltmeter is a high resistance instrument.

5.4 Conversion of Galvanometer into an Ammeter

Suppose that the galvanometer has a resistance R_G, and the full-scale deflection (f.s.d) is produced by a current of I_G through the galvanometer. The maximum p.d that should be applied across the terminal is given by

$$V_G = I_G R_G \quad (5.3)$$

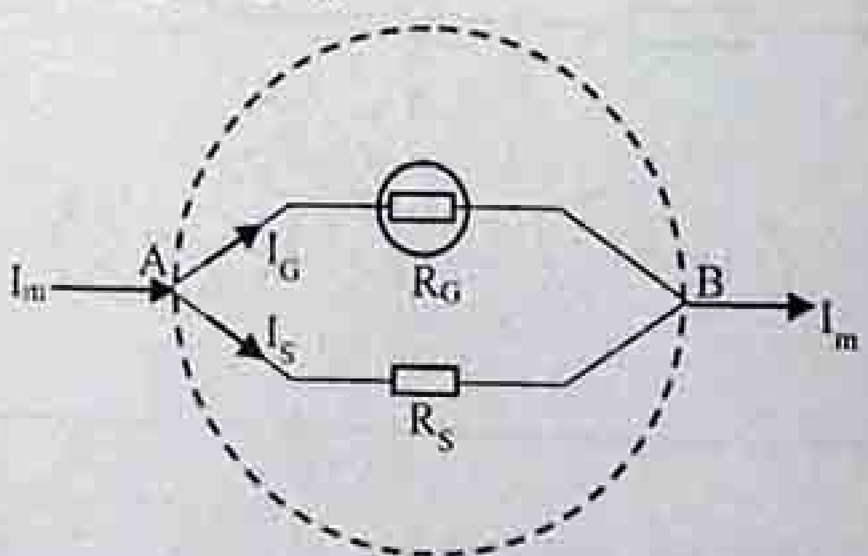


Fig. 5.7: Circuit showing conversion of the galvanometer into an ammeter, using a shunt resistance

To measure up to I_m, we add a shunt (resistor of small resistance, R_S) in parallel. Since R_G and R_S are parallel, the p.d across them is the same. That is, from Figure 5.7,
From the diagram above, note that I_m is the main circuit current which produces the full scale deflection.

$$V_{AB} = I_G R_G + I_S R_S \tag{5.4}$$

where $I_S = (I_m - I_G)$

This implies that;
$$R_S = \frac{I_G}{I_S} R_G = \frac{I_G R_G}{I_m - I_G} \tag{5.5}$$

The ammeter as a whole has the following properties:
(i) It has a very low resistance.
(ii) It is connected in series with the circuit (so one has to break the circuit to insert the ammeter).

Example 5.7
A galvanometer has a coil of resistance 5Ω and gives full scale deflection with a current of 10 mA. How may it be made into an ammeter reading up to 10 A?

Solution
The galvanometer must be provided with a shunt of resistance R_S such that when 10 A passes through the pair in parallel, 10 mA (0.01 A) goes through the galvanometer and 9.99 A through the shunt.

P.d across the shunt = P.d across the coil

i.e. $I_S R_S = I_G R_G$

$$9.99R_S = 0.01 \times 5; \qquad R_S = \frac{0.01 \times 5}{9.99} = 0.005\Omega$$

The shunt is hereby said to have a power of 1000 because $\dfrac{I}{I_G} = 1000$.

5.5 Conversion of a Galvanometer into a Voltmeter

A voltmeter measures the p.d between two points, and its terminals must be connected to these points. To measure V_{AB}, add a multiplier (a resistor of high resistance, R_m) in series with the coil. If V_m is the largest p.d to be measured, then we have that

$$V_m = I_G R_G + I_G R_m \tag{5.6}$$

or
$$R_m = \frac{V_m}{I_G} - R_G \tag{5.7}$$

The voltmeter as a whole has the following properties:
i. It has a very high resistance.
ii. It is connected across the points A and B whose p.d is to be measured (i.e. in parallel with the device concerned).

Example 5.8
How can a galvanometer of coil resistance 20Ω, which gives a fill-scale deflection when a current of $10mA$ passes through it, be converted into a voltmeter reading up to $100V$.

Solution
When $V_m = 100V$, the current in the cell must be $0.01A$. Therefore, using the equation (5.6),

i.e. $$R_m = \frac{V_m}{I_G} - R_G = \frac{100}{0.01} - 20 = 9980\Omega$$

A multi-range voltmeter uses a series of multipliers of different resistances, which can be chosen by a switch or plug and socket arrangement. Multi-meter is an instrument which is adapted for measuring current, voltage and resistance.

5.6 The Potentiometer

The potentiometer consists of a uniform resisitance wire AB in Figure 5.8 (a) of length $100cm$. A source of e.m.f, E maintains a steady current, I in AB and since the wire is uniform, if R is the resistance per centimeter, then

$$R \propto l$$

And since the current is constant, then $V_{AC} = kl$ where k is the constant.

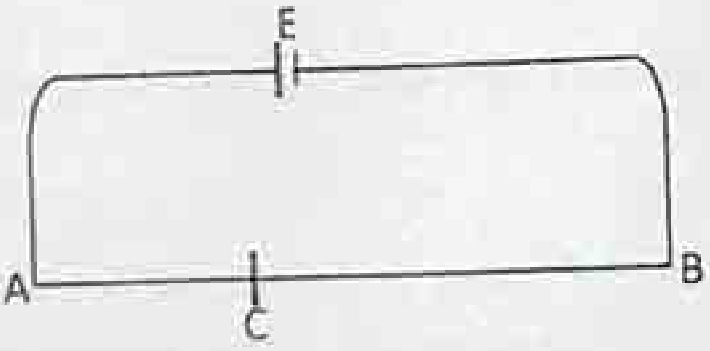


Fig. 5.8(a): Potentiometer

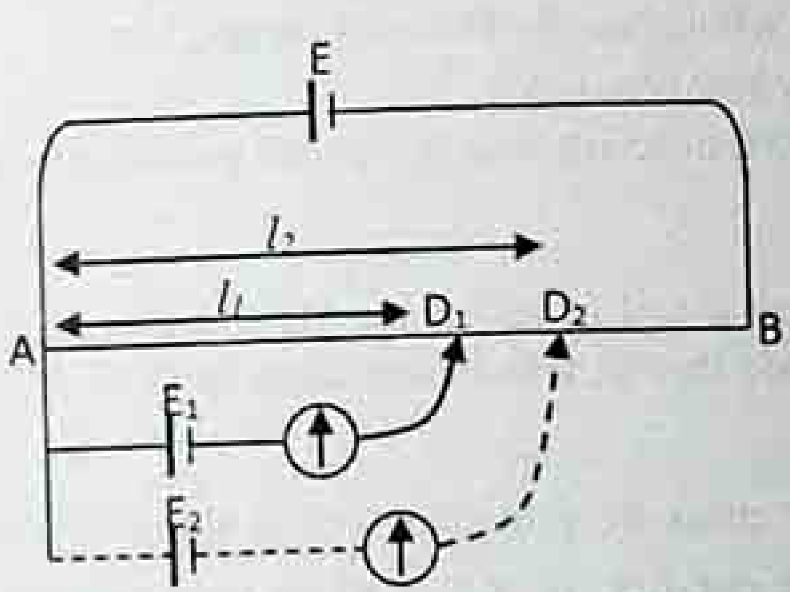


Fig. 5.8(b): Comparison of e.m.f of cells

5.7 Comparison of e.m.fs of two cells

In the circuit shown in Figure 5.8(a), the battery E provides a steady current through the wire. A cell C_1 of e.m.f E_1 is connected as shown in Figure 5.8(b) through a galvanometer and a balance point is obtained at D_1. This means that the e.m.f $E_1 = kl_1$ for C_2. With a second cell C_1 in place of C, a new balance, l_2 is obtained i.e the e.m.f $E_2 = kl_2$ for C_2.

Therefore,
$$\frac{E_1}{E_2} = \frac{l_1}{l_2} \qquad (5.8)$$

Hence, the e.m.fs of the two cells can be compared. If one of them is a standard cell of known e.m.f, the e.m.f of the other can be calculated.

Accuracy

1. When the potentiometer is used to compare the e.m.f of cells, no errors are introduced by internal resistances since no current flows at balance.
2. It is more accurate than the moving coil voltmeter for measuring e.m.f.
3. For every sensitive galvanometer, a high precision of the balance point of the potentiometer is possible.

Example 5.9

A potentiometer wire carrying a steady current is $100cm$ long. When a standard cell of e.m.f $1.5V$ is connected to a balance length of $60.0cm$ was obtained. Calculate the e.m.f of a cell of a cell that gives a balance length of $80.0cm$.

Solution

Using equation 5.8, $$\frac{E_1}{E_2} = \frac{l_1}{l_2}$$

$$\frac{1.5}{E_2} = \frac{60}{80};\ E_2 = \frac{4 \times 1.5}{3} = 2.0V$$

5.8 Wheatstone Bridge

About 1843, Wheatstone designed a bridge circuit which gave an accurate method for measuring resistances. It consists of four resistors of resistances R_1, R_2, R_3 and R_4 connected as shown in Figure 5.9.

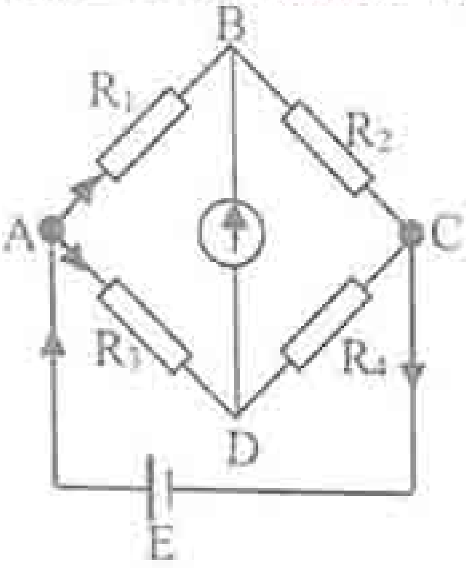


Fig.5.9: Wheatstone Bridge

The current I from the battery E divides at A along the two branches ABC and ADC. By adjusting one or more of the resistances, a condition is reached when no current passes through the galvanometer, G and indicates no deflection.

This is "balance" condition and

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \tag{5.9}$$

When no current flows through G, the points B and D are at the same potential. Therefore, $V_{AB} = V_{AD}$

i.e. $I_1 R_1 = I_2 R_3$

Also, $V_{BC} = V_{DC}$

$I_1 R_2 = I_2 R_4$

So, if R_1, R_2 and R_3 R are known, R_4 can be determined.

5.9 Meter Bridges

The meter-bridge is a simple form of Wheatstone bridge used in the laboratory for measurement of resistance. It is a practical application of Wheatstone bridge network principle in which the ratio of two resistances is deduced from the ratio of their balance points.

When the balance point is achieved at D, say, the jockey divides the wire into two parts with resistances R_{AD} and R_{CD} corresponding to lengths l_1 and l_2, respectively, on the wire. R_{AD}, R_{CD}, X and R thus form the four arms as in the case of a wheatstone bridge. Consider the diagram above.

At balance,
$$\frac{X}{R} = \frac{R_{AD}}{R_{CD}} = \frac{l_1}{l_2} \tag{5.10}$$

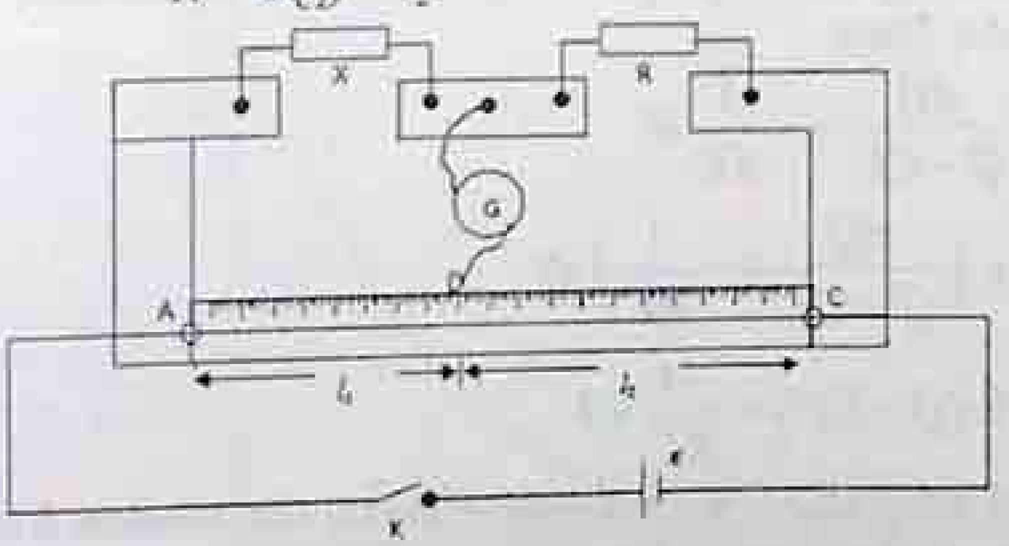


Fig. 5.10: The Meter Bridge

Example 5.10

In a meter wire bridge experiment, a balance point was found on the $55cm$ on the wire corresponding to a resistance, R when a resistance of 30Ω was connected to the other arm of the bridge. Find the value of R?

Solution

Let $l_2 = 55cm$ which correspond to R, then $l_1 = 100 - 55 = 45cm$ will correspond to $X = 30\Omega$

When the bridge is balanced, we use equation (5.10),

i.e. $\dfrac{X}{R} = \dfrac{l_1}{l_2}$ or $R = \dfrac{Xl_2}{l_1} = \dfrac{30\Omega \times 55cm}{45cm} = 36.67\Omega$

5.10 RC Circuits

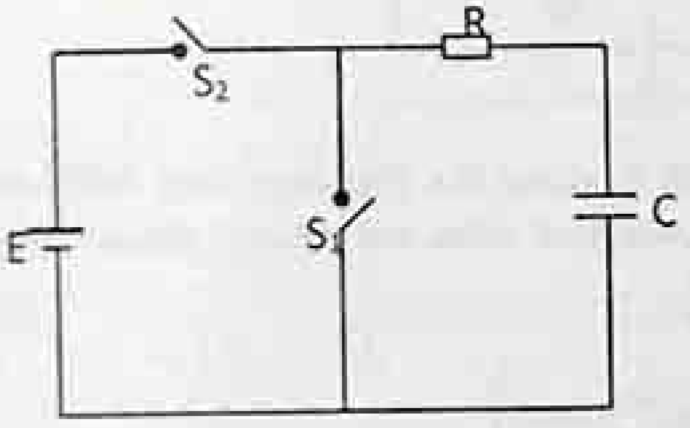


Fig. 5.11: A typical RC

Resistors and capacitors are usually found in electric circuits. When a circuit contains both resistor (R) and a capacitor (C), it is known as a RC circuit. Figure 5.11 shows a typical RC circuit, it contains a capacitor C, a resistor R, a battery with e.m.f E and two switches S_1 and S_2.

Initially, the switches S_1 and S_2 are open and the initial charge on the capacitor is zero. When S_2 is closed (S_1 is still open), electrons will flow from the negative terminal of the battery and accumulate at the lower plate of the capacitor while electrons will flow into the positive terminal of the battery through the resistor R leaving a positive charge on the upper plate of the capacitor. As the charge accumulates on the capacitor, the current is reduced until eventually the voltage across the capacitor is equal to the e.m.f E of the battery and no further current flows. The current, therefore, is not steady but transient. Suppose I is the current at any instant and Q is the charge on the capacitor at the same instant, using loop rule we obtain;

$$E = IR - \frac{Q}{C} = 0 \text{ or } E = IR + \frac{Q}{C}$$

Recalling that $I = dQ/dt$ and substituting this in the above expression we have

$$E = R\frac{dQ}{dt} + \frac{Q}{C} \tag{5.11}$$

The equation can be solved by multiplying through by C and by rearranging terms to separate the variables which are Q and t. Thus:

$$\frac{dQ}{Q - CE} = -\frac{1}{RC}dt$$

which we can integrate as $\displaystyle\int \frac{dQ}{Q - CE} = -\frac{1}{RC}\int dt$

$$\ln(Q - CE) = \frac{-t}{RC} + k$$

where k is the constant of integration. We can now evaluate the constant k by noting that at $t = 0$ (i.e. when the switches are open), $Q = 0$, so we have:

$$\ln(Q - CE) = 0 + k \text{ or } k = \ln(-CE)$$

Substituting this value of k into the above relation, we obtain

$$\ln(Q-CE)=\frac{-t}{RC}-\ln(CE) \text{ or } \ln\left(1-\frac{Q}{CE}\right)=\frac{-t}{RC}$$

Taking the antilog or exponent of both sides gives $1-\frac{Q}{CE}=e^{-t/RC}$ or

$$Q=CE\left(1-e^{-t/RC}\right) \tag{5.12}$$

Equation 5.12 is the solution to equation 5.11 and we see that the charge accumulating on the capacitor increases from $Q = 0$ (at $t = 0$) to a maximum value of CE (at $t = \infty$). Equation 5.12 is plotted in Figure 5.12. We should note that the charge rises to $(1-1/e)$ or 63% times its final value in time $t = RC$. This time RC is known as time constant of the circuit and it represents the time required for the capacitor to reach 63 percent of its full value.

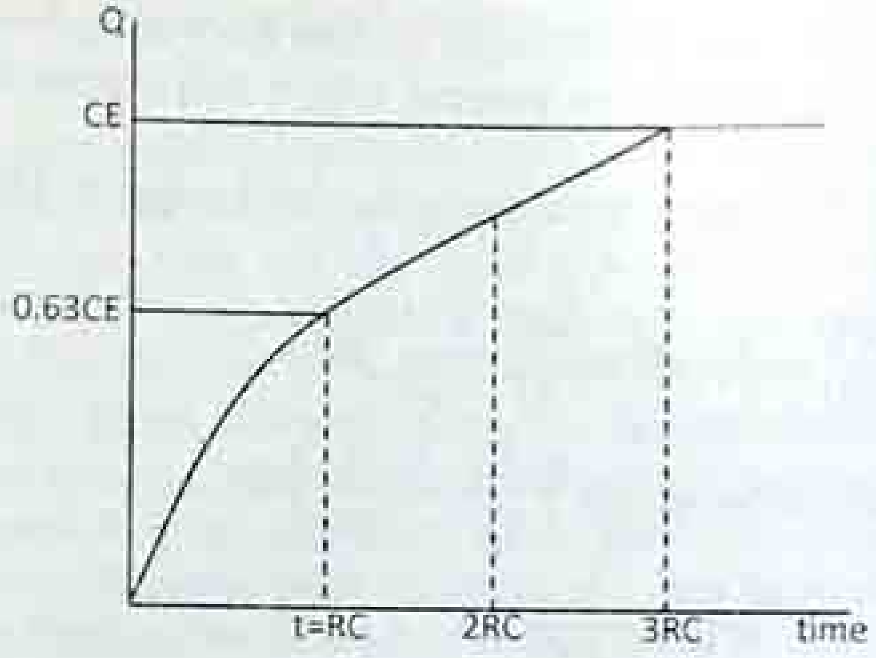


Fig. 5.12: Accumulation of charge on a capacitor when connected to a battery of e.m.f E.

It is therefore a measure of how quickly a capacitor gets charged and since every circuit contains some resistance, no capacitor can be charged instantaneously when connected to a battery. The charging current I through the resistor R at anytime t can be obtained by differentiating equation 5.12.

Thus, $$I=\frac{dQ}{dt}=\frac{E}{R}e^{-t/RC}$$

The current starts with a value E/r at $t = 0$, since at $t = 0$ the capacitor is uncharged- the whole battery voltage appears across the resistor. However, as the capacitor becomes charged, less of the voltage appears across R and the current begins to drop from its initial value to I_0/e at $t = RC$ and to zero at $t = \infty$. The variation of the current as a function of time is shown in Figure 5.13.

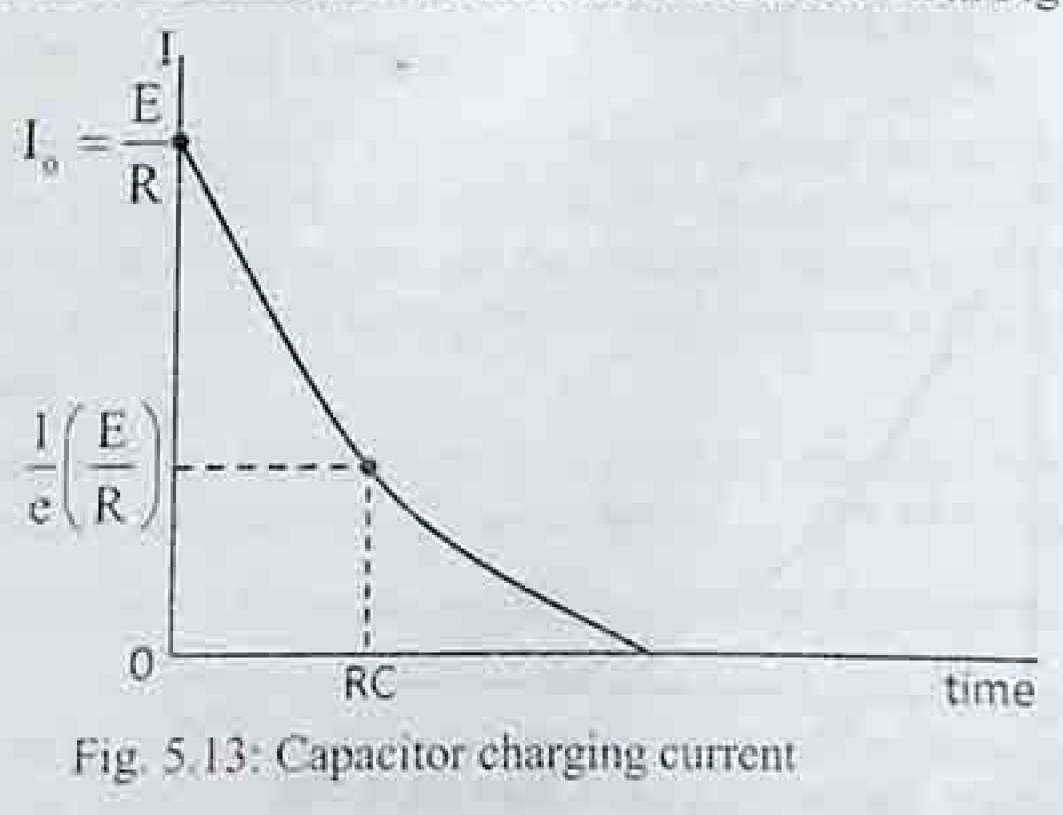


Fig. 5.13: Capacitor charging current

Example 5.11
How many time constants must elapse before a capacitor in an RC circuit is charged to within 1.0 percent of its equilibrium charge?

Solution
We know from equation 5.12 that $Q = CE\left(1 - e^{-t/RC}\right)$
But our capacitor is charged to only 1 percent of CE, we have $0.01C = CE\left(1 - e^{-t/RC}\right)$

$\therefore \quad -0.99 = -e^{-t/RC}$ or $\ln 0.99 = -t/RC$

$t = 4.6RC$

Let us now return to Figure 5.11, the capacitor is fully charged to some final value Q_0. Suppose we open the switch S_2 and close S_1. Charge will begin to flow from one plate of the capacitor through the resistor to the other. This is a discharging process which will continue until the capacitor is fully discharged.
Notice that when S_2 is open, the battery is no longer part of the circuit. Hence, the e.m.f, E in equation 5.11 is zero and it becomes

$$R\frac{dQ}{dt} + \frac{Q}{C} = 0 \text{ or } \frac{dQ}{Q} = -\frac{1}{RC}dt$$

Integrating the above expression we obtain $\int \frac{dQ}{Q} = -\frac{1}{RC}\int dt$

$$\ln Q = -\frac{1}{RC}t + k$$

Again k is constant of integration; k can be determined by the noting that the charges is Q_0 at $t = 0$.
Therefore, at $t = 0$ we have $\ln Q_0 = k$.
Substituting this value of k in the above expression we obtain

$$\ln Q = -\frac{1}{RC}t + \ln Q_0 \text{ or } \ln\frac{Q}{Q_0} = -\frac{1}{RC}t$$

Taking the exponent of both sides we find

$$Q = Q_0 e^{-t/RC} \tag{5.13}$$

The charge on the capacitor therefore decreases exponentially in time with a time constant RC. This is shown in Figure 5.14. The current through R during discharge is obtained by differentiating equation 5.13. Thus,

$$\frac{dQ}{dt} = I = \frac{Q_0}{RC}e^{-t/RC}$$

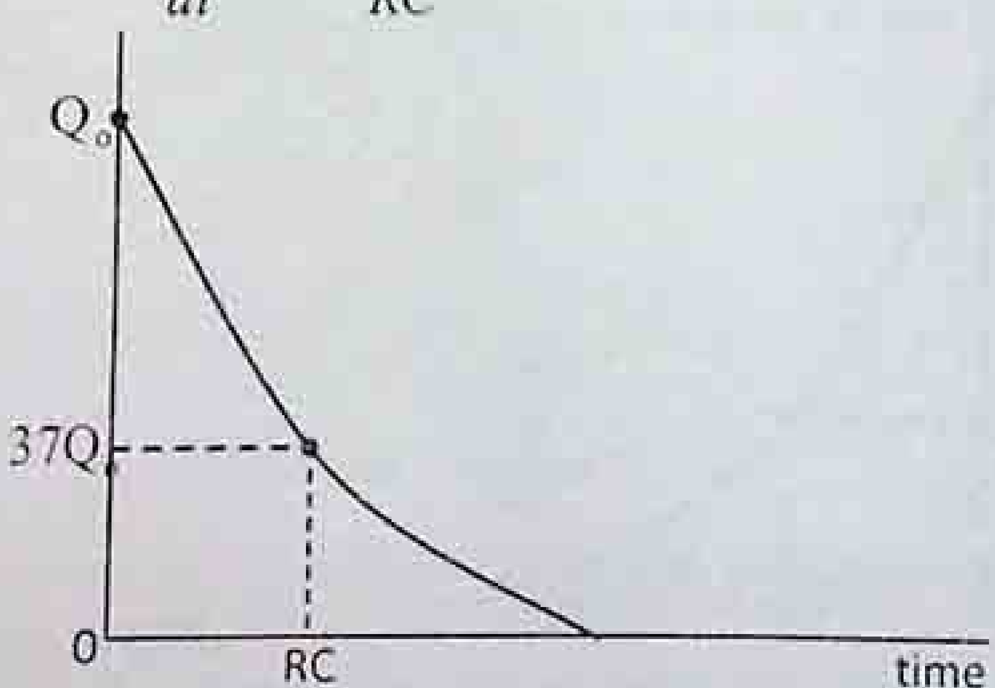


Fig. 5.14: Discharging curve for a capacitor

or $$I = I_0 e^{-t/RC} \tag{5.14}$$

where, $$I_0 = -\frac{Q_0}{RC}$$

The current I, from equation 5.14, decreases exponentially much like the result we obtained earlier in equation 5.12.

5.11 The Cathode Ray Oscilloscope

This is an essential instrument in any electronic or electromagnetic laboratory. It is used to visually follow and estimate transient variations of voltage which are of too high a frequency for any mechanical device to follow. The oscilloscope tube was first developed by the Physicists Braun, in 1897. The essentials of a modern tube are shown in Figure 5.15. Electrons from the thermionic cathode C pass through a small hole in a control grid and are accelerated by the voltage applied to the anode A.

The electron beam, if there were nothing else in the highly evacuated tube, would be undeflected and strike the fluorescent screen S, producing a point of light. However, it passes between two vertical plates X, X' which in normal operation, are connected to a capacitor which is being charged in the manner prescribed in section 5.6. The accumulated charge on the capacitor or the voltage between the plates varies as Figure 5.12. The electron beam is deflected more and more with time up to a certain value, after which it returns almost instantaneously to the undeflected position and begins the deflection process again. The point of light therefore moves uniformly across the tube, flicks back to the starting position, and begins to move uniformly again. Since the fluorescence lasts for finite time, the trace on the screen is a line. What has been produced is a time base, since the voltage between the X – X' plates, and thus the deflection, varies linearly with time.

If a varying signal is also applied to a pair of horizontal plates Y and Y', the spot is deflected vertically as well as horizontally. The line traced by the spot of light thus shows the variation of the wave form against time. If this is sinusoidal, the trace will be as shown on the CRO screen in Figure 5.15.

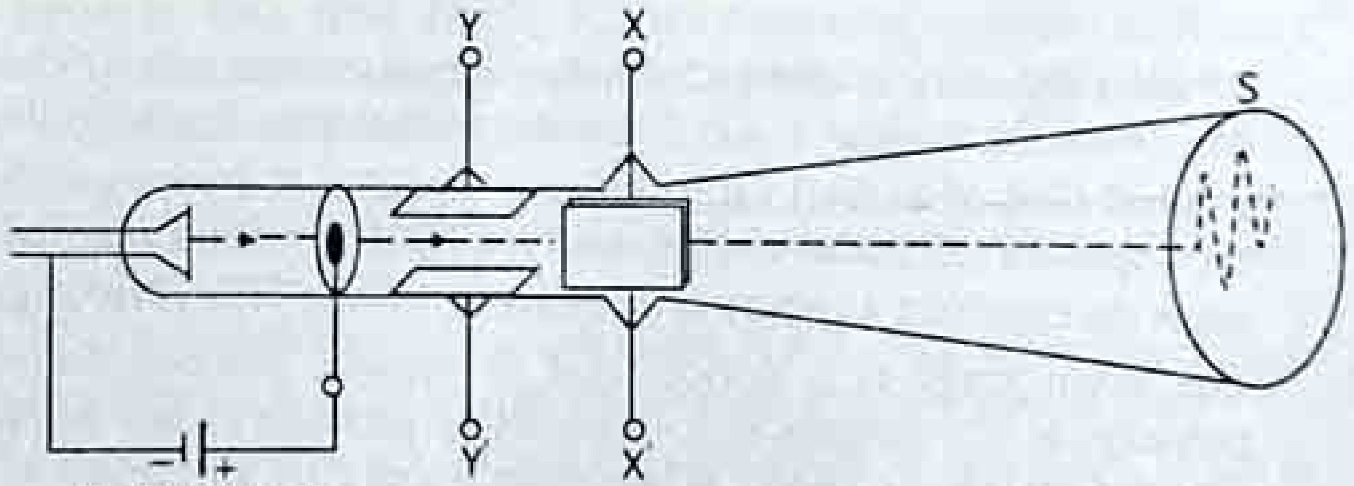


Fig. 5.15: Cathode ray tube

There are of course various controls on the CRO. Some of these are concerned with centering the trace, but there are others which alter time it takes for the trace to move across the screen. There are also devices which lock the incoming signal and the time base so that a stable pattern appears on the screen, or which starts the trace only if a signal arrives. This latter control is useful for studying pulses of irregular repetition rate.

The principle of the cathode ray oscilloscope is utilised in television tubes as well, except that magnetic deflecting forces are used instead of electric forces.

Other important D.C. instruments include strain-gauge, electrostatic voltmeter and digital voltmeter.

Summary

1. The equivalent resistance of a number of resistors connected in series is given by
$$R = R_1 + R_2 + R_3 + \cdots$$

For resistances connected in parallel we have $\frac{1}{R}=\frac{1}{R_1}+\frac{1}{R_2}+\frac{1}{R_3}+\cdots$

2. Kirchhoff's rules are:
   (i) At any junction in a circuit the sum of all currents directed towards the junction is equal to the sum of all currents directed away from the junction:
   $$\sum I = 0$$
   (ii) Around any closed path in a circuit, the algebraic sum of all the changes of potential is zero. The loop rule may be written as $\sum(E + IR) = 0$.

3. A galvanometer is a device that detects a small current that passes through it and gives a small scale deflection that is proportional to the current.
   The ammeter is a galvanometer plus a parallel resistor called a shunt resistor. It is used for measuring current. It can be connected in series in a circuit to measure the current without unduly affecting the current to be measured, because it has a very small resistance. The voltmeter which is used to measure the potential difference consists of a galvanometer plus a large series resistance. The voltmeter is connected between two points in a circuit to measure the p.d. between those points. It has a high resistance, therefore, the little current that flows through it will not affect the p.d. to be measured.

4. A Wheatstone bridge is a device for the comparison of known precision resistances R_1, R_2, R_3 in the determination of an unknown resistance R_x. The working equation is $R_x = R_1(R_3/R_2)$.

5. The potentiometer is an instrument that accurately compares an unknown potential difference with a standard potential difference. An advantage of the potentiometer over the voltmeter is the fact that the potentiometer takes no current from the source being measured and hence gives a reading of terminal potential difference equal to the e.m.f. The basic equation is
   $$\frac{E_1}{E_2}=\frac{l_1}{l_2}$$

6. When a capacitor is charged through a resistor, the rate of charging decreases exponentially with time. At a time $\tau = RC$, known as the time constant, the charge on the capacitor has reached 63% of its final value. When a capacitor is discharged through a resistor, the charge on the capacitor decreases exponentially with time with a time constant RC. This is the time it takes the capacitor's charge to decrease to $1/e$ or 37% of its original value.

7. The strain-gauge is a device used in medical practice for measuring pressure e.g. blood pressure. It is a transducer for converting pressure effects to electrical effects which are then detected accurately using a Wheatstone bridge type arrangement.

8. The electrostatic voltmeter is an instrument for measuring very high voltages while the cathode ray oscilloscope (CRO) is used to follow visually and estimate transient variations of voltage. Digital instruments have high accuracy, and enable information to be fed directly into digital computer.

Exercises 5

5.1 Three resistors $R_1 = 1\Omega$, $R_2 = 2\Omega$ and $R_3 = 3\Omega$ are connected in parallel. Find the equivalent resistance, R_{eq}.

A. 0.55Ω **B.** 0.65Ω **C.** 0.75Ω **D.** 0.85Ω

5.2 Which of the following is/are correct about Kirchhoff's rules?

I. Kirchhoff's rules are generalized form of Ohm's law.
II. Kirchhoff's first rule is based on the law of conservation of charge.
III. Kirchhoff's second rule based on conservation of momentum.
IV. Kirchhoff's second rule states that the sum of average drops in p.d. across all resistances in a closed circuit (closed loop) is equal to the sum of the emfs in the circuit.

A. I & III only **B.** I, II & IV only **C.** II & IV only **D.** I & II only

5.3 In converting galvanometer into an ammeter, which of the following is correct?

A. A shunt, R_s is connected in series with the galvanometer.

B. A multiplier is connected in parallel with the galvanometer.

C. A shunt, R_s is connected in parallel with the galvanometer.

D. A multiplier is connected in series with the galvanometer.

5.4 Which of the following are correct uses of potentiometer?

I. Comparison of emfs,

II. Measurement of internal resistance of cell,

III. Measurement of current,

IV. Comparison of resistance.

A. I, II & IV only **B.** I & II only **C.** II, III & IV only **D.** I, II, III & IV only.

5.5 For a standard cell of 6 V, the null point is obtained if the length of the potentiometer wire is 40*cm*. What is the emf of a known battery if the null point is obtained at 60*cm*?

A. $3V$ **B.** $6V$ **C.** $7V$ **D.** $9V$

5.6 The emf of a cell, $E = 3V$ which is balanced across $l = 100cm$ of a potentiometer wire. The cell is then shunted by the resistance $R_s = 30\Omega$. The required balance length is $l_s = 80cm$. What is the value of the current flowing through the shunt?

A. $0.125A$ **B.** $0.335A$ **C.** $0.225A$ **D.** $0.250A$

5.7 In a meter bridge experiment, a balance point was found in the wire corresponding to a resistance R when a resistance of 40Ω was connected to the other arm of the bridge. Find the value of R.

A. 50Ω **B.** 60Ω **C.** 70Ω **D.** 75Ω

5.8 Describe how a potentiometer circuit can be used to compare the emfs of two cells. A cell of emf $6V$ is connected to a potentiometer wire and the galvanometer indicates null condition at $l_1 = 15.5cm$. Another cell of unknown emf is used to replace the first cell and a new balance point is found at $l_2 = 35cm$. What is the emf of the second cell?

5.9 Describe the Wheatstone bridge and show how it can be used to determine the unknown resistances of a wire. Resistances 40Ω, 65Ω, 140Ω, RΩ are connected in order to form a Wheatstone bridge circuit. A galvanometer is connected to the junction of 40Ω, RΩ and a cell of emf of 12Ω is connected to the other pair junctions. Find the value of R required to balance the bridge?

5.10 A galvanometer has a coil of resistance 4Ω and gives f.s.d with a current of $10mA$. How may it be made into an ammeter reading up to $10A$?

5.11 How can a galvanometer of coil resistance 10Ω which gives a full scale deflection (f.s.d) when a current of $10mA$ passes through it be converted into voltmeter reading up to $100V$?

5.12 Describe how you can use the meter bridge to compare the resistances of two wires. What is the ratio of R_1/R_2 of the two resistance wires which produces balance for 42*cm* and 62*cm* respectively of the meter bridge circuit?

5.13 State Ohm's law and describe how you can verify it experimentally.

5.14 Does the relation $V = IR$ apply to resistors that do not obey Ohm's law?

5.15 How does the resistivity ρ of a semiconductor vary with temperature? Explain why semiconductors are used as thermistor. A strip of carbon has a resistivity $1.7\times10^{-8}\,\Omega m$ at 30°C. Calculate its temperature coefficient of resistance if it has a resistivity of $1.5\times10^{-8}\,\Omega m$ at 100°C.

5.16 Find the equivalent resistance of the network of identical resistors, each having R as shown in Figure 5.20.

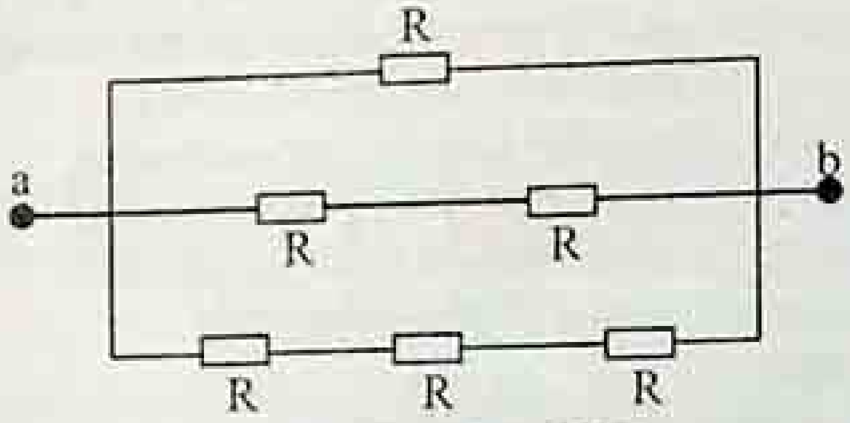


Fig. 5.20: Problem 5.16

5.17 Three 100Ω resistors are connected as shown in Fig. 5.21. The maximum power that can be dissipated in any one of the resistors is $25W$. (i) what is the maximum voltage that can be applied to the terminals a and b? (ii) for the voltage determined in (i), what is the power dissipated in each resistor? What is the total power dissipated?

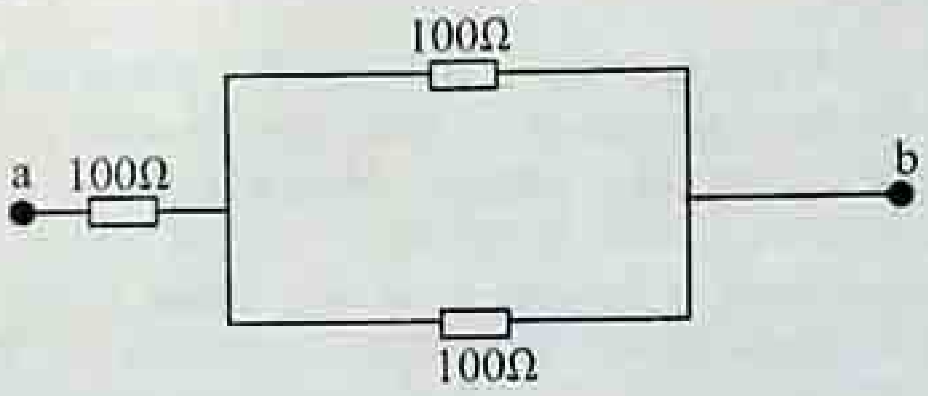


Fig. 5.21: Problem 5.17

5.18 A battery whose emf is $6.0V$ and whose internal resistance is 0.50Ω is connected to a circuit whose equivalent resistance is 11.6Ω. What is the terminal voltage of the battery?

CHAPTER 6
MAGNETISM

6.0 Introduction

In this chapter we shall encounter our third type of field. We have seen gravitational and electric fields previously. This third type of field is the magnetic field. However, we shall see in later chapters that magnetism and electricity are closely related. But this relationship was not discovered until the nineteenth century. The history of magnetism dates back to ancient civilization. It was in Magnesia, a region in Asia Minor, where lodestone was found. It was observed that these lodestones attract objects made of iron.

6.1 Magnets and Magnetic Fields

Magnets have two regions where the magnetic effect is strongest. These regions usually near the ends of the magnet are known as poles. When a bar magnet is dipped into a box of iron fillings, more iron fillings are observed to cling to the magnetic poles. This is illustrated in Figure 6.1.

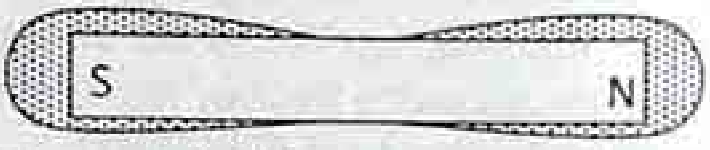


Fig. 6.1: Attraction of iron fillings by a straight bar magnet; it shows greater attraction at the poles

If the same bar magnet is suspended by a fibre as shown in Figure 6.2, a particular pole of the bar will always point towards the north of the earth. This pole is called the North Pole while the other end or pole is called the South Pole. This means that a bar magnet that is free to rotate about a vertical axis will always point in a definite direction. This is the principle of a simple compass.

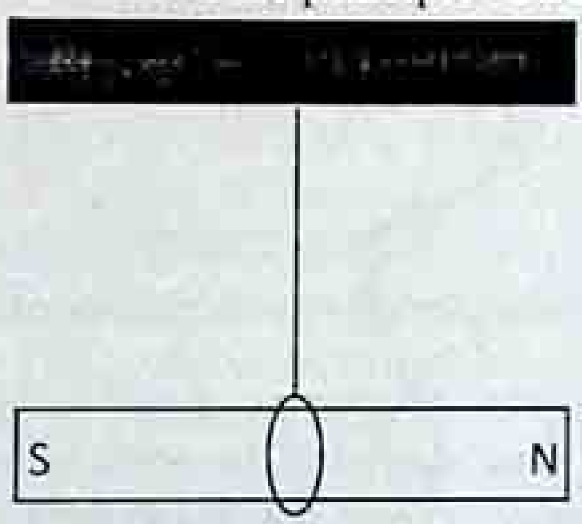


Fig. 6.2: The north-south direction of a freely suspended bar magnet

It is known that by the eleventh century, Chinese navigators were using this principle as a navigational aid. A compass consists of a steel needle which has been magnetized and mounted to rotate freely about a vertical axis. If there are no other magnets around, the needle points in the north-south direction.

It is a well-known fact that the north pole of two magnets repel each other while the south pole of one magnet attracts the north pole of another. Of course, the south pole of one repels the south pole of the other. Figure 6.3 illustrates the 'like poles repel and unlike poles attract' rule. We have seen from above that the north pole of a freely rotating bar magnet is attracted to the north of the earth. This means that the north pole of the earth is really the magnetic south pole.

The poles exist in pairs; if a magnet is broken in the middle in an attempt to separate the poles, one finds two new magnets each having a north and a south pole. If one of the new magnets is again broken into two pieces, each piece is again found to have two poles of opposite kind. This process can be repeated several times but a magnetic pole is always accompanied by a pole of opposite kind. It is possible to isolate a positive or negative electric charge but an isolated single magnetic pole referred to as magnetic monopoles have never been observed.

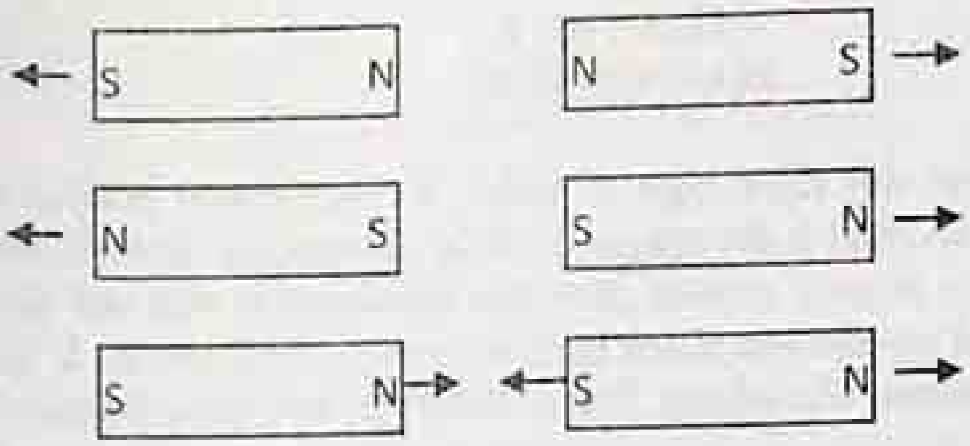


Fig. 6.3: Like poles repel and unlike poles attract

We introduced the concept of electric field in order to understand the electric interaction between two point charges. We assumed that each charge is surrounded by an electric field and the electric force one exerts on the other is described as the interaction between one charge and the electric field of the other. In exactly the same way, we introduce the concept of magnetic field. The magnetic field surrounds a magnet and the force one magnet exerts on another can then be described as the interaction between one magnet and the magnetic field of the other. Just as we drew the electric field lines or lines of force we can also draw the magnetic field lines.

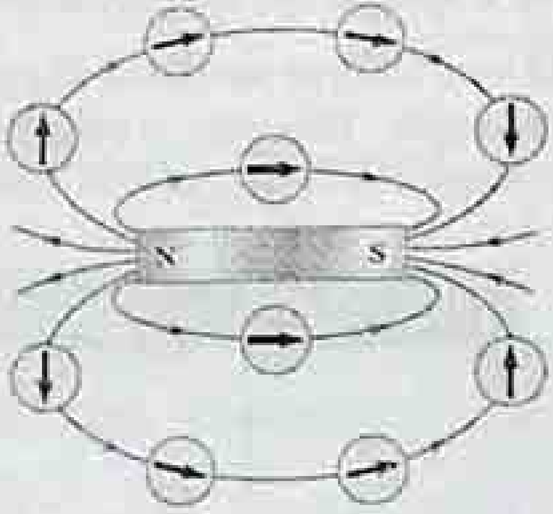


Fig. 6.4: Plotting magnetic field lines of a bar magnet

Magnetic field lines are drawn by means of compass needle placed near the magnet. The direction of the magnetic field at a given point, is the direction the north pole of the compass needle points. In this way, magnetic field lines can be drawn. Figure 6.4 shows how two magnetic field lines around a bar magnet are plotted using compass needles. We notice that the lines originate from the north pole, move towards and enter the south pole. Magnetic field B at any point in space is a vector. The terms magnetic flux density and magnetic induction are often used for B rather than just magnetic field. We shall use these terms interchangeably. The magnetic field lines are drawn so that the tangent to a field line at any point gives the direction of B at that point and they are also drawn so that the number of lines per unit cross-sectional area is proportional to the magnitude of the magnetic field. Typical magnetic field lines are illustrated in Figure 6.5.

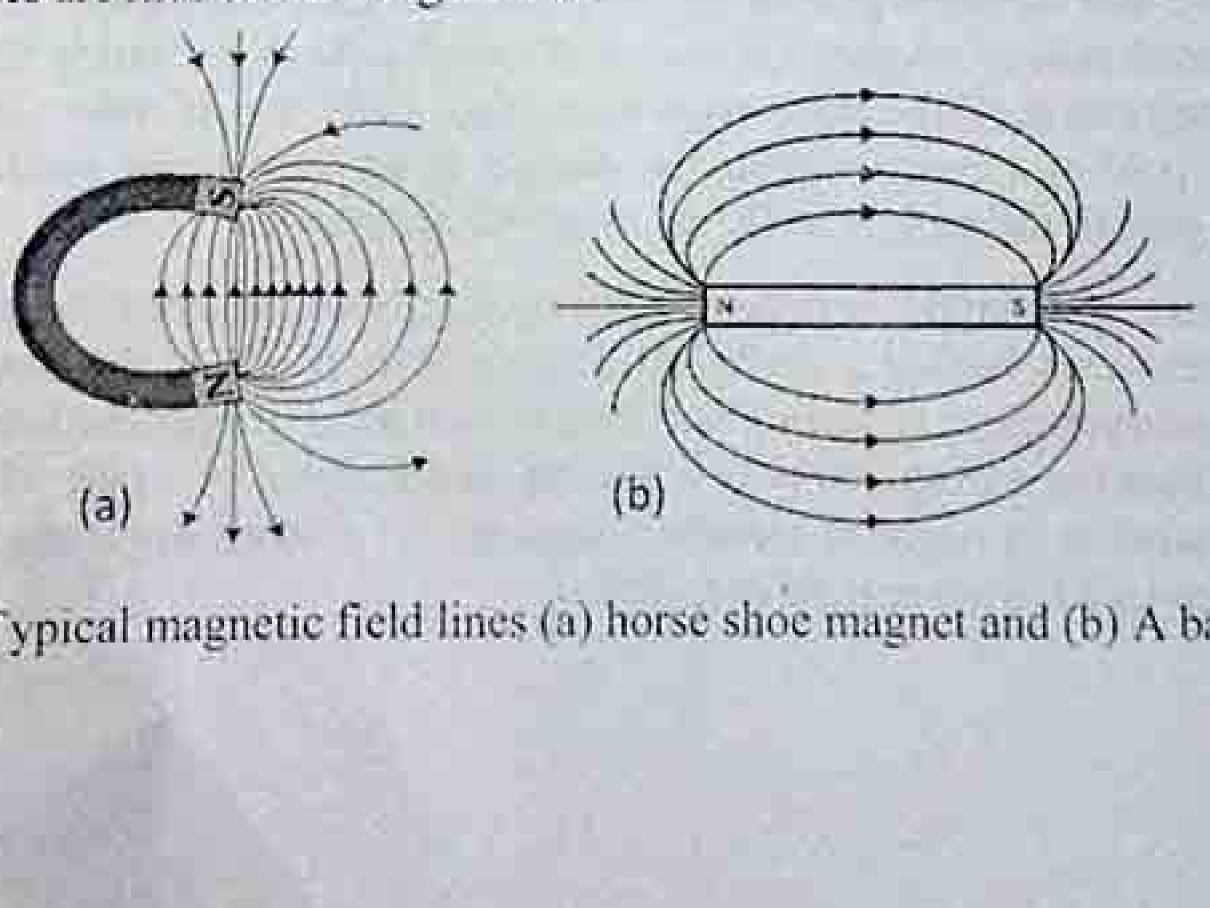


Fig. 6.5: Typical magnetic field lines (a) horse shoe magnet and (b) A bar magnet

6.2 Electric Currents as Sources of Magnetism

Electricity and magnetism were thought to be distinct and unrelated, but many attempts during the eighteenth century were made to find a connection between electricity and magnetism. Hans Christian Oersted (1771-1851), Danish scientist, was the first to discover a significant connection between electricity and magnetism. Oersted discovered in 1820 that electric currents produce magnetic field. This can easily be demonstrated by the experiment shown schematically in Figure 6.6(a). If the switch is open (that is, there is no current flowing), the compass needles point to the north.However, when the switch is closed, the needles line up as shown in Figure 6.6(a).

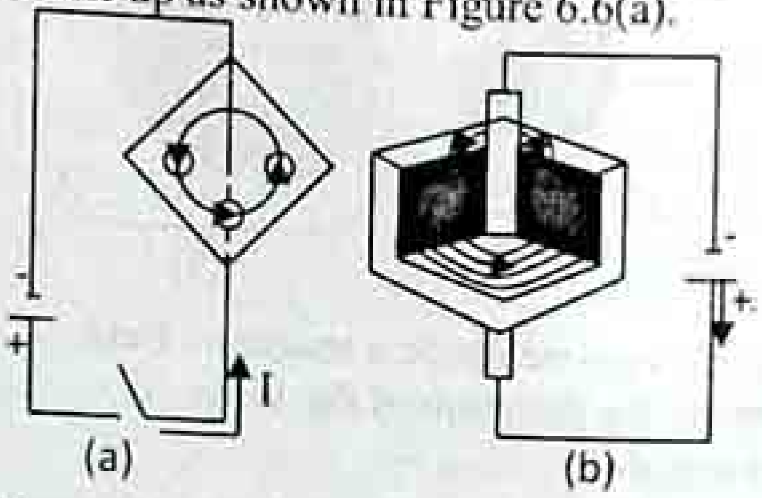


Fig. 6.6: Magnetic field lines due to a wire carrying current I

We know from section 6.1 that only a magnetic field can deflect the compass needle. Thus, when there is current in the wire, there is a magnetic field around the wire. Therefore, electric current produces magnetic field. This is the connection that was long sought after. The compass needles in Figure 6.6(a) line up so that they are tangent to a circle drawn around the wire with the wire at the centre. This means that the magnetic field lines of a wire carrying current are in the form of circles around the circle. These concentric circles of magnetic field lines are shown in Figure 6.6(b). The direction of these lines is indicated by the north pole of the compass. The symbols **O** and **X** represent the direction of the magnetic field coming out of the page towards the recorder and away from the reader into the page, respectively. The symbols represent the tip and tail of an arrow.

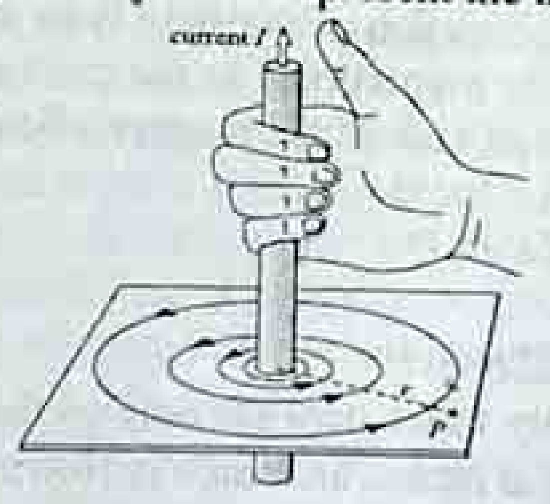


Fig. 6.7: The right hand rule

Now let us learn a simple way of remembering the direction of magnetic field lines around wires carrying current. It is called right hand rule. The rule states that if one grasp the wire with the right hand in such a way that the thumb is in the direction of the current, then the fingers will encircle the wire in the direction of the magnetic field lines. This is shown in Figure 6.7..

6.3 Magnetic Forces on Wire Carrying Currents

We saw in the last section that wires carrying currents produce magnetic fields around them. In 1821, Michael Faraday discovered that, when a wire carrying current is placed in a magnetic field, a force is exerted on the wire. Now let us look at this force in detail. Figure 6.8(a) shows a straight wire placed between the pole pieces of a permanent magnet. When current flows in the wire, a force F is exerted on the wire. The wire is deflected upwards as shown in the figure (Figure 6.8(a)); the deflection is not towards one or the other pole of the magnet. Instead the force is perpendicular to both the magnetic field B and to the current I. If we reverse the direction of the magnetic field, (That is by turning the magnet over, thereby interchanging the north and south poles), the deflection will be downwards.

Now if we reverse the direction of current flow, the force is again in the opposite direction. Thus, reversing either the magnetic field or the direction of current flow reverses the direction of the force on the wire.

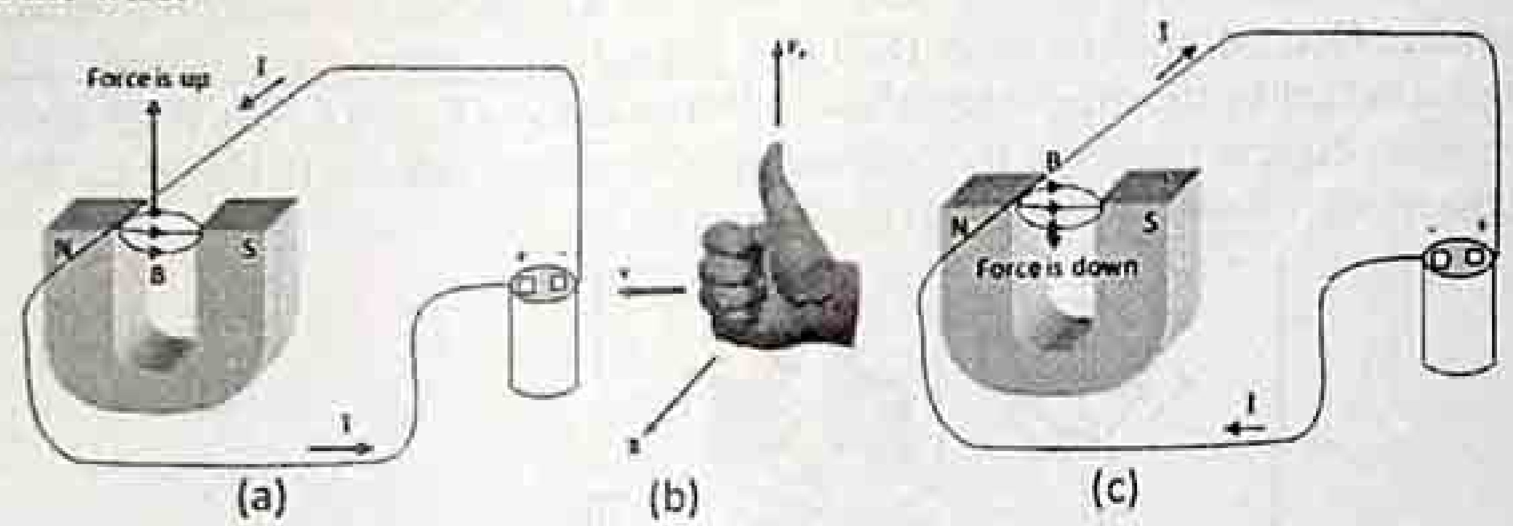


Fig 6.7: (a) Force on a current carrying wire inside a magnetic field
(b) The right-hand rule to determine the direction of the force
(c) Direction of the current is changed.

The direction of the force on the wire is determined by the right hand rule shown in Figure 6.8(b). Orient your right hand until your stretched fingers can point in the direction of the current I, and when you bend your fingers they point in the direction of the magnetic field lines B. Then your outstretched thumb will point in the direction of the force F on the wire. Trying this rule in Figure 6.8(a), it is clear how the force pushes upward. In using this rule, we first establish the plane formed by the current and the magnetic field lines; the force as we have seen is perpendicular to the plane. Figure 6.8(c) shows that the direction of the force changes when the direction of the current is changed.

We now know how to determine the direction of the force exerted on a wire carrying current I inside a magnetic field B. Hence, we wish to find the magnitude of the force which will lead to the precise definition of the magnetic field. Consider Figure 6.9 which shows a current-carrying wire of length L inside a uniform magnetic field B. The direction of the field is as shown. It is found experimentally that the magnitude of the force is directly proportional to the current I in the wire, to the length L of the wire in the magnetic field, to the magnitude of the magnetic field B and to the sine of the angle θ between the wire and the magnetic field. That is

$$F \propto ILB \sin\theta$$

In an equation form, this becomes

$$F = KILB \sin\theta \qquad (6.1a)$$

where K is the constant of proportionality which depends only on the choice of units. In SI units, current is measured in Amperes, length in metres and force in Newtons.

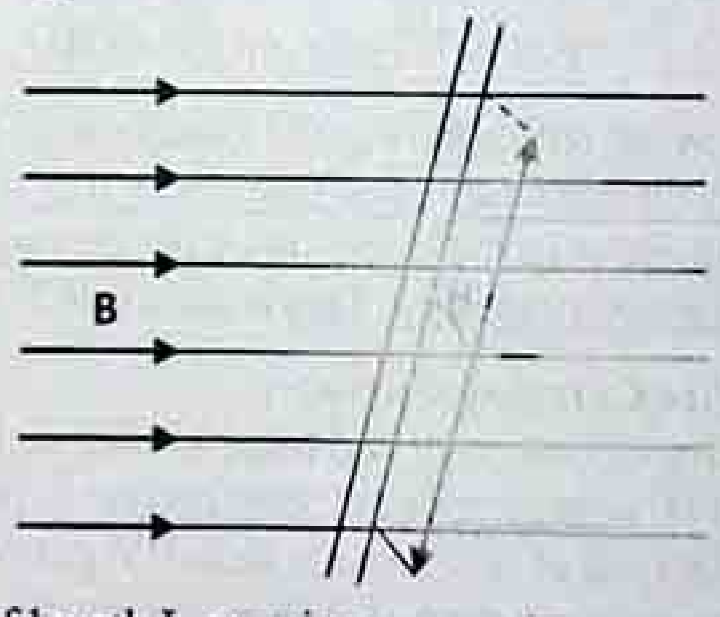


Fig. 6.9: A piece of wire of length L carrying current I in a magnetic field B.

Let us choose the unit of the magnetic field such that the proportionality constant K is precisely 1. In other words, the unit of B is $N/A(m)$. This is known as Tesla (T). This means that $1T = 1N/A(m)$. In older text books you will find that the unit of the magnetic field is Weber per metre square

(Wb/m^2). But $1T = 1Wb/m^2$. The unit of the magnetic field in CGS unit is Gauss (G): $1G = 10^{-4}T$. Ofcourse, a magnetic field given in Gauss must always be changed to Tesla. Let us return to equation 6.1(a); with our choice of unit of the magnetic field B, K is equal to 1. Equation 6.1(a) becomes

$$F = ILB\sin\theta \tag{6.1b}$$

If the wire is perpendicular to the field B, the force is then a maximum and we have

$$F_m = ILB \tag{6.2}$$

If the wire is in the same direction as the field, that is the length L is parallel to the field, then the force is zero. We may now define the magnitude of the magnetic field B from equation 6.2 as

$$B = \frac{F_m}{IL} \tag{6.3}$$

where F_m is the maximum force exerted on the wire carrying current I. As stated above, the maximum force is obtained when the wire is perpendicular to the magnetic field.

Equation 6.1(b) is the magnitude of a vector product between two vectors, namely B and L. Hence, Equation 6.1(b) may be in a vector form. Thus, we have

$$F = IL \times B \tag{6.4}$$

where L is a vector whose magnitude is the length of the wire and its direction is along the wire in the direction of the current. Equation 6.4 holds for a uniform magnetic field and a straight line segment L. However, if the wire is not a straight line, it can be divided into dL segments that are straight. The Equation 6.4 can be written in a differential form. Thus, we have

$$dF = IdL \times B \tag{6.5}$$

where dF is the infinitesimal force acting on the segment dL of the wire. Therefore, the total force on the wire is the vector sum of all such dF which can be found by integration.

Example 6.1

Calculate the magnetic force on a 240m length of wire stretched between two towers carrying a 150A current. The earth's magnetic field of $5 \times 10^{-5}T$ makes an angle of 60° with the wire.

Solution

From equation 6.1b we obtain $F = ILB\sin\theta = 150\,A \times 240\,m \times (5 \times 10^{-5}T)\sin 60° = 1.56N$

6.4 Forces on Moving Electrical Charges in a Magnetic Field

Experimentally, it is observed that when a charge q has a velocity, v in a magnetic field, there is a force on it that depends on q and on the magnitude and direction of the velocity. Let us assume that we know the direction of the magnetic field B at a point in space from a measurement with a compass. Experiment with various charges moving with various velocities at such a point give the following results for the magnetic force.

(i) The force is proportional to the charge q. The force on a negative charge is in the direction opposite that on a positive charge with the same velocity.

(ii) The force is proportional to the speed v.

(iii) The force is perpendicular to both the magnetic field and the velocity.

(iv) The forces are proportional to $\sin\theta$ where θ is an angle between the velocity and the magnetic field, B. If v is parallel or anti-parallel to B, the force is zero.

The experimental results can be summarized as follows. When a charge moves with velocity v in a magnetic field B, the magnetic force F on the charge is

$$F = qv \times B$$

Since F is perpendicular to both v and B, it is perpendicular to the plane defined by these two vectors. We now attempt to derive the above equation.

We saw in the last section 6.3 that a current-carrying wire experiences a force when placed in a magnetic field. The force exerted on the wire is expressed by equation 6.4.

Since the current I is the motion of charged particles in the wire, it is natural to expect that the magnetic field acts directly on individual charged particles. Suppose we have N charged particles, each with charge q, passing through a given point at time t, then the current $I = Nq/t$. Each particle moves with a velocity v and t is the time it takes a charged particle to travel the distance L in the magnetic field B, then $L = vt$. Hence, equation 6.4 becomes.

$$F = \frac{Nq}{t}(vt) \times B = Nqv \times B \tag{6.6}$$

Equation 6.6 is the force on the N charged particles all moving with velocity v. Therefore, the force on an individual charged particle is

$$F_1 = F/N = qv \times B \tag{6.7}$$

The magnitude of this force is

$$F_1 = qvB\sin\theta \tag{6.8}$$

where θ is the angle between the velocity and the magnetic field. It is clear from equation 6.8 that if the velocity of the particle is perpendicular to the field, the force exerted on the particle is maximum. On the other hand, the force experienced by the particle is zero if its velocity is parallel to the magnetic field.

The direction of the force is perpendicular to the magnetic field B and to the velocity v of the particle. It is determined by the use of the right hand rule. The outstretched fingers point in the direction of motion of the positive charge, when you bend your fingers they point in the direction of the field and then the thumb points in the direction of the force.

Since the force is perpendicular to the velocity v, the force has no component in the direction of motion of the particle. This means that the force does no work on the charge. The force merely alters the direction of the velocity. Let us consider a positively charged particle ($+q$), shown in Figure 6.10, travelling with a velocity v and entering a uniform magnetic field B. The field is coming out of the page. The magnetic force is perpendicular to both the magnetic field and the velocity; it therefore merely alters the direction of the velocity without changing its magnitude. Thus, the particle moves with a constant magnitude (v, B and θ as in equation 6.8). The part of such a particle is circular.

So far we have limited our discussions on the force on a moving positive charge. Actually the right hand rule given above is for a positive charge. Equation 6.7 is valid for negative charges also. In this case when q is negative, the direction of the force is opposite to that on the positive charge. To find the direction of the force on a negative charge, we first consider it to be a positive charge and find the force. This same force, reversed in direction is the force on the negative charge.

Figure 6.11 shows an electron with a velocity v entering a uniform magnetic field B. The direction of the force is exactly opposite the force on a positively charged particle travelling in the same direction.

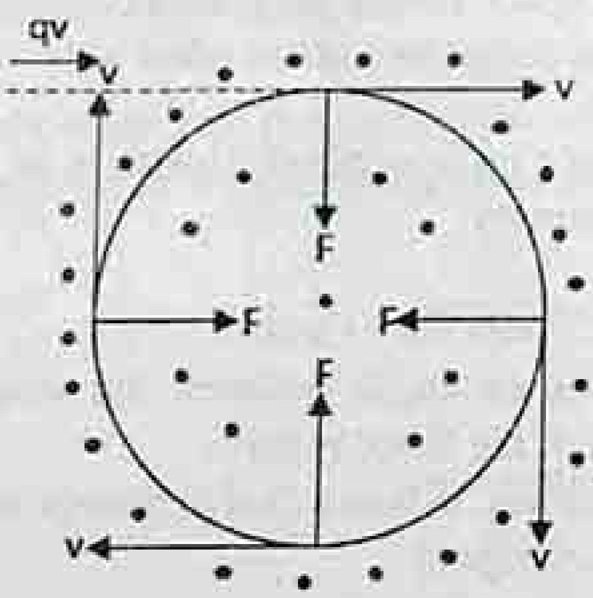


Fig. 6.10: A charged particle entering a uniform magnetic field.

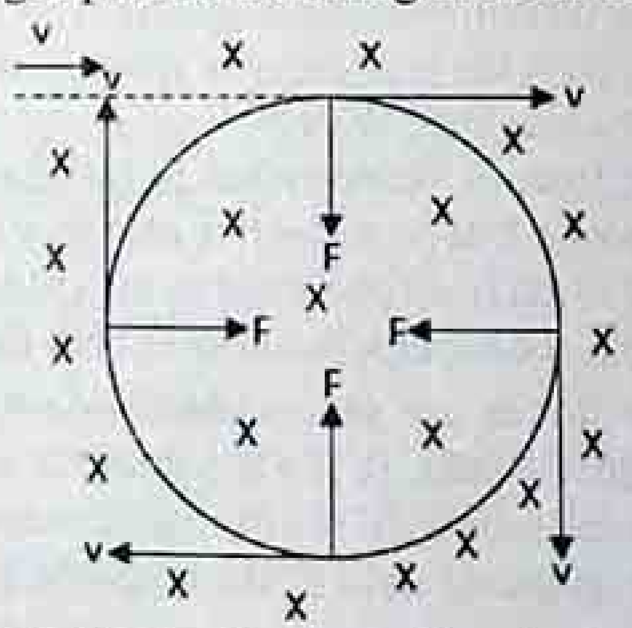


Fig. 6.11: An electron with velocity v inside a magnetic field B which is into the page

Example 6.2

An electron experiences a force $F = (3.2i - 2.7j) \times 10^{-13} N$ when passing through a magnetic field $B = 0.72kT$. What is the velocity of the electron?

Solution

Let us write the velocity of the electron in its component forms. Thus,

$v = v_x i + v_y j + v_z k$ and $F = F_x i + F_y j + F_z k$

From equation 6.7 we obtain, $F = qv \times B$

$$F_x i + F_y j + F_x k - q(v_x i + v_y j + v_z k) \times (B_x i + B_y j + B_z k)$$

But $F_z = 0$, $B_x = 0$ and $B_y = 0$

Hence,

$$F_x i + F_y j = q(v_x i + v_y j + v_z k) \times (0i + 0j + B_z k)$$

$$F_x i + F_y j = q(v_y B_k i - v_x B_k j)$$

$$F_x = qv_y B_k \text{ and } F_y = -qv_x B_k$$

$$v_y = F_x / qB_k$$

Therefore, $v_y = \dfrac{3.2 \times 10^{-13}}{1.15 \times 10^{-19}} ms^{-1} = 2.78 \times 10^6 ms^{-1}$

Similarly, $v_x = -F_y / qB_k$

$$v_x = -\frac{(-2.7 \times 10^{-13})}{1.15 \times 10^{-19}} ms^{-1} = 2.34 \times 10^6 ms^{-1}$$

$$v = (2.34i + 2.78j) \times 10^6 ms^{-1}$$

Example 6.3

A proton of charge q is accelerated through a potential difference of 100 volts and then enters a region where it is moving perpendicular to a magnetic field $B = 0.20T$. Find the radius of the circular path in which it will travel.

Solution

The energy acquired by a proton of charge q when moved through a potential difference V is given by $U = qV$. This energy is equal to the kinetic energy acquired by the proton. Hence,

$$qV = \tfrac{1}{2} mv^2$$

where m is the mass of the proton.

The velocity of the proton just before it enters the field is

$$v = \sqrt{\frac{2qV}{m}} = \sqrt{\frac{(2)(1.6 \times 10^{-19})(100)}{1.67 \times 10^{-27}}} = 1.38 \times 10^5 \, m/s$$

Since the velocity is perpendicular to the field, the force the proton experiences is

$$F = vBq$$

But this force maintains the proton in a circular orbit. Hence, we obtain

$$\frac{mv^2}{R} = F = qvB$$

Therefore, $R = \dfrac{mv}{qB} = \dfrac{(1.67 \times 10^{-27} kg)(1.38 \times 10^5 ms^{-1})}{(0.20T)(1.6 \times 10^{-19} C)} = 0.71cm$

Example 6.4

An electron enters a uniform magnetic field $B = 0.20T$ at a 30° angle to B. Determine the radius and pitch of the electron's helical path assuming its speed v_o is $3.0 \times 10^7 \, m/s$.

Solution

The situation is shown in Figure 6.12(a); the field is along the x-axis.
Let us resolve the velocity into components, v_x or parallel component and v_y or perpendicular component to the magnetic field.

$$v_x = v_o \cos 30° = 2.598 \times 10^7\ m/s \text{ and } v_y = v_o \sin 30° = 1.5 \times 10^7\ m/s$$

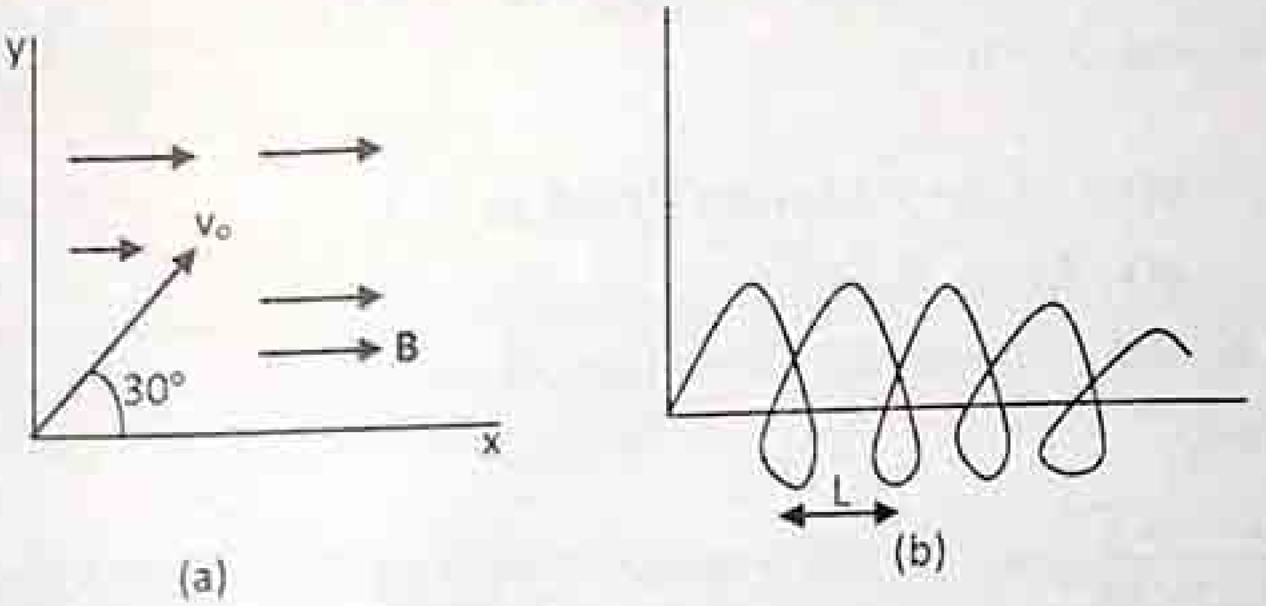


Fig. 6.12: Example 6.4

Since the field is in the x-direction, the x-component of the velocity is unaffected by the field. The perpendicular component (v_y) of the velocity is affected by the magnetic field. The particle will go into circular orbit because of the interaction between v_y and B. Therefore,

$$\frac{mv_y^2}{R} = qv_y B$$

Or $$R = \frac{mv_y^2}{qB} = \frac{9.1\times10^{-31}\,kg\times(3.0\times10^7)^2}{(1.6\times10^{-19})(0.20)} = 1.71\times10^{-15}\,m$$

The time it takes to complete one circle is $t = \frac{2\pi R}{v_y} = \frac{2\pi m}{qB} = \frac{2\pi\times9.1\times10^{-31}}{1.6\times10^{-19}\times0.20} = 7.14\times10^{-12}\,s$

During this time the electron will travel with $v_x = 2.598\times10^7\,ms^{-1}$ parallel to the x-axis therefore forming a helical path as shown in Figure 6.12(b). The pitch is the distance between loops. Hence,

$$L = v_x t = (2.598\times10^7)(7.14\times10^{-12}) = 1.85\times10^{-4}\,m$$

6.5 Hall Effect

The Hall Effect experiment which was reported in 1879 by E.H. hall enables us to determine the sign of a charge carrier. Suppose we have a flat conductor strip shown in Figure 6.13(a) and a magnetic field is set up at right angles to the strip. Suppose the current is of positive charges as shown in Figure 6.13(a), the magnetic field exerts an upward force on the positive charges.

The positive charges tend to move nearer point M than point N on the surface of the strip. Thus, there is a potential difference between points M and N. This continues to build up until the electric field it produces exerts a force on the moving charges that is equal and opposite to the magnetic force. That is

$$qE = qvB$$

or $$v = \frac{E}{B} \tag{6.9}$$

When equation 6.9 holds, the charges will now travel through strip without deflection.
The effect described above is known as Hall Effect named after E.H. Hall, the first scientist to observe it. The potential difference across the strip V_{MN} is called the Hall Emf. If the charge carrier is positive as in Figure 6.13(a), then point M is at a higher potential than point N. It is clear that when a current of positive charges is moving from left to right as shown in Figure 6.13(a), the upper plate (point M) is at a higher potential than the lower plate (point N).

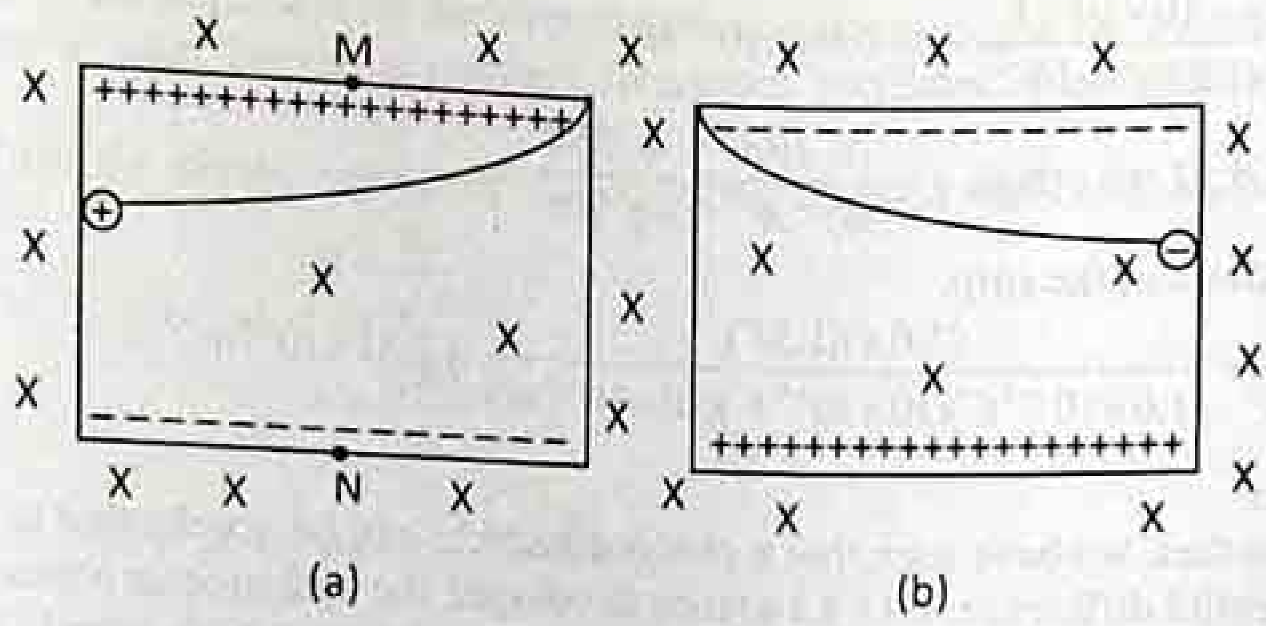


Fig. 6.13: The Hall Effect

However, the motion of positive charges from left to right is equivalent to the movement of negative charges from right to left. This is shown in Figure 6.13(b). The negative charges are deflected toward the upper plate making it more negative than the lower plate. The upper edge is therefore negatively charged and the lower edge is positively charged. There is a Hall voltage V_{MN} but this time point N is at a higher potential than point M. The direction of the Hall Emf reveals that it is negatively charged particles that move in most conductors.

We see that, because of the deflections of the carriers by the magnetic field, electric field is established between the lower and the upper edges of the strip. The Hall e.m.f is related to this electric field as follows:

$$V_{MN} = -\int E \cdot dy = Ed$$

where d is the width of the strip.

From equation 6.9 we know that, $E = vB$

Hence, $V_{MN} = vBd$ (6.10)

The Hall e.m.f (V_{MN}) in Equation 6.10 can be measured, and knowing the field B and width of strip d, we can determine the drift speed of the charge carriers. For positive charges moving from left to right (Figure 6.13(a)), as agreed above, point M is at a higher potential than N, but if current really consists of negative charges moving from right to left (Figure 6.13(b)), then v is negative in equation 6.10 and N is at a higher potential than M.

Current I can be expressed in terms of drift velocity v, area A and number of charges per unit volume n. Thus, $I = qnAv$ (6.11)

or

$$v = \frac{I}{nqA}$$

Substituting this in equation 6.9 we obtain

$$v_{MN} = \frac{1}{nq}\frac{IBd}{A} \qquad (6.12)$$

The factor $1/qn$ is called the Hall coefficient. All the quantities in equation 6.12 are measurable except n, it is therefore used to find the number of charge carriers per unit volume.

Example 6.5

In a Hall effect experiment, a current of $3.0A$ lengthwise in a conductor $1.0cm$ wide, $4.0cm$ long and $10\mu m$ thick produced a transverse Hall e.m.f. of $10\mu V$ when a magnetic field of $1.5T$ passes perpendicularly through the thin conductor. Find (a) the drift velocity of the charge carriers and (b) the number of charge carriers per unit volume.

Solution

From Equation 6.10 we have $V_{MN} = vBd$

or $$v = \frac{v_{MN}}{Bd} = \frac{10\times10^{-6}V}{(1.5T)(1\times10^{-2}m)} = 0.66\times10^{-3}m$$

And from equation 6.12 we obtain $v = \frac{1}{nq}\frac{IBd}{A} = \frac{1}{nq}\frac{IBd}{dh}$

where h is the thickness of the strip.

$$n = \frac{IB}{qv_{MN}h} = \frac{(3.0A)(1.5T)}{(1.6\times10^{-19}C)(10\times10^{-6}V)(10\times10^{-6}m)} = 2.81\times10^{29}m^{-3}$$

6.6 Cyclotrons

From our earlier studies, we have seen that a charged particle can be accelerated to a high speed if it falls through a potential difference. E.O. Lawrence developed the cyclotron as a means of accelerating charged particles such as protons, deuterons and helium nuclei to a high speed. These high speed particles are used as projectiles in high-energy nuclear experiments. The cyclotron shown in Figure 6.14 consists of two hollow metal containers held between the pole faces of a large and powerful magnetic field. The two hollow metals are known as 'dees' because they are more or less D-shaped. A positively charged particle is introduced into the cyclotron at point O; the edge of the left hand 'dee' is negatively charged at this instant. As a result, the particle is attracted to the left 'dee'. However, once the particle is inside the hollow metal, it does not experience any further force because there is no electric field inside the hollow metal which is perpendicular to and into the page. The field maintains the charged particle in a circular path and returns it to the edge of the 'dee'. By the time the particle is ready to cross the gap, the polarity of the voltage source is reversed and the particle is accelerated across the gap to the right 'dee'. If the potential difference between the edges of the 'dees' is V, the energy of the particle increases by qV each time it crosses the gap. If the particle crosses the gap n times, its energy is nV electron volts, assuming that the particle carries unit charge. This increases the speed of the particle and therefore the radius of curvature of its path. After many revolutions, the charged particle acquires high energy and reaches the outer edge of the cyclotron. Finally, the high speed particle is deflected out of the cyclotron with the help of a bending magnet or by a deflector plate.

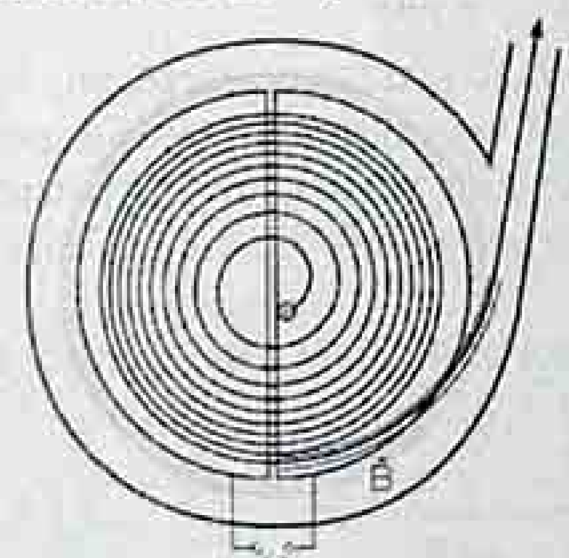


Fig. 6.14: Elements of a Cyclotron

Let us now derive the equations for this device. The centripetal force needed for any of the circular paths is provided by the magnetic force. Thus,

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

Solving for R we obtain $$R = \frac{mv}{qB}$$

Now the time taken to cover half a cycle is

$$T = \frac{\pi R}{v}\frac{\pi m}{qB} \qquad (6.13)$$

Equation 6.13 gives the time the polarity of the voltage must be changed since this is the time for the particle to move through one 'dee'. It is interesting to note that this time (equation 6.13) does not depend upon how fast the particle is moving or the radius of its circular path. This is so because the increase in speed of the particle with each revolution compensates for the fact that the circular path has also increased in size. Hence, a constant frequency oscillator is possible at moderate particle speeds.

We can also deduce the kinetic energy of the emerging charged particle from the above equations.

Since $$R = \frac{mv}{qB},$$

Then $$v = \frac{qBR}{m}$$

$$K.E = \frac{1}{2}mv^2 = \frac{1}{2}\left(\frac{q^2B^2}{m}\right)r^2$$

Example 6.6

A cyclotron for accelerating protons has a magnetic field of $1.5T$ and a maximum radius of $0.5m$.

(a) What is the cyclotron frequency?

(b) Find the kinetic energy of the protons when they emerge.

Solution

(a) The cyclotron period is $T = \frac{2\pi m}{qB}$

Therefore, the frequency $f = \frac{1}{T} = \frac{qB}{2\pi m}$

$$f = \frac{(1.6\times10^{-19}C)(1.5T)}{2\pi(1.67\times10^{-27}kg)} = 2.29\times10^{7}Hz = 22.9MHz$$

(b) The kinetic energy of the emerging proton is

$$K.E = \frac{1}{2}\frac{(1.6\times10^{-19}C)^2(1.5T)^2}{2\pi(1.67\times10^{-27}kg)} = 4.3\times10^{-12}J = 26.9MeV$$

Since $1eV = 1.6\times10^{-19}J$

6.7 Torque on a Current Loop

In section 6.3, we found that a piece of wire of length L carrying current I experiences a force when placed inside a uniform magnetic field, B. In equation 6.4, the force is

$$F = IL\times B$$

When a current I flows in a closed loop of wire that is placed inside a uniform magnetic field, the magnetic field on the loop produces a torque. Many devices, as we shall see in the next section, such as electric motors and moving coil meters make use of this fact.

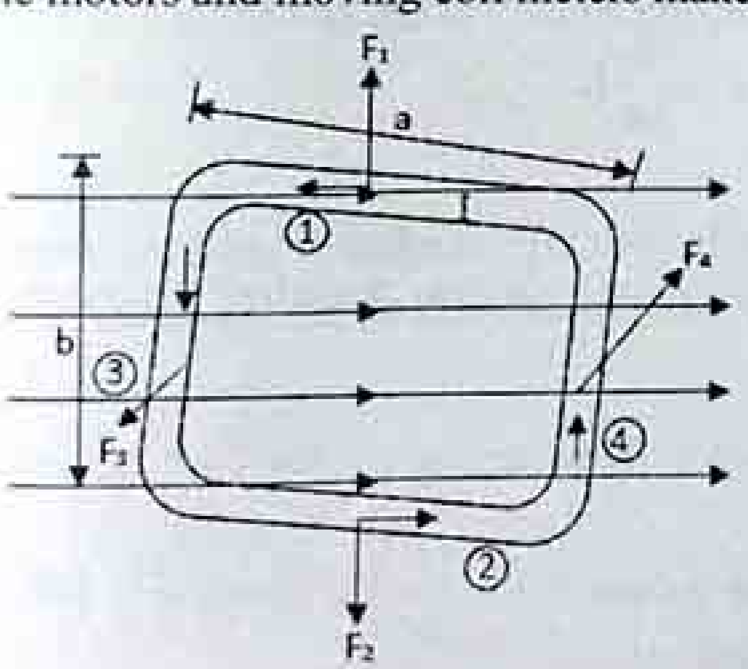


Fig. 6.15: A rectangular loop of wire carrying current I in a horizontal magnetic field

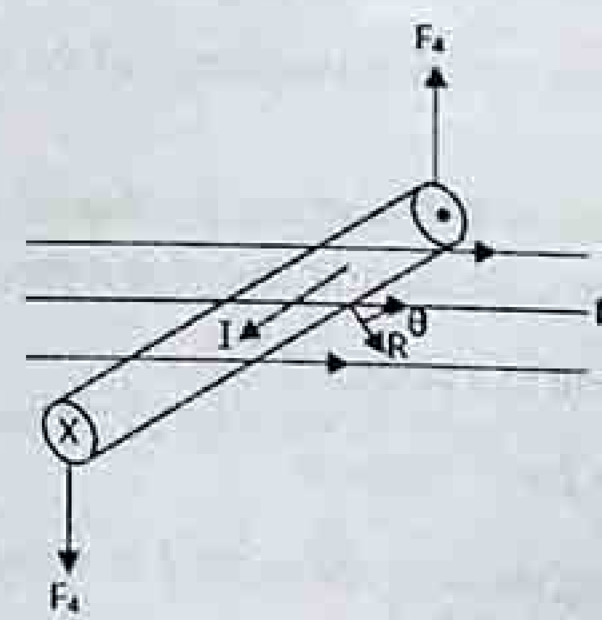


Fig. 6.16: A side view of a rectangular loop in a uniform magnetic field B

Now let us consider a flat rectangular loop of wire of length b and width a as shown in Figure 6.15. The loop is placed in a horizontal magnetic field B; the plane surface of the loop makes an angle θ with the magnetic field B. The electrical wires that carry current in and out of the loop are twisted together so that the magnetic forces on the lead-in wires can be ignored. The loop is suspended so that it is free. For the purpose of determining the forces exerted on the loop by the magnetic field B, we label the side of the rectangle (1) to (4) as shown in Figure 6.15. The force F_1 on side (1) is equal in magnitude to force F_2 on side (2). F_1 and F_2 are oppositely directed (as can be seen by the use of the right hand rule) and have a common line of action. Therefore, the resultant of F_1 and F_2 is zero and since they have a common line of action, the net torque due to these forces is also zero.

F_3 and F_4, the forces on sides (3) and (4) respectively, are equal in magnitude. The common magnitude is IbB since the vertical sides are perpendicular to the magnetic field. The forces are again oppositely directed. The resultant of F_3 and F_4 is zero, so these forces do not give the loop any translational motion. However, they do not have the same line of action; hence they can produce rotation.

Since F_1 and F_2 have a common line of action, it is convenient to choose this line as the axis about which to calculate the torque. Figure 6.16 shows the side view of Figure 6.15. The moment arm of each of these forces is $\frac{a}{2}\sin\theta$; and each force tends to rotate the loop in anticlockwise direction.

Therefore, the net torque is $\tau = \left[\frac{a}{2}\sin\theta IbB\right] + \left[\frac{a}{2}\sin\theta IbB\right] = IBab\sin\theta$

But the area of the loop is A = ab. Substituting this in the above expression we have

$$\tau = IAB\sin\theta \qquad (6.14)$$

The quantity IA in equation 6.14 is called the magnetic dipole moment of the coil. The magnetic dipole moment is a vector denoted by μ. Thus,

$$\mu = IAn \qquad (6.15)$$

where n is a normal unit vector to the surface of the loop. With this definition of μ, equation 6.14 may now be written in a vector form. Thus,

$$\tau = \mu \times B \qquad (6.16)$$

Even though equation 6.16 is derived for a flat rectangular loop, it is valid for flat loops of any shape and about any axis. If the loop contains N coils of wire, then the current in each section is NI. Then equation 6.15 becomes $\mu = NIAn$

Example 6.7

Show that the magnetic dipole moment μ of an electron orbiting the proton of a hydrogen atom is related to the orbital momentum L of the electron by $\mu = \frac{2}{2m}L$.

Solution

From Equation 6.15 we have $\mu = IAn$

But current is the electric charge that passes a given point per unit time, therefore the rotating electron is equivalent to a current $I = \frac{e}{T} = \frac{ev}{2\pi r}$

where $T = 2\pi r/v$ (period)

Substituting the expression for I in equation 6.15 we obtain $\pi = \left(\frac{ev}{2\pi r}\right)(\pi r^2)n = \left(\frac{ev}{2\pi r}\right)(mvr)n$ if m is introduced.

But $mvr = L$, is the angular momentum. Hence, taken $n=1$, $\mu = \left(\frac{e}{2m}\right)L$.

6.8 Galvanometers and Motors

When we discussed ammeters and voltmeters in section 6.4, it was pointed out that the galvanometer forms the main operating mechanism of most ammeters and voltmeters. The galvanometer is a device that makes use of the magnetic torque discussed in section 6.7 to measure currents. As shown in Figure 6.17, the galvanometer consists of a coil of wire wrapped around a cylindrical soft iron core suspended between the curved poles of a magnet. The coil is permitted to rotate about an axis passing through the centre of curvature of the magnet. In this way, the magnetic field B is always perpendicular to the vertical sides of the coil. A spring holds the coil in the centre of the magnet. When the coil is rotated by an angle ϕ with respect to the centre position, the spring exerts a restoring torque τ_S which is approximately proportional to the angle ϕ.

Thus,

$$\tau_S = K\phi$$

where K is the spring constant. If a current now passes through the coil, a magnetic torque $\tau_B = NIAB$ rotates the coil which has N turns or windings. The coil will rotate until the magnetic torque is counter-balanced by the spring torque.

Then we obtain

$$\tau_B = \tau_S$$

$$NIAB = K\phi$$

or

$$I = \frac{K\phi}{NAB} \tag{6.17}$$

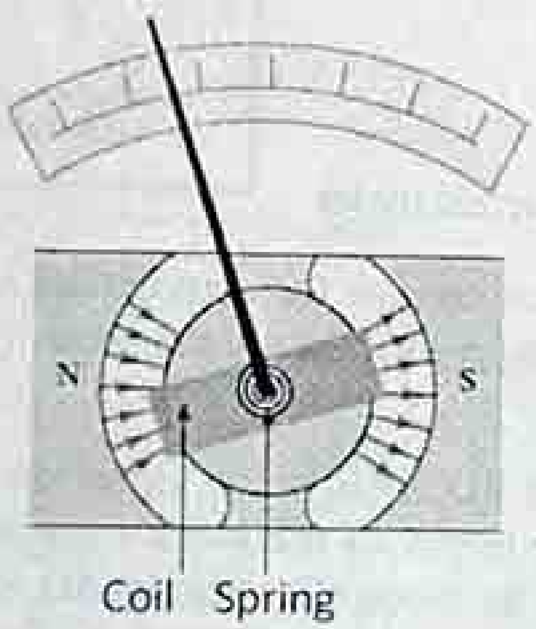


Fig. 6.17: Basic elements of a galvanometer

Thus, the current I that passes through the coil is proportional to the angle of deflection ϕ. Therefore, if a pointer is attached to the device to indicate the angular deflection ϕ, one can calibrate a scale to read the current.

Example 6.8

A galvanometer needle deflects full scale for a $34.0\mu A$ current. What current will give full-scale deflection if the magnetic field weakens to 0.90 of its original value?

Solution

From Equation 6.17 we know that initial value of the current I_0 and initial magnetic field B_0 are related thus: $I_0 = \frac{K\phi}{NA}\frac{1}{B_0}$

Similarly, later values are related, thus, $I = \frac{K\phi}{NA}\frac{1}{B}$

Hence, $\frac{I_0}{I}=\frac{B}{B_0}$ or $I=\frac{I_0B_0}{B}=(34\mu A)\left(\frac{B_0}{0.9B_0}\right)=38\mu A$

An electric motor is a device by which electric energy is transformed into (rotational) mechanical energy. A motor works on the same principle as the galvanometer described above. However, the coil of a motor is larger and is mounted on a large cylinder called armature. Figure 6.18, shows the schematic diagram of an electric motor. The armature is mounted on a shaft which is free to rotate inside a uniform magnetic field provided by a permanent magnet. Since continuous rotation is desired, unlike the galvanometer, the current cannot be led into and out of the rotating coil with fixed wires. Current is passed to the rotating coil by the use of commutators and brushes. The brushes are stationary electrical contacts that rub against the conducting commutators mounted on the motor shaft. When the current is sent through as shown in Figure 6.18, the coil rotates clockwise until its plane is vertical. At that point, however, the current automatically reverses by the switching of the connections as the commutator gaps pass the contacts. This reversal of the current compels the coil to make another half turn, at which point reversal occurs again, and so on. Thus, the current in the coil reverses every half revolution which results in a continuous rotation in one direction.

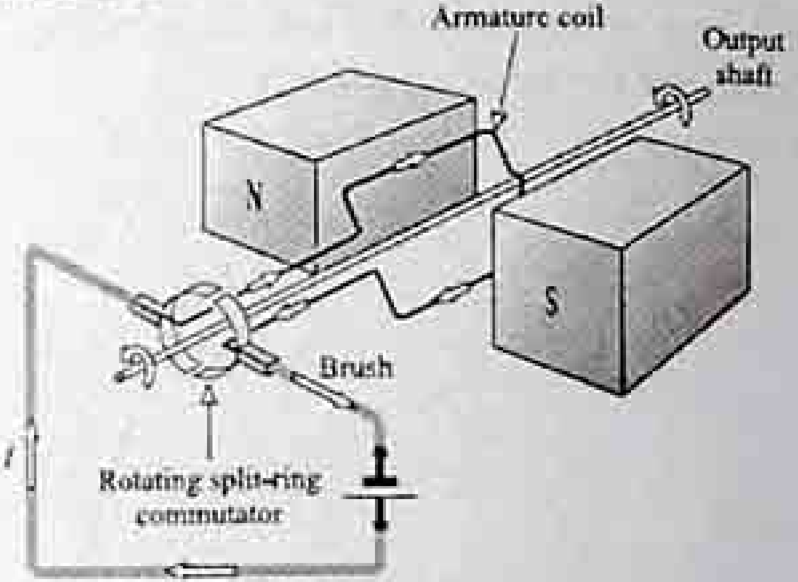


Fig. 6.18: Schematic diagram of an electric motor

We must warn that the design of most motors for commercial applications is far more complicated than that described here, but the general principles are the same.

6.9 The Earth's Magnetism

A compass needle always points in the north-south direction. It must be that the earth itself is surrounded by a magnetic field. The origin of the earth's magnetism is not clearly understood, but broadly speaking, the earth behaves as though it contained a relatively short bar magnet inside it, inclined at a small angle to its axis of rotation, and with its south pole in the northern hemisphere, while its north pole is in the southern hemisphere (Figure 6.19).

We note the following in this subject.

The magnetic elements

The magnetic elements of any given place are the angle of declination, the angle of dip and the horizontal component of the Earth's magnetic field.

(a) Angle of declination

A compass needle pivoted so that it swings in a horizontal plane will point to the magnetic north pole when allowed to move freely.

This is not the same as the geographical north pole, which is the point where the earth's imaginary axis of rotation would emerge from the surface in the northern hemisphere and which can be found from the direction of the Pole star at night or the direction in which a shadow falls at midday.

The angle of declination is the angle between the direction of magnetic North and the direction of geographical North at a particular point.

The angle of declination (or variation) has to be allowed for when navigating. Maps are marked with true North, while a compass used in finding direction will give magnetic North. The correct angle of declination, marked on the map, has to be added or subtracted.

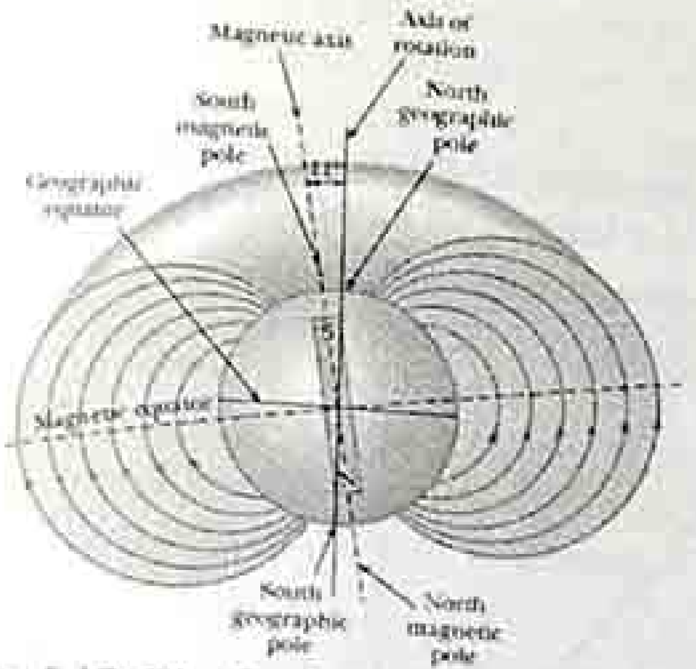


Fig. 6.19: The Earth's Magnetism

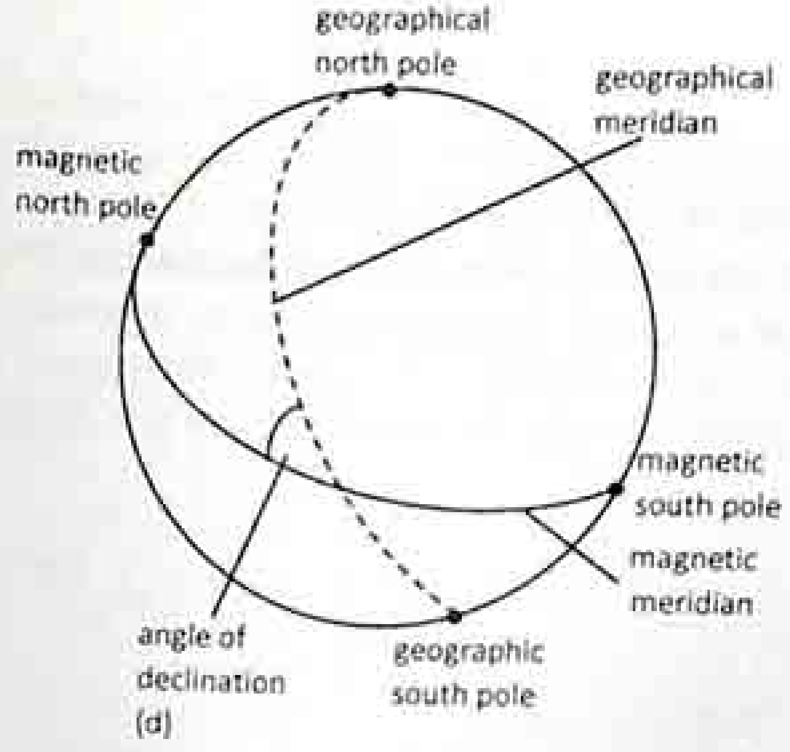


Fig. 6.20: Angle of declination

(b) Angle of dip

The lines of force due to the earth's magnetic field are shown in Figure 6.20. It will be seen that the Earth's magnetic flux lines are not along its surface. They rise up out of the surface in the southern hemisphere and dip down into the surface in the northern hemisphere. A compass needle pivoted so that it can swing in a vertical plane will set itself along the lines of magnetic force due to the earth's magnetism. The angle between the direction of the setting of the compass needle and the horizontalis called the *angle of dip*. (Figure 6.21)

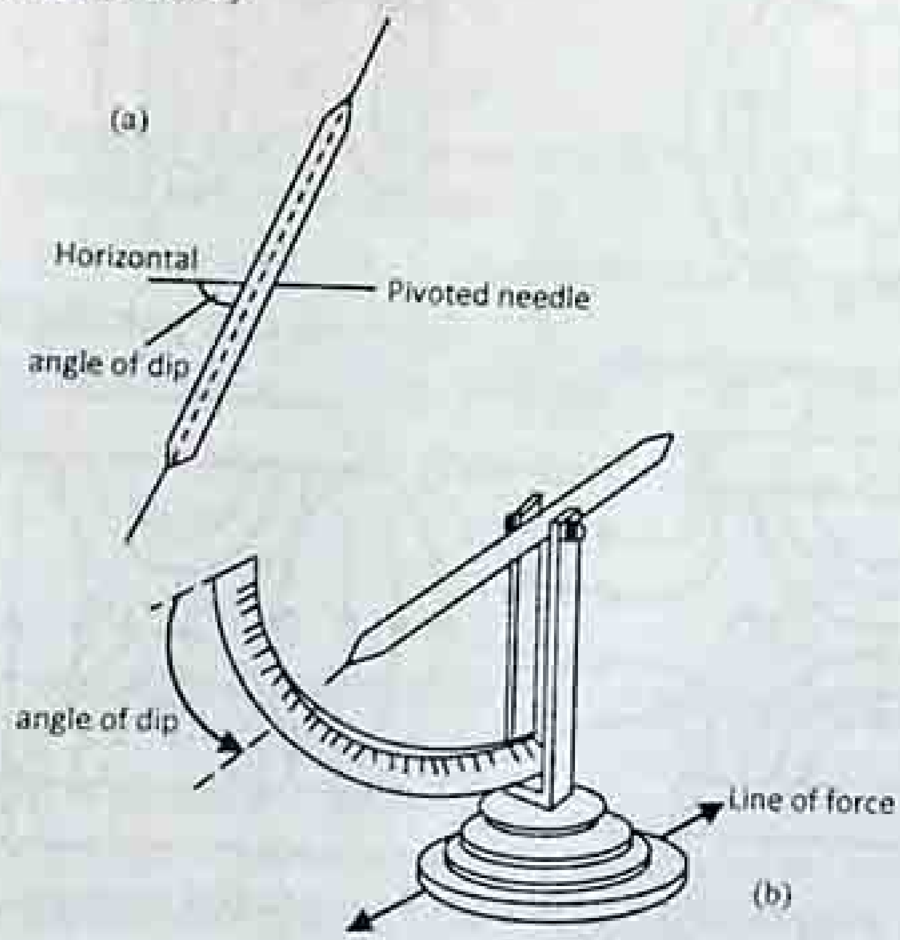


Fig. 6.21: Angle of dip.

Measurement of the angle of dip: The dip circle is an instrument used to measure the angle of dip at a point on the earth's surface. Before being used, the dip circle has to be set in the magnetic meridian. The magnetic meridian at a place is a vertical plane containing the magnetic axis of a free suspended magnet at rest at that place under the action of the earth's field.

The dip circle is turned until the needle sets vertically. The plane of the needle is then at right angles to the magnetic meridian. By rotating the apparatus through 90° from this position, the needle will lie in the plane of the magnetic meridian and the angle of dip may be read directly from the scale.

(c) Resolving the earth's magnetic field into components

The earth's magnetic field has a certain magnitude and makes a certain angle (angle of dip) with the horizontal. The exact values depend on the position on the earth's surface.

If OI represents the total flux (Figure 6.22), then the angle of dip, ϕ will be given by

$$\tan\phi = \frac{\text{vertical component}}{\text{horizontal component}} = \frac{OV}{OH}$$

OH is the horizontal component of the earth's field at the point. The horizontal component of the earth's field at a point is the strength of its magnetic flux acting in a horizontal direction.

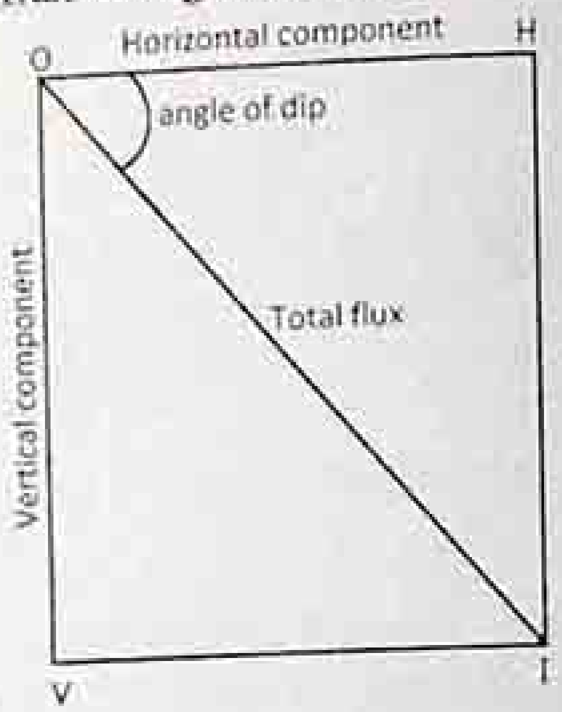
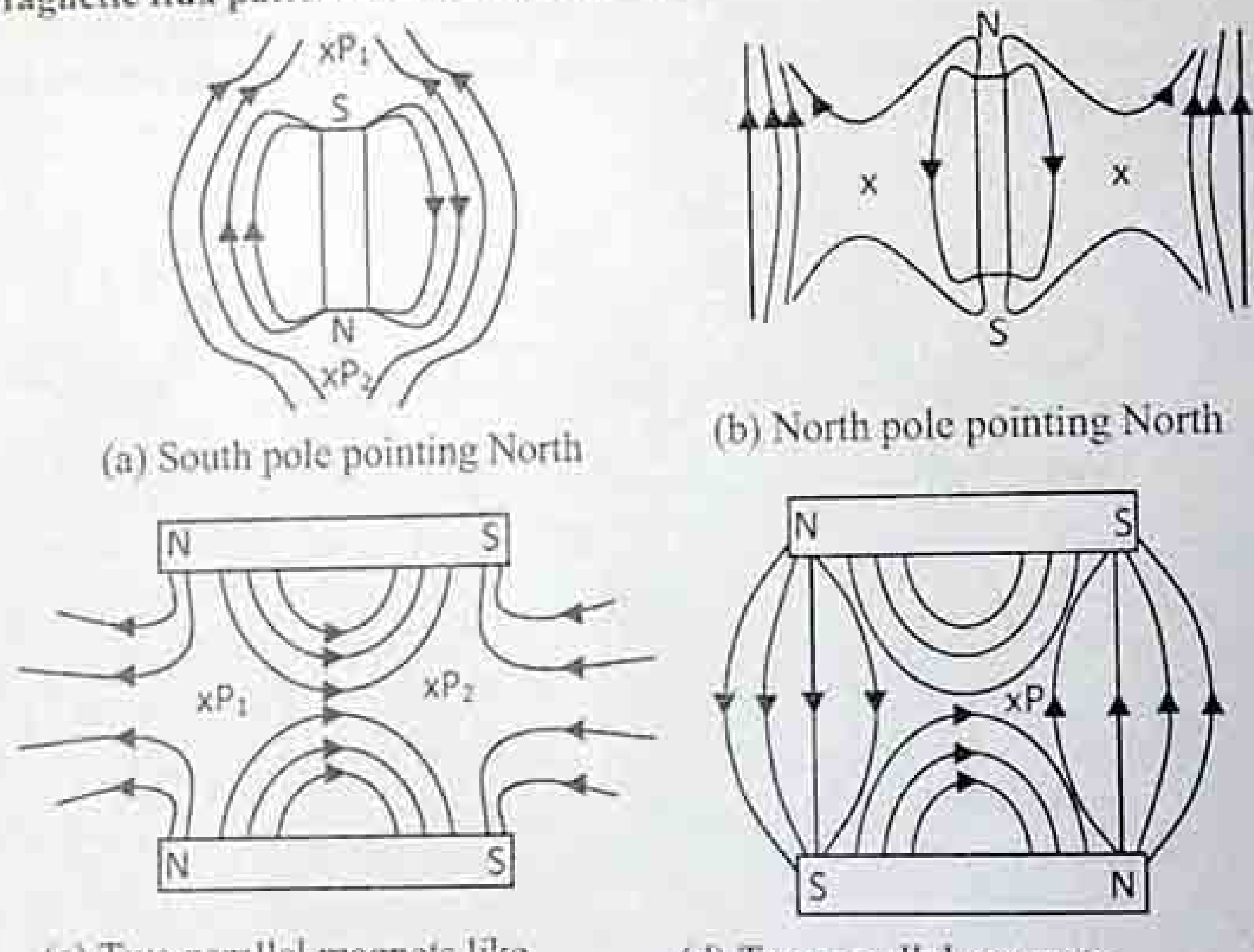


Fig. 6.22: Resolving the earth's magnetic field

6.10 Magnetic flux patterns in the Earth's field



(a) South pole pointing North

(b) North pole pointing North

(c) Two parallel magnets like poles adjacent

(d) Two parallel magnets, unlike poles attract

Fig. 6.23: Magnetic Flux pattern in the earth's field

When a bar margent is placed on a table, the field pattern will be a combination of the field due to the magnet and that due to the earth's horizontal component. The pattern depends on the direction in which the magnet is lying. Figure 6.23 shows the field patterns obtained,

(a) When the axis of the magnet is in the magnetic meridian and its south pole is pointing north,

(b) When the axis of the magnet is in the meridian and its north pole is pointing north.

Neutral points' can also be observed under suitable conditions. In Figure 6.23(a), we observe that along the axis of the magnet, the earth's flux lines and those of the magnet are in opposite directions. While the earth's flux remains constant, that of the magnet is strong near the magnet and weakens rapidly with increasing distance from the poles. At the point marked x in the figure, the horizontal component of the earth's field and the field due to the magnet are exactly equal and opposite. The resultant flux is zero. These points are called neutral points.

A neutral point is defined as a point at which the resultant magnetic flux density is zero.

In Figure 6.23(b), the neutral points are formed on the magnetic equator of the magnet.

Figure 6.23(c) and (d) show the magnetic field around pairs of magnets, placed with like poles together, (c), and with opposite poles together (d). These fields also have neutral points, as shown.

Example 6.9

At a point on the earth's surface, the horizontal component of the earth's main magnetic field is $26\mu T$ and the inclination is $59°$. Find the magnitude of the resultant (total) field at this point.

Solution

The horizontal components is $H = 26\mu T = 26\times 10^{-6} T$

The inclination is $I = 59°$

The resultant field is F

From Fig. 6.22. It follows that $F^2 = H^2 + Z^2$

Where $Z = H \tan I$.

Hence

$$F^2 = (26\times10^{-6})^2 + (4.327\times10^{-5})^2 = 6.76\times10^{-10} + 1.872\times10^{-9} = 2.5482\times10^{-9} T^2$$

$$F = 5.048\times10^{-5} T.$$

Another approach is to say that: $\frac{H}{F} = \cos I$ or $F = H/\cos I$

i.e. $F = 26\times10^{-6}/\cos 59° = 5.048\times10^{-5} T.$

Summary

1. All magnetic fields are caused by charges in motion. When a charge q moves with velocity v in a magnetic field B, it experiences a force. $F = qv\times B$
2. The SI unit of magnetic field is tesla (T). A commonly used unit is the gauss (G), which is related to the tesla by $1T = 10^4 G$.
3. A particle of mass m and charge q moving with speed v in a plane perpendicular to a magnetic field moves in a circular orbit of radius r given by $r = \frac{mv}{qB}$. The cyclotron period $T = \frac{2\pi m}{qB}$ and frequency $f = \frac{qB}{2\pi m}$.
4. A wire of length L carrying current I when placed in a magnetic field B experiences force.
   $F = IL\times B$
5. Torque on a current loop is given by $\tau = \pi \times B$
   where $\mu = NIAn$ and n is a unit vector normal to the surface of the loop.
6. At any given place on the earth, we can define 3 magnetic elements: the angle of declination, the angle of dip, and horizontal component of earth's magnetic field (see text for definitions).

Exercise 6

6.1 Which of the following produces a magnetic field?
A. Static charges **B.** Accelerating charges
C. Decelerating charges **D.** Moving charges at constant speed

6.2 An electron moves in a certain direction in the horizontal plane where a magnetic field is oriented in a particular direction. In this situation, there
A. is one possible direction for the magnetic force on the electron.
B. are infinite possible directions for the magnetic force on the electron.
C. are two possible directions for the magnetic force on the electron.
D. are three possible directions for the magnetic force on the electron.

6.3 What is the force on a wire of length $200cm$ and carrying a current of $10{,}000mA$ when it is placed at right angles in a field of $0.45T$?
A. $0.09N$ B. $0.0N$ C. $4.5N$ D. $9.0N$

6.4 Which of the following is not true concerning the force on electrical charges moving in a magnetic field?
A. The force is proportional to the charge q. B. The force is proportional to the speed v.
C. The force is perpendicular to both the magnetic field and the velocity. D. The force is always zero.

6.5 What is the force per meter of length on a straight wire carrying a $20.5A$ current whenperpendicular to a $1.50T$ magnetic field?
A. $0.75Nm^{-1}$ B. $30.75Nm^{-1}$ C. $7.5Nm^{-1}$ D. $3.075Nm^{-1}$

6.6 The force on an electric charge moving in a magnetic field is
A. independent of the speed of the charge. B. inversely proportional to the charge.
C. directed perpendicular to the velocity of the charge. D. directed along the magnetic field.

6.7 The Hall coefficient is dependent on
A. Magnetic field B. Current C. Number of charges per unit volume D. Area of the strip

6.8 The cyclotron frequency is given by the expression:
A. $\frac{2\mu m}{qR}$ B. $\frac{2\pi m}{qR}$ C. $\frac{qR}{2\pi m}$ D. $\frac{qB}{2\pi m}$

6.9 A coil of 500 turns has area of $5cm^2$. If the current in the coil is $10A$, calculate the magnetic moment of the coil.
A. $2.5Am^2$ B. $5\times10^{-3}Am^2$ C. $0.25Am^2$ D. $25Am^2$

6.10 Angle of Dip is
A. the angle between the direction of the magnetic North and the direction of geographic North.
B. the angle between the earth's magnetic field and the vertical.
C. the angle between the magnetic meridian and magnetic north pole.
D. the angle between the direction of the earth's magnetic flux and the horizontal.

6.11 Calculate the magnetic force on a $240m$ length of wire stretched between two towers carrying a $150A$ current. The earth's magnetic field of $5\times10^{-5}T$ makes an angle of 60° with the wire.

6.12 A wire carries a steady current of $2.0A$. A straight section of the wire, with a length of $0.75m$ along the x-axis, lies within a uniform magnetic field, $B=(1.6k)T$. If the current flows in the +x direction, what is the magnetic force on the section of wire?

6.13 A current $I = 15A$ is directed along the positive x-axis in a wire perpendicular to a magnetic field. The current experiences a magnetic force per unit length of $0.63\,N/m$ in the negative y-direction. Calculate the magnitude and direction of the magnetic field in the region through which the current passes.

6.14 A long wire parallel to the x-axis carries a current of $8.5A$ in the positive x-direction. There is a uniform magnetic field $B = 1.65j\ T$. Find the force per unit length on the wire.

6.15 A long rigid conductor, lying along the x-axis, carries a current of $5.0A$ in the negative x-direction. A magnetic field B is present, given by $B = 3i + 8x^2j$, with x in meters and B in mT. Calculate the force on the $2.0m$ segment of the conductor that lies between x = 1.0 and x = $3.0m$.

6.16 The magnetic field over a certain region is given by $B = (4i - 11j)T$. An electron moves in the field with a velocity $v=(-2i+3j-7k)ms^{-1}$. Write out in unit vector notation, the force exerted on the electron by the magnetic field.

6.17 An electron experiences the greatest force as it travels at $3.9\times10^5 ms^{-1}$ in a magnetic field when it is moving westward. The force is upward and of magnitude $8.2\times10^{-13}N$. What are the magnitude and direction of the magnetic field?

6.18 The electrons in the beam of a television tube have energy of $12 keV$. The tube is oriented so that the electrons move horizontally from magnetic south to magnetic north. The vertical component of the earth's magnetic field points down and has a magnitude of $55 \mu T$. (i) In what direction will the beam deflect? (ii) What is the acceleration of a given electron due to the magnetic field? (iii) How far will the beam deflect in moving $20 cm$ through the television tube?

6.19 A singly charged positive ion has a mass of $3.2 \times 10^{-26} kg$. After being accelerated through a potential difference of $833 V$, the ion enters a magnetic field of $0.92 T$ along a direction perpendicular to the direction of the field. Calculate the radius of the path of the ion.

6.20 A beam of electrons whose kinetic energy is K emerges from a thin-foil window, at the end of an accelerator tube. There is a metal plate a distance l from the window and at right angle to the direction of the emerging beam. Find the minimum magnetic field that we need to apply to prevent the beam from lifting the metal.

CHAPTER 7
SOURCE OF MAGNETIC FIELDS

7.0 Introduction

The fact that electric currents (electric charges in motion) exert magnetic effects was first observed by the Danish scientist H. Oersted in 1820. He found that a compass needle is deflected from its normal north-south orientation by an electric current in a conductor. This was a landmark experiment in the history of physics because it marked the first experimental connection between the sciences of electricity and magnetism which hitherto developed as quite separate subjects. One month after Oersted's discovery, Biot and Savart announced the results of their measurements of the force on a magnet near a long, current-carrying wire and analyzed these results in terms of the magnetic field by each element of the current. Ampere extended these experiments and showed that current elements also experience a force in the presence of a magnetic field and that two currents exert forces on each other.

We will explore in this chapter, methods of calculating the magnetic fields produced by some common current configurations such as a straight wire segment, a long straight wire, a current loop and a solenoid. Next, we will discuss the force between two parallel conductors and finally, discuss Ampere's law, which relates the line integral of the magnetic field around a closed loop to the total current that passes through the loop.

7.1 The Biot-Savart Law

Consider a conductor of arbitrary shape carrying current I as shown in Figure 7.1. According to the Bio-Savart law, the magnetic field dB produced by a current element Idl at point P is given by

$$dB = \left(\frac{\mu_0}{4\pi}\right)\frac{Idl\sin\theta}{r^2} \tag{7.1}$$

where r is the distance from the point P to the element and θ is the angle between the element and the line joining it to P. The constant $\mu_0 = 4\pi\times10^{-7}\,Hm^{-1}$ is permeability of free space. The formula cannot be proved directly since we cannot experiment with infinitesimally small conductor, however, deductions from large practical conductors are found to be true.

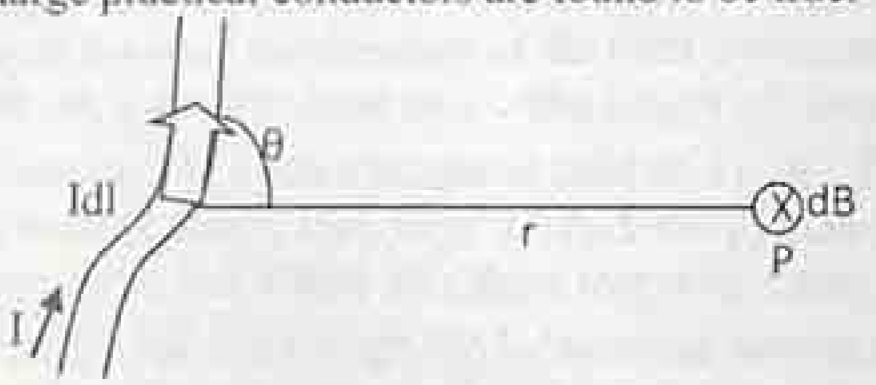


Fig. 7.1: Biot and Savart Law

The direction of dB is that of the vector $dl\times r$, that is at right angles to the plane of the page. We can obtain the resultant field B at point P by summing (integrating) the contributions of dB from all current elements Idl that make up the arbitrary current distribution. We will now illustrate the use of the Biot-Savart law to compute magnetic fields for simple circuit geometries. For such geometries, the integration procedure will present little difficulty.

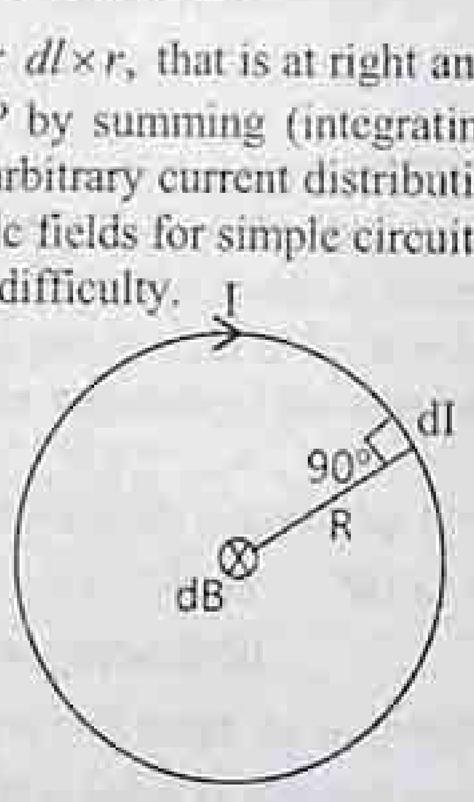


Fig. 7.2: Magnetic field at the centre of a current loop

7.1.1 B due to Current Loop

In this example, we will calculate the magnetic field at the centre of a circular loop. As seen in Figure 7.2, the radius R is constant for all the elements dl, and the angle θ is constant and equal to 90°. If the coil has N turns, the length of the wire is $2\pi RN$. Using equation 7.1, the field at the centre of the coil is

$$B = \int dB = \frac{\mu_0}{4\pi}\int_0^{2\pi RN} \frac{Idl\sin 90}{R^2} = \frac{\mu_0}{4\pi}\frac{I}{R^2}\int_0^{2\pi RN} dl = \frac{\mu_0}{4\pi}\int_0^{2\pi RN} \frac{\mu_0 I}{4\pi R^2}2\pi RN$$

Therefore

$$B = \frac{\mu_0 NI}{2R} \tag{7.2}$$

Example 7.1

A magnetic field of $1.4\times10^{-3}T$ exists at the centre of a circular coil of radius $20cm$ having 15 closely-wound turns. Find the current in the coil.

Solution

$$B = \frac{\mu_0 NI}{2R}$$

$$I = \frac{2RB}{\mu_0 N} = \frac{2\times0.20m\times1.4\times10^{-3}T}{(4\pi\times10^{-7}Hm^{-1})\times15} = 30A$$

7.1.2 B near a long straight wire

Consider a long straight wire carrying current I. We want to find the field at point P very close to the wire that the wire looks infinitely long (Figure 7.3). The field dB due to element dl at point P is normal to the plane of the figure; dB due to all the current elements will have this same direction. To find B at P we simply integrate equation 7.1.

$$B = \int dB = \frac{\mu_0 I}{4\pi}\int_{l=-\infty}^{l=\infty} \frac{\sin\theta}{r^2}dl \tag{7.3}$$

Notice from Figure 7.3 that l, θ and r are related, thus

$$r = (l^2+a^2)^{1/2} \quad \text{and} \quad \sin\theta = \sin(\pi-\theta) = a/(l^2+a^2)^{1/2}$$

where a is the perpendicular distance of P from the wire.

Therefore, equation 7.3 becomes

$$B = \frac{\mu_0 I}{4\pi}\int_{-\infty}^{\infty}\frac{adl}{(l^2+a^2)^{3/2}} = \frac{\mu_0 I}{4\pi}\left|\frac{l}{(l^2+a^2)^{1/2}}\right|_{l=-\infty}^{l=\infty} = \frac{\mu_0 I}{2\pi a} \tag{7.4}$$

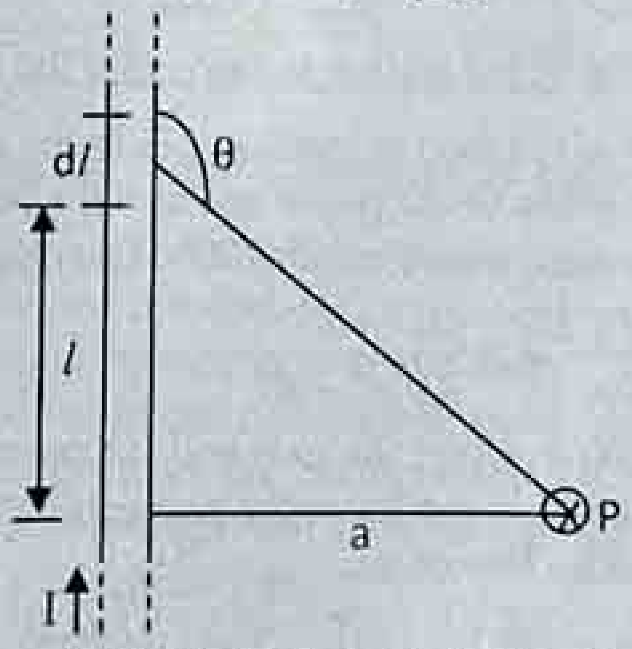


Fig. 7.3: Field of a long straight wire

Equation 7.4 shows that the magnetic field of a long, straight current-carrying wire at a point near it, is inversely proportional to the distance of the point from the wire.

At any point in space, the magnetic field lines of a long, straight current-carrying wire are at tangent to a circle of radius a about the wire, where a is the perpendicular distance from the wire to the field point. The direction of B can be determined by applying the right-hand rule as shown in Figure 7.4. The magnetic field lines thus encircle the wire. Note that the result of equation 7.4 was found experimentally by Biot and Savart using a vibration magnetometer method, and led to their general formula for the magnetic field due to a current element given in equation 7.1.

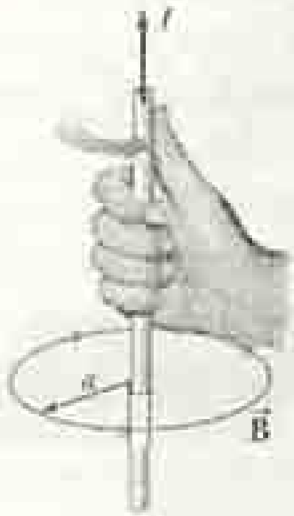


Fig. 7.4: Right hand rule

Example 7.2

A long straight wire 2mm in diameter is carrying a current of $5A$. Find the magnetic field at a distance of $40cm$ from the wire.

Solution

$$B = \frac{\mu_0 I}{2\pi a} = \frac{(4\pi \times 10^{-7} Hm^{-1}) \times 5A}{2\pi \times 0.4m} = 2.5 \times 10^{-6} T$$

7.1.3 B along the axis of a circular current loop

We want to obtain the magnetic field at a point P on the axis of a circular loop at distance x from its centre (Figure 7.5).

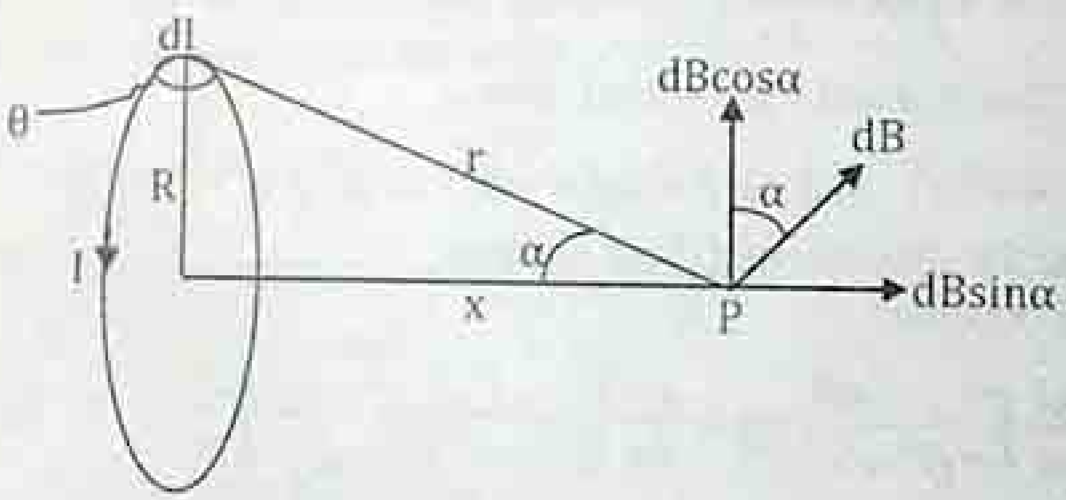


Fig. 7.5: Field on the axis of a flat coil

Consider the element at the top of the loop at right angles to the plane of the paper. The magnetic field dB due to this element is in the plane of the paper and at right angles to the radius vector r. The radius vector and dl are also at right angles, therefore, $\theta = 90°$. The magnitude of dB becomes

$$dB = \frac{\mu_0}{4\pi}\frac{Idl}{r^2} = \frac{\mu_0}{4\pi}\frac{Idl}{x^2 + R^2}$$

This field has components $dB\sin\alpha$ along the axis and $dB\cos\alpha$ perpendicular to the axis. When we sum around all the current elements in the loop, the perpendicular components will cancel out leaving only the component along the axis. Therefore,

$$dB\sin\alpha = \frac{\mu_0}{4\pi}\frac{Idl}{r^2} = \frac{\mu_0}{4\pi}\frac{Idl}{x^2 + R^2}\frac{R}{\sqrt{x^2 + R^2}} = \frac{\mu_0 IR}{4\pi(x^2 + R^2)^{3/2}}dl$$

We obtain the field due to the entire loop by integration.

$$B = \int dB\sin\alpha = \frac{\mu_0 IR}{4\pi(x^2 + R^2)^{3/2}}\int_0^{2\pi R} dl = \frac{\mu_0 R^2 I}{2(x^2 + R^2)^{3/2}} = \frac{\mu_0 I\pi R^2}{2\pi(x^2 + R^2)^{3/2}} \qquad (7.5)$$

At great distances from the loop, $x \gg R$

$$(x^2 + R^2)^{3/2} = (x^2)^{3/2} = x^3$$

Therefore,
$$B = \frac{\mu_0 I \pi R^2}{2\pi x^3} = \frac{\mu_0 IA}{2\pi x^3} = \frac{\mu_0 m}{2\pi x^3} \tag{7.6}$$

where A is the cross-sectional area of the loop and $m = IA$ is the magnetic moment of the current loop. Note that equation 7.6 is similar to the equation for the electric field on the axis of an electric dipole of moment P, viz.

$$E = \left(\frac{1}{4\pi\varepsilon_0}\right)\left(\frac{2P}{x^3}\right)$$

Therefore, a current loop behaves like a magnetic dipole, producing a magnetic dipole field far away from it.

7.1.4 B on the axis of a long Solenoid

A solenoid is a long wire wound in a close packed helix and carrying current I, as illustrated in Figure 7.6. It is used to produce a strong uniform magnetic field in the region surrounded by its loops in analogy to the electric field produced between the plates of a parallel-plate capacitor. The magnetic field of a solenoid is essentially that of a set of N identical current loops placed side by side. Figure 7.7 shows what the field looks like. Inside the solenoid, the field lines are parallel to the axis and closely spaced indicating strong field; outside the solenoid, the lines are less dense. The field looks like that of a bar magnet of identical shape.

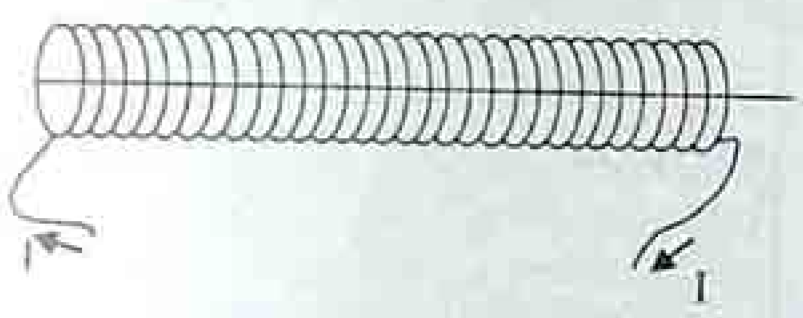


Fig. 7.6: A tightly wounded solenoid

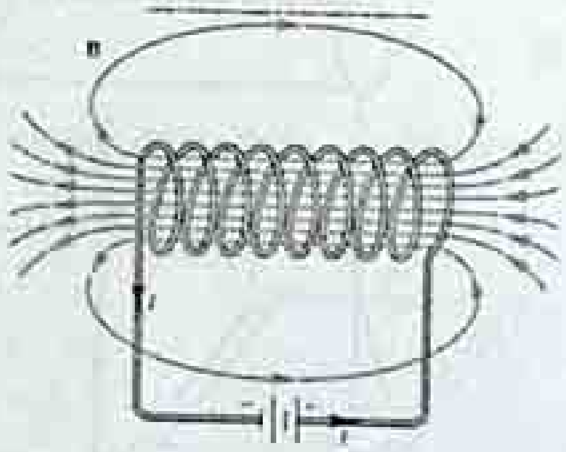


Fig. 7.7: Magnetic field lines of a solenoid

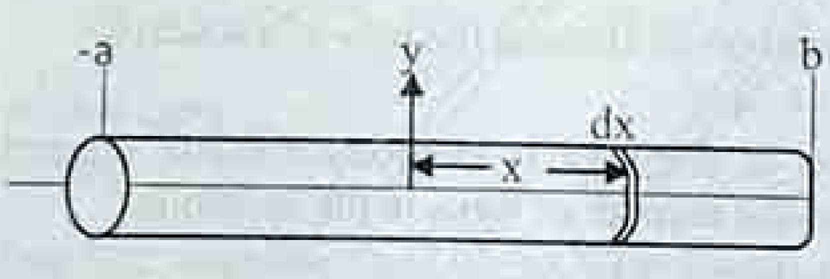


Fig. 7.8: Geometry for calculating the magnetic field inside a solenoid on the axis

We will make use of equation 7.6 to determine the magnetic field of a solenoid along the axis, at a point between the ends. From Figure 7.8, let the left end of the solenoid be at $-a$ and the right end be at $+b$. The magnetic field will be calculated at the origin of coordinates. The element of the solenoid dx is at distance x from the origin. If the solenoid has n turns per metre, then the element dx has ndx turns with each turn carrying current I. Therefore, the element can be regarded as a single loop carrying current $nIdx$. The magnetic field at the origin due to a loop at the point distance x from the origin is obtained by replacing I with $nIdx$ in equation 7.5.

Thus,
$$dB = \frac{\mu_0 R^2 nIdx}{2(x^2 + R^2)^{3/2}}$$

The magnetic field due to the entire solenoid is obtained by integrating this expression from $x = -a$ to $x = b$.

$$B=\int dB=\frac{\mu_0 R^2 nI}{2}=\int_{-a}^{b}\frac{dx}{(x^2+R^2)^{3/2}} \tag{7.7}$$

The integral in equation 7.7 is standard and can be found in Appendix III. It gives

$$\int_{-a}^{b}\frac{dx}{(x^2+R^2)^{3/2}}=\left.\frac{x}{R^2(x^2+R^2)^{1/2}}\right|_{-a}^{b}=\frac{b}{R^2(b^2+R^2)^{1/2}}+\frac{a}{R^2(a^2+R^2)^{1/2}}$$

Substituting this into equation 7.7, we get

$$B=\frac{\mu_0 nI}{2}\left(\frac{b}{(b^2+R^2)^{1/2}}+\frac{a}{(a^2+R^2)^{1/2}}\right) \tag{7.8}$$

For a long solenoid, a and b are much larger than R, making each term in the parenthesis above tend to 1, therefore, the resultant magnetic field is

$$B=\mu_0 nI \tag{7.9}$$

We can obtain the field at one end of the solenoid by making the origin at that end (that is a or b equal to zero), and then making the other end far away compared with R. in that case, one term in the parenthesis is zero, and the other tend to 1. This gives the field at either end as

$$B=\frac{1}{2}\mu_0 nI$$

In effect, the field is uniform near the middle of the solenoid and drops to half that value at the end as shown in Figure 7.9. Note that in practice, solenoids cannot be made infinitely long; but if the length of the solenoid is about ten times the diameter, the field near the middle is very uniform.

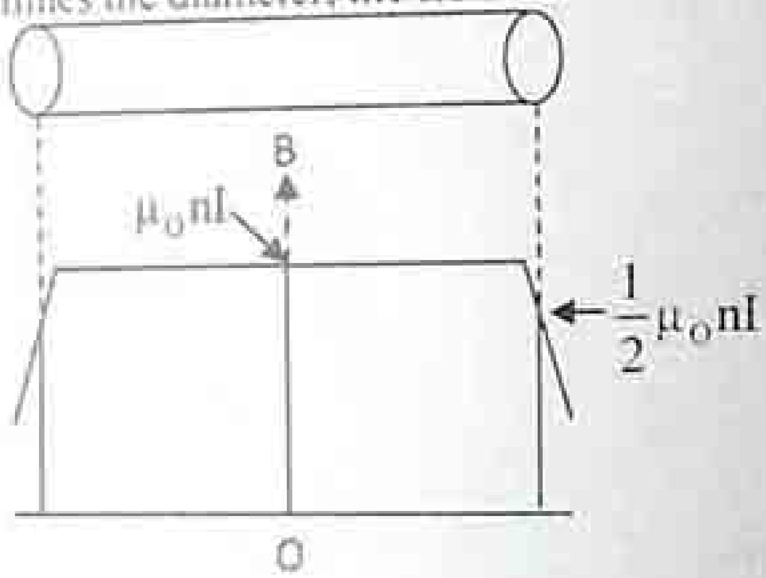


Fig. 7.9: Variation of B along the axis of a solenoid

Example 7.3

A $32cm$ long solenoid, $1.2cm$ in diameter is to produce $0.20T$ magnetic field at its centre. If the maximum current is $3.7A$, find the number of turns in the solenoid.

Solution

Equation 7.9 gives the expression for the field inside the solenoid.

$$B=\mu_0 nI \text{ or } n=\frac{B}{\mu_0 I}=\frac{0.20T}{(4\pi\times10^{-7}\,Hm^{-1})(3.7A)}=4.3\times10^4\,m^{-1}$$

$$N=nl=(4.3\times10^4\,m^{-1})(32\times10^{-2}\,m)=1.4\times10^4\,loops$$

7.1.5 Magnetic Field B in Toroid

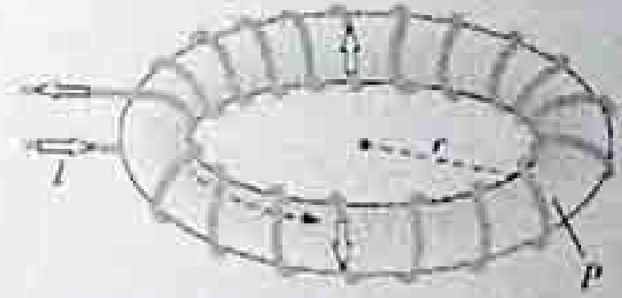


Fig. 7.10: A Toroid

As shown in Figure 7.10, a toroid is a solenoid of N turns and length L metres wound on a circular support instead of a straight one. It is equivalent to an infinitely long solenoid and the field strength at all points within it is given by $B=\mu_0 nI$.

The magnetic field within the toroid is very nearly uniform because the coil has no ends unlike the straight solenoid. In the above equation,

$$n = \frac{N}{L} = \frac{N}{\pi D}$$

where D is the average diameter of the toroid.

There is a lot of interest now in the toroids because they form an important component of the tokamak – a device showing promise as the basis for a fusion power reactor. The toroidal space in a tokamak is filled with hot gaseous hydrogen ions and electrons (plasma) and are confined there by magnetic fields within which the fusion reactions occur.

7.2 Forces Between Two Parallel Conductors

One week after the news of Oersted's experiment in 1820 got to Paris, Ampere showed that two long parallel wires carrying current in the same direction (Figure 7.11) attract each other. The force between these two conductors is purely magnetic. We will calculate the force as follows. The two wires are separated by distance d and carry currents I_1 and I_2. Wire 1 will produce a magnetic field B_1 at all nearby points. The magnetic field of B_1 at the site of wire 2 is, from equation 7.4,

$$B = \frac{\mu_0 I_1}{2\pi d}$$

From the right hand rule, we obtain that the direction of B_1 at wire 2 is down as shown in the figure. Wire 2 which carries current I_2 finds itself in an external field B_1. The length L of wire 2 will experience a sideways magnetic force $(il \times B_1)$ of magnitude.

$$F_2 = I_2 L B_1 = \frac{\mu_0 L I_1 I_2}{2\pi d} \qquad (7.10)$$

The rule for vector products gives that F_2 lies in the plane of the wires and points to the left as shown in Figure 7.11.

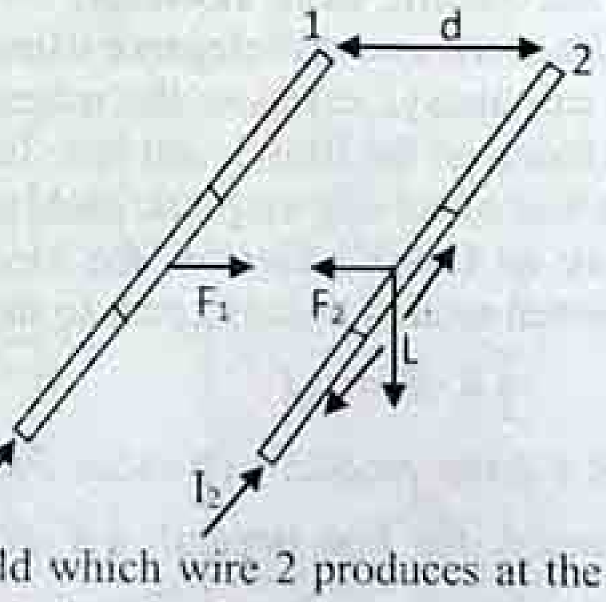


Fig. 7.11: Two parallel wires that carry parallel currents attract each other

If we had started by calculating the magnetic field which wire 2 produces at the site of wire 1, and then computed the force on wire 1, we will find that the force F_1 on wire 1 would, for parallel currents point to the right as shown in the figure. The forces that the two wires exert on each other are equal and opposite. Following this method, we can show that if the currents are antiparallel in the two wires, there will be repulsion between them. Therefore, like currents (same direction) attract each other, while unlike currents (opposite direction) repel each other'.

The attraction between long parallel wires is used to define the unit of current, the ampere. The official definition of the ampere as given by the International Committee on Weights and Measures is as stated. *The ampere is that constant current which, if maintained in two straight parallel conductors of infinite length, of negligible circular cross-section, and placed one metre apart in vacuum, would produce between these two conductors a force equal to* 2×10^{-7} *Newton per metre of length.*

Substituting these values in equation 7.10 we get

$$\frac{F}{L} = \frac{\mu_0 I^2}{2\pi d} = \frac{(4\pi\times10^{-7}\,Hm^{-1})(1A)^2}{2\pi\times1m} = 2\times10^{-7}\,Nm^{-1}$$

as expected from the definition of the ampere. Once the ampere is defined, the unit of charge, the Coulomb, is defined in terms of the ampere and the second from the relation $Q = It$.
It is the electric charge which passes any cross-section in a wire in 1 second when there is a current of 1 ampere in the wire.

Example 7.4
Two straight rods $1m$ long and $3mm$ apart in a current balance carry currents of $30A$ each in opposite directions. Find the mass that must be placed on the upper rod to balance the magnetic force of repulsion.

Solution
The force exerted by the lower rod on the upper rod of length L has the magnitude

$$F = \frac{\mu_0}{2\pi}\frac{I_1 I_2}{R}L = \frac{4\pi\times10^{-7}\,Hm^{-1}}{2\pi}\frac{(30\,A)(30\,A)}{0.003m}(1.0m) = 6.0\times10^{-2}\,N$$

This force can be balanced by a weight mg. Therefore, $mg = 6.0\times10^{-2}\,N$

$$m = \frac{6.0\times10^{-2}\,N}{9.81kg^{-1}} = 6.12\times10^{-3}\,kg = 6.12g$$

7.3 Ampere's Law

Recall that no fundamental magnetic 'charges' or poles playing a similar role to that of electric charges have ever been observed. Instead, the fundamental source of magnetic fields is electric current. The magnetic fields arising from current do not originate or end at points in space, but instead form closed loops encircling the current. This contrasts with electric field lines which begin and end on electric charges.
A major problem in electrostatics is how to calculate the electric field set up at various points by a given charge distribution. These problems can always be solved by direct integration methods using Coulomb's law for the electric field. However, we find that problems with enough symmetry can usually be solved with more ease and elegance using Gauss's law. In the same way, in magnetism as we saw above, we can always calculate the magnetic fields set up by current distributions using integration methods based on the Biot-Savart law. Just as in electrostatics, the Ampere's law provides a simple and elegant way of solving magnetic problems where there is high degree of symmetry.
In an analogous way to Gauss's law for the electric field, Ampere's law relates the tangential component of B summed around a closed path to the current I enclosed by the path. Mathematically, Ampere's law is

$$\oint B\cdot dl = \mu_0 I \tag{7.11}$$

The product $B\cdot dl$ is a scalar product with value $Bdl\cos\theta$ where θ is the angle between B and dl. In Figure 7.12 for instance, the line integral $\oint B\cdot dl$ round the curve is obtained by summing the products $Bdl\cos\theta$ round the loop. The algebraic sum of the currents enclosed by the path is $(I_1 + I_2 - I_3)$, where we have neglected I_4 which is not enclosed and also where, since we are traversing the path anticlockwisely, currents flowing upwards are positive and downwards negative. Thus, applying equation 7.11,

$$\oint Bdl\cos\theta = \mu_0(I_1 + I_2 - I_3)$$

We will now apply Ampere's law to compute magnetic fields in two high symmetry special cases of current-carrying conductors in order to illustrate the elegance.

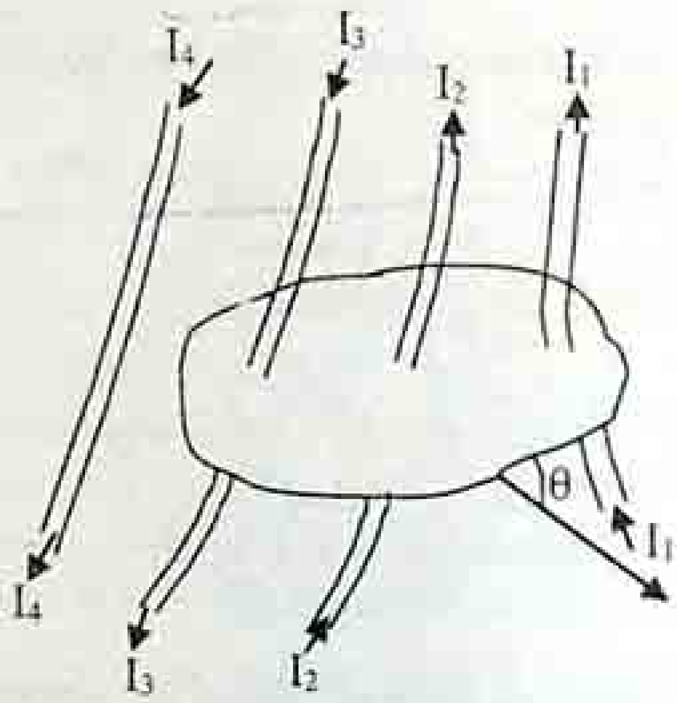


Fig. 7.12: Ampere's Law

7.3.1 Straight wire

We want to find the magnetic field at distance r from the wire carrying current I (Figure 7.13a). B is tangent to the circle of radius r (therefore $\theta = 0$) and constant in magnitude everywhere on the circle.

Therefore, $\oint B \cdot dl = B\oint dl = B \cdot 2\pi r$

The current bounded by the circle is I, therefore

$$B \cdot 2\pi r = \mu_0 I; \quad B = \frac{\mu_0 I}{2\pi r}$$

Suppose the radius of the wire is a, and $r > a$, what we want to find out is the field inside the wire. In this case, the current passing through the cross-section of radius r is

$$I(\pi r^2)/(\pi a^2) = r^2 I/a^2 \text{ (see Figure 7.13b).}$$

Therefore, applying Ampere's law

$$\oint B \cdot dl = B \cdot 2\pi r = \mu_0 \frac{r^2}{a^2} I$$

$$B = \frac{\mu_0}{2\pi} \frac{I}{a^2} r \qquad (r < a)$$

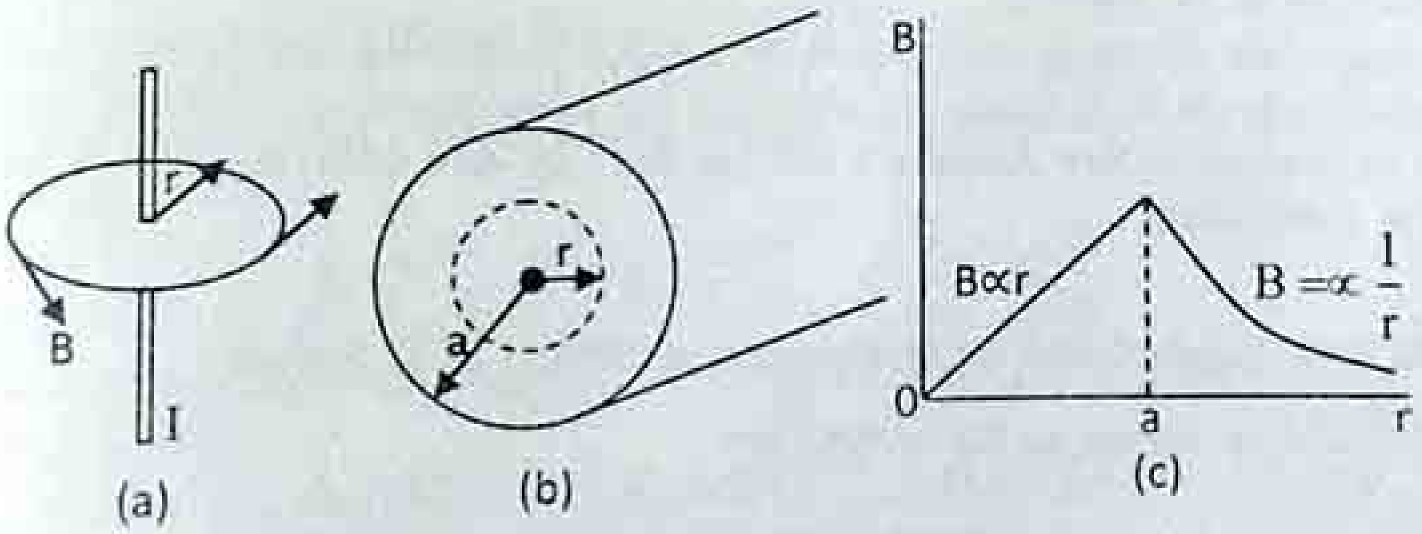


Fig. 7.13: (a) Field due to a long current carrying wire using Ampere's law
(b) Long wire of radius a, carrying current uniformly distributed over its cross-sectional area (c) Graph of B versus r for a wire of radius a carrying current

A curve showing variation of B with r is given in Figure 7.13c. The field B is zero at the centre ($r = 0$), then it increases linearly with r until the surface of the wire ($r = a$). For $r > a$, the field decreases as $1/r$.

7.3.2 Toroid

Let a and b be the inner and outer radii respectively of the toroid. By symmetry, the lines of B form concentric circles inside the toroid (Figure 7.14).

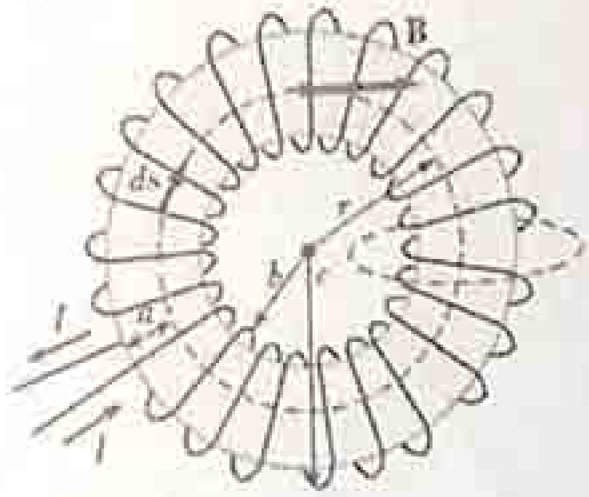


Fig. 7.14: Magnetic field of Toroid by Ampere's law

To obtain B, we calculate the line integral $\oint B \cdot dl$ around a circle of radius r centred at the centre of the toroid, thus,

$$\oint B \cdot dl = B \cdot 2\pi r = \mu_0 IN \tag{7.12}$$

where I is the current in the toroid windings. If there are N turns, then total current through the circle of radius r is NI.

Therefore,

$$B = \frac{\mu_0 IN}{2\pi r} \tag{7.13}$$

This is the situation in the toroid where $a < r < b$. When $r < a$, there is no current through the circle of radius r, therefore the right hand side of equation 7.12 is zero, leading to $B = 0$. Similarly when $r > b$, the total current through the circle of radius r is zero because for each current I into the page at the inner surface of the toroid, there is an equal current I out of the page at the outer surface. Therefore, the right hand side of equation 7.12 is zero, again leading to $B = 0$. Thus there is no field outside the toroid, it is all inside.

Note from equation 7.13 that the magnetic field inside a toroid is not uniform but decreases with r, it has the largest value along the inner edge. However, if the diameter of the loops of the toroid $b - a$ is much less than the radius of the form, the variation of the field from $r = a$ to $r = b$ is small and B is approximately uniform as in the solenoid.

We have used the above two examples to illustrate the simplicity and elegance in using Ampere's law to compute magnetic fields. In both cases, we obtained the same results as when we used the Biot-Savart rule. As we mentioned earlier, Ampere's law can only be used when there is a high degree of symmetry and that is a limitation.

Summary

1. The magnetic field dB at a distance r from a current element Idl is given by $dB = \left(\frac{\mu_0}{4\pi}\right)\frac{Idl \sin\theta}{r^2}$ is known as Biot-Savart law.

2. The magnetic field at the centre of a current loop is given by $B = \frac{\mu_0 NI}{2R}$, where N is the number of turns in the loop.

3. The magnetic field near a long straight wire is $B = \frac{\mu_0 I}{2\pi a}$.

4. The magnetic field along the axis of a circular current loop is $B = \frac{\mu_0}{2}\frac{R^2 I}{(x^2 + R^2)^{3/2}}$.

At great distances from the loop $B = \frac{\mu_0 m}{2\pi x^3}$. Where m is the magnetic moment of the loop.

5. The magnetic field inside a long solenoid and far from the ends is uniform and given by $B = \mu_0 n I$.
   where n is the number of turns per unit length of the solenoid.
6. The magnetic field inside a toroid is $B = \mu_0 n I$.
7. Parallel wires carrying currents in the same (opposite) direction attract (repel) each other. The magnitude of the force per unit length on either wire is $F/L = \mu_0 I_1 I_2 / 2\pi d$ where d is the separation of the wires. The Ampere is defined from this relation.
8. The Ampere is defined such that two long parallel wires each carrying a current of one ampere and separated by 1m exert a force of exactly $2\times10^{-7}\,Nm^{-1}$ on each other.
9. For current distribution of high symmetry. Ampere's law $\oint B \cdot dl = \mu_0 I$ can be used in place of the Biot-Savart law to calculate the magnetic field. This was illustrated by computing the magnetic field due to a long, straight current-carrying wire and also a tightly-wound toroid.

Exercise 7

7.1 A circular coil of wire has a diameter of 20cm. It contains 10 loops carrying currentof 3A. If the coil is placed in an external magnetic field of 2T, determine the maximum torque exerted on the coil by the field.

A. 0.0Nm B. 0.188Nm C. 1.88Nm D. 18.8Nm

7.2 An electric wire carries a dc current of 25A. What is the magnetic field due to the current at a point 10cm from the wire?

A. $2.5\times10^{-5}T$ **B.** $5.0\times10^{-3}T$ **C.** $40\times10^{-7}T$ **D.** $5.0\times10^{-5}T$

7.3 The magnetic field due to a long straight wire at a point near it is

A. directly proportional to the current. **B.** inversely proportional to the current. **C.** Directly proportional to the distance from the wire. **D.** Dependent only on permeability.

7.4 Which of the following statements are correct?

I. Parallel currents in the same direction exert an attractive force on each other.
II. Parallel currents in the same direction exert a repulsive force on each other.
III. Antiparallel currents in opposite directions exert an attractive force on each other.
IV. Antiparallel currents in opposite directions exert a repulsive force on each other.

A. II and III only **B.** I and IV only **C.** II and IV only **D.** I and III only

7.5 The field B at a point r inside a long straight cylindrical wire of radius R carrying a current I is given as,

A. $\frac{\mu_0 I}{2\pi r}$ **B.** $\frac{\mu_0 I r^2}{2\pi R^2}$ **C.** $\frac{\mu_0 I r}{2\pi R^2}$ **D.** $\frac{\mu_0 I}{2\pi R r}$

7.6 Calculate the field inside near the center of a solenoid of 400 turns, 10cm lond and carrying a current of 2.0A.

A. 0.01T B. 0.1T C. 1.6T D. 0.16T

7.7 Which of the following is not true concerning a toroid.

A. The magnetic field is not uniform within the toroid.
B. The magnetic field lines inside the toroid will be circles concentric with the toroid.
C. The magnetic field outside the coils of the toroid is zero.
D. The magnetic field is directly proportional to the radius of the toroid.

7.8 Two long straight wires are carrying currents of different magnitudes. If the two wires are parallel to each other and the amount of current flowing in each wire is doubled, the magnitude of the force between the two wires will be

A. half the magnitude of the original force. **B.** the same as the magnitude of the original force. C. four times the magnitude of the original force. **D.** twice the magnitude of the original force.

7.9 If the length of a solenoid is halved and the number of turns is doubled, the magnetic field will

A. Remain unchanged B. be halved C. be quadrupled D. be doubled

7.10 Which of the following will increase the magnetic field inside a solenoid in which the wires are wound such that each loop touches the adjacent ones?
A. increasing the radius of the wire. B. decreasing the radius of the wire. C. making the radius of the loop smaller. D. increasing the radius of the solenoid.

7.11 A long straight wire carries a current of $20A$. Determine the magnetic field due to the current at a distance of $4mm$ from the wire.

7.12 A coil carrying a current of $10A$ produces a magnetic field of $2\pi \times 10^{-5}T$ at the centre of the coil. Calculate the radius of the coil.

7.13 Show that the magnetic field at a distance r from a long straight wire carrying current I is given by $B = \mu_0 I/2\pi r$.

7.14 Two long parallel wires are separated by $10cm$. They carry the same current in the same direction. If the force between the wires per unit length is $2\times10^{-4}\,Nm^{-1}$, determine the current carried by the wires. Is the force attractive or repulsive?

7.15 Show that the magnetic induction at the centre of a long solenoid having n closely wound turns per unit length is given by $B = \mu_0 nI$, where I is the current in the winding.

7.16 The long straight wire AB in Figure 7.15 carries a current of $20A$. The rectangular loop whose long edges are parallel to the wire carries a current of $10A$. Find the magnitude and direction of the resultant force exerted on the loop by the magnetic field of the wire.

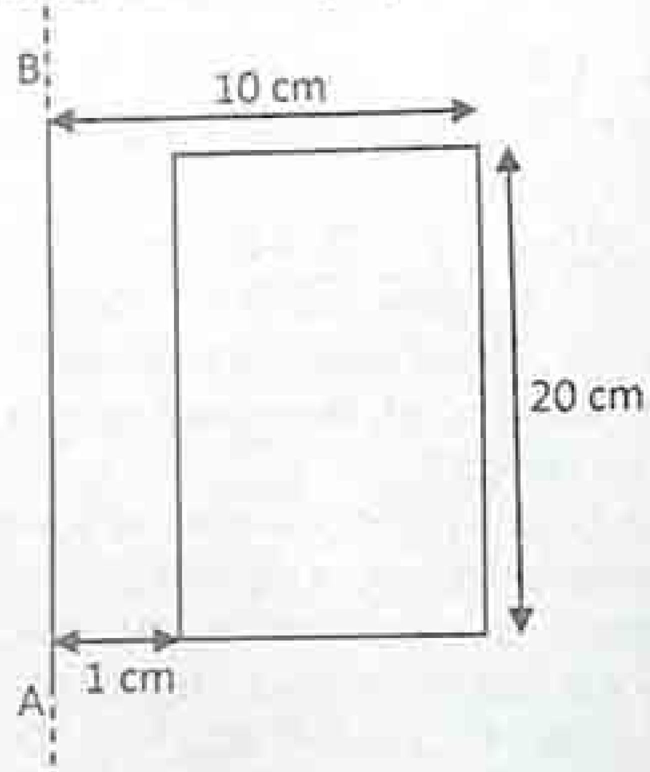


Fig. 7.15: Problem 7.16

7.17 A solenoid of length $30.2cm$ with 6000 turns produces a magnetic field of $0.1T$. Calculate the current in the windings.

7.18 A current of $1.0A$ circulates in a round thin wire loop of radius $10cm$. Find the magnetic induction (i) at the centre of the loop (ii) at a point lying on the axis of the loop at a distance of $10cm$ from its centre.

7.19 A square loop of wire of edge L carries a current I. Show that the value of the magnetic field at the centre is given by $B = 2\sqrt{2}\,\mu_0 I/\pi L$.

7.20 A long straight cylindrical copper tube having an inside radius of $1.0cm$ and an outside radius of $2cm$ carries a current of $200A$. Determine the magnetic induction at the following distances from the axis of the tube: $0.5cm$, $1.5cm$ and $4cm$.

CHAPTER 8
ELECTROMAGNETIC INDUCTION

8.0 Introduction

In the previous chapters, we saw that a wire carrying current produces magnetic field around it and that a magnetic field exerts a force on a moving electric charge or a piece of wire carrying current. Do magnetic fields produce electric currents since electric current produce magnetic field?

Two scientists, Joseph Henry, an American, and an Englishman Michael Faraday, independently, in 1831, found that a changing magnetic field actually gives rise to an electric current. The actual discovery was first made by Joseph Henry but Faraday first published his results before him hence, Faraday is credited with the discovery which is known as the electromagnetic induction. This chapter considers the phenomenon in some detail and states the laws that govern it. We shall also examine the profound effect it has on our world.

8.1 Induced EMF

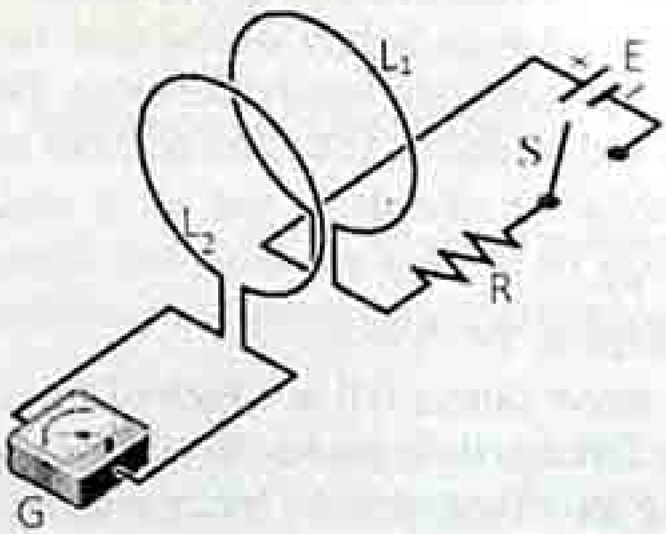


Fig. 8.1: Faraday's experiment to induce an e.m.f

Faraday in his attempt to produce electric fields from magnetic fields used the apparatus shown in Figure 8.1. A coil of wire L_1 was connected to a battery and to switch S. A second coil of wire L_2 was connected to a galvanometer but did not contain a battery this time. When the switch S was closed, a steady current produced a magnetic field and the second coil L_2 was inside this magnetic field. Faraday had hoped that the magnetic field would produce current in the second circuit which would be detected by the galvanometer.

It is clear that the current was produced in the second circuit only when the current in L_1 was starting or stopping. This current in L_2 is the induced current and an e.m.f. exists in the circuit for the current to flow.

Faraday carried out other experiments on the electromagnetic induction. Figure 8.2 shows a coil of wire connected to a galvanometer. If the bar magnet is quickly moved into the coil as shown in Figure 8.2(a), the galvanometer deflects indicating induced current in the coil. The deflection is zero if the magnet is held stationary with respect to the coil. If on the other hand the bar magnet is quickly pulled away from the coil as shown in Figure 8.2(b), the galvanometer again deflects, but in the opposite direction. If the bar magnet is held steady and the coil is moved toward or away from the magnet, the galvanometer deflects.

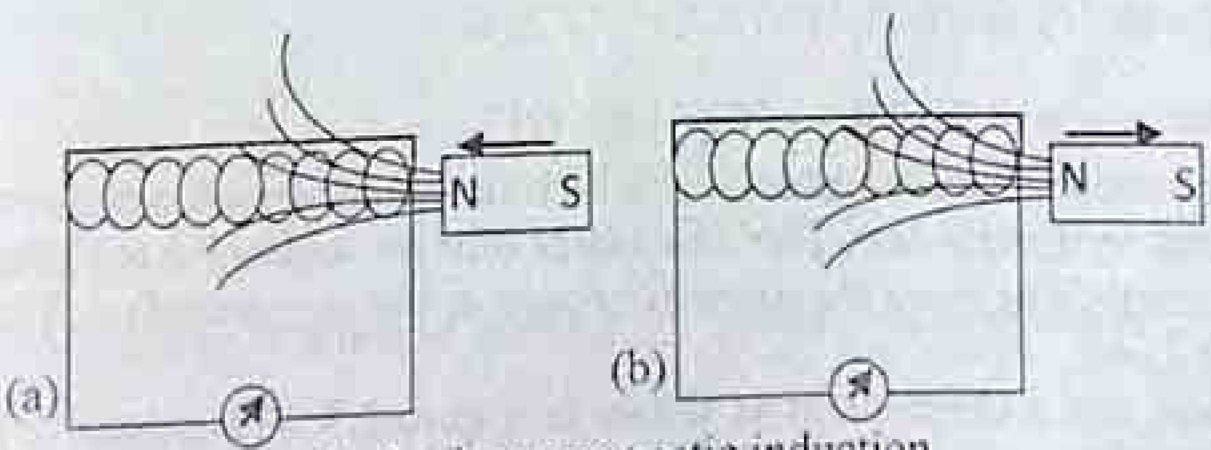


Fig 8.2: Faraday's electromagnetic induction

What is important therefore is the relative motion of the coil and the bar magnet. It does not matter whether the coil and the bar magnet.

Based on these experiments, Faraday concluded that steady magnetic fields do not induce current in closed loops or circuits but changing magnetic fields produce currents in closed loops. A changing magnetic field produces induced e.m.f.

8.2 Faraday's law of induction and Lenz's law

Faraday's law of induction states that the induced e.m.f. ε in a circuit is equal to the negative rate at which the flux through the circuit is changing. If the change of flux is in Weber's/sec the e.m.f. ε will be in volts.

$$\varepsilon = -\frac{d\phi_B}{dt} \tag{8.1}$$

This equation is called Faraday's law of induction. The minus sign is an indication of the direction of the induced e.m.f. If equation 8.1 is applied to a coil of N turns and e.m.f appears in every turn and these e.m.fs are to be added. If the coil is so tightly wound that each turn can be said to occupy the same region of space, the flux through each turn will be the same. Note that the flux through each turn is also the same for (ideal) toroids and solenoid. The induced e.m.f in all such devices is given by

$$\varepsilon = -N\frac{d\phi_B}{dt} = -d\frac{(\phi_B)}{dt} \tag{8.2}$$

where $N\phi_B$ measures the flux linkage in the device.

The coil connected to the galvanometer cannot tell in which of these experiments it is participating; it is aware only that the flux passing through its cross-sectional area changing. The flux through a circuit can be changed also by changing its shape, that is, by squeezing or stretching it. Suppose a loop enclosing an area A lies in a uniform magnetic field. The magnetic flux through the loop is equal to $BA\cos\theta$, hence, the induced e.m.f can be expressed as

$$\varepsilon = -\frac{d}{dt}(BA\cos\theta) \tag{8.2a}$$

Example 8.1

A coil of 600 turns is threaded by a flux of $8.0\times10^{-5}Wb$. If the flux is reduced to $3.0\times10^{-5}Wb$ in $0.015s$, find the average induced e.m.f.

Solution

From equation 8.2, $\varepsilon = -N\dfrac{d\phi_B}{dt} = -600\dfrac{(8.0-3.0)\times10^{-5}Wb}{0.015s} = -2.0V$

Example 8.2

The magnetic flux through each loop of a 35-loop coil is given by $(3.6t-0.71t^3)\times10^{-2}Tm^2$, where the time t is in seconds. Determine the induced e.m.f, ε at $t=5.0s$.

Solution

The total flux through the coil is $N\phi_B = 35\times10^{-2}(3.6t\times0.71t^3)$

Hence, $d\dfrac{(N\phi_B)}{dt} = 35\times10^{-2}(3.6t\times0.71t^3) = (0.75t^2-1.26)V$

Then at $t=5s$, we obtain $\varepsilon = (0.75(25)-1.26)V = 17.49V$

The minus sign in equation 8.2 indicates in which direction the induced e.m.f acts. The direction of the e.m.f is determined by *Lenz's law*. *This law states that an induced e.m.f always gives rise to a current whose magnetic field opposes the original change in the magnetic flux.* This law was deduced by Heinrich Friedrich Lenz.

A bar magnet is moved into a closed loop as shown in Figure 8.3. The magnetic field due to the bar magnet is into the page. Since the magnetic flux through the loop is changing, according to Faraday's law of induction, an e.m.f. is induced in the coil. The induced e.m.f. produces a current which sets up a magnetic field that opposes the change in the magnetic flux through the coil. The magnetic field set up by the induced current as shown in Figure 8.3 is out of the page. Hence, Lenz's law tells us that the induced current must flow in the direction shown in the figure so that the magnetic field due to this current is coming out of the page toward the reader.

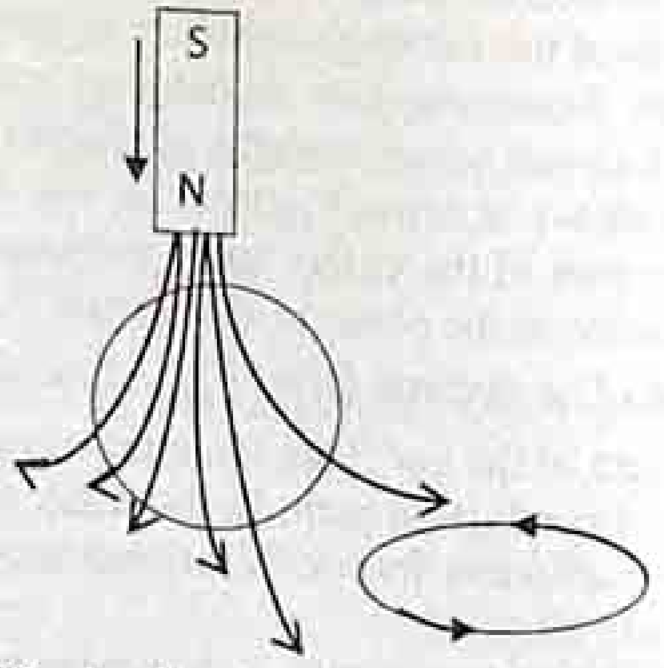


Fig 8.3: Application of Lenz's law

The flux through the loop decreases if the bar magnet is moved away from the loop. The induced current must now produce a magnetic field into the page in an attempt to oppose the change and to maintain the statuesque. The direction of the current must be removed.

8.3 Motional EMFs

In this section, we describe motional e.m.f., which is the e.m.f. induced in a conductor moving through a constant magnetic field. The straight conductor of length l shown in Figure 8.4 is moving through a uniform magnetic field directed into the page. For simplicity, let's assume the conductor is moving in a direction perpendicular to the field with constant velocity under the influence of some external agent. The electrons in the conductor experience a force

$$F_B = qv \times B$$

that is, directed along the length l, perpendicular to both v and B. Under the influence of this force, the electrons move to the lower end of the conductor and accumulate there, leaving a net positive charge at the upper end. As a result of this charge separation, an electric field E is produced inside the conductor.

The charges accumulate at both ends until the downward magnetic force, qvB, on charges remaining in the conductor is balanced by the upward electric force qE. The condition for equilibrium requires that the forces on the electrons balance: $qE = qvB$ or $E = vB$

The electric field produced in the conductor is related to the potential difference across the ends of the conductor according to the relationship $V = El$.

Therefore, for the equilibrium condition,

$$\Delta V = El = Blv \tag{8.3}$$

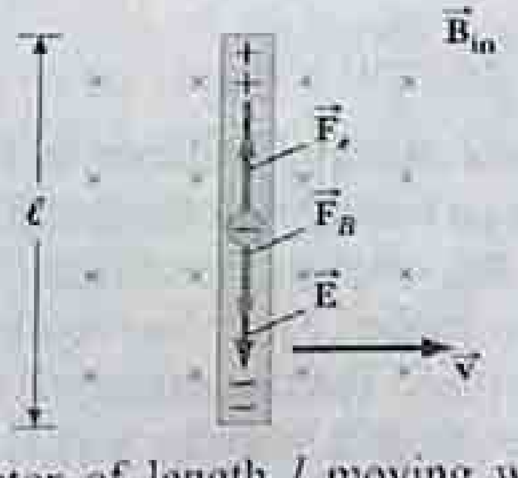


Fig 8.4: A straight electrical conductor of length l moving with a velocity v through a uniform magnetic field B.

where the upper end of the conductor in Figure 8.4 is at a higher electric potential than the lower end. Therefore, a potential difference is maintained between the ends of the conductor as long as the conductor continues to move through the uniform magnetic field. If the direction of the motion is reversed, the polarity of the potential difference is also reversed.

A more interesting situation occurs when the moving conductor is part of a closed conducting path. This situation is particularly useful for illustrating how a changing magnetic flux causes an induced current in a closed circuit. Consider a circuit consisting of a conducting bar of length l sliding along two fixed parallel conducting rails as shown in active Figure 8.5a. For simplicity, let's assume the bar has zero resistance and the stationary part of the circuit has a resistance R. A uniform and constant magnetic field B is applied perpendicular to the plane of the circuit. As the bar is pulled to the right with a velocity v under the influence of an applied force F_{app}, free charges in the bar experience a magnetic force directed along the length of the bar. This force sets up an induced current because the charges are free to move in the closed conducting path. In this case, the rate of change of magnetic flux through the circuit and the corresponding induced motional e.m.f. across the moving bar are proportional to the change in area of the circuit.

Because the area enclosed by the circuit at any instant is lx, where x is the position of the bar, the magnetic flux through that area is $\phi_B = Blx$.

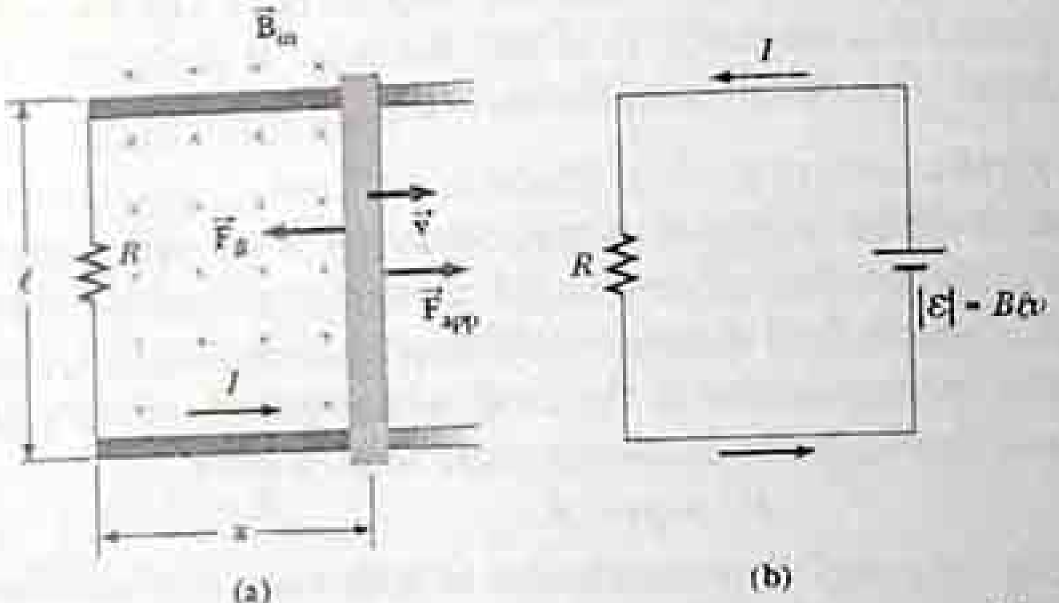


Fig. 8.5: (a) A conducting bar sliding with a velocity along two conducting rails under the action of an applied force the magnetic force opposes the motion, and a counterclockwise current I is induced in the loop.
(b) The equivalent circuit diagram for the setup shown in (a).

Using Faraday's law and noting that x changes with time at a rate $v = dx/dt$, we find that the induced motional e.m.f is

$$\varepsilon = -\frac{d\phi_B}{dt} = \frac{d}{dt}(Blx) = -Bl\frac{dx}{dt}$$

$$\varepsilon = -Blv \qquad (8.4)$$

Because the resistance of the circuit is R, the magnitude of the induced current is

$$I = \frac{\varepsilon}{R} = \frac{Blv}{R} \qquad (8.5)$$

The equivalent circuit diagram for this example is shown in Figure 8.5b. Let's examine the system using energy considerations. Because no battery is in the circuit, you might wonder about the origin of the induced current and the energy delivered to the resistor. We can understand the source of this current and energy by noting that the applied force does work on the conducting bar. Therefore, we model the circuit as a non isolated system. The movement of the bar through the field causes charges to move along the bar with some average drift velocity; hence, a current is established. The change in energy in the system during some time interval must be equal to the transfer of energy into the system by work, consistent with the general principle of conservation of energy.

Let's verify this mathematically. As the bar moves through the uniform magnetic field B it experiences a magnetic force F_B of magnitude IlB. Because the bar moves with constant velocity, it is modeled as a particle in equilibrium and the magnetic force must be equal in magnitude and opposite in direction to the applied force, or to the left in Figure 8.5a. (If F_B acted in the direction of motion, it would cause the bar to accelerate, violating the principle of conservation of energy.) Using equation 8.5 and $F_{app} = F_B = IlB$, the power delivered by applied force is

$$P = F_{app}v = (IlB)v = \frac{B^2l^2v^2}{R} = \frac{\varepsilon^2}{R} \tag{8.6}$$

This power input is equal to the rate at which energy is delivered to resistors.

Example 8.3

A wire which moves with a speed of $4.6ms^{-1}$ is $17.0cm$ long and has negligible resistance. The magnitude of the magnetic field is $0.25T$ and the resistance of the U-shaped conductor is 25.00Ω at a given instance. Calculate the induced e.m.f and the current flowing in the U-shaped conductor.

Solution

From equation 8.4, we have $\varepsilon = -Blv = (0.25T)(17\times10^{-2}m)(4.6ms^{-1}) = 0.20V$

Since the resistance in the U-shaped conductor is 25Ω, the current I is

$$I = \frac{\varepsilon}{R} = \frac{0.20V}{25\Omega} = 8.0mA$$

8.4 Induced Electric Fields

We have seen that a changing magnetic flux induces an e.m.f. and a current in a conducting loop. In our study of electricity, we related a current to an electric field that applies electric forces on charged particles. In the same way, we can relate an induced current in a conducting loop to an electric field by claiming that an electric field is created in the conductor as a result of the changing magnetic flux. We also noted in our study of electricity that the existence of an electric field is independent of the presence of any test charges. This independence suggests that even in the absence of a conducting loop, a changing magnetic field generates an electric field in empty space.

This induced electric field is non conservative, unlike the electrostatic field produced by stationary charges. To illustrate this point, consider a conducting loop of radius r situated in a uniform magnetic field that is perpendicular to the plane of the loop as in Figure 8.6. If the magnetic field changes with time, an e.m.f. $\varepsilon = -\frac{d\phi_B}{dt}$ is, according to Faraday's law, induced in the loop. The induction of a current in the loop implies the presence of an induced electric field E, which must be tangent to the loop because that is the direction in which the charges in the wire move in response to the electric force. The work done by the electric field in moving a test charge q once around the loop is equal to $q\varepsilon$.

Because the electric force qE acting on the charge is the work done by the electric field in moving the charge once around the loop is $qE(2\pi r)$, where $2\pi r$ is the circumference of the loop. These two expressions for the work done must be equal; therefore,

$$q\varepsilon = qE(2\pi r)$$

$$E = \frac{\varepsilon}{2\pi r}$$

Using this result along with Equation 8.1 and that $\phi_B = BA = B\pi r^2$ for a circular loop, the induced electric field can be expressed

$$E = \frac{1}{2\pi r}\frac{d\phi_B}{dt} = -\frac{r}{2}\frac{dB}{dt} \tag{8.7}$$

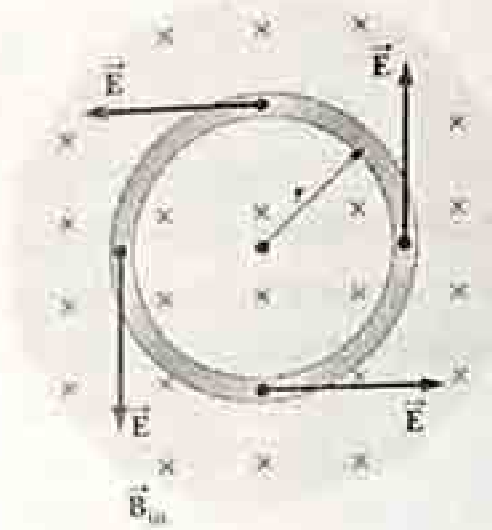


Fig. 8.6: A conducting loop of radius r in a uniform magnetic field perpendicular to the plane of the loop.

If the time variation of the magnetic field is specified, the induced electric field can be calculated from equation 8.7. The e.m.f for any closed path can be expressed as the line integral of $E \cdot dS$ over that path: $\varepsilon = \oint E \cdot dS$. In more general cases, ε may not be constant and the path may not be a circle.

Hence, Faraday's law of induction, $\varepsilon = -\frac{d\phi_B}{dt}$, can be written in the general form

$$\oint E \cdot dS = -\frac{d\phi_B}{dt} \tag{8.8}$$

The induced electric field E in equation 8.8 is a non conservative field that is generated by a changing magnetic field. The field E that satisfies equation 8.8 cannot possibly be an electrostatic field because were the field electrostatic and hence conservative, the line integral of $E \cdot dS$ over a closed loop would be zero, which would be in contradiction to equation 8.8.

8.5 Alternating-Current Generator

Electric generators take in energy by work and transfer it out by electrical transmission. To understand how they operate, let us consider the alternating-current (AC) generator. In its simplest form, it consists of a loop of wire rotated by some external means in a magnetic field. In commercial power plants, the energy required to rotate the loop can be derived from a variety of sources. For example, in a hydroelectric plant, falling water directed against the blades of a turbine produces the rotary motion; in a coal-fired plant, the energy released by burning coal is used to convert water to steam, and this steam is directed against the turbine blades.

As a loop rotates in a magnetic field, the magnetic flux through the area enclosed by the loop changes with time, and this change induces an e.m.f. and a current in the loop according to Faraday's law. The ends of the loop are connected to slip rings that rotate with the loop. Connections from these slip rings, which act as output terminals of the generator to the external circuit are made by stationary metallic brushes in contact with the slip rings.

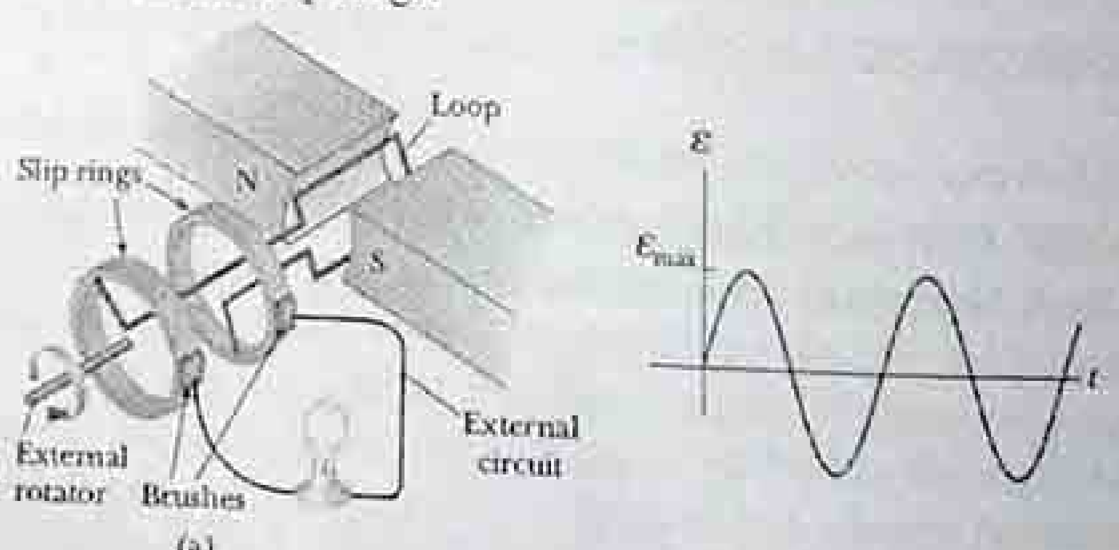


Fig. 8.7: (a) Schematic diagram of an AC generator. An e.m.f is induced in a loop that rotates in a magnetic field. (b) The alternating e.m.f. induced in the loop plotted as a function of time.

Instead of a single turn, suppose a coil with N turns (a more practical situation), with the same area A, rotates in a magnetic field with a constant angular speed ω. If θ is the angle between the magnetic field and the normal to the plane of the coil as in Figure 8.8, the magnetic flux through the coil at any time t is

$$\phi_B = BA\cos\theta = BA\cos\omega t \tag{8.9}$$

$$\varepsilon = -N\frac{d\phi_B}{dt} = -NAB\frac{d}{dt}(\cos\omega t) = NAB\sin\omega t \tag{8.10}$$

This result shows that the e.m.f. varies sinusoidally with time as plotted in active Figure 8.7b. Equation 8.10 shows that the maximum e.m.f. has the value

$$\varepsilon_{\max} = NAB\omega \tag{8.11}$$

which occurs when $\omega t = 90°$ or $270°$. In other words, $\varepsilon = \varepsilon_{\max}$ when the magnetic field is in the plane of the coil and the time rate of change of flux is a maximum. Furthermore, the e.m.f. is zero when $\omega t = 0$ or $180°$, that is, when B is perpendicular to the plane of the coil and the time rate of change of flux is zero.

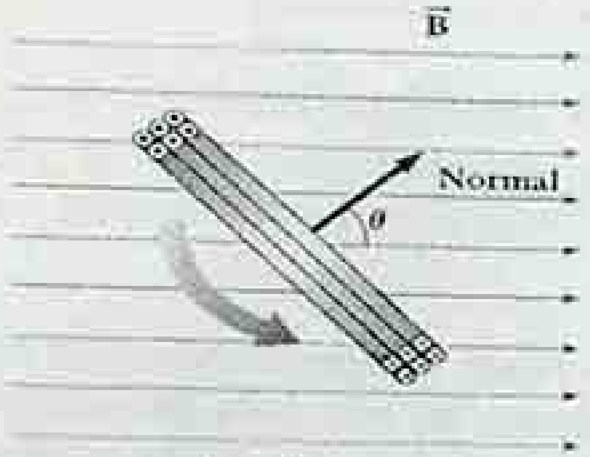


Fig. 8.8: A loop enclosing an area A and containing N turns, rotating with constant angular speed ω in a magnetic field. The e.m.f induced in the loop varies sinusoidally in time.

Example 8.4

A simple generator has 200-loop square coil, 8.0cm on a side. How fast must it turn in a 0.40T field to produce a 12.0V peak output?

Solution

Equation 8.11 gives the peak voltage, $\varepsilon_{\max} = NAB\omega$;

$$12V = (200)(64\times10^{-4}m^2)(0.40T)\omega$$

$$\omega = \frac{12}{200\times64\times10^{-4}\times0.40} = 23.44rads^{-1}$$

But, $\omega = 2\pi f$; $f = \frac{\omega}{2\pi} = 3.7Hz$

8.6 Transformers

A transformer is an electrical device that transfers energy between two or more circuits through electromagnetic induction. It is used for increasing or decreasing AC voltage. Figure 8.9 shows a typical transformer; two coils are wound around a core of magnetic material. The primary coil, of N_p turns, is connected to an AC voltage which produces a time-varying current I_p in the primary coil. This current I_p produces time varying magnetic flux $\phi_B(t)$ in the core upon which both the primary and secondary coils are wound. This changing magnetic flux induces an AC voltage of the same frequency in the secondary coil. According to Faraday's law, the voltage or e.m.f. induced in the secondary coil is

$$V_S = N_S \frac{d\phi_B(t)}{dt} \tag{8.12}$$

If we assume that the resistance of the coil is negligible, that is, there is no power loss in the primary coil, then the changing magnetic flux produces a back e.m.f. in the primary that is equivalent in magnitude to the applied voltage. Therefore,

$$V_P = N_P \frac{d\phi_B(t)}{dt} \tag{8.13}$$

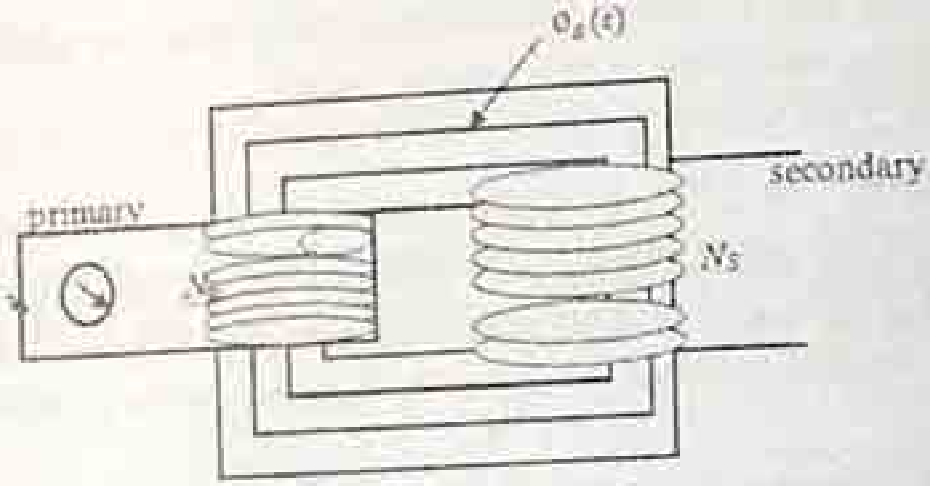


Fig. 8.9: A typical transformer

Dividing equation 8.12 by equation 8.13, we obtain

$$\frac{V_S}{V_P} = \frac{N_S}{N_P} \tag{8.14}$$

From equation 8.14, we see that the induced voltage is proportional to the relative number of turns on the two coils. This proportionality arises because the two coils are coupled with the same varying magnetic flux. If N_S is greater than N_P, V_S is greater than V_P. In this case, the transformer is called a step-up transformer since the induced voltage is increased by it. If N_P is greater than N_S, V_P is greater than V_S, then the transformer is said to be a step-down transformer. Equation 8.14 is known as the transformer equation.

The power output of a transformer cannot be greater than the power input although a transformer can increase or decrease voltage. If the secondary coil is connected to an external load, current I_S is drawn and the power dissipated is $V_S I_S$. The loss in power at the secondary coil causes the primary coil to act as though it has resistance, then the power input is $V_P I_P$. If we assume that the transformer is 100 percent efficient and no power loss occurs when it operates, we can equate the power input to the power loss in the secondary coil. Thus,

$$V_S I_S = V_P I_S$$

$$\frac{V_S}{I_P} = \frac{N_S}{N_P} = \frac{I_P}{I_S} \tag{8.15}$$

We see from equation 8.15 that the current in the coil are inversely proportional to the voltages as long as no power is lost in the transfer.

Example 8.5

A model-train transformer delivers $14V$ to the engine. If the input voltage to the transformer is $100V$ and the primary coil contains 400 turns, how many turns are on the $14V$ secondary coil? If the transformer is ideal and the train draws $4A$, what is the current drawn from the circuit?

Solution

From the transformer equation we have $\frac{V_S}{V_P} = \frac{N_S}{N_P}$

But $N_P = 400\,turns,\ V_P = 100V\ and\ V_S = 14V$

Therefore,

$$N_p = \frac{V_p}{V_S} N_S = \frac{100}{14}(400) = 2857\, turns$$

From equation 8.15 we have $\frac{V_S}{V_p} = \frac{N_S}{N_p} = \frac{I_p}{I_S}$

$$I_p = \frac{N_S}{N_p} I_S = \frac{400(14)}{100(400)}(4) = \frac{56}{100} = 0.56\,A$$

Transformers play an important role in the transmission of electricity. Electricity generation companies of Nigeria (PHCN) locates its generating plants some distances from the city limits. This is because electricity must be transmitted over long distances; and there is always some power loss in the transmission lines. This loss can be minimized if electricity is transmitted at high voltage. For this reason, electricity is transmitted at $167kV$ in Nigeria. When this high voltage arrives in towns, it is then stepped down in stages ($165kV$ (National Grid) to $415V$ in electric substations) prior to the distribution to homes, factories and to schools.

Example 8.6

If $80kW$ is to be transmitted over two 0.055Ω lines, estimate how much power is saved if the voltage is stepped up from $120V$ to $1200V$ and then down again, rather than simply transmitting at $120V$. Assume the transformers are each 99 percent efficient.

Solution

At $120V$: $P_{in} = VI_1$

$$I_1 = \frac{P_{in}}{V} = \frac{80\times 10^3 W}{120V} = 0.67\times 10^3\,A$$

Therefore, $P_{loss} = I_1^2 R = (0.67\times 10^3\,A)^2(2)(0.055\Omega) = 4.94\times 10^4\,W$

At $1200V$: $P_{in} = 0.99(80\times 10^3) = VI_2 = 1200\,I_2$

$$I_2 = \frac{(0.99)(80\times 10^3)}{1200}A$$

$$P_{loss} = I_2^2 R = \left[\frac{(0.99)(80\times 10^3)}{1200}\right]^2 \times (2\times 0.055) = 479W = 4.79\times 10^2 W$$

$$P_{loss} = 4.79\times 10^2 W + (0.001)(79.5\times 10^2) + (0.001)(80\times 10^2) = 2.1kW$$

Power saved $= 49.4kW - 2.1kW = 47.3kW$

8.7 Application of Electromagnetic Induction

To conclude this chapter, we will discuss one application of electromagnetic induction in the field of medicine. If the heartbeat stops, external electrical stimulation can start it again. If electrodes are fixed to the chest on the long axis of the heart, pulses may be passed through the chest at regular intervals corresponding to the repetition time of the normal heartbeat. This may by be achieved by charging a capacitor through a resistor up to $60V$, at which point it discharges through the electrodes. The artificial stimulation is very painful to a conscious patient and must be regarded as purely emergency treatment to restart the heart.

With certain heart conditions, this emergency treatment would have to be continued indefinitely when the heart is incapable of resuming its electrical activity. For such ailing hearts, pacemakers are implanted routinely to assist in pumping blood through the body. Electrodes usually of platinum to prevent chemical reaction taking place are fixed to the heart or surrounding tissue, and pulses are sent through them from an electrical circuit as described above. The power is provided by mercury cells.

The patient must be operated on at intervals of around two years to renew the cells. Although the cells are encased in a substance as epoxy resin, at times, the fluids in the body find a way inside and react with substance which composes the cells themselves. This can lead to failure of the pacemaker. Several pacemakers have therefore been devised in which the power supply is external to the body and strapped to it. The power is supplied to a coil taped to the chest and a current is induced in a coil taped to the chest and attached to the heart. This method which employs induced e.m.f leads to ease of changing the cells and eliminates arising from the attack of cells by body chemicals.

Summary

1. Faraday's law of induction states that the induced e.m.f which appears in a coil of wire containing N loops is equal to the rate at which the magnetic flux through the coil is changing.
2. The flux ϕ_B for a given surface immersed in a magnetic field B is given by $\phi_B = \int_S B \cdot dS$.

   If B is uniform then $\phi_B = Ba\cos\theta$

   where θ is the angle between the field and the normal to the surface and A is the area of the surface.
3. Lenz's law states that the induced e.m.f is always in such a direction so as to oppose the change producing it.
4. When a conducting bar of length l moves at a velocity v through a magnetic field B, where B is perpendicular to the bar and to the motional e.m.f induced in the bar is $\varepsilon = -Blv$.
5. Electric fields induced by changing magnetic fields differ from those associated with static charges in that their lines of force form closed loops. The betatron uses induced electric fields to accelerate electrons to high energy.
6. The instantaneous e.m.f in a rotating coil is $\varepsilon = NAB\sin\omega t$.
7. The maximum e.m.f in a rotating coil is given by $\varepsilon_{\max} = NAB\omega$. When $\omega t = 90°\, or\, 270°$.
8. Therefore when the magnetic field is in the plane of the coil and the time rate of change of flux is a maximum $\varepsilon = \varepsilon_{\max}$.
9. Mechanical energy is transformed to electrical energy by transformers. A generator consists of a rectangular coil of wire rotating in a uniform magnetic field. The e.m.f induced in such a generator is alternating in character.
10. The basic transformer equation is $\frac{V_S}{V_P} = \frac{N_S}{N_P} = \frac{I_P}{I_S}$.
11. Electromagnetic induction is employed in artificial external pacemakers which aid the ailing heart in pumping blood to the body.

Exercises 8

8.1 A coil of area $10cm^2$ is in a uniform field of magnetic induction $0.1T$. The field is reduced to zero in $1ms$. Determine the induced e.m.f?

A. $0.093V$ **B.** $0.1V$ **C.** $0.1mV$ **D.** $0.01V$

8.2 A transformer has $N_1 = 350\,turns$ and $N_2 = 2000\,turns$. If the input voltage is $V(t) = 170\cos(\omega t)V$, what rms voltage is developed across the secondary?

A. $876V$ **B.** $687V$ **C.** $68V$ **D.** $68.7V$

8.3 A $2cm$ by $1.5cm$ rectangular coil has 300 turns and rotates in a magnetic field B at $60Hz$. What must the value of B be so that the maximum e.m.f generated is $24V$?

A. $7.07T$ B. $0.707T$ C. $0.0707T$ D. $0.335T$

8.4 A helicopter has blades of $3m$, rotating at $2rev/s$ about a central hub. If the vertical component of the earth's magnetic field is $0.5\times10^{-4}T$, what is the e.m.f induced between the blade tips and the centre hub?

A. $2.83V$ **B.** $1.41mV$ **C.** $1.41V$ **D.** $2.83mV$

8.5 The coil in an AC generator consists of 8 turns of wire, each of area $A = 0.090m^2$, and the total resistance of the wire is 12.0Ω. The coil rotates in a 0.500T magnetic field at a constant frequency of $60.0Hz$ .Find the maximum induced e.m.f in the coil.

A. $78V$ **B.** $13.6V$ **C.**$136V$ **D.** $28V$

8.6 What is the maximum induced current in the coil in exercise 9.05 when the output terminals are connected to a low-resistance conductor?

A. $11.3A$ **B.** $1.13A$ **C.** $22.6A$ **D.** $11.3A$

8.7 A coil consists of 200 turns of wire. Each turn is a square of side $d = 18cm$, and a uniform magnetic field directed perpendicular to the plane of the coil is turned on. If the field changes linearly from 0 to $0.50T$ in $0.80s$, what is the magnitude of the induced e.m.f in the coil while the field is changing?

A. $0.4V$ **B.** $7.0V$ **C.** $4.0V$ **D.** $0.8V$

8.8 Find the magnitude of the induced current in the coil in exercise 9.07 while the field is changing? **A.** $8.0A$ **B.** $0.5A$ **C.** $2.0A$ **D.** $2.26A$

8.9 A general form of Faraday's law of induction is $\oint E \cdot dS = -\frac{d\phi_B}{dt}$.

Where E is the non conservative electric field that is produced by the changing magnetic flux.

A. True **B.** False

8.10 A solenoid with 20turns/cm has a radius of 2.4cm. The magnitude of the induced electric field $2cm$ from the axis is $5\times10^{-3}\,V/m$. At what rate is the current in the solenoid changing?

A. $1.99\times10^{-3}A$ **B.** $1.99\times10^{-3}\,A/s$ **C.** $1.99\times10^{3}\,A/s$ **D.** $1.99\times10^{3}A$

8.11 Explain what is meant by electromagnetic induction and show that energy is conservedduring the process.

8.12 A rectangular loop of N turns and of area A is rotated at a frequency f in a uniform magnetic field B. Show that the induced e.m.f is given by $\varepsilon = \varepsilon_{max}$, where $\omega = 2\pi f$ and $\varepsilon_{max} = NAB\omega$.

8.13 Explain the principle of operation of a simple alternating current generator.

8.14 Explain the operating principle of a transformer and derive the transformer equation.

8.15 A voltage of $120V$ a.c is supplied to a calculator which uses only $24V$. What type of transformer is required? If the resistance is 10ohms, calculate the current in the secondary and primary circuits of the calculator.

8.16 Draw a diagram of the cross section of a betatron and state the functions of the magnetic field inside it if properly shaped and controlled.

8.17 Find the magnetic flux through a solenoid of length $25cm$, radius $1cm$ and 400turns, that carries a current of $3A$.

8.18 A transformer has 1800 primary turns and 120 secondary turns. The input voltage is $120V$ and output current is $8.0A$. What is the secondary voltage and primary current?

8.19 A faraday disc dynamo of radius $R = 20cm$ generates $1.2V$ in a $0.08T$ magnetic field directed perpendicularly to the plane of the disc. What is the frequency of rotation in rpm?

8.20 A motor contains a coil with a total resistance of 10Ω and is supplied by a voltage of $120V$. When the motor is running at its maximum speed, the back e.m.f is $70V$. (a) Find the current in the coil at the instant the motor is turned on. (b) Find the current in the coil when the motor has reached maximum speed.

CHAPTER 9
INDUCTANCE AND ENERGY STORAGE IN MAGNETIC FIELDS

9.0 Introduction

We saw in chapter 8 that a changing magnetic flux through a circuit induces e.m.f in that circuit and that a wire carrying current produces a magnetic field around it. Using Faraday's law of electromagnetic induction, we can calculate this induced e.m.f. if we combine these two ideas. We will expect a circuit carrying a time varying current to induce an e.m.f in a nearby second circuit.

9.1 Mutual Inductance

Let us consider two coils of wire which are placed near one another as shown in Figure 9.1. Coil A contains N_1 closely package loops, similarly coil B contains N_2 loops. Coil A carries a time varying current I_1.

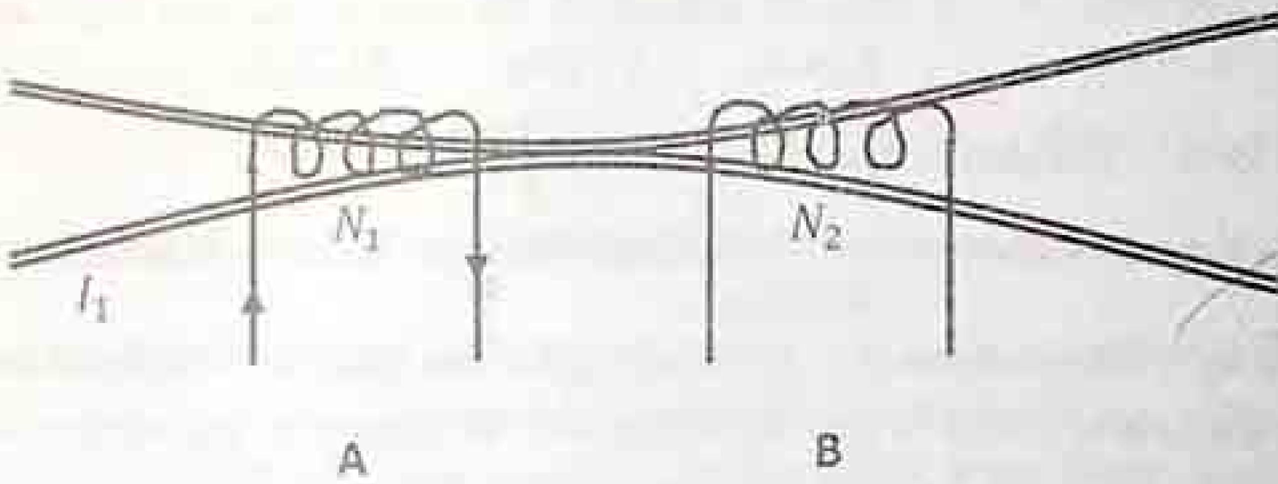


Fig. 9.1: Two coils of wire, coil A carrying time-varying current I_1

The induced e.m.f. ε_2, in coil B is proportional to the time rate of change of flux passing through it. This change of flux through coil B is as a result of the current I_1 in coil A. it is therefore more convenient to express the induced e.m.f. ε_2 in terms of current I_1.

Now let the magnetic flux in each loop of coil B due to the current I_1 in coil A be $\Phi_{2,1}$. Then $N_2\Phi_{2,1}$ is the total number of flux linkages in coil B. If the two coils maintain their positions in space, $N_2\Phi_{2,1}$ is proportional to I_1

Thus,
$$N_2\Phi_{2,1} \propto I_1$$
or
$$N_2\Phi_{2,1} = M_{2,1}I_1 \tag{9.1}$$

where $M_{2,1}$ in equation 9.1 is the constant of proportionality and is called mutual inductance. $M_{2,1}$ is the mutual inductance of coil B with respect to coil A. from equation 9.1. We have

$$M_{2,1} = \frac{N_2\Phi_{2,1}}{I_1} \tag{9.2}$$

The induced e.m.f ε_2 in coil B, which is due to the changing current in coil A maybe expressed according to Faraday's law as

$$\varepsilon_2 = -N\frac{d\Phi_{2,1}}{dt} = -M_{2,1}\frac{dI_1}{dt} \tag{9.3}$$

To obtain equation 9.3, we have used the fact that $M_{2,1}I_1 = N_2\Phi_{2,1}$.

Note that the mutual inductance depends on geometric factors like shape, size, number of turns and relative positions of the coils, these factors are said to be constant and is valid as long as the coils do not contain iron ore or any ferro-magnetic materials.

$$\varepsilon_1 = -M_{1,2}\frac{dI_2}{dt} \tag{9.4}$$

Mutual inductance of coil A with respect to coil B is $M_{1,2}$.

From equation 9.3 and 9.4, we observe that the induced e.m.f in coil A is proportional to the rate of change of current in the coil B and vice versa.

Since they are the same for a given arrangement, the subscripts are not required. Thus,

$$M = M_{1,2} = M_{2,1}$$

Therefore, equations 9.3 and 9.4 are rewritten as

$$\varepsilon_2 = -M\frac{dI_1}{dt} \text{ and } \varepsilon_1 = -M\frac{dI_2}{dt}$$

The SI units of mutual inductance M is Henry (H), where $1H = \frac{V.S}{A} = 1\Omega s$.

$V.S$ = volt-second, A = Ampere and Ω = Ohm

9.2 Self Inductance

An isolated coil of wire containing N turns or loops can also induce an e.m.f in itself. When a changing current passes through the coil, it in turn induces an e.m.f. The induced e.m.f. according to Faraday's law is

$$\varepsilon = -\frac{d}{dt}(N\Phi_B) \qquad (9.5)$$

The number of the flux linkages $(N\Phi_B)$ is the most important quantity in equation 9.4. $N\Phi$ is proportional to the current I as long as iron or other magnetic materials are not nearby. Thus

$$N\Phi \propto I$$

or

$$N\Phi = LI \qquad (9.6)$$

where L is the constant of proportionality and it is called self inductance. Equation 9.5 becomes

$$\varepsilon = -L\frac{dI}{dt} \qquad (9.7)$$

The self inductance has a similar S.I. unit as mutual inductance which is Henry and its magnitude depends on geometric factors.

Example 9.1

Determine the self inductance of a solenoid that contains n loops per unit length. The solenoid is l meter long and has a radius R.

Solution

For a close packed coil with no iron nearby, we have $L = \frac{N\Phi}{I}$.

The number of flux linkages in the length l near the center of the solenoid is $N\Phi = (nl)BA$,

$B = \mu_0 nI =$ Magnetic field, $A = \pi R^2 =$ cross-sectional area.

$\therefore$ $N\Phi = \mu_0 n^2 lI\pi R^2$

The inductance from equation 9.6, $L = \frac{N\Phi}{I} = \frac{\mu_0 n^2 Il\pi R^2}{I} = \mu_0 n^2 l\pi R^2$.

9.3 Energy Stored in a Magnetic Field

From our study of classical mechanics we recall that when a stone is lifted from the ground to a height h work is done. The work is equal to the potential energy of the stone. This energy is stored in the gravitational field between the stone and the earth. The energy is withdrawn from the field when the stone is lowered to its original position. Similarly the work done in charging up a capacitor ($W = \frac{1}{2}Q/C$), is stored in the electric field in the capacitor as energy. This energy is retrieved from the field when the capacitor is discharged.

When an inductor of self-inductance L, is carrying a current i, which is changing at a rate of di/dt, the induced e.m.f in the inductor opposes this change. Work must therefore be done on the inductor if this change is to continue. Let us now calculate the work required to increase the current in an inductor from zero to some value I. The induced e.m.f. in the coil, because of this change, is

$$\varepsilon = L\frac{di}{dt}$$

But power is $P = i\varepsilon$. Hence, $dW = Pdt = i\varepsilon dt = Lidt$.
Then the total work done to increase the current from zero to I is

$$W = \int dW = L\int_0^I i\,di = \frac{LI^2}{2}$$

This work is equal to the energy stored in the inductor when it is carrying a current I. This energy stored in the inductor can be considered to be stored in the magnetic field in the coil.
Let U be the energy stored in the magnetic field of the inductor. Thus,

$$U = \frac{LI^2}{2} \qquad (9.8)$$

where we assume that $U = 0$ when $I = 0$. Now to express equation 9.8 in terms of the magnetic field, let us use the results of Example 9.1. In that example, we determined the expression for the self inductance of a solenoid of length l which is $L = \mu_0 n^2(\pi R^2 l) = \mu_0 n^2 lA$. Where the cross-sectional area A is equal to πR^2. But the magnetic field at the centre of the solenoid carrying current i is

$$B = \mu_0 ni, \text{ but } i = I.$$

Then equation 9.8 becomes

$$U = \frac{1}{2}(\mu_0 n^2 lA)\left(\frac{B^2}{\mu_0^2 n^2}\right) = \frac{1}{2}\frac{B^2}{\mu_0}(Al) \qquad (9.9)$$

Here Al is the volume of the solenoid. Therefore, the energy density u of a solenoid is

$$u = \frac{U}{Al} = \frac{1}{2}\frac{B^2}{\mu_0} \qquad (9.10)$$

Equation 9.10 states that *the energy density in any region of space is proportional to the square of the magnetic field in that region*. Even though equation 9.10 is derived for a special case of a solenoid, it is valid for any region of space where a magnetic field exists. Note that equation 9.10 is analogous to the equation below.

$$u = \frac{1}{2}\varepsilon_0 E^2$$

Example 9.2

Typical large values of electric and magnetic fields attained in laboratories are about $10^4\, V/m$ and $2.0T$. (i) Determine the energy density for each field and compare (ii) what value of electric field would produce the same energy density as the $2.0T$ magnetic field.

Solution

(i) From equation 9.10, we have $u_B = \frac{1}{2}\frac{B^2}{\mu_0} = \frac{0.5\times 2^2}{1.25\times 10^{-6}\,Hm^{-1}} = 1.6\times 10^6\, J/m^3$.

Similarly, $u_E = \frac{1}{2}\varepsilon_0 E^2 = \left(0.5\times 8.85\times 10^{-12}\, C^2/Nm^2 \times (10^4\, V/m)^2\right) = 4.4\times 10^{-4}\, J/m^3$.

(ii) Equate u_E to u_B to determine the value of E that would produce the same energy as 2.0 magnetic field. Thus,

$$u_E = u_B$$

$$\frac{1}{2}\varepsilon_0 E^2 = \frac{1}{2}\frac{B^2}{\mu_0}$$

$$E = \frac{B}{\sqrt{\varepsilon_0 \mu_0}} = 6.0 \times 10^8\, V/m$$

9.4 LR Circuits

A practical electronic circuit consists of wire with self-inductance L that prevents the current in the circuit from increasing or decreasing instantaneously. Such a coil is known as an inductor or choke. We use this symbol $\text{-}\!\!\!\text{ⲙⲙ}\!\!\!\text{-}$ to represent an inductor in circuit diagrams. Any inductor will have some resistance, so we usually draw this resistance separately from its inductance L. This is shown in Figure 9.2. The resistance R could also include any resistor connected in series with the inductor. As soon as the switch K is thrown, current I will begin to increase in the circuit and due to this increasing current, the inductor will produce e.m.f that will oppose this change. In other words, the inductor acts like a battery whose polarity is opposite that of the working battery. The induced e.m.f. in the inductor is $\varepsilon_L = -L\frac{dI}{dt}$.

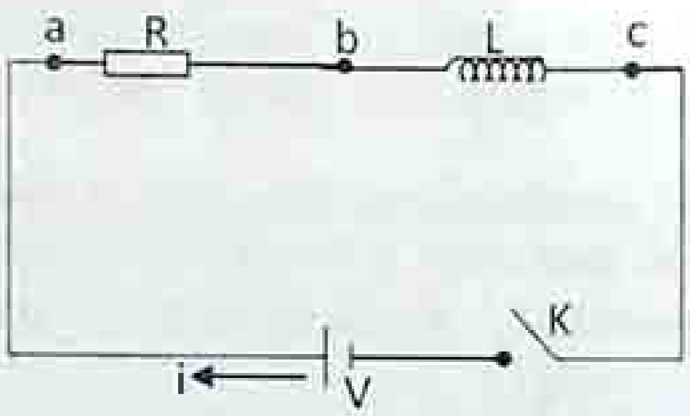


Fig. 9.2: An LR Circuit

If I is increasing in the circuit, dI/dt is positive, therefore the induced e.m.f in the inductor is negative and opposes the working battery voltage V. Now we are in a position to write down the loop equation for the circuit shown in Figure 9.2 with the switch closed, we apply Kirchhoff's loop equation to the circuit and we have

$$V - IR - L\frac{dL}{dt} = 0$$

where I is the current in the circuit at any instant and IR is the voltages drop across the resistance R. Rearranging the last expression, we obtain

$$L\frac{dL}{dt} + IR = V \qquad (9.11)$$

This is a differential equation and it is similar to the equation we solved in section 5.10 when we discussed RC circuits. Let us rearrange the terms in equation 9.11 in order to separate the variable I and t. Thus,

$$\frac{dI}{V - IR} = \frac{dt}{L}$$

Now integrating both sides of the above expression, we obtain

$$-\frac{1}{R}\ln\left[\frac{V - IR}{V}\right] = \frac{t}{L}$$

$$\ln\left[\frac{V - IR}{V}\right] = -\frac{R}{L}t \qquad (9.12)$$

Let us define inductive time constant of the RL circuit as $\tau = \frac{L}{R}$.

Therefore equation 9.12 becomes $\ln\left[\frac{V - IR}{V}\right] = -\frac{t}{\tau}$

or
$$I = \frac{V}{R}\left[1 - \exp - \frac{t}{\tau}\right] \tag{9.13}$$

Equation 9.13 is plotted in Figure 9.3, where $I = 0$ at $\tau = 0$, note that the equilibrium value occurs at $\tau = \omega$ and it is given by V/R.

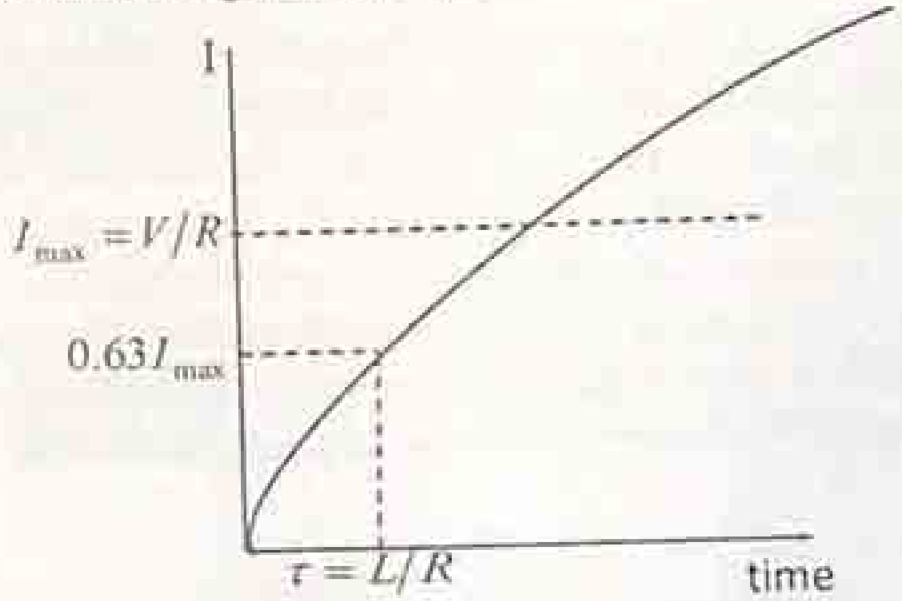


Figure 9.3: Growth of current in LR circuit when connected to a battery

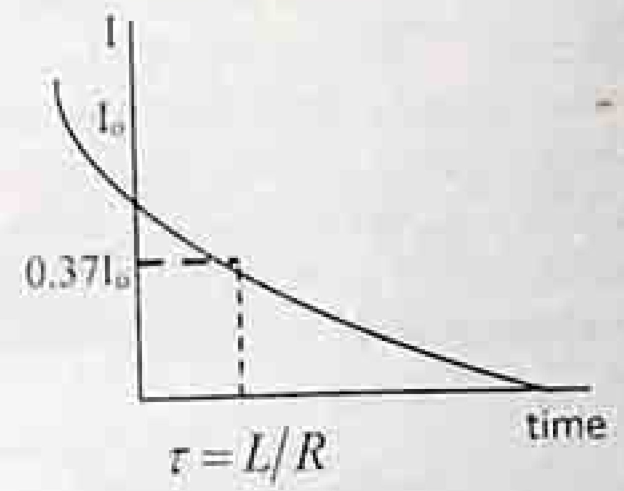


Figure 9.4: Decay of current in an LR circuit when the battery is removed

We recall from our studies of RC circuits that the time constant represents the time required for the current I to reach 63 percent of its final value. If we remove the battery ($V = 0$) in Figure 9.3 at the instant when the current in the circuit is I_0, then equation 9.11 becomes

$$L\frac{dI}{dt} + IR = 0$$

Rearranging the terms and integrating, we have $\int_0^I dI = -\frac{R}{L}\int_0^t dt.$

where $I = I_0$ at $t = 0$ and $I = I$ at some time t. The above expression is integrated to give us

$$\ln\left[\frac{I}{I_0}\right] = -\frac{R}{L}t$$

or
$$I = I_0 \exp - \frac{t}{\tau} \tag{9.14}$$

Equation 9.14 shows that the current decays exponentially to zero. The plot of the current versus time is shown in Figure 9.4 as the current decreases continuously with time. The slope (di/dt) of this curve is negative; this means that $\varepsilon_L = -L\frac{di}{dt}$ is positive but it still opposes the change in current.

Example 9.3

It takes 1.56mins for the current in an LR circuit to increase from zero to half its maximum value. Determine (a) the time constant of the circuit and (b) the resistance of the circuit if the inductance is $L = 310H$.

Solution

From equation 9.13 we have $I = \frac{V}{R}\left[1 - \exp - \frac{t}{\tau}\right]$

Let $\quad I_0 = V/R$

Hence, $I = I_0\left[1 - \exp - \frac{t}{\tau}\right]$

$$\frac{I}{I_0} = 1 - \exp - \frac{t}{\tau}$$

$$\frac{I_0 - I}{I_0} = \exp - \frac{t}{\tau}$$

Now taking the natural log of both sides, we have $\ln\left[\frac{I_0 - I}{I_0}\right] = -\frac{t}{\tau}$

$$\tau = \frac{t}{\ln(I_0 - I/I_0)} = \frac{1.56\times10^{-3}s}{\ln 2} = 2.25\times10^{-3}s$$

(b) $R = L/\tau = 1.38\times10^{-5}\Omega$

Example 9.4

A coil has a resistance of 15.0Ω and a self inductance of $5.0mH$. It is placed across the terminals of a $12V$ battery of negligible internal resistance.

(i) Determine the final current attained.

(ii) Calculate the current after $100\mu s$.

Solution

(i) The final current is $I_f = \frac{\varepsilon_0}{R} = \frac{12V}{15\Omega} = 0.800\,A$

(ii) The time constant for the circuit is $\tau = \frac{L}{R} = \frac{5\times10^{-3}H}{15\Omega} = 333\,\mu s$

The current after $100\,\mu s$ is given by

$$I = \frac{\varepsilon_0}{R}\left(1 - e^{-t/\tau}\right) = 0.80\,A\left(1 - e^{-100/333}\right) = 0.80\,A\times(1 - 0.741) = 0.207\,A$$

9.5 LC Circuit

We have already discussed both RC and LR circuits. Now we wish to look at the circuit that contains only a capacitor and an inductor L. Such a circuit is known as an LC circuit. An LC circuit is shown in Figure 9.5. Initially, the capacitor C is fully charged and it carries the total charge Q_m, the current I in the inductor is zero; at this instant, the total energy U stored in the electric field of the capacitor is

$$U_E = \frac{Q_m^2}{2C} \tag{9.15}$$

When the switch K is closed, the capacitor begins to discharge through the inductor; positive charge carried move clockwise as shown in Figure 9.5. As the capacitor discharges, the current $I = dI/dt$ in the inductor is established and the energy stored in the electric field decreases. As the current increases, some energy is stored in the magnetic field of the inductor. The energy stored in the magnetic field of the inductor is

$$U_B = \frac{1}{2}LI^2 \tag{9.16}$$

As the charge q decreases, the energy U_E decreases but the energy U_B stored in the magnetic field increases. Thus, the energy is transferred to the magnetic field around the inductor. At the instant when the capacitor is fully charged ($q = 0$), the electric field is zero and the current I, in the inductor has reached its maximum value. The energy stored in the capacitor has been transferred entirely to the magnetic field of the inductor. Then the current begin to decrease, transporting positive charges to the opposite plate of the capacitor.

The energy stored in the magnetic field begins to decrease and is being transferred to the electric field that is being built up in the capacitor. When the current fully drops to zero, the capacitor will be fully charged with a maximum charge Q_m. Now that the energy is transferred entirely back to the electric field of the capacitor, U_B is zero.

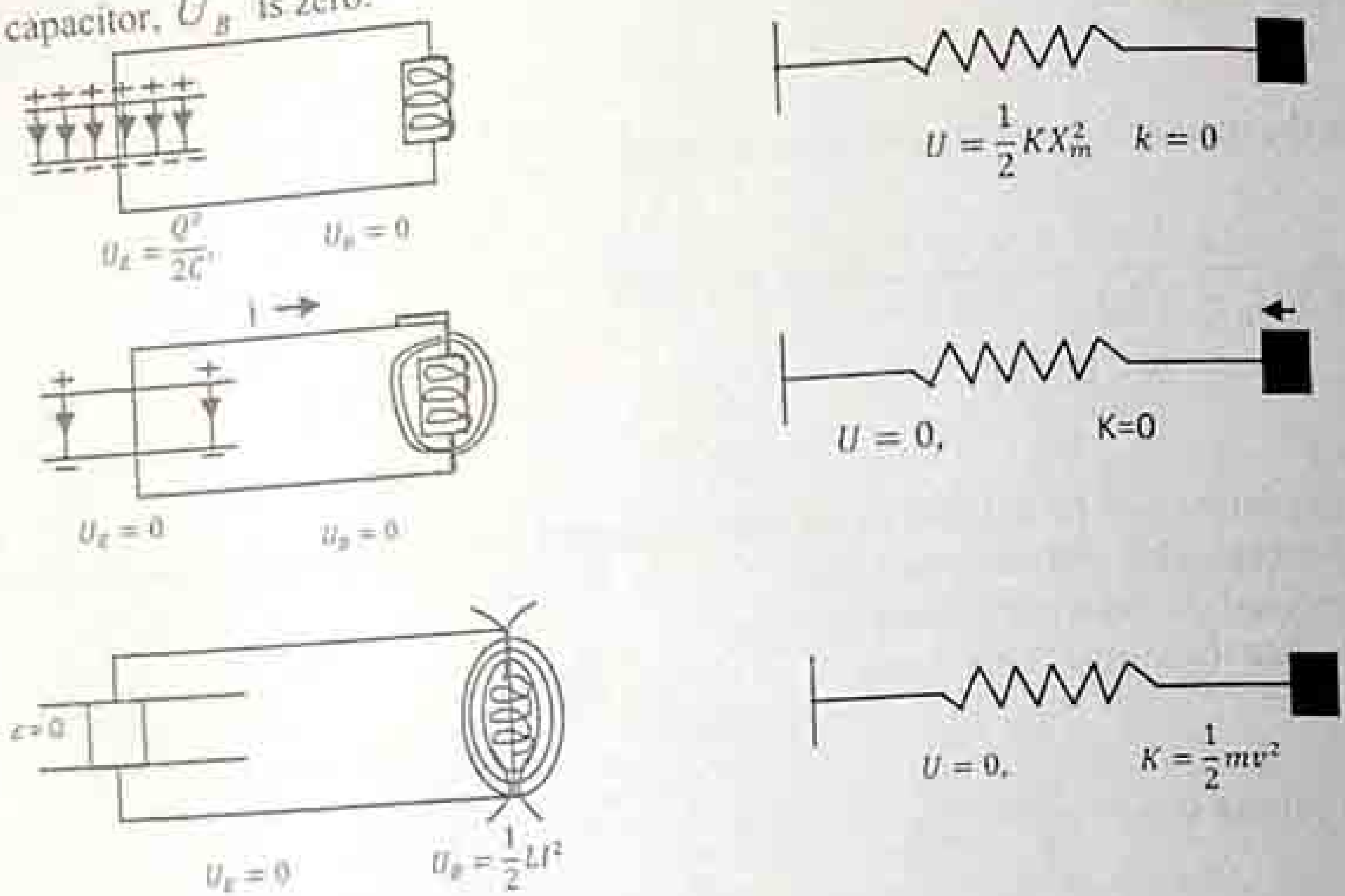


Fig. 9.5: An oscillating LC circuit

The capacitor will now start to discharge again but this time, the current will flow counter clockwise. The process of transferring charge from one plate of the capacitor to the other or transferring energy back and front between the electric field in the capacitor and magnetic field in the inductor continues at a definite frequency. This process is called an LC oscillation, and for the ideal case being discussed, it will continue definitely.

This electromagnetic oscillation is analogous to the oscillating mass spring system that we studied in mechanics. The potential energy of the mass spring system is $\frac{1}{2}kx_m^2$ when the spring is fully stretched which is analogous to the energy stored in the capacitor $(Q_m^2/2C)$ when it is fully charged. The kinetic energy of the mass spring system (when the spring is at its equilibrium position) is $\frac{1}{2}mv^2$ which is analogous to $\frac{1}{2}LI_0^2$ - the energy stored in the magnetic field, when this capacitor is fully discharged. At intermediate points, part of the energy is potential energy, and part is kinetic energy. Analogously, when the capacitor is not fully discharged and the current in the inductor has not reached its maximum value, part of the energy is stored in the capacitor and the other part, in the magnetic field of the inductor.

Now let us derive an expression for the natural frequency of the LC oscillation. From the conservation of energy principle, the original energy U_E stored in the electric field of the capacitor oscillates between the capacitor and the inductor. At any instant, the total energy of the system is equal to the sum of the U_E stored in the capacitor and the energy U_B stored in the inductor. Thus,

$$U_T = U_E + U_B = \frac{q^2}{2C} + \frac{1}{2}LI_0^2 \qquad (9.17)$$

where q is the charge on the capacitor at that instant and $I = dI/dt$ is the current flowing through the capacitor. Since total energy is a constant, taking the time derivatives of equation 9.17, we obtain,

$$\frac{U_T}{dt} = 0 = \frac{q}{C}\frac{dq}{dt} + LI\frac{dI}{dt} \qquad (9.18)$$

But we know that, $$I = \frac{dq}{dt}$$

Therefore, $$\frac{dI}{dt} = \frac{d^2q}{dt^2}$$

Substituting these expressions into equation 9.18, we obtain

$$\frac{d^2q}{dt^2} + \frac{q}{LC} = 0 \qquad (9.19)$$

Equation 9.19 is the differential equation that describes the oscillations of an LC circuit. Let us recall our study of simple harmonic motion and the differential equation that describe the oscillation of the mass spring system. The equation is

$$\frac{d^2x}{dt^2} + \frac{k}{m}x = 0 \qquad (9.20)$$

where x represents the displacement from equilibrium of the mass m on the end of a spring whose spring constant is k. Since both equations (equations 9.19 and 9.20) have the same form except for the variables, they will have the same solutions. The solution to Equation 9.20 from our study of mechanics is

$$x = x_m \cos[\omega t + \phi]$$

Similarly, $$q = q_m \cos[\omega t + \phi] \qquad (9.21)$$

where q_m is the maximum charge on the capacitor.

Now let us test whether equation 9.21 is indeed a solution to Equation 9.19 by substituting it and its second derivative into equation 9.19. Thus,

$$\frac{dq}{dt} = q_m \omega \sin[\omega t + \phi]$$

$$\frac{d^2q}{dt^2} = -q_m \omega^2 \cos[\omega t + \phi] \qquad (9.22)$$

Substituting Equation 9.21 and 9.22 into Equation 9.19 we obtain

$$-\omega^2 q_m \cos[\omega t + \phi] + \frac{1}{LC} q_m \cos[\omega t + \phi] = 0$$

$$\left[\frac{1}{LC} - \omega^2\right] q_m \cos[\omega t + \phi] = 0$$

The above expression can be zero at all times if only $\left[\frac{1}{LC} - \omega^2\right] = 0$

Therefore, $$\omega = \frac{1}{LC} \qquad (9.23)$$

Thus, if ω is given by equation 9.23, then equation 9.21 is indeed a solution of equation 9.19. Therefore, the charge on the capacitor in an LC circuit oscillates sinusoidally. It is important to note that the angular frequency ω of the oscillation depends only on the inductance and the capacitance of the circuit.

Since the charge varies sinusoidally, the current also varies sinusoidally. We can show this by taking the time derivative of equation 9.21. Thus,

$$I = \frac{dq}{dt} = -\omega q_m \sin[\omega t + \phi] \qquad (9.24)$$

Let us determine the phase angle for the ideal case that we are considering. Our situation requires that at $t = 0$, $I = 0$ and $q = q_m$. Setting $I = 0$ at $t = 0$ in equation 9.24, we obtain

$$0 = -\omega q_m \sin\phi$$

Which case that $\phi = 0$. Therefore, for this ideal case, the time variation of q, and that of I are given by

$$q = q_m \cos \omega t \qquad (9.25)$$

$$I = -I_m \sin \omega t \qquad (9.26)$$

Where $I_m = \omega q_m$ in equation 9.26.

Equation 9.25 and 9.26 are plotted in Figure 9.6. We notice in Figure 9.6(a) that the charge on the capacitor oscillate between q_m and $-q_m$ and the current (Figure 9.6 (b)) oscillates between I_m and $-I_m$. The current is 90° out of phase with the charge.

Let us return to energy considerations of the circuit. We know from equation 9.15 that the energy stored in the electric field of the capacitor at time t is

$$U_E = \frac{q^2}{2C}$$

Substituting equation 9.25 into the above expression, we have $U_E = \frac{q_m^2}{2C}\cos^2(\omega t)$.

Similarly, energy stored in the magnetic field of the inductor (after substituting equation 9.26) at the same time is $U_B = \frac{LI_m^2}{2}\sin^2(\omega t)$.

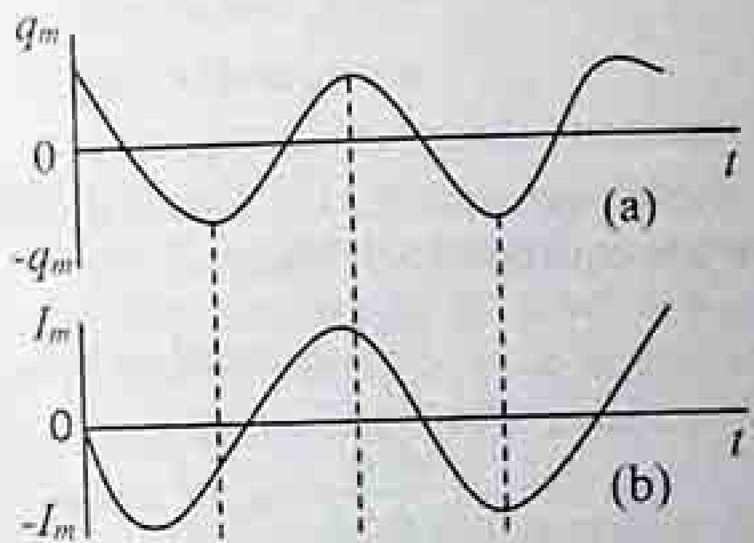


Figure 9.6: plot of charge versus time and current versus time for an ideal LC circuit.

Then from Equation 9.17, we obtain the total energy of the system. Thus,

$$U_T = U_E + U_B = \frac{q_m^2}{2C}\cos^2(\omega t) + \frac{LI_m^2}{2}\sin^2(\omega t) \qquad (9.27)$$

Equation 9.27 shows that the total energy of the system continuously oscillates between the energy stored in the electric field and the field of the inductor. When the energy in the capacitor is maximum $(q^2/2C)$, the energy in the inductor U_B, is zero, and when the energy in the inductor is maximum $(LI^2/2)$ the energy in the capacitor, U_E is zero. However, the sum of U_B and U_E is constant and is equal to $q_m^2/2C$.

Example 9.5

At $t = 0$, $q = q_m$ and $I = 0$ in an LC circuit. At the first moment when the energy is shared equally by the inductor and the capacitor, what is the charge on the capacitor?

Solution

When $U_E = U_B$, we have $\frac{1}{2}LI^2 = \frac{q^2}{2C}$

$$q = \sqrt{LC} = \frac{1}{\omega}$$

For the initial conditions, $q = q_m \cos(\omega t)$

Thus, $I = -q_m \omega \sin(\omega t)$

Therefore, when $U_E = U_B$, we obtain $q_m^2 \cos^2(\omega t) = q_m^2 \sin^2(\omega t)$

$$\cos^2(\omega t) = \sin^2(\omega t)$$

$$\tan(\omega t) = 1$$

But, $\quad q = q_m \cos(\omega t)$

Therefore, we obtain $\dfrac{q_m}{\sqrt{2}}$.

9.6 LRC Circuit

In the last section, we discussed an LC circuit where we idealized that the inductor has no resistance. It is not a realistic assumption, there is always some resistance R, in any practical circuit. Now let us turn our attention to a realistic situation and connect a resistance R, in series with C and L. This is shown in Figure 9.7 this circuit is known as an LRC circuit.

As before, we give the capacitor an initial charge of q_m and when the switch is closed at $t = 0$, current I begins to flow clockwise through the resistor and inductor. Now we expect some energy to be converted to thermal energy at the resistance R. Therefore, a total transfer of energy from electric field in the capacitor to the magnetic field in the inductor (as in the case of the LC circuit) does not occur. Applying the loop rule to the circuit shown in the Figure 9.7, we have

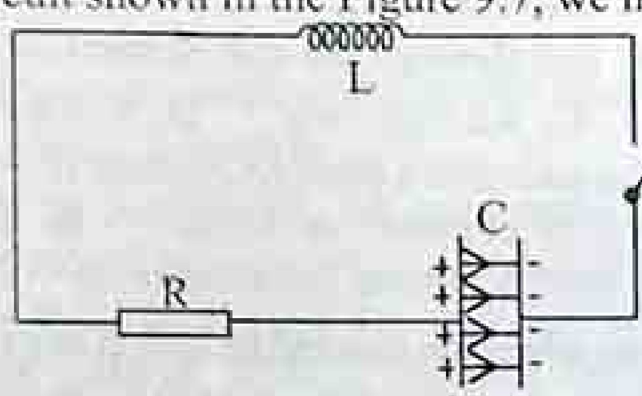


Fig. 9.7: An LRC Circuit

$$L\frac{dI}{dt} + IR + \frac{q}{C} = 0$$

Substituting dq/dt for I and d^2q/dt^2 for dI/dt in the above expression, we have

$$L\frac{d^2I}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = 0 \qquad (9.28a)$$

If we set $R = 0$, this equation reduces to equation 9.19.

Let us recall our study of damped harmonic oscillator Figure 9.9, show a mass spring system moving in a viscous medium. The equation of motion of this damped system is

$$m\frac{d^2x}{dt^2} + b\frac{dx}{dt} + kx = 0 \qquad (9.28b)$$

where b is the damping factor and k is the spring constant.

Fig. 9.8: A mass-spring system moving in a viscous medium.

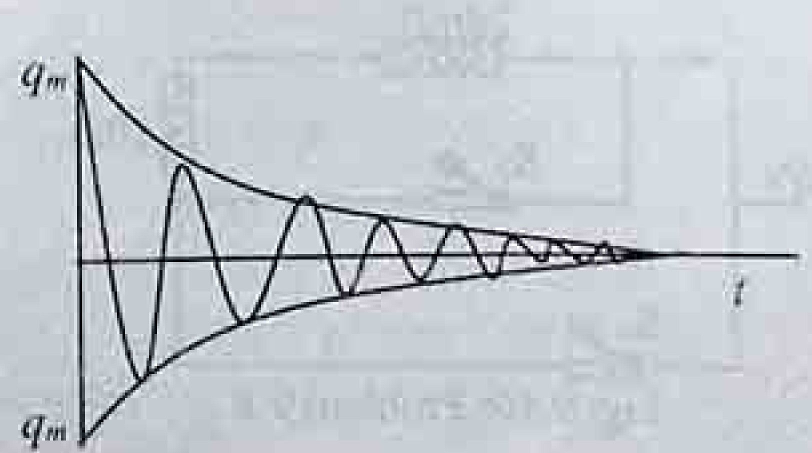


Fig. 9.9: A damped oscillation

Comparing equation 9.28(a) and 9.28(b), we observe that the two equations are mathematically identical with the following correspondences: q correspond to x: L corresponds to m: R corresponds to b and C corresponds to $1/k$. Therefore, an LRC circuit is a damped harmonic oscillator.

We have already solved equation 9.28(b) (the mathematical equivalent of LRC circuit) in our studies of mechanics, therefore, the solution of equation 9.28(a) follows at once by method of correspondence. The solution is

$$q = q_m \exp - \frac{Rt}{2L} \cos[\omega t + \phi] \qquad (9.29)$$

where $\omega = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2}$

Equation 9.29 means that the charge in an LRC circuit oscillates with damped harmonic motion analogous to a mass-spring system moving in a viscous fluid. We note that equation 9.29 reduces to equation 9.21 when $R = 0$. Then $\omega = \sqrt{1/LC}$. Equation 9.29 oscillates with decreasing amplitude. This is illustrated in Figure 9.8.

Exercises 9

9.1 The current in a coil of wire is initially zero but increases at a constant rate; after $10.0s$ it is $50.0A$. The changing current induces an e.m.f of $40.0V$ in the coil. Find the inductance of the coil. **A.** $0.8H$ **B.** $6.0H$ **C.** $8.0H$ **D.** $4.0H$

9.2 Determine the total magnetic flux through the coil in exercise 9.1 when the current is $50A$. **A.** $400Wb$ **B.** $40.0Wb$ **C.** $4.2 \times 10^{-2} Wb$ **D.** $0.004Wb$

9.3 How much resistance must be added to a pure LC circuit ($L = 200mH$, $C = 1200pF$) to change the oscillator's frequency by 0.10 percent? **A.** 115Ω **B.** 1150Ω **C.** 575Ω **D.** 2300Ω

9.4 Two nearby coils, A and B, have a mutual inductance $M = 28mH$. What is the e.m.f induced in coil A as a function of time when the current in coil B is given by $I = 3t^2 - 4t + 5$, where I is in amperes and t in seconds? **A.** $(-0.168T + 0.112)V$ **B.** $(0.168T + 0.112)V$ **C.** $(0.168T - 0.112)V$ **D.** $(-0.168T - 0.112)V$

9.5 A fixed inductance $L = 1.05\mu H$ is used in series with a variable capacitor in tuning section of a radio. What capacitance will tune the circuit into the signal from Radio Nigeria broadcasting at a frequency of $99.1MHz$? **A.** $3.732pF$ **B.** $7.332pF$ **C.** $6.245pF$ **D.** $2.456pF$

9.6 It is proposed to store $1kWh = 3.6 \times 10^6 J$ of electrical energy in a uniform magnetic field with magnitude $0.5T$. What volume must the magnetic field occupy to store this amount of energy? **A.** $36.2m^3$ **B.** $18.1m^3$ **C.** $1.8m^3$ **D.** $3.62m^3$

9.7 If instead this amount of energy is to be stored in a volume of $0.125m^3$, what magnetic field is required? **A.** $0.04T$ **B.** $2.04T$ **C.** $0.089T$ **D.** $8.51T$

9.8 For the circuit shown in Figure 9.10, if the current in the circuit at a time $0.5ms$ after the switch S is closed is $113mA$. What is the maximum value of the current in the circuit? **A.** $60A$ **B.** $600A$ **C.** $600mA$ **D.** $300mA$

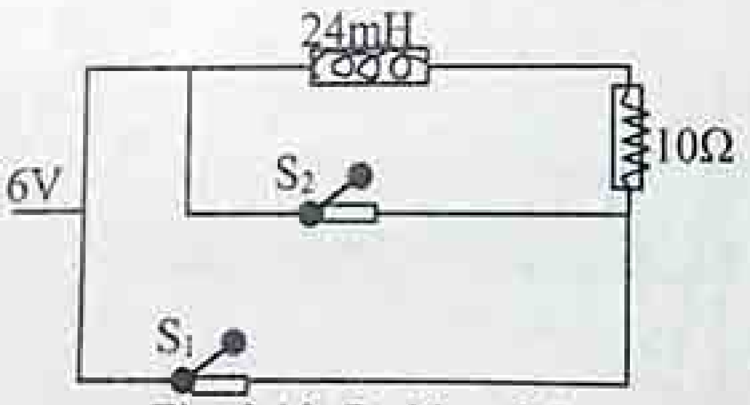


Fig. 9.10: Problem 9.8

9.9 The energy stored in an inductor carrying a current I is $U_m = \frac{1}{2}LI^3$. TRUE or FALSE

9.10 In an LR circuit, the e.m.f source is $6.3V$; the total circuit resistance is 175Ω and the current takes $58\mu s$ to build up from zero to $4.9mA$. Find the final current and inductance. **A.** $72mA$, $195\mu s$ **B.** $36mA$, $390\mu s$ **C.** $3.6mA$, $3900\mu s$ **D.** $7.2mA$, $3900\mu s$

9.11 Explain what you understand by inductance and distinguish between self inductance and mutual inductance.

9.12 The earth's magnetic field strength is approximately $10^{-4}T$ near the surface. (i) What is the magnetic field strength density? (ii) A solenoid of length $10cm$ and radius $1cm$ has 100 turns. What current would produce the energy density found in (i).

9.13 (a) Two solenoids of inductance $2mH$ and $6mH$ are connected in series. They are well separated. What is the total magnetic energy stored in them for a current in each of $1A$? (b) The two coils are connected in parallel and a total current of $1A$ enters the junction with the two inductors. The inductors are still well separated. What is the total energy stored in the inductors?

9.14 A small thin coil with N_2 loops, each of area A_2, is placed inside a long solenoid, near its centre. The solenoid has N_1 loops in its length L and area A_1. Find the mutual inductance as a function of θ, the angle between the plane of the small coil and the axis of the solenoid.

9.15 In an LRC circuit $R^2 << 4L/C$, show that the total energy stored in C and L is given by $U = \frac{q_m^2}{C}\exp-\left[\frac{Rt}{L}\right]$.

9.16 Explain what you understand by understand by the time constant of a circuit. Show that the time constant of an LR circuit is equal to L/R.

9.17 In the tuning circuit of an AM radio, the inductance is $5mH$. Over what range must the capacitance vary for the circuit to detect the AM band from $550kHz$ to $1600kHz$?

9.18 An LC circuit oscillates at $10.4kHz$ (i) if the capacitance is $340\mu F$, what is the inductance? If the maximum current is $7.20mA$, what is the total energy in the circuit? (iii) Calculate the maximum charge on the capacitor

9.19 Discuss the concept of energy storage in an inductor and show that the stored energy is proportional to the square of the current.

9.20 A fixed inductance $L = 1.05\mu H$ is used in series with a variable capacitor in the tuning section of a radio. What capacitance will tune the circuit into the signal from Radio Rivers broadcasting at a frequency of $99.1MHz$?

CHAPTER 10
MAGNETIC MEDIA

10.0 Introduction

The aim of the chapter "Magnetic Media" is to enable students improve their knowledge on the magnetic properties of matter, types of magnetic materials, magnetic properties of these magnetic materials and how to enhance or improve the magnetization of magnetic materials. Also solving questions with equations involving magnetic field, field strength, permeability, susceptibility and magnetization would be geared at.

We have seen in the previous chapters that magnetic fields are generated by electric currents. As far as we know, this is the only source of magnetic fields. The existence of magnetic monopoles- the equivalent of electric charges was predicted by P.A.M. Dirac in 1931. Physicists have searched for magnetic monopoles since Dirac's prediction, but they have not been found yet. The nonexistence of magnetic poles destroys the symmetry in Maxwell's equations. We shall say more about these equations in a later chapter. Since currents are the only sources of magnetic fields, any alteration of magnetic fields by any substance can be explained in terms of currents within the substance.

10.1 Magnetic Properties of Matter

Suppose we have a long solenoid consisting of n turns per metre and carrying a current I. The magnetic field inside the solenoid is

$$B_0 = \mu_0 nI \tag{10.1}$$

The magnetic field inside the solenoid is proportional to μ_0, the permeability of free space. If we fill the solenoid with material substance, the magnetic field B within the solenoid changes in magnitude so that the ratio B/B_0 is a constant for that particular substance. Thus,

$$\frac{B}{B_0} = K_m$$

or

$$B = B_0 K_m \tag{10.2}$$

where K_m is the relative permeability of the substance. Substituting equation 10.1 into equation 10.2 we obtain :

$$B = \mu_0 K_m nI \tag{10.3}$$

The permeability μ of the substance filling the solenoid may be defined as

$$\mu = K_m \mu_0$$

Therefore, the magnitude of the magnetic field within the filled solenoid is

$$B = \mu nI \tag{10.4}$$

In some substances, K_m is approximately unity; this means that the change from B_0 is negligible. The use of magnetic susceptibility is therefore more convenient and is defined as

$$\chi_m = K_m - 1 = \frac{B}{B_0} - 1$$

The susceptibility χ_m and K_m of a given substance are not constant but depend upon temperature. We list typical values of χ_m in Table 10.1.

A careful study of Table 10.1 shows some substances when present in the solenoid increase the magnetic field substantially (χ_m positive). These substances such as cerium, ferric chloride, liquid oxygen, iron, nickel and cobalt and their alloys are known as ferromagnetic materials.

Substances that have negative magnetic susceptibilities (χ_m negative) are diamagnetic. Bismuth, carbon (graphite), and mercury are typical examples of diamagnetic substances. If a ring of bismuth forms the core of our solenoid, the magnetic field B, will be less than B_0. Other substances such as aluminium, platinum and oxygen (gas) have very small (even though positive) magnetic susceptibilities χ. If a ring of aluminium forms the core of the solenoid, the magnitude of the magnetic field B will be slightly higher than B. It is useful to define a quantity H, the magnetic field strength as $B = \mu H$, or $B_0 = \mu_0 H$.

Table 10.1: Magnetic susceptibilities of various substances (At approximately room temperature)

Substance	$\chi_m \times 10^5$
Air	0.04
oxygen	0.18
Aluminium	2.10
Cerium	130.00
Ferric Chloride	306.00
Oxygen (liquid at -219°C)	490.00
Copper	-0.96
Lead	-1.60
Mercury	-2.80
Diamond	-2.20
Carbon (graphite)	-9.90

10.2 Ferromagnetism

Iron, nickel and cobalt are known as 3S transition elements; they and their alloys react very strongly in a magnetic field. They are ferromagnetic and have very high values of χ_m. To understand all these, let us recall that a close loop of wire carrying current I, sets up a magnetic moment $\bar{\mu}$ which is given as

$$\bar{\mu} IAn \quad (10.4b)$$

where A is the area of loop. We must not confuse the magnetic moment $\bar{\mu}$ with permeability μ. The electrons associated with the atoms or ions in these substance set tiny current loops in the substances. These current loops are known as Amperean currents. Figure 10.1 shows a cross sectional area of a magnetic substance with the current loops.

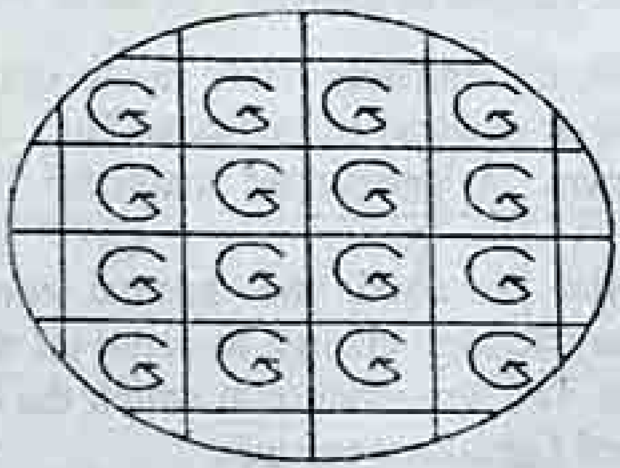


Fig 10.1: Amperean currents

Each current loop sets up magnetic moments $\bar{\mu}$, but the magnetic moment do not necessarily align in the same direction. However, in ferromagnetic materials, there exists a force known as exchange coupling between adjacent atoms that tend to align their magnetic moments together. The nature of this force can only be described by quantum mechanics. But it suffices here to know that such force exists.

The alignment of the magnetic moments of amperean current loops does occur without application of external magnetic field. This spontaneous alignment extends over a volume of about $1mm^3$ but which contains millions of atoms. The regions of complete alignment of magnetic moments are known as domains. In each domain, the magnetic moments of amperean current loops are completely aligned. However, the net magnetic moments in the domains are at random. Figure 10.2 shows schematic diagram of domain patterns of a single crystal of ferromagnetic materials. It has no overall magnetic field associated with it because the magnetic moments in the domains are randomly oriented. The domains in Figure 10.2 vary in sizes and the sample has not been subjected to a significantly large external magnetic field.

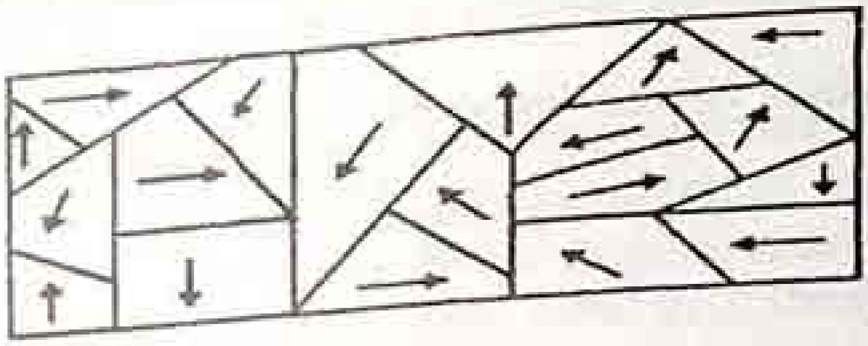


Fig. 10.2: Randomly oriented domains in a ferromagnetic material

Now let us apply an external magnetic field B_{ext} to this sample; Figure 10.3(a) shows four touching domains of unmagnetised sample. When an external magnetic field of small value is applied, as in Figure 10.3(b), the domains aligned with the field direction begin to grow at the expense of the non-favourably aligned domains. The growth of the favourably aligned domains is due to the motion of domain walls. These domain walls will continue to move until they strike an impurity in the crystal or another domain wall. As a result of this movement, more magnetic moments become aligned in the direction of the field and the net magnetic field of the domains enhances the magnetic field. When strong magnetic field is applied, in addition to considerable domain wall movement, domain rotation occurs. This is shown in Figure 10.3(c). The external field has rotated some of the domains so that their magnetic moments have components in the direction of the applied field. A further increase in the strength applied will cause complete alignment within the sample and the complete domain is said to be saturated. This is shown in Figure 10.3 (d).

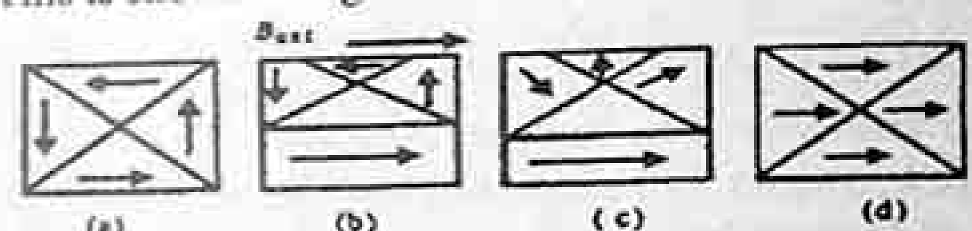


Fig. 10.3: Domains of ferromagnetic sample as magnetic fields of various magnitudes are applied.

When the external field is removed, the domains rotate in the opposite direction, the domain walls move back but not to their original orientation or position. They return to a certain orientation that leaves a resultant magnetic moment. A permanent magnetic field remains and the sample is permanently magnetized.

As the temperature of the material is raised, the thermal energy is increased. At very high temperatures, the thermal energy may be high enough to cause all the atoms to break loose from their neighbors, hence loading alignment. The temperature at which the atomic dipoles of a ferromagnetic material lose their alignment is called *Curie temperature*.

10.3 The Magnetization Curve

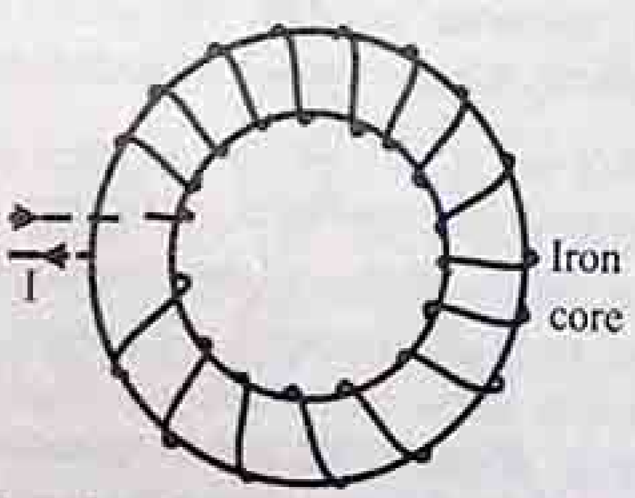


Fig. 10.4: A torus with iron core

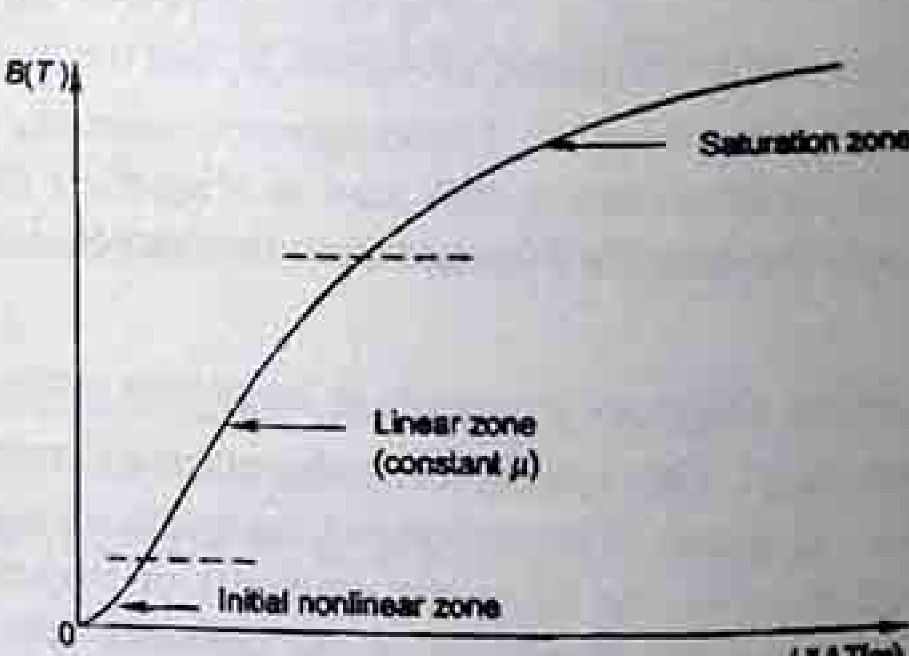


Fig. 10.5: Magnetizations curve

Now we are in a position to look, at least graphically, at the magnetic field B as H is increased. Suppose we have a torus or toroid that has iron in its core and we wish to find out how the field B in the core vanes with the magnetizing field H. This is illustrated in Figure 10.4. The iron core is initially unmagnetized and there is no current in the windings of the torus. When the current I is turned on and is slowly increased, H increases linearly with the total field, it follows the curved line shown in Figure 10.5. This is known as the magnetizations curve and it is highly non-linear. At very high applied fields, the magnetic moments are essentially all aligned with the field. The maximum B contributed by the Ferromagnetic material occurs at that point and it is known as the *saturation field*.

We plot a typical experimental curve of B versus B_0 in Figure 10.6; this figure shows the effect of past history on the magnetizations of a ferromagnetic sample.
We start with an unmagnetized material at point **O** and increase the current in our torus thereby increasing B_0. The ferromagnetic material is then magnetized along **OA** until saturation is reached at point **A**. After reaching point **A**, the current is decreased. Reducing B_0 to zero, the magnetic field B also decreases but along a different part. When B_0 is zero, B has the value called *remanence* indicated at point **C** in Figure 10.6.

This is so because the domain walls and orientations have not returned to their original positions or orientations. This gives the state of permanent magnetization of the sample. To reduce B in the toroid to zero, we must forcibly dis-align the domains by using a reverse external field. Point **D** gives the *coercive force*- the reverse field necessary to demagnetize the sample. As we continue in the reverse direction, B follows **DE** path until it reaches saturation in the opposite direction at point **E**. If the current is again reduced to zero and then increased in the original direction, the total field follows the path **EGKA**, and approaches again saturation at point **A**.

We note that the curve did not pass through the origin at point 0. The curve **ACDEGKA** is called a hysteresis loop. In such a cycle, a lot of energy is used up in realigning the domains; this energy is proportional to the area under the hysteresis loop.

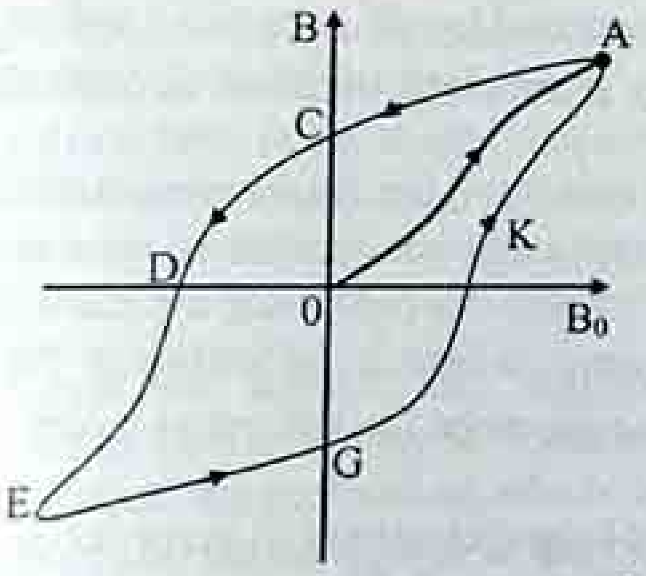


Fig. 10.6: Hysteresis Curve

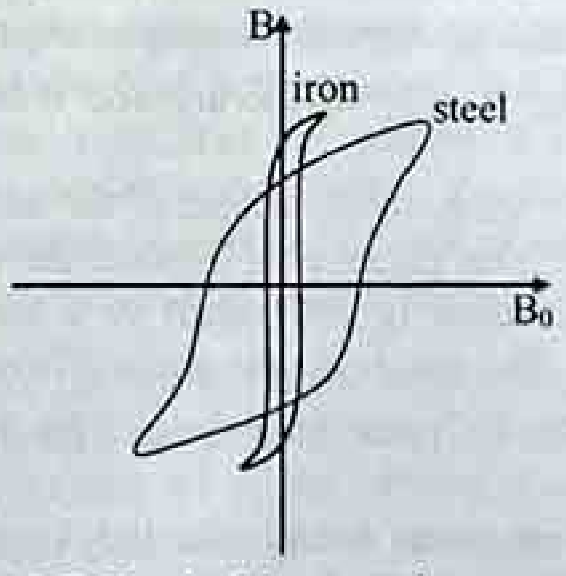


Fig. 10.7: Typical hysteresis curves for iron and steel

10.4 Properties of Magnetic Materials

In Figure 10.7, we show typical hysteresis loops for iron and steel. Notice that the coercive force for steel is much higher, therefore, steel is more suitable for permanent magnets because it will not be easily demagnetized by shaking. The fact that the remanence of iron is a little greater than that of steel is completely overweighed by its smaller coercivity which makes it easy to demagnetize. However, iron is more suitable for electromagnets which have to be switched on and off as in relays. Even the cores of transformers and the armatures of machines are also better made using iron. These machines go through the magnetizing cycles continuously. The work done by the field against the internal friction of the domains is dissipated as heat. *The heat dissipated in this way per cycle is known as hysteresis loss and is proportional to the area of the hysteresis loop.*

In a large transformer, the hysteresis loss and the heat developed by the current in the resistance of the windings is so large that the transformer must be cooled by circulating oil which itself is cooled by the atmosphere. Table 10.2 gives the properties of some typical magnetic materials.

Table 10.2: Properties of some magnetic materials

	Materials	Relative Permeability K_m	Remanence 10^{-4}T	Coercivity H_c Am^{-1}	Hysteresis Loss Jm^{-3} per cycle	Application
Soft Materials	Iron	5500	13000	81.0	500	Armature, relays,
	Nickel	600	3600	272	30	Transformer
	Cobalt	240	5000	800	200	Core,
	Silicon-iron	6700	12000	40	350	telephone diaphragms
	Mumetal	80000	6000	4	20	magnetic shields
Hard Materials	Carbon	-	8000	4800	20000	loudspeakers, Microphones,
	Steel	-	8000-10000	12000-19200	-	Telephone earpiece
	Cobalt	-	12500	44000	-	magnets

10.5 Paramagnetism and Diamagnetism

From our earlier studies, we know that an e.m.f. is induced in current loops by changing magnetic fields. The induced e.m.f. causes induced current in the loop; and according to Lenz's law this induced current sets up a magnetic field that opposes the changing magnetic field. In order words, if we impress a magnetic field B upon a loop, the induced current in the Loop will generate an opposite field B^1. Therefore, the field through the loop is always decreased by B^1, the field generated by the loop. This effect of course, is only temporary and it lasts only while the flux through the loop is changing.

We shall see later when we study the structure of the atom that electrons travel around the nucleus in circular orbits. These electrons circulating around the nucleus can be considered as current loops. Each circulating electron which is equivalent to a tiny current loop has an associated magnetic moment of magnitude $\bar{\mu}$. The magnetic moments of the many circulating electrons in the substance are randomly oriented with the result that the net magnetic moment for the orbits taken together is zero.

Now if we apply external magnetic field to the tiny current loops; electrons in some loops will speed up while others will loose speed. Their orbits will remain unchanged. The magnitudes of associated magnetic moments will also increase from $\bar{\mu}$ to $(\bar{\mu}+\Delta\bar{\mu})$ while others will decrease from $\bar{\mu}$ to $(\bar{\mu}-\Delta\bar{\mu})$.

The net effect, therefore, is a non-zero magnetic moment directed opposite the applied field. From this we see that all atoms will tend to reduce the magnetic field impressed upon them. Substances in which this effect is dominant are known as diamagnetic substances. The magnetic susceptibilities of such substances are negative as mentioned earlier.

We have seen that magnetic effects of electrons in certain atoms exactly cancel so that the atoms are not magnetic. But this does not hold true for other atoms such the rare earths. In these atoms, the magnetic effects of the electrons do not cancel, so as a whole has a magnetic dipole moment $\bar{\mu}$. Such atoms exhibit the property of paramagnetism. When an external field is applied to an atom with a permanent magnetic moment μ, the applied field tends to align the magnetic moment with the field. As a result, when a magnetic field is applied a substance containing N atoms, each with permanent magnetic moment $\bar{\mu}$, $N\bar{\mu}$ magnetic dipole moments will align with the field. This will increase the measured magnetic field slightly. However, thermal motion of atoms prevents complete alignment of the magnetic dipole moments.

The atoms which have permanent magnetic dipole moments are also subject to diamagnetic effect. But usually the magnetic effects predominate. However, the paramagnetic tendencies are temperature dependent and decrease with increasing temperature. Diamagnetic on the other hand, occurs both at low and high temperatures.

We can determine the extent to which a substance subjected to an external magnetic field is magnetized by dividing its measured magnetic dipole moment by its volume. This useful quantity is known as magnetization $\vec{M}$, it is the magnetic dipole moment per unit volume. This can be written as

$$\vec{M} = \frac{\vec{\mu}}{V}$$

where $\vec{\mu}$ is the magnetic dipole moment of the sample and V is its volume. Pierre Curie, in 1895, discovered experimentally that the magnetization $\vec{M}$ of a sample is directly proportional to the magnetic field B that is trying to align the dipole and inversely proportional to the temperature (the temperature expressed in Kelvin). Thus,

$$\vec{M} = C\frac{B}{T} \qquad (10.5)$$

where C is a constant. Equation 10.5 is known as [illegible] and the constant C is called Curie's constant. This law makes sense because increasing B tends to align the magnetic dipoles in the sample, this increases $\vec{M}$ while increasing the temperature. T tends to randomize the magnetic dipole directions thereby reducing $\vec{M}$. The magnetization $\vec{M}$ does not increase unlimitedly but approaches a maximum $\vec{M}_{max}$; this corresponds to a complete alignment of all the dipoles in the sample.

Electron spin alignment also plays a role in determining the magnetization properties of a material. If the spin alignment of neighboring atoms is parallel, then the material is ferromagnetic. If the favored spin alignment results in a net zero macroscopic magnetic moment, then the material is said to be anti-ferromagnetic. If the spin structure is comprised of both spin-up and spin-down components that results in a non-zero macroscopic magnetic moment, then the material is said to be ferromagnetic.

In a perfect magnetic material, all electronic motion is confined to the atomic structure in the form of atomic currents and their moments are randomly oriented in the absence of an external applied magnetic field. They are then either paramagnetic or diamagnetic materials.

Ferro-magnets will tend to stay magnetized to some extent after being subjected to an external magnetic field. This tendency to "remember their magnetic history" is called *hysteresis*. The fraction of the saturation magnetization which is retained when the driving field is removed is called the remanence of the material, and is the basis of data storage on audio and video tapes and computer hard drives. The recording head of a tape recorder or the write head of a disk drive applies a field that magnetizes a small portion of the tape (or disk). The magnetism in each portion remains until another magnetic field changes it. When each magnetized section is moved under the playback head or read head, the moving magnetic field induces small currents which are amplified and turned into either music or data bits. If the domains were unable to remember the field that had been applied to them, none of this would be possible. The magnetic domains will remain aligned until randomized by thermal agitation or by some other external force which can do work in rotating the domains within the material. (For example, heating up a magnet or whacking it with a hammer can remove the material's magnetic effects!)

Ferromagnetic materials will respond mechanically to an impressed magnetic field, changing length slightly in the direction of the applied field. This property, called magnetostriction, leads to the familiar humming of transformers as they respond mechanically to $60Hz$ AC voltages. In paramagnetic materials, the electron orbits do not cancel out, but the electron fields don't reinforce each other as much as those in ferromagnetic materials. They therefore have permanent dipole moments that try to line up with the magnetic field, but are prevented from remaining aligned by random thermal motion. When a paramagnetic material is placed in a strong magnetic field, it becomes a magnet, and as long as the strong magnetic field is present, it will attract and repel other magnets in the usual way. But when the strong magnetic field is removed, the net magnetic alignment is lost as the dipoles relax back to their normal random motion.

Summary

1. If a solenoid is filled with magnetic material, the magnetic field changes to B from the original field B_0. Hence,

$$B = K_m B_0$$

where K_m is called relative permeability of the substance. For diamagnetic materials $K_m < 1$, for paramagnetic materials $K_m > 1$, and for ferromagnetic materials $K_m >> 1$. Also, $K_m = \mu / \mu_0$.

2. Magnetic susceptibility of a substance χ_m is defined as $\chi_m = K_m - 1 = \dfrac{B}{B_0} - 1$.

3. The electric field strength H is defined as $H = \dfrac{B_0}{\mu_0} = \dfrac{B}{\mu}$.

4. The magnetic properties of materials can be explained in terms of permanent magnetic dipole moments. The small region in which all the atoms have their magnetic moments aligned are called domains. The alignment of these domains in the direction of an applied field results in ferromagnetism.

5. The temperature at which the atomic dipoles of a ferromagnetic material lose their alignment is called the Curie temperature. Above the Curie temperature, the substance exhibits only paramagnetic properties.

6. According to the Curie law, for paramagnetic materials, the magnetization $\bar{M}$ is given by

$$\bar{M} = C\frac{B}{T}$$

where C is Curie constant.

7. The plot of B against $H(B_0/\mu_0)$ gives a closed loop called hysteresis loop. The area within a hysteresis loop is a measure of the energy per unit volume lost during a cycle of magnetization.

8. From a study of the hysteresis loop, it is concluded that while steel is adequate for permanent magnets, iron is more suited for electromagnets and transformer cores.

Exercise 10

10.1 The relative permeability of a panama genetic substance is
A. Unity B. Slightly more than unity C. Zero D. Less than unity

10.2 Ferroelectric materials are characterized by
A. Very high degree of pollination B. A sharp dependence of organization on temperature. C. Non-linear dependence of the charge q on the applied voltage D. All of the above

10.3 Which of the following metals is not affected by magnets?
A. Cobalt B. Nickel C. Iron D. Tin

10.4 Which of the following statements about making magnets is true?
A. Soft iron object that has been magnetized retain its magnetism.
B. A steel object that has been magnetized retains its magnetism.
C. Stroking the object back and forth instead of stroking it in only one direction.
D. Steel is easier to magnetize than iron.

10.5 The relative permeability of cobalt is 240 and that of copper is -0.96. What is the ratio of their magnetic susceptibility? A. 293/-1.96 B. -121.94 C. 239 D. -0.96

10. 6 Due to non-existence of magnetic dipoles, magnetic fields are generated by
A. Potential difference B. Battery C. Electric current D. Waves

10.7 A substance has magnetization property M, and a magnetic field of $0.6T$ at a temperature of 10°C. Find M? (Take the curie constant to be 1.5K)

A. $3.18\times10^{-3}Am^{-1}$ **B.** $3.18\times10^{3}Am^{-1}$ **C.** $0.09Am^{-1}$ **D.** $9.0\times10^{-2}Am^{-1}$

10.8 A long solenoid having 500 turns carrying current of $2.5A$ is filled with a magnetic material with a magnetic field of $0.85T$. Find the relative permeability of the magnetic material.

A. 6.8×10^{4} **B.** 1.06×10^{3} **C.** 1.06×10^{-3} **D.** 6.8×10^{-4}

10.9 What is the permeability of iron filling solenoid, if its relative permeability is 5500? ($\mu_0=1.26\times10^{-6}$). **A.** 5500 **B.** 6.93×10^{-3} **C.** 1.26×10^{-6} **D.** 4.37×10^{9}

10.10 Magnetic Susceptibility is related to the relative permeability of a magnetic material by

A. $\chi_m=K_m$ **B.** $\chi_m+K_m=1$ **C.** $K_m=1/\chi_m$ **D.** $\chi_m-K_m=1$

10.11 A solenoid produces a magnetic field of $0.02T$ when there is a current of $2A$ and the total number of turns is 4000. Find the length of the solenoid?

10.12 Define the magnetic susceptibility of a substance and hence distinguish between diamagnetic, paramagnetic and ferromagnetic materials.

10.13 Consider a toroid with 1000 turns and a radius of $25cm$. If there is a current of $1250A$, what will be the magnetic field inside the toroid?

10.14 The magnetic susceptibility of iron ammonium alum is 7.54×10^{-3}. Determine its relative permeability and its permeability?

10.15 Define intensity of magnetization and magnetic susceptibility, and with the aid of a diagram, explain what is meant by hysteresis.

$B_o(10^{-4}T)$	0	0.13	0.25	0.5	0.63	0.78	1.3
$B(T)$	0	0.0042	0.010	0.028	0.043	0.095	0.67
$B_o(10^{-4}T)$	1.9	2.5	6.3	13.0	130	1300	104
$B(T)$	1.01	1.18	1.44	1.58	1.72	2.26	3.15

10.16 Draw the diagram of a typical hysteresisloop for steel and comment on the features. Use the curve to explain what you understand by retentivity and coercive fore.

10.17 Discuss the domain theory of magnetization.

10.18 The relative permeability of bismuth, iron ammonium alum, liquid oxygen and nitrogen gas are 0.9999832, 1.00754, 1.00385 and 0.9999957 respectively. Find out which of the materials are paramagnetic and which are diamagnetic?

10.19 State and explain the Curie law.

10.20 Draw and explain the magnetization curve for a material. What is the significance of the saturation field? What are the characteristics required of a material for use as a permanent magnet?

CHAPTER 11
ALTERNATING CURRENT

11.0 Introduction

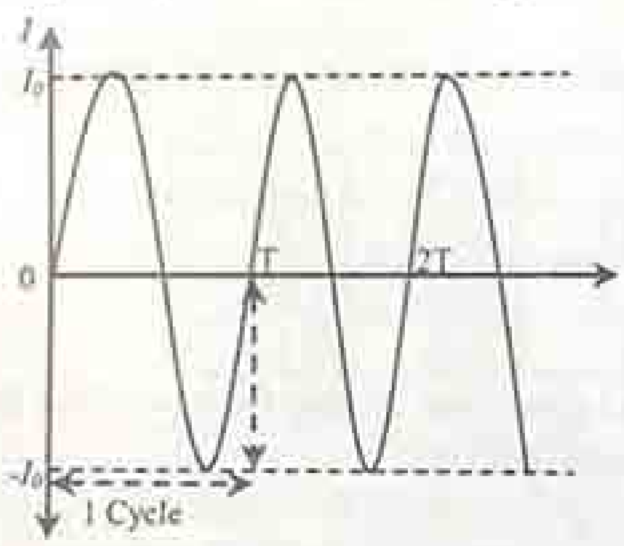
Resistors, Capacitors and Inductors all play vital roles in modern electronic circuits. In some cases, they were connected to direct current source; in other cases, they were not connected to any source at all, as in the case of capacitor discharging across inductor in LC circuit. In this chapter, we shall consider the behaviour of these elements when connected to a sinusoidal source of alternating current like those from generating plants.

11.1 Definition of Alternating current

A current that is an oscillating function of time is called an alternating current, a.c. There are various forms but the one which we will be concerned with in this write up, which is of practical important to homes and industries is sinusoidal alternating current.

For example, a coil of wire rotating with constant angular velocity in a magnetic field develops a sinusoidal alternating current. The main feature of a sinusoidal alternating current is that the direction of flow and the magnitude of current changes periodically. The potential difference, V that supplies the alternating current also varies in the same manner. See Figure 11.1 (a) and (b).

As shown in the figure, a complete change of current or p.d from a particular value and back to the same value in the same direction is known as a cycle. The time taken for a complete cycle is known as the period, T of the alternating current or p.d.



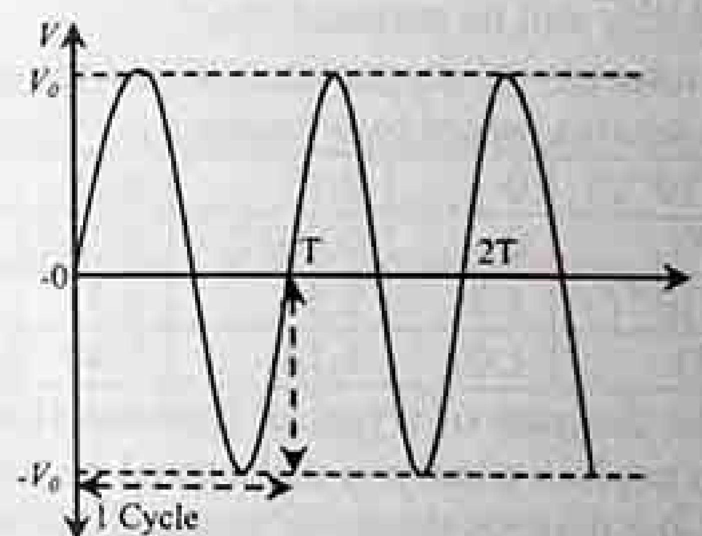


Fig. 11.1(a): Sinusoidal alternating current. Fig. 11.1(b): Sinusoidal alternating p.d

The number of complete cycles in one second is known as frequency f, of the alternating current. In most countries, the value of $f = 50Hz$. The period and the frequency are related by

$$T = \frac{1}{f} \qquad (11.1)$$

The peak current I_0 is the maximum current. The sinusoidal current and p.d can be represented by

$$I = I_0 \sin(2\pi f t) \qquad (11.2)$$

$$V = V_0 \sin(2\pi f t) \qquad (11.3)$$

(Remember that $\omega = 2\pi f$)

11.2 The root mean square value of alternating current, I_{rms}

We define the root mean square (r.m.s) value of an alternating current, I_{rms} as the value of the steady current which when flowing through the same resistor produces heat at the same rate as the mean rate produced by the alternating current. It can be proved that, these are related by the equation

$$I_{rms} = \frac{I_0}{\sqrt{2}} \qquad (11.4)$$

$$V_{rms} = \frac{V_0}{\sqrt{2}} \qquad (11.5)$$

They are considered as effective current and voltages respectively. Proofs of equations can be seen in Appendix 1.

The voltmeter in an a.c. circuit usually reads the rms of the voltage. The line voltage in Nigeria is 240 volts which is the rms value of the voltage.

Example 11.1

In engineering practice, if a supply voltage is rated $240V$, what does it mean?

Solution

This means that $V_{rms} = 240V$

The peak voltage p.d, $V_0 = V_{rms} \times \sqrt{2} = 240\sqrt{2}V$

Similarly, if a device is rated $10A$, it can safely carry up to $10 \times \sqrt{2} = 10\sqrt{2} = 1.4A$

Example 11.2

An alternating current of r.m.s value $3.0A$ and frequency $50Hz$ flows in circuit containing resistor. What is the peak current? What is the value of the current $10^{-3}s$ after it changes direction? Leave your answer in π.

Solution

Peak current $I_0 = \sqrt{2}I_{rms} = \sqrt{2} \times 3 = 3\sqrt{2}A$

From equation 11.2: $I = I_0 \sin(2\pi ft) = 3\sqrt{2} \times \sin(\pi \times 50 \times 10^{-3}) = 3\sqrt{2} \times \sin(0.05\pi)A$

Example 11.3

A voltage source delivers at $240V_{rms}$ value to a $300W$ hot wire. Calculate:

(a) The peak voltage V_0.

(b) The resistance of this hot wire.

(c) The r.m.s value of electric current I_{rms}.

(d) The peak value of current I_0.

Solution

(a) Peak voltage $V_0 = \sqrt{2}V_{rms} = \sqrt{2} \times 240 = 339.4V$

(b) Power $= \frac{V_{rms}^2}{R}$, $\therefore 300 = \frac{240^2}{R}$, $\quad R = 192\Omega$

(c) $I_{rms} = \frac{V_{rms}^2}{R} = \frac{240}{192} = 1.25$ A

(d) $I_o = \sqrt{2}I_{rms} = \sqrt{2} \times 1.25 = 1.76A$

11.3 Uses of Alternating Current

Alternating currents are of utmost importance in technology, industry and medicine. Transmission of power over long distances is very much easier and more economical with alternating current than with direct current. Alternating voltages can be stepped up or down by means of transformer. All the appliances connected to ordinary outlets in homes involve circuits with oscillating currents. Even electronic devices such as radio receivers, transmitters and communication equipment involve a variety of circuits with oscillating currents of high frequency. Many of these circuits have natural frequencies of oscillation. These circuits exhibit the phenomenon of resonance when the natural matches the frequency of a signal applied to the circuit. For example, the tuning of a radio relies on an oscillating circuit whose frequency of oscillation is adjusted by means of a variable capacitor (connected to the tuning knob) so that it matches the frequency of the radio signal.

Many life processes involve alternating voltages and currents. The beating of the heart induces alternating currents in the surrounding tissues; the detection and study of these currents called electrocardiography, provide doctors with valuable information concerning the health or pathology of the heart, Electroncephalograms, which are recordings of alternating currents in the brain function. Both electrocadiograms and electrocephalograms are invaluable diagnostic tools in modern medicine.

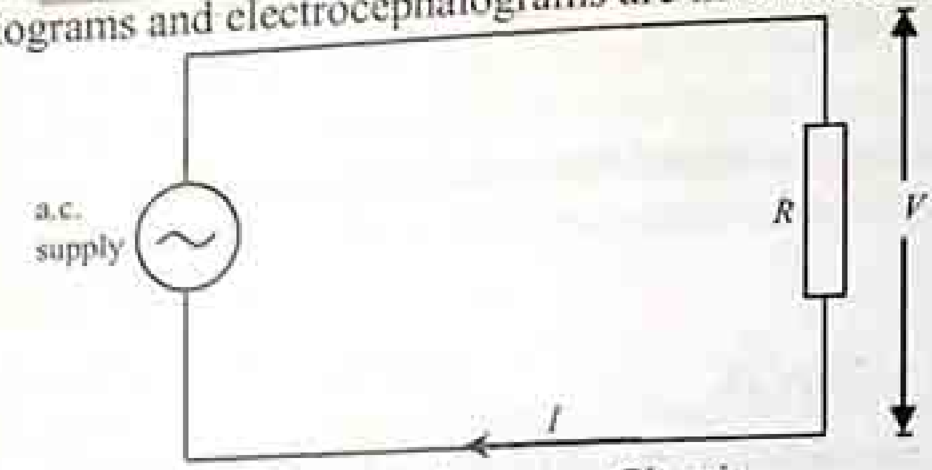


Fig. 11.2: Resistor *a.c.* Circuit

11.4 Pure Resistors in a.c Circuit

A pure resistor is one whose capacitive and inductive effects are negligible. Figure 11.2 shows a pure resistor in an a.c. circuit. If the current in the circuit is $I = I_0 \sin(2\pi ft)$, the potential difference across the resistor of resistance R is

$$V = I_0 R \sin(2\pi ft) = V_0 \sin(2\pi ft) \tag{11.6}$$

$$V_0 = I_0 R \tag{11.7}$$

11.5 Pure Capacitor in a.c. circuit

A pure capacitor has no resistance or self induction. Suppose that the alternating p.d. applied to the capacitor is given by,

$$V = V_0 \sin(2\pi ft) \tag{11.8}$$

Then the instantaneous charge on the capacitor is

$$Q = CV = CV_0 \sin(2\pi ft)$$

From the equation, $I = dQ/dt$, the current in the circuit is

$$I = \frac{d}{dt}(CV_0 \sin(2\pi ft)) = CV_0(2\pi f)\cos(2\pi ft) = I_0 \sin\left(2\pi ft + \frac{\pi}{2}\right) \tag{11.9}$$

$$\frac{V_0}{I_0} = \frac{1}{2\pi fC} \tag{11.10}$$

where $I_0 = CV_0(2\pi ft)$ is the peak current.

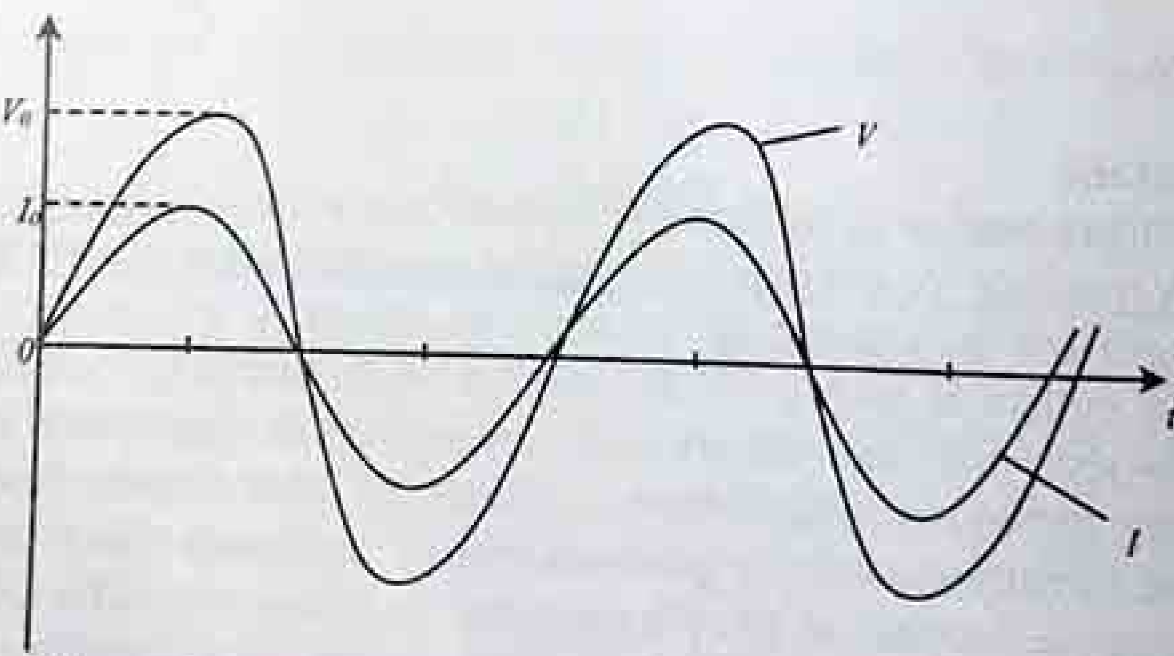


Fig. 11.3: The variation of the current I and the potential difference V with time t. The current and potential difference are in phase with each other, both achieving their respective maximum at the same instant. A phasor diagram can be drawn to visualize the phase of two alternating quantities such as the alternating potential difference.

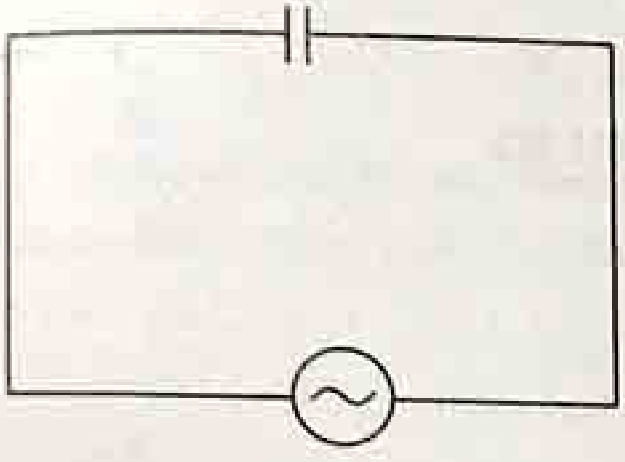


Fig. 11.4: Acapacitor in an a.c. circuit

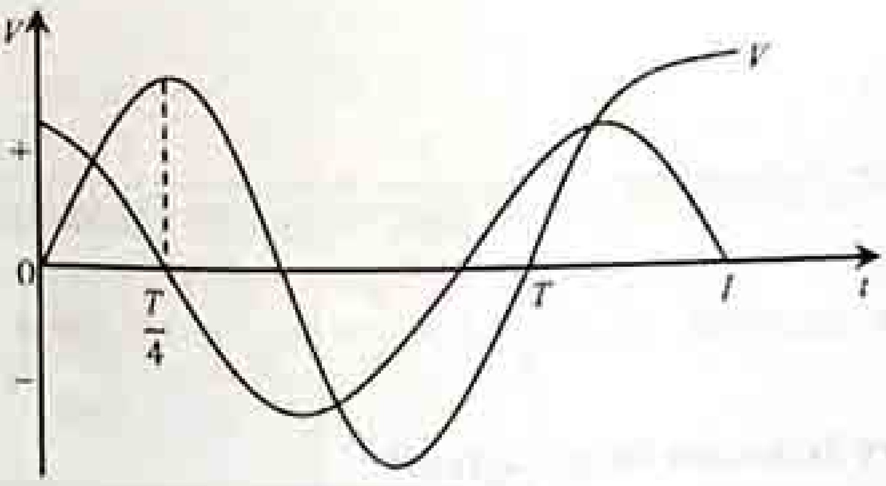


Fig. 11.5: The Current (I) leads the $p.d.$ (V) by $\pi/2$ radius

Comparing equations 11.8 and 11.9, it can be deduced that the potential difference V and the current I are not in phase. There is a phase difference of $\pi/2$ radians between I and V. The current I leads the p.d. V by $\pi/2$ radians since $I = I_0 \sin(2\pi ft + \pi/2)$.
Alternatively, the p.d. V is said to lag behind the current by $\pi/2$ radians.
Figure 11.5 is the graph of I against t and V against t. The current is at its peak when $t = 0$ whereas the p.d. is at its peak only after a time $T/4$ which is equivalent to a phase difference of $\pi/2$ radians. Hence the current I leads the p.d. V by $\pi/2$ radians. In summary, since

$$I_0 = CV_0(2\pi f)$$

$$\frac{V_0}{I_0} = \frac{1}{2\pi fC}$$

Also, $V_0 = \sqrt{2}V_{rms}$ and $I_0 = \sqrt{2}I_{rms}$

Therefore,
$$\frac{V_0}{I_0} = \frac{\sqrt{2}V_{rms}}{\sqrt{2}I_{rms}} = \frac{V_{rms}}{I_{rms}}$$

This ratio is known as the *reactance*, X_c of the capacitor. That is, the reactance X_c of the capacitor is given by

$$X_c = \frac{V_0}{I_0} = \frac{V_{rms}}{I_{ms}} = \frac{1}{2\pi fC} \qquad (11.11)$$

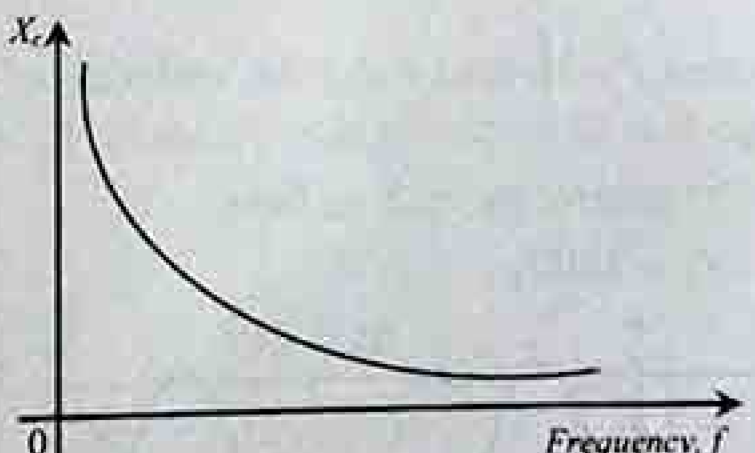


Fig. 11.6: Variation of X_c with frequency, f

The variation of the capacitor reactance X_c with frequency is shown in Figure 11.6. The unit for reactance is the ohm (Ω).

Example 11.4
A $100\,\mu F$ capacitor is connected to an a.c. potential difference of $12V$ and frequency $50Hz$. Calculate
(i) the reactance of the capacitor,
(ii) the current in the circuit.

Solution

(i) Reactance of capacitor, $X_c = \dfrac{1}{2\pi fC} = \dfrac{1}{2\pi \times 50 \times (100 \times 10^{-6})} = 31.8\Omega$

(ii) From the equation, $X_c = \dfrac{V}{I}$; $I = \dfrac{V}{X_c} = \dfrac{12}{31.8} = 0.336\,A$

11.6 Pure Inductor in a.c. circuit

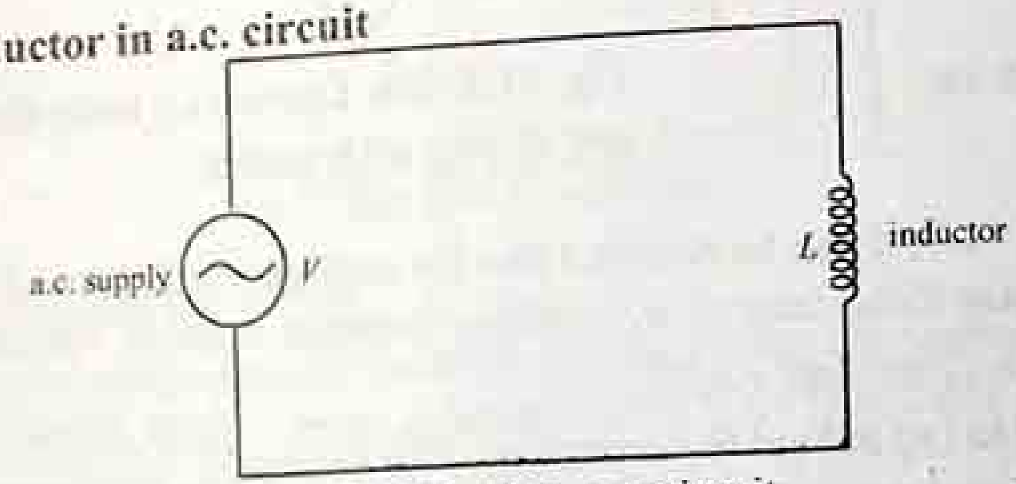


Fig. 11.7: Induction *a.c.* circuit

A pure inductor has no resistance or capacitor effect. Figure 11.7 shows the connection.
Suppose the sinusoidal current flowing in the pure inductor is

$$I = I_0 \sin(2\pi ft) \qquad (1.12)$$

When a changing current flows in the inductor, a back e.m.f. is produced.

Back e.m.f $= -L\dfrac{dI}{dt}$.

Applying Kirchoff's second law to the circuit in Figure 11.7, $V - L\dfrac{dI}{dt} = 0$.

Therefore, $\quad V - L\dfrac{dI}{dt} = 0 = L\dfrac{d}{dt}(I_0 \sin(2\pi ft)) = 2\pi fLI_0 \cos(2\pi ft)$

$$V = V_0 \sin(2\pi ft + \pi/2) \qquad (11.13)$$

where the peak voltage $\quad V_0 = 2\pi fLI_0 \qquad (11.14)$

Comparing equations 11.12 and 11.13, it can be deduced that the voltage V leads the current I by $\pi/2$ radians.

The variation of the current I with time t and the variation of the voltage V with time t are shown in Figure 11.8. The voltage V is at its peak value I_0, only ¼ of a cycle later. This illustrates the fact that the voltage V leads the current by $\pi/2$ radians.

From equation 11.14, $\quad V_0 = 2\pi fLI_0$,

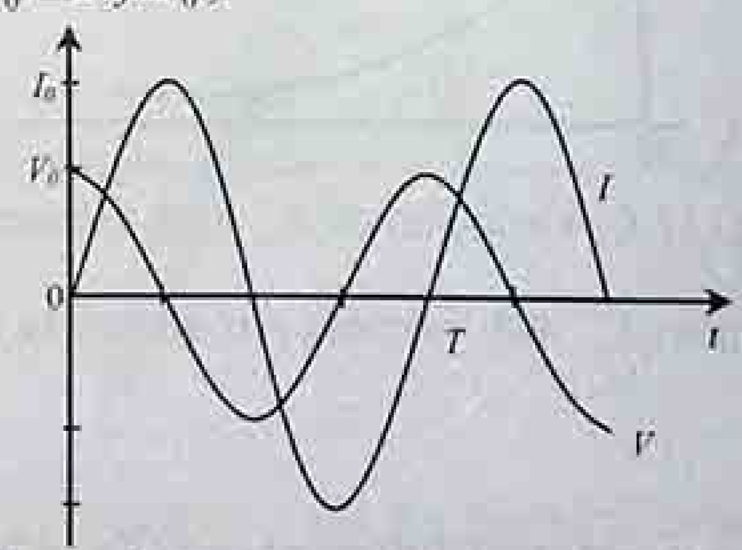


Fig. 11.8: Voltage leads current

the reactance X_L of the inductor is given by

$$X_L = \frac{V_0}{I_0} = \frac{V_{rms}}{I_{rms}} = 2\pi fL \qquad (11.15)$$

Figure 11.9 shows how the reactance X_L of the inductor varies with the frequency. If the r.m.s value of the sinusoidal voltage applied to the inductor is constant, then the r.m.s. current

$$I_{rms} = \frac{V_{rms}}{X_L} = \frac{V_{rms}}{2\pi fL} \propto \frac{1}{f}$$

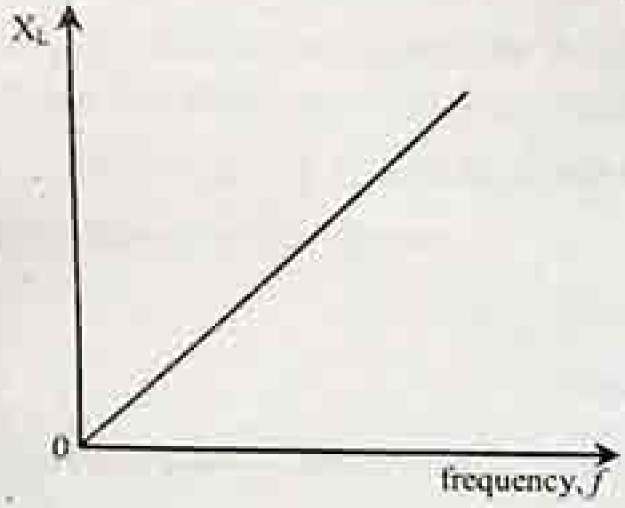


Fig. 11.9: Plot of X_L against f

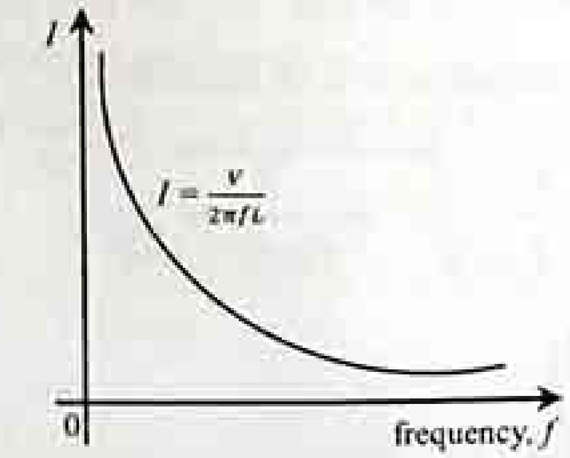


Fig. 11.10: Characteristics curve of I againt f

Example 11.5

What is the potential difference across an inductor of inductance $5H$ when an alternating current of 20 mA and frequency $50Hz$ flows through it?

Solution

Reactance of inductor, $X_L = 2\pi fL = 2\pi \times 50Hz \times 5H = 1570\,\Omega$

Potential difference across inductor $V = IX_L = 20 \times 10^{-3} A \times 1570\,\Omega = 31.4V$

11.7 Resistors and Capacitors in Series

Figure 11.11 shows a resistor R connected in series with a capacitor C to a sinusoidal a.c. supply. Since the resistor R and capacitor C are in series, the same current I flows in R and C. Hence in drawing the phase diagram (Figure 11.11), the phase for the current I is used as a reference. For the p.d. V_R across the resistor R, $V_R = IR$ and this is in phase with the current. Therefore, the phasor for V_R is drawn parallel to the phasor for I.

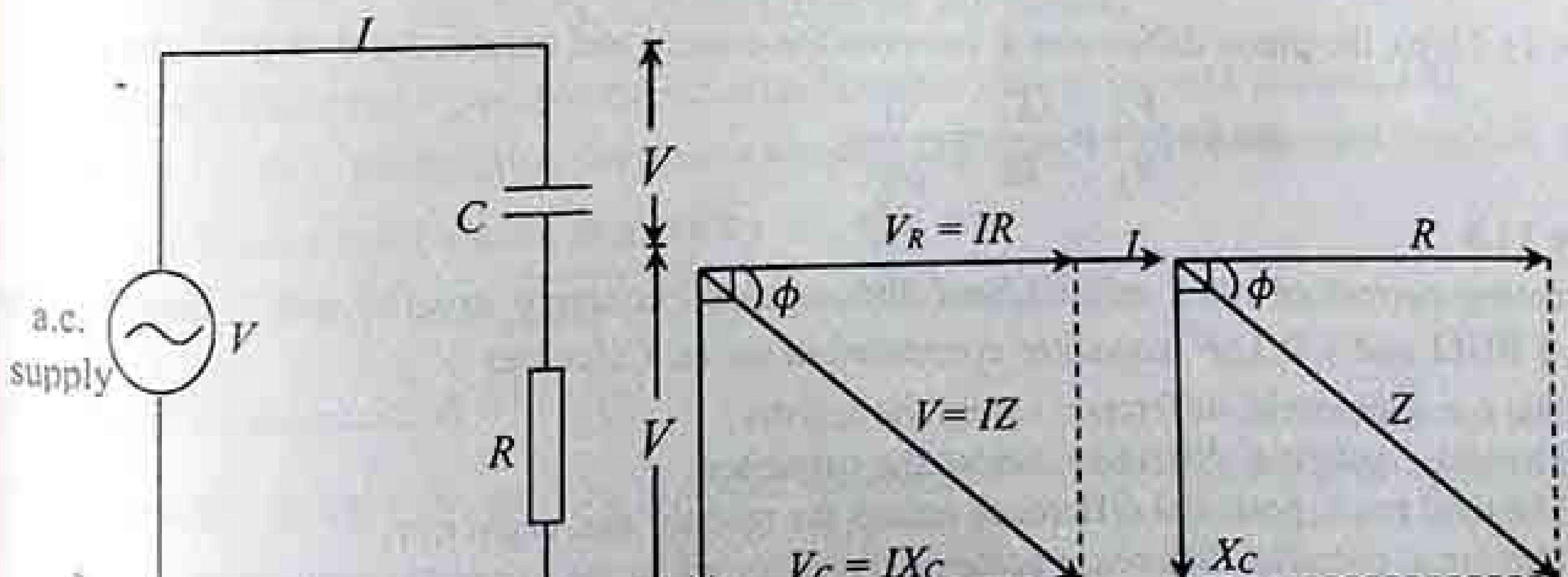


Fig. 11.11: A resistor and capacitor in series

For the capacitor C, the p.d. across it, $V_C = IX_C = 1/2\pi fC$ lags behind the current I by $\pi/2$ radians. Hence V_C lags behind V_R by $\pi/2$ radians. The resultant p.d. V across the resistor R and capacitor C is found by vector addition as shown in Figure 11.11. Hence

$$V^2 = V_R^2 + V_C^2 = I^2(R^2 + X_C^2)$$

$$V = I\sqrt{R^2 + \frac{1}{(2\pi fC)^2}}$$

$$\frac{V}{I} = \sqrt{R^2 + \frac{1}{(2\pi fC)^2}} \tag{11.16}$$

The ratio $\frac{V}{I}$ is known as the impedance Z of the circuit.

Hence impedance, $$Z = \frac{V}{I} = \sqrt{R^2 + \frac{1}{(2\pi fC)^2}} \tag{11.17}$$

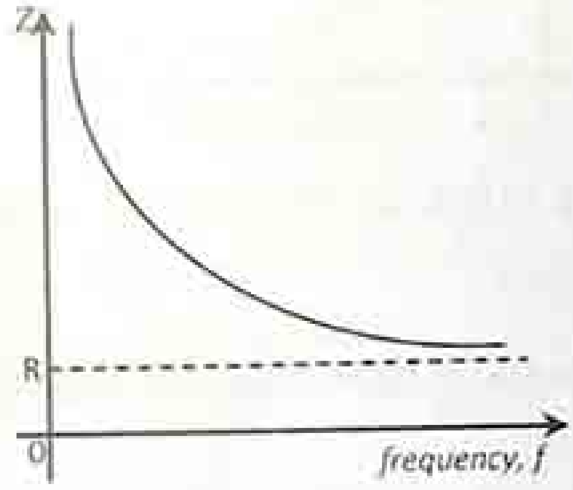


Fig. 11.12: Variations of impedance Z with frequency

The unit for impedance is the ohm (Ω). Figure 11.12 shows the variations of the impedance Z with frequency. For a very high frequency ($f \to \infty$), the impedance Z equals the resistance R of the resistor.
Figure 11.11(c) shows the relationship between the resistance R, reactance of the capacitor X_C and the impedance Z in the circuit.

Also, $$Z^2 = R^2 + X_C^2$$

$$Z = \sqrt{R^2 + X_C^2} = \sqrt{R^2 + \frac{1}{(2\pi fC)^2}}$$

In Figure 11.11(b), the phase difference ϕ between the current and resultant p.d. is given by

$$\tan\phi = \frac{V_C}{V_R} = \frac{IX_C}{IR} = \frac{1}{2\pi fC}$$

Example 11.6

An alternating current of r.m.s value $1.5mA$ and angular frequency $\omega = 100\,rads^{-1}$ flows through a resistor of $10k\Omega$ and a $0.5\mu F$ capacitor connected in series. Calculate

(a) The r.m.s. potential difference across the resistor,
(b) The r.m.s. potential difference across the capacitor,
(c) The total r.m.s. potential difference across the resistor and capacitor,
(d) The r.m.s. potential difference across the resistor when the current is maximum,
(e) The impedance in the circuit.

Solution

(a) Potential difference across the resistor. $V_R = IR = (1.5 \times 10^{-3})(10 \times 10^3) = 15V$
Potential difference across the capacitor

$$V_C = IX_C = I \times \frac{1}{2\pi fC} = \frac{1}{\omega C} = \frac{(1.5\times10^{-3})}{100\times0.5\times10^{-6}} = 30V$$

(b) The total p.d, V across the resistor and capacitor is given by

$V^2 = V_R^2 + V_C^2; V = \sqrt{15^2 + 30^2} = 33.5V$

(c) The p.d. across the resistor is in phase with the current whereas the p.d, V_C across the capacitor lags behind the current by $\pi/2$ radians. Therefore, when the current is maximum, the p.d. across the capacitor is zero and the p.d. across the resistor is maximum. Hence p.d. across the resistor when the current is maximum is

$V = I_0R \quad (I_0 = \text{peak current} = \sqrt{2}I_{rms}) = (1.5\times10^{-3})\sqrt{2}\times(10\times10^3) = 21.2V$

(d) Impedance in the circuit,

$$Z = \frac{V_{r.m.s}}{I_{r.m.s}} = \frac{33.5}{(1.5\times10^{-3})} = 2.23\times10^4\Omega$$

11.8 Resistors and Inductors in Series

Figure 11.13 shows a resistor of resistance R and an inductor of inductance L connected in series to an alternating voltage V. The potential difference V_R across the resistor is $V_R = IR$ and is in phase with the current I.

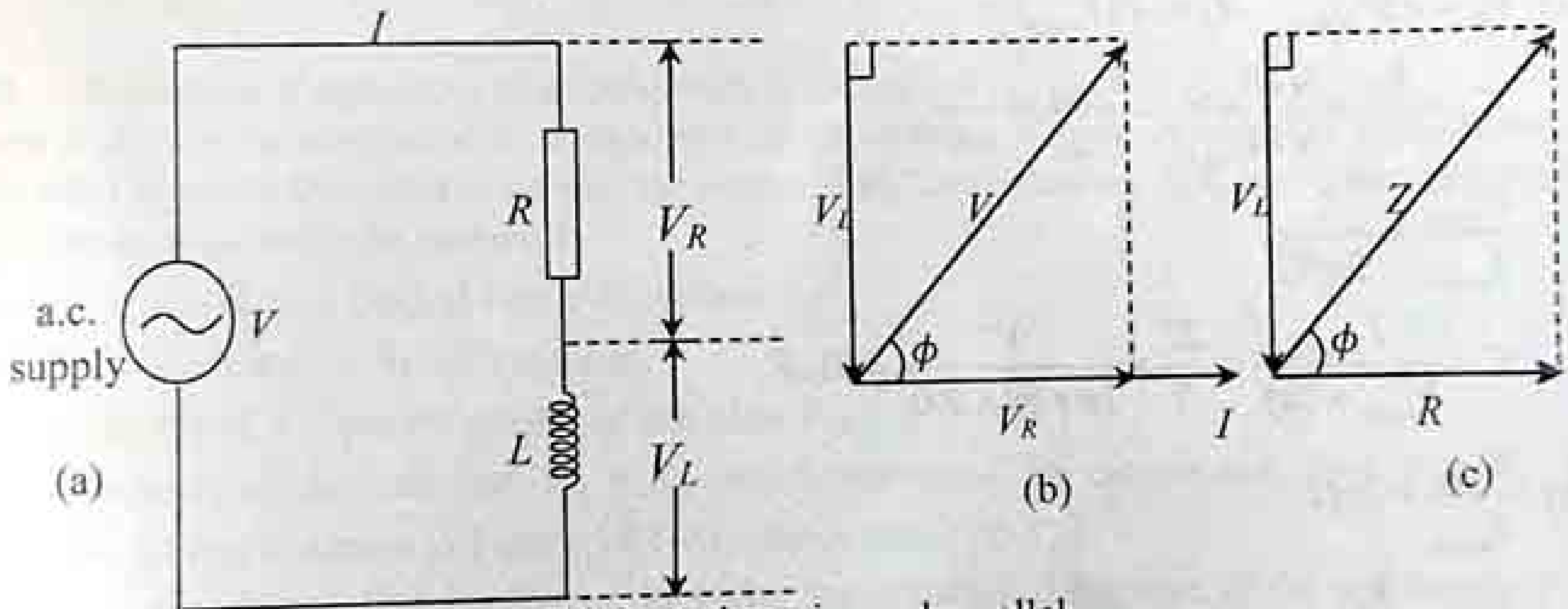


Fig. 11.13: Resistors in series and parallel

The potential difference V_L across the inductor leads the current by $\pi/2$ radians, and $V_L = IX_L = I(2\pi fL)$. Hence from the phasor diagram in Figure 11.13(b), the total potential difference V across the resistor and inductor is given by $V^2 = V_R^2 + V_L^2$

$$V = \sqrt{(IR)^2 + (IX_L)^2} = I\sqrt{R^2 + (2\pi fL)^2}$$

Impedance of the circuit, $Z = \frac{V}{I} = \sqrt{R^2 + (2\pi fL)^2}$

The phase angle ϕ between the total potential difference V and the current I is given by

$$\tan\phi = \frac{V_L}{V_R} = \frac{I(2\pi fL)}{IR} \frac{2\pi fL}{R}$$

The potential difference V leads the current I.

Example 11.7

A solenoid has inductance $2.5H$ and resistance 6.0Ω. At what frequency is the resistance of the solenoid equal to 1% of its reactance?

Solution

When $R = 1\%$ of reactance, $6 = \frac{1}{100} \times 2\pi f \times 2.5$

$$f = 38.2Hz$$

Example 11.8

A $4\mu F$ capacitor is used in a radio circuit through which $5mA$ of alternating electric current of frequency $60Hz$ flows. Determine: (a) the reactance of the capacitor, (b) the p.d. across the capacitor.

Solution

(a) The reactance $X_C = \frac{1}{\omega C} = \frac{1}{2\pi fC} = \frac{1}{2\pi \times 60 \times 4 \times 10^{-6}} = 663.15A$

(b) The p.d. across the capacitor, V_C is $V_C = I_C \times X_C = 5 \times 10^{-3} \times 663.15 = 3.3V$

Example 11.9

The minimum alternating current required to light a given lamp is $22/7\,A$. If this lamp is connected in series to a capacitor of capacitance C, and a $200V$ alternating power supply of frequency $50Hz$.

(a) What must be the value of C that will give the required current?

(b) What value of the alternating current will flow through the circuit if the capacitor is replaced with an inductor $7/22\,H$?

Solution

(a) $V_0 = \sqrt{2}V_{r.m.s}, \quad I_0 = \sqrt{2}I_{r.m.s}$

$$X_0 = \frac{V_0}{I_0} = \frac{\sqrt{2}V_{r.m.s}}{\sqrt{2}I_{r.m.s}} = \frac{V_{r.m.s}}{I_{r.m.s}} = \frac{1}{\omega C}$$

$$\frac{V_{r.m.s}}{I_{r.m.s}} = \frac{1}{2\pi fC}$$

$$C = \frac{I_{r.m.s}}{V_{r.m.s} \times 2\pi f} = \frac{22}{7} \times \frac{1}{2\pi \times 50 \times 200} = 50\mu F$$

(b) $\frac{V_{r.m.s}}{I_{r.m.s}} = 2\pi fL$

$$I_{r.m.s} = \frac{V_{r.m.s}}{2\pi fL} = \frac{200}{2 \times 22/7 \times 50 \times 7/22} = 2.0A$$

Example 11.10

At what frequency in Hz will a $10\mu F$ capacitor have a reactance of $1{,}000\Omega$?

Solution

$$X_C = \frac{1}{2\pi fC} = \frac{1}{2\pi \times f \times 10 \times 10^{-6}}; \qquad \frac{1000}{1} = \frac{1000}{1} = \frac{1}{2\pi \times f \times 10 \times 10^{-6}}$$

$$f = \frac{1}{1000 \times 2\pi \times 10^{-5}} = \frac{1}{2\pi \times 10^{-2}} = 50/\mu\,Hz$$

Example 11.11

What is the inductance of an inductor whose reactance is 1Ω at $50/\pi\,Hz$?

Solution

$$X_L = \omega L = 2\pi fL; \qquad 1 = 2\pi \times 50/\pi \times L$$

$$1/100 = L; \; L = 10^{-2}H$$

Example 11.12

An inductor with $L = 4.0\times10^{-2}\,H$ is connected to an oscillating source of emf. This source provides an emf $\varepsilon = \varepsilon_0 \sin \omega t$ with $\varepsilon_0 = 0.20V$ and $\omega = 6.0\times10^3\, rad/s$.

(a) What is the reactance of the inductor?

(b) What is the maximum current in the circuit?

(c) What is the current at time $t = 0$ and time $t = \pi/4\omega$?

Solution

(a) $X_L = \omega L = 6\times10^3 \times 4\times10^{-2} = 240\Omega$

(b) $X_L = \dfrac{V_0}{I_0}$; $\qquad I_0 = \dfrac{V_0}{X_L} = \dfrac{0.2}{240} = 8.3\times10^{-4}\,A$

(c) $I = I_0 \sin\left(\omega t - \dfrac{\pi}{2}\right) = \dfrac{V_0}{\omega L}\sin\left(\omega t - \dfrac{\pi}{2}\right)$ since $\dfrac{V_0}{I_0} = \omega L$

$I = -\dfrac{V_0}{\omega L}$ at $t = 0$, $\qquad I = \dfrac{0.2}{240}A = 8.3\times10^{-4}\,A$

and at $t = \dfrac{\pi}{4\omega}$; $I = \dfrac{V_0}{\omega L}\sin\left(\dfrac{\pi}{4} - \dfrac{\pi}{2}\right)$

$$I = \frac{V_o}{\omega L}\sin\left(-\frac{\pi}{4}\right) = -\frac{V_o}{\omega L}\times\frac{\sqrt{2}}{2} = \left(-\frac{0.2}{240}\times\frac{\sqrt{2}}{2}\right) = 5.9\times10^{-4}\,A$$

11.9 Resistors, Capacitors and Inductors in Series

When a resistor of resistance R, a capacitor of capacitance C and an inductor of inductance L are connected in series with an a.c. supply, the potential difference across R, C and L are, respectively

$V_R = IR$, in phase with the current I;

$V_C = IX_C$, which lags behind I by $\pi/2$ radians;

$V_L = IX_L$, which leads I by $\pi/2$ radians.

The magnitude of V_L can be greater or less than V_C.

If the reactance of the inductor $X_L > X_C$, the reactance of the capacitance, then $V_L > V_C$ and vice versa. The phasor diagram in Figure 11.14(a) shows that $V_L > V_C$.

Also, from the phasor diagram Figure 11.14(b), the potential difference of the a.c. supply V is given by $V^2 = V_R^2 + (V_L - V_C)^2 = I^2\left[R^2 + \left(2\pi f L - \dfrac{1}{2\pi f C}\right)^2\right]$.

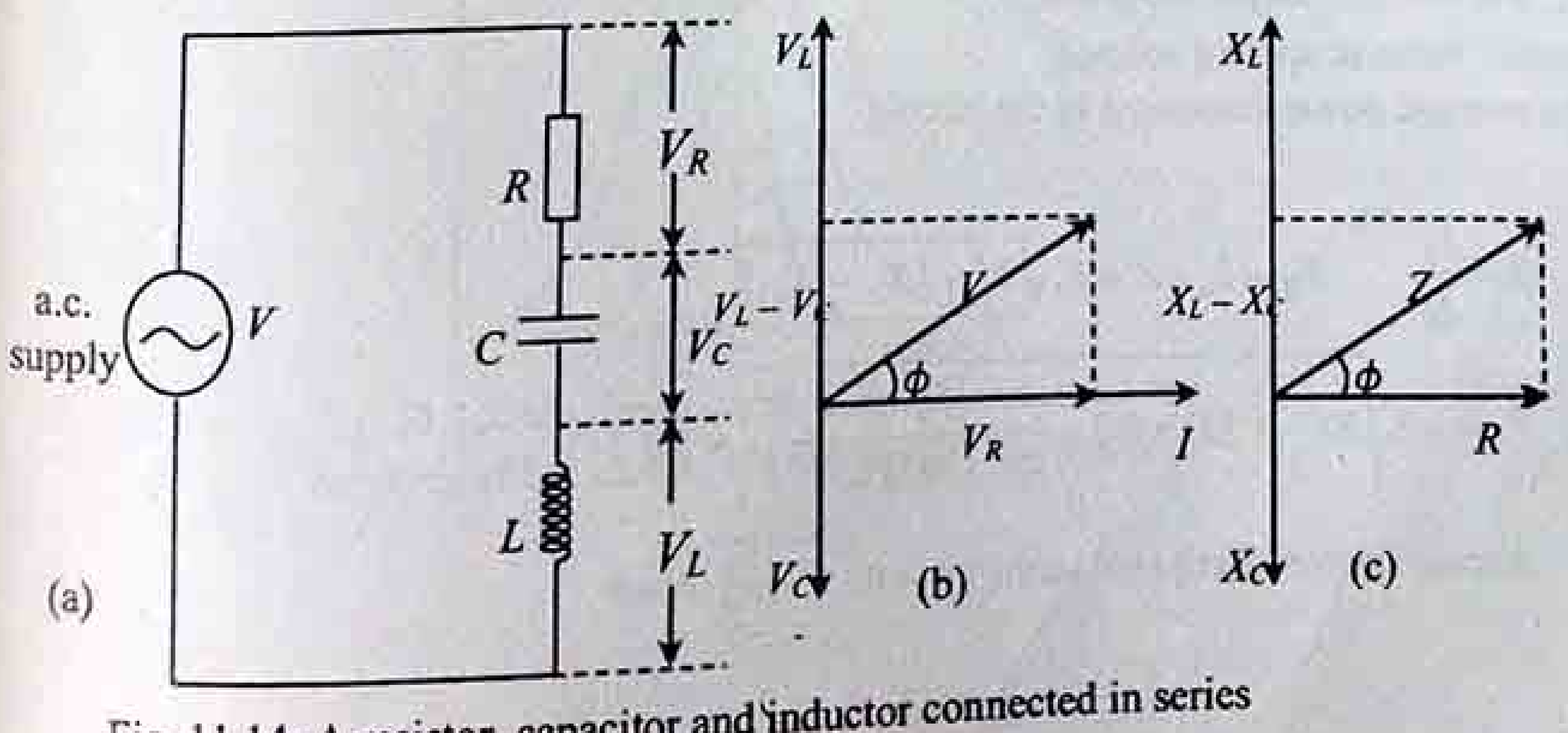


Fig. 11.14: A resistor, capacitor and inductor connected in series

Hence the impedance in the circuit

$$Z = \frac{V}{I} = \sqrt{R^2 + \left(2\pi fL - \frac{1}{2\pi fC}\right)^2} \qquad (11.18)$$

The phase angle ϕ between the p.d. V of the supply and the current I is given by

$$\tan\phi = \frac{(V_L - V_C)}{V_R} = \frac{2\pi fL - \frac{1}{2\pi fC}}{R} \qquad (11.19)$$

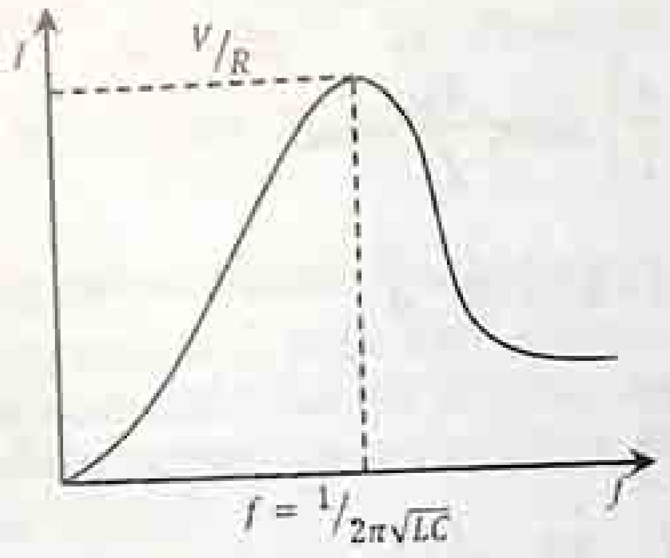


Fig. 11.15

From equation 11.18,

The minimum impedance, $Z = R$ occurs when $X_L = X_C$

$$2\pi fL = \frac{1}{2\pi fC}$$

i.e. when $f = f_0 = \dfrac{1}{2\pi\sqrt{LC}}$, the impedance is minimum.

Figure 11.16 shows the variation of the current I with the frequency. When the frequency is $f_0 = \dfrac{1}{2\pi\sqrt{LC}}$, the impedance is least and the current is maximum. The frequency f_0 is known as the *resonant frequency* of the series RLC circuit.

Example 11.13

An alternating current with a peak value of $6.8A$ flows through a series RLC circuit. If $R = 50\Omega$, $L = 40\mu F$, and $\omega = 337\ rad/s$, calculate

(i) The peak value of applied voltage,

(ii) The average power dissipated in the circuit.

Solution

(i) $\quad Z = \dfrac{V_0}{I_0} \qquad V_0 = I_0 \times Z = I_0\sqrt{R^2 + (X_L - X_C)^2}$

$$V_0 = 6.8\sqrt{50^2 + \left(337 \times 0.2 - \frac{1}{337 \times 40 \times 10^{-6}}\right)^2} = 345.56V$$

(ii) $\quad$ Average power dissipated in the circuit, $P = \dfrac{V_0^2}{Z}\cos\phi$

where $\tan\phi = \dfrac{X_L - X_C}{R} = \dfrac{V_o^2 R}{R^2 + (\omega L - 1/\omega C)^2}$

$$= \frac{345.56^2 \times 50}{50^2 + \left[(337 \times 40 \times 10^{-6}) - (1/337 \times 0.2)\right]^2} = 2388.24W$$

11.10 Resonance Circuit

In an RLC circuit, a resonance circuit is one through which maximum alternating current flows. Recall that for an RLC circuit, the current flowing is in general given by:

$$I = \frac{V}{Z} = \frac{V}{\sqrt{R^2 + (\omega L - 1/\omega C)^2}}$$

The current I is maximum when $\omega L = 1/\omega C = 0$

That is $\omega = \omega_0 = 1/\sqrt{LC}$

or $f_0 = \dfrac{1}{2\pi\sqrt{LC}}$ since $\omega_0 = 2\pi f_0$

where f_0 is called *resonance frequency*. The maximum current at resonance is therefore: $I_{max} = \dfrac{V_{max}}{R}$.

Bandwidth: The width of the peak measured between the points below and above resonance at which the curve has one-half of its maximum height, is called the bandwidth of the RLC circuit. This terminology is adopted from radio engineering, where frequency intervals are commonly called bands.

Applications of Resonance

The phenomenon of resonance is basically the same as in sound, mechanics and other branches of physics. The series resonance circuit, just discussed is used for tuning a radio receiver. Thus, we are tuned to a station when an incoming radio wave is in resonance with the electrical RLC circuit.

The incoming radio wave of frequency, f, induces a voltage, V, of the same frequency, in a coil and capacitor, C, is varied, the resonance frequency is changed and at one setting of C, the resonance frequency becomes f, the frequency of the incoming wave. Maximum current is then obtained, and the station is now heard very loudly. In other words, we are now tuned to the station.

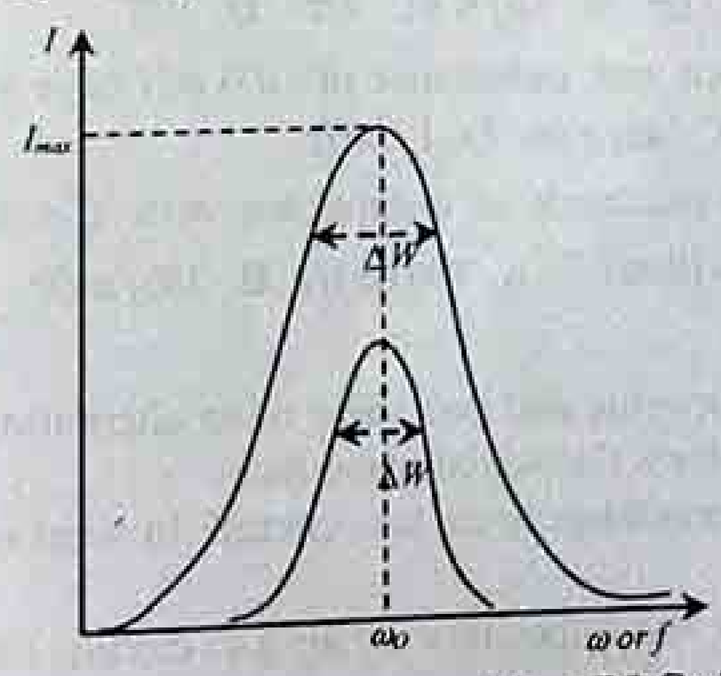


Fig. 11.16: Maximum current in a RLC circuit as a function of Bandwith (W) for two curves shown

Example 11.14

A series RLC circuit with $R = 10\Omega$, $L = 0.2mH$ and $C = 0.008\mu F$ is connected to an alternating voltage. If the amplitude of the alternating voltage source is $3.0V$, calculate: the resonance frequency f_0 of the circuit and the amplitude of the a.c. at resonance.

Solution

$$f_0 = \frac{1}{2\pi(LC)^{1/2}} = \frac{1}{2\pi(2\times10^{-4}\times8\times10^{-8})^{1/2}} = 39.78Hz$$

$$\text{Max current } I_0 = \frac{V_0}{R} = \frac{3}{10} = 0.3A$$

Exercise 11

11.1 Calculate the peak voltage of a mains supply of $240V_{rms}$ in volts.
A. 121 B. $240(2)^{1/2}$ C. 40 D. $240(2)^{-1/2}$

11.2 An alternating current flows through a circuit containing a resistor.If the current is $4.0A$ r.m.s, and frequency is $50Hz$, calculate the peak current in A.
A. 5.66 B. 2.33 C. 1.2 D. 0. 33

11.3 An alternating frequency $50Hz$ is connected to a capacitor of $100\mu F$ across a p.d of $12V$ Calculate the capacitance of the capacitor in ohms **A.** 200 **B.** 100 **C.** 31.8 **D.** 10.8

11. 4 What is the peak value of current in A whose r.m.s value is $10A$?
A. 7.1 **B.** $10\times(2)^{1/2}$ **C.** $20\times(2)^{1/2}$ **D.** $10\times(2)^{-1/2}$

11.5 A $10\mu F$ capacitor is connected in series to 100Ω resistor. Calculate the total impeadance at a frequency of $159Hz$. **A.** 10Ω **B.** 60Ω **C.** 141Ω **D.** 200Ω

11.6 A sinusoidal supply of frequency $100Hz$ and r.m.s voltage $12V$ is connected to a $2.2\mu F$ capacitor. What is the r.m.s current? **A.** $100mA$ **B.** $50mA$ **C.** $30mA$ **D.** $17mA$

11.7 An inductor dispates heat at the rate of $10W$ when an a.c of $0.5A$ flows in it.If the reactance of the inductor at a frequency supply is 30Ω. What is the impeadance?
A. 50Ω B. 30Ω **C.** 20Ω **D.** 10Ω

11.8 When an A.C flows through a resistor of 5Ω heat is dissipated at the rate of $20W$. What is the rms value of the a.c? **A.** $50A$ **B.** $20A$ **C.** $5A$ **D.** $2A$

11.9 In a series L.C. circuit the inductance and capacitance are $9H$ and $100\mu F$ respectively the resonant frequency in Hz is **A.** $5/3\pi$ **B.** $3\pi/5$ **C.** 2π **D.** 10π

11.10 At what frequency will an inductor with inductance of 10.0 mH have a resistance of 8000Ω?
A. $40000/\pi\ Hz$ B. $4000/\pi\ Hz$ C. $40/\pi\ Hz$ D. $1.6\pi Hz$

11.11 At what frequency in Hz is the reactance of a capacitor with $C = 100\mu F$ be equal to the reactance of an inductor with $L = 10mH$? **A.** $100/\pi\ Hz$ **B.** $250/\pi\ Hz$ **C.** $370/\pi\ Hz$ **D.** $500/\pi\ Hz$

11.12 Explain what is meant by the peak value and *rms* value of an alternating current. Establish the relation between these quantities for a sinusoidal wave form.

11.13 What do you understand by the reactance in an a.c. circuit? In what way does it differ from resistance?

11.14 Explain what you understand by the impedance of an a.c. circuit. Obtain the condition for resonance in a series RLC circuit.

11.15 The voltage across the terminals of an a.c power supply varies with time according to $v = V_m \cos\omega t$. The peak voltage $V_0 = 50V$. What is the root mean square potential difference V_{rms}?

11.16 A $5.0kW$ electric cloth dryer runs on $240V_{rms}$, I_{rms} and I_{max}. Find the same quantities for a dryer of the same power that operates at $120V_{rms}$.

11.17 What is the *rms* value of the current through an incandescent bulb rated at $1000W$ and $240V$, but connected to an alternating $120V_{rms}$ supply.

11.18 A $50\mu F$ capacitor is connected to a $60Hz$ source that provides $24V_{rms}$. Find (i) the peak charge on the capacitor and (ii) peak current in the wires.

11.19 A $100V_{rms}$ voltage is applied to a series RC circuit. The voltage across the capacitor is $80V$. What is the voltage across the resistor?

11.20 An inductor is connected to a $20Hz$ power supply that produces a $50V_{rms}$ voltage. What inductance is needed to keep the instantaneous current in the circuit below $80mA$?

CHAPTER 12
MAXWELL'S EQUATIONS AND ELECTROMAGNETIC WAVES

12.0 Introduction

It has been shown that a current carrying conductor sets up a magnetic field around it and if such a conductor is placed in a magnetic field, it experiences a force. It is also known that two current carrying loops repel or attract each other. These experimental results show that a current carrying conductor behaves just like a magnet. They also show that electricity and magnetism are inter-related. Hence the name "electromagnetism" is used to show that electrostatic and magnetic phenomena are indistinguishable.

12.1 Maxwell's Equations

It was James Clerk Maxwell who found in 1860 that the experimental laws of electricity and magnetism-the Coulomb's, Gauss's, Biot-Savart's, Ampere's and Faraday's which we have already studied could be summarized in a concise mathematical form known as Maxwell's equations, and showed that these equations predict the possibility of electromagnetic waves. Maxwell's equations relate the electric and magnetic field vectors E and B to their sources, which are electric charges, currents and changing fields. These equations, which we will list below, play a role in classical electromagnetism analogous to that of Newton's laws in classical mechanics. In principle, all problems in classical electricity and magnetism can be solved using Maxwell's equations just as all problems in Classical Mechanics can be solved using Newton's Laws. However, Maxwell's equations are considerably more complicated than Newton's laws and their applications to most problems involve mathematics beyond the scope of this book. We will therefore not attempt either to derive the wave equations or go into complicated applications. Our main emphasis in this chapter is to list these equations, indicate their physical basis, and possibly discuss the nature of various electromagnetic waves.

Maxwell's equations are:

1. Gauss' law for electricity:

$$\oint_s E_n dA = \frac{1}{\varepsilon_o} Q \tag{12.1}$$

2. Gauss' law for magnetism:

$$\oint_s B_n dA = 0 \tag{12.2}$$

3. Faraday's law for Induction:

$$\oint_c E.dl = \frac{-d}{dt}\int_s B_n dA \tag{12.3}$$

4. Ampere's-Maxwell's law:

$$\oint_c B.dl = \mu_o I + \mu_o \varepsilon_o \frac{d}{dt}\int_s E_n dA = \mu_o\left[I + \varepsilon_o \frac{d}{dt}\phi\right] \tag{12.4}$$

In the above equations:

$\oint_s$ = integrating over the entire surface

$\oint_c$ = integrating around the closed curve, c

E_n = normal component of electric field intensity (volt/metre)

B_n = normal component of magnetic flux density (Webers/m^2) or Tesla

dA = element of area

Q = total charge enclosed

ε_0 = permittivity of free space $= 8.854\times10^{-12}$ F/m

μ_0 = permeability of free space $= 4\pi\times10^{-7}$ H/m

Maxwell showed that these equations could be combined to yield a wave equation for electric and magnetic field vectors E and B. Such electromagnetic waves are caused by accelerating charges, for example, the charge in an alternating current in antenna. Maxwell further deduced that the speed of electromagnetic waves in free space is

$$C = \frac{1}{\sqrt{\mu_o \varepsilon_o}} \cong 3\times10^8\, ms^{-1}$$

12.2 The Physical Basis for Maxwell's Equations

We now review the physical basis for Maxwell's equations listed above. The first equation 12.1 is the statement of Gauss' law which relates the flux of the electric field intensity E through a closed surface to the net charge Q, enclosed within that surface. It therefore states that *the product of the normal component of an electric field and the area of the surface enclosed is proportional to the total charge enclosed*. Its experimental basis is the Coulomb's law.

The second equation 12.2 is sometimes called Gauss' law of magnetism. It states that *the flux of the magnetic field vector B is zero through any closed surface, i.e. the number of magnetic field lines that enters any closed surface is exactly equal to the number that leaves the surface*. The equation describes the experimental observations that magnetic field lines do not diverge from any point or converge on any point. That is, it implies that isolated magnetic poles do not exist. Whereas we have isolated charges such as protons and electrons, magnetic monopoles (the magnetic equivalent of electric charges) have not been found in nature. The existence of magnetic monopole however, is not forbidden by any physical law. As we say in physics, "unless it is forbidden, it exists". The search for magnetic monopoles is continuing.

Equation 12.3 is a statement of Faraday's law of electromagnetic induction which states that *the sum of induced e.m.f. equal to the time rate of change of magnetic flux*. The mathematical law states that *the integral of the electric field around any closed curve C, which is the e.m.f. equals the negative rate of change of magnetic flux through any surface S bounded by the curve*-this is not a closed surface, so that magnetic flux through S is not necessarily zero. Faraday's law describes how electric field lines encircle an area through which magnetic flux is changing and it relates the electric field vector E to the rate of change of magnetic field vector B.

Equation 12.4 is the Ampere's law which states that the total magnetic field is a measure of the total current flowing in the circuit. It was Maxwell who came up with a modified version of the law which is now known as Maxwell-Ampere's law. It states that the line integral of the magnetic field B around any closed curve C equals μ_0 times the current through any surface bounded by the curve plus $\mu_0\varepsilon_0$ times the rate of change of electric flux through the surface. This law describes how the magnetic field lines encircle an area through which a current is passing or the electric flux is changing.

12.3 The Electromagnetic Spectrum

It can be shown using Maxwell's equations that accelerating charges produce electromagnetic waves which transport energy and momentum from source (e.g. oscillating dipole) to a receiver. The electromagnetic waves travel in a vacuum at the speed of light. There are many forms of electromagnetic waves which are distinguishable only by their wavelengths λ and their frequencies f. The frequency f and the wavelength λ of any electromagnetic wave are related by this expression,

$$c = f\lambda \qquad (12.5)$$

where c, is the speed of light in vacuum. All frequencies and wavelengths that satisfy equation 12.5 are allowed; there are no restrictions, upper or lower limits. The various types of electromagnetic wave are listed in Figure 12.1; each is labelled by its wavelength and frequency. But we must point out that there is a considerable overlap between the various regions.

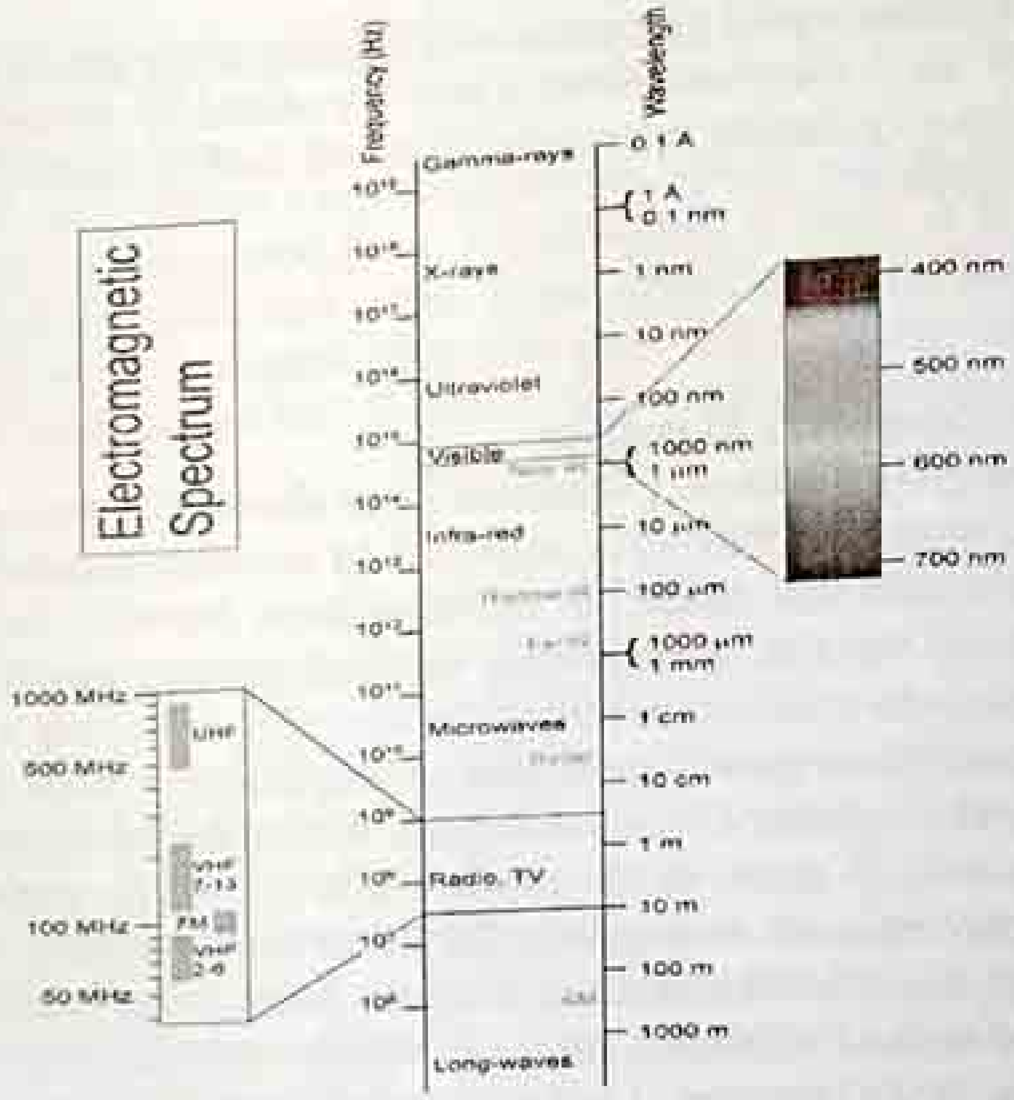


Fig. 12.1: The Electromagnetic Spectrum

Visible Light

The visible light of the spectrum of Figure 12.1 is further expanded and shown in Figure 12.2. Its wavelengths range from $400nm$ to $700nm$. An approximate range of wavelengths is associated with each colour: violet ranges from $400nm$ to $450nm$, blue from $450nm$ to $520nm$; green from $529nm$ to $560nm$; yellow from $560nm$ to $600nm$; orange from $600nm$ to $625nm$ and red from $625nm$ to $700nm$. Light is produced when electrons in atoms and molecules undergo transitions between energy levels. The sensitivity of the human eye is a function of the wavelength of light; the sensitivity is maximum at about $560nm$ which is the wavelength of green-yellow colour.

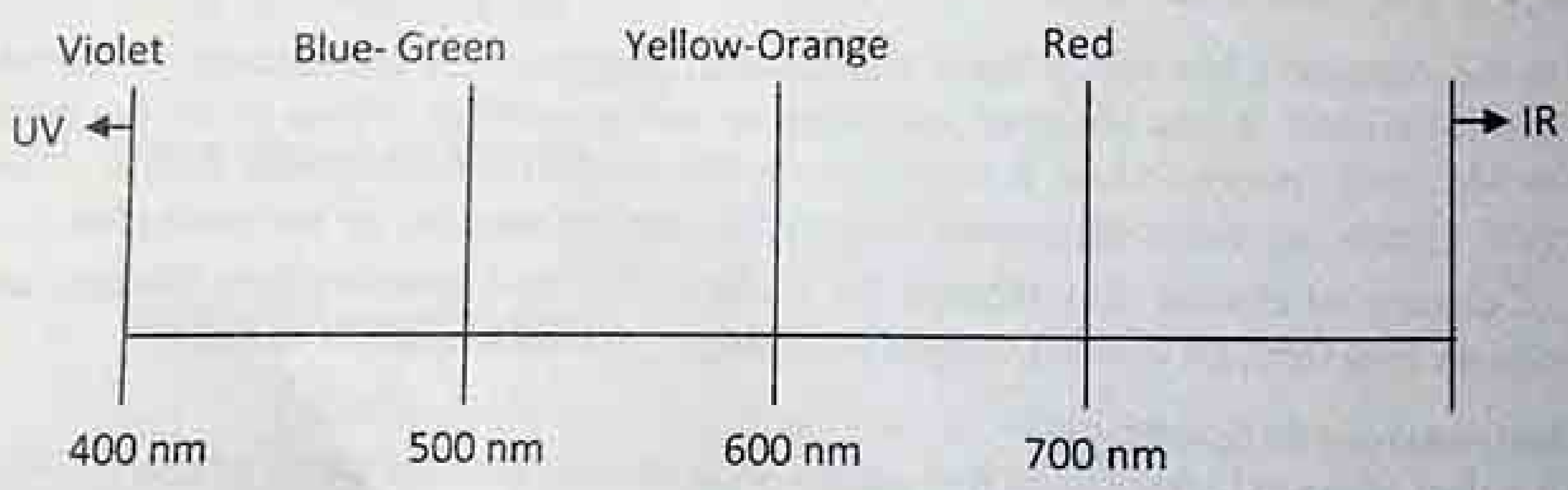


Fig. 12.2: The wavelength of the visible spectrum

Infrared Radiation (IR)

Infrared radiation starts from the longest wavelength of the visible light (red $700nm$) to $1mm$. It is at times referred to as heat waves. Infrared radiation is produced by hot bodies and molecules. It has many practical uses; infrared sensitive films can detect IR radiation emitted by warm bodies. This is the basis on which snakes and night vision instruments can detect objects at night.

Microwaves

Microwaves have wavelengths that range from $1m$ to $15cm$. They are generated by the oscillation of electrons in electronic devices. Klystron is such a device and it can produce microwaves of up to $30GHz$ which has a wavelength of $1cm$. Microwaves are used in radar detection system because they have short wavelengths. The radiation of microwave oven found in some kitchens has a frequency of about $2450Hz$.

Radio and TV Signals

The radio and TV waves range from 15*cm* to 2000*m*. They are generated by accelerating charges through conducting wires. They are used in radio and TV communication systems.

UV Radiation

The ultraviolet (UV) region extends from 380*nm* to 60*nm*. The sun is the most important source of UV radiation; the UV radiation helps in producing vitamin D and tanning in humans. However, a large dose of UV radiation can produce skin cancer in humans. Fortunately, the ozone (O_3) in the earth's atmosphere absorbs most of the UV radiation from the sun. This ozone protection against this deadly radiation is being depleted by chlorofluorocarbon (CFC's). If it is completely depleted, we expect to see an increase in skin diseases.

X-Rays

The X-ray region is adjacent to the UV region as we can see from Figure 12.1; this region extends from 1 nm to 0.01*nm*. X-rays are produced when fast moving electrons are allowed to strike a heavy metal target. X-rays are used as diagnostic tool in medicine and scientific studies. X-rays can damage living tissues and organisms, therefore, every effort should be made to avoid over exposure to it.

Gamma Rays

Gamma rays range from 0.01*nm* to less than $10^{-5}nm$. They are produced by radioactive nuclei and during certain nuclear reactions. Gamma rays, like X-rays, can damage living tissues, but unlike X-rays they are highly penetrating.

Summary

1. Equations 12.1 to 12.4 give a summary of Maxwell's equations. They relate electric and magnetic field vectors E and B to their sources which are electric charges, currents and changing fields.
2. These equations can be shown to lead to a wave equation known as electromagnetic waves that originate from accelerating charges such as oscillating dipole.
3. All electromagnetic waves travel with speed of light c. They only differ in wavelength λ, and frequency, f, which are related by $f = c/\lambda$.
4. Electromagnetic waves include visible light, infrared radiation (IR), microwaves, radio and TV signals, UV radiation, X-rays, and gamma rays.
5. Visible light has wavelength ranging from 400 – 700*nm*.
6. Infrared ranges from 700*nm* – 1*mm*.
7. Microwave ranges from 1*m*-15*cm*.
8. TV/Radio waves ranges from 15*cm* – 2000*m*.
9. UV radiation has wave lengths 380*nm* – 60*nm*.
10. UV from the sun is dangerous but we are protected by the ozone layer which absorbs UV
11. X- rays range from 1nm – 0.01*nm*. They have very useful applications in medicine
12. Gamma rays range from 0.01*nm* – $10^{-5}nm$. It is very penetrating and can damage tissues.

Exercises 12

12.1 State Gauss' Law

A. Flux of magnetic field vector through closed surface = 0.
B. Flux of electric field vector is zero.
C. Number of magnetic lines of force is zero.
D. That magnetic flux is divergent.

12.2 Which electromagnetic waves have the greatest frequencies?

A. Light waves **B.** X-rays **C.** Gamma rays **D.** UV rays.

12.3 What kind of waves has wavelengths of order of few meters?

A. Gamma rays **B.** Radio waves **C.** Optical light **D.** UV rays.

12.4 Which of the following about X-rays is not correct?

A. X-rays are electromagnetic radiation. **B.** They have short wavelengths.
C. They are charged. **D.** They are not deflected by magnetic fields.

12.5 Whichof the following about X-rays is correct?

A. X-rays do not affect photographic films.

B. X-rays have high penetrating power.

C. X-ray photons do not cause excitation of the atoms in the fluorescent materials.

D. X-rays do not exhibit wave properties such as interference

12.6 Which of the following is not correct?

A. X-rays are used to detect problem in human bones through photographic film.

B. X-rays are used in destroying cancerous cells.

C. In engineering, X-rays are used to locate cracks in metals.

D. X-rays are very important in optical Astronomy

12.7 Which of the following about Magnetic monopoles is not correct?

A. It is an equivalence of electric charge

B. Does not exist

C. Its existence is not forbidding by any physical law

D. The existence is still under investigation

12.8 What is the experimental basis for Gauss law of magnetism?

12.9 What is the experimental basis for Gauss law of electricity?

12.10 What is the major interpretation of Faradays law of e.m induction?

12.11 Find the wavelength for:

(a) A typical AM radio wave with frequency $1000kHz$

(b) A typical FM radio with frequency $100MHz$.

(Hint: Use $\lambda = v/f$) where v is the velocity of light

12.12 (a) What is the frequency of $3cm$ microwave?

(b) What is the frequency of an X-ray with wavelength of 0.1 nm?

(Hint: For both questions use $f = v/\lambda$) where v is the velocity of light.

12.13 Compute the frequency of the following electromagnetic waves:

(i) Microwaves ($\lambda = 1cm$), (ii) infrared radiation ($\lambda = 1\mu m$), (iii) ultraviolet light ($\lambda = 100nm$), and (iv) X-rays ($\lambda = 1pm$). (Hint: $f = v/\lambda$)

12.14 Discuss the characteristics of radiation in different sections of the E.M spectrum.

MODERN PHYSICS

CHAPTER 13
ATOMIC PHYSICS

13.0 Introduction

Atom is not visible to the human eye; it is therefore a microscopic quantity. Microscopic processes are best studied using quantum mechanics and theory of relativity, which is above the subject of this chapter and will be introduced to you later in your carrier, though some basic quantum theory will be used in the calculations of the energies involved in the electronic transitions in the atom. We will use atomic models to learn about the internal structures and processes of an atom. The J. J. Thomson model of the atom assumes that the atom consists of a uniform spherical mass having positive and negative charges distribute in it. Ernest Rutherford postulated the nucleus of atom which was deduced from the experiment by Geiger and Marsden.

In this chapter, we will study the structure and properties of the atom using the atomic theories and experiment. The line spectrum of hydrogen atom will be discussed. We will also discuss the production of X-rays by bombarding metals of high atomic number with high speed electrons.

13.1 Geiger and Marsden α-scattering Experiment

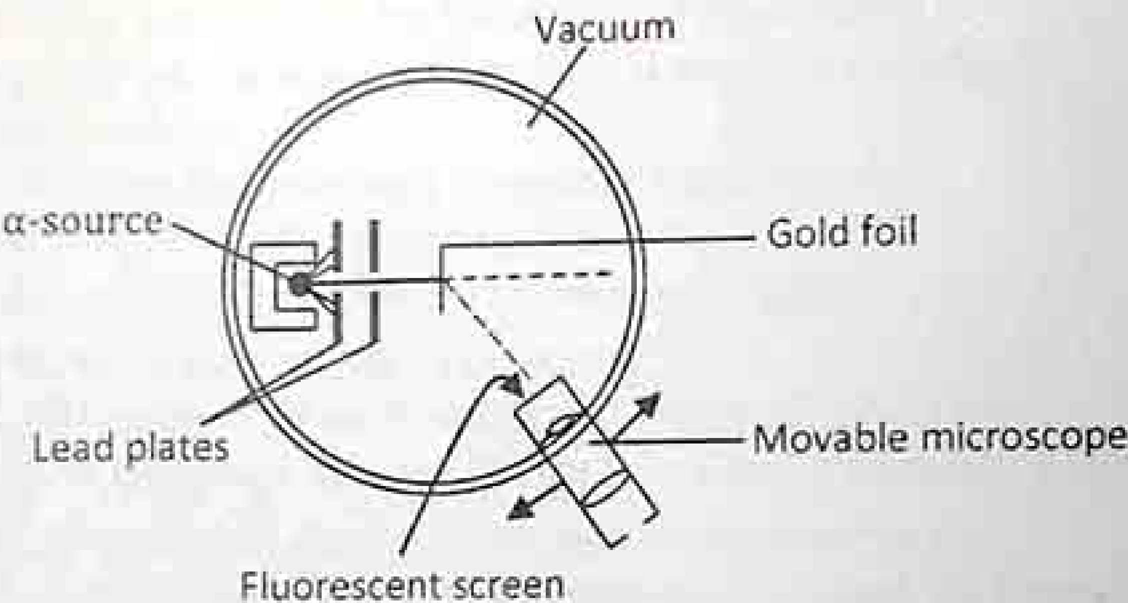


Fig. 13.1: α-scattering experiment

Geiger and Marsden used the arrangement shown in Figure 13.1, in which α-particles from a radioactive source were used to bombard a thin gold foil placed in a vacuum. After bombarding the gold foil, the α-particles were detected using a movable microscope fitted with a fluorescent screen. specks of light could be seen through the microscope.

The experimental results turned out to be quite surprising. Although the gold foil was a few hundreds of atomic layers thick and free of holes, Geiger and Marsden found that most of the α-particles were not deflected through small angles. Thus, the α-particles behaved as though the gold foil was not there.

Another unexpected result was less than 1% of the α-particles were deflected through very large angles of more than 90°. These α-particles seemed to have struck something massive but very few of the α-particles were deflected through nearly 180° (i.e. back towards where it came from).

In order to explain these observations, Rutherford postulated the nuclear atom. He postulated a model of the atoms consisting of a central massive core which he called the nucleus. The nucleus carries the entire atom's positive charge and nearly all its mass.

Assuming the inverse square law for the electrostatic forces, Rutherford derived from his model of the atom, a mathematical formula predicting the proportion of α-particles that should be scattered through any given angle. Taking measurements using the apparatus shown in Figure 13.1, Geiger and Marsden were able to prove the correctness of Rutherford's model.

Qualitatively, the large number of α-particles passing through the gold foil undeflected suggests that there exists a large empty space in an atom. Those α-particles (positively charged) that were deflected through large angles had come very close to the positively charged nucleus as postulated by Rutherford (see Figure 13.2). The diameter of the nucleus is of the order of $10^{-14}m$ whereas the diameter of the atom as computed from the density is of the order of $10^{-10}m$. 1 To help you visualize the amount of empty space in an atom: if the nucleus is drawn as a sphere of radius $1mm$, then on the same scale, the radius of the atom is $10m$.

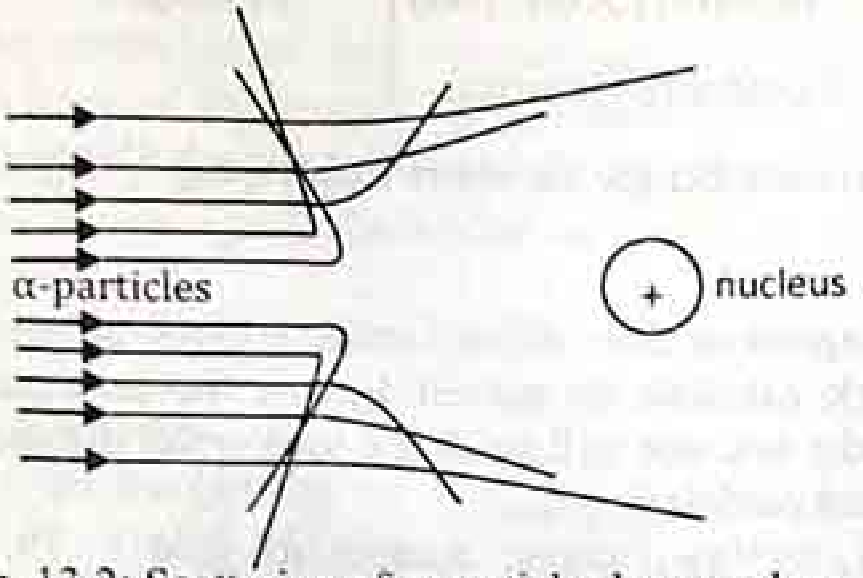


Fig. 13.2: Scattering of **α**-particles by a nucleus

The elements had been arranged in the periodic table based on their chemical properties. The number of electrons in an atom is equal to its atomic number which was introduced initially to specify the position of element in the periodic table. Thus, gold, being the seventy-ninth element in the periodic table, has an atomic number of 79 and, therefore has 79 electrons. Since an atom is neutral, the nucleus must carry a charge of $+79e$.

13.2 Protons and Neutrons

Protons were discovered by Rutherford when he bombarded nitrogen atoms with α-particles. The proton was found to be positively charged and carried a charge $+e$. Rutherford attributed the positive charge in the nucleus to the proton as the number of protons was equal to the number of electrons in a neutral atom. From Moseley's experiments on characteristic X-rays, it was established that the periodic table should be arranged according to the number of protons in the nucleus. This number is known as the atomic number Z of the atom

Another component of the nucleus is the neutron. Neutrons were discovered by Chadwick when he bombarded beryllium with α-particles. The particle emitted was found to be uncharged. Further experiments involving the bombardment of neutrons with nitrogen atoms and protons showed that the mass of the neutron was roughly equal to the mass of the proton. The total number of nucleons, i.e. protons and neutrons, in a nucleus is known as the mass number A. The symbol used to represent a nucleus X is $^A_Z X$.

Where the subscript Z represents the atomic number and the superscript A represents the mass number.

Examples: ^1_1H -Hydrogen

^4_2He - Helium

- A proton is also represented by ^1_1H since the nucleus of hydrogen contains only one proton.
- A neutron may be represented by ^1_1H and an electron by $^0_{-1}e$.
- A particle structure of a nucleus is known as a nuclide, e.g. the nuclide $^{17}_8O$ is a nucleus which has 8 protons and 9 neutrons while $^{16}_8O$ is another nuclide having 8 protons and 8 neutrons.

Nuclides which have the same atomic number but different mass number such as $^{16}_8O$ and $^{17}_8O$ are known as isotopes of the same element, that is oxygen in this case. Hence isotopes of an element have the same number of protons but a different number of neutrons in the nucleus. Isotopes have similar chemical properties because they have the same number of electrons but different physical properties as their masses are different. Other examples of isotopes are:

$$\left.\begin{array}{l} {}^{35}_{17}Cl: 17 \text{ protons and } 18 \text{ neutrons} \\ {}^{37}_{17}Cl: 17 \text{ protons and } 20 \text{ neutrons} \end{array}\right\} \text{Chlorine}$$

$$\left.\begin{array}{l} {}^{1}_{1}H: 1 \text{ proton only} \\ {}^{2}_{1}H: 1 \text{ proton and } 1 \text{ neutron (deuterium)} \\ {}^{3}_{1}H: 1 \text{ proton and } 2 \text{ neutrons (tritium)} \end{array}\right\} \text{Hydrogen}$$

Nuclides with the same mass number are known as isobars, e.g. ${}^{111}_{47}Ag$ and ${}^{111}_{48}Cd$.

Example 13.1

α-particles travelling at a speed of $3.0\times10^{6}\,ms^{-1}$ strike a block of gold. By assuming that the gold atoms are fixed in the block, calculate the nearest distance that an α-particle came to a gold nucleus. State your assumption of the law that still holds for such small distances. What is your assumption about the path of the incident particles?

[Mass of α-particle $=6.8\times10^{-27}\,kg$, atomic number of gold = 79, permittivity of free space, $\varepsilon_o = 8.9\times10^{-12}\,Fm^{-1}$]

Solution

Kinetic energy of α-particle

$$E_o = \frac{1}{2}mv^2 = \frac{1}{2}(6.8\times10^{-27})(3.0\times10^{6})^2$$

Let r = nearest distance that α-particle came to the gold nucleus.

Electrical potential energy of α-particle at a distance r from the gold nucleus is

$$U = \frac{qQ}{4\pi\varepsilon_o r} \qquad (q = \text{charge of α-particle, } Q = \text{charge of gold nucleus})$$

From the principle of conservation of energy,

$$U = E_o; \qquad \frac{qQ}{4\pi\varepsilon_o r} = E_o$$

$$\therefore r = \frac{qQ}{4\pi\varepsilon_o E_o} = \frac{(2\times1.6\times10^{-19})(79\times1.6\times10^{-19})}{4\pi\times(8.9\times10^{-12})\times\frac{1}{2}(6.8\times10^{-27})(3.0\times10^{6})^2} = 1.2\times10^{-12}\,m$$

These assumptions were made:

1. The inverse square law holds for the electrostatic force between the α-particle and the nucleus, i.e. $F \propto \frac{1}{r^2}$.
2. The path of α-particles is a straight line directed towards the centre of the gold nucleus.

13.3 The Bohr Theory of the Hydrogen Atom

Rutherford postulated the nuclear atom in which the number of electrons around the nucleus and the number of protons in the nucleus are the same. Rutherford's model of the atom did not mention about the manner in which the electrons are arranged around the nucleus.

Thus, the question remains: why are the negatively charged electrons not attracted by the positive nucleus and fall into it? In 1913, Niels Bohr postulated his classical model of the simplest atom: the hydrogen atom. The hydrogen atom has a proton as its nucleus and one electron. According to Bohr, the electron can only orbit in certain allowed discrete orbits. The electron can only be in a certain orbit or another, but not in between these allowed orbits. These allowed orbits are known as stable energy levels. The electron in a hydrogen atom normally occupies energy level available and is said to be in the *ground state* where it has the least amount of energy. In this manner, the electron is also in its most stable state.

The electron in the ground state may be boosted to one of its many other allowed stable energy levels in several ways. An electron in the higher energy level is said to be in the *excited state*. The energy absorbed by the electron to jump from the ground state to one of the excited states is known as the *excitation energy*. If the excitation energy is expressed as eV, then V is known as the *excitation potential*.

Sometimes the electron in the ground state may receive sufficient energy to enable it to escape completely from the attractive force of the nucleus. When this occurs, the atom is said to be *ionized*. The energy required by the electron in the ground state to escape completely from the attraction of the nucleus is known as the *ionization energy*. If the ionization energy is expressed as $eV^{'}$, then $V^{'}$ is known as the *ionization potential*.

In general, the various means by which an excitation or ionization may occur are:

1. Bombardment with fast electrons

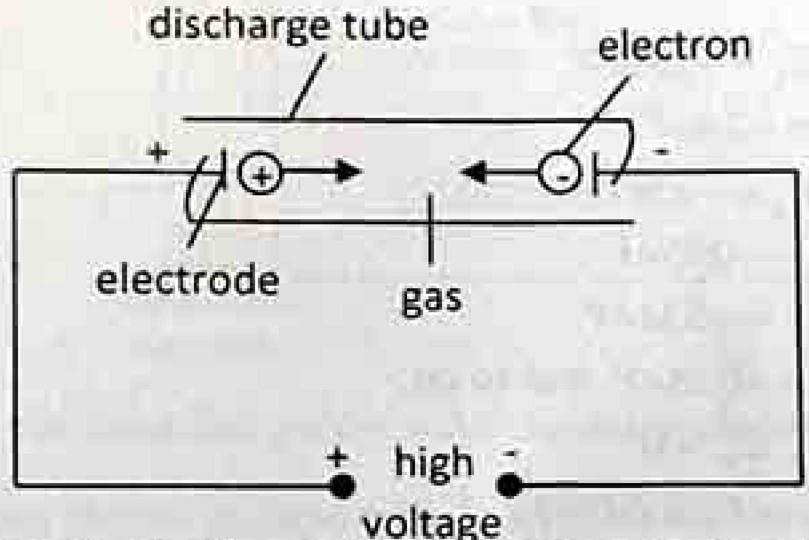


Fig. 13.3: Excitation and ionization in a discharge tube

One of the most common means of exciting or ionizing the gaseous atom is to use a discharge tube as shown in Figure 13.3. The tube contains a gas at low pressure. When a high potential difference is applied across the tube, ionization initially occurs near the electrodes. The electrons and ions produced are then accelerated by the high potential difference. When a fast electron collides with the electron in the ground state of gaseous atom, energy is transferred to the electron in the ground state. This causes either excitation or ionization. In the same way, the bombardment of electrons on the target atom in an X-ray tube causes excitation and ionization of the target atoms.

2. Collision between molecules in a flame

Collisions between molecules in a flame which have high thermal energy can also cause excitation and ionization. A very good example of this is the collision of molecules in the surface a hot compact white dwarf in binary stars which produces He-like (6.70*keV*) and H-like (7.00*keV*) Fe lines.

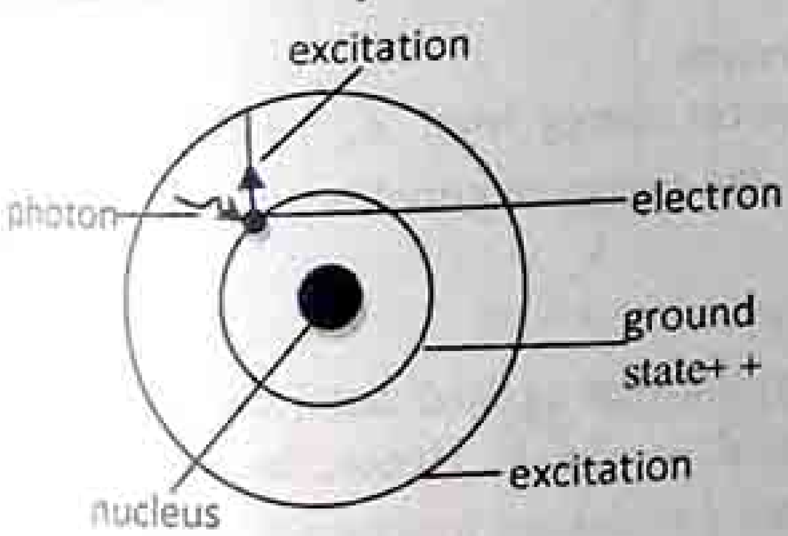


Fig 13.4: Excitation due to the collisions of photons

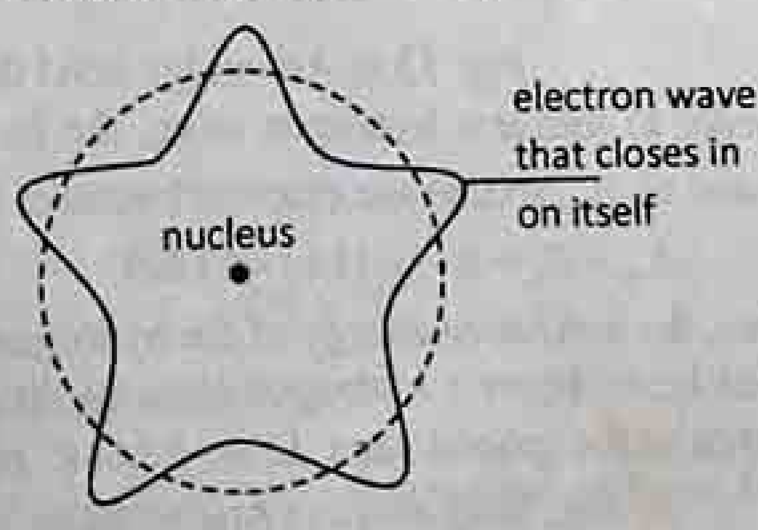


Fig. 13.5: A possible stable orbit

3. Collision with Photons

A photon of wavelength λ has momentum $p = \hbar/\lambda$. When this photon collides with an electron in the ground state, sufficient energy may be gained by the electron to enable it to jump to a higher energy level. Excitation caused by incident photons occurs in a fluorescent lamp. The walls of such a lamp are coated with a phosphorescent material, and mercury vapour at low pressure fills the lamp. The discharge of the mercury vapour at low pressure produces ultra-violet radiation, and the absorption of the ultra-violet photons by atoms of the phosphorescent material causes excitation.

Bohr's assumption of the existence of discrete stable energy levels was explained only in 1923 when De Broglie suggested the wave properties of electrons. Why the electron in the hydrogen atom occupies only discrete stable energy levels is explained by considering the electron as a wave and not a particle. An orbit for the electron exists where an electron wave closes in on itself and in phase. The wavelength of the electron wave must fit evenly into the circumference of the orbit as shown in Figure 13.5.

When the electron in a hydrogen atom is at the nth stable energy state, its energy is given by

$$E_n = -\frac{13.6}{n^2} eV \qquad n = 1, 2, 3, 4, \ldots$$

When $n = 1$, i.e. the ground state, $E_1 = -13.6eV$

When $n = 2$, i.e. the first state, $E_2 = -3.4eV$

When $n = 3$, i.e. the second state, $E_3 = -1.5eV$

When $n = 4$, i.e. the third state, $E_4 = -0.85eV$

When $n = 5$, i.e. the fourth state, $E_5 = -0.54eV$

When $n = 6$, i.e. the fifth state, $E_6 = -0.38eV$ and so on;

When $n = \infty$, i.e. for a free electron, $E_\infty = 0$

Notice that $E_\infty > \ldots > E_6 > E_5 > E_4 > E_3 > E_2 > E_1$

The allowed energy levels of the hydrogen atom can be represented by Figure 13.6.

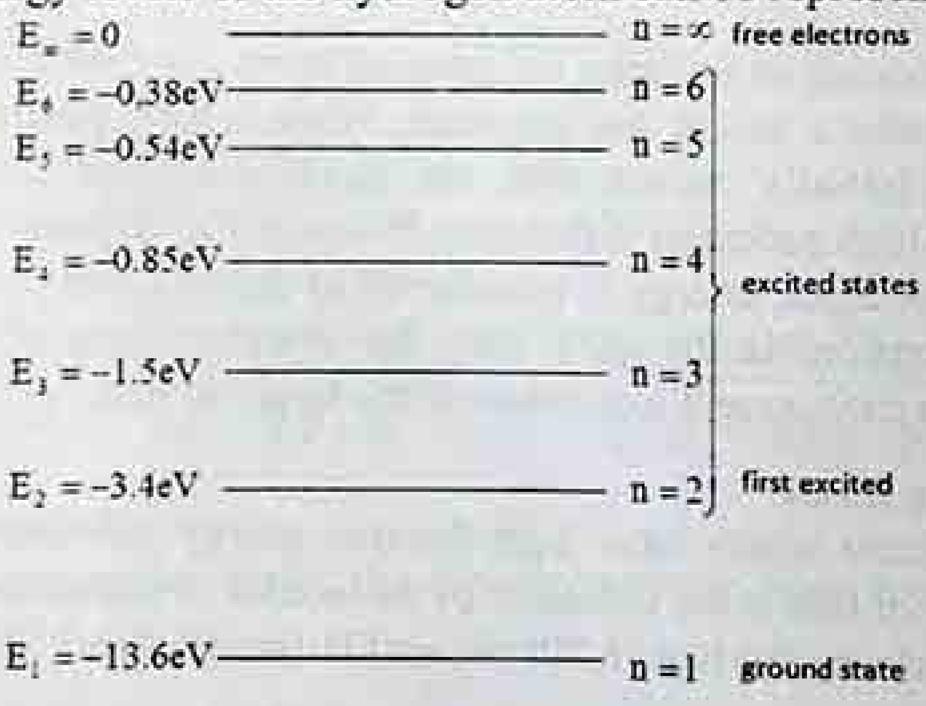


Fig. 13.6: An energy level diagram for hydrogen

In order to ionize a hydrogen atom, the electron in the lowest energy level $E_1 = -13.6eV$ must be boosted to E_∞, so that the electron becomes free. Therefore, the energy required is

$$E_\infty - E_1 = 0 - (-13.6) = 13.6eV$$

Hence, the ionization energy of the hydrogen atom is $13.6eV$.

Similarly, to boost a hydrogen atom in the ground state to the first excited state, the energy of the electron in the ground state $E_1 = -13.6eV$ must be raised to $E_2 = -3.4eV$. Hence, the energy required is $\quad E_2 - E_1 = -3.4 - (-13.6) = 10.2eV$

Conversely, when an electron in the first excited state drops back to the ground state, the difference in the energy of the electron is radiated as a photon, i.e.

$$E_2 - E_1 = \hbar v$$

where h = Planck's constant and v = frequency of the photon radiated. In general when an electron in an energy level E' drops to a lower energy level E'', the difference in energy, $(E' - E'')$ is related as a photon of frequency v, where $E' - E'' = hv$

13.4 The Line Spectrum of Hydrogen

The existence of discrete energy levels in a hydrogen atom can be deduced from the characteristic line spectrum of hydrogen which consists of various lines of different discrete frequencies. Figure 13.7, shows the transitions of the electrons that give rise to the various lines in the characteristic line spectrum of hydrogen.

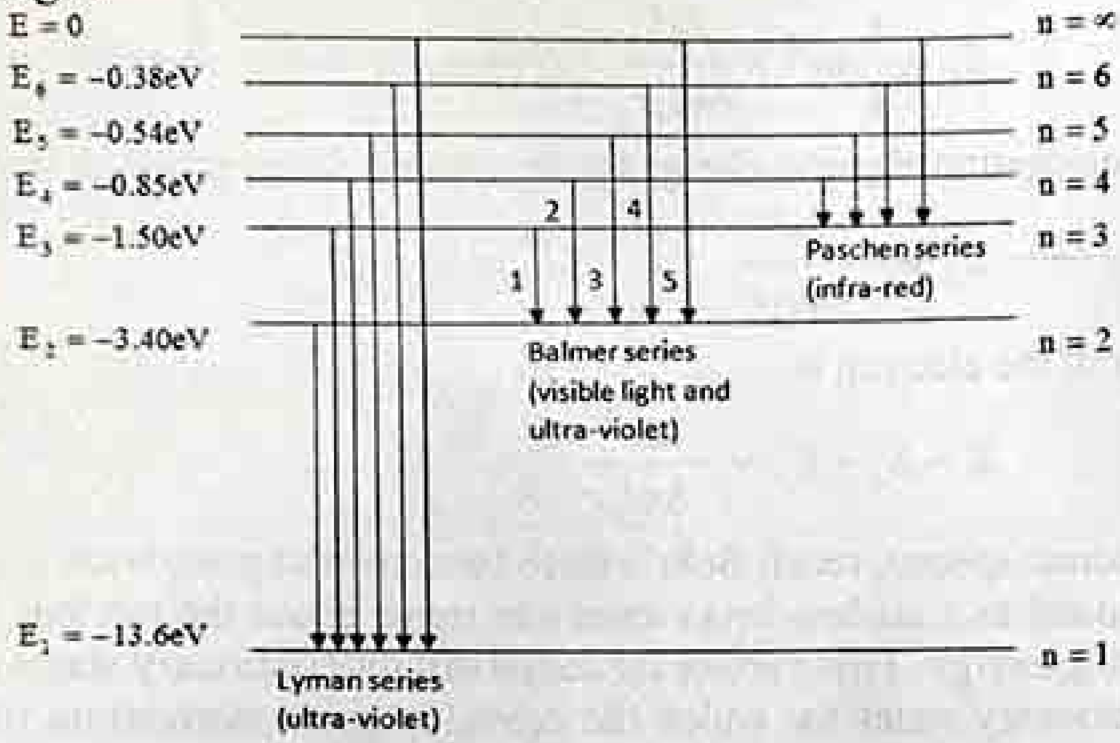


Fig 13.7: Transitions of electrons that produce the characteristic line spectrum of hydrogen

Of the three series of lines shown in Figure 13.7, only the *Balmer series* is visible light and hence can be seen by the eye. The *Balmer series* of spectral lines is emitted as the electron in the hydrogen atom falls from higher excited energy levels to the energy level $n = 2$. The red line (1) in Figure 13.8, is produced when the electron falls from the energy level $n = 3$ to $n = 2$. The blue-green (2) line shows the electron transition from $n = 4$ to $n = 2$. The line of highest frequency is the violet line (5) when the electron falls from $n = \infty$ to $n = 2$.

The spectral lines in the *Lyman series* have high frequencies because of the high energy differences involved, and they are in the ultra-violet region of the electromagnetic spectrum. All the lines in the Lyman series are due to electron transitions that end at the energy level $n = 1$.

On the other hand, the energy differences involved in the *Paschen series* are small. The spectral lines in the Paschen series are produced by the transitions of electrons that end at the energy level n = 3. These lines are not visible as they fall in the infra-red radiation region.

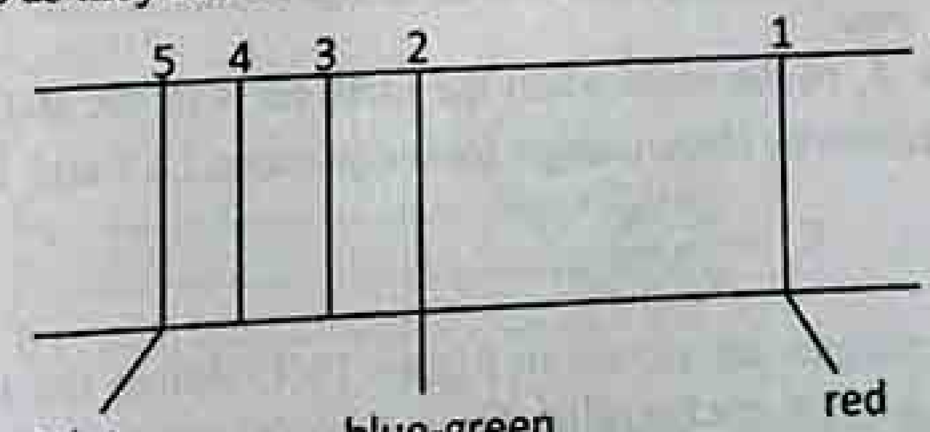


Fig. 14.8: Visible lines in the Balmer series

Assuming the Rutherford model, the coulomb force of attraction, F_e, between the nucleus and the electron provides the centripetal force, F_c, necessary to keep the electron of mass m in orbit. Therefore $F_c = F_e$, and

$$\frac{Ze^2}{4\pi\varepsilon_o r^2} = \frac{mv^2}{r} \tag{13.1}$$

The electron has tangential velocity v, the charge of the nucleus is Ze and ε_o is permittivity of free space, while r is orbital radius. Therefore

$$mv^2 = \frac{Ze^2}{4\varepsilon_o r} \tag{13.2}$$

The kinetic energy of the electron is

$$E_k = \frac{1}{2}mv^2 = \frac{Ze^2}{8\pi\varepsilon_o r} \tag{13.3}$$

The potential energy of the negative electron charge is

$$E_p = -\frac{Ze^2}{4\pi\varepsilon_o r} \tag{13.4}$$

Therefore, the total energy of the electron is

$$E = E_k + E_p = -\frac{Ze^2}{8\pi\varepsilon_o r} \tag{13.5}$$

To explain the observed atomic spectra, recall Bohr's three fundamental postulates:

(i) An electron bound to a nucleus in an atom can move about the nucleus in certain orbits without radiating energy. These orbits are called discrete stationary states of the atom.

(ii) Only those stationary states for which the orbital angular momentum of the electron is equal to an integral multiple of $\hbar = h/2\pi$ are permissible. That is

$$mvr = n\hbar \tag{13.6}$$

where n = 1, 2, 3, . . . and n is called the principal quantum number: $\hbar$ is the Planck's constant.

(iii) Whenever an electron makes a transition, that is jumps from an initial high energy state E_i To a final state E_f, a photon of energy $\hbar f$ is emitted, so that

$$\hbar f = E_i - E_f \tag{13.7}$$

where $f = c/\lambda$ is the frequency of the emitted radiation? From equations 13.2 and 13.6, the radii of the permissible orbits are:

$$r = \frac{n^2\hbar^2\varepsilon_o}{Z\pi\pi m^2} \tag{13.8}$$

Note that when $n = 1$, $r = \frac{\hbar^2\varepsilon_o}{\pi me^2} = 0.0529\,nm$ which is the radius of the smallest Bohr orbit of the hydrogen atom. The energy E_n of the atom when the electron is in the stationary orbit characterized by quantum number n is obtained by eliminating r from equations 13.5 and 13.8 giving;

$$E_n = \frac{-Z^2me^4}{8\varepsilon_o^2 n^2\hbar^2} \tag{13.9}$$

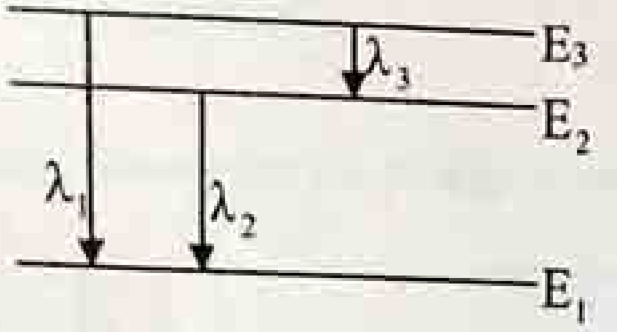
The relative values of the energies are shown in Figure 13.7. Where the lowest energy, the ground state is for $n = 1$, and the highest energy, that for $n = \infty$ corresponds to zero total energy as the electron becomes free from the atom. Using equation 13.7, for transition from orbit n_i to n_f,

$$f = \frac{E_i - E_f}{\hbar} \tag{13.10}$$

Example 13.2

Transitions between three energy levels in a hydrogen atom give rise to three spectral lines of wavelengths, in increasing magnitudes, λ_1, λ_2 and λ_3. What is the relationship between λ_1, λ_2 and λ_3?

Solution



The transitions that give rise to the wavelengths λ_1, λ_2 and λ_3 can be represented by the figure above.

$$E_3 - E_2 = \frac{\hbar c}{\lambda_3} \qquad (1)$$

$$E_2 - E_1 = \frac{\hbar c}{\lambda_2} \qquad (2)$$

(1) ÷ (2):

$$E_3 - E_1 = \frac{\hbar c}{\lambda_3} + \frac{\hbar c}{\lambda_2} \qquad (3)$$

Also

$$E_3 - E_1 = \frac{\hbar c}{\lambda_1} \qquad (4)$$

(3) = (4):

$$\frac{\hbar c}{\lambda_1} = \frac{\hbar c}{\lambda_2} + \frac{\hbar c}{\lambda_3}$$

$$\frac{1}{\lambda_1} = \frac{1}{\lambda_2} + \frac{1}{\lambda_3}$$

Example 13.3

The first excitation energy of the hydrogen atom is $10.2eV$. Explain what is meant by this statement. Find the speed of the slowest electron that would cause excitations of a hydrogen atom.

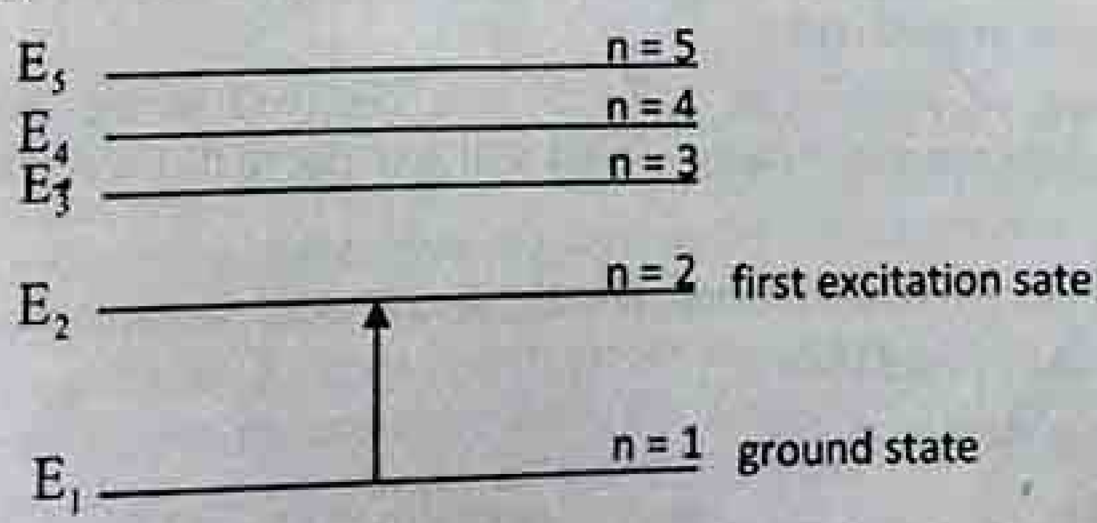
Solution

The statement means that $10.2eV$ of energy must be absorbed by the electron at the energy level $n = 1$ (ground state) to raise it to the energy level $n = 2$ (first excitation state).

Let v = speed of the slowest electron, then

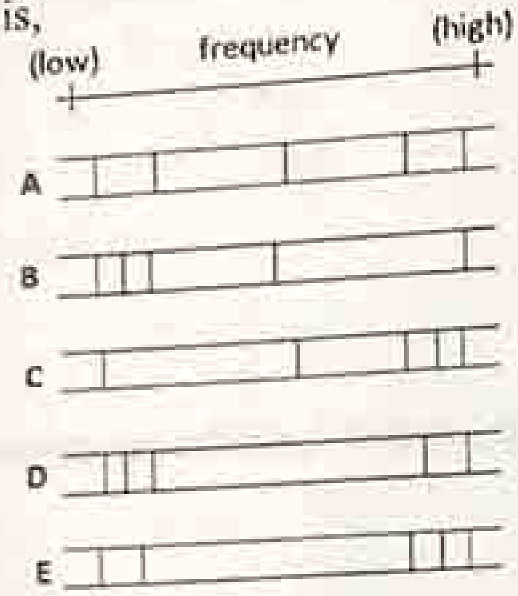
$$\frac{1}{2}mv^2 = 10.2eV$$

$$v = \sqrt{\frac{2 \times 10.2 \times (1.6 \times 10^{-19})}{9.11 \times 10^{-31}}} = 1.89 \times 10^6 ms^{-1}$$



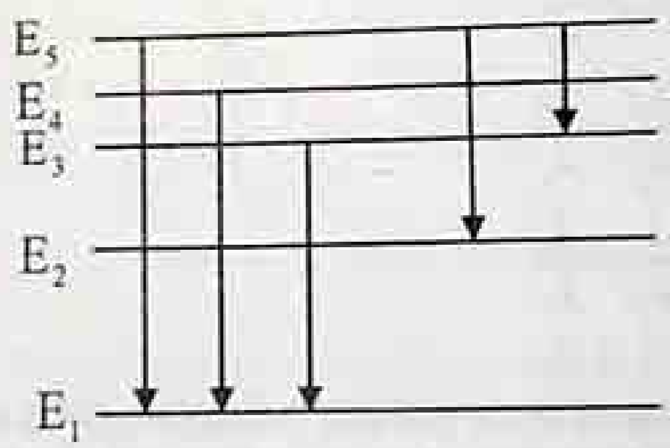
Example 13.4

The figure (on the left) shows five energy levels of an atom. Five transitions between the levels are indicated, each of which produces a photon of definite frequency. The spectrum which corresponds to the transitions indicated is,



Solution

The first line on the right has the highest frequency and is due to the transition $E_5 \rightarrow E_1$. The second line represents the transition $E_4 \rightarrow E_1$. Since the difference in energy between $(E_5 - E_1)$ and $(E_4 - E_1)$ is small, the difference in frequency between the first and second line is also small. Thus, the two lines are close.



Similarly, the distance between the second and third line is directly proportional to $(E_4 - E_3)$; the distance between the third and fourth line is directly proportional to $(E_2 - E_1)$; and the distance between the fourth and fifth line is directly proportional to $(E_3 - E_2)$. Thus, the best representation of the line spectrum is E.

Example 13.5

An atom is assumed to have zero energy in the ground state; and its energy in the first, second and third excited states are $1.635 \times 10^{-18} J$, $1.936 \times 10^{-18} J$ and $2.024 \times 10^{-18} J$ respectively.

(a) What is the wavelength of the photon which would excite the atom from the first excited state to the second excited state?

(b) A blue line of wavelength $5.17 \times 10^{-7} m$ is observed in the spectrum of this atom. The transition between which energy levels will give rise to this spectral line?

Solution

(a) If λ = wavelength of the incident photon, then to raise the atom from the first excited state E_2 to the second excited state E_3,

Energy of photon = $E_3 - E_2$

$$\frac{hc}{\lambda} = E_3 - E_2;\ \lambda = \frac{hc}{E_3 - E_2} = \frac{(6.63 \times 10^{-34})(3 \times 10^{8})}{(1.936 - 1.635) \times 10^{-18}} = 6.61 \times 10^{-7} m$$

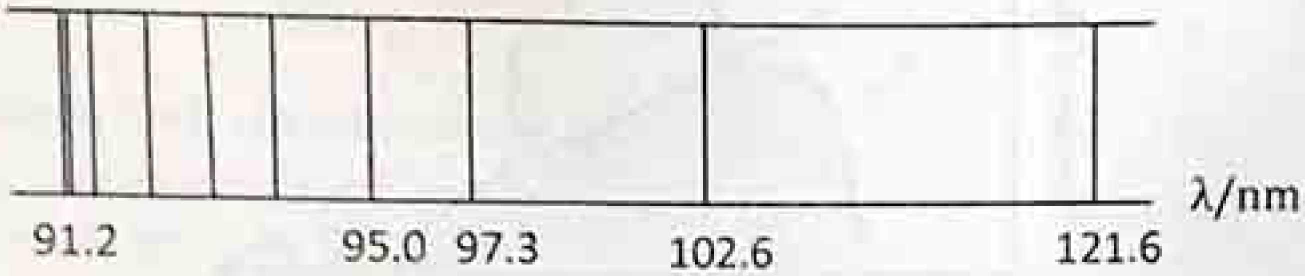
(b) Energy of the blue light photon $\frac{hc}{\lambda} = \frac{(6.63\times10^{-34})(3\times10^{8})}{5.17\times10^{-7}} = 3.85\times10^{-19}J$

Also, $E_4 - E_2 = (2.024 - 1.635)\times10^{-18}J = 3.89\times10^{-19}J$

Hence, the transition which gives rise to this blue line is from E_4, the third excited state, to E_2, the first excited state.

Example 13.6

What is a photon? Show that the energy E of a photon and its wavelength λ are related by $E\lambda = 1.99\times10^{-16}\,Jnm$.



The figure above represents part of the emission spectrum of atomic hydrogen. It consists of a series of lines, the wavelengths of some of which are shown. The shortest wavelength in the series measures 91.2*nm*.

13.5 Production of X-rays

X-rays, which are also called Rontgen rays after the discoverer, are produced when high speed electrons are suddenly stopped, especially by metals of high atomic number. An arrangement for producing X-rays is shown in Figure 13.9 (Coolidge tube). It consists of a high vacuum tube with the stream of electrons obtained from a heated filament F placed in the tube. The intensity of the X-rays which depends only on the rate of emission of electrons from the filament can be controlled by the filament current. The electrons are accelerated towards the anode target by a potential of up to 100*kV*. The target T consists of a hard core of high melting point metal, such as tungsten, embedded in a block of copper.

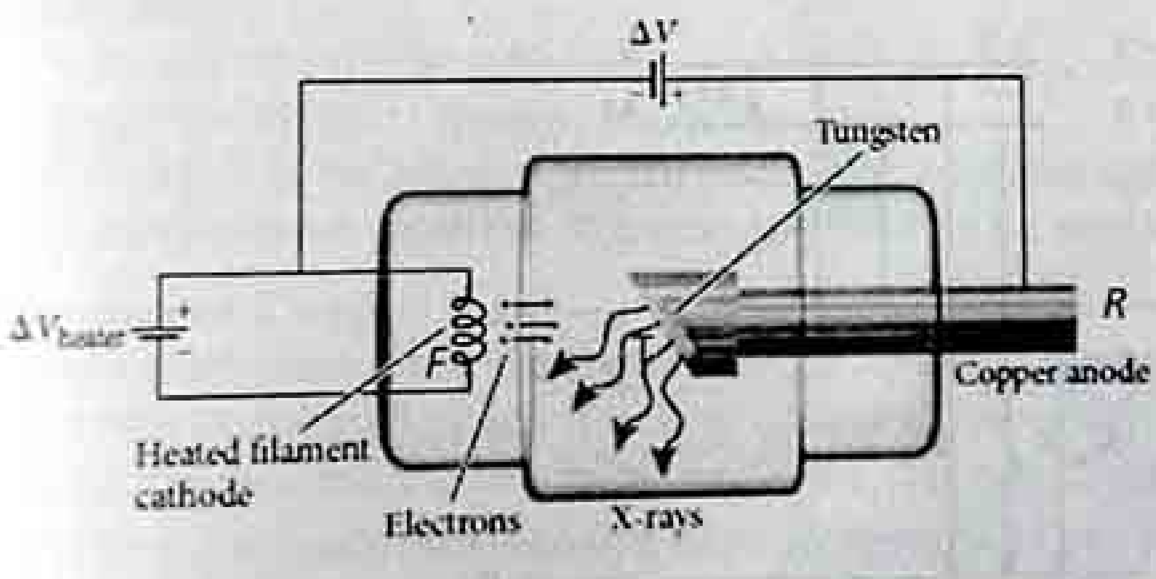


Fig 13.9: X-rays Production

Only about 1 percent of the energy of the electrons incident on the target is converted into X-rays. The rest of the electron energy is lost in collisions hence producing heat. The targets are therefore very hot and must be cooled by circulating a stream of cold water through it or by providing it with external radiator fins R as in the refrigerator. Apart from the intensity of X-rays, another parameter of interest is the quality or penetrating power of the X-rays. This depends partly on the nature of the target but mainly on the accelerating potential V. Since X-rays are penetrating radiation, they constitute a health hazard. To protect operators from its harmful effects, the X-ray tube is usually enclosed in a lead shield of thickness at least 5 mm which absorbs all the rays except those allowed to emerge through window W.

13.6 X-ray Spectra

A plot of the intensity of X-rays emitted by a target against wavelength has the general shape shown in Figure 13.10. There is (1) a continuous spectrum and (2) a sharp line spectrum super-imposed on the continuous spectrum. The continuous spectrum results from the radiation emitted by the incident electrons as they are accelerated in the Coulomb field of force of the nuclei of the target atoms. The term *bremmstrahlung* is used for this type of radiation.

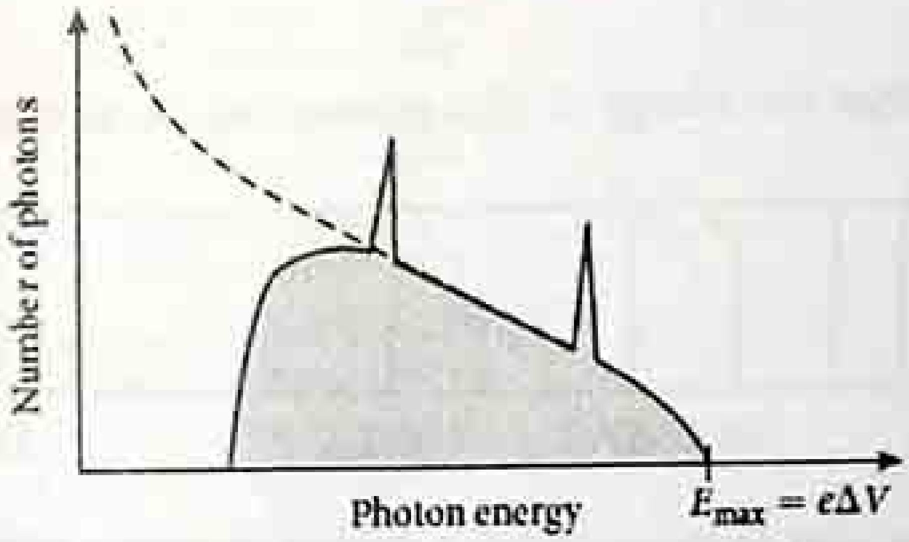


Fig. 13.10: Typical X-ray Spectrum

If V is the accelerating potential for the electrons, kinetic energy with which they strike the target is V_e. The X-rays are emitted with a continuous range of frequencies up to a maximum f_{max}. If an incident-electron uses all its energy to produce X-rays, $V_e = \hbar f_{max}$.

Therefore, $$f_{max} = \frac{V_e}{\hbar}$$

Since $$c = f_{max} \times \lambda_{min}$$

The minimum wavelength from the tube is $\lambda_{min} = \dfrac{\hbar c}{V_e}$

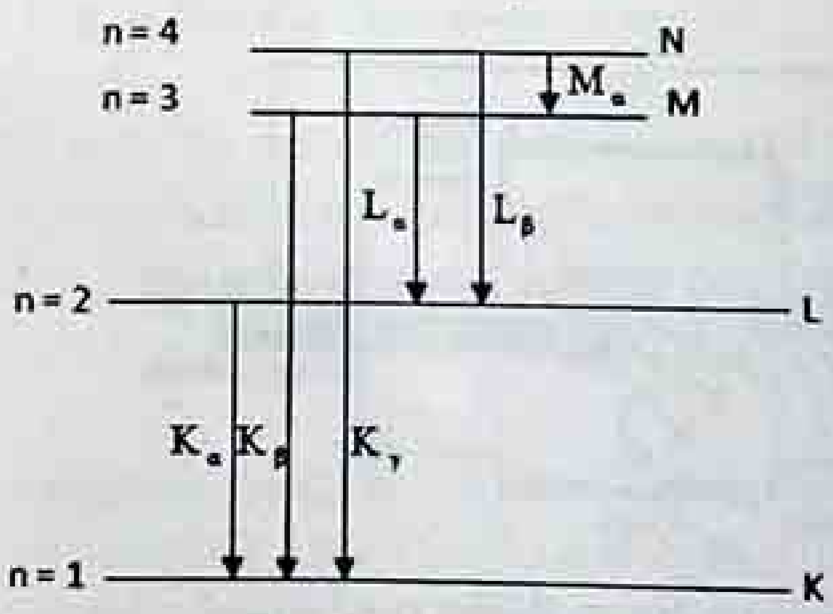


Fig. 13.11: Origin of the characteristic X-ray lines

From the formula, this minimum wavelength decreases as the accelerating potential increases, which means more energetic or penetrating X-rays. Note that λ_{min} is independent of the target material.

The sharp line spectrum is characteristic of the target element. To consider the origin of this characteristic spectrum, we must use the spectroscopic notation for the atomic energy levels. The n = 1, 2, 3, 4... levels are designated, K, L, M, N,... levels, respectively. Suppose an incident electron knocks off, for example, a K electron from the target, the hole thus created has to be filled and the energy changes that take place as a result of the rearrangement of the electrons in the various electronic levels leads to the production of the characteristic lines. If the gap is filled by an L electron, the K_α line is produced; if by an M electron, the K_β line is produced, e.t.c. The possible transitions and corresponding lines are illustrated in Figure 13.11. The spectrum consists of discrete and characteristic lines because the atomic energy levels have discrete fixed values for any atom; the energy of the emitted X-rays is obtained by simple subtraction involving these discrete energy values. For example, if υ_a is the frequency of the K_a line, then

$$h\upsilon_a = E_K - E_L,$$

where E_K, E_L are the energies of the K and L levels, respectively?

Example 13.7

The K absorption limit of palladium is 50*pm*. Find the minimum potential difference which must be used across an X-ray tube to excite the K series.

Solution

To excite the K series, the K electron must be knocked off.

K absorption edge occurs at $50 \times 10^{-12} m$.

The corresponding energy of the K level;

$$E = \frac{hc}{\lambda} = \frac{6.6 \times 10^{-34} \times 3 \times 10^{8}}{50 \times 10^{-12}} J = 3.96 \times 10^{-15} J$$

$$= \frac{3.96 \times 10^{-15}}{1.6 \times 10^{-19}} eV = 24.8 keV$$

This is the ionization energy. Therefore, the minimum potential difference to be used across the X-ray tube = 24.8*keV*.

13.7 Properties of X-rays

The properties which have been identified for X-rays include

(i) X-rays are electromagnetic waves of the same nature as light waves but with much shorter wavelength. The wavelength ranges from 0.01 to 10*nm*.

(ii) X-rays affect photographic plates just as light waves.

(iii) X-rays are not influenced by electric or magnetic fields and are therefore not charged particles.

(iv) X-rays discharge electrified bodies whether charged positively or negatively.

(v) X-rays ionize gases through which they pass.

(vi) X-rays can be reflected, refracted and diffracted just as light waves.

(vii) X-rays penetrate matter to an extent governed by the density of the material.

13.8 Applications of X-rays

1. We have already mentioned the application of X-rays in the extensive subject crystallography leading to the understanding of crystalline structure. Some of the highlights in this type of work include (a) the fact that the determination of the structure of penicillin and vitamin B_{12}, when chemical methods had failed, allowed these compounds to be synthesized for widespread use in medicine. (b) the X-ray work on fibrous proteins, on puriness and pyrimidines, and on DNA led to the double-helix model of Crick and Watson.
2. The condition in (vii) above has been used for the diagnostic use of X-ray photographs in medicine in identifying broken bones or locating the position of metallic objects inadvertently swallowed. As earlier mentioned, the principle of this radiographic examination is that in the structure under investigation, the presence of any defect like bone fracture is shown by the differential absorption of the incident X-rays by various parts of the body using radiation of uniform intensity. The transmitted beam falls on photographic film which is subsequently processed. Since many organs show no greater absorption than the surrounding tissue, it is often necessary to insert into the organ to be studied material which has a much greater absorption. Such materials e.g. barium salts and iodine compounds are called contrast media and lead to better contrast in the radiograph. For example, in the examination of gastrointestinal tract, the contrast medium of barium sulphate may be fed from the patient. With that, pathological problems in the gastrointestinal tract such as suspected tumour, swallowed foreign bodies,

ulcers and inflammations may be diagnosed by X-ray examination.This method is also applicable in detecting cracks and flaws in metal castings in metallurgy.

3. Still with radiographic examination, X-rays are used to detect multiple pregnancies, foetal deformity or confirmation of pregnancy. Gamma and X-rays freely penetrate all the growing cells of the body causing abnormality of almost any system in the infant. X-ray examinations should therefore be avoided before the 20th week of gestation.It should be done cautiously if at all required.

4. Another use of X-rays is in the photographic copying process employing X-rays as a source and known as xerography. Light is passed through the document to be copied and falls on an electro-statically charged plate which is previously coated with a semiconductor material. The effectiveness of this process depends on the intensity of the radiation. A powder with an opposite electric charge is then spread into the plate and sticks to the dark areas where the plate has not been discharged by the incident X-rays. The powder is then transferred from the plate to a charged paper where it is fixed for annealing.

5. Note that all the applications of radio-isotopes given in section also apply to X-rays. For therapeutic purposes, X-rays have the advantage over radioactive isotopes in that no implant operation is necessary before their use. A collimated beam of X-rays can be directed to the target volume, machines operating under $250kV$ being used for tumours on or near the skin, and others operating at around $4MW$ for deep-seated tumours. Since rapidly dividing cancer or tumorous cells are more vulnerable to radiation than healthy ones, this controlled dose of X-rays can be applied to destroy a growing tumour.It must be emphasized that in all applications of X-rays, extreme care must be exercised because the hazards associated with radioisotopes apply to equally to X-rays.

Summary

1. Rutherford postulated a model of the atoms consisting of a central massive core which he called the nucleus. The nucleus carries the entire atom's positive charge and nearly all its mass.
2. The total number of nucleons, i.e. protons and neutrons, in a nucleus is known as the mass number A. the symbol used to represent a nucleus X is

$$^{A}_{Z}X$$

   Where the subscript Z represents the atomic number and the superscript A represents the mass number.
3. Nuclides which have the same atomic number but different mass number such as $^{16}_{8}O$ and $^{17}_{8}O$ are known as isotopes of the same element, that is oxygen in this case.
4. Bohr's three fundamental postulates:
   (i) An electron bound to a nucleus in an atom can move about the nucleus in certain orbits without radiating energy. These orbits are called discrete stationary states of the atom.
   (ii) Only those stationary states for which the orbital angular momentum of the electron is equal to an integral multiple of $h/2\pi$ are permissible. That is

$$mvr = n\hbar$$

   where $n = 1, 2, 3, \ldots$ and n is called the principal quantum number: $\hbar$ is the Planck's constant.
   (iii) Whenever an electron makes a transition, that is jumps from an initial high energy state E_i to a final state E_f, a photon of energy $\hbar f$ is emitted, so that

   $\hbar f = E_i - E_f$ where $f = c/\lambda$ is the frequency of the emitted radiation.
5. X-rays are electromagnetic waves and can be produced in the laboratory by bombarding a hard target such as copper anode with highly accelerated electrons and has wide applications in Industries and Medicine.

Exercise 13

13.1 The deviation of α-particles by thin metal foils through angle of 90° - 180° shows that:
A. α-particles struck a massive target. **B.** α-particles struck a negatively charged target
C. α-particles struck a positively charged target. **D.** α-particles struck a neutral target.

13.2 The experiment by Geiger and Marsden where a thin gold foil is bombarded by α-particles shows that:
A. Atom has a nucleus **B.** Atom has positively charged protons
C. Atom has negatively charged electrons **D.** Elements are made up of atoms.

13.3 What is the magnitude of the electric field strength at a distance r from an isolated stationary nucleus X represented by the symbol ${}^{A}_{Z}X$?

13.4 An α-particle of energy $5.30MeV$ moves directly towards a lead nucleus ${}^{206}_{82}Pb$ which is stationary. Calculate the nearest distance of approach of the α-particle from the lead nucleus. State any assumptions that you make in your calculation.

13.5 An atom of sodium atom is represented by ${}^{23}_{11}Na$. Which of the following statements is/are correct? (i) The nucleon number (mass number) is 12.
(ii) The sodium atom is determined by mass number.
(iii) The atomic number is 11. (iv) The number of protons is 11.
A. i only **B.** i & ii, **C.** ii, iii, iv **D.** all of the above.

13.6 The large number of α-particles passing through the gold foil undeflected suggests that:
A. there exists large empty spaces in an atom.
B. α-particles are not deflected by the atom.
C. The gold foil is transparent to α-particles.
D. α-particles are opaque

13.7 When a fast electron collides with the electron in the ground state of gaseous atom, energy is transferred to the electron in the ground state and this causes:
(i) excitation (ii) ionization (iii) collisional ionization (iv) photoionization
A. all of the above B. i only C. i & ii D. i, ii, & iii

13.8 The Lyman series are due to electron transitions that end at the energy level
A. $n = 1$ **B.** $n = 2$ **C.** $n = 3$ **D.** $n = 5$

13.9 These are properties x-rays <u>except</u>:
A. X-rays are electromagnetic waves with wavelength range from 0.01 to 10nm.
B. X-rays affect photographic plates just as light waves.
C. X-rays are influenced by electric or magnetic fields.
D. X-rays ionise gases through which they pass.

13.10 Collision of molecules in the surface of a hot compact white dwarf in binary stars produces:
(i) He-like (6.70keV) (ii) H-like (7.00keV) Fe lines, (iii) 6.4keV line (iv) O_2 ion line.
A. i & ii only, **B.** i, ii, & iii **C.** iv only **D.** all of the above.

13.11 Describe Millikan's oil drop method for estimating electronic charge. What is the significance of this experiment?

13.12 Derive the Bohr formula for the allowed energy levels of the hydrogen atom. Discuss the assumptions made in the derivation.

13.13 Describe one method for the generation of X-rays. Explain the factors that influence the intensity and penetration power of generated X-rays.

13.14 Distinguish between continuous and characteristics x-ray spectra.

13.15 Determine the speed of an electron whose kinetic energy equals its rest energy.

13.16 Discuss the use of absorption spectra.

13.17 Write notes on the properties of X-rays.

13.18 Discuss the application of X-rays in medicine. Comment briefly on the hazards.

13.19 A proton has a mass of $1.67\times10^{-27}kg$. Determine its rest energy in (i) joules (ii) MeV.

13.20 If the proton is moving with a velocity $0.5c$, find (a) its kinetic energy (b) its total energy in MeV.

CHAPTER 14
WAVE-PARTICLE DUALITY OF MATTER

14.0 Introduction

Electromagnetic waves such as light exhibit dual nature as they posses both waves properties and particle-like properties. Conveniently, particles like electrons have wave like properties as well as particle like properties. This is referred to as the wave-particle duality and forms the basis of what is known as the quantum theory. Electromagnetic waves behave as wave motion in diffraction and interference but as particles (photons) in photoelectric effects. Fast moving electrons posses mass and momentum. According to Planck's and Einstein, electromagnetic radiation such as light comes in packets of quanta and only whole number of quanta can exist. The quantized radiation is given by:

$$E = hf \tag{14.1}$$

where h = Planck's constant = 6.63×10^{-34} Js and f is the frequency of radiation. The quantum of light or electromagnetic radiation is known as photons.

Put in another language, Wave-particle duality of matter implies that matter has both the nature of wave and of a particle (a supposed tiny solid element). This phenomenon can be exhibited by elementary particles of the atom, like the electron, and also by electromagnetic radiation. Therefore we can talk about wave-particle behaviour of electrons and electromagnetic radiation or light.

The Bohr's theory of the hydrogen atom which we treated in the last chapter, involved the use of the concept of emission and absorption of radiation in whole quanta. This was really an extension of Planck's hypothesis. In 1901, Planck, in proposing a successful theory of black body radiation, introduced a revolutionary assumption regarding the way in which radiation was emitted and absorbed by atoms. He postulated that radiation was emitted unit of energy hf is called a quantum of energy or a photon, if we are referring to electromagnetic radiation of frequency f. Note that this concept of discrete set of energy values or energy levels for a system emitting radiation is completely at variance with classical ideas. Classically, the energy changes are smooth and the total energy can have continuous values. In effect, one is assigning particle characteristics to waves (i.e. electromagnetic waves).

In our everyday sense, there is no difficulty distinguishing between waves and particles. Waves are associated with such parameters as wavelength and frequency while particles have mass, momentum and charge (where applicable). However, in the microscopic world, the division is not clear-cut. We have already met the case of wave manifesting as particle; on the other hand, particles such as electrons can exhibit wave characteristics. In this chapter we will study a few phenomena with a view to elucidating the dual characteristics exhibited by waves and particles.

14.1 Photoelectric Effect

Electrons are emitted from an insulated metal surface when light of sufficiently high frequency falls on the surface. This phenomenon is called the photoelectric effect. Millikan investigated the effect in detail with equipment of the type shown in Figure 14.1. A variable potential difference V is maintained.

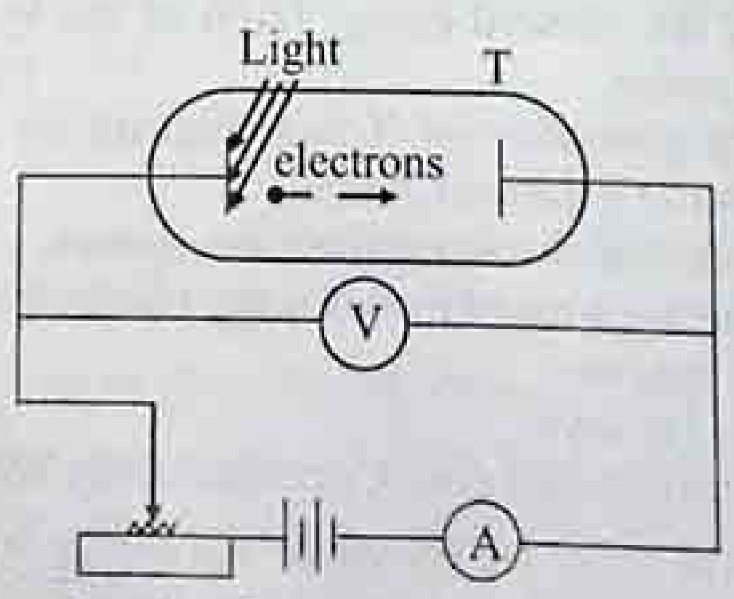


Fig. 14.1: Apparatus for investigating photoelectric emission between two electrodes in an evacuated tube T.

The main features photoelectric effects are:

(a) At a given frequency of the incident light, the photoelectric current (i.e. number of electrons emitted per second) is proportional to the intensity of incident light.

(b) The energy of the emitted electrons is independent of the intensity of the incident light; it is related only to the frequency of the incident light.

(c) There is no detectable time lag between the irradiation of the surface and the ejection of the electrons.

(d) For each material, there is a threshold frequency f_0 of the incident light, below which no photoelectrons are emitted.

These features could not be explained using the wave theory of light. Under that theory, the kinetic energy of the emitted electrons should increase with the intensity of the incident light. Also, the photoelectric effect should occur for any frequency of the incident light provided the light is intense enough. In addition, under the wave theory, for a very feeble source, the part of the wave front that is intercepted by an electron in the irradiated material will be so small that it will take the electron a long time to absorb enough energy to get emitted. Therefore, there should be a time lag between the incidence of the light and the electron emission.

As a result of the failure of the wave theory of light to explain the observed effects in photoelectricity, Albert Einstein (1905) enunciated a mechanism for the phenomenon based on Planck's quantum theory of black-body radiation. Einstein assumed that light is not only emitted a quantum at a time but is also propagated as individual quanta called photons. A photon of light of frequency f carries an amount of energy $E = hf$ where h is Planck's constant. In Einstein's explanation of the photoelectric effect,

(i) The entire energy of a photon is transferred to a single electron in the metal, which gets emitted instantaneously. This immediately removes the difficulty regarding time lag.

(ii) When the electron comes out of the metal surface it will have a maximum kinetic energy given by

$$\frac{1}{2}mv^2 = hf - \phi \qquad (14.2)$$

This equation shows that no electron can be emitted if the frequency of the incident light is so small that $hf < \phi$. Thus, there is a threshold frequency in agreement with experimental results. If the frequency is f_0 then $\phi = hf_0$. ϕ is the work function of the surface and is the minimum energy needed to dislodge an electron from the metal surface being illuminated. It is of the order of a few electron volts. A plot of electron energy T against frequency of the incident light is of the form shown in Figure 14.2. Planck's constant can be obtained from the gradient.

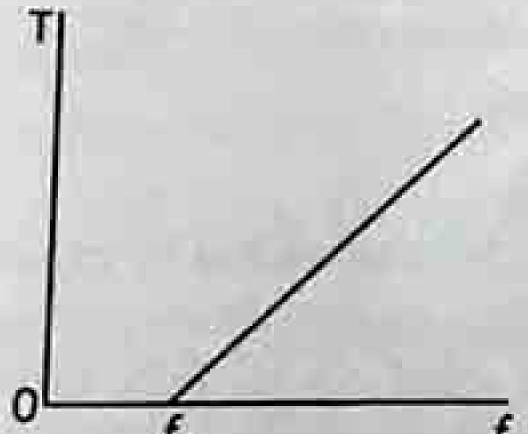


Fig. 14.2: Electron kinetic energy against frequency of incident light

Example 14.1

When a copper surface is illuminated by radiation of wavelength 2537 Å the value of the stopping potential is found to be 0.24 volts. Calculate (i) the work done by the electron in escaping through the surface of the copper (ii) the wavelength of the threshold frequency of copper, (iii) the velocity of the most energetic electron ejected by the incident light.

Solution

Energy of incident radiation $= hc/\lambda$

$$= \frac{6.6\times10^{-34}\times3\times10^{8}}{2537\times10^{-10}}J = \frac{6.6\times10^{-34}\times3\times10^{8}}{2537\times10^{-10}\times1.6\times10^{-19}}eV = 4.88eV$$

K.E of the emitted electron $= 0.24eV$

(i) Work done by electron escaping through surface= 4.88 – 0.24 eV = 4.64 eV

(ii) $\frac{hc}{\lambda_0} = 4.64eV = 4.64\times1.6\times10^{-19}J$

$$\therefore \lambda_0 = \frac{6.6\times10^{-34}\times3\times10^{8}}{4.64\times1.6\times10^{-19}} = 2667\times10^{-10}m$$

(iii) $\frac{1}{2}mv^2 = 0.24\times1.6\times10^{-19}J$

$$v^2 = \frac{2\times0.4\times1.6\times10^{-19}}{9.31\times10^{-31}}$$

$$v = 2.9\times10^{5}ms^{-1}$$

14.2 Compton Effect

The Compton effect or Compton scattering is said to occur when a high-energy photon of wavelength λ, collides with a target-atom or molecule, thereby releasing loosely bound electron as well as a scattered X-ray photon of greater wavelength λ'. The change in wavelength, $\Delta\lambda$ of the radiations is given by

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{mc}(1-\cos\theta). \qquad (14.3)$$

Equation 14.3 is the Compton formula which conforms with experimental observation in which case the wavelength λ' of the scattered X-rays is always greater than the wavelength of the incident radiation, see Figure 14.3. This is because the recoil electron carries away some energy. Also, the wavelength change is independent of the scattering material, but depends on the angle of scattering. However, the formula does not account for the observed unmodified line. This is because the derivation considers only free electrons, but there are also electrons in the inner shells which are tightly bound. When such electrons are hit by incident photons, they experience little or no recoil. The energy of the photon is therefore affected minimally. This accounts for the observed unmodified line: it is more prominent with heavier elements (Figure 14.3). The quantity h/mc in equation 14.3 is known as the Compton wavelength of the electron; it has value $0.002426nm$. Note also that the maximum wavelength shift in Compton scattering occurs when $\theta = 180°$.

The derivation of the above equation is shown below.

Explanation of Compton Effect

The scattered radiation experiences a wavelength shift that cannot be explained in terms of classical wave theory, thereby lending support to Einstein's photon theory. Probably the most important implication of the effect is that it showed light could not be fully explained according to wave phenomena and further confirmed the particle nature of radiation. When X-rays were scattered by a target with loosely bound electrons e.g. carbon, the scattered radiation was found to consist of two components- one having the same wavelength as the incident beam (unmodified line) and the other having a slightly longer wavelength (modified line). According to the classical electromagnetic theory, there should be no change in wavelength or frequency. The incident radiation was expected to set the atomic electron vibrating with the frequency of the incident radiation, and then produce radiation emitted in all directions with the same frequency (scattered radiation). Compton in 1923 provided the explanation to the observed effects by treating the incident radiation as a stream of individual photons each of which could interact with a single electron. This is the Compton effect and the change in wavelength is given as $\Delta\lambda$.

Derivation of Compton Effect

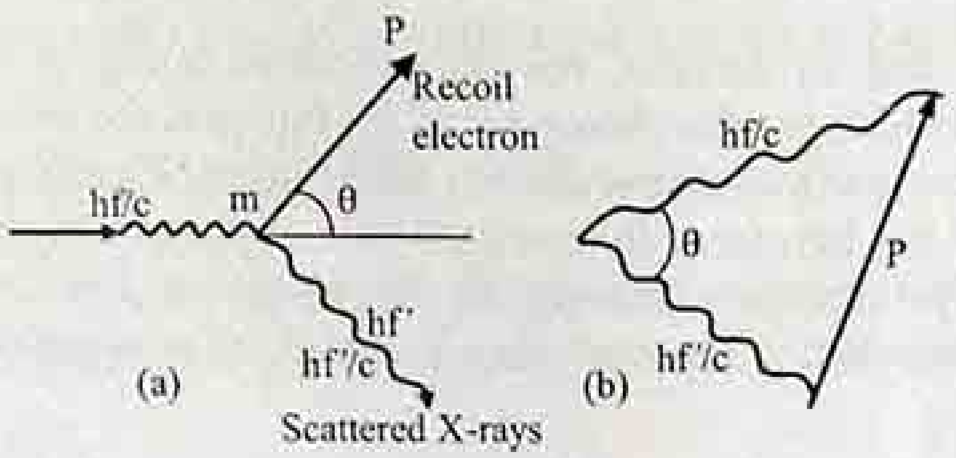


Fig. 14.3: (a) the Compton effect (b) momentum triangle for the Compton effect

Suppose the frequency of the incident photon is f, the momentum is hf/c. The photon strikes a relatively free electron of mass m which recoils with momentum $P = mv/\left(1 - v^2/c^2\right)^2$ while the incident X-ray photon is scattered through angle θ with new energy hf', the corresponding momentum being hf'/c. This is shown in Figure 14.3(a). The interaction is treated as a simple collision problem in mechanics in which case the initial momentum vector hf/c of the X-ray is equal to the two vectors P and hf'/c as shown in Figure 14.3(b). From energy conservation,

$$hf + mc^2 = hf' + E$$

But
$$E = mc^2/\left(1 - v^2/c^2\right)^{1/2}$$

Therefore,
$$h(f - f') + mc^2 = mc^2\left(1 - v^2/c^2\right)^{1/2}$$

Dividing both sides by mc^2, squaring and subtracting 1;

$$\left[\frac{h}{mc^2}(f - f') + 1\right]^2 - 1 = \frac{v^2}{c^2\left(1 - v^2/c^2\right)}$$

Multiplying both sides by m^2c^4,

$$m^2c^4\left\{\left[\frac{h}{mc^2}(f - f') + 1\right]^2 - 1\right\} = \frac{m^2v^2c^2}{1 - v^2/c^2} \qquad (14.4)$$

From the vector triangle (Figure 14.3b)

$$P^2c^2 = (hf')^2 + (hf)^2 - 2h^2 ff'\cos\theta \qquad (14.5)$$

Recoil
$$P = mv/(1 - v^2/c^2)^2 \qquad (14.6)$$

Substituting for P from equation 14.6 with the LHS of 14.5.

$$\frac{m^2v^2c^2}{1 - v^2/c^2} = (hf')^2 + (hf)^2 - 2h^2 ff'\cos\theta \qquad (14.7)$$

Comparing equation 14.4 and 14.7 and simplifying, we get

$$mc^2(f - f') = hff'(1 - \cos\theta) \qquad (14.8)$$

If this is written in terms of wavelength where $\lambda = c/f$ and $\lambda' = c/f'$, we finally obtain

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta) \qquad (14.9)$$

Equation 14.9 is the derivation of Compton effect earlier written in equation 14.3.

14.3 Wave-Particle Duality

The wave theory of light is successful in explaining the variation of the velocity of light as well as interference and diffraction effects. These effects give a picture in which there are alternate regions of

Describe a typical traditional marriage

brightness (maximum) and darkness (minimum). The particle (corpuscular) theory of light could not explain these patterns.

However, the wave theory had its failures, the most notable being in the explanation of the photoelectric effect. As we saw above, this theory could not interpret the observed effects. It was left to Einstein to adopt the photon (particle) theory introduced earlier by Planck inexplaining the features of the photoelectric effect. It has become a fundamental part of present-day physical theory that the two characters (wave and particle) are inherent in the nature of light. It never exhibits both characteristics in any one experiment: it behaves either as a wave or as a particle. This is the so-called wave-particle duality.

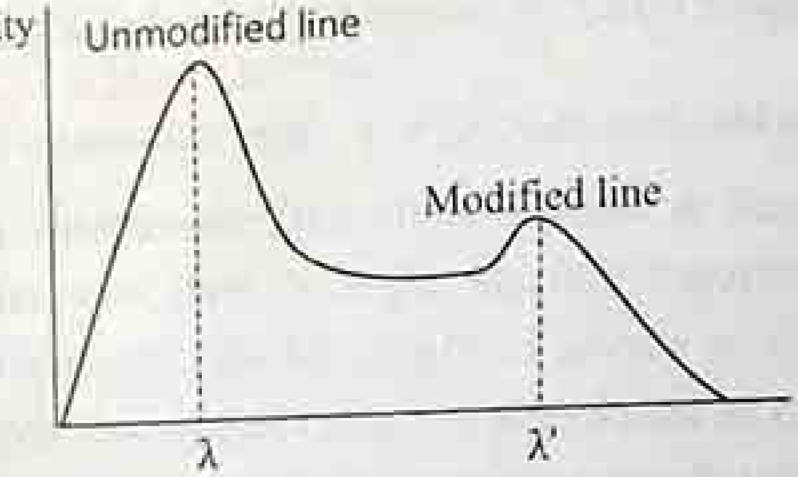


Fig. 14.4: Intensity against wavelength of scattered x-rays with medium

According to the De Broglie hypothesis (1924), this dual character applies not only to radiation but to all fundamental entities of physics.On this hypothesis, electrons, protons neutrons, atoms and molecules should have some type of wave motion associated with them. The wavelength of the motion is given by

$$\lambda = \frac{h}{p} \tag{14.10}$$

where p is the momentum of the particle. This relationship has been confirmed experimentally for material particles as described below.

Interference and diffraction patterns provide conclusive tests for wave character. Davisson and Germer (1927) and Thomson (1928) showed that elections could be diffracted by crystalline materials. This diffraction pattern produced by Thomson was most striking. In the experiment, a beam of high speed cathode rays produced in a discharge tube operated at between 10 and $60eV$ was passed through a thin metal foil. After passing through the foil made of eithergold, aluminum, copper, platinum, lead or iron, the cathode rays were received on a photographic plate (Figure 14.5).

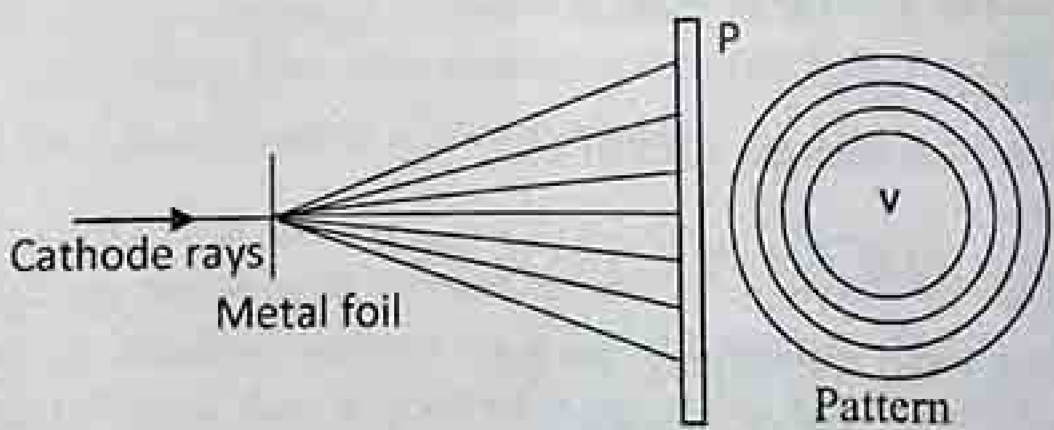


Fig.14.5: Schematic diagram for electron diffraction experiment (Thomson, G. P., 1928).

The pattern on the photographic plate consisted of a series of well-defined concentric rings about a central spot. The metal foil consists of a large number of randomly oriented crystallites. The observed pattern is similar to the X-ray power diffraction pattern. By measuring the diameter of the rings and using known values of the grating space of the crystal, the wavelength associated with the electron can be estimated using the well-known Bragg equation. The significant point is that the wavelength determined this way agrees with the wavelength computed using the equation 14.10 thereby confirming that other fundamental particles e.g. neutrons exhibit this phenomenon. Therefore, wave-particle duality is a fundamental characteristic of all atomic and nuclear species. It is interesting to note that while G.P. Thomson demonstrated the wave nature of the electron, the discovery of the electron by his father J.J. Thomson in 1897 amounted to a demonstration of the particle character of the electron. Each of them received the Noble prize for their individual discovery.

Example 14.2

Determine the wavelength associated with $216eV$ electrons.

Solution

The energy of the electron is small compared with the rest energy $0.511MeV$ of the electron, therefore, we ignore relativistic correction. The kinetic energy $T = \frac{1}{2}mv^2$. Therefore, momentum is

$$p = mv = (2mT)^{1/2}$$

$$= (2 \times 9.1 \times 10^{-31} \times 216 \times 1.6 \times 10^{-19})^{1/2} = 8.0 \times 10^{-24} kgms^{-1}$$

The wavelength is $= \frac{h}{p} = \frac{6.63 \times 10^{-34} Js}{8.0 \times 10^{-24} kgms^{-1}} = 8.3 \times 10^{-11} m$

14.4 Uncertainty Principle

Heisenberg observed that the very act of measuring physical parameters like position and momentum of an electron disturbed the electron because of interaction between the apparatus and the electron. This invariably introduced uncertainties in the precision of measurement. These uncertainties concern the nature of matter and are not related to errors introduced by the limited precision of the measuring device. In the macroscopic world the uncertainties are not noticeable, but in the microscopic world they are significant.

According to the Heisenberg uncertainty principle, it is impossible to know (or measure) precisely and simultaneously both the momentum and position of a particle. Mathematically, the principle is represented as

$$\Delta x \cdot \Delta p \approx \frac{h}{2\pi} = \hbar \qquad (14.11)$$

where Δx is uncertainty in position and is Δp uncertainty in momentum. The relation indicates that, the more precisely position is known i.e Δx small, the less accurately momentum will be known. Another representation of the uncertainty principle is $\Delta E \cdot \Delta t \sim \hbar$ where ΔE is uncertainty in energy and Δt uncertainty in time.

One of the areas in which the ideas inherent in the uncertainty principle have been employed is in tackling the vexed problem of atomic stability. Recall that we mentioned the use of Schrodinger's wave mechanics to explain the concept of the stationary atomic orbit. However, one cannot preclude Coulomb attraction between the nucleus and the atomic electrons. The electrostatic forces pull the electrons as close to the nucleus as possible, but the electrons are compelled to stay out in space over a distance given by the uncertainty principle. Thus, if the electron is confined to too small a space (Δx small or electron close to the nucleus), it would have a great uncertainty in momentum. This means that it would have a high expected energy which it would use to escape from the electrical attraction. The net result is an electrical equilibrium, with the electron not having a fixed orbit (definite r) as such, but different probabilities of orbital radius. For example, for the IS electron of the hydrogen atom, this probability is highest for $r = 0.053nm$ which corresponds to the Bohr orbit radius.

Example 14.3

Find the uncertainty in the position of an electron that has an uncertainty in its speed of $10^5 ms^{-1}$.

Solution

According to the uncertainty principle,

$$\Delta x \cdot \Delta p = h/2\pi; \qquad \Delta p = m\Delta v$$

Therefore,

$$\Delta x = \frac{h}{2\pi m \Delta v} = \frac{6.63 \times 10^{-34} Js}{2\pi \times 9.31 \times 10^{-31} kg \times 10^{5} ms^{-1}} = 1.1 \times 10^{-9} m$$

Summary

1. Planck in 1901 postulated that photons are particles emitted with energy hf where f is frequency.
2. The ejection of electrons from a metal surface when electromagnetic radiation of sufficient frequency is incident on the surface is the photoelectric effect. The photon theory of radiation is required to explain the observed phenomenon.
3. The Einstein equation of the photoelectric effect is given by

   $$\frac{1}{2}mv^2 = hf - \phi$$

   where ϕ is the work function of the metal surface.
4. The Compton effect is another confirmation of the particle nature of radiation. It concerns the scattering of X-rays by matter. The wavelength change is given by the Compton formula

   $$\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta)$$

   where θ is the angle of scattering.
5. Radiant energy in some situations behaves like waves (e.g. indiffraction and interference) and in some situations, has particle characteristics (e.g. photoelectric effect, Compton effect). According to the De Broglie hypothesis, this dual character applies not only to radiation but to all fundamental entities of physics e.g. electrons, protons, neutrons. The wave-length of the motion of such particles is given by $\lambda = h/p$ where p is momentum.
6. The wave property of the electron has been demonstrated in diffraction experiments e.g. G.P. Thomson experiment.
7. Heisenberg showedthrough the Uncertainty Principle thatin quantum physics, there is an inherentlimitation on our ability to make measurements. Thus, the mathematical representation of the Uncertainty Principle is given in the formula $\Delta x \cdot \Delta p = {}^{h}/_{2\pi}$.

Exercises 14

14.1 Which statement is incorrect: According to Planck's and Einstein, e.m. radiation

A. is quantized **B.** The quantized radiation is given by $E = hf$

C. Where h = Planck's constant = 10^{20}s.

D. e.m. radiation is known as photons.

14.2 In Photo electric effect

A. Electrons are emitted from metal surface.

B. Light must have sufficient high frequency.

C. Light should have low frequency.

D. Millikan investigated this effect.

14.3 The main features of Photoelectric effect are:

A. At a given frequency of the incident light, the photoelectric current is proportional to the intensity of incident light.

B. The energy of the emitted electrons is independent of the intensity of the incident light.

C. There is a big time lag between the irradiation of the surface and the ejection of the electrons.

D. For each material, there is a threshold frequency f_0 of the incident light, below which no photoelectrons are emitted

14.4 The wave property of the electron

A. has been demonstrated in diffraction experiments.

B. Has been demonstated in Thomson experiment.

C. Has been demonstrated in sound experiment.

D. Electron can behave as a wave as well as particle.

14.5 The emission of photoelectrons from a metal surface ceases when the wavelength of the incident light exceeds $5.6\times10^{-7}m$. Calculate the maximum kinetic energy (in eV) of electrons emitted from the surface when the incident light has a wavelength of $4.2\times10^{-7}m$

A. $0.74eV$ B. $1.4eV$ C. $3.0eV$ D. $5.0eV$

14.6 Photoelectrons of energy $1.5eV$ are ejected from a tungsten surface. If the threshold wavelength for photoelectric emission in tungsten is $2.3\times10^{-7}m$, determine the wavelength of the incident light.

A. $1.8\times10^{-7}m$ **B.** $3.8\times10^{-7}m$ **C.** $5.8\times10^{-7}m$ **D.** $9.8\times10^{-7}m$

14.7 A beam of x-rays is scattered through 45° by free electrons. If the wavelength of the scattered beam is $0.35\overset{o}{A}$, calculate the wavelength of the incident beam.

A. $0.3429\overset{o}{A}$ **B.** $1.325\overset{o}{A}$ **C.** $4.3401\overset{o}{A}$ **D.** $6.1234\overset{o}{A}$

14.8 An x-ray photon of initial frequency 3×10^{19} sec^{-1} collides with an electron and is scattered through 90°. Find its new frequency in s^{-1}?

A. 9.4×10^{19} **B.** 5.4×10^{19} **C.** 2.4×10^{19} **D.** 0.4×10^{19}

14.9 X-rays of wavelength $1\overset{o}{A}$ are incident on a carbon target. Calculate the maximum wavelength change in the scattered radiation.

A. $0.0485\overset{o}{A}$ B. $1.0546\overset{o}{A}$ C. $3.456\overset{o}{A}$ D. $6.0456\overset{o}{A}$

14.10 Calculate the De Broglie wavelength of an electron which has a velocity of $3\times10^{9}ms^{-1}$?

A. 7.28×10^{11} **B.** 5.28×10^{11} **C.** 3.28×10^{11} **D.** 1.28×10^{11}

14.11 Discuss the main features of the photoelectric effect. How are the features accounted for by the Einstein theory?

14.12 Whatare the expectations of the wave theory when light is incident on a metal surface? How do these expectations conflict with experimental results?

14.13 What do you understand by the Compton effect? Distinguish between the origins of the modified andunmodified lines in the scattering of X-rays by matter.

14.14 Describe an experiment that demonstrates the wave characteristics of particles.

14.15 State the De Broglie hypothesis. Write generally about the experimental confirmation of the hypothesis.

14.16 A student observes that a certain phototube requires $0.25V$ to serve as thestopping potential for light with a wavelength of $5000\overset{o}{A}$, the stopping potential is one volt. Estimate h/e from the data.

14.17 When light of wavelength $4.0\times10^{-7}m$ falls on a sodium surface, the maximum energy of the emitted electrons is $0.8eV$. Determine the work function of the surface.

14.18 The emission of photoelectrons from a metal surface ceases when the wavelength of the incident light exceeds $5.6\times10^{-7}m$. Calculate the maximum kinetic energy (in eV) of electrons emitted from the surface when the incident light has a wavelength of $4.2\times10^{-7}m$.

14.19 Photoelectrons of energy $1.5eV$ are ejected from a tungsten surface. If the threshold wavelength for photoelectric emission in tungsten is $2.3\times10^{-7}m$, determine the wavelength of the incident light.

14.20 A beam of x-rays is scattered through 45° by free electrons. If the wavelength of the scattered beam is $0.35\overset{o}{A}$, calculate the wavelength of the incident beam.

CHAPTER 15
NUCLEAR PHYSICS

15.0 Introduction: The Nucleus

In discussing radiation it is pertinent to refresh briefly our knowledge of the nucleus. Nuclear theory is not fully understood today because the nucleus is made up of minute entities. Consequently, our knowledge of the nucleus depends on indirect observations for which classical mechanics cannot fully explain. However, quantum mechanics can be applied but with some limitations. For example, we need to know the potential energy of a system for a possible solution to the Schrödinger's equation. In addition, if the energy is known then one must have an understanding of the forces that hold the nucleus together. Besides as pointed out earlier the exact form of the nuclear force is unknown but the study of radioactivity and nuclear stability has given us insight about the nature of the nucleus, the energy it possesses and how this energy can be harnessed and released.

15.1 The Nuclear Structure

The conclusions we draw from thermionic and photoelectric effect are;

(a) Atoms of materials contain electrons.

(b) Atoms contain equal number of positive and negative charges since atoms are usually electrically neutral.

(c) The mass of an electron is very small compared with the mass of the atom and it follows that most of the mass of the atom must be associated with the positive charge.

From modern atomic theory, we do know that the mass of the atom is concentrated within a small region of space and surrounded by a cloud of electrons. This proposal is actually the Rutherford's model and the size of the atomic nucleus can be obtained from the expression

$$r_a = \frac{4kZe^2}{mv^2} \qquad (15.1)$$

where e is the electronic charge, Z is the atomic number, the product Ze is the total charge, $k = 1/4\pi\varepsilon_0$ and ε_0 is the permittivity of vacuum. It has been found that r_a is of the order of $10^{-4}\ \overset{o}{A}$.

15.1.1 The Nuclear Force

The forces in the nucleus include the gravitational attractive force between the nucleons (combination of proton and neutron). However, the force is negligible compared to the repulsive electrical force between positively charged protons. If we consider this large force one would have expected that the nucleus would simply fly apart. But we do know that this does not happen since most atoms are very stable. The conclusion to draw is that there is an additional force that holds the nucleus together which is referred to as the *strong nuclear force*. It is also known simply as the nuclear force. It must be pointed out that the exact nature of the force is unknown. However, what we do know is that:

(i) The nuclear force is strongly attractive and much larger than the electrostatic and gravitational forces.

(ii)The nuclear force is short ranged; that is a nucleon interacts only with its nearest neighbours.

(iii) The nuclear force acts between any two nucleons; that is between two protons, a proton and a neutron or two neutrons.

15.1.2 Nuclear Notation

The notation used in describing the nuclei of different atoms is shown in Figure 15.1. In this notation the chemical symbol of the element is used with subscripts and a superscript. The left subscript is the atomic number (*Z*) which is the number of protons in the nucleus. In an atom the atomic number represents the number of electrons in the atom. We recall that the number of protonsis equal to the number of electrons in a neutral atom, which means that, this term is equally applicable to protons. Consequently, we used the term [proton number]. The left superscript A is called the mass number and is equivalent to the number of protons and neutrons in the nucleus. In addition, since $m_p \approx m_n$, it follows that the mass numbers of nuclei give a relative comparison of nuclear masses. For example for carbon $A = 12$ and since $p = 6$ and $n = 6$ the neutron number N is indicated by the subscript on the right hand side. Frequently, the neutron number is omitted since $N = A - Z$ and the symbol becomes ${}^{A}_{Z}C$.

$$^{A}_{Z}X_{N}$$

Fig. 15.1: Nuclear notation

Isotopes are atoms whose nuclei have the same number of protons but different number of neutrons. For instance some isotope of carbon are written as; ${}^{12}_{6}C$, ${}^{13}_{6}C$ *and* ${}^{14}_{6}C$. Isotopes of the same family have the same electronic properties. The isotopes are usually referred to by their mass number for example carbon twelve, carbon thirteen, carbon fourteen, etc. A particular nuclear species or isotope is called a nuclide. For carbon we have six nuclides the remaining ones being, ${}^{11}_{6}C$, ${}^{15}_{6}C$ *and* ${}^{16}_{6}C$. The only exception to this is the hydrogen isotopes which are referred to as ordinary hydrogen for ${}^{1}H_{2}$, deuteron for ${}^{2}H_{2}$ and tritium for ${}^{3}H_{2}$.

15.2 Radioactivity

Radioactivity is the spontaneous emission of radiations from the nuclei of some unstable isotopes. In other words these unstable nuclei disintegrate spontaneously (decay) of their own accord by emitting energetic particles. These types of isotopes are said to be radioactive. The name radio refers to the emitted nuclear radiation. The radiation may be a wave or a particle and a radioactive isotope is said to be "active" in doing so. It must be pointed out that only a small number of unstable nuclide actually exhibit radioactivity naturally and the vast majority of radioactivity are produced artificially.
Radioactivity is a quantum mechanical probability which occurs within the nucleus. There are three different types of radioactivity namely;

(i) Alpha (α) particles which are doubly charged (2^{+}) ions containing two protons and two neutrons. In fact they are identical to the helium atoms (${}^{4}_{2}He$).
(ii) Beta (β) particles which are electrons
(iii) Gamma (γ) particles which are particles or quanta of electromagnetic energy.

15.3 Properties and Detection of Emitted Radiations

Table 15.1 will aid us in the understanding of the properties of the three types of radiations.

Table 15.1 Properties of radioactive emissions

	α -radioactivity	β - radioactivity	γ - radioactivity
Penetrability	small	large	very large
Range in air	few *cm*	tens of *cm*	many metres
Charge	positive	+ or –	no charge
Deflection by electric or magnetic fields	some	Large	none
Ionizations	10,000	1,000	none
Mass	Heavy	Light	none
Velocity	0.05*c*	0.9*c*	*c*
Nature	Helium atom	electrons	e.m. radiation

The energy of radioactive particle is usually expressed in mega electron volts (*MeV*) where $1eV = 1.6 \times 10^{-19} J$ is the energy acquired by an electron when it is accelerated through a potential of $1V$. The various radiations from radioactive sources can be distinguished by passing them through a magnetic field as indicated in Figure 15.2.

15.4 Detectors

Radioactive decay cannot be detected directly by our senses hence an indirect method is used in the detection of nuclear radiation. The detectors used are based on the ionization or excitation of atoms by the passage of energetic particles through matter. In the case of alpha and beta particles their energy of interaction is transferred to atoms along their paths through Coulomb interaction. An indirect way to detect gamma rays is to recall that photons lose energy by Compton scattering or pair production and the energetic particles produced by these interactions cause the ionization in a detector.

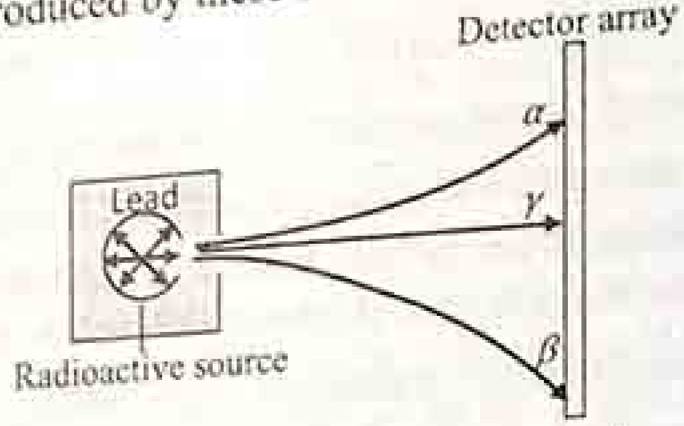


Fig 15.2: Separation of the three particles of radiation from radioactive source.

The Geiger Muller Counter (GM)

The basic principle of a GM counter involves a voltage of between 800 – 1000V which is applied across a wire electrode and a metal tube as shown in Figure 15.3. Inside the tube is pumped a gas such as argon at low pressure. When an ionizing radiation enters the tube through a small opening at one end, it ionizes a few atoms of the gas. The freed electrons are attracted and accelerated toward the positive wire anode. This process continues and multiplies and the result is an "avalanche" discharge which produces a voltage pulse. The pulse is amplified and sent to an electronic counter that counts the pulses or the number of particles detected. If a loudspeaker is installed, an audible click can be heard.

A major disadvantage of the GM counter is that it has a relatively long dead time which is the recovery time required between successive electrical discharges in the tube. This time is about 200μs and it limits the counting rate to a few hundred counts per second. For instance if a few particles are incident on the tube at a faster rate not all of them will be counted.

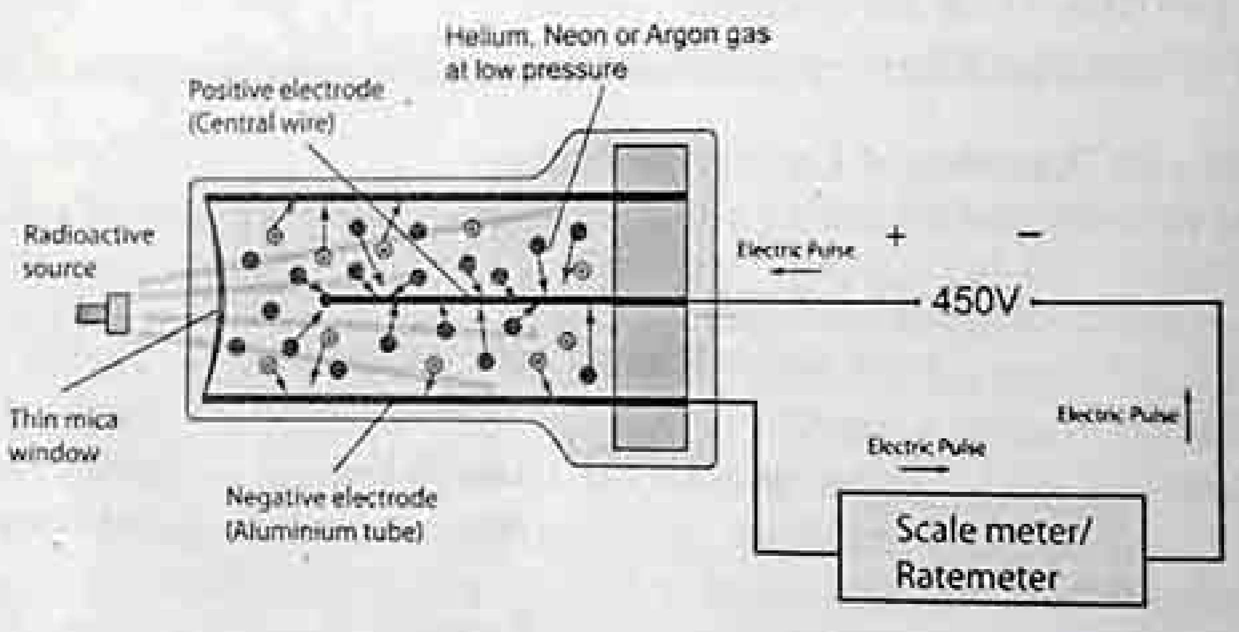


Fig. 15.3: Geiger Muller Counter

Scintillation Counter

These types of counters have much shorter dead time than the GM counters. In a scintillation counter, atoms of a phosphor material are excited by an incident particle and visible light is emitted. The light pulse is then converted to an electrical pulse by a photoelectric material and is further amplified by photomultiplier. The scintillation counter (see Figure 15.4) is made up of a series of successively higher potential electrodes. The photoelectrons are accelerated towards the first electrode and acquire energy to cause several secondary electrons to be emitted when they strike the electrode. This process continues and relatively weak scintillations are converted into sizeable electrical pulses which are converted electronically. These counters have dead time of about a few microseconds. In addition, the magnitude of the photomultiplier current pulse is proportional to the number of photons generated by the incident particle which in turn is proportional to its energy. It follows that the magnitude of the counted pulses gives a measure of the energy of the incident particle. Counters with these features are referred to as proportional counters.

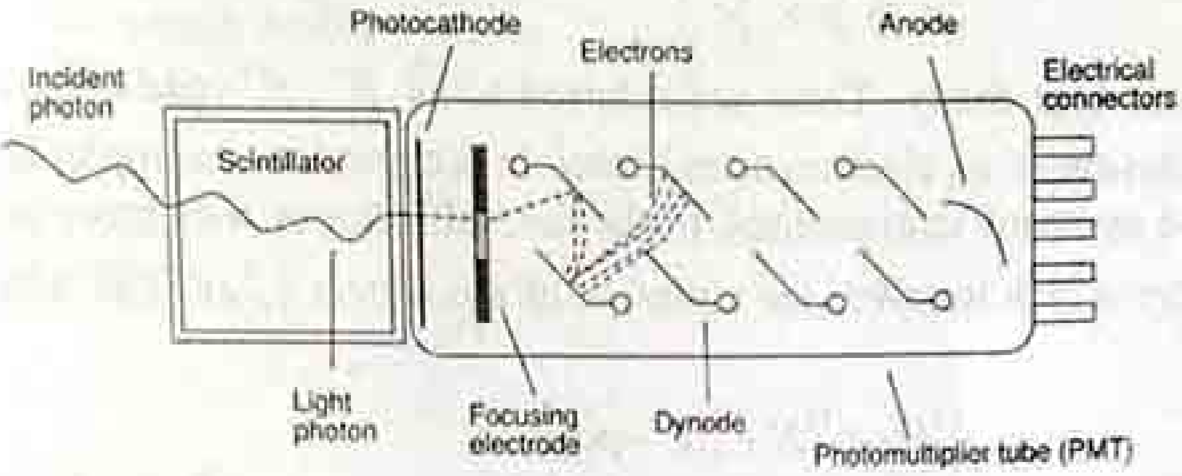


Fig. 15.4: The scintillation counter

Solid State or Semiconductor counters

The semiconductor counters are relatively new devices for radiation counters. They operate on the principle that when charged particles pass through a junction diode electron-hole pairs are produced. When these electron-hole pairs are subjected to an electric field electric signals are generated which may be amplified and counted. Solid state counters are very fast and are capable of very high counting rates. Some detectors actually provide visual records of decaying nuclei.

15.5 Nuclear Changes during Radioactive Decay

In this section we have a detailed look at the various radioactive decay starting with α -decay.

α- Decay

The ejection of a particle from a radioactive nucleus results in the nucleus losing two protons and two neutrons. Consequently, the mass number A is decreased by four, that is, $\Delta A = -4$ and the proton number is decreased by 2, that is, $\Delta Z = -2$. The conclusion we draw is that since the parent loses two protons, the resulting daughter must be the nucleus of another element. Note the original and resulting nuclei are frequently referred to as *parent* and *daughter* nuclei respectively.

In summary the decay process is a nuclear transmutation in which the heavier element changes into nuclei of a lighter element. The following process illustrates this.

$$ {}^{214}_{84}Po \rightarrow {}^{210}_{82}Pb + {}^{4}_{2}He \tag{15.2}$$

Polonium $\rightarrow$ lead + α –particle

From this reaction we observe that two conservation laws must be obeyed namely;

(i) Conservation of nucleons (A), and
(ii) Conservation of charge (p)

Example 15.1

Suppose ${}^{238}_{92}U$ nucleus undergoes α -decay, what will be the resulting nucleus?

Solution

For α -decay we have $\Delta Z = -2$ which means we have a loss of 2 protons by the uranium atom and as a consequence the daughter nucleus will have a proton number of $92 - 2 = 90$. This corresponds to the proton number for thorium. The equation for this process is

$$ {}^{238}_{92}U \rightarrow {}^{234}_{82}Po + {}^{4}_{2}He \tag{15.3}$$

β –Decay

In this type of decay an electron is created in the nucleus. As an example we look at the decay of carbon 14 which takes the form

$$ {}^{14}_{6}C \rightarrow {}^{14}_{7}N + {}^{-1}_{0}e \text{-beta particle (electron)} \tag{15.4}$$

Carbon $\rightarrow$ nitrogen

Under beta decay we notice that the neutron number of the parent nucleus decreases by one $(\Delta N = -1)$ while the proton number of the daughter nucleus increases by one $(\Delta Z = +1)$. However, the mass number is unchanged and the conclusion we draw from this is that the nucleus decays into a proton and an electron as follows;

$${}_0^1n \rightarrow {}_1^1p + {}_{-1}^{0}e \quad (15.5)$$

There are three types of beta decay. These are referred to as β^-, β^+ and electron capture. The β^- decay involves the emission of an electron as indicated in the previous example. Isotopes that decay in this manner have more neutrons than protons. In the case of isotopes with more protons than neutrons they decay by β^+ decay which involves the emission of a positron $({}_{+1}^{0}e)$. The following example will illustrate this

$${}_8^{15}O \rightarrow {}_7^{15}N + {}_{+1}^{0}e \text{ - positron} \quad (15.6)$$

In this type of decay a proton disintegrates into a neutron and a positron $({}_1^1p \rightarrow {}_0^1n + {}_{+1}^{0}e)$.
In the case of electron capture an atomic electron is absorbed by the nucleus for example

$${}_{-1}^{0}e \rightarrow {}_4^7Be \rightarrow {}_3^7Li \quad (15.7)$$

That is, electron + beryllium gives lithium

γ-decay
Gamma decay results when a nucleus is in an excited state after a previous decay or is due to an energetic collision with another particle. The gamma decay equation takes the form as shown in equation

$${}_{28}^{61}Ni^* \rightarrow {}_{28}^{61}Ni + \gamma \text{ - gamma ray} \quad (15.8)$$

Note that the asterisk on the ${}^{61}Ni$ nucleus indicates that it is an excited state. This means that the daughter nucleus is simply the parent nucleus with less energy.

15.6 Rate of Radioactive decay

An important factor to consider in radioactivity is the activity of a radioactive material which is defined as the number (ΔN) of disintegrations or decays per time $\Delta N/\Delta t$. For a given amount of material it is expected that the activity will decrease with time since fewer nuclei are now available after a time has elapsed. In addition, the activity is proportional to the number of nuclei (N) present at any particular time that is, $\Delta N/\Delta t \propto N$. If we expressed this in an equation we have

$$\frac{\Delta N}{\Delta t} = -\lambda N \quad (15.9)$$

If there are N_0 as the starting nuclei at $t = 0$ and we have N remaining at a later time t and on integrating equation 15.9, we have

$$\int_{N_o}^{N} \frac{dN}{dt} = -\lambda \int_0^t dt \text{ or } (\ln N)_{N_o}^{N} = -\lambda t$$

where $\ln = \log_e$

$\therefore$ $$\ln(N - N_0) = \ln(N/N_0) = -\lambda t$$

Hence, $$N = N_0 e^{-\lambda t} \quad (15.10)$$

where N_0 is the initial number of nuclei and λ is the decay constant. Frequently, the decay rate of an isotope is usually expressed in terms of its half-life.

Half-Life ($t_{1/2}$)

The half-life is the time it takes for half of the original nuclei in a radioactive sample to decay. What this means is that $N = N_0/2$ as shown in Figure 15.5.

From experimental observations it is found that the half-lives of radioactive isotopes vary greatly for example compare the have lives of some isotopes indicated in Table 15.2

It can easily be shown that a simple relationship exists between the disintegration constant and the half-life by recalling that the number of undecayed parent nuclei at a particular time is $N = N_0 e^{-\lambda t}$.

But at $t = t_{1/2}$, $N = \frac{N_0}{2}$ which makes $\frac{N}{N_0} = e^{-\lambda t_{1/2}} = \frac{1}{2}$

Since $e^{-0.693} = \frac{1}{2} = e^{-\lambda t_{1/2}}$

It follows that by comparison

$$t_{1/2} = \frac{0.693}{\lambda} \tag{15.11}$$

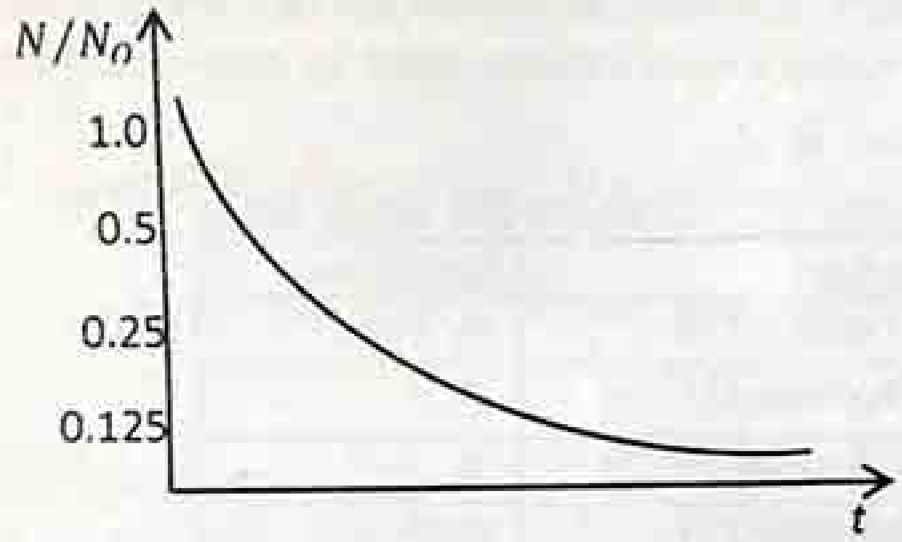


Fig. 15.5: Radioactive decay versus time

Table 15.2: Half-lives of some radioactive isotopes

Name of Isotope	Half-Life ($t_{1/2}$)
Be (α)	$1\times10^{-16} s$
Iodine (β^-)	8days
Oxygen (β^-)	27 s
Bismuth 214	$1.6\times10^{-4} s$
Helium 5	$6\times10^{-20} s$
Co (β^-)	5.3years
Platinum (α)	4.4×10^{4} years
Carbon 14	5530years
Uranium ($^{238}_{92}U$)	4.5×10^{9} years
Actinium	7.04×10^{8} years
Thorium ($^{232}_{90}Th$)	1.41×10^{10} years
Neptunium ($^{237}_{93}Np$)	2.14×10^{6} years
Plutonium -239	2.4×10^{4} years

15.6.1 Radioactive Decay Series

It is found that some isotopes that result from radioactive decay may itself be radioactive and can decay. If the others following it are radioactive the sequence of decays are referred to as radioactive decay series. In addition, each isotope in the series is a daughter product of the isotope above it in the series. For example, uranium 238 decays to form a series that ends with the stable lead -206 as shown in Table 15.3.

Under equilibrium, it can be shown that for a radioactive series;

$$\lambda_1 N_1 = \lambda_2 N_2 = \lambda_3 N_3 = \cdots \tag{15.12}$$

Unit of Radioactivity

The strength of a radioactive sample may be specified at a given time by its activity and the unit commonly employed is the curie (Ci) and is defined as:

$$1Ci = 3.70 \times 10^{10} \; decay/s$$

This unit is based on the activity of one gram of radium. However the SI unit is the Becquerel (Bq) which is defined simply as $1Bq = 1 \, decay/s$,

hence $1Ci = 3.70 \times 10^{10} \; Bq$

As a remark, we note that the curie is a relatively large unit and sub units are frequently used such as the milli- and the micro- (mCi and μCi). It must be pointed out that lower units may be employed for example the radon limit which is a lung cancer causing agent by most countries is $4 \, pCi/litre$ of air. This is equivalent to $4 \times 10^{-12} \, Ci/litre$ of air.

Table 15.3: Uranium-238 decay series

Isotope	Emissions
Uranium-238	α, γ
Thorium-234	β, γ
Protactinium-234	β, γ
Uranium-234	α, γ
Thorium-230	α, γ
Radium-226	α, γ
Radon-222	α
Polonium-218	α
Lead-214	β, γ
Bismult-214	β, γ
Polonium-214	α
Lead-210	β, γ
Bismult-210	β
Polonium-210	β
Lead-206	α stable

Example 15.2

The iodine used in medical test for thyroid disease has a half-life of 8 days. Assume that after some time the iodine 131 that accumulates in the patient's thyroid gland is 4×10^{10} nuclei; determine, (a) the observed activity in 24 hours and (b) the number of ^{131}I nuclei remaining at this time.

Solution

(a) Given $t_{1/2} = 8days = 8 \times 60 \times 60 \times 24 = 6.9 \times 10^5 s$; $N_0 = 4.0 \times 10^{22}$ nuclei; $\Delta t = 24hr = 8.64 \times 10^4 s$. We are required to determine $\Delta N/\Delta t$.

From equation 15.11, $t_{1/2} = \dfrac{0.693}{\lambda}$

hence $\lambda = 0.693/6.9 \times 10^5 = 1 \times 10^{-6} s$.

But $\dfrac{\Delta N}{\Delta t} = -\lambda N_0 = -1 \times 10^{-6} \times (4 \times 10^{22}) = -4 \times 10^{16} \; decay/s$.

The negative sign indicates that the activity is decreasing.

(b) The actual number of parent nuclei present can be found from the equation $N = N_0 e^{-\lambda t}$ and with $t = 1$ day and $\lambda = 0.693/8days$, then from

$N = N_0 e^{-\lambda t}$ we have, $N = 4\times10^{22}\times0.917 = 3.7\times10^{22}$ nuclei.

Example 15.3

Suppose an archaeologist discovered a bone fragment which registered a count rate of 20 counts per minute while a similar fresh bone gave a count rate of 25 counts per minute. Determine the age of the specimen.

Solution

The activity is proportional to the number of radioactive atoms within it, hence

$N = N_0 e^{-\lambda t}$ or $20 = 25e^{-\lambda t}$

On taking the natural log of both sides to we get $\lambda t = 0.223$ or $t = \dfrac{0.223}{\lambda}$

From equation 15.11, we know that $t_{1/2} = \dfrac{0.693}{\lambda}$ and from the table, the half-life of carbon is 5530 years

$$\therefore \qquad t = \frac{0.223\times5530}{0.693} = 1{,}779.5\ years.$$

Example 15.4

Determine the nucleus that would be formed if $^{238}_{92}U$ undergoes (i) α -decay (ii) β -decay and (iii) γ -decay.

Solution

(i) We recall that in any of these decay processes, the charge number, linear momentum, angular momentum, nucleon number and of course mass/energy must be conserved.

$\therefore$ The reaction for the α-decay takes the form $^{238}_{92}U \to {}^{A}_{Z}X + {}^{4}_{2}He$.

If the nucleon number is to be conserved then 238 = A + 4 or A = 234 and for conservation of charge we must have 92 + Z = 2 or Z = 90. A look at the chemical table shows that the daughter nucleus formed is $^{234}_{90}Th$.

(ii) For beta decay $^{238}_{92}U \to {}^{A}_{Z}X + {}^{0}_{-1}e$ and for nucleon conservation we must have 238 = A + 0 and for charge conservation 92 = Z + (-1) or Z = 93. Again from the periodic table this corresponds to $^{238}_{93}Np$.

(iii) The reaction for γ decay is $^{238}_{92}U \to {}^{A}_{Z}X + \gamma$

Since the γ ray does not have a nucleon number and carries no charge the daughter nucleus is given simply as $^{238}_{92}U$. This is just the same nucleus but with a reduced energy.

15.7 Applications of Radioactivity

From previous discussions it is found that with the knowledge of the half-life of a radioactive isotope it can be used to estimate how much a given amount of material will exist in future. With this knowledge scientists are able to project backward in time to determine the ages of objects containing radioactive isotopes. The principle hinges on the fact that living things or things that once lived contained a known amount of radioactive ^{14}C. If a comparison of the concentration of carbon 14 in dead matter relative to that in the living matter is made then the age of the dead matter can be ascertained.

Other areas where radioactivity has been used includes:

(i) dating geological specimen using uranium.
(ii) dating archaeological specimen using carbon 14.
(iii) thickness measurements by back-scattered beta radiation.
(iv) treatment of tumors.

(v) sterilization of food.
(vi) nuclear pacemakers for the heart.
(vii) liquid flow measurement.
(viii) tracing sewage or silt in the sea or rivers.
(ix) checking blood circulation and blood volume.
(x) atomic light using krypton 85.
(xi) checking the silver contents in coins
(xii) radiographs of castings and teeth.
(xiii) testing for leaks in pipes.
(xiv) tracing phosphate fertilizers using phosphorous 32.
(xv) sterilization of insects for pest control.

Example 15.5
A piece of bone from an archaeological site is found to give a count rate 15 counts/min. A similar sample of a fresh bone gives a count rate of 19 counts/min. Calculate the age of the specimen.

Solution
Recall that $N = N_0 e^{-\lambda t}$

$\therefore \quad 15 = 19e^{-\lambda t}$ or $t = \dfrac{0.236}{\lambda}$

From equation 15.11, $t_{1/2} = \dfrac{0.693}{\lambda}$

and for carbon 14 the half-life is 5530 years hence

$$t = \frac{0.236 \times t_{1/2}}{0.693} = 1883.2\,years$$

15.7.1 Medical Application

Nuclear radiation is used beneficially in the diagnosis and treatment of some diseases but we have to bear in mind that it can be very harmful if it is not properly handled and administered. Just note that nuclear radiation just like X-rays can penetrate human tissues without pain or any other sensation. However, it is important to know that large doses or repeated small doses can lead to skin burns, lesions and other conditions. The main hazard of radiation is damage to living cells as a result of ionization. Several effects can take place for instance ions particularly complex ions or radicals resulting from ionizing radiation can change the structure of a cell so that it can no longer perform its normal functions. In addition, the cell can also die. This in itself may not pose too serious a problem since similar cells would reproduce to make more cells. But the snag is that if large radiation doses damage enough cells, cell reproduction might not be fast enough.

Another problem that may arise is that damaged cells may reproduce cells that are different and defective and in most cases the reproduction occurs in an uncontrolled manner. The unregulated and fast production of abnormal cells is known as cancer. Benign tumour is the result when cancer cells grow slowly in numbers with little effect on the surrounding normal cells. However, when they grow at the expense of surrounding cells, the end result is referred to as malignant tumour. Skin cancer and leukemia (blood cancer) are associated with excessive radiation exposure.
An advantage of radiation is that even though it can cause cancer it can also be used to treat cancer. In essence cell damage is used to control the growth of cancer cells. For health care professionals when carrying out radiation therapy, they need to consider radiation safety; that is the amount of dose of radiation. Terms usually employed to describe this include exposure, absorbed dose and equivalent dose. One of the earliest units based on the exposure is the *roentgen* (R) which is defined in terms of the ionization produced in air. One roentgen is the quantity of X-ray or gamma rays required to produce an ionization charge of $2.58 \times 10^{-4}\ C/kg$ of air. This unit has virtually been replaced by the

rad (radiation absorbed dose). One rad is equivalent to an absorbed dose of $10^{-2} J$ of energy per kg in any absorbing material. The SI unit for absorbed dose is the gray (*Gy*) where $1Gy = 1J/kg = 100\,rad$.

In order to measure the biological damage, units like the rem is used to measure the effective dose. The rem stands for roentgen or rad equivalent man. The *relative biological effectiveness* (RBE) or quality factor (QF) is used to measure the different degrees of effectiveness of different particles. The relationship can be expressed as
Effective dose (in *rem*) = dose (in *rad*) × RBE.
The SI unit for effective dose is the Sievert (*Sv*) which means that,
Effective dose (in *S*) = dose (in *Gy*) × RBE
It also means that 1 *Sv* = 100 *rem*.
It is important to note that for humans, the standard permissible is 5 rem/year after the age of 18 with no more than 3 *rem* in a 3-month period.
In the medical field, nuclear radiation is used extensively such as positron emission tomography (PET), nuclear magnetic resonance (NMR), single-photon emission tomography (SPET), etc.

15.7.2 Domestic and Industrial Application
The major application of radioactivity in the home is the smoke detector. They are used as tracers in the industry to detect leaks in pipes, flow rate and to study corrosion and decay.

15.8 Nuclear Stability and Binding Energy
We do have stable isotopes for elements with proton number from 1 – 83 except for those with $Z = 43$ (technetium) and $Z = 61$ (promethium). Nuclear stability is arrived at by carrying out qualitative and quantitative guesses. Hopefully, this will give us some general rules for determining whether or not a particular nuclide would be stable or unstable.

Nuclear Population
It is easier to consider the reactive number of protons and neutrons in stable nuclei. Somehow nuclear stability is related in some way to dominance of the repulsive Coulomb force between protons or the attractive nuclear force between nucleons and this dominance depends on the relative numbers or ratio of protons and neutrons. For nuclei to be stable the mass numbers must be low. In other words A must be less than 40 ($A < 40$) and the number of neutrons must be or nearly equal to the number of protons for example ${}^{4}_{2}He$, ${}^{12}_{6}C$, ${}^{23}_{11}Na$ and ${}^{23}_{13}Na$. For higher mass number $A > 40$, the number of neutrons outnumber the protons. In addition, it is observed that the heavier the nuclei the more the neutrons exceed the number of protons. The conclusion here is that the extra neutrons of such heavy stable nuclei presumable allow the attractive forces among the nucleons to be greater than the repulsive forces between the numerous protons. When nuclei decay, there is a readjustment of proton and neutron numbers of an unstable isotope until a stable isotope is achieved.

Pairing Effects
A critical observation will reveal that many stable nuclei have even numbers of protons and neutrons and very few have odd numbers of both protons and neutrons as illustrated in Table 15.4.

Table 15.4: Pairing effects of stable nuclei

Proton No	Neutron No	Number of stable nuclei
Even	Even	168
Even	Odd	107
Odd	Even	107
Odd	Odd	4

This pairing effect gives a qualitative criterion for stability. Generally, the criteria for nuclear stability may be summarized as follows;

1. All isotopes with proton numbers greater 83 ($Z > 83$) are unstable.

2. (i) Most even-even nuclei are stable.
(ii)Many odd-even or even-odd nuclei are stable.
(iii) Only four odd-odd nuclei are stable namely; ${}^{2}_{1}H$, ${}^{6}_{3}Li$, ${}^{10}_{5}B$ and ${}^{14}_{7}N$.
3. (i) Stable nuclei with mass numbers less than 40 ($A < 40$) have approximately the same number of protons and neutrons.
(ii) Stable nuclei with mass numbers greater than 40 ($A > 40$) have more neutrons than protons.

Example 15.6

Discuss whether the sulphur isotope ${}^{38}_{16}S$ is likely to be stable.

Solution

On applying the general criteria we have;
Since $Z = 16$ the criterion for stability is satisfied.
It is satisfied to be stable since it is an even-even nucleus. However, with $Z = 16$ and $N = 22$, it does not satisfy the requirement that Z and N be approximately equal. Nevertheless, the isotope is likely to be stable. The nucleus is actually unstable due to beta decay.

Binding Energy (E_b)

The stability of the nucleus may be understood by measuring the nuclear binding energy of the nucleus. As a result of the fact that the masses we encounter are very small, another standard is used and is referred to as the unified *Atomic Mass Unit* (u). This has the exact value of 12.000000u for which it corresponds to a neutral atom of carbon-12. Consequently,

$$1u = 1.6606 \times 10^{-27}\,kg$$

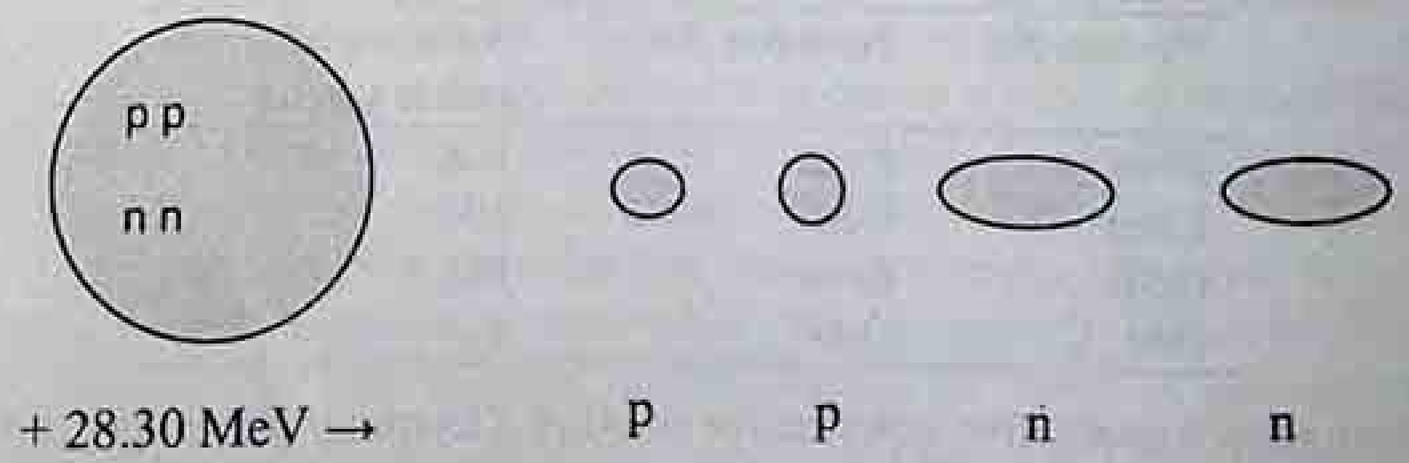
The atomic mass units of various atomic particles are indicated in Table 15.5 with their corresponding energies ($E = mc^2$).

Binding energy is defined as *the energy that would be required to separate the constituent nucleons into free particles.* A further insight into the nature of nuclear force is to consider the average binding energy per nucleon for stable nuclei. This is given by the total binding energy divided by its number of nucleons or E_b / A where A is the mass number. A qualitative way of looking at binding energy is to look at the helium atom with an average binding energy,

$$E_b / A = \frac{28.30\,MeV}{4}$$

Table 15.5: Particle Masses and Energy Equivalent Masses

Particle	u	kg	Energy (MeV)
	1	1.6606×10^{-27}	931.50
Electron	0.000548	9.109×10^{-31}	0.511
Proton	1.007276	1.67265×10^{-27}	938.79
Hydrogen atom	1.008665	1.67356×10^{-27}	938.79
Neutron	1.008665	1.67500×10^{-27}	939.57



The picture above indicates that $28.30MeV$ of work is required to separate a helium nucleus into free protons and neutrons. Conversely, if two protons and two neutrons could be combined to form a helium nucleus $28.30MeV$ of energy would be given up and this amounts to the binding energy. From experimental observation, it is found that for most nuclei $E_b \propto A$.

15.9 Nuclear Reaction

When nuclear reaction takes place, we find that the nuclei of isotopes are converted into the nuclei of other isotopes which are in general are isotopes of completely different elements. This is unlike in chemical reactions in which substances react with each other to form compounds. From studies of the nucleus, it is now possible to induce nuclear reactions by bombarding the nuclei with energetic particles. The first recorded induced nuclear reaction was carried out by Lord Rutherford in 1919. In his experiment, nitrogen was bombarded with α-particles resulting in the transmutation of the nitrogen into oxygen and the ejection of a proton. This reaction corresponds to the chemical equation shown in equation 15.13.

$$^{14}_{7}N + ^{4}_{2}He \rightarrow ^{17}_{8}O + ^{1}_{1}H \qquad (15.13)$$

Nuclear reactions may not be as straight forward as equation 15.13 since the reaction could have been written as

$$^{14}_{7}N + ^{4}_{2}He \rightarrow ^{18}_{9}F^{*} \rightarrow ^{17}_{8}O + ^{1}_{1}H \qquad (15.14)$$

The star in the fluorine means that it was formed in an excited state. The compound nucleus formed rids itself of the excess energy by ejecting a particle. This is what alchemist of old dreamt off when they tried to change common metal into gold. Nevertheless, it is now possible to convert one nuclei into another. Nowadays huge machines called particle accelerators use electric and magnetic fields to accelerate charged particles to very high energies. When these accelerated particles collide with the target nuclei, they initiate nuclear reactions. However, different reactions require different particle energies as illustrated in the following reaction.

$$^{200}_{80}Hg + ^{1}_{1}H \rightarrow ^{197}_{79}Au + ^{4}_{2}He \qquad (15.15)$$

199.968321u 1.007825u 196.96665u 4.002603u

In equation 15.15 even though gold was obtained in the reaction, it is however prohibitive because making such small amounts of gold in an accelerator would be worthless.

Nevertheless, a close scrutiny shows that the reactions have the general form

$$A + a \rightarrow B + b \qquad (15.16)$$

where the upper case letters represent the nuclei and the lower case represent the particles. These reactions are often written in a shorthand notation as $A(a,b)B$ for example $^{14}N(\alpha, p)^{17}O$ and $^{200}Hg(p,\alpha)^{197}Au$ stand for the two reactions discussed earlier. Let us recall that the periodic table has 105 elements but only 90 occur naturally on earth. Elements with proton numbers greater than uranium ($Z = 92$) including technetium ($Z = 43$) and promethium ($Z = 62$) are created artificially by nuclear reactions.

Conservation of Mass-Energy

In nuclear reactions, the total relativistic energy $(E = T + m_0c^2)$ is always conserved. For instance in the reaction $^{14}N(\alpha, p)^{17}O$, the energy conservation requires that

$$T_N + m_Nc^2 + T_\alpha + m_\alpha c^2 = T_0 + m_0c^2 + T_p + m_pc^2$$

where the subscripts refer to the energy and rest mass of a particular particle. On rearranging we obtain:

$$T_0 + T_p - T_N - T_\alpha = (m_N + m_\alpha - m_0 - m_p)c^2 \qquad (15.17)$$

The Q – value

In equation 15.18, the Q-value of the reaction is defined as

$$Q = T_0 + T_p - T_N - T_\alpha \qquad (15.18)$$

The Q-value is actually a measure of the total energy released or absorbed in a reaction. In addition it is the kinetic energy of the products of the reaction minus the kinetic energy of the reactants. This can also be expressed as $T_f - T_i$ where f and i stand for final and initial respectively. In equation 15.17 we find that

$$Q = (M_N + M_a - M_0 - M_p)c^2 \tag{15.19}$$

and in terms of the general reaction of the form $A(a,b)B$ equation 15.19 can be written as

$$Q = (M_N + M_a - M_0 - M_p)c^2 = (\Delta M)c^2 \tag{15.20}$$

From equation 15.20, the Q-value is just the difference in the rest energies of the reactants and the products of the reaction. What this means is that the mass is generally converted into energy and vice versa. The difference in mass (ΔM) is known as the *mass defect*. In fact $E_b = (\Delta M)c^2$.

Remarks

1. If $Q < 0$, the reaction is said to be endoergic or endothermic which means energy was absorbed. However, if $Q > 0$ energy is released and the reaction is said to be exoergic or exothermic.
2. $Q = \Delta M\,(931.5\,MeV/u)$.
3. In radioactive decay with $Q > 0$, the energy is sometimes called disintegration energy.

In summary if $Q > 0$, the effect is exoergic in which mass is converted into energy. This means the mass of the reactants is greater than the mass of the products. On the other hand if $Q < 0$, we have endoergic reaction in which energy is converted into mass and the mass of the products is greater than the mass of the reactants.

By definition, the minimum kinetic energy that an incident particle needs to have to initiate a reaction is called the threshold energy and it can be shown to be

$$T_{\min} = \left(1 + \frac{m_a}{m_A}\right)|Q| \tag{15.21}$$

where m_a is the mass of the incident particle and M_A is the mass of the stationary target nucleus.

Example 15.7

Calculate the threshold energy for the reaction $^{14}N(\alpha,p)^{17}O$, if $m_a = m_\alpha = 4.0002603\,u$, $M_A = M_N = 14.003074\,u$ and $Q = -1.193\,MeV$.

Solution

From equation 15.21, $T_{\min} = \left(1 + \frac{m_a}{m_A}\right)|Q| = \left(1 + \frac{4.0023074\,u}{14.00307\,u}\right) \times 1.193$,

Hence, $T_{\min} = 1.53\,meV$

Reaction Cross Section

For a reaction to take place, there is a minimum or threshold energy a particle must have to initiate a particular reaction. If however, a particle has more kinetic energy than the threshold energies, several possible reactions may occur. A measure of the probability of that a particular reaction will occur is called the cross section of the reaction. The cross section is determined experimentally and it is found that it actually varies with the kinetic energy of the incident particle. In addition, they may vary as $1/v$ where v is the speed of the neutron. In other words, the probability of a reaction appears to be proportional to the time a neutron spends in the vicinity of a nucleus ($t \propto 1/v$).

15.10 Nuclear Energy: Fission and Fusion

15.10.1 Fission

Fission is the process of splitting a heavy nucleus into two lighter nuclei with the emission of two or more neutrons. In addition, energy is emitted in the process which is carried off primarily by the neutrons and fission fragments. Some heavy nuclei do undergo spontaneous fission but at very slow rates. However, fission may be induced and this is the important process in energy production. For example the capture of a neutron by ${}^{235}_{92}U$ results in the formation of an excited uranium-236 nucleus as indicated in equation 15.22.

$$ {}^{235}_{92}U+{}^{1}_{0}n \rightarrow {}^{236}_{92}U^{*} \rightarrow {}^{140}_{54}Xe+{}^{94}_{38}Sr+2({}^{1}_{0}n) \qquad (15.22)$$

In this reaction the unstable nucleus undergoes violent oscillations and then breaks apart and emitting two or more neutrons. Sometimes some isotopes may capture neutrons without fissioning. Nevertheless, there are a variety of ways uranium 235 can fission such as equation 15.23(a) and (b).

$$ {}^{235}_{92}U+{}^{1}_{0}n \rightarrow {}^{141}_{56}Ba + {}^{92}_{36}Kr+3({}^{1}_{0}n) \qquad (15.23a)$$

$$ {}^{235}_{92}U+{}^{1}_{0}n \rightarrow {}^{150}_{60}Nd + {}^{81}_{32}Ge+5({}^{1}_{0}n) \qquad (15.23b)$$

The probability of a fission reaction for a fissionable isotope depends on the energy of the incident neutrons since only a certain nuclei undergo fission. The energy released in an exoergic fission reaction can be estimated from E_b/A curve for stable nuclei. For example when a uranium nucleus splits up energy of about $234MeV$ may be released. In relative terms this may seem to be a lot of energy. On the other hand this might not seem like much energy since $200MeV$ is only about $3\times10^{11}J$. If we compare this to the energy of about $5J$ one expends in picking up a textbook from a desk it is obvious that this is very small. However, we need to remind ourselves that there are billions of nuclei in a small sample of fissionable material. Therefore to produce practical amounts of energy we simply need to have enough fissioning materials. This is achieved by means of a chain reaction. Take for instance in ${}^{235}U$ chain splitting two neutrons are released. Ideally the neutrons may initiate two more fission reactions resulting in the availability of four neutrons. These other neutrons will then initiate more reactions and the process multiples. When this occurs uniformly with time, what we have is an exponential growth.

In order to have a sustained chain reaction, there must be an adequate quantity of fissionable material. For this reason we call the minimum mass required to produce a chain reaction is called *critical mass*. Several factors determine the critical mass. One important consideration is the amount of fissionable material. For example if there are no enough material, neutrons may escape from the sample before inducing a fission event and the reaction would die out. In addition, we have to take into consideration the nuclei of other isotopes contained in the sample which may absorb neutrons for a nuclear reaction other than fission. What this means is that the purity of fissionable isotope affects the critical mass. For instance the natural uranium is composed of 99.3% ${}^{238}U$ while the remaining 0.7% is the fissionable ${}^{235}U$ isotope. The ${}^{238}U$ may absorb neutrons and removes neutrons from contributing to a chain reaction. Consequently, in order to have adequate quantity of fissionable material in a sample and reduce the critical mass the fissionable material needs to be concentrated or enriched. This enrichment can be up to 99% for weapon-grade material and between 3-5% ${}^{235}U$ for reactor grade material for electrical generation.

In a nuclear bomb chain reaction takes place since the reactions proceed uncontrolled almost instantaneously. In this case there is a quick and enormous release of energy from billions and billions of fissioning nuclei which cause massive explosion. However, for the practical production of nuclear energy the chain reaction process must be controlled.

15.10.2 Fusion

Another way of producing nuclear energy is through fusion in which a light nuclei fuse together to form a heavier nucleus and releasing energy in the process. Take for example the fusion atomic hydrogen and deuterium to form helium as indicated in equation 15.24.

$$ {}^{2}_{1}H(2.014102\,u)+{}^{3}_{1}H(3.016049\,u) \rightarrow {}^{4}_{2}He(4.002603\,u)+{}^{1}_{0}n(1.008665\,u) \qquad (15.24)$$

The mass difference for this reaction is

$(\Delta M) = m_{2H} + m_{3H} - (m_{He} - m_n) = 0.018883\,u$, and $Q = 0.018883\,u\times 931.5MeV = 17.6MeV$

Remarks: (1) The energy released may seem small in comparison to the more than $200MeV$ released from a fission reaction.If we have equal masses of hydrogen and uranium then a given mass of isotopes have many more nuclei than that of an equivalent mass of a heavy fissionable isotope.

(2) There is critical mass to consider since no chain reaction is involved.

(3) Fusion is a source of energy for stars including the sun.

15.11 Elementary Particles

The search for the fundamental building blocks of matter has been going on for quite a long time now. Initially it was thought that matter was made up of only electrons, protons and neutrons. As of today it is found that there are a variety of particles associated with the nucleus and these run into hundreds. Even though some of them are excited states of other particles they are nevertheless regarded as particles.Furthermore, each particle excluding the photon, π^0 and η^0 is also associated with a so-called antiparticle of the same mass and spin but with opposite charge and opposite alignment between the spin and magnetic moment. An example is the electron which has as its antiparticle the positron with a positive charge and of same electronic mass. The antiparticle is usually denoted by a line over the symbol for the particle equivalent such as $\overline{p}$ for the antiproton.

Sources of elementary particles

The elementary particles can be observed from the following sources;

(a) Cosmic Rays: These are rays of extraterrestrial origin travelling with speeds of the order of the speed of light. The primary cosmic rays entering the atmosphere consist mostly of protons with peak energy of about $10^{20}eV$. As soon as the flux enters the earth's atmosphere multiple collisions take place with atmospheric molecules leading to the production of secondary particle like the muons, electrons, positrons, photons, etc.

(b) Particle Accelerators: When particles are accelerated to high energies up to $10^3 GeV$, the result is the production of elementary particles. A typical reaction is of the form

$$p+p \to p+n+\pi^+.$$

15.11.1 Classification of Elementary particles

(a) The many elementary particles of varying categories can be broadly classified in accordance with Table 15.7.

(b) The photon is in its own class. It is a massless particle with unit spin. We may wish to recall that a particle can be viewed as spinning about its axis with a spin angular momentum in units of $h/2\pi$.

(c) The Leptons: The leptons are made up of the electron and the muon (μ) with their neutrinos (υ) with spin quantum number of $\frac{1}{2}$. Consequently, they are fermions just like other particles with half integral spins.

(d) The mesons: Their mass is intermediate between the electron and the proton with spin. They are therefore bosons like other particles with integral spin.

(e) The baryons: These are composed of the nucleons and the hyperons. In general a baryon is a fermion whose mass is at least as great as the proton mass. Hyperons are baryons which cannot decay strongly into a nucleon. The decay is weak as determined by the lifetime. The baryons are assigned numbers and in any interaction the baryon number is always conserved. Baryons and muons are collectively referred to as hadrons.

A peculiar observation in cosmic ray photographs and subsequently in the laboratory experiments was that certain particles were always produced in pairs. Consequently, they were called strange particles. They could be found among the mesons and baryons and have been assigned strangeness quantum number is indicated in Table 15.6. Particle that are not strange have strangeness quantum number of zero. The strangeness number is conserved in strong interaction as can be verified in equation 15.25(a) and (b).

$$\pi^- + p \to K^- + \Sigma^+ \qquad (15.25a)$$

$$\pi^0 + n \rightarrow K^+ + \Sigma^- \qquad (15.25b)$$

15.11.2 The Forces of Nature

In physics, there are four fundamental forces of nature namely; gravitational, electromagnetic and the nuclear force which is further divided into the weak and strong nuclear forces. Gravity mutually attracts all matter and is not particularly noticeable in our daily experience. The weak nuclear force is responsible for particle decay in which neutrinos and leptons are involved notably beta decays. It has a decay time of the order of $10^{-10} s$. The electromagnetic force is responsible for interactions between all charged particles and also particles with electric and magnetic moments. In fact, matter in bulk is held together electromagnetically. Both the gravitational and electromagnetic forces obey the inverse square law force with the latter being much stronger. The strong nuclear force is attractive and is the binding force in the nucleus acting between the mesons, between baryons and between mesons and baryons. The nuclear force is the strongest of all the forces with a decay time of $10^{-23} s$. In addition, it has a very short range of the order of $10^{-15} m$. The relative strengths of the forces are indicated in Table 15.6. In this table we have also given the mediating particle in the different interactions. We can understand what this means by considering electromagnetic interaction. According to quantum electrodynamics, the interaction between two electrons takes place because one of the electrons emits a photon which is promptly absorbed by the other electron. This is vividly illustrated in Figure 15.6 known as Feynmann diagramme named after the discoverer. In the diagram, time runs from left to right. The photon transfers energy from the first electron A to the second electron B. On gaining this energy the electron is accelerated. Since force is required for acceleration, the photon effectively transmits a force whose origin is the original electron some distance away. In this sense therefore, we say that photons mediated electromagnetic forces or they are the quanta of the electromagnetic field.

Table 15.6: Comparison of the four basic interactions

Field	Relative Strength	Associated Particle	Characteristic Time
Gravitational	10^{-39}	Graviton	$10^{16} s$
Electromagnetic	10^{-3}	Photon	$10^{-20} s$
Strong Interaction	1	Meson	$10^{-23} s$
Weak Interaction	10^{-13}	Intermediate Boson	$10^{-10} s$

The pi meson (pion) is recognized as the mediating particle in the nuclear strong interaction force. The graviton which is supposed to be the quantum of the gravitational field has not been discovered. The intermediate boson $W^{\pm}$ is the mediator of the weak interaction force. A typical Feymann diagram is shown in Figure 15.7 for such a positively charged mediator. Note that because the weak interaction force is short ranged, the mediator is massive compared with the long range electromagnetic force with massless photon as the mediator in conformity with the uncertainty principle (range = $\hbar/mc$). In the usual weak interaction with the neutrino shown in Figure 15.7, the neutrino is always changed into a lepton. The conclusion to make is that there is some relationship between the electromagnetic and weak interactions (Figure 15.8a and 15.8b) as demonstrated by the unified electro-weak thoory due to Abdus Salam and Weinberg along with Glashow the 1979 Nobel prize in physics. The intermediate bosons W^+ and W^- were discovered in 1983 with a mass of $82GeV$ while Z^0 was found to be, in 1984, with a mass of $92GeV$ as predicted. With the unification of the weak and electromagnetic forces we should in fact be talking of only basic natural forces: gravitational, electro-weak and strong. However, with the belief that there is only one universe with one set of physical laws, theoreticians are working hard towards the development the Grand Unified Theory of all natural forces.

15.11.3 Quarks

From 1960, more than 100 elementary particles have been discovered and a recurring pattern had been observed among the particles suggesting the possibility of a more fundamental variety of matter out of which these nuclear particles and ultimately the nucleus are formed. These fundamental particles were called quarks. In 1964 Gell-Mann and Zweig independently suggested the existence of three types of quarks. These are the up (u) and down (d) and the strange (s) quarks. The u and d quarks are adequate to build the hadrons with zero strangeness but the strange hadrons contain the s quark. Quarks are very unusual in that they have electrical charges which are fractions of a proton's charge: the u quark has charge of $+2/3$ proton charge, the d quark has charge of $-1/3$ of proton charge, while the s quark also has a charge of $-1/3$ proton charge and strangeness of -1. The antiquarks $\bar{u}, \bar{d}$ and $\bar{s}$ have opposite charges and strangeness. To obtain, note that the quarks have spin of $\frac{1}{2}$, therefore, to obtain the baryons from quark clusters we require a combination of odd number (three) quarks, while to obtain the mesons a combination of even number (two) quarks is required.

In Table 15.7, we illustrate the combination of quarks that account for the properties of some hadrons. The deduction is that the protons, neutrons and other hadrons are clusters quarks, which means that the proton itself, instead of being a fundamental particle, has internal structure. The structure has been confirmed in the same way that Rutherford used large angle scatters of alpha particles by atoms to confirm the existence of a positively charged nucleus. In this case, violent deflections of electrons and neutrinos fired at protons confirmed the existence of spin $\frac{1}{2}$ entities trapped within the proton, the entities having charges $2/3$ and $-1/3$ of the proton charge. Other observations include that the quarks are light with mass less than $1/3$ proton mass and also that some electrically neutral particles called gluons existed in proton in addition to the quarks. It is suspected that the gluons are the carriers of the force that attracts quarks together forming the proton just as the photon is the quantum of electromagnetic field. So far neither the quarks nor the gluons have been isolated and identified.

Table 15 .7: Formation of hadrons from quark clusters

Particle Quark	Clusters	Strangeness Number	Charge = sum of quark charge
p	uud	0	1
n	udd	0	0
Λ^0	uds	-1	0
Σ^+	uus	-1	1
Ξ^-	dss	-2	-1
Ω^-	sss	-3	-1
π^+	$u\bar{d}$	0	1
π^-	$d\bar{u}$	0	-1
K^+	$u\bar{s}$	+1	1

What we have done is provide a superficial treatment of this profound frontier subject called elementary particles with the hope of exciting the interest of the reader. The hadrons are no more than clusters of quarks of all descriptions and 'colours or flavours'. The quarks are believed for now to be truly elementary or the fundamental building blocks of the hadrons which include protons and neutrons. In addition, since leptons (e.g. electrons) are light and with no recognizable structure they are also regarded as fundamental particles. With the current recognition of quarks and leptons as the only fundamental particles it is strongly speculated that they are intimately related to each other. Work is in progress to unravel the underlying basis for this intimacy.

Example 15.8

Explain why the following reactions are not possible. (a) $\pi^- + n \rightarrow \Xi^- + K^+ + K^-$
(b) $\pi^- + p \rightarrow K^+ + K^-$

Solution

The procedure is to find out whether the different conservation laws are obeyed.

(a) For $\pi^- + n \rightarrow \Xi^- + K^+ + K^-$, the charge is conserved since -1 + 0 → -1+ 1– 1 = -1

Baryon number gives for $\pi^- + n \rightarrow \Xi^- + K^+ + K^-$ gives 0 + 1 → 1+ 0+ 0 = 1.

It follows that the baryon number is conserved.

For the strangeness number we have for this reaction

$\pi^- + n \rightarrow \Xi^- + K^+ + K^-$, 0 + 0 →-2 + 1 – 1 = -2.

It follows that the strangeness number is not conserved which makes the reaction not possible.

(b) For the reaction $\pi^- + p \rightarrow K^+ + K^-$, we have for the charge -1 + 1 → 1– 1 = 0.

For this the charge is conserved.

For the strangeness number, we have for the reaction $\pi^- + p \rightarrow K^+ + K^-$

0 + 0 → 1 -1 = 0.

In this case the strangeness number is conserved.

For the baryon number we have for this reaction $\pi^- + p \rightarrow K^+ + K^-$; 0 + 1 → 0 + 0 which clearly show that the baryon number is violated.

The conclusion is that the reaction $\pi^- + p \rightarrow K^+ + K^-$ is not possible.

Summary

1. The atomic nucleus is made up positively charged protons and neutral neutrons and is held together by strong attractive short-ranged forces. If the nucleus is assumed spherical the average radius is given by $R = r_0 A^{1/3}$

   where $r_0 = 1.2 \times 10^{-5}\, m$.
2. The binding energy ($B.E$) is the energy required to form a nucleus and it can be expressed in terms of mass difference as $B.E = [Zm_p + (A-Z)m_n - M_{Z,A}]$.

   The mass spectrometer is used for estimating nuclear masses.
3. Radioactive materials emit alpha and beta particles and gamma rays. The alpha and beta particles are charged whereas the gamma rays are uncharged and besides they are in fact electromagnetic radiations.
4. Radioactive materials decay exponentially and they obey the expression

   $N = N_0 e^{-\lambda t}$. The half life is the time it will take for half of the atoms to disintegrate. The half life ($t_{1/2}$) can be obtained from $t_{1/2} = \dfrac{0.693}{\lambda}$ while the average life is given by $T_a = 1/\lambda$. In addition, the activity of a radioactive material is expressed as A =λN. The unit of radioactivity is the Curie (Ci) for which $1Ci = 3.7 \times 10^{10}$ disintegrations/second.
5. Radioactive materials find various uses in medicine for example in the treatment of cancer and in industry, for instance, food preservation. However, it must be stressed that radioactive materials are very hazardous and extra care must be exercised in handling and use of these materials. The biological effects of radiation may be categorized as either somatic or genetic.
6. In nuclear reaction the mass number and charge are always conserved. Nuclear fission is a process in which a heavy nucleus breaks into two more fragments with the release of an enormous amount of energy. Some heavy nuclei undergo fission but at very slow rates. However, most fission reactions are now induced as obtained in nuclear reactors. In nuclear reactors for generating energy chain reaction is sustained and fission is controlled. In nuclear weapons no such control is needed. In nuclear fusion two light nuclei fuse together to form a heavier nucleus releasing a large amount of energy. Practical fusion reactors are not available because very high temperatures are needed to initiate fusion reaction.
7. The study of elementary particles stems from man's age-long quest for the fundamental building blocks of matter. The particles are observed in cosmic rays and particle accelerators.

They can be classified into photon, lepton, meson and baryon. The four fundamental forces of nature are gravitational, electromagnetic, strong and weak interactions. Presently the quarks are recognized as the fundamental building blocks of the hadrons which include protons and neutrons.

Exercise 15

15.1 Given the nucleus $^{270}_{107}X$, determine the number of neutrons.
A. 270 **B.** 107 **C.** 377 **D.** 163

15.2 Which of this entity is a helium nucleus?
A. gamma particle **B.** α particle **C.** X-ray **D.** beta particle

15.3 The binding energy is the energy required to; **A.** to accelerate **B.** to assemble **C.** to tear it apart **D.** stop it from decaying.

15.4 If a 10g sample of a radioactive material decays for five half lives calculate how much of the radioactive material is left. **A.** 0.625g **B.** 1.25g **C.** 0.313g **D.** no enough information.

15.5 In a radioactive series it eventually end with
A. Lead **B.** hydrogen **C.** a stable nucleus **D.** a gamma decay.

15.6 In virtually all β^- decay processes there is emitted ------ along with the electron e^-.
A. antineutrino **B.** neutrino **C.** α **D.** γ

15.7 What is the binding energy per nucleon of **A.** 305MeV **B.** 236.5MeV **C.** 150MeV **D.** 50MeV

15.8 Polonium ($^{214}_{84}Po$) undergoes α-decay. Determine the daughter nucleus.
A. $^{214}_{82}Pb$ **B.** $^{310}_{82}Pb$ **C.** $^{210}_{82}Pb$ **D.** $^{110}_{82}Pb$

15.9 Here on earth we receive the least of the average yearly dose from the following sources of radiation. **A.** Cosmic rays **B.** medical and dental diagnostics **C.** radioactivity from the earth **D.** nuclear fallout.

15.10 In a nuclear reaction find the name, atomic number and the nucleon number of the compound nucleus if a nitrogen $^{14}_{7}N$ nucleus absorbs a deuterium ($^{2}_{1}H$).
A. $^{12}_{8}O$ **B.** $^{12}_{6}C$ **C.** $^{16}_{6}C$ **D.** $^{16}_{8}O$

15.11 Determine the count rate resulting from 0.1mg of cesium 137 with a half life of 28 years.

15.12 What would be the energy released if the following fission reaction occurs?

$$^{1}_{0}n + ^{235}_{92}U \rightarrow ^{141}_{56}Ba + ^{92}_{36}Kr + 3(^{1}_{0}n)$$

1.008665u 235.043915u 140.9139u 91.8973u 3(1.008665u)

15.13 The helium atom could be broken down to form 2 hydrogen atoms and two neutrons. By using atomic mass units determine the binding energy of the $^{4}_{2}He$ nucleus.

15.14 By means of β^- decay $^{234}_{90}Th$ changes into $^{234}_{91}Pa$. Calculate the energy released.

15.15 Find the number of protons and neutrons contained in the nucleus of the following atoms $^{35}_{17}Cl$, $^{56}_{26}Fe$ and $^{120}_{50}Sn$.

15.16 Given a sample of strontium-90, determine how long it would take for 2/3 of the sample to decay.

15.17 Determine the half-life of an isotope if after 2hrs, 1/8 of the radioactive isotope is left.

15.18 Suppose Thorium $^{232}_{90}Th$ decays by α decay to produce a daughter nucleus which in itself undergoes β^- decay. Determine the nucleus produced in the form $^{A}_{Z}X$.

15.19 When ^{226}Ra decays it emits a 4.706MeV of α-particle. Determine the velocity of recoil of the daughter nucleus from the stationary radium-226.

15.20 Calculate the minimum kinetic energy a proton will have in order to initiate the reaction $^{3}_{1}H(p,d)^{2}_{1}H$ where d the deuterium nucleus.

WAVES

CHAPTER 16
WAVE PHYSICS

16.0 Introduction

The study of wave motion is very important in physics because it involves the transmission of energy and momentum from a source to a detector.
Waves are disturbances usually set up by vibrating bodies, which result in the transfer of energy. They occur both in material media and in free space. When they occur in space they usually go by the name *radiations*, and they involve changing electric and magnetic fields with no particle motion. Light waves are examples of these. In material media, they may involve changes in some physical quantities like pressure and temperature. While the individual particles do not move far, the disturbance itself may travel great distances transferring energy and momentum. Sound is an example of such a media-important wave.

Waves can be generally defined as the transport or propagation of a physical quantity, from one point to another, due to the introduction of some form of energy in the medium.
When a vibrating object sets up waves, the waves travel away from it. The particles of the medium can move either in the same direction as the wave or at right angles to it, about a mean position. When the particles move in the same direction as the wave, the wave is said to have displacements along the wave direction. The wave is then said to be *longitudinal*. Example of this type of wave is sound wave. When the particles vibrate at right angles to the wave direction, the wave is said to be *transverse*. Example of this type of wave is ripples on water surface and electromagnetic wave. In the case of electromagnetic wave there are no particle vibrations; rather it is the interaction of the electric and magnetic fields which are oscillating at right angles to each other. These waves are called electromagnetic waves.

16.1 Characteristics of Waves

Wave Front

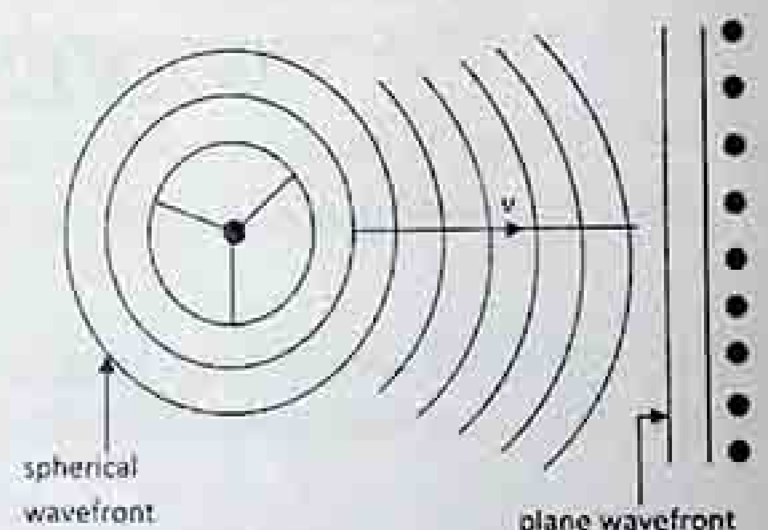


Fig. 16.1: Types of wavefronts

This is the surface in an advancing wave over which the disturbance or particles have the same phase at all points (i.e. the same amplitude and frequency). Figure 16.1 shows an example of a wave front. If S is the source of disturbance at time t, the radius of the sphere formed is vt. At another time t_1, the new radius is vt_1 and so on as the wave propagates, with v as the velocity of the wave. The spheres formed have surfaces known as the spherical *wavefronts.*
The most important characteristic of the wave fronts is that particles are in the same phase. At large distances the spheres disappear and plane wavefronts in two dimensions or points in one dimension are formed which are special types of spherical wavefronts.

Classification of wave according to production mode

Waves can be classified either according to their mode of production or according to their mode of propagation. Classification according to the mode of production includes two types.

i. **Transverse wave**: In this type, the wave is propagated by vibrations perpendicular to the direction of travel of the wave. Examples are waves on plucked stretch strings, surface water wave, secondary earthquake waves and electromagnetic waves which include light and radio waves.

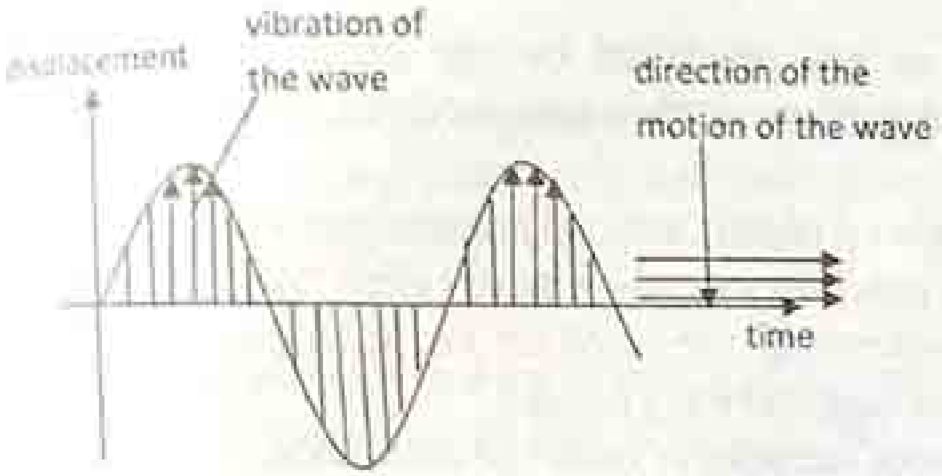


Fig. 16.2: Transverse Wave

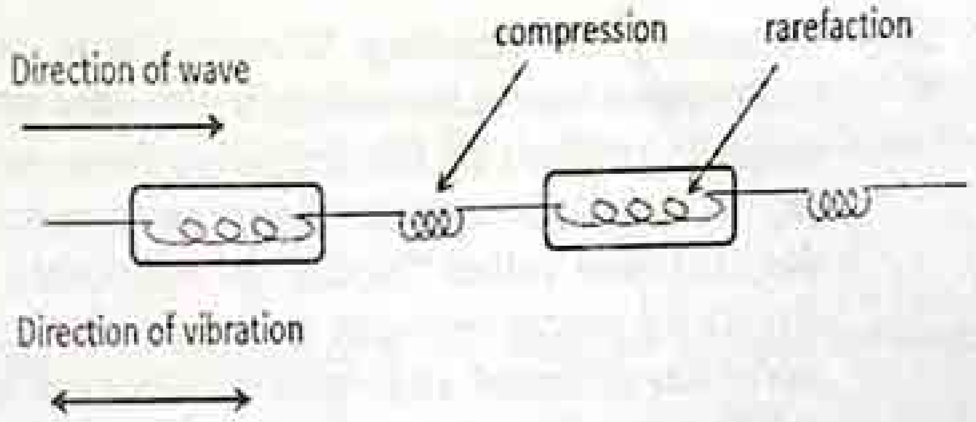


Fig.16.3: Longitudinal Wave

ii. **Longitudinal wave:** This is the type in which the vibrations occur in the same direction (or parallel) as the direction of propagation of the waves. Sound waves, earthquakes, seismic waves and waves from a tuning fork are longitudinal waves. This type of wave is characterized by a back and forth compression and stretching of the medium along the direction of propagation (Figure 16.3). At compression the particles move in the direction of the wave and at rarefaction, the particles move in opposite direction to the wave motion.

In summary, the direction of propagation of the wave may either be the same as that of the vibration or perpendicular to the vibration. If the direction of vibration is perpendicular to the direction of propagation of the wave, the wave is said to be transverse. If the direction of vibration is the same as the direction of propagation, the wave is said to be longitudinal. Apart from the above we can further classify waves into mechanical and electromagnetic waves. String, water and sound waves are called Mechanical waves, while light, radio and other radiations are called electromagnetic waves.

Many kinds of wave motion phenomena occur in nature, and they can actually be grouped. The waves whose motions are predictable and repeatable are called *periodic*, while unpredictable ones are called *non-periodic* or *random*. This latter group is referred to as *noise*. The shape of a wave, as it moves can also identify it, hence there are plane waves, spherical waves and harmonic waves. Whether a wave is heard or not also distinguishes it from others, hence there are *audible* or *audio* waves, *ultrasonic* (beyond the range of hearing) waves and *infrasonic* (below the range of hearing) waves.

Table 16.1: Some commonly encountered waves

Waves	Classification
Sound	Longitudinal
Speech	Longitudinal
Earthquake	Longitudinal
Tidal	Longitudinal
Water	Longitudinal and Transverse (because water particles have vertical And horizontal components)
Light	Transverse
String wave	Transverse
Heat	Transverse
Radio and TV	Transverse
X-rays and Gamma rays	Transverse

The classification of waves according to their mode of propagation includes:

i. Plane-progressive (Travelling) Waves: In this type of wave, the waveform moves along repeatedly with the speed of the wave and with same frequency. The waveform repeats itself in a given period within a distance called the wavelength. Plane-progressive waves are characterized by:

(a) particle vibrations that are of the same amplitude and frequency, but with phases of the vibrations changing at different points along the wave.

(b) pressure variations that are the same at every point in the medium.

ii. Stationary or Standing Waves: These waves are produced by the superposition of two progressive waves travelling in the opposite directions in the same medium. Stationary waves are characterized by the following features:

(a) There exist points along the wave where there are permanent zero or minimum displacement called "Nodes" N and points of maximum displacement called "Antinodes" A;

(b) At all points between successive nodes, vibrations are in phase and each point along the wave has different amplitude of vibration from neighbouring points;

(c) The wavelength is twice the distance between successive nodes or antinodes;

(d) In sound waves the pressure variations (compression and rarefaction) are always maximum at nodes and zero at antinodes.

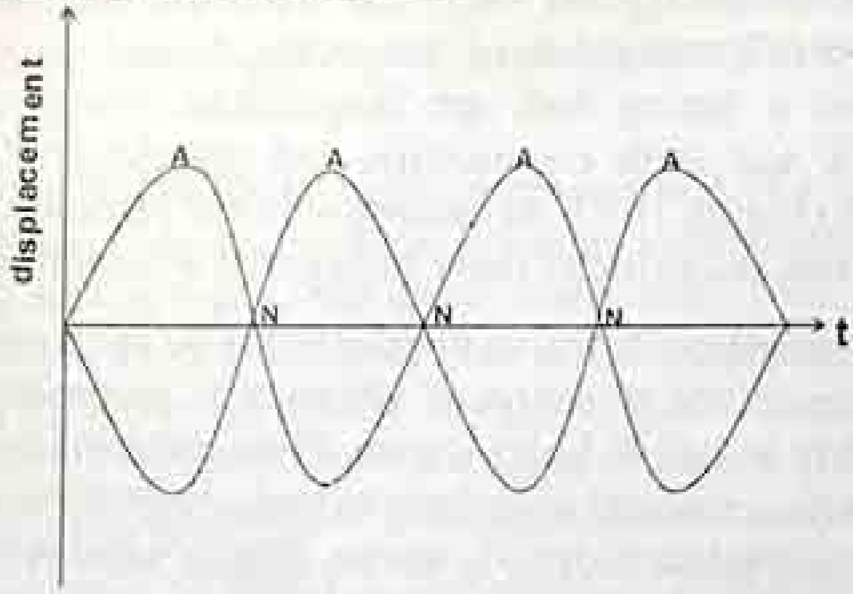
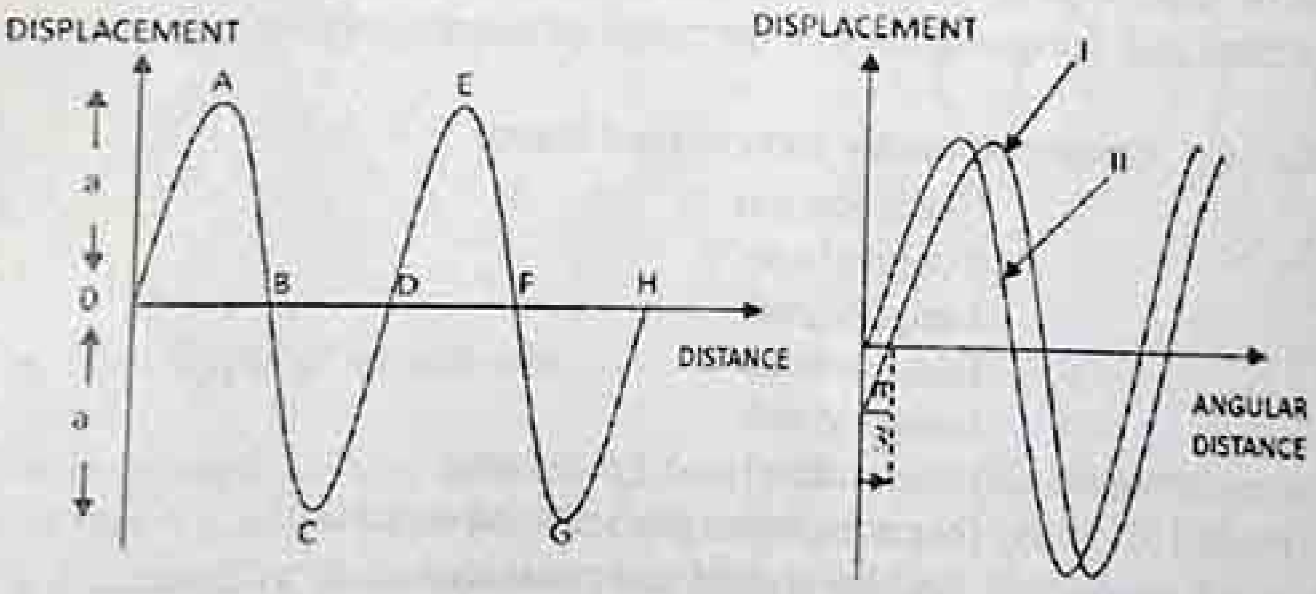


Fig. 16.4: Vibration pattern formed by a string fixed at both ends due to reflection

16.2 Wave Parameters

Periodic waves have several quantities, which characterize them. Figure 16.5(a), which is a plot of the displacement of a sinusoidal wave against distance along the x-axis. The same shape results if displacement is plotted against angular distance along the x-axis, as shown in Figure 16.5(b).



(a) Sinusoidal displacement (b) Two waves, phase difference of $\pi/3$ radians.

Fig. 16.5: Waveforms of sinusoidal displacements

Amplitude, *a*:

This is the maximum displacement about a mean position that a wave can undergo during its motion. Since it is distance, its unit is metre.

Wavelength, λ

The points A and E, in Figure 16.5 (a) are the peaks of the wave; so also are points C and G on the lower portion. The horizontal distance between the successive peaks is the wavelength (i.e. A to E, O to D, B to F and D to H). The wavelength is also the horizontal distance moved by the wave in a cycle. The unit is also metre.

Cycle

When a wave has undergone maximum displacement up and maximum displacement down and returns to its original position, it has gone through a cycle. In Figure 16.5 (a), this corresponds to O-A-B-C-D, or B-C-D-E-F, or D-E-F-G-H. The cycle has no unit, as it is just a number.

Period, T

This is the time it takes a wave to complete one cycle of its motion. It is also the time elapsed between successive wave crests. Its unit is the second.

Phase, ϕ

This is the angular distance between a wave that has zero displacement at time $t = 0$ and that which has non-zero displacement at same $t = 0$. Figure 16.6 (b) shows two waves differing in angular distance by $\pi/3$. Wave (I) lags wave (II) by $\pi/3$ or wave (II) is said to lead wave (I) by $\pi/3$ radians.

The phase ϕ is equal to $\frac{2\pi}{3}x$, where x is the linear distance travelled by the wave.

Frequency, f

This is the number of cycles a wave completes in one second. Its unit is thus cycles per second or Hertz (Hz). It is the inverse of period, thus

$$f = \frac{1}{T} \tag{16.1}$$

Velocity, v

This is the speed at which a wave crests or peak travels. If we consider the distance λ covered during one cycle, and the time T to cover it, then

$$v = \frac{\lambda}{T} = \lambda f \tag{16.2}$$

Example 16.1

A wave pulse on a string moves a distance of $8m$ in $0.04s$. Determine:

(a) The velocity of the pulse,
(b) The frequency of a periodic wave on the string if its wavelength is $0.5m$.

Solution

(a) $v = \frac{s}{t} = \frac{8}{0.04} = 200\,ms^{-1}$

(b) Since $v = f\lambda$, $f = \frac{v}{\lambda} = \frac{200}{0.5} = 400\,Hz$

Each type of wave has its own characteristic velocity, and this velocity can be predicted from the physical laws that describe the particular wave phenomenon-as we shall see in the discussion on the velocity of sound. Electromagnetic waves in vacuum have a velocity of $3\times10^{8}\,ms^{-1}$ and smaller values in matter. The velocity of sound, on the other hand, in air at temperature of 30°C is $344\,ms^{-1}$ but is much higher in solids. On the whole, the velocity of any wave depends on the type of wave, on the medium in which it is travelling and sometimes on its frequency.

16.3 Progressive Wave Equation

The transverse waves and longitudinal waves described above are progressive waves. In a progressive wave, energy is transferred from the source outwards. The wave profile can be seen to move in the direction of propagation of the wave.

An equation can be written to represent a progressive wave. The progressive wave equation represents the displacement y of a vibrating particle in a medium in which the wave passes.

Figure 16.6 shows a progressive wave moving with a velocity v from left to right when the source at O oscillates in simple harmonic motion with angular velocity ω.

The motion of a body in a sinusoidal pattern can be illustrated to show the angle made by the body at every instant.

For a small angle θ, the distance, s moved by the object is approximately a straight line, thus;

$$\sin\theta = \frac{s}{r} \tag{16.3}$$

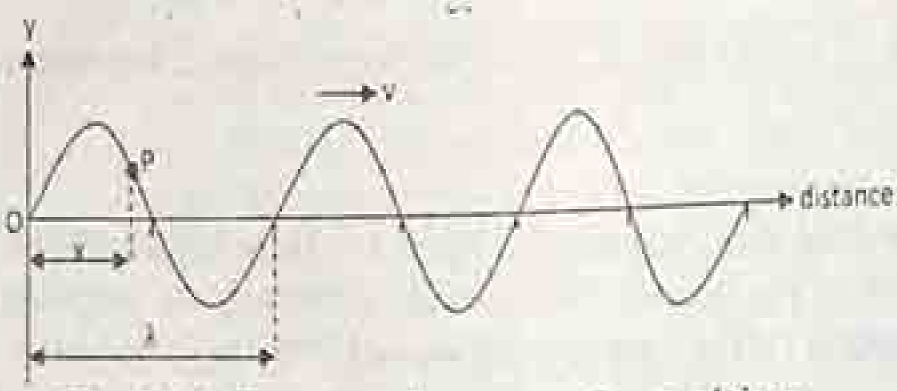


Fig.16.6 Progressive wave in positive direction

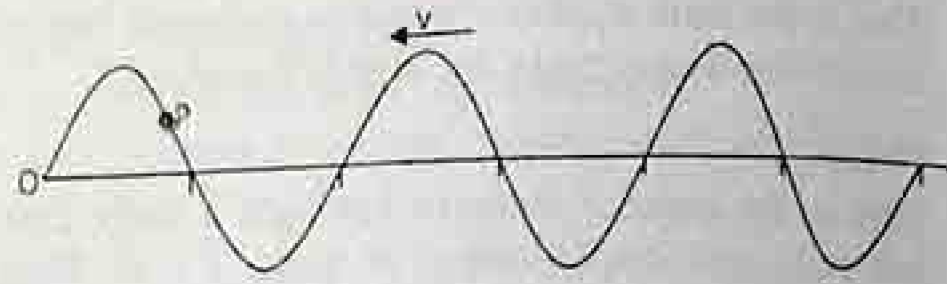


Fig. 16.7: Progressive wave in the negative direction

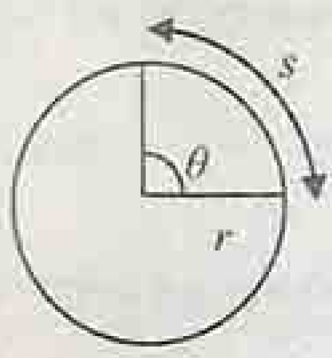


Fig. 16.8: Cyclic motion of the body

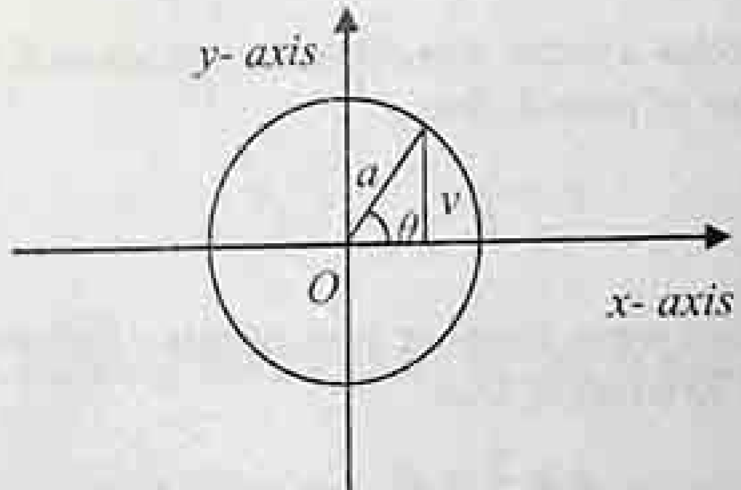


Fig. 16.9: Representation of sinusoidal motion in a circle

But at very small angles, $\sin\theta \approx \theta$, thus

$$s = r\theta \tag{16.4}$$

The angular frequency, ω, is a measure of the ratio of the angle θ made with time, t.

$$\omega = \frac{\theta}{t} \tag{16.5}$$

When the body moves one full cycle, it makes an angle of $2\pi\, rad$ which is equivalent to 360°.

Therefore,
$$\omega = \frac{2\pi}{T} = 2\pi f \tag{16.6}$$

where f is the frequency of the motion.

The oscillation of the body varies between a minimum amplitude $(y = 0)$ and a maximum $(y = a)$. The oscillation of the particle O about a mean position may be represented by the equation

$$\sin\theta = \frac{y}{a} \tag{16.7}$$

$$y = a\sin\theta = a\sin\omega t \tag{16.8}$$

where a is the amplitude of oscillation. For a particle P in the medium at a distance x from O, its oscillation lags behind that of O by a phase difference of ϕ.

Since a full wavelength λ corresponds to a phase difference of 2π, the distance x corresponds to a phase difference of

$$\phi = \frac{2\pi}{\lambda}x \tag{16.9}$$

Therefore, the displacement y of the particle P which is at a distance x from O can be written as

$$y = a\sin\left(\omega t - \frac{2\pi x}{\lambda}\right) \tag{16.10}$$

Substituting $\omega = 2\pi f$, where f = frequency into equation 16.10, we have

$$y = a\sin\left(2\pi ft - \frac{2\pi x}{\lambda}\right) \qquad (16.11)$$

$$y = a\sin 2\pi\left(ft - \frac{x}{\lambda}\right)$$

$$y = a\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) \qquad (16.12)$$

Substituting $f = v/\lambda$ into equation 16.11, we obtain

$$y = a\sin 2\pi\left(\frac{vt}{\lambda} - \frac{x}{\lambda}\right) \qquad (16.13)$$

$$y = a\sin \frac{2\pi}{\lambda}(vt - x) \qquad (16.14)$$

We summarise below the 3 key equations of motion of progressive wave. Just remember one and you can deduce the others and use them according to the parameters given in the problem

$$y = a\sin\left(\omega t - \frac{2\pi x}{\lambda}\right) \qquad (16.15)$$

$$y = a\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) \qquad (16.16)$$

$$y = a\sin \frac{2\pi}{\lambda}(vt - x) \qquad (16.17)$$

Stationary or standing waves are waves produced by the superposition of two waves travelling in the opposite directions in a medium.

Example 16.2

A progressive wave is represented by the equation $y = 0.1\sin\left(20\pi t - \frac{20\pi x}{17}\right)$

where y is in mm, t in seconds and x, the distance from the point O, in metres. Find

(i) The frequency,
(ii) The wavelength,
(iii) The speed of the wave
(iv) What is the phase difference between a point $0.25m$ from O and another point $1.00m$ from O?
(v) Write the wave equation for a progressive wave having twice the amplitude, twice the frequency and moving in the opposite direction in the same medium.

Solution

Comparing the equation $y = 0.1\sin\left(20\pi t - \frac{20\pi x}{17}\right)$ with the wave equation $y = a\sin\left(\omega t - \frac{2\pi x}{\lambda}\right)$,

(i) $\omega = 2\pi f = 20\pi$

$\therefore$ frequency, $f = \frac{20\pi}{2\pi} = 10Hz$

(ii) $\frac{2\pi x}{\lambda} = \frac{20\pi x}{17}$

$\therefore \lambda = \frac{17}{10} = 1.7m$

(iii) Speed of the wave, $v = f\lambda = 10 \times 1.7 = 17ms^{-1}$

(iv) Distance between the two points $= (1.0 - 0.25) = 0.75m$

One full wavelength λ corresponds to a phase difference of 2π radians.

∴ A distance of $0.75m$ corresponds to a phase difference.

$= \frac{2\pi}{\lambda} \times 0.75 = \frac{2\pi}{1.7} \times 0.75 = 2.77\,rad$

(v) Amplitude of first wave $= 0.1m$

∴ Amplitude of second wave $= 0.2m$

Frequency $= 20Hz$; $\omega = 2\pi f = 40\pi$

Wavelength, $\lambda = \frac{v}{f} = \frac{17}{20}$

$\frac{2\pi x}{\lambda} = \frac{2\pi x}{17/20} = \frac{40\pi x}{17}$

Therefore, the required wave equation is $y = 0.2\sin\left(\omega t - \frac{2\pi x}{\lambda}\right) = 0.2\sin\left(40\pi t + \frac{40\pi x}{17}\right)$.

Example 16.3

A plane progressive wave is represented by the equation $y = a\sin\left(100\pi t - \frac{100\pi x}{17}\right)m$. Determine:

(i) The frequency of the wave,
(ii) Its wavelength,
(iii) Its speed, and
(iv) Its phase angle at $34cm$ from the origin.

Solution

(i) $\omega = 2\pi f$, so $f = \frac{\omega}{2\pi}$, i.e. $f = \frac{100\pi}{2\pi} = 50\,Hz$.

(ii) $\frac{100\pi x}{17} = kx = \frac{2\pi x}{\lambda}$, hence, $\lambda = \frac{17}{50} = 0.34m$

(iii) $v = \lambda f = 50 \times 0.34 = 17ms^{-1}$

(iv) $\phi = \frac{2\pi}{\lambda}x = \frac{2\pi \times 0.34}{0.34} = 2\pi$

16.4 General Properties of Waves

Reflection of Waves

When a wave meets a boundary or an obstruction, a number of things can happen to the wave, depending on the nature of the boundary. It can be absorbed in part, it can be partly transmitted and it can be partly reflected. The laws governing reflection are similar to those discussed for light waves in the geometric Optics section.

In a string with one end fixed, the reflected wave is the inverted mirror-image of the original incident wave; the shape and frequency remaining the same. In the case of a string with the ends free, the reflected wave is still the mirror-image of the original wave, but is not inverted. The reasons for these results can be appreciated from the principles of mechanics.

Echoes and Reverberations are the effects of reflection of sound waves. Reverberation is the persistence of sound waves in an enclosure after the source of the sound has ceased. This effect is mainly due to (*i*) direct sound waves from the source; (ii) echoes produced by walls and ceilings; and (iii) sound waves diffused from walls, ceilings and other objects within the enclosure. Reflection can also result in the production of stationary or standing waves, when the reflected wave is the opposite direction to the incident wave in the same medium.

Refraction of Waves

The phenomenon called refraction occurs when waves pass from one medium into another; in other words, when they cross boundaries between two different media. The path of the wave is either bent away from the normal or towards the normal, depending on whether the second medium is less dense or denser, respectively, than the first medium when there is oblique incidence. This change in the direction of the wave is due to the change in velocity of the wave entering a different medium. As the velocity changes, there is a corresponding change in wavelength while frequency remains the same. The amount of refraction or bending of the wave depends on the ratio of velocities of the wave in the two media.

Diffraction of Waves

When a wave meets an obstacle or discontinuity, it has the ability to spread round it. This phenomenon is known as diffraction. The possibility of diffraction depends on the relationship between the size of the obstacle and the wavelength of the wave under consideration. The dimension of the obstacle must be comparable to the wavelength of the wave. This property is also the result of the superposition of waves on the same wavefront using Huygens principles of secondary wavelets. Generally, diffraction is easily observed in waves with longer wavelengths.

Interference Effects

This effect and diffraction are the two distinguishing properties of wave motion from other type of motions. The superposing waves must be coherent (i.e. must have same frequency and maintain a constant phase relationship). Waves are said to produce interference when they interact with one another or with boundaries. When two waves or more travel in a medium, they result in a single or complex wave, which is the sum of the displacements of the individual waves. This property of waves obeys the principle of superposition and applies to all waves. The principle of superposition states that when two or more waves propagating in the same medium interact, the resultant displacement at any point along the wave is the vector sum of the individual displacements at such points. Althou gh the resulting wave may be complex, the individual waves do not change form.

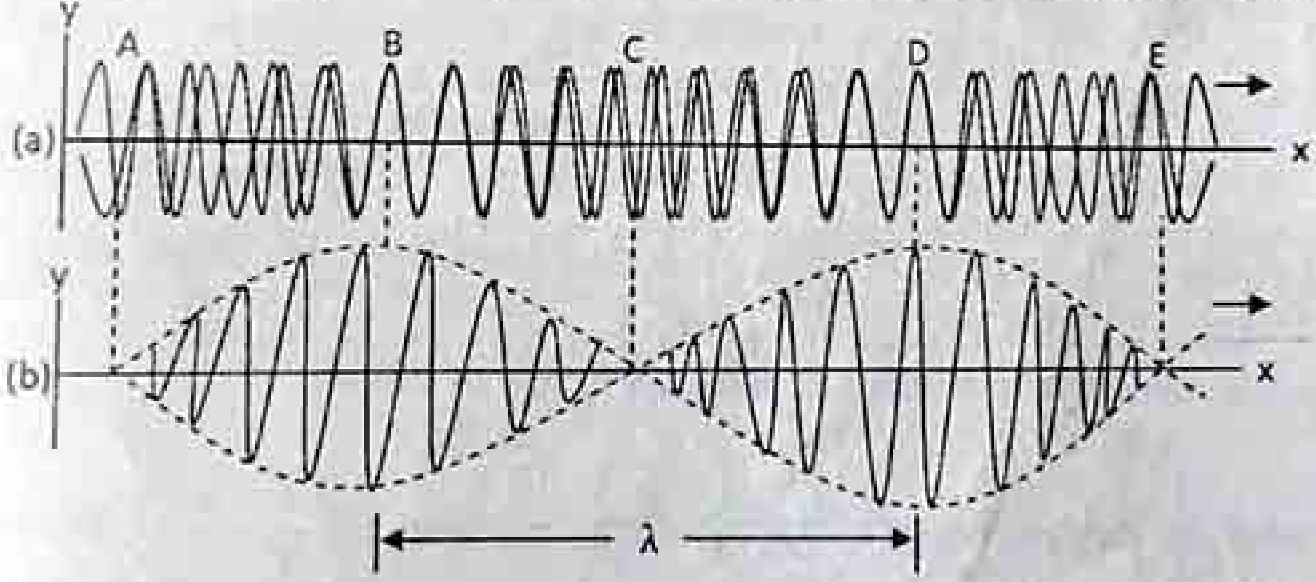


Fig. 16.10: (a) Two waves of slightly different frequencies, f_1 and f_2 add to form the resultant wave. (b) The dashed lines show how the amplitude changes.

The interference effects produced by waves depend on their phases. If the waves reaching a point have their maximum displacements at the same time, they are said to be in phase. Their displacements then add constructively and are said to interfere constructively. However, when one maximum coincides with the other's minimum displacement, the interference add destructively since they are not then in phase. Such waves are actually 180° out of phase. Figure 16.10 illustrates this.

Specifically, when two or more waves combine together to produce a maximum zero effect, they are accordingly said to have interfered destructively. Interference however is the superposition of different frequencies e.g. $y_1 = A\sin(\omega_1 t + \phi)$ and $y_2 = A\sin(\omega_2 t + \phi)$. By superposition, $y = y_1 + y_2$. The result will either be zero or more than individual intensities.

Supposing a wave travelling along the x-axis encounters a boundary and is reflected at rigid boundary (e.g. a wave pulse travelling on a string with the end of the string fixed), the incident wave (travelling from right to left) maybe described by the equation;

$$y_1 = A\sin(\omega t + kx) \tag{16.18}$$

while the reflected wave (travelling from left to right) is given by

$$y_2 = -A\sin(\omega t + kx) \tag{16.19}$$

Superposition of the two waves gives

$$y = y_1 + y_2 \qquad (16.20)$$
$$= A[\sin(\omega t + kx) - \sin(\omega t - kx)]$$

Applying trigonometry identity, this gives

$$y = (2A\sin kx)\cos \omega t \qquad (16.21)$$

Let $A_{sw} = 2A$,

$$y = (A_{sw} \sin kx)\cos \omega t \qquad (16.22)$$

At the nodes, the maximum amplitude is zero. So,

$$\sin kx = 0, \qquad kx = 0, \pi, 2\pi, \cdots, n\pi \qquad (16.23)$$

where n is a whole number.

While at the antinodes, the amplitude is maximum.

$$\sin kx = 1, \qquad kx = \frac{1}{2}\pi, \frac{3}{2}\pi, \frac{5}{2}\pi, \cdots, \left(n + \frac{1}{2}\right)\pi \qquad (16.24)$$

Summary

1. Waves are disturbances usually set up by vibrating bodies.
2. Waves are classified in terms of shape, movement, frequency range and behaviour.
3. All waves have certain characteristics in common, and when they meet boundaries, they suffer reflection, refraction and diffraction.
4. They also exhibit the characteristics of interference, beat phenomenon and Doppler Effect.
5. Two major types of waves exist: These are transverse and longitudinal.
6. The general equation of progressive wave is:

$$y = a\sin\left(\omega t + \frac{2\pi x}{\lambda}\right) \text{ or } y = a\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right) \text{ or } y = a\sin\frac{2\pi x}{\lambda}(vt + x)$$

Exercise 16

16.1 A wave travels a distance $30cm$ in $2s$. The distance between successive crests of the wave is 2 cm. What is the frequency of the wave in Hz? **A.** 5 **B.** 7.5 **C.** 3.5 **D.** 2.0

16.2 What is the period of a wave of wavelength $80m$ at a speed of $800\,ms^{-1}$?
A. $10s$ **B.** $5s$ **C.** $1s$ **D.** $0.1s$

16.3 A progressive wave is represented by the equation $y = a\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right)$. Write down an expression for the angular velocity ω. **A.** T/ϕ **B.** $2\pi/T$ **C.** $2\pi T$ **D.** $1/f$

16.4 What is the amplitude of the above wave in question 3? **A.** π **B.** a **C.** πx **D.** λ

16.5 Equation of a transverse wave along a string is $y = 6.0\sin(0.20x + 4.0t)$, where x and y are in cm and t in sec. Calculate the amplitude in $metres$?
A. 6.0 **B.** 0.60 **C.** 0.06 **D.** 0.006

16.6 What is the wavelength of the wave in metres in question 1.5?
A. 0.50 **B.** 0.31 **C.** 0.42 **D.** 0.60

16.7 A progressive wave has a wavelength of $20cm$. Calculate the phase difference between two points at distance of $5cm$ apart?
A. 0.2π **B.** 0.5π **C.** 2π **D.** π

16.8 Which of the following properties is a consequence of spreading of wave after passing through tiny openings? **A.** Interference **B.** Reflection **C.** Refraction **D.** Diffraction

16.9 Points (or regions) of maximum displacements of a wave are called
A. Nodes **B.** Crests **C.** Antinodes **D.** Troughs

16.10 The period of a wave is a $0.05s$. Calculate the wavelength if its speed is $330\,ms^{-1}$.
A. $15.6m$ **B.** $16.5cm$ **C.** $6.6m$ **D.** $14.5m$

16.11 A harmonic wave is travelling along a string? Where in this wave is the kinetic energy at maximum? The potential energy? The total energy?

16.12 You have a long thin steel rod and a hammer, how must you hit the end of the rod to generate a longitudinal wave?

16.13 What is the justification in using the x-axis to represent time as well as angles when plotting the displacement curve of a wave?

16.14 State giving reasons, whether these are transverse or longitudinal:
(a) Sound waves on a string (b) Water waves (c) Radio waves.

16.15 In section 16.2, it was explained that a travelling sinusoidal wave can be represented by the expression $y = A\sin(\omega t - \phi)$. If such a wave has an amplitude $2cm$, wavelength $4cm$ and frequency $0.5Hz$, calculate the values of the displacement, y for a linear distance along x-axis of $2cm$ at:
(a) $t = 0$ (b) $t = 0.5s$ (c) $t = 1.5s$ (d) $t = 2.5s$

16.16 A sea wave has a wavelength $140m$ and a period 9.80 milli-seconds. Calculate the frequency, the wave number and the velocity of the wave.

16.17 A sound wave has a frequency of $600Hz$ while the frequency of yellow light is about $5\times10^{14}\,Hz$. In air, sound travels at about $342\,ms^{-1}$ and light at $3\times10^{8}\,ms^{-1}$. Determine the wavelengths of the two waves.

16.18 An elastic wave on a string has amplitude of $4.0cm$, a wavelength of $2.4cm$, and a velocity of $12\,ms^{-1}$ in the positive x-direction. At $t = 0$, this wave has a crest at $x = 0$.
(a) What is the mathematical equation describing this wave as function of x and t?
(b) Determine the period, frequency, angular frequency, and wave number of the wave.

16.19 If a wave is represented by the expression $y = A\sin\left(1000\,\pi t - \dfrac{\pi x}{17}\right)cm$, where t is in seconds and y is the displacement, find:
(a) The wavelength (b) The velocity (c) The frequency

16.20 Radio Nigeria transmits music on a wavelength of $1500m$ and frequency of $200kHz$, calculate the velocity of the radio waves that carry the music. Determine also the frequency of a local programme transmitted on a wavelength of $330m$.

CHAPTER 17
NATURE, PRODUCTION AND PROPAGATION OF SOUND

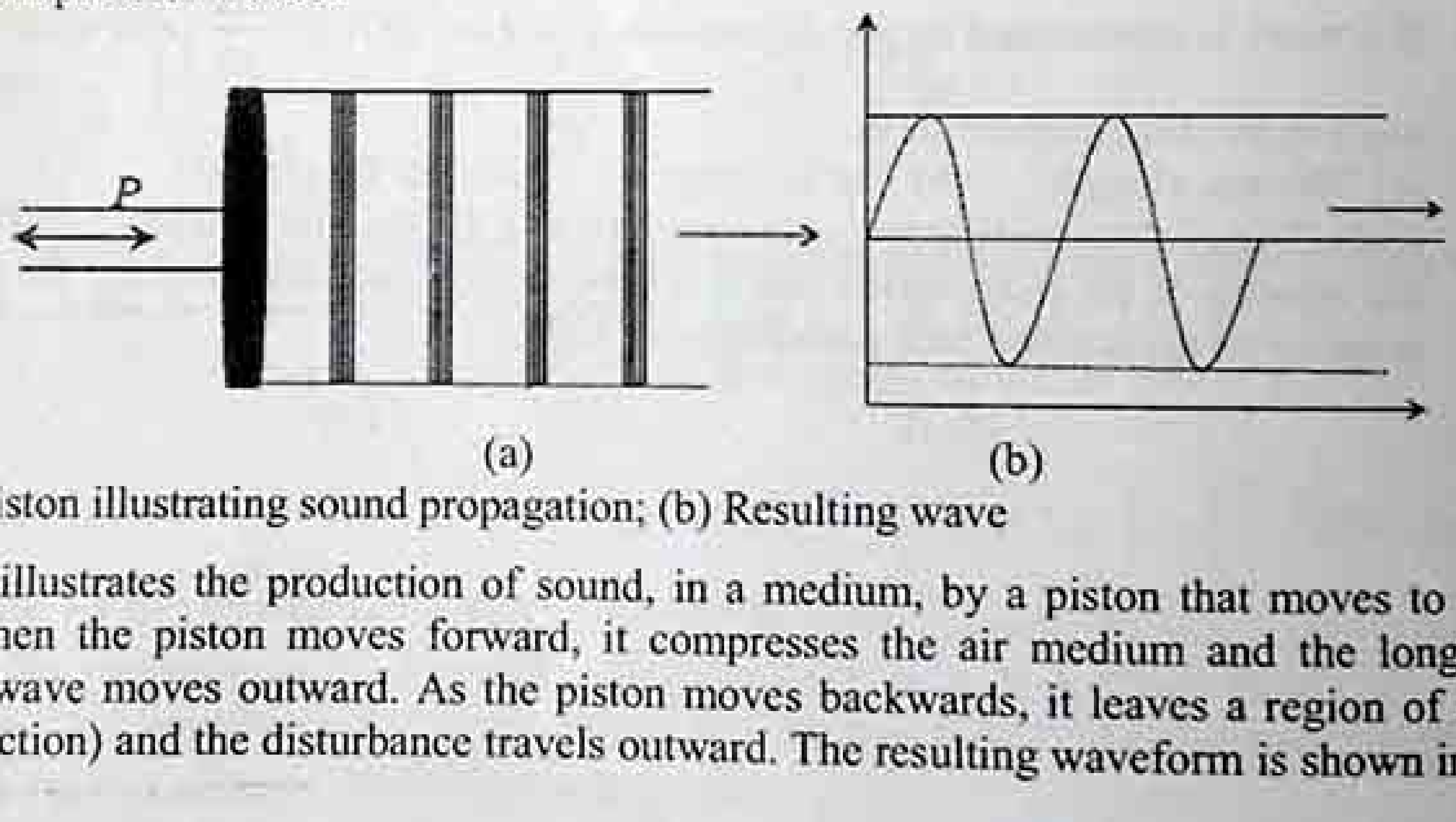
17.0 Introduction

Sound is a term used to describe the form of energy which is transmitted in a fluid (gas or liquid) or solid medium and which produces a sensation that enters and is detected by the car. Sound waves can be classified into two groups

(1) Shock waves which result when a material medium is strained beyond its elastic limit, and
(2) Elastic waves which are the common sound waves are produced when the materials medium still retains its elasticity

Most sounds we encounter result from mechanical vibrations. Examples include the ringing bell, the movement of the mouth and the beating of a drum. In some sources, however, there are no mechanical vibrations involved. Such sources are found in jet airflows, ducts and pipes where the sound produced results from fluid-velocity mixing. In any case, energy is transferred from one particle to another. Sound thus requires a medium to propagate and results in being simply the molecular transfer of motional energy, capable of travelling through any medium. Our discussions will centre on the second group, the elastic waves.

As the atoms or molecules of a medium are displaced from their normal positions, internal elastic restoring forces of stiffness arise. These forces as well as the system inertia enable the particles of the medium to exhibit oscillatory motions and thus transfer energy. Thus at any given time, the particles of the medium are alternately bunched together and spread apart in all directions around the source as far as the wave has propagated. Another way of putting it is to say that the particles of the medium undergo compressions and rarefactions. This is the main reason why sound waves are generally referred to as compression waves.



(a) (b)

Fig. 17.1: (a) Piston illustrating sound propagation; (b) Resulting wave

Figure 17.1(a) illustrates the production of sound, in a medium, by a piston that moves to and fro (oscillates). When the piston moves forward, it compresses the air medium and the longitudinal (compression) wave moves outward. As the piston moves backwards, it leaves a region of reduced pressure (rarefaction) and the disturbance travels outward. The resulting waveform is shown in Figure 17.1(b).

17.1 Characteristics of Sound

In section 16.4, we discussed the generally characteristic of waves. In this section, we shall focus attention on the characteristics that apply to sound waves.

Pressure

Force acting on a molecule over a given area constitutes pressure. When sound wave passes a point in the medium it appears as small fluctuation in the pressure about its mean position. Thus sound waves propagate by variations of fluctuations in pressure about the normal atmospheric pressure. The amplitude is therefore very small. The minimum discernible change in pressure by the human ears is called the *threshold of hearing*, and is 20 micropascal, the maximum pressure referred to as the *threshold of pain* is 20 Pascal.

Intensity

Intensity is the rate of flow of energy per unit area, perpendicular to the direction of propagation of the (sound) wave. It is thus mathematically equal to power, per unit area.

$$I = \frac{Power}{Area} = \frac{Watt}{4\pi r^2} \quad \text{(i.e. for a spherical surface)} \qquad (17.1)$$

Equation 17.1 shows that intensity depends on power and varies inversely with the square of the distance away from the source.

Intensity can also be defined in terms of pressure, such that

$$I = \frac{Peak\ pressure}{2\rho v} = \frac{P^2}{2\rho v} \qquad (17.2)$$

ρ is the density of the medium of propagation and v is the velocity of the sound wave.

Equation 17.2 shows that intensity depends on the pressure as well as the characteristic acoustic impedance of the medium (i.e. the product ρv). Intensity is measured in watts per square meter. Relative intensities or powers are expressed in decibels (*db*) or one-tenth of a *bel*. When there is an intensity or power change from P_1 to P_2, then the number of decibels is $= 10\log_{10}\left(\frac{P_2}{P_1}\right)$.

The minimum change of power detectible by the human ear is 1db, which is equivalent to a power change by 25%. Also, the intensity level of a sound source is its intensity relative to the threshold hearing, which corresponds to an intensity level $P_0 = 10^{-12}Wm^{-2}$. This is the zero intensity level. The difference in intensity levels of two sound sources of intensities P_1 and P_2 is

$$10\log_{10}\left(\frac{P_1}{P_0}\right) - 10\log_{10}\left(\frac{P_2}{P_0}\right) = 10\log_{10}\left(\frac{P_1}{P_2}\right) \qquad (17.3)$$

Similarly, as intensity obeys the inverse law, if d is the distance between a source of sound and an observer, then the difference in intensity levels at two points d_1 and d_2 from a sound source is $10\log_{10}\left(\frac{d_2^2}{d_1^2}\right)$.

Loudness

The loudness of a tone is determined largely by its intensity. This does not mean that intensity and loudness are the same. Loudness is the way in which an individual perceives the intensity of a note at a particular frequency. Loudness depends on the amplitude of vibration and the surface area set into vibration. It cannot be measured directly except by comparing it with a standard source of frequency of 1000*Hz*. The unit of loudness is the *Phon*, which is the loudness at the threshold of hearing.

Power

The amount of power generated by a source of sound is measured in watts, using 10-12*W* as the reference. This is the power that corresponds to the threshold-hearing intensity. The power output of most sources of sound is quite small. For example, the average power of human speech in conversion is only 20μW.

In electricity, the reference power is zero whereas in acoustics (sound) it is 10-12*W*. This explains why one can safely touch the bare wires carrying sound from a 10*W* amplifying device but would not dare do the same to wire carrying current from a mere 2*W* electrical device.

Timbre

Timbre is a measure of the quality of sound of a musical tone. This makes it possible for human beings to distinguish between two tones of the small intensity level and the same frequency but different waveforms. It expresses our ability to recognize the sound of a child as distinct from that of a young girl, for example. It also enables us to recognize a musical note from a violin or a guitar or a flute. Timbre is thus a complex function which depends on waveform, frequency and intensity.

Pitch

Pitch, like timbre is a complex characteristic of sound. No absolute definition has been assigned to it rather it is usually loosely defined in terms of the change in frequency. If the loudness of a tone remains constant, then pitch changes linearly with frequency. For such a condition, a high frequency note has a high pitch and a low frequency one has a low pitch.

Reflection of Sound Waves

Sound waves like any other wave motion can be reflected whenever they encounter an obstruction or discontinuity in the medium of propagation. All the laws applicable to light waves at plane and curved surfaces are similarly applicable to sound waves. Echoes and Reverberations are the results of reflection of sound waves. Echoes have several practical applications such as:

(i) The determination of velocity of sound in air employs the principle of the echo. If one stays a reasonable distance away from a wall and claps his hands, the sound of the clap will travel to the wall, a distance of d meters, and return as echo in time t sounds. Thus the velocity v of the sound in air is $v = 2d/t$.

(ii) Echo is used in medical diagnosis, where sound waves are sent to the desired part of the body and the echo is analyzed.

(iii) The same method is employed in determining the depth of a sea, an ocean or a river. This is known as echo sounding.

(iv) Echo sounding is also used in prospecting for minerals such as hydrocarbons. The returning sound wave will carry the characteristic of medium through which it had travelled, and these characteristics will, on analysis, reveal the contents of the depths as well the depth or distance, once its velocity and duration are known.

Reverberation is the persistence of sound waves in an enclosure after the source has ceased. This effect is mainly due to

(i) direct sound waves from the source;

(ii) echoes produced by walls and ceilings;

(iii) Sound waves diffused from wall, ceilings and other objects within the enclosure.

Sabine's investigation into the problem of reverberation shows that the time of reverberation is given as

$$T = \frac{\kappa V}{\alpha A} \qquad (17.4)$$

where V = volume of the enclosure; α = absorptive power of all surfaces; A = surface area of the enclosure and κ = a constant.

Reflection in air columns

Sound waves can be reflected at either the open or the closed end of a long narrow tube or air column. At the closed end, molecules cannot move freely, so the node appears at that end. At the open end, the molecules can move freely, so the open end is the antinodes of the standing wave. Figure 17.2 shows the longest wavelength corresponding to the smallest or fundamental frequency of the sound wave in the closed tube.

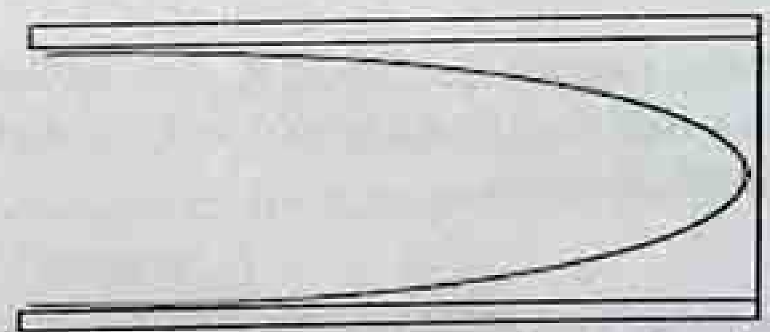


Fig. 17.2: Wave Pattern in a closed tube

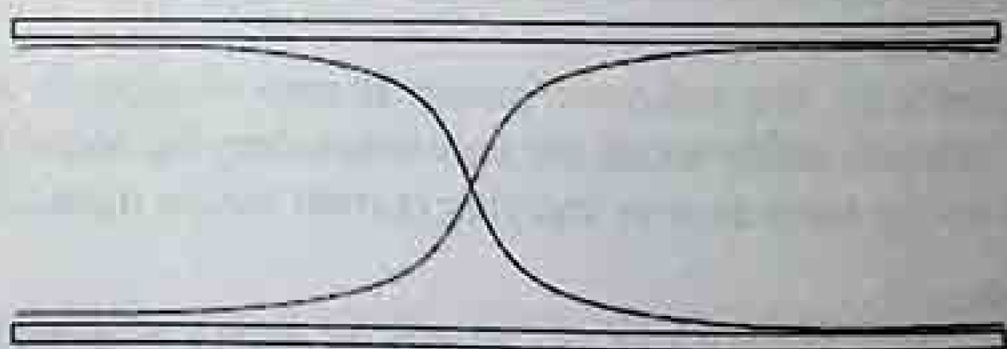


Fig. 17.3: Wave Pattern in an Open tube

In Figure 17.2, we can observe that the length of the tube is equal to one-quarter of the wavelength i.e. $l = \frac{1}{4}\lambda$ or $\lambda = 4l$.

Thus using equation 17.2, the fundamental frequency is

$$f_0 = \frac{v}{4l} \tag{17.5}$$

The above equation shows that frequency increases with decreasing length. Thus generally, for an air column closed at one end, the frequency is given as

$$f_{n-1} = \frac{v(2n-1)}{4l} \tag{17.6}$$

where $n = 1, 2, 3, 4,$ ....

When these values of n are put into equation 17.6 above, it becomes clear that only odd multiples (harmonics) of the fundamental frequency are possible for this type of air column.

For a column with the two ends open, there are antinodes at each (open) end. Thus from Figure 17.3, we observe that the relationship between the length of the column and the wavelength is,

$$\lambda_{n-1} = \frac{2l}{n} \tag{17.7}$$

Consequently, the fundamental frequency for this type of air column, when $n = 1$, is

$$f_0 = \frac{v}{2l} \tag{17.8}$$

When these values of $n = 1, 2, 3, 4$ ... are put into equation 17.7 above, it becomes clear that all multiples (harmonics) of the fundamental frequency are possible for this type of air column. The harmonics of fundamental frequencies are sometimes referred to as *overtones*, especially in music circles. Other corresponding lengths of tubes and lengths are shown in Figure 17.4.

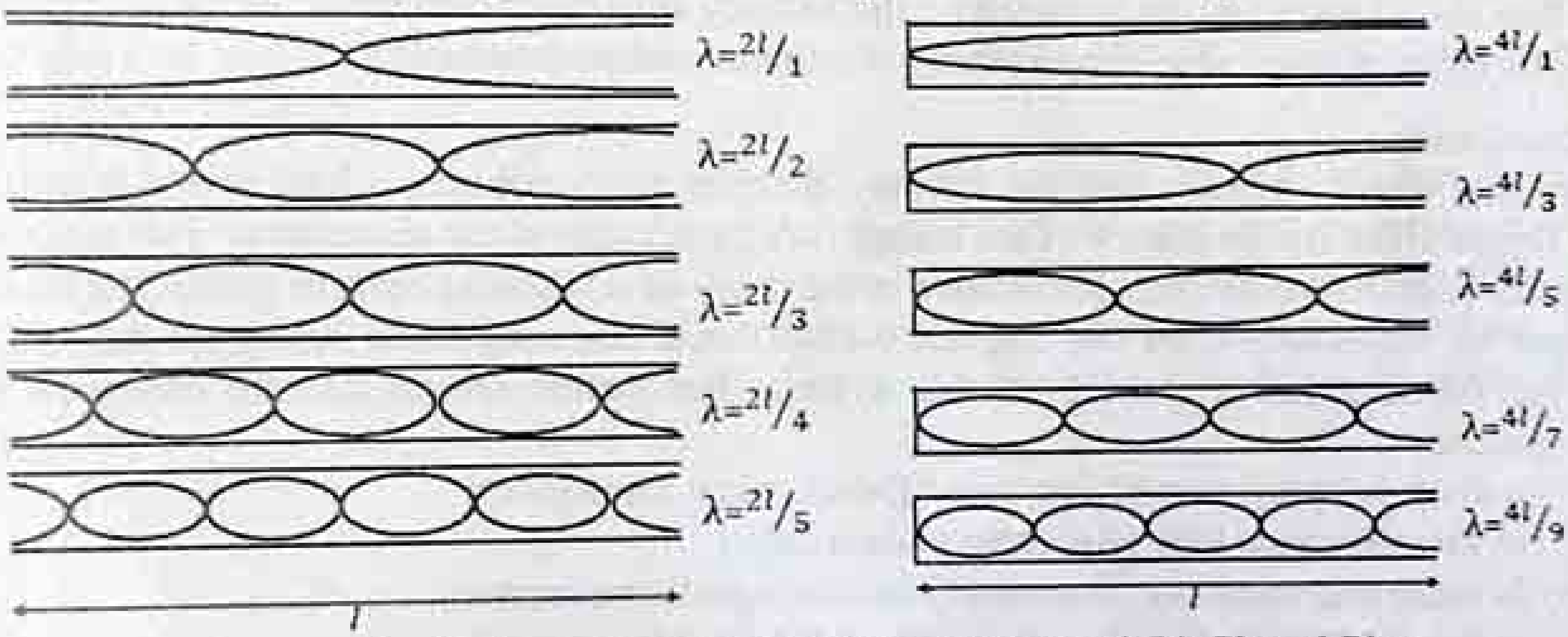


Fig. 17.4: Wave Patterns for higher harmonics in (a) Open Pipe and (b) Closed Pipe.

17.2 Reflection in Strings

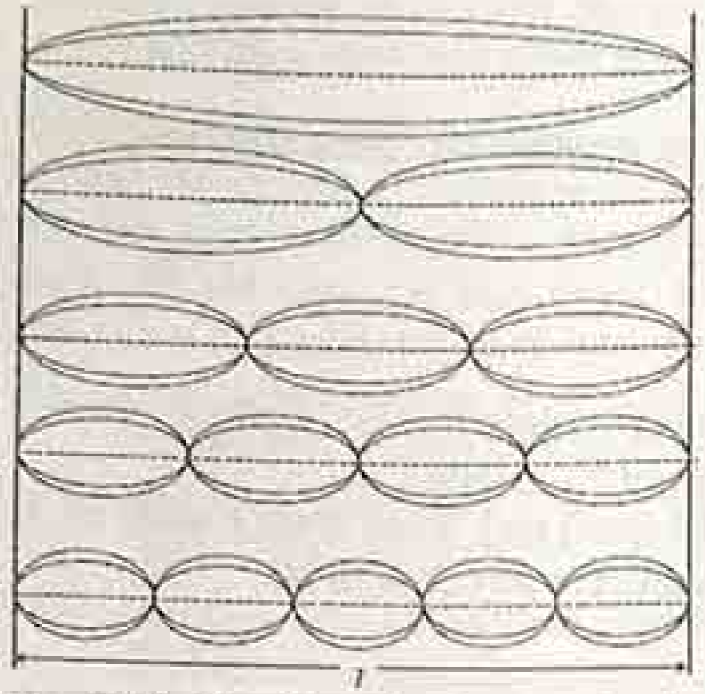
Consider a string of length l rigidly fixed at both ends, it is obvious that only nodes at the ends of the string can be produced by standing waves due to reflection at the ends. When a string fixed at both ends is plucked, transverse stationary waves are set up along the string. The simplest mode of vibration is set up when the stretched string is plucked in the middle. Figure 17.5 shows the standing waves on a string. The relationship between the length of the string and the wavelength is given by the equation

$$\lambda_{n-1} = \frac{2l}{n}; \text{ for n} = 1, 2, 3, \ldots \tag{17.9}$$

and the corresponding frequencies, called *eigenfrequencies* or *overtones*, are found from the equation

$$f_{n-1} = \frac{v}{\lambda_n} \tag{17.10}$$

The corresponding frequencies reveal that all harmonics or overtones are possible in string instruments. This explains why more melodious notes can beplayed using the guitar (string instrument) than by using a flute (air column instrument).



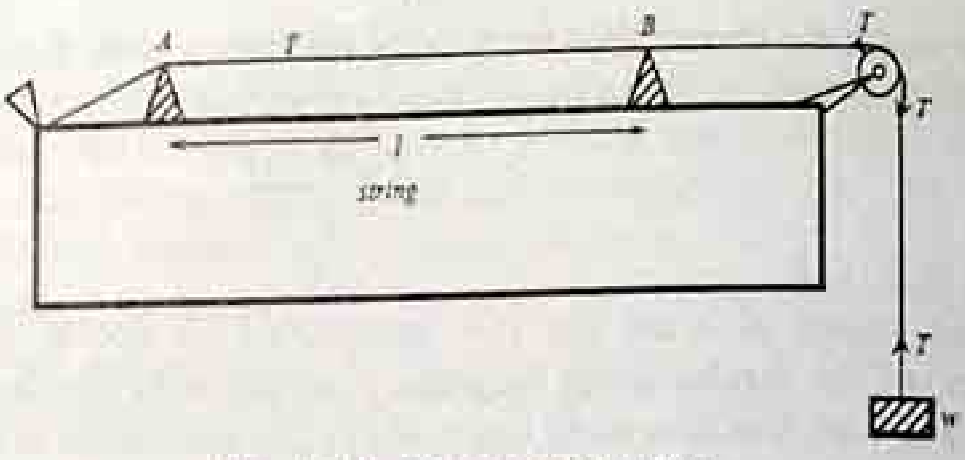


Fig.17.6: Sonometre setup

Fig. 17.5: The first five harmonics for standing waves on a stretched string fixed at both ends

The velocity of sound on a string is found to depend on the tension T, in the string and the mass per unit length μ (linear density) of the string. This is expressed as

$$v = \sqrt{\frac{T}{\mu}} \qquad (17.11)$$

The relationship in equation 17.11 is easily verified by the use of the Sonometer shown in Figure 17.6. The tension along the string is varied by the hanging of weights and the linear density varied by using various types of wires (e.g. copper, steel, aluminum, and constantan).

17.3 Resonance

It is known that when a vibrating tuning fork is held over an air column, a loud sound is heard at a particular length when the frequency of the tuning fork equals that of the air column. This coincidence of frequencies is called *resonance*. Resonance is the result of a vibrating system producing its largest displacement or amplitude when the frequency of external vibrating force acting on the system is equal to the fundamental frequency of that system. Resonance effects can be observed in the following systems:

(i) a diver jumping on a diving board (Mechanical resonance);
(ii) in the resonance tube (Acoustic resonance);
(iii) in radio and transistor receivers (Electromagnetic resonance);
(iv) in the absorption or emission spectrum (Optical resonance).

It is employed effectively in musical instruments, such that if air is blown into a flute or a clarinet, for example, and the holes of the musical instrument are blocked in turn at various points, musical notes result when resonance occurs.

Example 17.1

A violin string vibrates with a fundamental frequency of $420Hz$. What are the frequencies of its first two overtones?

Solution

$f_0 = 420\,Hz$

First overtone, $f_1 = 2f_0 = 2\times 420 = 840\,Hz$

Second overtone, $f_3 = 3f_0 = 3\times 420 = 1260\,Hz$

Example 17.2

What are the frequencies of the first three harmonics of a piano string of length $2.0m$ if the velocity of the wave on the string is $140\,ms^{-1}$?

Solution

$$f_0 = \frac{v}{2l} = \frac{140}{2\times 2} = 35\,Hz.$$

Since all harmonics are produced, the second and third harmonics are 70 and 105Hz respectively.

Example 17.3

An empty bottle can resonate by blowing across the top. If the bottle is 25cm deep, what would be the fundamental frequency? The speed of sound in air is $340\,ms^{-1}$.

Solution

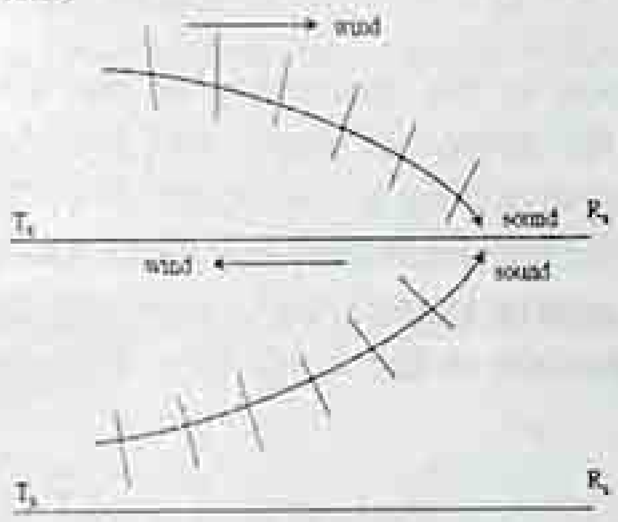
The empty bottle is a container with one end closed. So, $\lambda = 4l$

$$f = \frac{v}{\lambda} = \frac{340}{4\times 0.25} = 340\,Hz$$

17.4 Refraction

Sound waves are refracted (i.e. the bending of the path of wave) when they travel across the interface of two media of different densities. Recall that the velocity of sound in air is proportional to the square root of the absolute temperature. Under normal conditions, the temperature of air decreases with increasing height above the ground, thus the velocity of sound is faster near the ground. In the day time, the upper layers of the atmosphere are cooler and hence denser resulting in slower movement of sound waves, while the lower regions, which are hotter and less dense, enhance faster movement of sound waves. Consequently, sound waves are refracted away from the earth's surface. However, at night time due to temperature inversion, sound waves are refracted towards the earth and therefore travel longer distances along the surface of the earth. This explains why sound is heard far from the source at night.

Sound can be refracted upwards, in travelling from one layer of air into another, later it can be refracted down. In between these refractions, the sound can disappear, leaving a region of silence in the middle. During inversion, when air higher up is warmer than the one near the surface of water, sound will be refracted downward and will then travel long distances along the surface. This is one mode by which earthquakes travel. Figure 17.7 illustrates the refraction of sound waves due to wind motion.



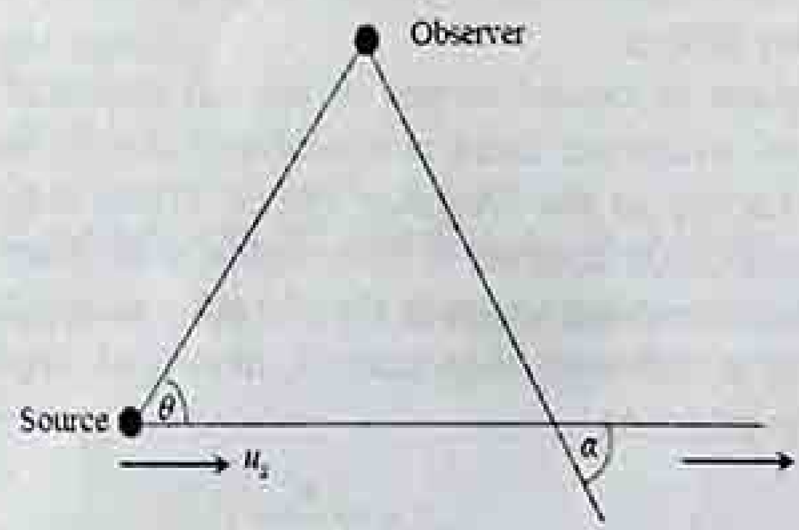


Fig. 17.7: Refraction of Sound

Fig. 17.8: Source moving at angle θ to line joining source and observer.

17.5 Interference

When sound waves from coherent sources (i.e. sound waves of same frequency, amplitude and having some constant phase relationship) travel in the same medium, they interfere either constructively or destructively according to the principle of superposition of waves. At any point along the direction of propagation where the path difference is zero or integer multiples of the wavelength, there is permanent loud sound. At such points the waves are in phase, having their initial frequency and four times the intensity of any of the sources. However, for path differences which are half-wavelength or odd-multiples of half-wavelength, the waves interfere destructively. The waves are 180° out of phase and there is permanent silence or zero displacement at such points.

17.6 Diffraction

Diffraction is another effect of superposition of waves. This is the interference between waves from coherent sources on the same undivided wavefront. This phenomenon is easily observed where the dimensions of the obstruction are comparable to the wavelength of sound. Sound in air has a velocity of about $344ms^{-1}$, while its frequency range is $20Hz$ to $20kHz$. This means that its lowest wavelength is $1.72cm$ while its highest is $17.2m$. Consequently, the lower frequencies of sound will be diffracted more in air where their wavelengths are comparable to the sizes of everyday objects that may be in the path of sound. The effect of the diffraction of sound is readily observed in public address systems, where horn loudspeakers are used. This is because both speech and music comprise a wide range of frequencies. Diffraction allows the lower frequencies to spread out and be heard over a wide area whereas the higher frequencies are not heard beyond the axis of the loudspeaker. Listeners near the principal axis of the horn hear all the sound fairly well reproduced whereas those far from the axis hear only the low frequency components. Diffraction makes it possible to detect sound waves around obstructions e.g. across walls and around corners of buildings.

17.7 Beats

Beats are produced when two notes or wave trains of nearly equal frequencies f_1 and $f_2 (f_1 > f_2)$ but equal amplitude travel in the same direction at the same velocity are superposed. A rise and fall in intensity of sound is heard. The beat frequency $(f_1 - f_2)$ is the frequency at which the amplitude pulsates. The beat effect is another unique interference effect of sound. Often when a string is disturbed or standing waves are set up in air column, both the fundamental frequency and harmonics (or overtones) are present, the shape of the waveform becomes complex. The interference of two waves with almost identical frequencies produces the phenomenon of beats. Figure 16.10 shows two travelling waves with slightly different frequencies. At points B and D waves add constructively and at the intermediate point C, the waves add destructively. This phenomenon is called *beating*. The frequency with which the nodes pass given points on the x-axis is called the beat frequency $(f_1 - f_2)$.
When two tones, $400Hz$ and $390Hz$, are played together, their resultant mixture is a wave with varying amplitude and pitch. It has a beat frequency of $400Hz$ minus $390Hz$ which is $10Hz$.
Skilled musicians employ the beat phenomenon to tune their instruments; in fact, the presence of beats between two notes enhances the pleasure of listening to musical notes.

17.8 Doppler Effect

In the propagation of sound waves in air, an apparent alteration in the frequency of the sound is always observed whenever there is a relative velocity between the source of sound and the observer. This effect is known as the *Doppler Effect*. When a train is approaching a stationary observer, the frequency of its whistle appears to increase. A stationary observer receives more crowded wavefronts when the source is moving towards the observer than when the source is stationary. This results in an apparent shorter wavelength than normal, hence, an apparent increase in the frequency. The apparent frequency f' is given as

$$f' = \frac{v}{v - u_s} f \tag{17.12}$$

where v is the velocity of sound; u_s is the velocity of the source and f is the original frequency. Similarly, when the source recedes from the observer, the wavefronts are farther apart than when the source is stationary. Hence the observer receives fewer waves per second and thus the apparent frequency is therefore lower. The apparent frequency f' in this situation is given as

$$f' = \frac{v}{v + u_s} f \tag{17.13}$$

For an observer approaching a stationary source with a velocity u_o, the relative velocity of the sound waves is $v + u_o$ and since the wavelength of the sound is not affected, the result is an apparent increase in the frequency given by the expression

$$f' = \frac{v+u_o}{v} f \quad (17.14)$$

When the observer moves away from a stationary source, he receives sound waves with reduced relative velocity $(v+u_o)$. As a result of this, there is an apparent decrease in the frequency of the received waves given by the expression

$$f' = \frac{v-u_o}{v} f \quad (17.15)$$

Finally, the movement of both observer and source at the same also brings about the Doppler Effect. When the source and the observer are moving in the same direction, but with different velocities (u_s as the source velocity), the apparent frequency is given by

$$f' = \frac{v-u_o}{v-u_s} f \quad (17.16)$$

This will result in either an increase or decrease in the frequency depending on the velocities of the source and observer. However, when the source and observer are moving in the opposite directions, there is an apparent increase in frequency given by

$$f' = \frac{v+u_o}{v-u_s} f \quad (17.17)$$

The velocity and the direction of wind relative to that of a source of sound can also bring about Doppler Effect. In all the cases discussed above, the effect of wind is represented by replacing the wave velocity v with $(v+u_w)$ when the wind is in the direction of the observer and by $(v-u_w)$ when the wind is blowing away from the observer. Similarly, when the source is moving at an angle to the line joining the source and the observer, as shown in Figure 17.9, the apparent frequencies are expressed as follows

$$f' = \frac{v}{v-u_s\cos\alpha} f \text{ and } f' = \frac{v}{v+u_s\cos\theta} f \quad (17.18)$$

Note that generally, the motion of the observer affects only the velocity of the waves reaching the observer, while the motion of the source affects only the wavelength of the waves received by the observer. When the source and the observer both move in the same direction, with equal velocities (i.e. $u_o = u_s$), then the Doppler Effect does not occur.

Example 17.3

A police siren has a frequency of $800Hz$. If the velocity of sound in air is $344ms^{-1}$, determine the frequencies heard by drivers of vehicles moving at $14ms^{-1}$, (i) towards the siren and (ii) away from the siren.

Solution

(i) $f' = \frac{(344-14)}{344} \times 800 = 767.44Hz$

(ii) $f' = \frac{(344+14)}{344} \times 800 = 832.56Hz$

17.9 Velocity of Sound

The speed with which a wave travels in a medium is determined by the quickness with which the molecules transfer energy from one to another. This quickness is in turn determined by the strength of the force among molecules. These forces are characterized by *elasticity* and *density* of the medium, as well as the surface tension and gravitational forces. These characteristics make the speed of sound differ in air, water, blood or solids, for example. When disturbances are large, other factors such as phase, attenuation and absorption affect the forces and thus the velocity of sound as well. Our dissuasions on sound will assume small disturbances only. The velocity v of sound in air is dependent on the modulus of elasticity E and the density ρ of the medium. This was first obtained by Newton in the form $v=\sqrt{E/\rho}$; where E the isothermal bulk modulus and is equal to the pressure P of the gas. However, Laplace later corrected the above relation by postulating that the bulk modulus of elasticity is the adiabatic bulk modulus. Hence, the velocity of sound in air or gas is given by the expression

$$v=\sqrt{\frac{\gamma P}{\rho}} \tag{17.19}$$

where γ is the ratio of the principalspecific heats of the gas and ρ is the density of the gas.
The change in pressure is assumed to be so rapid that no heat transfer takes place. In other words, the process is adiabatic for the small disturbance.
From Boyle's law for gases,

$$PV = MRT \tag{17.20}$$

where V = volume of the gas, M = molecular weight, T = the absolute temperature and R = the gas constant.
Inserting equation 17.20 into equation 17.19, the later becomes

$$v=\sqrt{\gamma RT} \tag{17.21}$$

Table17.1: Velocity of Sound in various materials

Materials	Speed at 20°C (ms^{-1})	Density (kgm^{-3})
Air	344	1.05
Lead	1200	
Hydrogen	1270	0.0899(0°C)
Vulcanized rubber	54	
Ethyl alcohol	1210	791.00
Pure water	1480	1000.00
Blood	1570(37°C)	1056
Aluminum	5100	2700
Glass	5600	2500

The above shows that the velocity of sound in a gas medium is of pressure at constant temperature of the medium, but is proportional to the square root of the absolute temperature. Generally, the equation of the velocity of sound in any medium is given as

$$v=\frac{K}{\rho} \tag{17.22}$$

For liquids, $K = B$ (bulk modulus), and for solids $K = E$, the Young's modulus of elasticity.
This velocity is the speed with which each wave crest travels. It is thus the speed of the movement of the phase position of the wave. It is therefore a *phase velocity* wave. Table 17.1 shows a list of velocities of sound in different media together with their corresponding densities. Generally, sound travels faster in solid than in fluids.

Example 17.4

The velocity of sound in a medium at a temperature 20°C was $244ms^{-1}$, determine the velocity when the temperature changes to 45°C.

Solution

Since $v \propto \sqrt{T}$; where T is the absolute temperature, then $\frac{v_1}{v_2}=\frac{\sqrt{T_1}}{\sqrt{T_2}}$

Thus $v_2 = \dfrac{344\times\sqrt{318}}{\sqrt{293}} = 358\,ms^{-1}$

Particles Velocity

Recall that the particles of a medium in which sound propagates do not travel with the wave but rather vibrate about their mean positions. We observe their phase relationships as waves. Thus the velocities, u, of these particles are those of simple harmonic motion as distinct from the phase velocities.

The equation of a progressive wave is usually given in term of its displacement, as we saw in equation 16.3.

$$y = A\sin(\omega t - kx) \tag{17.23}$$

differentiating with respect to t

So that velocity, u is given by

$$\frac{dy}{dt} = u = \omega A\cos(\omega t - kx) \tag{17.24}$$

This particle motion depends on the impressed pressure, and the relationship between pressure and velocity

$$P = \rho V u \tag{17.25}$$

Thus,

$$\frac{P}{u} = \rho V = Z \tag{17.26}$$

The above is referred to as the characteristic impedance of the medium and it determines how sound waves can propagate in the medium.

Example 17.5

Sound pressure of $24.9\,Nm^{-2}$ was impressed on a fluid medium and its particles velocity was found to be $6cms^{-1}$. Determine the characteristic impedance of the fluid.

Solution

$$Z = \frac{P}{U} = \frac{24.9}{0.06} = 415\,ray/s$$

Summary

Sound waves are produced as a result of mechanical disturbances in materials. The speed of sound depends on the modulus of elasticity of the medium, which in turn depends on the elastic properties of the medium. Sound wave has velocity which is proportional to the square root of the absolute temperature. Other properties of sound includes: reflection which results in echo, refraction, interference, diffraction, Beat phenomenon and Doppler effect.

Exercise 17

17.1 When a periodic wave is produced on a stretched string, which of the following factors influences the speed of the wave?
A. Tension; **B.** Wavelength; **C.** Frequency; **D.** Linear density of the string.

17.2 An intensity level of 0 dB is equivalent to a sound intensity of
A. $1.0Wm^{-2}$ **B.** $1.0\times10^{-4}Wm^{-2}$ **C.** $1.0\times10^{-12}Wm^{-2}$ **D.** $0Wm^{-2}$

17.3 What happens to the velocity of sound in an ideal gas when the absolute temperature of the gas is doubled?
A. It increases to 1.4 times its original value; **B.** It doubles; **C.** It quadruples; **D.** None of the above.

17.4 An observer approaches a stationary source of sound emitting frequency of 444*Hz* at one-half the speed of sound. What is the frequency of the sound received by the observer?
A. 666*Hz*; **B.** 222*Hz*; **C.** 888*Hz*; **D.** 444*Hz*.

17.5 A standing sound wave in a tube opened at one end has a displacement ______ at the open end and a displacement ________ at the closed end.
A. node, node; **B.** node, antinode; **C.** antinode, antinode; **D.** antinode, node.

17.6 Given that the lowest frequency of standing waves that can exist in a tube opened at both ends is $250Hz$, what is the length of the tube?
A. $0.34m$; B. $0.69m$; C. $1.4m$; D. $2.8m$.

17.7 Standing waves are produced by
A. the superposition of identical waves travelling in the same direction;
B. the superposition of waves which travel with different speeds;
C. the superposition of identical waves travelling in opposite directions;
D. the superposition of otherwise identical waves of slightly different frequencies.

17.8 When the tension in a wire is tripled, what happens to the speed of waves on the wire?
A. It triples; B. It is reduced by $\frac{1}{3}$; C. It is increased by $\sqrt{3}$; D. It is reduced by $\sqrt{3}$

17.9 A wave is described by the expression $y = (2.0cm)\sin[2\pi ft - (15m^{-1})x]$ and is travelling with a speed of $15ms^{-1}$, what is the frequency of the wave?
A. $0.50Hz$; B. $3.1Hz$; C. $6.3Hz$; D. $36Hz$.

17.10 Why can't a tube closed at one end play all tunes whereas the guitar can?

17.11 Why are our radio receivers able to get many distant stations at night but very few during the day?

17.12 Two waves of frequencies $250Hz$ and $280Hz$ are sounded together. Describe what a listener hears.

17.13 What are the chief characteristics of a progressive wave? What conditions must be satisfied in order that stationery - wave system may be obtained.

17.14 Give one example of musical instruments that emit only odd harmonics; give one example of those that emit all members of the harmonic series.

17.15 Explain (a) Doppler effect (b) Resonance (c) Longitudinal and Transverse vibrations.

17.16 Find the velocity of the source along the line joining the source to a stationary observer if, as a result of the motion, the frequency of the note heard is (a) increased in the ratio 18 : 15; (b) decreased in the ratio 15 : 18, with reference to the true frequency. Take the velocity of sound in air to be $345ms^{-1}$.

17.17 Taking the speed of sound in air at 0°C to be $330ms^{-1}$, find the tube-lengths at 27°C for the first two resonance positions for a turning fork of frequency $256Hz$.

17.18 A string has a length of $3m$ and a mass of $24g$. If this string is subjected to a tension of $500N$, what is the speed of transverse waves?

17.19 A siren produces $100W$ of sound. If the energy spreads out spherically, determine the intensity of the sound at a distance of $100m$.

17.20 Lightening flash was seen by an observer and the accompanying thunder was heard $6s$ later. If the velocity of sound was $344ms^{-1}$, find the approximate distance of the observer from the storm cloud.

CHAPTER 18
SOURCES OF SOUND AND SOUND DETECTORS

18.0 Introduction

Any material that can vibrate is capable of producing sound. Thus there are many possible sources of sound. The sources include mechanical, electrical, electroacoustic and electromechanical devices as well as human beings. The sources are natural as well as man-made, and they all have two things in common; each has a vibration-producing mechanism and a resonant structure. We shall see this later. Among the natural sources are the human speech organs and the human heart-beat. Others include wind and rainfall. The man-made sources are numerous. They can be grouped into mechanical (like music instruments), electrical (such as radio sets, fan, power-generating machines) and electro-acoustic (which include loudspeakers). Let us examine how some of these operate.

18.1 Sound Sources

18.1.1 Musical Instruments

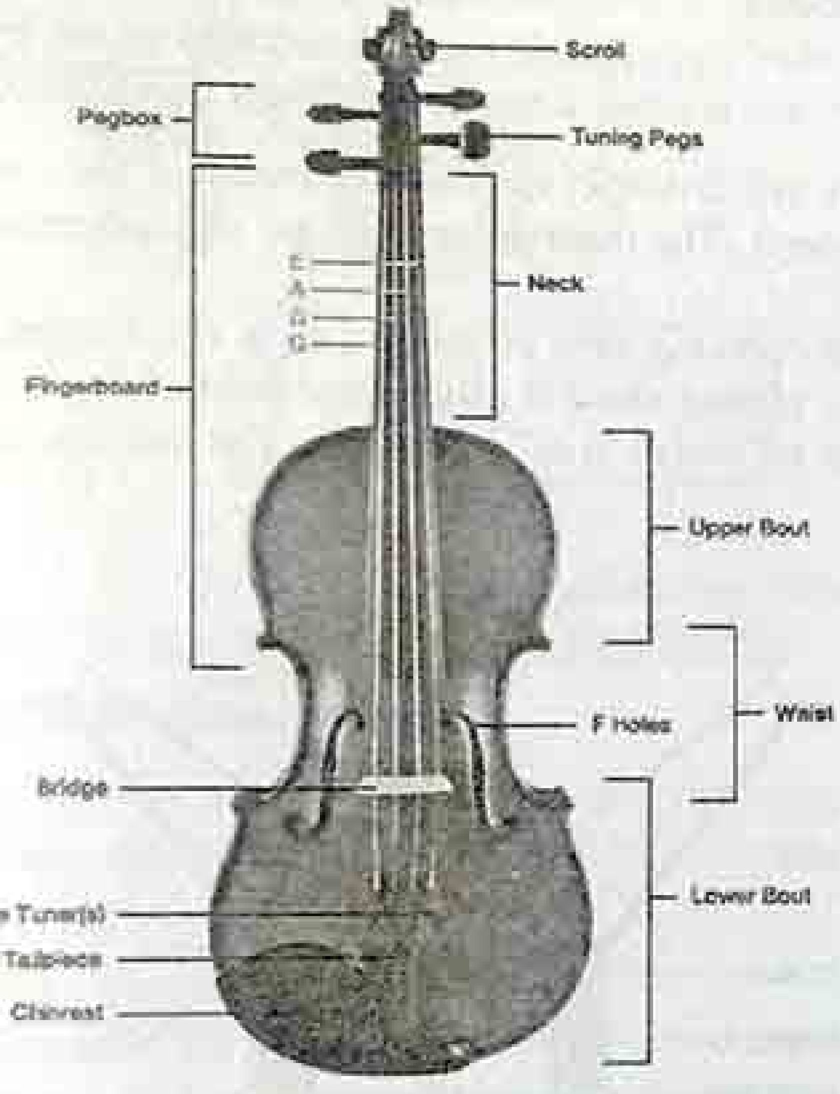


Fig. 18.1: A front view of the violin

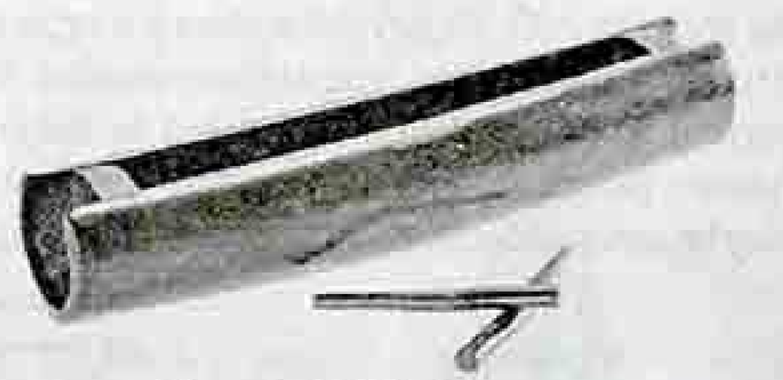


Fig. 18.2: The Gong

Musical instruments include string types like the guitar, membranes such as drums and gongs, and wind instruments like flute and the trumpet. Every of these consist of some sources of vibrations as well as some arrangements for transmitting energy the air with reasonable efficiency. In *stringed instruments*, the basic frequency depends on the length of the string L, on the tension T, put on it and on the mass (per a unit length) m of the string material. Equation 18.1 shows the relationship between the frequency and the different parameters.

$$f = \frac{1}{2L(T/m)^{1/2}} \qquad (18.1)$$

The vibrating systems are connected to resonating systems whose responses to the various harmonics are what we hear as music. Figure 18.1 shows a violin with this complex resonant structure. As the violin strings are bowed or plucked, they vibrate and transfer these vibrations to the body through the bridge.

In the case of the drum or gong, the resonant structure is the hollow cavity; it reinforces the sound of the beating. A typical gong is shown in Figure 18.2.

For the wind instruments, the vibration of frequency depends on the hole opening along the length of the pipe. On the whole, musical instruments present a variety of arrangements for producing sound.

18.1.2 The Human Speech Organs

The human speech is one natural sources of sound. Figure 18.3 shows the essential components of the human speech mechanism. Air from the lungs, resulting from chest muscle contraction, is set into vibrations by vocal cords at the glottis. The nose, throat and mouth cavities modify these vibrations and thus act as resonant structures. Many different speech sounds are possible because the tensions of the vocal cords can be varied. Besides, the shapes of cavities of the nose, throat and mouth can be altered at will. Speech sounds occupy a range of frequencies from 16*Hz* to 16*kHz*, although research has revealed that the really useful range is 300*Hz* to 400*Hz*. In conservation, the average power of speech varies from less than one microwatt (whisper) to one milliwatt (shout).

18.1.3 The Loudspeaker

One electro-acoustical source of sound is the loudspeaker. It is often mistaken for a sound detector which it is not. Sound that has been modified by electronic systems is reconverted into sound by the loudspeaker. Since it thus converts one form of energy into another form, it is called a *transducer*.

There are three basic types of loudspeakers; these are electrostatic, the moving iron and the moving coil. There are others, all based on the three basic groups. The most common type of loudspeaker is the moving coil. Figure 18.4 shows a typical one.
It uses a permanent magnet surrounded by a current-carrying coil connected to a diaphragm. The current of an amplified signal fed to the coil makes it vibrate since it also in the field of bar magnet (motor principle). As it vibrates, so does diaphragm to which it is connected. The motion of the diaphragm produces sound waves in the air.

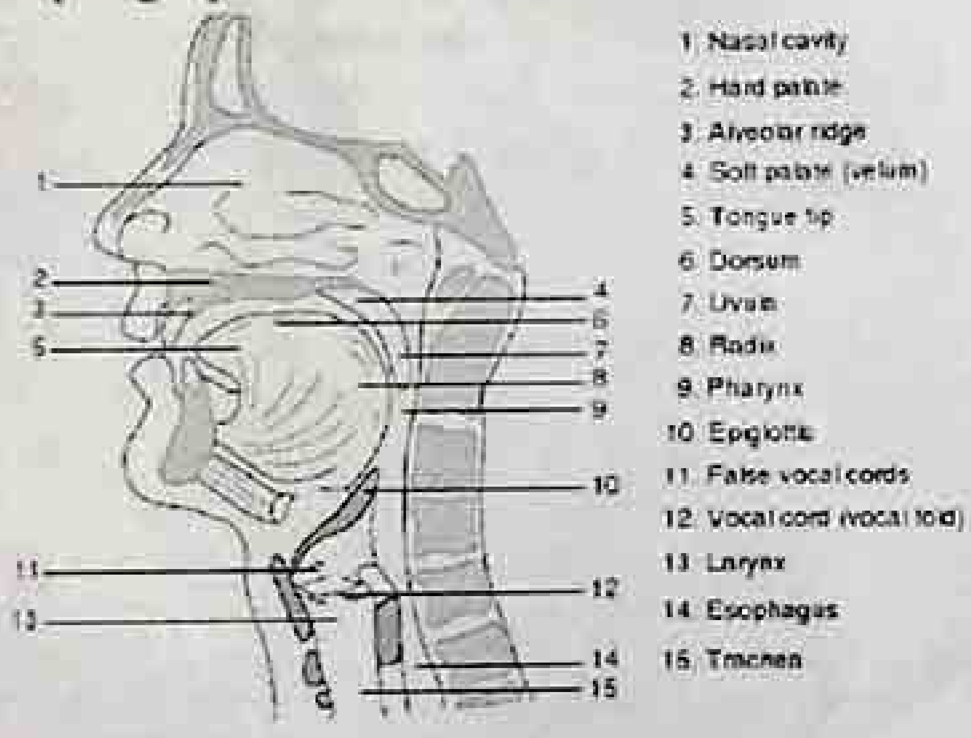


Fig. 18.3: Components of the speech mechanism

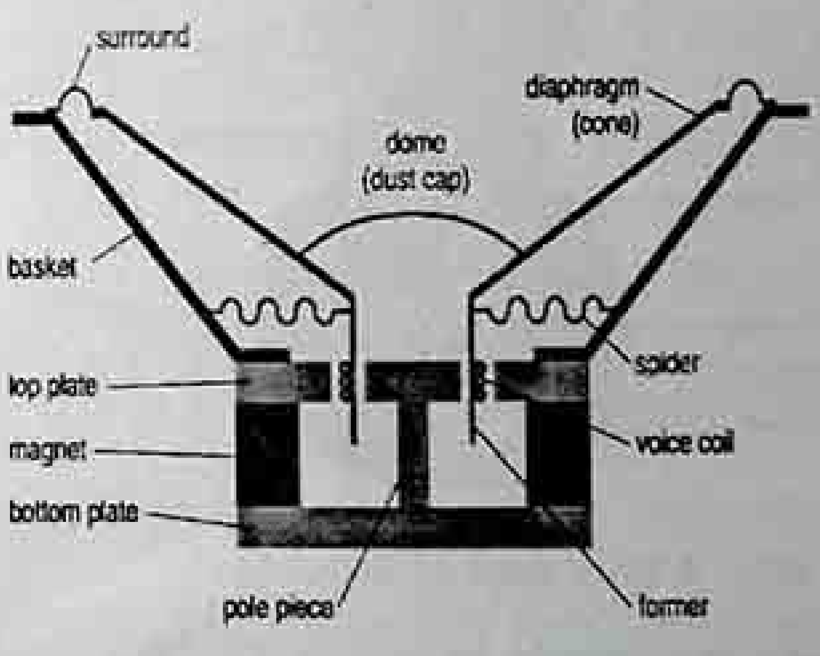


Fig 18.4 Moving coil loudspeaker

18.2 Sound Detectors

18.2.1 Microphones

Microphones are electroacoustic transducers which convert acoustic energy into electrical one. The conversion is very important where amplification of sound is contemplated. The electrical energy equivalent of sound is amplified, after detection and made to drive a loudspeaker which then converts this equivalent into (bigger amplitude) natural sound.
There are types of microphone, and they drive their classifications largely from their modes of operation. They include the condenser, moving coil, velocity ribbon, crystal bimorph and carbon granule. It does not matter what type a microphone is, it must respond accurately to vibrations in the frequency and intensity of sound input. We shall discuss, here, the detection of sound by carbon granule type, and highlight the principles of operation of others.

18.2.2 Carbon Granule Microphone

Figure 18.5a shows a typical carbon granule type of microphone. It contains carbon granules between two carbon blocks. One block is connected to a diaphragm which is conical and protected from damage by a protective plate (perforated).

When a person speaks into the microphone, sound waves impinge on the diaphragm and make it vibrate. It moves very slightly to and fro in accordance with the varying pressure of the input speech waves. As it moves towards the particles are loosened. The movements are rather like the one described in section 17.1 executed by the piston. The outcome is a variation in the resistance of the carbon granules which then result in a varying current flow (Figure 18.5b). The magnitude of this current flow is altered at exactly the same frequency as the sound pressure variations that started the movements. Thus sound energy is detected and changed into electrical energy.

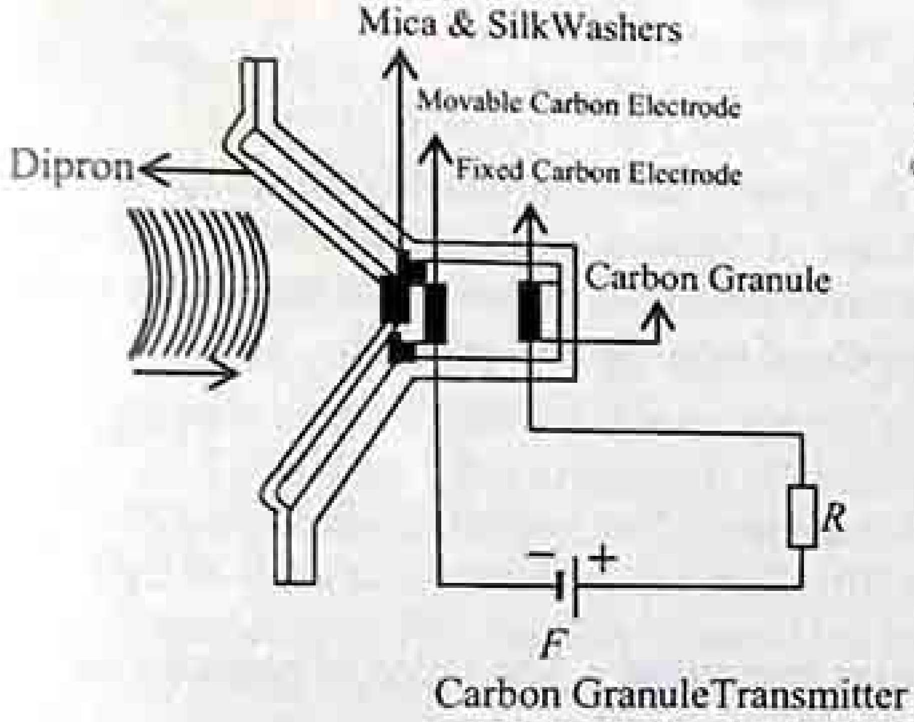


Fig. 18.5a: Simple principle of carbon granule microphone

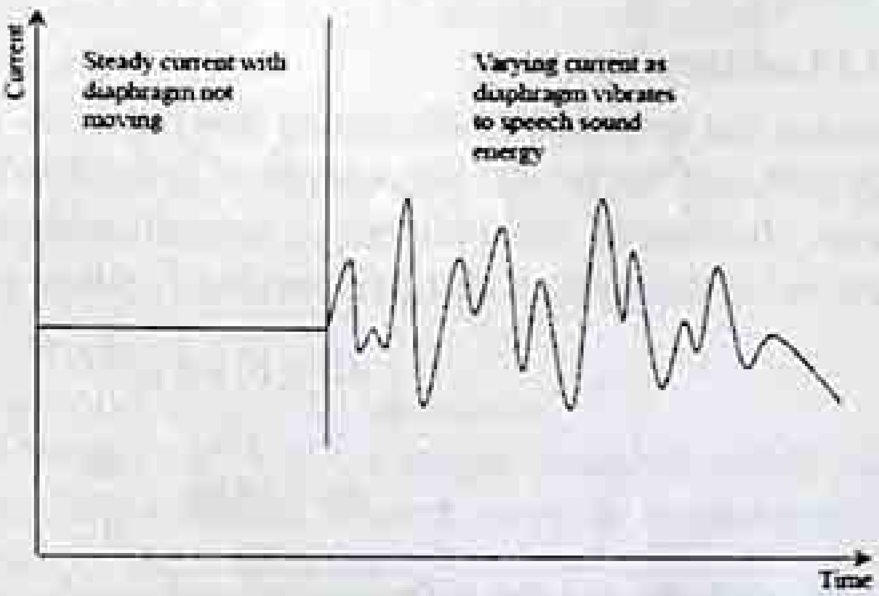


Fig. 18.5b: Varying d.c. produced by carbon granule microphone

This type of microphone finds extensive usage in the telephone handsets because of its fairly-good sound-reproducing qualities within the useful range of frequencies 300 to 340*Hz*. It is also comparatively cheap and robust.

18.2.3 The Human Ear

The human ear is a unique sound-detecting tool. It detect external sound well despite the fact that there are distractions around it. For example, there are sounds of blood flow, those of heart beat, body movements and of other internal organs. These distractions do not reduce its efficiency. Apart from just detecting sound, the human ear does locate sound sources even in confused sound environments. Figure 18.6 shows the anatomy of a human ear.

It is divided into three distinct parts; the outer ear, the middle ear and the inner ear. The outer ear consists of the 'aerial' pinna, the air-filled ear canal and the ear drum. The middle ear is made up of three bony structures – malleus, incus and stapes – called *ossicles*. The inner ear is liquid-filled and is made up of the cochlea and auditory nerves.

Sound waves from outside are focused into the air-filled ear canal and hence to the ear drum, a distance of about 25*mm* in the average adult. The ear drum is setinto vibration by the sound waves which it then transmits to the ossicles. These bony structures then retransmit the waves mechanically to the oval window of the inner ear, having magnified the vibrations of the ear drum up to sixty times.

The Eustachian tube, and approximately 36*mm* long tube, ensures that pressure is equalized on both sides of the ear drum. Sound now enters the inner ear by way of the oval window into the cochlea. The cochlea is in a spiral form (see Figure 18.6) and is divided into three parts by Basilar membrane and Reiassner's membrane. The Basilar membrane supports the organ of Corti which contains some thirty thousand hair cells that end in about eighteen thousand nerve fibers. These convey the sound wave stimuli detected by the ear to the brain for interpretation.

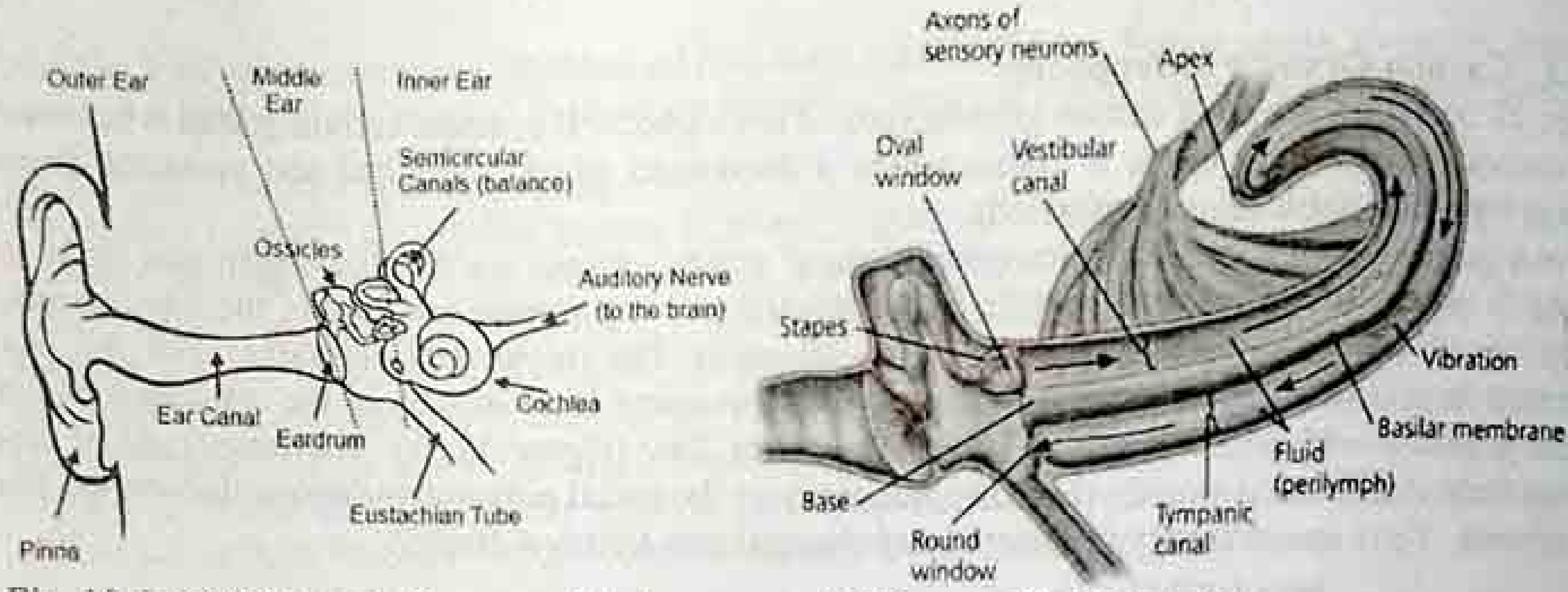


Fig. 18.6: (a) Schematic diagram of the human ear (b) The uncoiled inner ear.

18.2.3.1Auditory Response

The human ear responds to intensities ranging from $10^{-12}Wm^{-2}$ called threshold of hearing to $1Wm^{-2}$ usually referred to as the *threshold of pain*. The normal hearing range lies between these two extremes. Because of this wide range, intensities are usually measured in decibels, logarithmic scales. The decibel is thus not an absolute unit but is rather a ratio of intensities of powers.

$$dB = 10\log_{10}\frac{I}{I_0} \qquad (18.2)$$

$$dB = 10\log_{10}\frac{W}{W_0} \qquad (18.3)$$

Where I_0 and W_0 are the reference intensity and power respectively. The ratios of pressure are also expressed in decibels. Since intensity is proportional to the square of pressure, as shown in section 17.4, the sound pressure level can be given as

$$dB = 10\log_{10}\frac{P}{P_0} \qquad (18.4)$$

Where the reference sound pressure $P_0 = 2\times10^{-5}\,Pascal$.

Example 18.1

Find intensity level of sound from a radio set when it gives out sound at $10^{-8}Wm^{-2}$.

Solution

$I = 10^{-8}Wm^{-2}$ while $I_0 = 10^{-12}Wm^{-2}$

Ratio $I/I_0 = 10^{4}$, hence intensity level $= 10\log_{10}(10^4) = 40dB$

Example 18.2

Two radio sets each of average intensity level $30dB$ are played together but tuned to different stations. Find the total average intensity level of the sound heard.

Solution

$30dB = 10\log_{10}$ (ratio), hence ratio = antilog of 3, since total $I = 2I$,

Level, $dB = 10\log_{10}\frac{2I}{I_0} = 10\log_{10}2 + 10\log_{10}\frac{I}{I_0} = 3 + 30dB = 33dB$

These answer illustrates the fact that sources of sound, even with the same intensity, may increase, the intensity level does not increase correspondingly.

Summary

Sound detection involves the conversion of mechanical vibrations made by objects into forms that permit the analysis of their frequency and intensity. Detectors include the human ear and microphone.

The microphone converts natural sound energy into electrical one to facilitate (sound) amplification, while the ear not only detects sound, but also localizes the sources.

The human ear responds to a wide range of sound intensities, so to accommodate all, the intensity levels are measured in decibel, a logarithmic scale.

Exercise 18

18.1 Which of the following is not a natural source of sound
A. human speech **B.** vibrating string **C.** loud speakers **D.** wind

18.2 In stringed instruments, the basic frequency depends on :
I. length of the string II. Tension III. Mass of the string material IV. Area of the string
A. I, II and III only **B.** I and III only **C.** II and IV only **D.** I and II and IV

18.3 Find the pressure level of sound from a radio set when it gives out sound at 2×10^{-2} paschal.
A. $60dB$ **B.** $40dB$ **C.** $30dB$ **D.** $20dB$

18.4 Which of the following are essential components of the human speech mechanism?
I. Velum II. Tongue III. Glottis IV. Larynx
A. I, II and III only **B.** I and II only **C.** I, II, III, and IV only **D.** II only

18.5 A string of $60cm$ has a tension of $80N$ put on it. Calculate the frequency of the string if the mass per unit length of the string is $20\,kg/cm$.
A. $6.20\times10^{-3}\,Hz$ **B.** $4.17\times10^{-3}\,Hz$ **C.** $5.15\times10^{-3}\,Hz$ **D.** $3.25\times10^{-3}\,Hz$

18.6 Which of the following is not a type of microphone
A. Condenser **B.** moving coil **C.** velocity ribbon **D.** transducer

18.7 Which of the following are true about the human ear. I. it is a unique sound detecting tool II. it detects external sound well III. it locates sound sources even in a confused sound environment IV. cochlea is one of the parts of the human ear.
A. I and II only **B.** I, II and III only **C.** I, II,III and IV only **D.** I and III only

18.8 Find the intensity level of sound from a radio set when it gives out sound at $10^{-9}\,Wm^{-2}$?
A. $30dB$ **B.** $40dB$ **C.** $60dB$ **D.** $50dB$

18.9 The psychological effects of noise on human beings include the following except **A.** Disturbs conversation **B.** Disturbs sleep **C.** Causes mental fatigue **D.** It causes cracking of the wall

18.10 The inner ear is made of
A. malleus and incus **B.** cochlea and auditory nerves **C.** ear drum and ear canal

18.11 What are the frequencies of the first five harmonics of a violin tuned to $96Hz$?

18.12 Explain what is meant by
(a) Pressure level in sound (b) Two pressure levels differ by $3dB$ (c) Threshold of hearing.

18.13 Distinguish intensity from loudness of a sound.

18.14 Draw a typical carbon granule microphone and explain its operation. Why is it called a transducer?

18.15 Two loud speakers face each other and emit sound of frequency $150Hz$. If they are $150m$ apart, explain the variation of the intensity of sound along the line joining them.

18.16 The length of a violin is $50cm$ and the frequency is $500Hz$. What is the velocity of the wave in the string? (The speed of sound in air at room temperature is $344\,ms^{-1}$).

18.17 The tension on the longest string of a grand piano is $1098N$ and the velocity of a wave on this string $130\,ms^{-1}$. Calculate the mass per unit length of the string.

18.18 If the average intensity level of each two radio receivers is $48dB$ what is the average intensity level when both radios are on and tuned to different stations.

18.19 An outdoor public address system is adjusted to a level of $70dB$ for listeners $10m$ away. What intensity level is heard at $50m$ away?

18.20 The sound intensity level of a quiet typewriter is $40dB$. What is the sound intensity in Wm^{-2}?

OPTICS

$v = f\lambda$

CHAPTER 19
REFLECTION OF LIGHT AT PLANE AND CURVED SURFACES

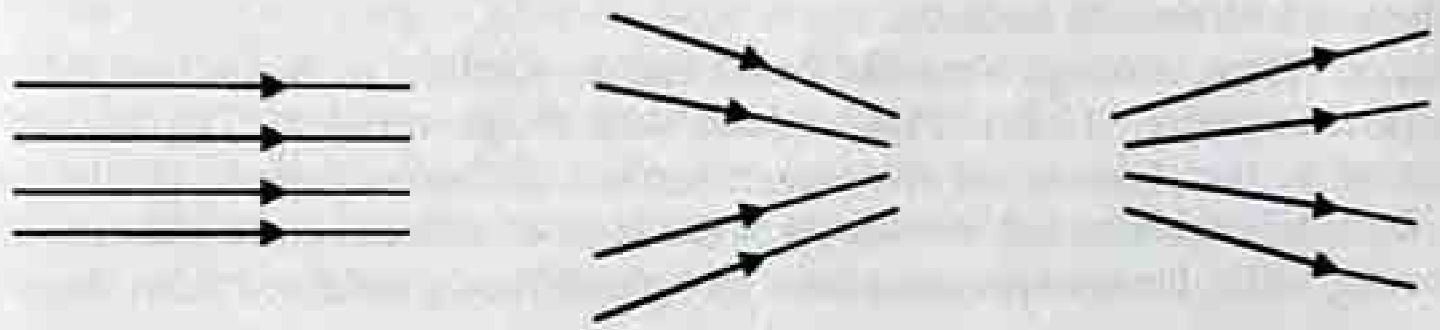
19.0 Introduction

Visible light is that aspect of radiant energy which stimulates the retina of the eye to create a sensation of vision. Light rays and waves are electromagnetic in nature. Other components of the electromagnetic spectrum include X-rays, radio waves, infrared, ultraviolet rays, etc. They have the same velocity and differ only in wavelength and frequency. They are connected by the relation $c = f\lambda$, where c is the common velocity of electromagnetic radiation, f is frequency and λ is the wavelength. For example, the wavelength of visible light ranges from 390*nm* to 760*nm* while that of the ultraviolet rays ranges between 100 and 390*nm*. One nanometer (*nm*) equals $10^{-9}\,m$.

The study of light is generally referred to as Optics. Optics is further divided into Geometrical and Physical Optics.

Geometrical Optics deals with rays; a ray is the path along which light energy travels. A ray can also be viewed as the motion of a stream of particles (photons) emanating from a source. Using the ray concept simplifies calculations. Rays propagate in straight lines (rectilinear propagation of light) and this has enabled the explanation of the formation of shadows of objects in the path of sight, formation of photographic images in pin-hole cameras and the eclipses of the sun and moon. When a bundle of rays proceed from a source in a particular direction it is a beam. The beam could be

(a) Parallel if the rays are parallel. This is the situation for rays from distant sources like the sun (Figure 19.1a)
(b) Convergent if the rays approach each other. A typical example is when a source is behind a lens like in a projection lantern (Figure 19.1b)
(c) Divergent if the rays are separating, for example rays from a lamp (Figure 19.1c).



(a) Parellel beam (b) Convergent beam (c) Divergent beam
Fig. 19.1: Beams of Light

Physical Optics on the other hand considers light as a wave motion arising at a source and advancing forward through a medium (Figure 19.2).

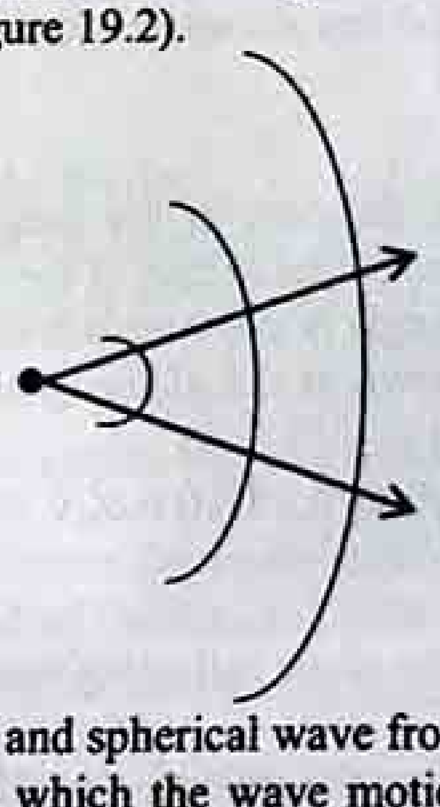


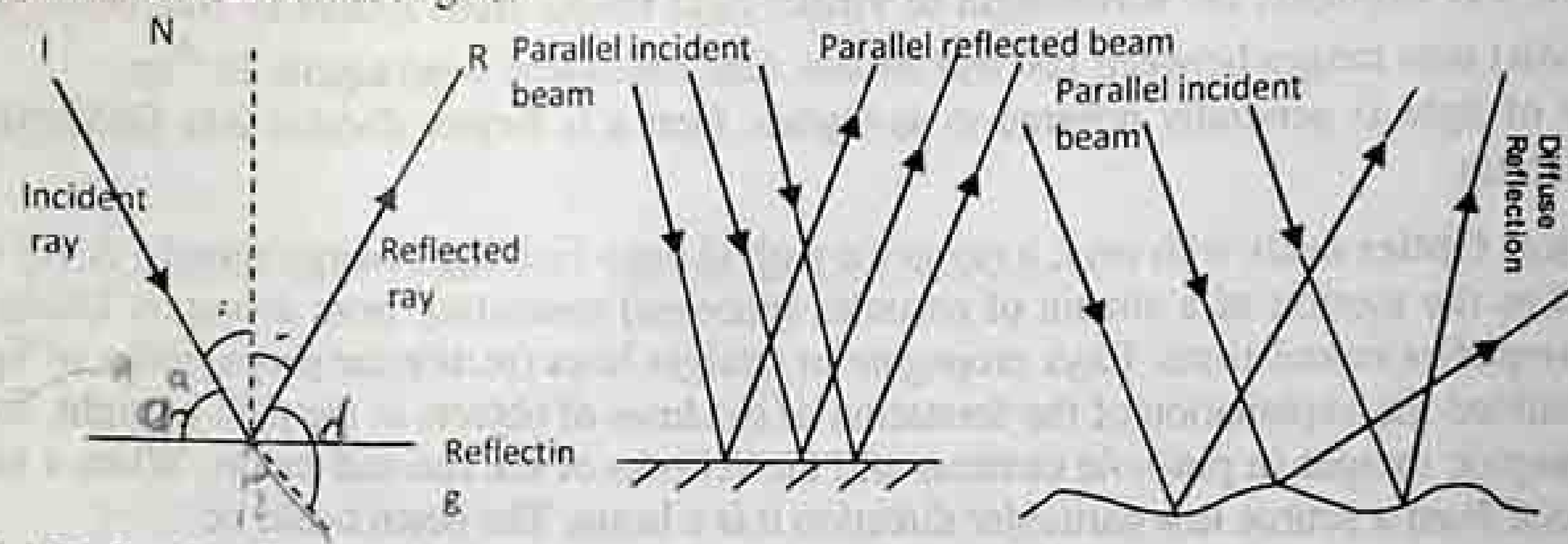
Fig. 19.2: Rays and spherical wave front from a source

A line draw in the direction in which the wave motion is travelling is called a *ray*. Some physical phenomena e.g. reflection and refraction can be explained using both the wave and ray concepts of light while others e.g. interference and diffraction can be explained using only the wave concept.

19.1 Reflection of Light Waves (Plane Surfaces)

What happens to any beam of light incident on a surface depends on the nature of the surface. Part of the beam could be thrown back or reflected into the same medium. The light could be absorbed if the surface is opaque like wood, or it could be transmitted if the surface is transparent e.g. clear glass. If the body transmits light but such that an object cannot be seen through it, the body is translucent e.g. some types of Louvre blades.

Ordinary glass reflects only about 4% of the light incident on it while polished metal surfaces reflect about 90% of the incident light. To ensure optimum reflection, mirrors are normally made by depositing silver on the back of glass.



(a) Reflection from a plane (b) Regular Reflection (c) Diffuse

Fig. 19.3: Reflections of light

Reflection

There are two main types reflection:

(i) Regular reflection: In this case, a parallel beam is reflected as a parallel beam (Figure; 19.3b). This is the situation with smooth surfaces.

(ii) Diffuse reflection: Here, although a parallel beam may be incident on a surface, the reflected beam is not parallel (Figure 19.3c). This happens with rough surfaces. The reason for the diffuse reflection is that because of the rough surface, different angles of incidence are presented to the incident beam, but at each point the laws of reflection are obeyed. Different objects e.g. books, walls, human beings are seen by light diffusely reflected from them into the eye.

The laws of Reflection

(i) The incident ray I, the reflected ray R, and the normal N, all lie in the same plane.

(ii) The angle of incidence, i, is equal to the angle of reflection, r. The angles are those between the light ray and the normal. The normal is the line drawn perpendicular to the reflecting surface. (Figure 19.3a)

Angle of Deviation and Glancing Angle

In the reflection of plane mirror, Figure 19.3a, the path of the incident beam of light is changed through an angle d to become the reflected beam, by the action of the reflecting surface. The angle d is called the angle of deviation. Another useful angle is that between the incident ray (or the reflected ray) and the reflecting surface. This angle is known as the glancing angle, a. From Figure 19.3a we can see easily that $d = 2a$.

$d = 180 - (i + r)$, $d = 180 - 2i$

19.2 Image Formation in Plane mirrors

The characteristics of reflection in plane mirrors are:

(i) The image of an object is as far behind the plane reflecting surface (mirror) as the object is in front. $U = V$

(ii) The object and image lie on a line normal to the reflecting surface.

(iii) The size of the image is the same as the size of the object. This is evident when we stand in front of a plane mirror. The magnification, (defined by size of image/size of object) is unity. $m = 1:1$

(iv) The image is laterally inverted.

(v) The image is virtual, i.e cannot be formed on a screen.

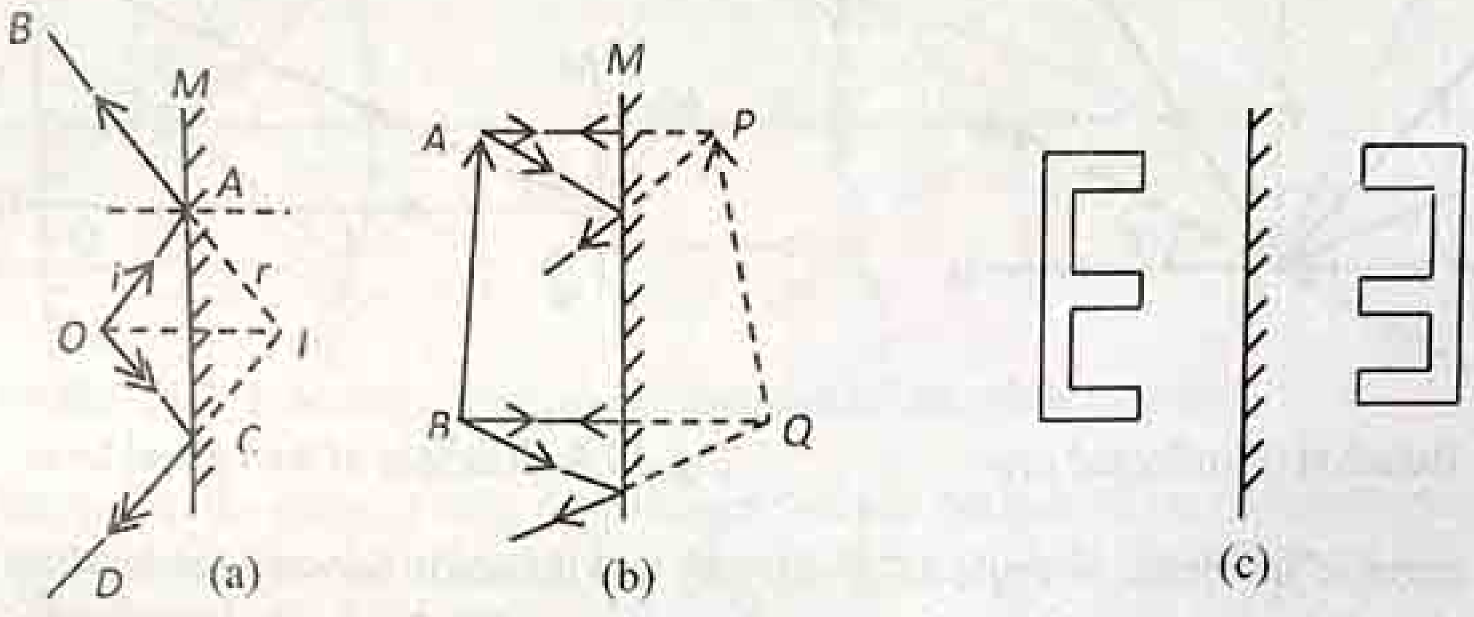


Fig. 19.4: (a) Location of the image of a point object in a plane mirror
(b) Location of the image of an extended object in a plane mirror
(c) Lateral inversion of image in a plane mirror

Image formation in plane mirrors can be easily demonstrated using ray diagrams. Suppose a point source O is located in front of the mirror M as shown in Figure 19.4(a), Rays OA and OC emanating from O will be reflected as AB and CD respectively. When BA and DC are extended backwards they meet at I where the image is formed. In Figure 19.4(b) we illustrate how the image PQ of an extended object AB can be located in a plane mirror M.

The rays shown for the two end points of the object can be reproduced for every point on the object leading to the extended image PQ seen. Note that a left hand in front of a mirror looks like a right hand in the mirror (Figure19.4c). The image is said to be perverted or laterally inverted.
One obvious point from Figure 19.4a and b is that the reflected rays do not actually pass through the image but appear to do so. Therefore if a screen were to be placed at I or PQ, no image would be received. Such an image is called virtual image as opposed to real images which can be located on a screen. Images of real objects in a plane mirror are always virtual.

Example 19.1
A man is standing in front of a vertical plane mirror $0.5m$ away. He then walks away from the mirror in a direction normal to the mirror at a velocity of $0.6ms^{-1}$. Determine the distance between the man and his image after $5s$.

Solution
Using the relation, $d = v \times t$
In $5s$ the man walks distance $0.6 \times 5 = 3m$
Distance of man from mirror $3m + 0.5m = 3.5m$
$\therefore$ Distance between man and image $3.5 \times 2 = 7m$.

19.3 Rotating Mirror
When a mirror rotates through an angle, the reflected beam rotates through twice that angle. This can be proved using Figure 19.5. The initial position of the mirror is MN; AO is the incident ray and OB the reflected ray. The angle between the incident and reflected rays is angle $AOB = 2i$. PO is the normal to the mirror at the point of incidence. The mirror is then rotated through angle θ to $M'N'$; correspondingly the normal is rotated through the same angle θ to QO.

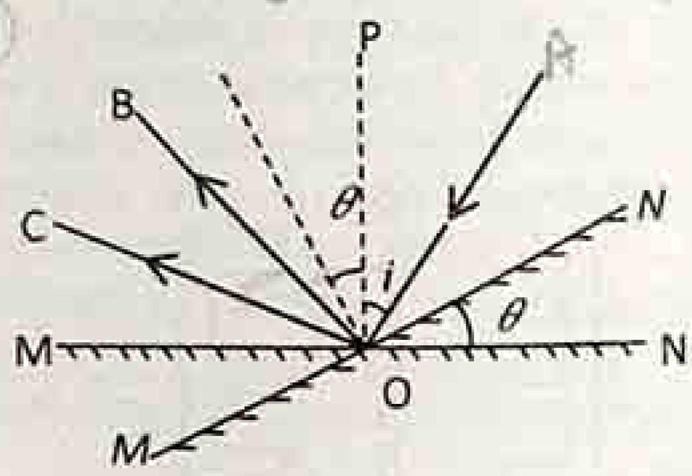


Fig. 19.5: Rotation of reflected ray

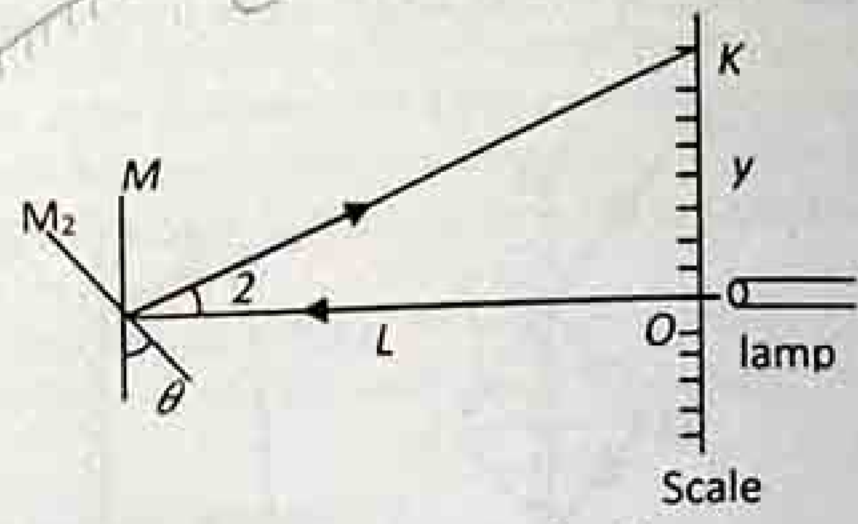


Fig. 19.6: Principle of the optical lever

The new angle of incidence is angle $AOQ = (i+\theta)$ and the angle between the incident and reflected rays is angle $AOC = 2(i+\theta)$, where the new reflected ray is OC. The angle between the reflected rays (= angle BOC) is equal to angle AOC - angle AOB, *i.e.* $= 2(i-\theta) - 2i = 2\theta$. Therefore the reflected ray is rotated through twice the angle of rotation of the mirror.

This principle is applied in the mirror galvanometer for measuring small currents. Light from a lamp (Figure 19.6) is used as a 'pointer' of negligible weight. The unit has the further advantage that it effectively doubles the rotation of the system to which the mirror M is attached. This system rotates when current flows, therefore the mirror turns from M_1 to M_2 through angle θ.

The angles involved are small and $\tan 2\theta = y/L$; therefore $\theta = y/2L$ radians. θ is related to the current flow.

19.4 Reflection at curved surfaces

Curved mirrors are polished surfaces made from portions of spherical surfaces. The same laws of reflection hold as for the plane mirror, but image formation differs. There are two general types of spherical mirror called the concave (or converging) mirror and convex (or diverging) mirror. The converging mirror converge a beam of parallel light rays to a real focus, while the diverging mirror diverges a beam of light from a virtual focus.

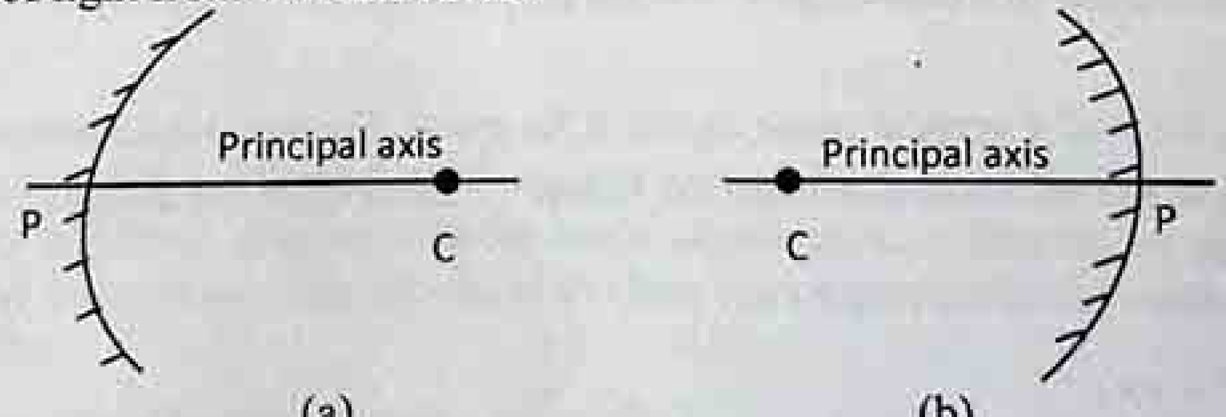


Fig. 19.7: Principal axis of (a) concave mirror (b) convex mirror

Definitions

The following definitions are illustrated in Figure 19.7 and 19.8:

(i) Centre of curvature, C, is the centre of the sphere from which the mirror is taken. This is a point.

(ii) Radius of curvature, R, is the radius of the sphere from which the mirror is taken. This is also the distance between the centre of curvature and the mirror, $CP = R$.

(iii) Principal axis, CP, is the axial plane, which divides the mirror into two equal halves.

(iv) Pole, P, is the point where the principal axis intersects the mirror.

(v) Principal focus, F, is the point where rays parallel to the principal axis converge (real focus), or diverge from (virtual focus), after reflection from the mirror.

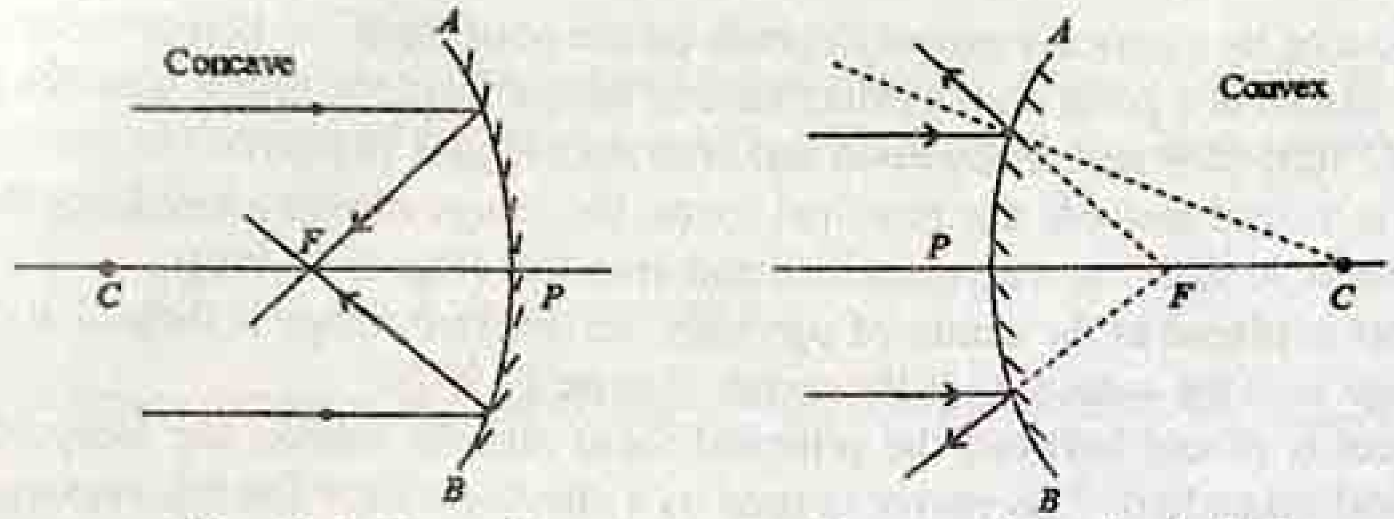


Fig. 19.8: Essential parameters of two special spherical mirrors

(vi) Focal length, f, is the distance from the principal focus to the pole of the mirror, measured along the principal axis. The focal length of a curved mirror is half the radius of curvature i.e. $PF = FC = f$, and $f = R/2$.

(vii) The positions u and v, represent the distance s of the object and image, respectively, from the mirror.

(viii) Magnification, m, is the ratio of height of image to that of object.

(ix) Real and virtual images: A real image can be received on a screen, whereas a virtual image is not obtainable on the screen.

19.5 Locating Images by Ray diagrams

We can make use of ray diagrams to locate images formed by mirrors and lenses. Graph sheets can be used in order to reflect the actual dimensions given in such problems.

Once the principal focus and center of curvature of a mirror are located, any two or three of the following rays can be drawn to fix the position of the image.

(i) A ray from the tip of the object drawn parallel to the principal axis should pass through the principal focus after reflection.

(ii) A ray from the tip of the object drawn through the principal focus emerges parallel to the principal axis after reflection.

(iii) A ray from the tip of the object passing through the centre of curvature will hit the mirror normally (This means that angle of incidence is zero) and will be reflected back along the same path.

(iv) A ray hitting the pole or any other part of mirror will be reflected such that the laws of reflection are obeyed.

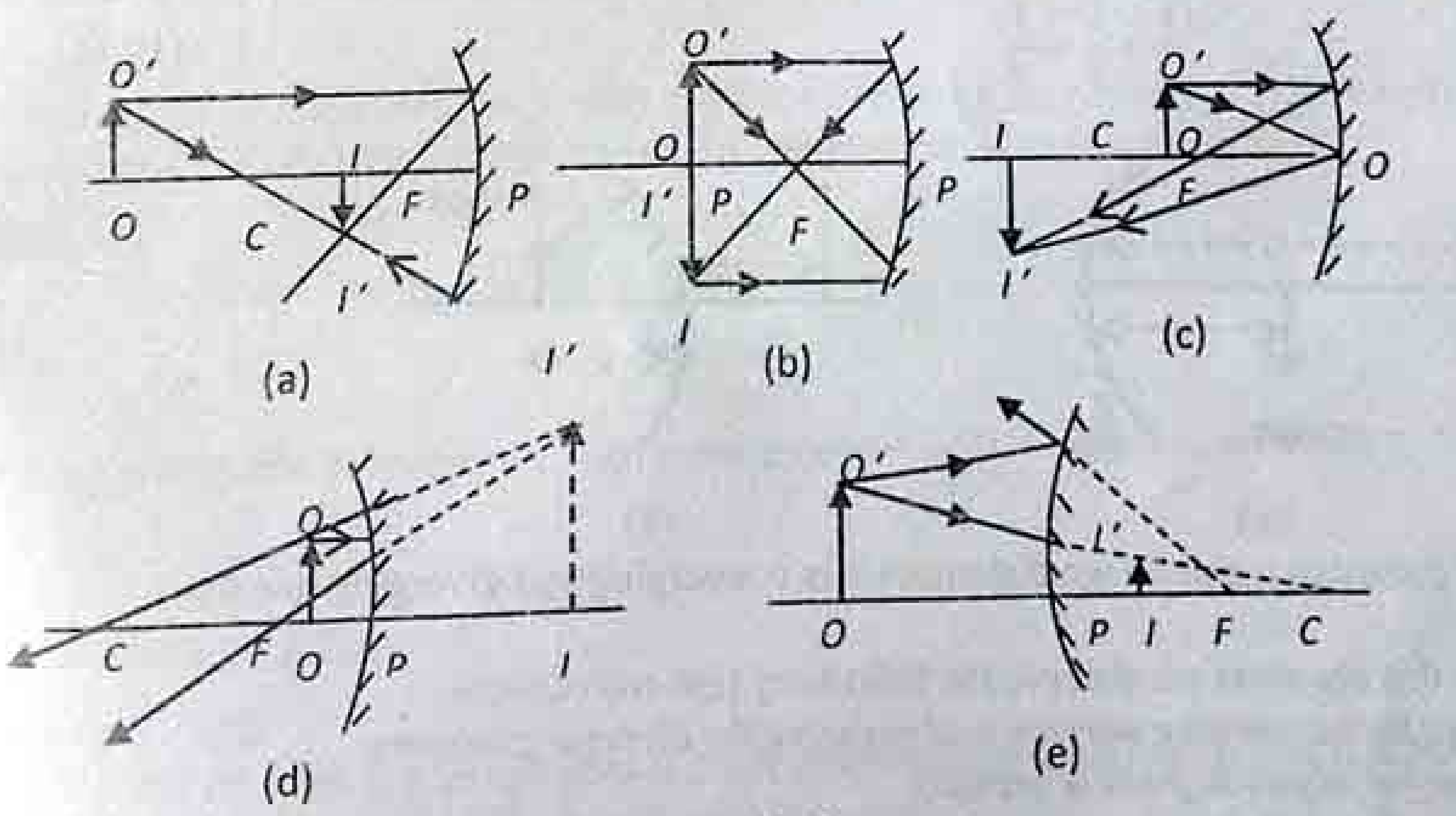


Fig. 19.9: Formation of images in spherical mirrors
(a), (b), (c) and (d)- Converging (concave) mirror: (e) - diverging(convex) mirror.

The type of image produced by a concave mirror depends on the position of the object:

(i) If the object is placed at a point on the principal axis farther from the mirror than the principal focus, the image formed is real, diminished and inverted. (Figure 19.9(a))

(ii) As the object is moved towards the principal focus, the image moves away from the mirror and becomes larger in size (i.e. real, magnified and inverted) (Figure 19.9(c))

(iii) When the object is placed at the centre of curvature, an inverted image is formed at the same place. The image is of the same size as the object. (Figure 19.9(b))

(iv) When the object is placed between the principal focus and the mirror, the image becomes virtual, erect and magnified. This mirror is used as a shaving mirror for this reason. (Figure 19.9(d))

(v) In general, the magnification for a converging mirror can be less than, equal to or greater than unity.

Diverging (convex mirror): This class of mirror produces images that are always virtual, erect and diminished, irrespective of the position of object from the mirror. Hence, the magnification for a diverging mirror is always less than unity. (Figure 19.9(e))

19.6 Mirror Equation

The mirror equation relates the object and image distances to the focal length, f or radius of curvature, r of the mirror (convex or concave).

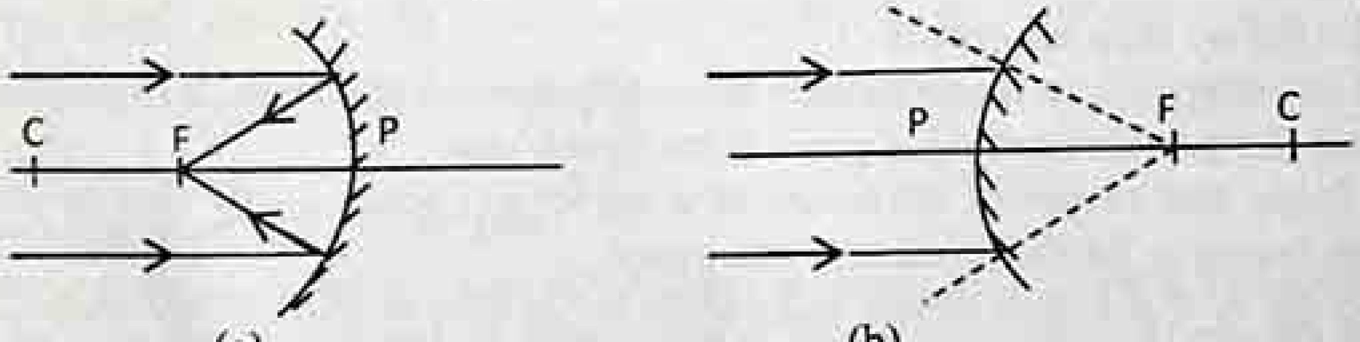


Fig 19.10: Rays in converging and diverging mirrors

C = Centre of curvature, F = The focus of the mirrors, P = The pole of the mirrors

In each case CF = FP = f, the focal length

CP = R= (radius of curvature) = $2f$

In both Figure 19.10(a) and (b) parallel rays near the principal axis CFP converge or appear to diverge from the focus F after the reflection from the mirror. It is easy to prove that u and v are related to the focal lengths of the mirror by the general mirror equation

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f} \qquad (19.1)$$

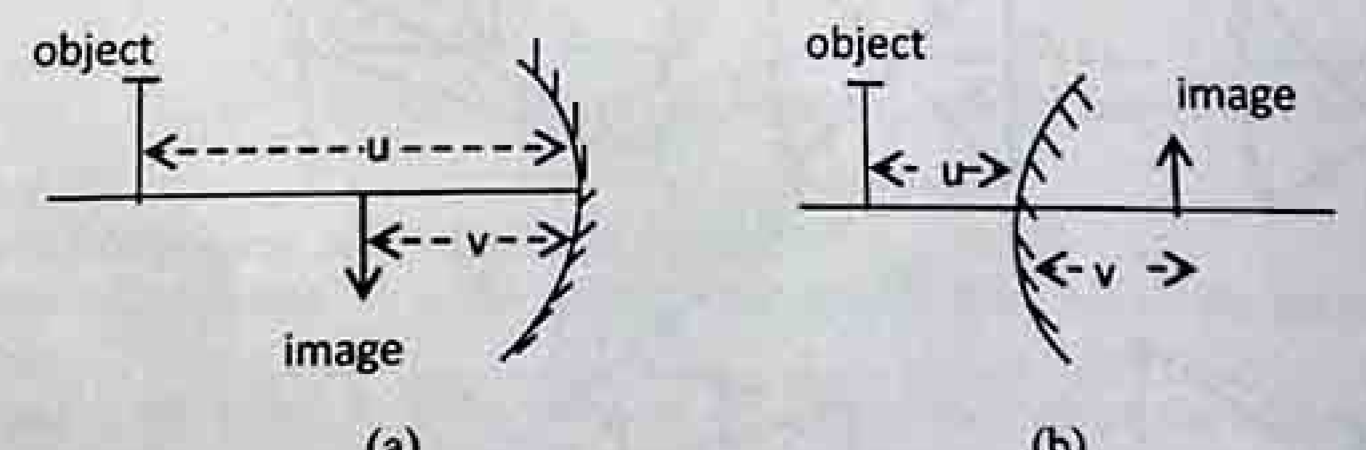


Fig 19.11: Object distances u and image distamces v in converging and diverging mirrors.

To properly apply this equation we observe the following sign convention:

(i) *f is positive for converging mirrors and negative for diverging mirrors*

(ii) *u is positive for object in front of mirrors*

(iii) *v is negative for virtual images and positive for real images..*

Magnification

The size of image formed by a concave or convex spherical mirror is usually called the magnification. It is expressed as the ratio of the image height to object height, i.e. h_i/h_o. It can also be expressed as a ratio of distances as follows:

$$linear\ magnification = \frac{v}{u} \tag{19.2}$$

where v = image distance from mirror and u = object distance from mirror

Therefore, $m = \frac{h_i}{h_o} = \frac{v}{u}$

Example 19.2

If light rays intersect a plane mirror at an angle of 25° from the vertical line to the surface, find the angle between the incident and reflected rays

Solution

Since the angle of incidence equals the angle of reflection, the incident and reflected rays will be separated by an angle equal to twice the angle of incidence. Hence,

$$2i = 2 \times 25° = 50°$$

Example 19.3

A tree stands one meter from a concave mirror with radius of curvature $20cm$. Find

(i) the position of the image
(ii) the magnification, and
(iii) the nature of the image

Solution

(i) Applying the mirror equation,i.e. $\frac{1}{u}+\frac{1}{v}=\frac{2}{r}$

This implies; $\frac{1}{100}+\frac{1}{v}=\frac{2}{20}$

Therefore; $v = 11.11cm$

(ii) Magnification, $m = \frac{v}{u} = \frac{11.11}{100} = 0.11$

(iii) The image is real and inverted.

Example 19.4

Where must an object be placed with respect to a convex mirror with a focal length of $25cm$, in order to obtain an image two-third as far behind the mirror as the object is in front?

Solution

For image behind mirror, v is negative. So,

$$v = \frac{-2u}{3}$$

Also, f is negative for a convex mirror. So, from the mirror equation, we have,

$$\frac{1}{u} - \frac{3}{2u} = -\frac{1}{25}$$

Therefore, u = 12.5cm

Example 19.5

An object 3 cm high is placed 16cm in front of a converging mirror so that it produces an image of 40cm from the mirror. Find the height of this image.

Solution

Magnification, $m=\frac{v}{u}=\frac{40}{16}=2.5$

But magnification, $m=\frac{h_2}{h_1}$

Therefore, $2.5=\frac{h_2}{3}$

Hence, $h_2=7.5cm$

Example 19.6
An object $2cm$ in height is placed $8cm$ from a spherical mirror and produces a virtual image $3.5cm$ in height. Where is the image located and what type of mirror is this?

Solution:

Magnification, $m=\frac{h_2}{h_1}=\frac{-3.5}{2}$

So, $y=1.75\times u=-1.75\times 8=-14cm$

Since v is negative. The image is located $14cm$ behind the mirror.

Hence, $\frac{1}{f}=\frac{1}{u}+\frac{1}{v}=\frac{1}{8}+\frac{1}{-14}$

Therefore, $f=18.7cm$

Since f is positive, the mirror is concave, and it has a radius of curvature of $37.4cm$.

Example 19.7
An object is placed $30cm$ in front of a concave mirror whose radius of curvature is $20cm$. Determine the position, nature and magnification of the image.

Solution:

The mirror formula is $\frac{1}{u}+\frac{1}{v}=\frac{1}{f}=\frac{2}{R}$

Therefore; $\frac{1}{v}=\frac{2}{R}-\frac{1}{u}=\frac{2}{20}-\frac{1}{30}=\frac{1}{15}$

Therefore the image distance is $15cm$. This is positive indicating the image is real.

Magnification, $m=\frac{v}{u}=\frac{15}{30}=0.5$

Example 19.8
At what distance must a dentist place a concave mirror from a tooth if she wishes to observe the tooth magnified four-fold?

Solution

Magnification, $m=4=\frac{v}{u}$

Therefore; $v=4u$

The image is virtual. Using the mirror formula; $\frac{1}{u}-\frac{1}{4u}=\frac{2}{4}=\frac{1}{2}$

Hence, $u=0.5cm$

Example 19.9
An object is placed in front of a convex mirror of focal length $10cm$. If the image is formed $8cm$ from the mirror, calculate the object distance.

Solution

The mirror formula is $\frac{1}{u}+\frac{1}{v}=\frac{1}{f}$

Therefore; $\frac{1}{u}=\frac{1}{f}-\frac{1}{v}$

This gives; $\frac{1}{u}=-\frac{1}{10}-\left(-\frac{1}{8}\right)=\frac{1}{8}-\frac{1}{10}=\frac{1}{40}$

Hence, $u=40cm$

Both the focus and the image are virtual.

19.7 Spherical Aberration in Mirrors

Using curved mirrors, only paraxial rays are reflected to pass through the principal focus. Rays far from the principal axis are reflected to meet the principal axis nearer the pole than the focus as exaggerated in Figure 19.12(a). The reflected ray touch a curve *ABF* called acaustic. Therefore, rather than a sharp well-defined image at the principal focus, we have a blurred and imperfect image. It is this phenomenon that is referred to as spherical aberration. By the principle of reversibility of light which states that if a ray is reversed, it always travels along its original path, any luminous source placed at F will not produce parallel rays on reflection from points far from the principal axis.Spherical aberration can be corrected using parabolic mirrors.

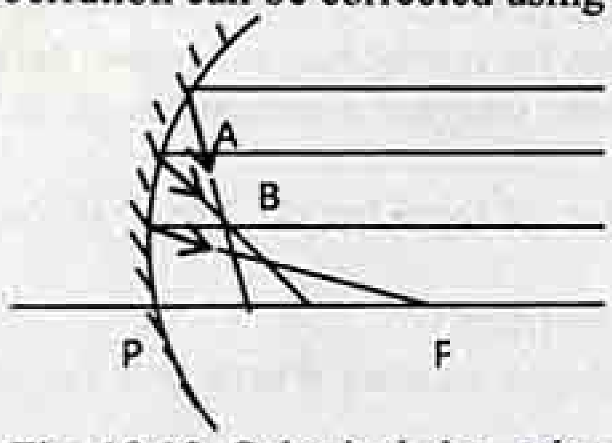


Fig. 19.12: Spherical aberration in mirror

Summary

(i) The laws of reflection are:
- The incident ray, the reflected ray and the normal to the reflecting surface at the point of incidence are in the same plane.
- The angle of incidence is equal to the angle of reflection.

(ii) When a mirror is rotated through an angle, the reflected ray is rotated through twice that angle. This principle is applied in the mirror galvanometer for measuring small currents. .

(iii) The mirror formula for reflection at a curved surface is $\frac{1}{u}+\frac{1}{v}=\frac{1}{f}=\frac{2}{R}$

where R is the radius of curvature of the mirror. A consistent sign convention must be used in applying this formula. Rays which are far from the principal axis of a mirror are reflected to cut the axis nearer the mirror than those close to the principal axis. Therefore all the reflected rays do not pass through a single focus. This defect is known as spherical aberration.

(iv) Parbolic mirrors can be used to correct this defect.

Exercise 19

19.1 What is the nature of the image produced by the convex mirror irrespective of the position of object from the mirror?

A. Virtual, erect and magnified **B.** Virtual inverted and diminished
C. Virtual, erect and diminished **D.** Real, inverted and diminished

19.2 A man stands 50*cm* from a concave mirror with radius of curvature 10cm. Find the position of the image and the magnification

A. 5.56*cm*, 0.11 **B.** 6.65*cm*, 0.12 **C.** 7.55*cm*, 0.13 **D.** 7.80*cm*, 0.14

19.3 Where must an object be placed with respect to a convex mirror with a focal length of 30*cm*, in order to obtain an image one-third as far behind the mirror as the object is in front?

A. 50.00*cm* B. 60.00*cm* C. 65.00*cm* D. 55.00*cm*

19.4 An object 5cm high is placed 20*cm* in front of a concave mirror so that it produces an image 60*cm* from the mirror. Find the height of this image.

A. 16.00*cm* B. 15.00*cm* C. 17.00*cm* D. 18.00*cm*

19.5 An object 5*cm* in height is placed 10*cm* from a curved mirror and produces a virtual image of 6.5*cm* height. Where is the image located and what type of mirror is this?

A. -13.00*cm*, concave mirror B. -14.00*cm*, convex mirror

C. -13.00*cm*, convex mirror D. -13.00*cm*, convex mirror

19.6 If light rays intercept a plane mirror at an angle of 35° from the vertical to the surface, find the angle between the incident and reflected.

A. 70° B. 80° C. 60° D. 50°

19.7 When an object is placed between the principal focus of a concave mirror and the mirror, the image formed is ....

A. Real, erected and magnified B. Virtual, inverted and magnified

C. Virtual, erect and magnified D. Virtual, erect and dimished

19.8 Which of the following is TRUE about the essential parameters of curved mirrors?

A. Pole is the centre of the sphere from which the mirror is taken

B. Radius of curvature, R is the distance from the principal focus to the pole of the mirror

C. Principal axis is the axial plane, which divides the mirror into two equal halves

D. A real image cannot be received on a screen, whereas a virtual image is obtainable on a screen.

19.9 An object is placed in front of a convex mirror of focal length 15*cm*. If the image is 10*cm* from the mirror, calculate the object dstance.

A. 30.00*cm* B. 25.00*cm* C. 35.00*cm* D. 40.00*cm*

19.10 An object is placed at infinity in front of a concave mirror. Where will the image be formed and what is the nature?

A. Principal focus, real, diminished and inverted

B. Centre of curvature, real, magnified and inverted

C. Principal focus, virtual, diminished and erect

D. Centre of curvature, real, diminished and inverted

19.11 Explain what you understand by opaque, transparent and translucent materials. Give examples of each.

19.12 Explain how a real image can be distinguished from a virtual image.

19.13 Why is the principal focus of a convex mirror called a virtual focus?

19.14 A distant object is brought towards a concave spherical mirror. Describe the changes in the size and nature of the image as the object distance varies from infinity to zero.

19.15 Explain the conditions under which a spherical mirror which may be concave or convex, can form (a) a virtual image (b) an erect image and (c) an image larger than the object.

19.16 Describe a laboratory methods for measuring the focal length of a concave mirror. What precautions are necessary to obtain an accurate result?

19.17 A man 2*m* 50*cm* tall stands a distance of 3*m* in form of a large vertical plane mirror,

(i) Determine the size of the image of the man formed by the mirror.

(ii) How far from the man is the image?

(iii) What is the shortest length of mirror that will enable the man see himself fully?

(iv) What is the answer to (iii) above if the man were 5*m* away?

19.18 A plane mirror 20*cm* long is held vertically 25*cm* from the eye. If it is completely filled by the image of a building 200*m* away, find the height of the building. What is the magnification of the image?

19.19 Two plane mirrors are inclined at an angle of 30°. A ray of light which makes an angle of incidence of 50° with one of the mirrors undergoes two successive reflections at the mirrors. Calculate the angle of deviation.

19.20 An object of height 2*cm* is placed at a distance of (a) 10*cm* (b) 6*cm* from a concave mirror of radius of curvature 15*cm*. Find the position and size of the image in each case.

CHAPTER 20
REFRACTION AT PLANE AND CURVED SURFACES

20.0 Introduction

When light rays strike a smooth flat interface between two media, the direction of the light changes as it enters the second medium, except when the rays strike the interface at right angles. This change in direction, i.e. bending of the light rays on entering another medium is called refraction of light (Figure 20.1). Refraction is caused by the change in speed of light beam in the new medium which is accompanied by a change in wavelength at a constant frequency.

The amount of light reflected or refracted depends on the media involved. If the first medium is air, and the second is water or glass, most of the light is transmitted into the second medium and bent towards the normal to the surface at the point of incidence; this is illustrated in Figure 20.1 (a).

In general, if the first medium is lighter than the second, refraction is towards the normal, otherwise, it is away from the normal (Figure 20.1b). Refraction enables us see transparent objects. Refraction effects are too evident in everyday life. For example, a pond appears less deep than it really is because rays from the bottom are refracted away from the normal on getting to less dense air. Therefore, to an eye E in air, the bottom of the pond A appears to be at B (Figure 20.1c)

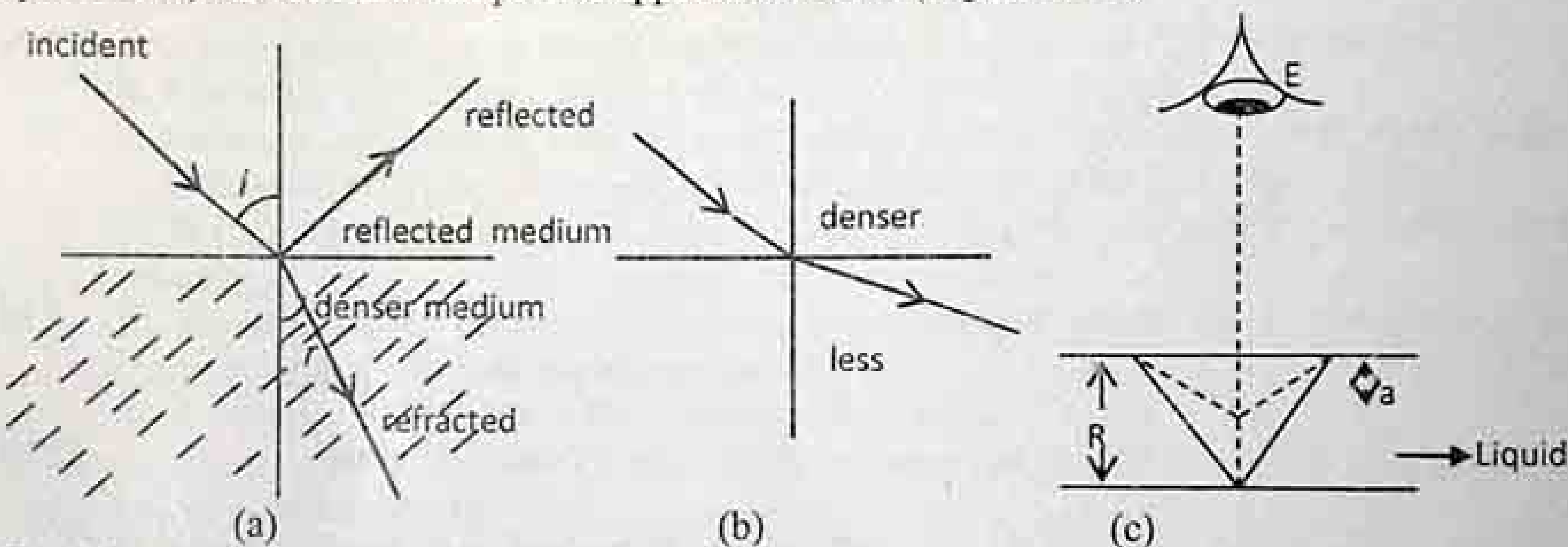


Fig. 20.1: (a) Refraction at a plane surface (b) Refraction from a denser to less dense medium (c) Real and Apparent depth.

20.1 The Laws of Refraction

The laws of refraction state that;

1. The incident ray, the refracted ray and the normal at the point of incidence, all lie in the same plane.
2. The ratio of sines of the angle of incidence and refraction is a constant for light of a given wavelength.

This is called the *Snell's law*; which can be expressed as follows;

$$\mu = \frac{speed\ of\ light\ in\ medium\ 1}{speed\ of\ light\ in\ medium\ 2} = \frac{\sin i}{\sin r} \qquad (20.1)$$

where the constant μ is the *index of refraction* of medium 2 relative to medium 1 (or the relative refractive index). The laws of refraction of light hold for any homogenous and isotropic medium, in the absence of the absorption of light.

Refractive Indices

The ratio of the speed of light in empty space, c, to its speed v in a given medium is called the *absolute refractive index, n*, of this medium, i.e.

$$n = \frac{velocity\ of\ light\ in\ vacuum}{velocity\ of\ light\ in\ vacuum} = \frac{c}{v} \qquad (20.2)$$

It follows that $n > 1$ for any medium except a vacuum for which $n = 1$. The absolute refractive index n depends on the frequency of the light and the state of the medium (i.e. its density and temperature).

Refractive index can also be defined as the ratio of the velocities of light v_1 and v_2 in the first and second media respectively;

$$n_{21} = \frac{n_2}{n_1} = \frac{c/v_2}{c/v_1} = \frac{v_2}{v_1} = {}_1\mu_2 \tag{20.3}$$

where n_1 and n_2 are the absolute refractive indices of the first and second media, respectively. For a series of media with absolute refractive indices n_1, n_2, n_3 and n_4.

$$n_{41} \text{ or } {}_1\mu_4 = \frac{n_4}{n_1} = \frac{n_2}{n_1} \times \frac{n_3}{n_2} \times \frac{n_4}{n_3} \tag{20.4}$$

$$= {}_1\mu_2 \times {}_2\mu_3 \times {}_3\mu_4$$

We can also prove that refractive index of a liquid is real depth divided by apparent depth $= R/a$ as shown in Figure 20.1

Example 20.1

Consider a ray of light in air incident at an angle of 42° on the boundary with water. Calculate the angle of refraction if the relative refractive index of air to water is 0.75.

Solution

Let i and r be the angles of incidence and refraction respectively. Then by *Snell's* law,

$n_1 \sin i = n_2 \sin r$, n_1

$$\frac{\sin r}{\sin i} = \frac{n_1}{n_2} = 0.75$$

Therefore; $\sin r = \sin i \times 0.75 = 0.67 \times 0.75$

$$\sin^{-1}(0.5025) = 30.17°$$

20.2 Optical Invariant

Consider refraction through a parallel sided glass block located in air (Figure 20.2a). At M, the refractive index of the glass relative to air is given by

$${}_a n_g = \frac{\sin i}{\sin r} \tag{20.5}$$

Using the principle of reversibility of light, at N

$${}_a n_g = \frac{\sin e}{\sin r} \tag{20.6}$$

Therefore $e = i$.

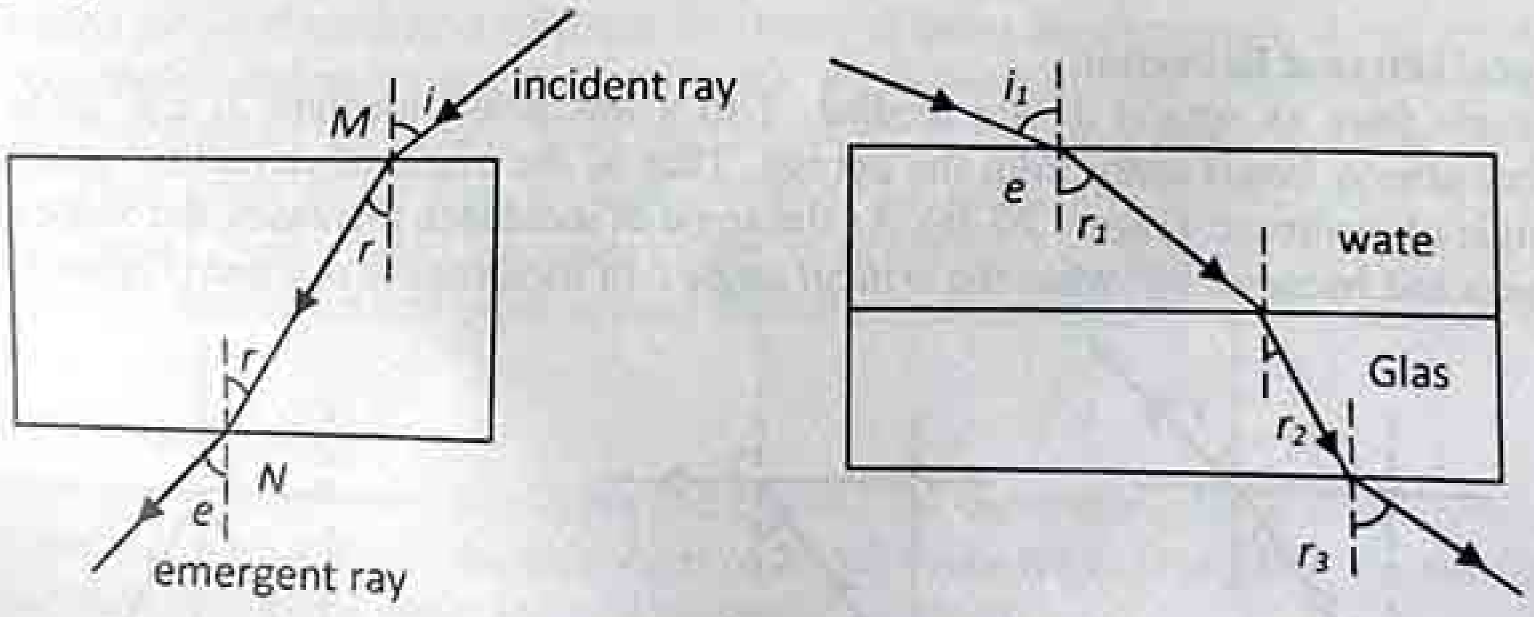


Fig. 20.2 (a) Refraction through Parallel-sided Glass block,
(b) Refraction through several Parallel-sided media.

This means that the emergent ray is parallel to the incident ray. Therefore, the angle of deviation is zero and the ray is not deviated but laterally displaced.

Next consider refraction through several parallel -sided media, for example water and glass as shown Figure. 20.2(b).

$$ {}_a n_w = \frac{\sin i_1}{\sin r_1} \qquad (20.7)$$

$$ {}_w n_g = \frac{\sin r_1}{\sin r_2} \qquad (20.8)$$

$$ {}_a n_g = \frac{\sin r_3}{\sin r_2} = \frac{\sin i_1}{\sin i_2} \qquad \text{(since } i_1 = r_3\text{)} \qquad (20.9)$$

Therefore, $$ {}_w n_g = \frac{{}_a n_g}{{}_a n_w} = \frac{\sin r_1}{\sin r_2} $$

This gives $$ {}_a n_w \sin r_1 = {}_a n_g \sin r_2 \; {}_a n_w \sin r_1 = {}_a n_g \sin r_2 \qquad (20.10)$$

Therefore at *A*, the refractive index of water multiplied by the sine of the angle made in water by the ray with the normal is equal to the refractive index of glass multiplied by the sine of the angle made by the ray with the normal. In general, for refraction through several media:

$$ n \sin\theta = constant \qquad (20.11)$$

This is known as *optical invariant*. It is an important relation with wide application - valid for plane and curved surfaces.

Example 20.2

In Figure 20.2 (b) calculate the angle r_2 if the refractive index of water is 1.33, that of glass is 1.6 and the angle of incidence in air i_1 is 60°.

Solution

Using the optical invariant at the air-water interface

$$1 \times \sin 60° = 1.33 \sin r_1$$

Similarly, at the water-glass interface

$$1.33 \sin r_1 = 1.6 \sin r_2$$

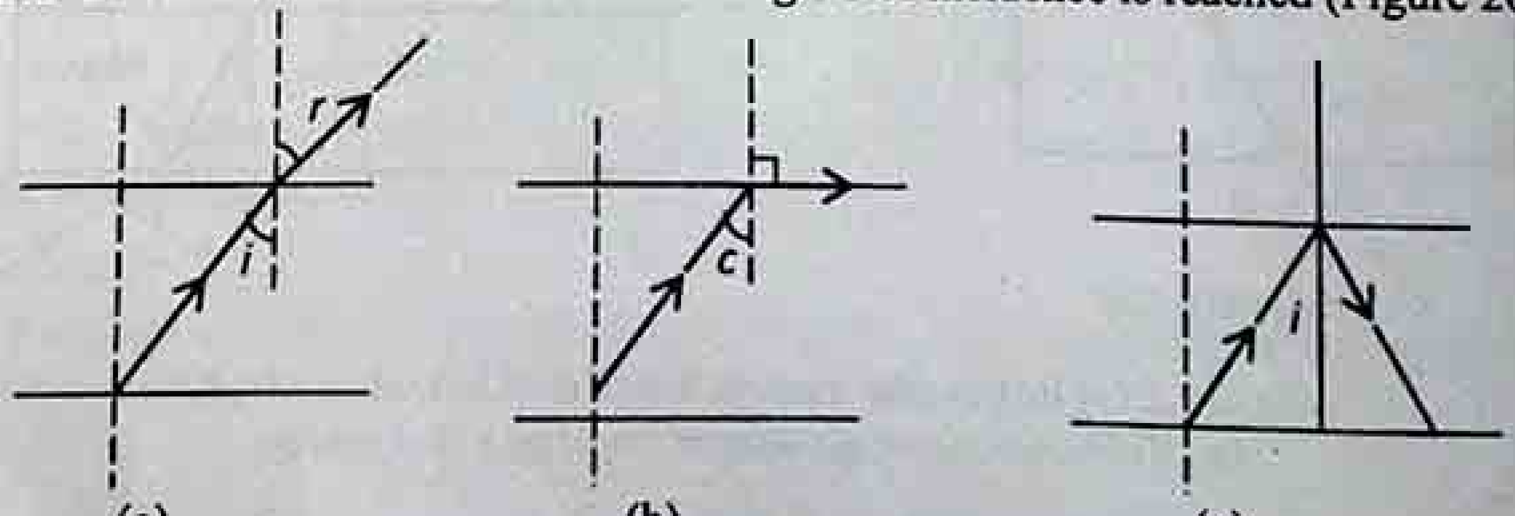
Therefore; $1.6 \sin r_2 = 1 \times \sin 60°$ 1.6

$$\sin r_2 = \frac{\sin 60°}{1.6} = 0.5413$$

Hence, $r_2 = 32°46'$

20.3 Total Internal Reflection

If light travels from an optical denser medium 1 to a less dense medium 2, e.g. glass to air, the refracted ray always bends away from the normal. That is, the angle of refraction is always greater than the angle of incidence, (Figure 20.3a). As the angle of incidence increases, the angle of refraction also increases and becomes 90° when the *critical angle c* of incidence is reached (Figure 20.3b).



(a) (b) (c)

Fig 20.3: Normal refraction — Critical angle — Total internal reflection

That is, the refracted ray travels along the boundary separating the two media when the angle of incidence equals the critical angle. As the angle of incidence increases beyond the critical angle, there is no refracted ray, rather all the light are reflected Figure 20.3c. This phenomenon is known as *total internal reflection.*

Computation of Critical Angle

Taking the case of light traveling from glass to air, the critical angle for glass, c is obtained as follows: The refractive index for glass, $n_1 = 1.50$ and that for air, $n_2 = 1$. By Snell's law,

$$\frac{\sin i}{\sin r} = \frac{n_2}{n_1} \tag{20.12}$$

$$\frac{\sin c}{\sin 90} = \frac{1}{1.5} \tag{20.13}$$

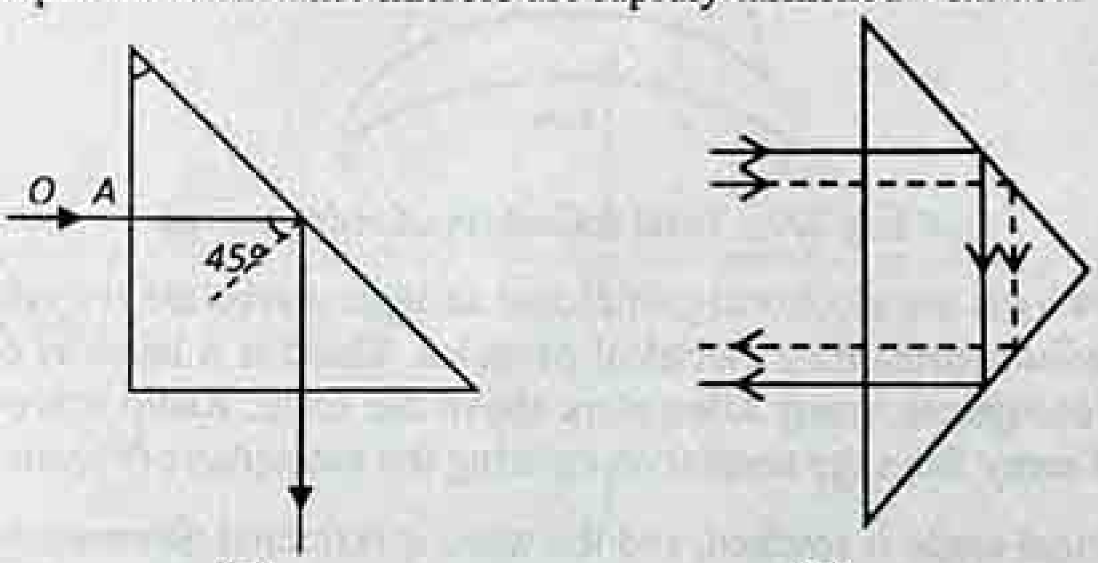
i.e.

$$\sin c = \frac{1}{n} \tag{20.14}$$

Hence, $\sin c = 0.67$ and $c = 41.8°$. Now, we have that the critical angle for glass is 41.8°.

20.4 Right-angled Prisms and Reflectors

In Optics, a prism is a transparent body bounded by plane, polished surfaces which intersect in parallel straight lines. As a reflector, the plane mirror has the disadvantages if the mirror is thick, multiple reflections can lead to multiple confusing images, some light is usually lost to absorption or refraction and finally unprotected metallic mirrors are rapidly tarnished with use.



(a) (b)

Fig. 20.4: Totally reflecting prisms

The right-angled prism if used as a reflector has none of these disadvantages. A ray incident on the prism at right angles makes an angle of 45° with the hypotenuse and is totally reflected (Figure 20.4a). This is because 45° is greater than the critical angle for most glass. In effect the ray is bent through 90°. Thus there is no loss of light intensity to refraction or absorption. Figure 20.4(b) illustrates the use of right angled isosceles prism to invert rays or turn the ray through $180°$. This principle is used in binoculars and other optical instruments where it is desired to change the direction of light.

Light Pipe for Transmitting Light

Fibres made of glass of high refractive index can be used for transmitting light. If the light enters one end of the fibre at a small angle to the axis, most of it will strike the curved glass-air interface at angles greater than the critical angle and will be totally reflected.

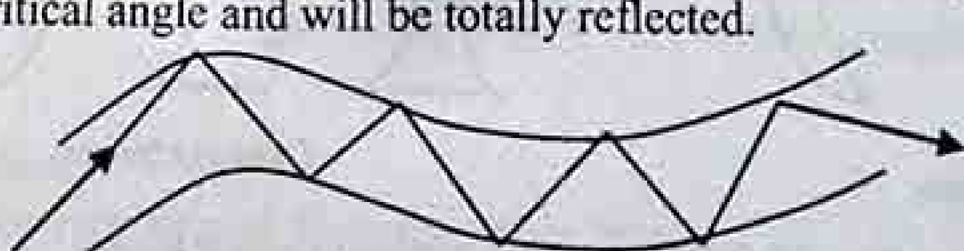


Fig. 20.5: Light pipe made of a transparent material

Thus, the light will be guided from one end of the fibre to the other (Figure 20.5). Pulling thousands of thin fibres together result in a flexible light pipe. Apart from light, pictures can also be transmitted by this device. Each fibre transmits a small portion of the picture placed in front of it and illuminated. At the other end, and image is seen consisting of an assemblage of individual dots of light, each one coming from a single fibre. The image is therefore formed by a pattern of light and shade. It is clear that to form a sharp image, the diameter of the fibre must be small- usually of the order of $2\times10^{-6}m$. Also, fibre pipes together with high intensity light from lasers are used in surgery and cancer therapy, where lamps producing too much heat cannot be used. Inaccessible parts of the body such as under the skin and in the lungs can be examined by looking at the image transmitted using fibre optics. A bundle can be enclosed in a hypodermic needle for the study of tissues and blood vessels far beneath the skin. Special fibres with low-loss characteristics are now used in the communication industry. A bundle of fibres can carry hundreds of lines of telephone calls using light signals more than a bundle of copper wires of the same size using electrical signals.

20.5 Total Reflection of Radio Waves

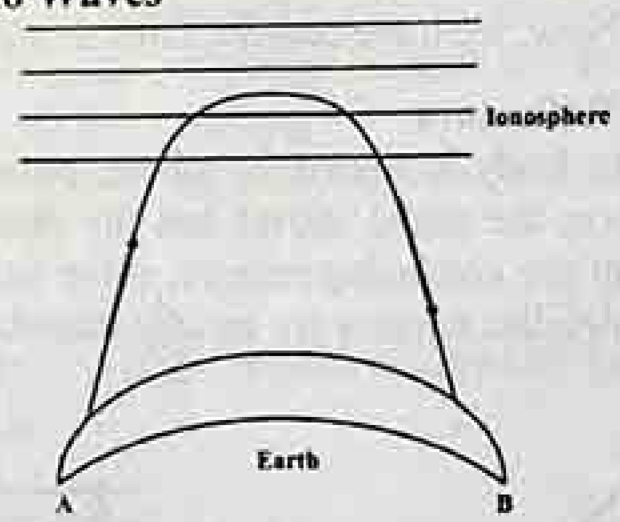


Fig 20.6: Total reflection of radio waves

Both radio and light waves are electromagnetic. Just as light waves are refracted, so radio waves are refracted but in a medium containing electrical particles. There is a layer of considerable density of electrons called the ionosphere, many kilometers above the earth. Radio waves sent from location A on earth are refracted away from the normal on entering the ionosphere (Figure 20.6).

At some height, a critical angle is reached, and the wave is refracted downwards to be received on earth by B.

20.6 Refraction at Spherical Surfaces-Thin Lenses

A lens is a transparent body, usually made of glass, with regular curved surfaces. When it is thicker in the middle than at the edges, it is called *convex* or *converging lens*; when it is thinner in the middle than at the edges, it is called *concave* or *diverging lens.*

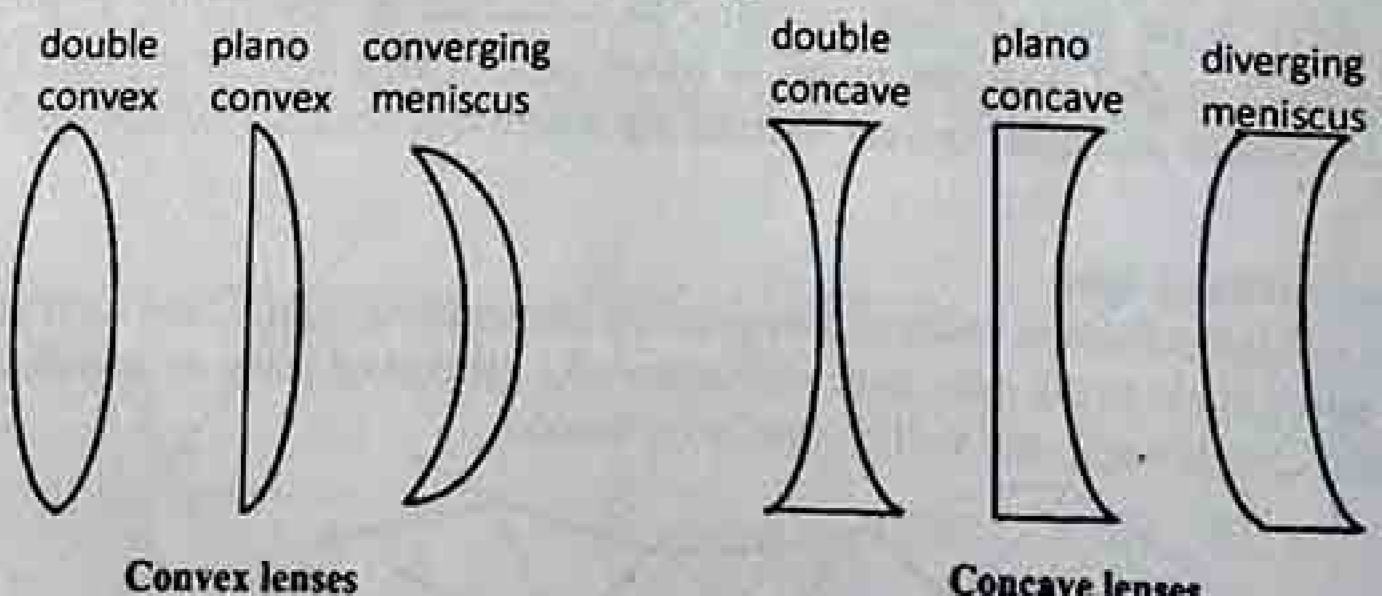


Fig. 20.7: Converging and diverging lenses

Typical examples are shown in Figure 20.7. For a thin lens, the thickness is small compared with the other dimensions. The principal axis of the lens is the line joining the centres of curvature of the surfaces and passing through the middle of the lens. If a parallel beam is incident on a convex lens the rays will converge at F_2 called the principal focus (Figure. 20.8a).

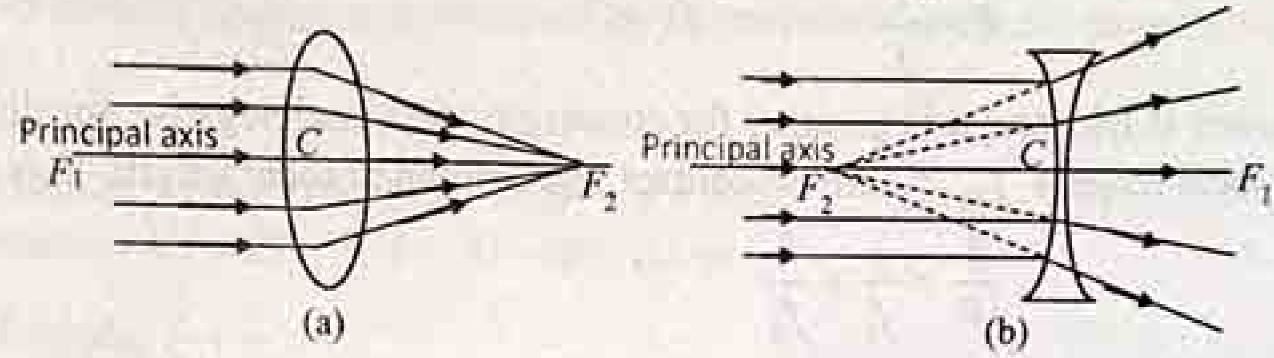


Fig. 20.8: Focal point of (a) converging and (b) diverging lenses

The distance from the centre of the lens to F_2 : is called the focal length of the lens. In this case the rays actually pass through F_2, therefore, the focal point is real and the focal length positive, On the other hand, with a concave lens, the rays diverge and the focus is virtual (Figure 20.8b). Note that there are two focal points, one at either side of a thin lens equidistant from the centre of lens.

20.7 Image Formation in Thin Lenses

The essential rays for determining the size and position of the image of an object in a thin lens are:

(i) any ray parallel to the principal axis is refracted to pass through the corresponding focal point
(ii) any ray through the centre of the lens is undeviated.
(iii) any ray through the near focal point will emerge parallel to the principal axis after refraction by the lens.

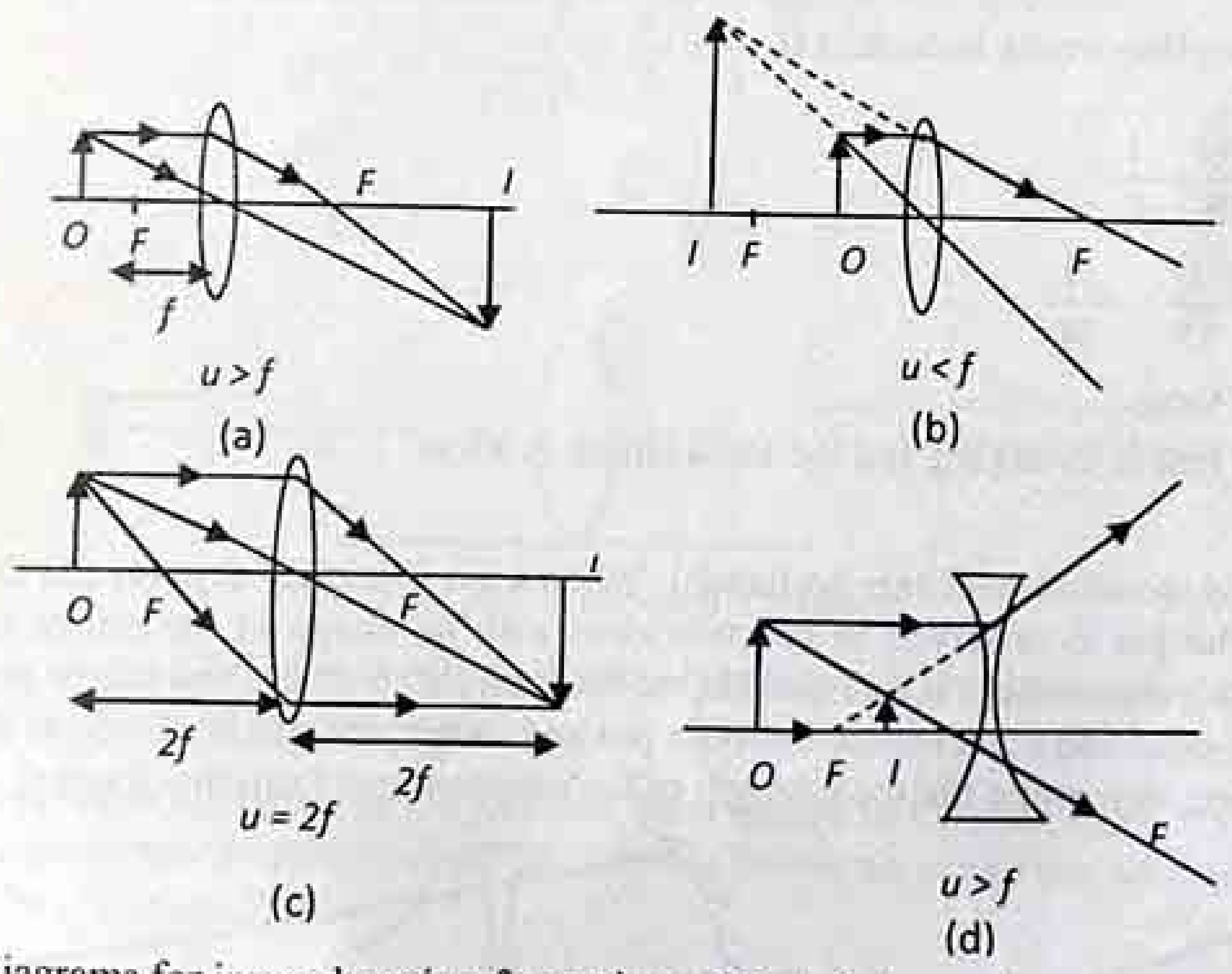


Fig. 20.9: Ray diagrams for image location for various object distances in convex (converging) and concave (diverging) lenses.

In Figure 20.9 (a, b, c, and d) we provide ray diagrams for a few representative object positions in thin lenses. Just as for spherical mirrors, note that the intersection of at least two rays in the image space is required to locate the image of any point on the object. The virtual images are dotted and cannot be received on a screen.

20.8 Thin Lens Equation

It can be proved that for converging and diverging sthin lenses

$$\frac{1}{u}+\frac{1}{v}=\frac{1}{f}$$

where u is object distance, v in image distance and f in focal length.

This equation applies equally to concave and convex lenses provided a consistent sign convention is applied. If the focal length of a lens is in metres, the power, P of the lens is given by

$$P = \frac{1}{f} \tag{20.15}$$

The unit of lens power is Dioptre; it is positive for converging lens and negative for diverging lens. If two thin lenses of focal lengths f_1, f_2 are in contact, the combined focal length, F is given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} \tag{20.16}$$

To apply the formula correctly account must be taken of the sign of the focal lengths.

Example 20.3
Two thin lenses are in contact. An object placed at a distance of 45*cm* from the lenses forms a real image at a distance of 30*cm* from the lenses. One of the lenses is converging with a focal length of 15 *cm*. Determine the nature and focal length of the second lens.

Solution

Using the lens equation $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

$$\frac{1}{30} + \frac{1}{45} = \frac{1}{f}; \ \frac{1}{f} = \frac{3+2}{90}$$

Therefore; $f = 18cm$. This is the focal length of the combination.

The formula for two thin lenses in contact is $\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$

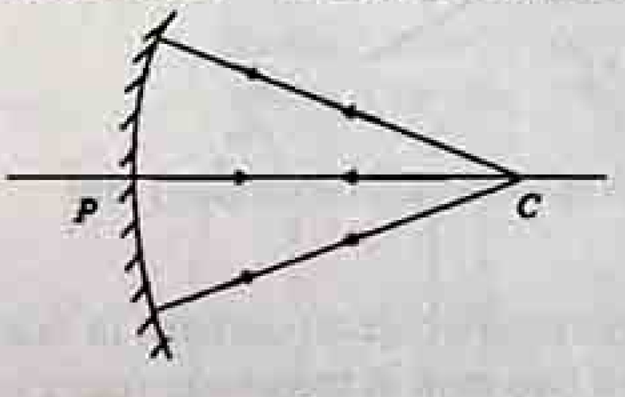
Therefore; $\frac{1}{18} = \frac{1}{15} + \frac{1}{f_2}$

$$\frac{1}{f_2} = \frac{1}{18} - \frac{1}{15} = -\frac{1}{90}$$

Therefore; $f_2 = -90cm$

Hence, the second lens is diverging and the focal length is 90*cm*

Example 20.4
A concave mirror is mounted with axis horizontal. When a pin is located at point C a distance of 18*cm* from the mirror, the pin is observed to be coincident with its image in the mirror (see diagram in Figure 20.10 (a)). A concave lens is interposed (see diagram (b)) between the mirror and C, where PM = 10 cm, and to re-establish coincidence between pin and image, the pin is moved to a new location I such that PI = 30 cm, determine the focal length of the concave lens from the diagram.



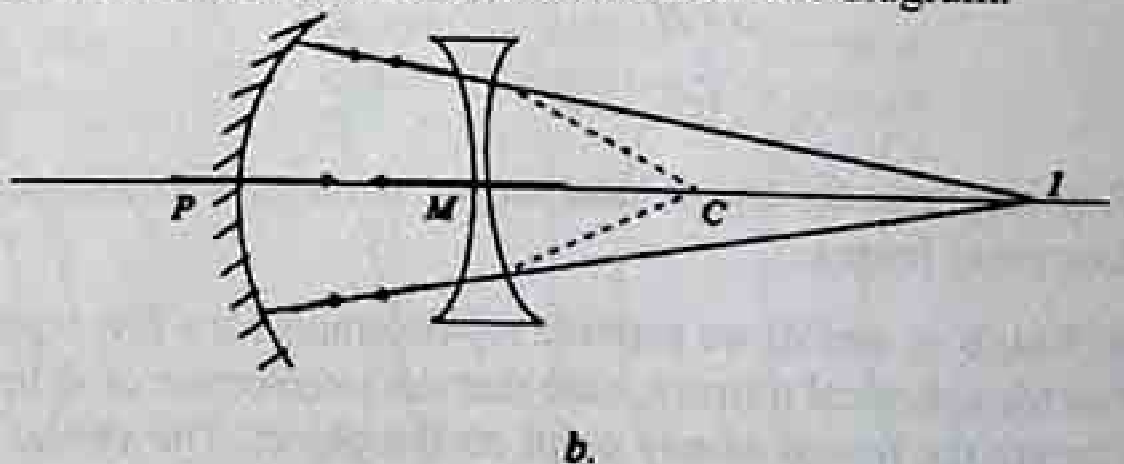


Fig. 20.10: Example 20.4

Solution
This problem demonstrates a method for determining the focal lengths of a concave mirror and a concave lens. Initially when the object and image coincide at C, the rays are hitting the mirror normally and are hence returned along the same path.
∴ C is the centre of curvature of the mirror.

$\therefore$ Focal length of mirror $= 18/2 = 9cm$. Similarly, for coincidence at I when the concave lens is interposed the rays hit the mirror normally.
$\therefore$ When produced backwards, the rays reflected by the mirror will cut the principal axis at C as shown in the diagram (b).
$\therefore$ For the concave lens.
Virtual object distance $= MC = PC - PM = 18 - 10 = 8cm$
Real image distance $= MI = PI - PM = 30 - 10 = 20cm$

Using the lens formula, $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

$$-\frac{1}{8} + \frac{1}{20} = \frac{1}{f}$$

$$\frac{1}{f} = -\frac{3}{40}$$

Therefore; $f = -\frac{40}{3} = -13.3cm$

20 .9 Conjugate Points- Newton's Relation

If object Ohas image at I, since the principle of reversibility of light holds, and object at I will give image at O (Figure 20.11), the points O, I are interchangeable and are called conjugate points. Planes through these points perpendicular to the axis are called conjugate planes. If object and image distances are measured from the focal points as shown, using the lens equation

$$\frac{1}{x+f} + \frac{1}{y+f} = \frac{1}{f}$$

$$f(y+x+2f) = (y+f)(x+f)$$

$$xy = f^2 \qquad (20.17)$$

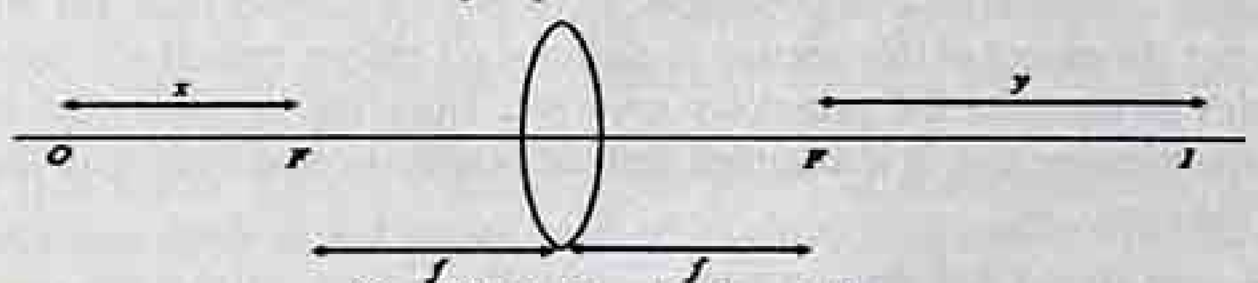


Fig. 20.11: Newton's relation

This is known as *Newton's relation*. It shows that as object approaches the lens (x decreasing), the image recedes from the lens (y increasing).

20.10 Minimum Distance between Object and Real Image in Convex Lens

There is a minimum distance between object and screen before an image can be observed on a screen. Let the distance be d (Figure 20.12).

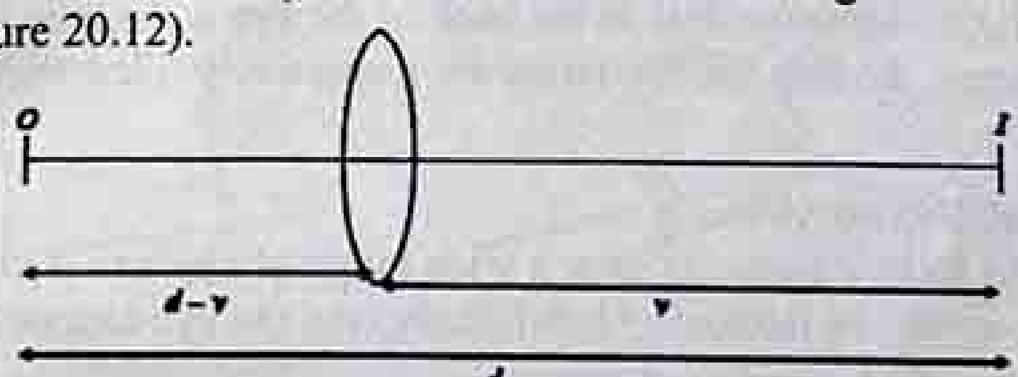


Fig. 20.12: Minimum distance between object and real image

Using the lens equation, $\frac{1}{d-v} + \frac{1}{v} = \frac{1}{f}$

$$v^2 - dv + df = 0$$

Using the general rules for solution of quadratic equation, v is real if

$$d^2 - 4df \geq 0$$

or

$$d \geq 4f \quad (20.18)$$

Therefore the minimum distance between object and screen to obtain image on screen is $4f$.

20.11 Spherical Aberration in Lenses

As with mirrors, spherical aberration is also observed in lens images. For wide angle beam, a point image is not observed for a point object. The best image for point object O (Figure 20.13) is found between I and I_1.

The circular image is called the circle of least confusion. The defect may be minimized by allowing light to pass only through the centre of the lens. This will lead to a sharp image but the brightness will diminish because of the limited amount of light passing through the lens. The defect is most marked if the angle involved in the refraction is large, and it is minimum if the angles are kept small. It is best to arrange that the bending be shared between two surfaces of the lens, In fact to reduce spherical aberration telescope eyepieces are made of a combination of two lenses usually planoconvex, in which case four surfaces are involved in the deviation. It can be shown that spherical aberration is minimum if the separation of the two lenses is equal to the difference between the focal lengths.

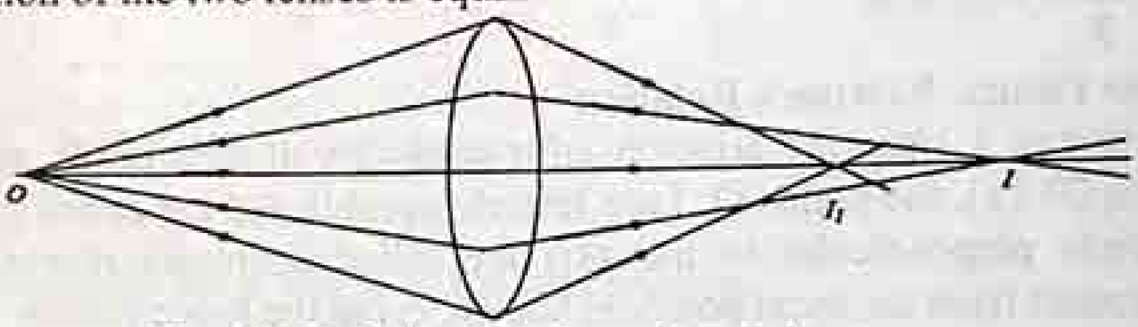


Fig. 20.13: Spherical aberration in lens.

20.12 Optical Instruments

The design of efficient optical instruments is the ultimate purpose of Geometrical Optics. Optical instruments include microscopes, telescopes (reflecting and refracting), binoculars, cameras, etc. They employ a combination of lenses and are designed to focus images of objects either on the retina of the eye or on a film. For microscopes and telescopes, objects viewed with their aid are made to appear much larger. It is not the object of this section to describe all known optical instruments. We will just use a few examples to illustrate the principles involved. However, we will start by defining a few terms relevant to the viewing eye; it is assumed that the reader is conversant with the operation of the optical system of the eye.

20.13 The Human Eye

The human organ of vision is the eye. The eye has an outer covering, the *sclera*, which protects the eyeball and the interior of the eye. Behind this layer is the iris, a circular diaphragm whose central aperture is the pupil. The amount of light entering the eye via the pupil is regulated by the iris (a camera-type diaphragm). The eye has a crystalline lens held in place by the ciliary muscles which can contract or expand to vary the shape, curvature, power and focal length of the lens, a process called accommodation. The light –sensitive area at the back of the eye is the retina, where images form. The spaces between the cornea, the lens and the retina are respectively filled with fluids called the aqueous and vitreous humor.

Far and near points of distinct vision

The ability of the eye to focus on objects over a wide range of distance is called accommodation. The focal length cannot, however, be decreased indefinitely, and the eye cannot focus on the retina images of objects closer or farther than a certain distance:

(i) The far point is the position of the farthest object that can be focused by an unaided eye. The far point for a normal eye is at infinity.

(ii) The near point is the least distance of distinct vision, D. It is normally about $25cm$ for a young adult, i.e. for a normal eye.

The Power of a normal eye

The power of a normal eye is $1/D$, where D is the diameter of the eyeball when relaxed. $\frac{1}{f}=\frac{1}{d}+\frac{1}{0.25}m$ when focused on the near-point. The power of accommodation of the eye is thus $\frac{1}{f}-1=4diopters$. However, the entire optical system of the eye is comparable to a lens of power about $58.5diopters$.

Optical Defects of the Eye

The inability of the eye to focus objects properly due to age and other factors is considered a defect. The defects of vision include the following:

(a) Myopia (short sight). When the eyeball is too long, it results in the focal length of the eye being too short. Thus, distant objects are brought by the relaxed eye to a focus in front of the retina, as shown in Figure 20.14. The image becomes blurred.

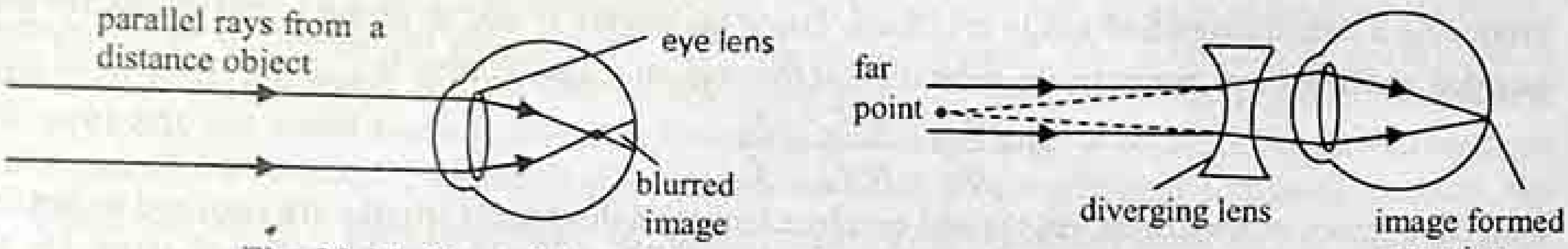


Fig. 20.14: Short sight and its correction

This defect can be corrected with spectacles using a suitable diverging lens as illustrated in Figure 20.14 (b). The focal length of the correcting lens is computed from the lens equation. For the bespectacled eye the least distance of distinct vision is no longer d but x, where $\frac{1}{x}=\frac{1}{d}-\frac{1}{f}$.

(b) Hypermetropia (long sight): When the eyeball is too short or the focal length of the eye is too long, nearby objects are brought to a focus behind the retina, causing blurred vision (Figure 20.15). Correction of this is effected with spectacles using a converging lens of suitable focal length, as shown in Figure 20.15(b).

(c) Presbyopia (old sight): This is a defect in the power of accommodation of the eye. That is, where the eye lens hardens and losses some of its elasticity (usually with the onset of old age), the ciliary muscle is also weaken. Correction can be affected by the use of spectacles with biofocal lenses to correct for both the near and far points.

(d) Astigmatism: This defect arises from lack of symmetry of the cornea, with varying curvatures in its different planes. Thus, various images are formed thereby straining the eye. Point objects do not from point images on the retina, and horizontal and vertical lines will appear to vary in intensity.It can be corrected with cylindrical lenses placed at the suitable orientation.

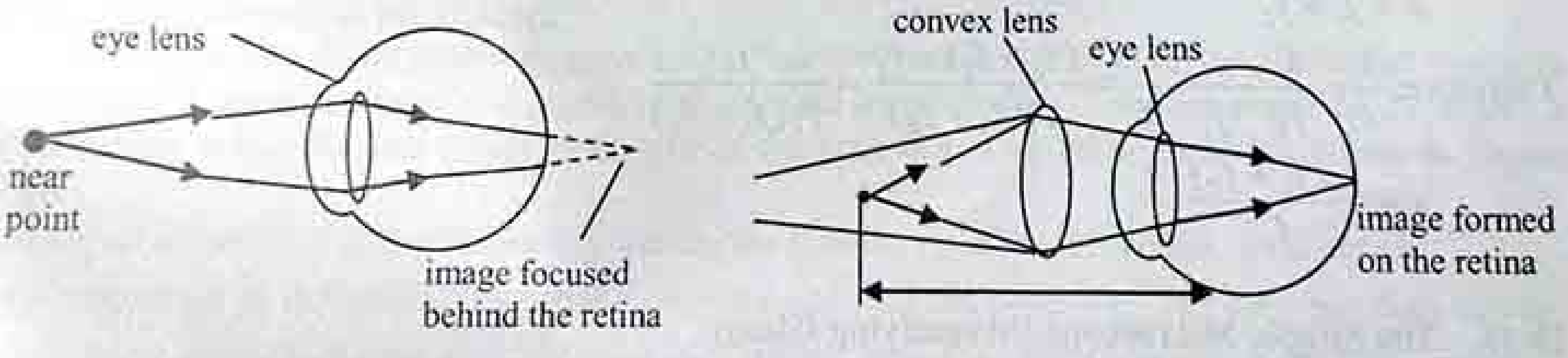


Fig. 20.15: Long sight and its correction

20.14 Refracting Telescopes (Astronomical Telescope)

A telescope is an optical device which enables the details of very distant objects (e.g stars) to be examined. A very simple telescope is the astronomical telescope shown in Figure 20.16. It consists of two converging lenses (compare with the compound microscope) but the objective has a long focal length, f_o, which coincides with the short focal length, f_e, of the eyepiece.

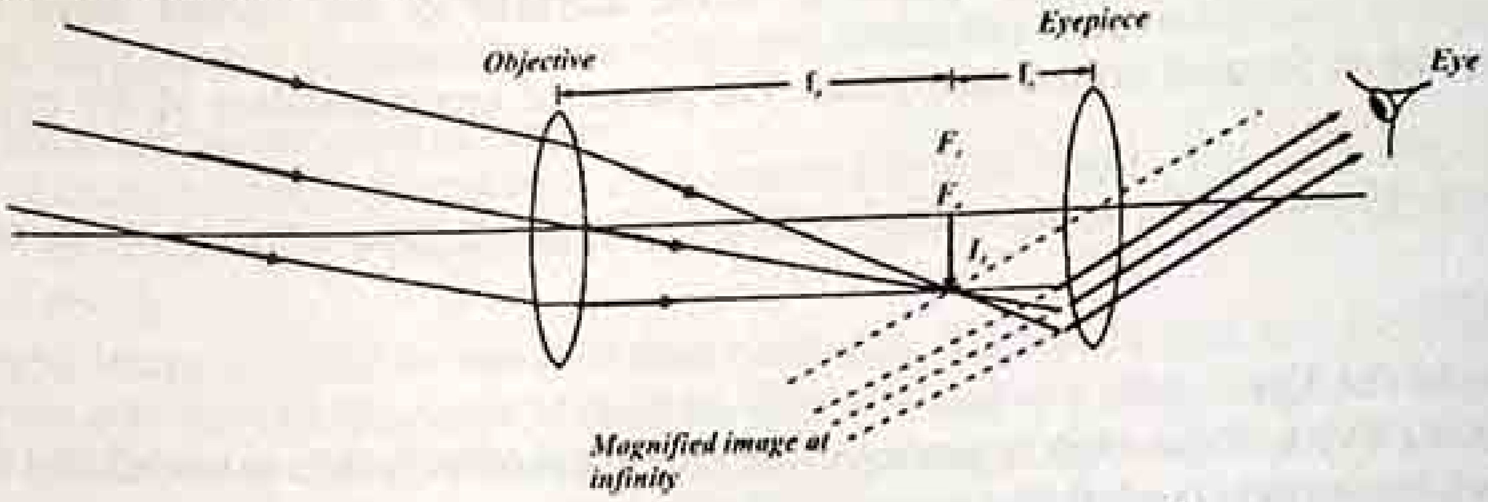


Fig. 20.16: Image formation in the refracting telescope

Parallel rays coming from an object at infinity will converge at the second point of the objective which coincides with the focal point of the eyepiece. The eyepiece then acts as a magnifying glass producing a virtual magnified image at infinity. Using the angles θ and ϕ as the visual angles for the unaided and aided eye, respectively, we have that the angular magnification is

$$m = \frac{\phi}{\theta} = \frac{-h/f_e}{-h/f_o} = \frac{f_o}{f_e} \qquad (20.19)$$

Objective lenses of long focal lengths and eyepiece lenses of short focal lengths are required to obtain high magnifications. In practice, telescopes have objective lenses with focal lengths of about $1m$ to $5m$. In the design of high performance telescope, the required *field of view*, the resolving power of the lenses and the ways of eliminating aberration, must be considered. More than two lenses are used in some telescopes, while mirrors (plane, parabolic, etc.) and other reflectors are employed to serve specific functions in most professional telescope.

Example 20.5

A telescope in normal adjustment has objective of local length f_1 and an eyepiece of focal length f_2. Show that the eye-ring is located at a distance of $f_2(f_1 - f_2)/f_1$ from the eyepiece.

Solution

For telescope in normal adjustment, the final image is formed at infinity.

In Figure. 20.16, f_e and f_o coincide.

Distance between objective and eyepiece $= f_1 + f_2$. The eye-ring is the image of the objective in the eyepiece.

Suppose the distance is V from the eyepiece.

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{f_2}$$

$$u = f_1 + f_2$$

Therefore; $\frac{1}{v} = \frac{1}{f_2} - \frac{1}{f_1 + f_2} = \frac{(f_1 + f_2) - f_2}{f_2(f_1 + f_2)} = \frac{f_1}{f_2(f_1 + f_2)}$

$$v = \frac{f_2(f_1 + f_2)}{f_1}$$

20.15 The Simple Microscope (Magnifying Glass)

The microscope is an instrument that aids the eye in obtaining considerable angular magnification of close micro-sized objects. Ordinarily, two rays from a nearby object with an angular separation of about 0.03°, or less, can hardly be distinguished by a normal eye. This threshold angular separation is called the visual acuity, which can be enhanced with a simple magnifier.

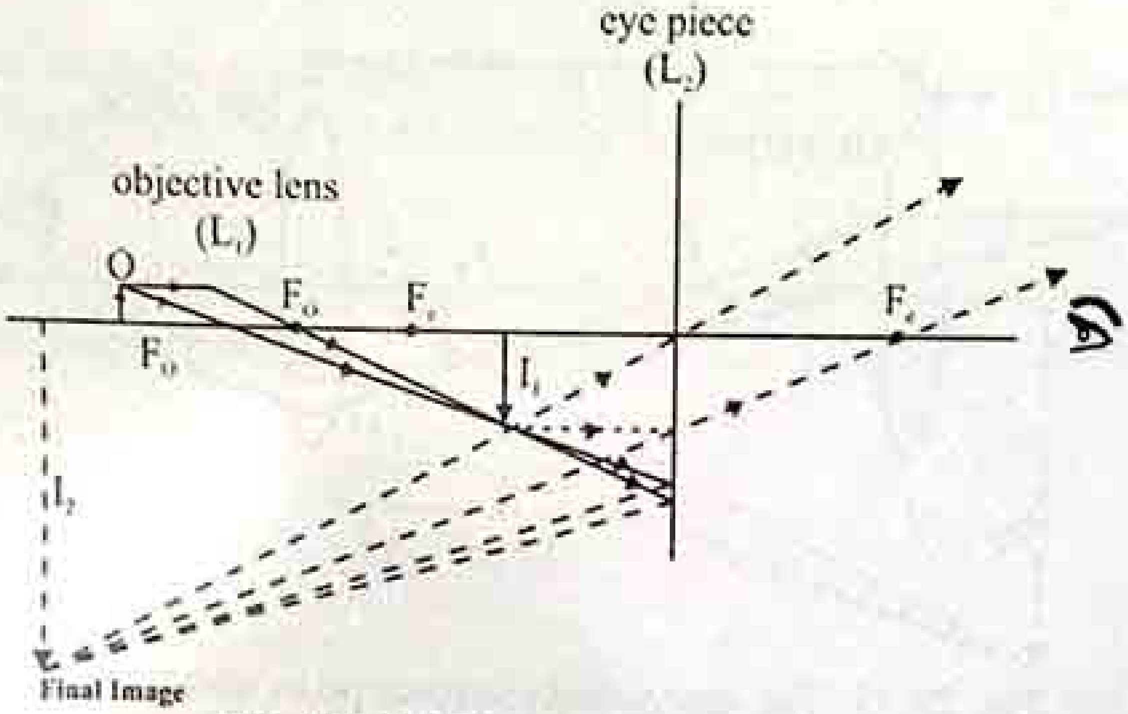


Fig. 20.17: The Simple microscope

A simple microscope consists of a single convex lens which produces an erect and magnified virtual image of an object placed between the lens and its focal point. The maximum magnification is obtained when the image is formed at the eye's near point, D. The size of the image on the retina is measured by the visual angle θ. Suppose a normal unaided eye sees the object of height h at an angle of vision θ, $\tan\theta = h/D$, where D is the near point. Then looking through a magnifying glass, the eye sees the object at the visual angle θ, where $\tan\phi = h/f$. Assuming small angles, $\tan\theta \approx \theta$, we then have, $\theta = h/D$ and $\phi = h/f$.

So that, angular magnification is,

$$m = \frac{D}{f} = \frac{25}{f} \tag{20.20}$$

Where f is the focal length of the magnifier measured in centimeters. In practice, however, the angular magnification for the magnifying glass is limited by lens aberration and image quality is also impaired.

Linear magnification, $$m_L = \frac{h_1}{h} = \frac{v}{u} \tag{20.21}$$

From the lens equation, $\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$

And multiplying by v gives, $\frac{v}{u} = \frac{v}{f} - 1 = m_L$

But since $v = D$, we obtain, $m_L = \frac{D}{f} - 1$

20.16 The Compound Microscope

This device permits us to obtain angular magnifications much higher than the limits for a simple microscope. A combination of two lenses of short focal length is used. The lens nearer to the object is the objective, while the lens closer to the eye is the eyepiece. This arrangement is shown in Figure 20.18.

If an object of height h is placed just beyond the focal point F_o of the objective, an inverted magnified image h_1 is formed between the eyepiece and its focal point F_e. The eyepiece then acts as a magnifying glass producing a highly magnified virtual image h_2. This image may be formed at the near point, D, of the observer.

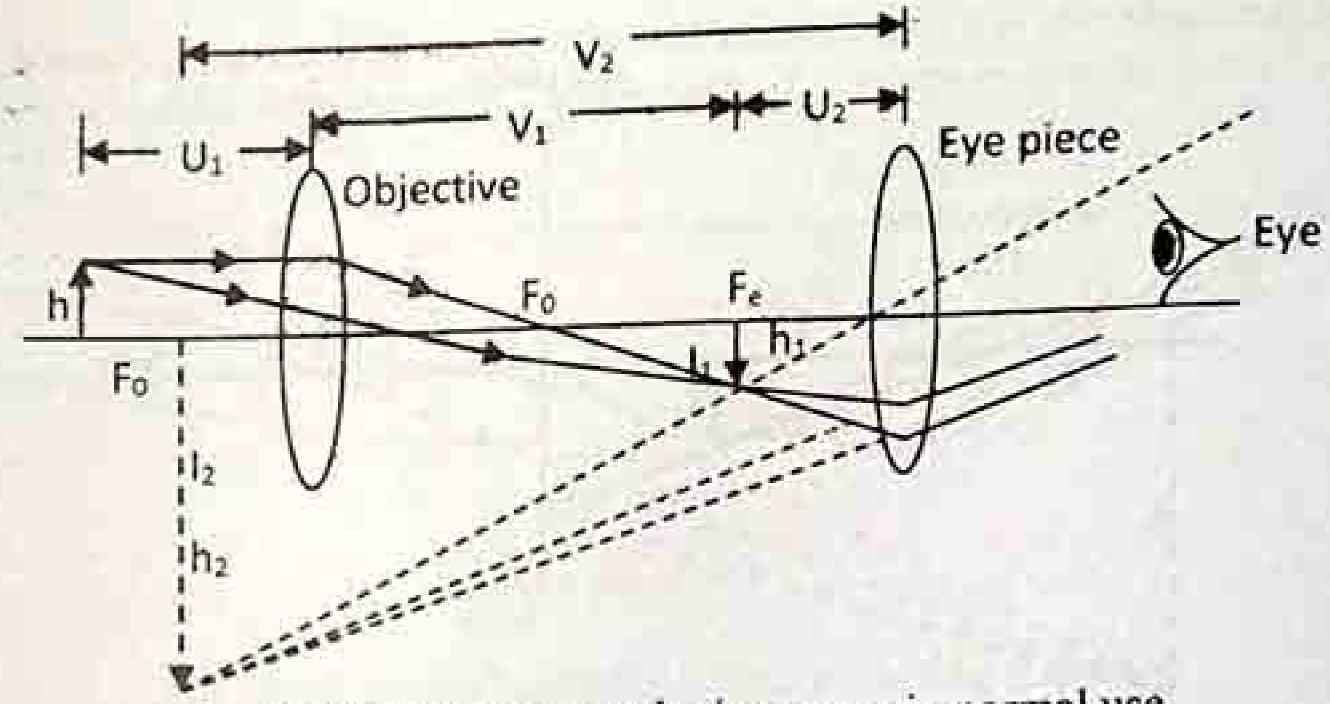


Fig. 20.18: The compound microscope in normal use

The magnifying power of the microscope is defined as the lateral magnification produced by the objective lens multiplied by the angular magnification produced by the eye lens.

Thus; $m = \frac{v_1}{u_1} = \frac{D}{u_2}$

Using $\frac{1}{u_1} + \frac{1}{v_1} = \frac{1}{f_o}$, $\frac{1}{u_1} = \frac{1}{f_o} - \frac{1}{v_1} = \frac{v_1 - f_o}{v_1 f_o}$

Therefore, $m = -\left(\frac{v_1 - f_o}{f_o}\right) \times \frac{D}{u_2} \approx -\left(\frac{F'_o F_e}{f_o}\right) \times \frac{D}{u_2}$

The negative sign implies that the image is virtual. Also, if the final image is formed at infinity, $u_2 = f_e$ and if formed at home other finite distance, u_2 is still very close to f_e. Thus,

$$m = -\left(\frac{F'_o F_e \times D}{f_o f_e}\right)$$

The distance $F'_o F_e$ is called the *tube length L* of the microscope. Thus,

$$m = -\left(\frac{D \times L}{f_o f_e}\right) \quad (D = 25cm) \qquad (20.22)$$

where the focal lengths are measured in *centimeters*. In high-resolution microscopes, the objective and eyepiece lenses are multicomponent lenses which reduce lens abreactions, and magnification as high as 2,500 can be achieved.

20.17 The Camera

The camera works on the principle of Figure 20.9(a) where the light sensitive film is located at*I*. Here, the image distance is approximately equal to the focal length of the lens. A facility is usually available for adjusting the distance between the lens and the film to ensure a sharp and well-focused image of the object. The light affects the silver halide grains of the film and when the film is developed these grains are converted into 'black grains of metallic silver. When the film is put in a fixing bath the unaffected silver halide grains are removed; the final result is the negative as we know it which can then be used to obtain photographic prints.

The amount of light falling on the film will be proportional to the area of the lens, while the length of the image will be proportional to *f*. Therefore the amount of light per unit area of image or brightness is proportional to D^2/f^2 where *D* is the diameter of the lens. The relative time of exposure *r* required is inversely proportional to *D*. Therefore

$$t \propto \frac{f^2}{D^2} \qquad (20.23)$$

It is common practice to define a term *f*-number equal to the ratio. Therefore

$$t \propto (f-number)^2 \qquad (20.24)$$

The f-number indicates the light-gathering power; hence it is referred to as the speed of the lens. Thus, from equation 20.24, the smaller the f-number the higher the speed.

Example 20.6

A photograph is taken with a lens of speed $f/4$ and an exposure time of $1/100\,s$. If the same picture is to be retaken with an exposure of $1/50\,s$, what should be the speed of the lens?

Solution

From equation 20.24, $\frac{t_1}{(f_1-number)^2}=\frac{t_2}{(f_2-number)^2}$

Therefore; $\frac{1/100}{4^2}=\frac{1/50}{(f_2-number)^2}$

This implies that $(f_2-number)^2=2\times 4^2$

$$f_2-number=4\sqrt{2}=5.6$$

Therefore, speed of lens $=f/5.6$

20.18 The Periscope

This is a simple optical device utilizing the principle of total internal reflection. It is used for viewing (otherwise) inaccessible objects. The principle of the periscope is illustrated in Figure 20.19. The reflectors shown are glass prism.

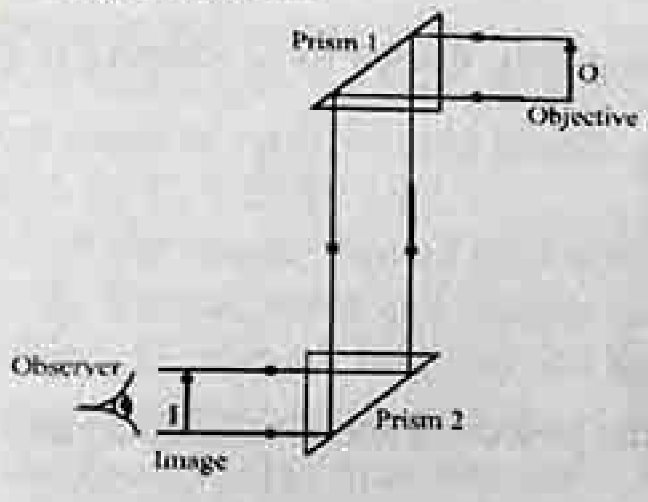


Fig.20.19: The periscope

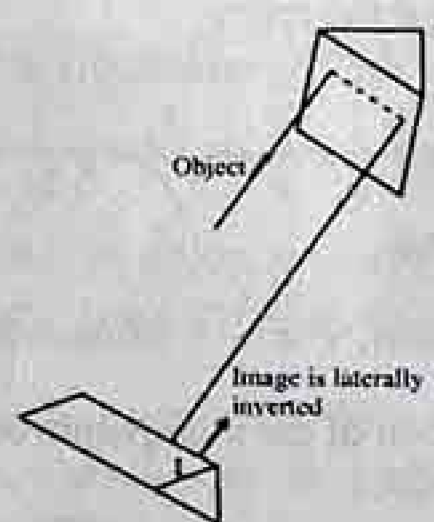


Fig. 20.20: Prism binoculars

20.19 Prism Binoculars

These class of optical instruments are widely used as field glasses. They consist of short astronomical telescopes containing two right-angled isosceles prisms. One of the prisms has its refracting edge vertical while that of the other prism is horizontal, as shown in Figure 20.20. This arrangement of the prisms between the two lenses produces an erect final image of the object.

The optical path of the ray is about 3 times the objective to eyepiece distance, since the direction of the ray is reversed twice. Thus, the angular magnification of this arrangement is the same as that of astronomical telescope having three times the focal lengths of the objective and the eyepiece.

20.20 The Projector

This is also called the projection lantern. Its function is to project slides or films onto a screen with a reasonable degree of clarity and magnification. The essential components of a projector are arranged as shown in Figure 20.21. It consists of a powerful *source of light*, S; a *condensing lens* (comprising two plane-convex lenses) arranged to have conjugate foci and S at the location of the *objective lens*, and a screen.

The slide or film is placed between the condensing lens and the objective, and the latter is an adjustable achromatic lens concentrates and focuses the light on the screen.

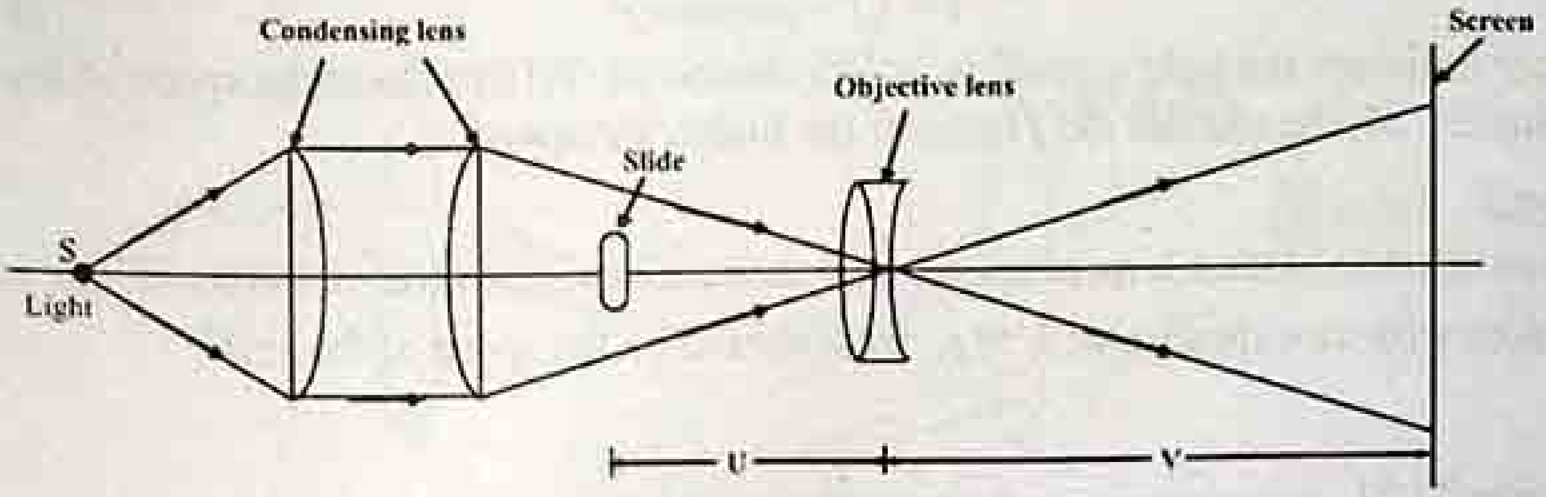


Fig. 20.21: The Projector

The linear magnification, m of the slide $= v/u$ where u and v are the respective distances of the screen and the slide from the objective. Using the lens equation, we can write m is equal to $v/u = v/f - 1$ where f is the focal length of the objective. Hence, by using an objective whose focal length is small compared with v, a high magnification is obtained.

Example 20.7

The near point of a defective eye is $100cm$ in front of the eye

(i) What does this mean?

(ii) What lens should be used to enable an object at $25cm$ in front of the eye to be seen clearly?

Solution

(i) The defect is long sight or that the focal length of the eye is too long. Thus, a blurred image is formed behind the retina.

(ii) Correction of the Hypermetropia is by converging lens, whose focal length f is given by

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \text{ or } \frac{1}{f} = \frac{1}{0.25m} + \frac{1}{1m}$$

Therefore, $f = 0.33m$

The power of the lens, $P = \frac{1}{f} = 3diopters(D)$.

Note that the location of the near points depends on the power of accommodation and hence on age, as shown in Table 20.1.

Table 20.1: Variation of Near Point of the Eye with Age

Age of year	10	20	30	40	50	60
Near points (cm)	7	10	14	22	40	200

Summary

(i) Snell's laws of refraction are:
The incident ray, the refracted ray and the normal at the point of incidence all lie in one plane

$$\frac{\sin i}{\sin r} = \text{constant}$$

(ii) The optical invariant ($n \sin \theta$ = constant) is a most important relation when considering refraction through several media.

(iii) Total internal reflection occurs when light passes from a denser to a less dense medium. At a certain angle of incidence, called critical angle c, the refracted ray just grazes the interface. For angles of incidence greater than c, the ray is not refracted but is totally internally reflected.

(iv) The lens equation is $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

(v) When two thin lenses of focal lengths f_1 and f_2 are in contact, the combined focal length is given by $\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2}$.

(vi) The minimum distance between an object and its real image in a convex lens is $4f$ where f is the focal length of the lens.

(vii) Just as for mirrors, spherical aberration is observed in lenses with wide angle beam. It is minimized by combing two lenses instead of a single lens.

(viii) The magnification of a hand lens is $D/f-1$, where D is the least distance vision. For a compound microscope, the magnification $M = m_1 \times m_2$ where m_1 is magnification due to the objective, and m_2 is the magnification due to the eyepiece.

(ix) The telescope is used for viewing distant objects. The angular magnification is given as

$$M_a = \frac{\alpha_2}{\alpha_1} = \frac{f_{objective}}{f_{eyepiece}}$$

(x) The *f-number* of a camera, often referred to as the speed of the lens, is given by f/D, where D is the diameter of the lens.

Exercise 20

20.1 A ray of light in air incident at an angle of 45° on the boundary with water. Calculate the angle of refraction if the relative refractive index of air is 0.55?
A. 25.60° **B.** 22.89° **C.** 23.75° **D.** 30.17°

20.2 Which of the following statements is not correct about total internal reflection
A. For a total internal reflection to occur, the light must come from a denser medium to a less medium.
B. At the critical angle c, the refractive index of the glass ${}_a n_g$ is given as ${}_a n_g = 1/\sin c$.
C. For total internal reflection to occur, the critical angle must not be exceeded.
D. At the critical angle, the angle of reflection is equal to 90°.

20.3 The critical angle at an air–liquid interface is 50°. Calculate the refractive index of the liquid?
A. 1.60 **B.** 1.50 **C.** 1.31 **D.** 1.42

20.4 The nature of the image formed by a convex lens when the object is placed closer to the lens than f is **A.** Virtual, erect and magnified **B.** Virtual, erect and diminished **C.** Real, erect and magnified **D.** Virtual, inverted and magnified

20.5 A near sighted woman needed to see at 100*cm* away from her eyes. What type of corrective lens and the focal length must she use? **A.** 33.50*cm*, convex lens **B.** 33.50*cm*, concave lens
C. 43.50*cm*, convex lens **D.** 33.33*cm*, concave lens

20.6 Which of the following is not a correct equation of thin lens? Where the notations have their usual respective meaning.

A. $\frac{n_1}{u}+\frac{n_2}{v}=\frac{n_2-n_1}{R}(convexlens)$ **B.** $\frac{n_1}{u}-\frac{n_2}{v}=\frac{n_2-n_1}{R}(concavelens)$

C. $\frac{1}{f}=\frac{\mu}{u}+\frac{1}{v}$ **D.** $\frac{1}{f}=(n_2H)\left(\frac{1}{R_1}+\frac{1}{R_2}\right)$

20.7 A compound microscope consists of objective and eye piece lenses of focal length 0.50*cm* and 1.00*cm* respectively. An object is placed 0.55*cm* from the objective lens and the final image is formed at infinity. Calculate the magnifying power of the glass and the tube L.
A. -275, 5.5*cm* **B.** -625, 12.5*cm* **C.** -275, 7.5*cm* **D.** -5.75, 6.5*cm*

20.8 Which of the following is not an optical instrument?
A. The human eye **B.** Camera **C.** Periscope **D.** Pyranometer

20.9 A light ray travelling from glass to air undergoes refraction. Calculate the angle of refraction to the nearest whole number if the incidence angle is 30° and the refractive for air and glass are 1.0 and 1.5 respectively. **A.** 75° **B.** 65° **C.** 49° **D.** 55°

20.10 A student is only able to read the lecture notebook when it is held at 65.00*cm* from his eyes. What focal length and power must the lenses of his reading glasses have to bring his near point to the normal value?
A. 4.00*cm*, 0.025*D* **B.** 5.00*cm*, 0.045*D* **C.** 30.00*cm*, 0.030*D* **D.** 45.00*cm*, 0.035*D*

20.11 (a) Explain clearly what you understand by the critical angle between two media.
(b) Explain the conditions under which total internal reflection will occur. Give examples of uses made of total internal reflection.

20.12 By means of a ray diagram, describe the nature, size and location of the image formed of an object placed between a thin converging lens and its focal point.

20.13 The difference between the apparent and real depths of a swimming pool filled with water is $1.5m$. Calculate the depth of the swimming pool if the refractive index of water is 1.33.

20.14 A ray of light passes from air into glass of refractive index 1.5. If the angle of refraction is $40°$, calculate the angle of incidence.

20.15 A ray of light is incident on the surface of water at an angle of $30°$. Calculate the deviation suffered by the ray in the water.

20.16 The critical angle at an air-liquid interface is $45°$. Calculate the refractive index of the liquid.

20.17 An object is placed (a) $15cm$ (b) $25cm$ from a thin diverging lens of focal length $20cm$. Determine the nature, position and magnification of the image in each case.

20.18 Two thin lenses are in contact. One is a converging lens of focal length $25cm$ and the other is a diverging lens of focal length $20cm$. What is the focal length of the lens combination? Is the combination converging or diverging?

20.19 A thin converging lens of focal length $30cm$ is in contact with another thin lens. An 'object at a distance of $25cm$ from the lenses forms an image on the same side of the lenses at a distance of $10cm$. Determine the nature and focal length of the second lens.

20.20 A compound microscope consists of an objective of focal length $0.5cm$ and an eyepiece of focal length $1cm$. The two lenses are separated by distance $14cm$. What is the position of the object it the final virtual image is formed at a distance of $25cm$ from the eyepiece? Calculate the linear magnification of the instrument.

CHAPTER 21
PRISM AND DISPERSION

21.0 Introduction

A prism is an object that is transparent, made of glass, and possesses two plane refractive surfaces that are inclined to each other. It is used in prism binoculars, measure refractive index of glass, submarine periscope etc.

21.1 Refraction at Prism Surface

Consider a prism with refracting angle A. This is the angle between the refracting surfaces; the line of intersection of the refracting surfaces is the refracting edge.

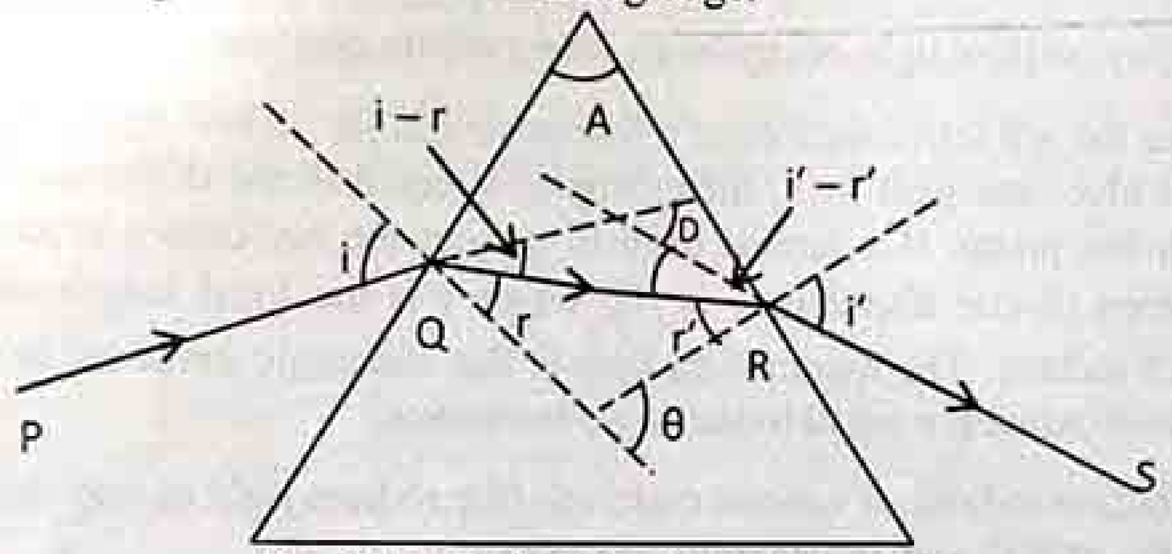


Fig. 21.1: Refracting through a prism

A ray PQ incident on one of the refracting surfaces emerges in air as RS after refraction through the prism (Figure 21.1). Recall that for refraction through a parallel-faced glass block, the ray is undeviated. It is deviated in a prism since the emergent ray is bent towards the base of prism. The angle of deviation is the angle D through which the direction of the incident ray is turned to arrive at the direction of the emergent ray. If i is the angle of incidence and the other angle involved in the refraction are as given in Figure 21.1, we have, using properties of triangles and quadrilaterals

$$D=(i-r)+(i'-r')=(i+i')-(r+r')$$

Also, $\theta=r+r'=A$. At Q, if n is the refractive index of the prism material,

$$n=\frac{\sin i}{\sin r}$$

The smallest deviation is called minimum deviation $D_{\min}$ obtained when the ray passes symmetrically through the prism. In that case, $i=i', r+r'$. Therefore from equations above, we have

$$i=\frac{1}{2}(A+D_{\min});\ r=\frac{1}{2}A$$

Therefore, $$n=\frac{\sin\frac{1}{2}(A+D_{\min})}{\sin\frac{1}{2}A} \qquad (21.1)$$

This formula is well known for determining the refractive index of prism material. Apart from experiments in the laboratory using optical pins, the angle of minimum deviation and the refracting angle of the prism could be determine to great precision using the spectrometer.

21.2 Dispersion by a Prism

Suppose a narrow beam of white light is incident on a prism as shown in Figure 21.2. After passing through the prism, the beam is split into a band of colours called a spectrum. The spectrum of white light consists of Red, Orange, Yellow, Green, Blue Indigo and Violet (ROYGBIV). The splitting of white light into its component colour is known as dispersion. The rainbow which we see is due to dispersion of sunlight by the water droplets in the air, which behave like tiny prism.

From Figure 21.2 it is clear that the red rays are the least deviated and the violet rays the most deviated. It is obvious that at the point of incidence P, the angle of refraction r is least for the violet ray.

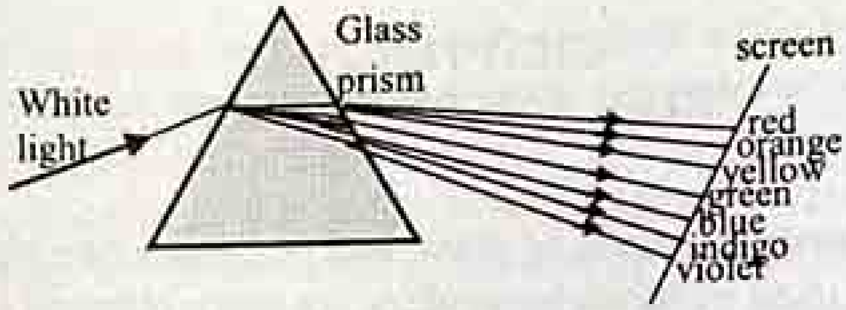


Fig. 21.2: Dispersion of white light by a prism

Since the angle of incidence i is the same for all the rays, it entails that the refractive index of the material, $n = \frac{\sin i}{\sin r}$ depends on the colour or wavelength of the light. It is least for the red rays and largest for the violet ray; yellow light occupies an intermediate position

In Table 21.1 we give the refractive indices of a few materials for yellow light. The angular dispersion between the red and blue emergent rays, for example, is equal to the difference in deviation of two colours produced by the prism. If a lens is used to project the ray onto a screen, the length of the spectrum on the screen (linear dispersion) $= f\theta$, where f is the focal length of lens and θ is the angular dispersion in radians. The assumption here is that the angle involved is small; therefore the sine of the angle is approximately equal to the angle in radians.

Table 21.1: Refractive indices of various materials (for sodium light of wavelength 589.3 mm)

A **Gases at 0°C and 760 mmHg pressure**

Gas	*n*
Air	1.000293
Chlorine	1.000781
Hydrogen	1.000132

B **Liquids at 20°C**

Liquid	*n*
Acetone	1.359
Benzene	1.501
Ethyl alcohol	1.362
Glycerin	1.472
Methyl iodide	1.756
Paraffin	1.430
Water	1.333

C **Solids**

Solid	*n*
Canada balsam	1.530
Diamond	2.417
Glass (dense Flint)	1.655
Glass (ordinary crown)	1.517
Ice (H_2O)	1.310
Plastic (Polystyyrene)	1.598
Quartz	1.458
Rock salt	1.544

Note that if a second prism with its vertex to the base of the first was placed to receive the emergent light, a white spot is produced on the screen because of the recombination of the differently coloured beams into white beam (Figure 21.3)

The phenomenon of dispersion was first discovered by Newton (1666). However, Newton's spectrum of sunlight was an impure spectrum because the different coloured images overlapped. A pure spectrum contains monochromatic images only.

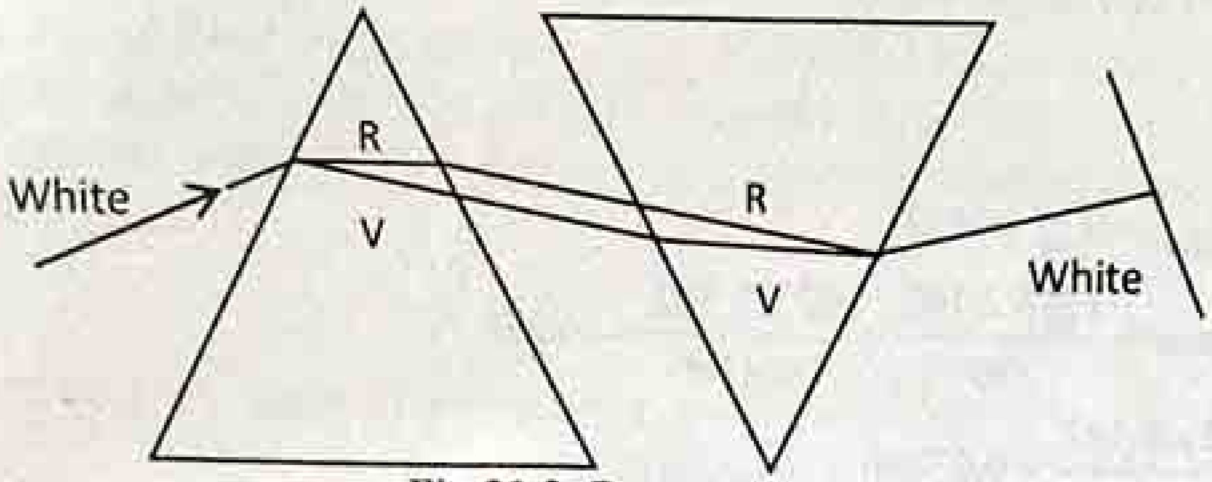


Fig.21.3: Recombination of dispersed light

Overlapping can be eliminated if a narrow pencil of white light is used, and if the rays of particular colours emerge from the prism parallel, so that each coloured beam can be brought to a separate focus. The spectrometer can be used to achieve this objective (Figure 21.4). The spectrometer consists essentially of a rotatable table and two horizontal tube, collimator and telescope.

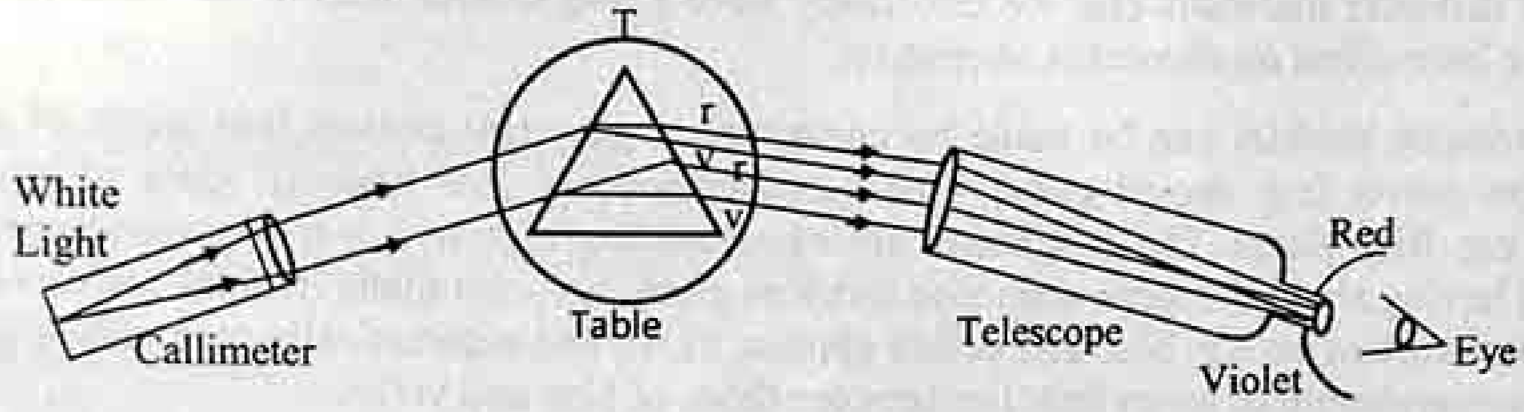


Fig. 21.4: The Spectrometer

The telescope can also be rotated about an axis perpendicular to the table, but the collimator is fixed. The telescope carries a vernier scale which moves over a circular scale attached to the table so that its angular position may be noted accurately. The function of the collimator is to render parallel the rays from the source. It consists of a hollow metal tube of adjustable length having a variable slit at the source end and an achromatic lens at the other end.The function is achieved by adjusting the length of the tube to equal the focal length of the lens. The telescope is used to collect the light emerging from the prism. The achromatic objective lens forms an image of the spectrum which is viewed through the eyepiece. To observe the spectrum, the prism is placed on table T while the collimator and telescope are adjusted for parallel light. This means that parallel rays which emerge from the collimator are incident on the prism while the different coloured parallel rays emerging from the prism are brought to separate foci to be viewed through the telescope eyepiece. Therefore a pure spectrum is observed.

Example 21.1

A parallel beam of white light incident on a 60° prism passes symmetrically through a prism. Calculate the deviations of the blue and red light caused by refraction through the prism assuming refractive indices for red and blue light of 1.604 and 1.620 respectively. The light emerging from the prism is focused on to a screen by means of a lens of focal length 25*cm*. calculate the length of the spectrum seen on the screen

Solution

Since the beam passes symmetrically through the prism we have a case of minimum deviation

$$n = \frac{\sin\frac{1}{2}(A+D)}{\sin A/2}$$

For red ray, $1.604 = \dfrac{\sin\frac{1}{2}(A+D_{red})}{\sin 30}$

$\sin\frac{1}{2}(60+D_{red}) = 1.604\sin 30 = 0.802$

$\therefore \sin\frac{1}{2}(60+D_{red}) = 54°6'$

$D_{red} = 46°40'$

For blue ray, $\sin\frac{1}{2}(60+D_{red}) = 1.620\sin 30 = 0.810$

$\therefore \sin\frac{1}{2}(60+D_{blue}) = 54°20'$

$D_{blue} = 48°12'$

The angular dispersion $\theta_d = D_{blue} - D_{red} = 1°32'$

This angle converted to $\theta\, radians = \frac{1\frac{32}{60}\times 3.142}{180} = \frac{92\times 3.142}{60\times 180} radians$

Linear dispersion, $f\theta = \frac{25\times 92\times 3.142}{60\times 180} = 0.67\,cm$

21.3 Chromatic Aberration in Lenses

The image defect of having different coloured images when white light is incident on a lens is known as chromatic aberration. Since the convex and concave lenses refract rays in opposite direction, it is possible to annul chromatic aberration by combining two such lenses. However, the lenses must be made of different materials else we effectively have a single thick lens of the same material which will have little effect on chromatic aberration.

An achromatic doublet can be made by combining a powerful convex lens made of glass of low dispersive power (e.g. crown glass) with a weaker concave lens made of glass of high dispersive power (e.g. flint glass). The two lenses are cemented together with a transparent material (Canada balsam) having about the same refractive index as glass. An achromatic doublet corrects only for two different wavelengths e.g. blue and yellow (Figure 21.6); this automatically corrects for green as well since the wavelength of green light lies between those of blue and yellow.

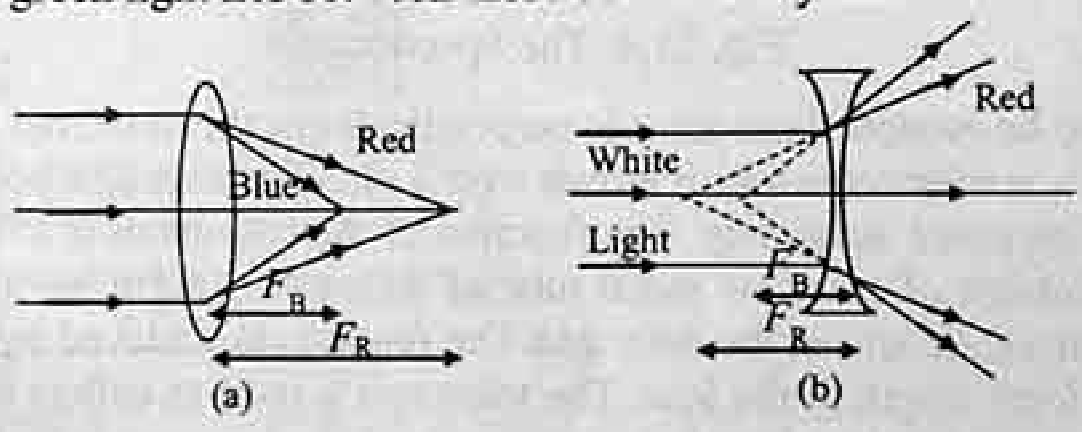


Fig. 21.5: Dispersion produced by (a) convex lens (b) concave lens

NB: The quantity $$\frac{n_B - n_R}{n_y - 1} = w \qquad (21.2)$$

Equation 21.2 is called the dispersion power of the glass. where n_B, n_R, and n_y are refractive index for blue, red and yellow lights respectively.

The focal lengths of the lenses are chosen such that the combination has a net positive focal length- it is converging. In this mode the achromatic doublet is used as objective lens in microscopes, telescope, cameras and other optical instruments to reduce colour effect.

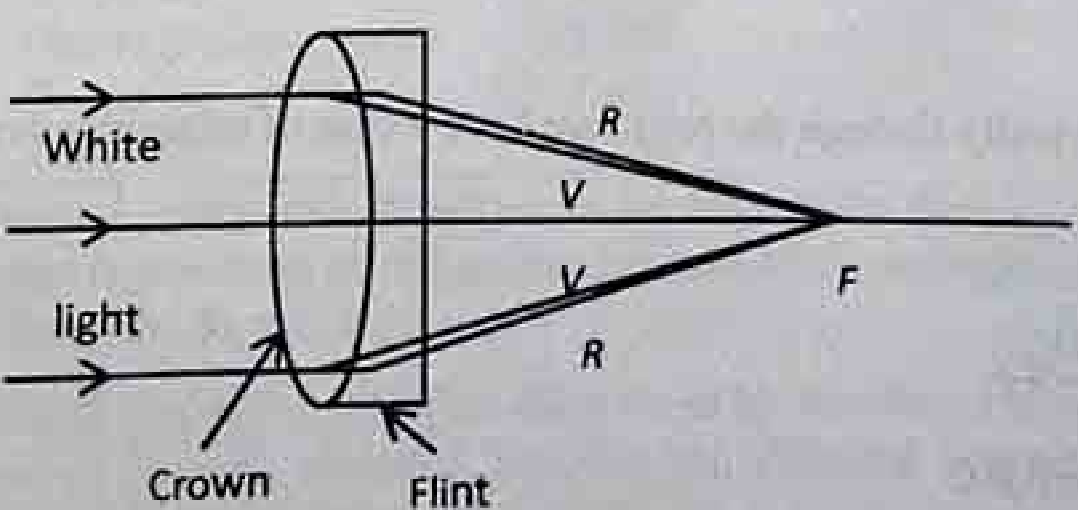


Fig. 21.6: Achromatic lens combination

Example 21.2

A concave lens of focal length 60*cm* is made of material whose refractive index for red light is 1.641 and refractive index for blue light is 1.659. That for white light = 1.65. It is combined with a convex lens of dispersive power 0.0173 to form an achromatic doublet. Calculate the focal length of the achromatic lens.

Solution

For the concave lens, mean refractive index $n = \dfrac{1.659 + 1.641}{2}$

$\therefore$ dispersive power, $w_2 = \dfrac{n_b - n_r}{n-1} = \dfrac{1.659 - 1.641}{1.65 - 1} = 0.0277$

Use formula for achromatic combination t obtain focal length of convex lens f_1

$$\frac{f_1}{f_2} = -\frac{w_1}{w_2};\ f_1 = -\frac{w_1}{w_2} f_2 = -\frac{0.0173}{0.0277} \times (-60)$$

$$\therefore \qquad f_1 = \frac{1.73}{2.7} \times 60$$

The focal length of the combination is F where

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} = \frac{2.77}{1.73 \times 60} - \frac{1}{60} = \frac{2.77 - 1.73}{1.77 \times 60} = \frac{1.04}{1.73 \times 60}$$

$$F = \frac{1.73 \times 60}{1.04} = 100\,cm$$

21.4 Types of Spectra

The spectrometer which produces a pure spectrum can be used for estimating the wavelength of the lines in the spectrum. If in place of the telescope there is a camera, we have a spectrograph in which the dispersed radiation leaves permanent impressions on a photographic plate. However, because ordinary glass absorbs wavelengths in the infrared and ultraviolet region, the prism of the instrument must be made of quartz for spectral observation in the ultraviolet region, and rocksalt for work in the infrared region. In spectral investigations, at times it is necessary to know the intensities of the different lines in a spectrum. An instrument which can measure and compare the intensities emitted from or passing through two samples is called a photometer. A spectrometer combined with a photometer is called a spectrophotometer. It is now known that spectral lines originate from the movement of atomic electrons from states of higher to lower energy. Every element has a unique spectrum, and therein lies the importance of spectroscopy as a tool in the identification of samples in many scientific fields. For example in biochemical research this technique is used for studying products formed during complex reaction. Even in medicine the state of working of a human organ can be diagnosed by using the spectrophotometric method to determine the quantity of a particular chemical present in some part of the body. There are three types of spectra.

(i) The first type is the discrete or line spectra. When the light emitted by the atoms of a luminous substance e.g. vaporized sodium or hydrogen gas at low pressure, is examined by means of a spectrometer bright lines of various wavelengths are observed. Note that the lines are images of the collimator slit; a circular slit will yield circular images. The important point is that the wavelengths are discrete for materials in which the atoms are so far apart that they exert little or no influence on one another. Typical line spectra are the visible part of the hydrogen spectrum shown in Figure. 21.7.

(ii) The second type of spectra is the band spectra. This originates from excited molecules. It consists of a series of bands each sharp at one end but fading at the other end as illustrated in Figure 21.8, the bands consist of many lines very close to each other; it is thought that the bands owe their origin to rotations and vibrations of the atoms in an electronically excited molecule.

(iii) The third type of spectra is the continuous spectrum. This is observed for incandescent solids, liquids or gases at high pressure. In this case all wavelengths are present- the atoms are so close together that the electron orbital changes of a particular atom are influenced by neighbouring atoms.

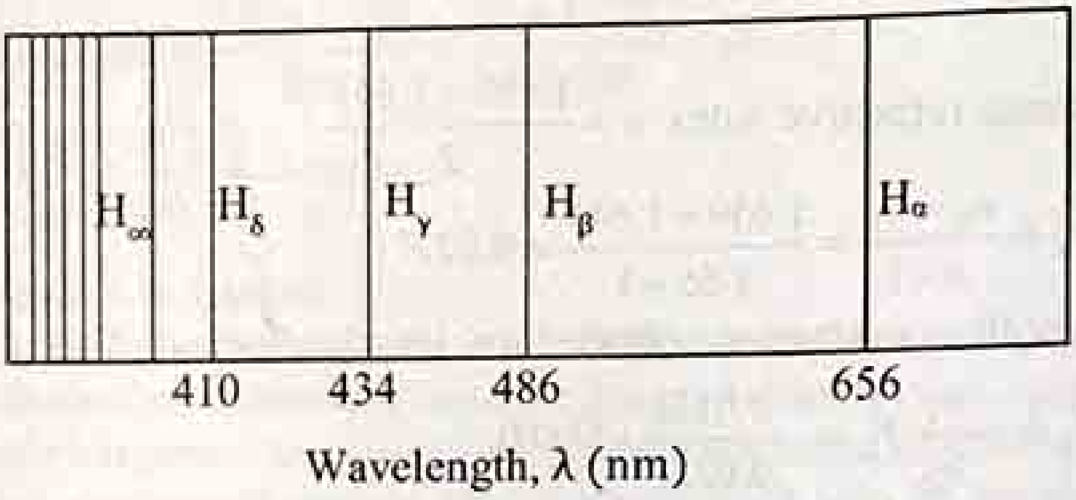


Fig. 21.7: Visible line spectrum of hydrogen

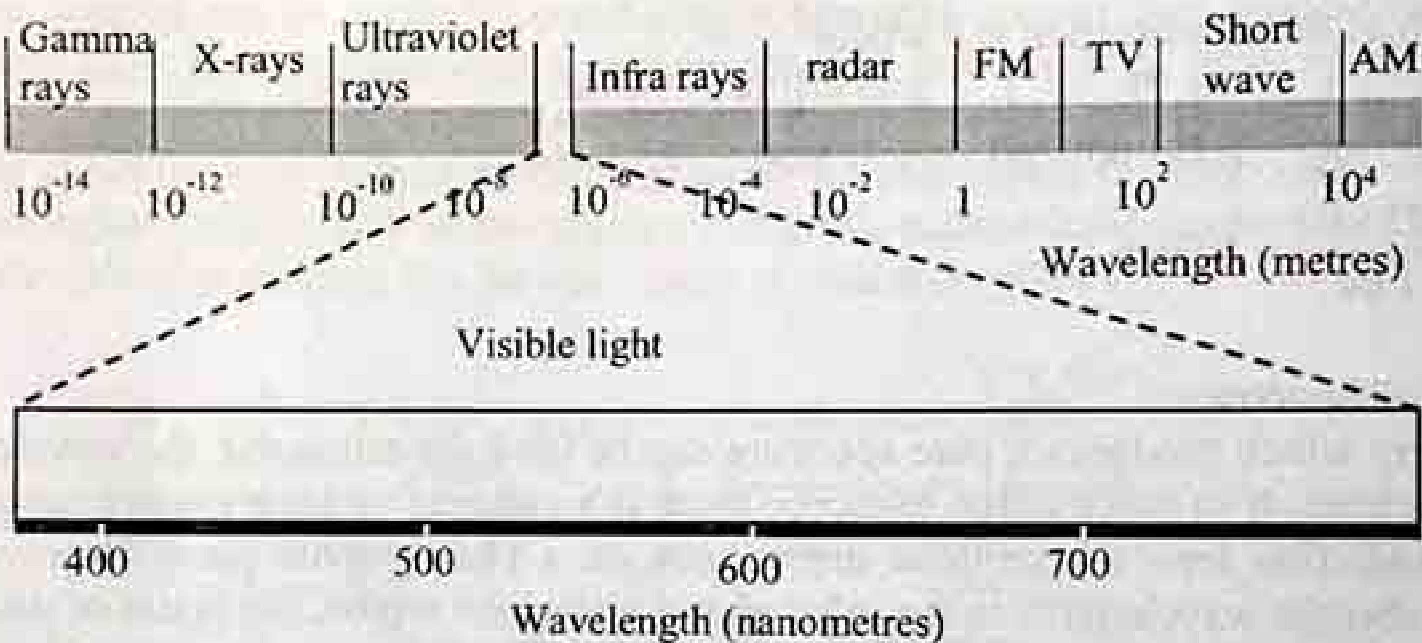


Fig. 21.8: Band spectra

These three types of spectra are called emission spectra because what is observed is light emitted by the substances. In the spectrometer bright line are seen in a dark background. But there is also the absorption spectrum. For example, if light from a source having a continuous spectrum is passed through sodium vapour before it gets to the slit of the collimator of the spectrometer, the observed continuous spectrum will have dark lines corresponding to positions of the yellow lines of the sodium emission spectrum. This is an absorption spectrum which is characteristic of the absorbing substance. According to Kirchhoff's law, a substance which emits light of a certain wavelength at a given temperature can also absorb light of the same wavelength at that temperature.

When the white light from the sun is examined in greater detail, the continuous spectrum is found crossed by some dark lines, called Fraunhofer lines. It is obvious that the vaporized elements in the sun's atmosphere absorb their characteristic wavelengths from the continuous spectrum emanating from the interior of the sun. From such studies, it was possible to identify that hydrogen and helium existed round the sun.

Summary

- Minimum deviation is the smallest deviation obtained when light passes symmetrically through a glass prism. The refractive index of the prism material is given by.

$$n = \frac{\sin \frac{1}{2}(A + D_{\min})}{\sin \frac{1}{2} A}$$

- The refractive index of a material depends on the colour or wavelength of the light. It is least for red rays and largest for violent light.
- A spectrometer can be used to produce a pure spectrum of white light.
- The image defect of having different coloured images when white light is incident on a lens is known as chromatic aberration.
- The achromatic doublet is used to correct for this defect in lenses. The condition for achromatic combination of lenses is

$$\text{Dispersive power of a lens } w = \frac{n_B - n_R}{n_y - 1}$$

- Emission spectra are classified as line, band or continuous according to their appearance and origin. The emission and absorption spectra are characteristics of a material. Kirchhoff' law is applicable. A study of the Fraunhofer lines has led to the confirmation of the existence of hydrogen and helium around the sun.

Exercise 21

21.1 Minimum deviation is obtained when ray passess
A. obliquely through the prism. **B.** symmetrically through the prism.
C. asymmetrically through the prism. **D.** with velocity *c* through the prism

21.2 Apart from experiment using optical pins, the angle of minimum deviation and the refracting angle of the prism could be determined to great precision using the
A. Speedometer **B.** Manometer **C.** Spectrometer **D.** Binoculars

21.3 The splitting of white light into its component wavelengths (colours) is called
A. Dispersion **B.** Aspersion **C.** Repercusion **D.** Interference

21.4 Rainbow which we see in the sky is due dispersion of sunlight by the ____ which behaves like tiny prisms. **A.** Beam **B.** Water droplets **C.** Acid droplets **D.** Particles

21.5 When considering the spectrum of white light, the most and least deviated rays are ____ and ______ respectively **A.** Violet, red **B.** Red, green **C.** Orange, blue **D.** Blue and red

21.6 The phenomenon of dispersion was first discovered by **A.** Albert Einstein, 1866 **B.** Fredrick Foresight, 1746 **C.** Isaac Newton, 1666 **D.** Max Planck 1920

21.7 Refractive index of a material depends on the _______ of the light.
A. Colour **B.** Dispersion **C.** Energy **D.** Displacement

21.8 The image defect of having different coloured images when white light is incident on a lens is known as **A.** Sideal aberration **B.** Chromatic aberration **C.** Coma **D.** Myopia

21.9 When a converging lens and diverging lens are combined to form one lens the following is not correct

A. Image of distant object may not converge **B.** Combined focal length is $\frac{1}{F} = \frac{1}{f_c} + \frac{1}{f_d}$

C. $\frac{1}{F} = \frac{f_c}{f_d}$ **D.** Combined focal length F will be less than f_c and f_d

21.10 Types of spectra include the following, Which of the following is incorrect?
A. Line Spectra **B.** Band Spectra **C.** Continuous Spectra **D.** Fraunhofer Spectra

21.11 The refractive index of a material for red light is less than for violet light. Explain.

21.12 What conditions are essential to obtain a pure spectrum? How are these conditions realized when a spectrometer is used to view the spectrum of white light?

21.13 A prism has a refracting angle of 60°. If the refractive index of the material is 1.5, calculate the angle of minimum deviation?

21.14 A 60° glass prism has an angle of angle of refractive of 30°. Calculate the refractive Index of the glass.

21.15 A prism made of flint glass of refractive index 1.6 has an angle of minimum deviation of 40°. Determine the refracting angle of the prism.

21.16 The refracting angle of a glass prism is 58°. A ray is incident on one of the refracting surfaces at an angle of 60° and emerges through the adjacent surface. Calculate the deviation suffered by the ray if the refractive index of the glass is 1.6.

21.17 A ray at grazing incidence on a refracting surface of a 60° prism is refracted through the prism and finally emerges from the second refracting surface. Calculate the angle of deviation of ray. The refractive index of the prism material is 1.6.

21.18 A convex crown-glass lens has for blue light a refractive index of 1.523 and a focal length of 30.0*cm*. If the refractive index for red light is 1.515, find the corresponding focal length.

21.19 A parallel beam of white light is incident on a biconvex glass lens whose surface has radii of curvature 20*cm* and 15*cm* respectively. Determine the linear chromatic aberration that is the distance between the violet and the red focal points. The refractive index of the glass for violet light is 1.663 while that for red light is 1.622.

21.20 A concave lens of flint glass of dispersive power 0.045 has a focal length of 100*cm*. Calculate the focal length of a crown glass lens of dispersive power 0.018, which will form an achromatic combination with the flint glass lens. Find the focal length of combination?

CHAPTER 22
WAVE THEORY OF LIGHT
REFLECTION, REFRACTION, INTERFERENCE, DIFFRACTION AND POLARIZATION

22.0 Introduction

A wave is a means of transmitting energy from one point to another without any net transfer of matter. Thus it is a means of transmitting information from one point to another. In this chapter we will show that waves like light obey allthe laws of reflection and refraction and in addition wave theory can explain the concept of phenomena: interference, diffraction and polarization, reflection, refraction, interference, diffraction and polarization

Wavefronts and Rays

Wavefronts are surfaces or lines in the path of a wave motion on which the disturbances at every point have the same phase. Thus every point on the wavefront vibrates in step or in phase with every other point on it. The shape of the wavefront depends upon the shape of the light source used. There are three types namely: spherical, cylindrical and plane wavefronts.

22.1 Reflection and Refraction of Waves

If a wave encounters a boundary between two different media, part of the energy is reflected back into the medium of original energy and the rest of the energy is transmitted into the other medium with a sudden change in the direction of the propagation at the boundary. Refraction of waves involves a change in the direction of waves as they pass from one medium to another. Refraction, or the bending path of the waves, is accompanied by a change in speed and wavelength of the wave. Note that the speed of a wave is independent upon the properties of the medium through which the waves travel. Thus for a wave the following laws are obeyed like light rays:

$$\frac{\sin i}{\sin r} = \frac{v_1}{v_2} = \text{constant} \tag{22.1}$$

$$n = \frac{c}{v} = \frac{\lambda_o}{\lambda} \tag{22.2}$$

where the symbols have their usual meaning.

Example 22.1

Find the absolute refractive index of a medium in which the speed of light is $2\times10^8\, ms^{-1}$. The speed of light in vacuum is $3\times10^8\, ms^{-1}$.

Solution

Absolute refractive index is $n = \dfrac{c}{v}$

$$\therefore \quad n = \frac{3\times10^8}{2\times10^8} = 1.5$$

Example 22.2

A wavefront is incident in water on a water-crown glass plane interface at an angle of incidence of 30°. Find the angle of refraction in the crown glass, where the absolute refractive indices of water and crown glass are 1.333 and 1.515 respectively.

Solution

Using Snell's law, $n_1 \sin i_1 = n_2 \sin i_2$

$$\therefore \quad 1.333 \sin 30^\circ = 1.515 \sin i_2$$

$$\sin i_2 = \frac{1.333 \sin 30^\circ}{1.515} = 0.4399$$

$\therefore$ Angle of refraction in the crown glass is $i_2 = 26°6'$.

Example 22.3

The velocity of light in air is $3\times10^{8}\,ms^{-1}$. Find the velocity and wavelength of sodium light ($\lambda = 5893\,\mathring{A}$) in glass of refractive index 1.658.

Solution:

We know that $n = \frac{c_1}{c_2}$ or $c_2 = \frac{c_1}{n}$

$$c_2 = \frac{3\times10^{8}}{1.658} = 1.809\times10^{8}\,ms^{-1}$$

Let the wavelength of sodium light in glass be λ_2. As frequency f remains constant, we have $c_1 = f\lambda_1$ and $c_2 = f\lambda_2$.

$$\frac{c_1}{c_2} = \frac{\lambda_1}{\lambda_2} = n \text{ or } \lambda_2 = \frac{\lambda_1}{n}$$

$$\lambda_2 = \frac{5893}{1.658} = 3554\,\mathring{A} = 3554\times10^{-8}\,cm.$$

22.2 Interference of Waves

It is observed in all waves, that two wave motions may combine together such that in some places they destroy each other while in others they strengthen each other. In other words, two or more isolated portions of a wave front of the same source or waves from *independent* coherent sources can be superimposed to produce a combined field effect referred to as *interference.*

In more detail, when two wave motions meet and overlap they interfere with each other, the resultant amplitude at a point being the sum of the amplitudes of two waves at that point. For two waves that are in phase the resulting intensity will be large and this effect is referred to as *constructive interference*. On the other hand if the two wave motions are out of phase by 180° and of equal amplitude then the resulting intensity will be zero, this is referred to as *destructive interference.*

22.3 Physical Conditions for interference

In order to obtain a static interference pattern at a point the following important conditions are necessary:

(a) The light sources must have the same frequency.
(b) The light sources must have a constant phase difference between them.
(c) The wave trains must overlap in space

The first two conditions form the requirements for *coherence.*

The condition for coherence is that the waves should have the same form, the same frequency and a fixed phase difference between them.

Note: Two separate light sources cannot be used as sources for static interference pattern because light consist of radiation from each individual atoms such that the emissions are in a random manner of pulses which have phases that are different from one atom to another.

22.4 Mathematical Superposition of two Waves

Consider the superposition of the following electromagnetic waves of the same frequency but having different phase: The waves are:

$$E_1 = E_{01}\sin(\omega t - \alpha_1) \text{ and } E_2 = E_{02}\sin(\omega t - \alpha_2) \qquad (22.3)$$

It can be shown mathematically that these two waves can combine to give:

$$E_1 + E_2 = A\sin(\omega t - \beta) \qquad (22.4)$$

Where A is the resultant amplitude which is related to the intensity I, by:

Intensity, $$I \approx A^2 = 2a^2(1+\cos\delta) = 4a^2\cos\frac{\delta}{2} \qquad (22.5)$$

If $\delta = 0, 2\pi, 4\pi, \ldots\ I \sim 4a^2$, and constructive interference occurs. (22.6)

If $\delta = \pi, 3\pi, 5\pi, \ldots\ I \sim 0$, and destructive interference occurs. (22.7)

For intermediate values of δ, the intensity varies between these limits according to the factor $\cos^2 \delta/2$. These modifications of intensity obtained by combining waves are known as interference effects.

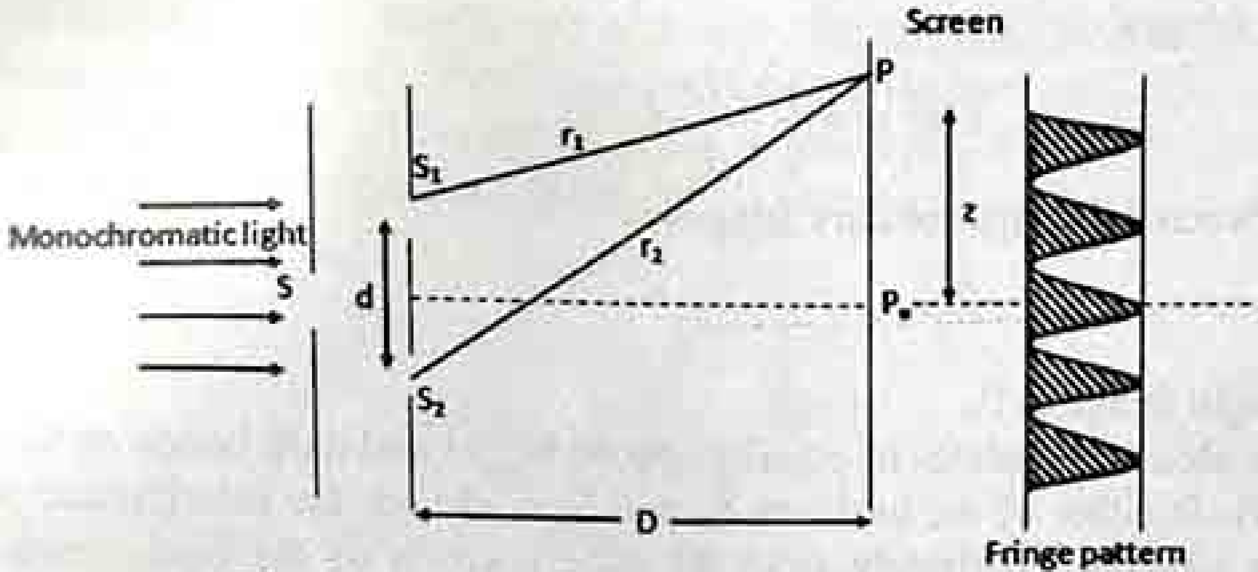


Fig. 22.1: Young's Double-slit Experiment

For two coherent sources (same wave amplitude, frequency and in phase) of same wave number k, the phase difference, δ, between waves from these sources at any observation point is k times their path difference. For example in Figure 22.1, $r_2 - r_1$ is the path difference between the waves arriving at P from S_1 and S_2. That is, path difference $\lambda\delta/2\pi$ since $k = 2\pi/\lambda$. In other words, if path difference is a whole number of wavelengths $(\delta = 0, 2\pi, 4\pi, \ldots)$ constructive interference occurs, while if path difference is an odd number of half wavelengths $(\delta = \pi, 3\pi, 5\pi, \ldots)$ destructive interference occurs.

22.5 Young's Double Slit Experiment

We will illustrate interference effects with a few examples, starting with Young's historic experiment. In the original experiment, Young used sunlight as the source of light.

However, any bright source such as tungsten filament or an arc could as well be used satisfactorily. The coherent sources are two narrow slits S_1 and S_2 which are close to each other and are illuminated by a monochromatic light source S (an illuminated narrow slit) equidistant from S_1 and S_2. The light waves arriving at S_1 and S_2 have equal phase and amplitude.

If P is at a distance z from the central point P_o on the screen and if d is the slit separation,

$$r_2^2 - r_1^2 = \left(z + \frac{d}{2}\right)^2 - \left(z - \frac{d}{2}\right)^2$$

$$r_2 - r_1 = \frac{2zd}{(r_2 - r_1)} = \frac{2zd}{2D} = \frac{zd}{D}$$

Therefore

Phase difference, $\delta = \dfrac{2\pi}{\lambda}(r_2 - r_1) = \dfrac{2\pi}{\lambda}\left(\dfrac{2d}{D}\right)$

Using equation 22.5, we obtain intensity

$$I \approx A^2 = 4a^2 \cos^2 \frac{\delta}{2} = 4a^2 \cos^2 \frac{\pi z d}{D}$$

Intensity is maximum when $\dfrac{\pi z d}{\lambda D} = 0, \pi, 2\pi, \ldots$

i.e. $\dfrac{zd}{D} = 0, \lambda, 2\lambda, 3\lambda, \ldots$

Therefore for bright fringes

$$z = \frac{m\lambda D}{d} \quad (22.8)$$

Where $m = 0, 1, 2 \ldots$ is the order of interference.

Intensity is minimum when $\frac{\pi zd}{\lambda D} = \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots$

That is $\frac{zd}{D} = \frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \ldots$

Therefore for dark fringes,

$$z = \left(m + \frac{1}{2}\right)\frac{\lambda D}{d} d \tag{22.9}$$

The separation between two bright or dark fringes is

$$\frac{\lambda D}{d} \tag{22.10}$$

with the central bright fringe at P_o.

Thus we expect as observed a series of equally spaced bright and dark bands or fringes on the screen on either side of P_o. If either of the two slits S_1 and S_2 is closed, the interference pattern disappears, indicating that the pattern is definitely from the superposition of the light waves from S_1 and S_2. Apart from its historically providing conclusive evidence for the wave nature of light, Young's double slit experiment has been used to determine the refractive indices of gases and liquids, and to measure the wavelength of light as well as the thickness of thin films of transparent materials.

Example 22.3

A monochromatic beam of light illuminates Young's experiment producing a fringe pattern with 5*mm* separation between consecutive bright bands. If the distance between the plane containing the slits and the plane of observation is 20*m* and if the two slits are separated by 2*mm*, find the wavelength of the light?

Solution

From equation 22.10, the separation between two bright or dark fringes is $\lambda D/d$.

$$\therefore \quad 5\times10^{-3} = \frac{20}{2\times10^{-3}}\lambda$$

$$\therefore \quad \lambda = \frac{5\times10^{-3}\times2\times10^{-3}}{20} m = 5\times10^{-7} m$$

Example 22.4

In Young's experiment, two slits spaced 0.2*mm* apart are at a distance of 1.6*m* from a screen. If the 5th bright fringes is located 25*mm* from the central fringe, find the wavelength of the light used.

Solution

Let λ be the wavelength of the light. If the slit spacing is *d* and the fring distance from central position is *L* and $m = 5$, we have.

$$\lambda = dL/mD$$

$$\therefore \quad \lambda = 0.2\times10^{-3}\times2.5\times10^{-3}/5\times1.6 = 6.25\times10^{-7} m$$

22.6 Interference in Thin parallel Films and Thin wedge-shaped Film

Thin Parallel Films

Without going into the mathematics it can be shown thatif a thin parallel film of refractive index *n* and thickness *d* is illuminated by white light, the thin film appears coloured.

Thin Wedge-Shaped Film

On the other hand for a thin wedge-shaped film of angle α and refractive index *n*, at any thickness of the film $d \approx a$, there is a relative phase shift of π between the ray reflected internally at the top of the wedge and that reflected externally at the bottom of the wedge.

Thus if a thin wedge is illuminated by a parallel beam of monochromatic light the interference pattern is a series of alternating straight bright and dark bands running parallel to the line of intersection of the two planes forming the wedge. These fringes are therefore fringes of equal thickness.

22.7 Newton's Rings

The same conditions for the thin wedge-shaped film also apply for the interference effects obtained by shining light on a system composed of a spherical surface resting on an optical flat or two spherical surfaces of different curvatures in contact. Bright and dark bands concentric circles about the point of contact are observed. Interference effects occur in the thin layer of varying thickness enclosed by the two surfaces.

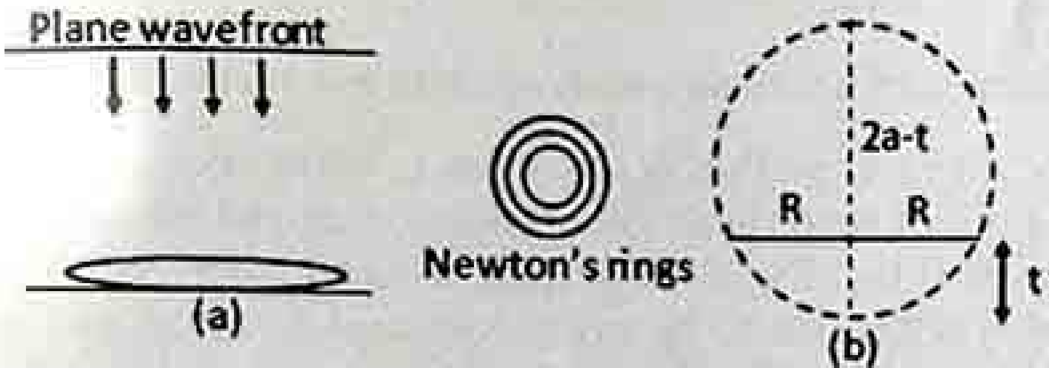


Fig.22.2: (a) Newton's ring experiment, (b) Geometrical construction for Newton's ring experiment

Consider the near-normal incidence of light of wavelength λ on a convex spherical lens resting on a horizontal flat glass plate (Figure. 22.2). If the gap or film enclosed has a refractive index n, then at the gap thickness t, as a result of interference between the light waves reflected at the top and bottom of the film, bright rings are formed when

$$2nt = \left(m + \frac{1}{2}\right)\lambda,$$

And dark bands when $2nt = m\lambda$

where m is the order of the fringe. If the radius of the ring is R and the radius of curvature of the lower surface of the spherical lens is a then by geometry (Figure 22.2b)

$$R^2 = t(2a - t)$$

$$\frac{R^2}{a} = 2t \qquad \text{(since } t << a\text{)}$$

Thus, for bright rings,

$$\frac{nR^2}{a} \approx \left(m + \frac{1}{2}\right)\lambda \qquad (22.11)$$

These rings, called Newton's rings are therefore fringes of equal thickness. The diameter $2R$ of a particular ring is measured with a travelling microscope, and the wavelength of the light or the radius of curvature of the lens calculated from equation 22.11.

These interference methods have several further applications, for instance, by means of Newton's ring or the Thin Wedge-Shaped Film. Interferometer as a very sensitive displacement measurer, thermal expansion, magnetostriction, elastic constant of materials, for example, has been precisely measured. The interference patterns are also used to confirm perfect polishing of lenses or the production of optically flat glass or metal plates. For such plates, the requirement is that the variation in the thickness be less than a quarter of the wavelength of light to be used. The fringe contours produced are irregular if the surfaces are not perfect. Polishing must therefore be contained until perfection is achieved.

22.8 Diffraction of Waves

Diffraction is the term used to describe the ability of waves to bend around obstacles placed in their path. The diffraction of light is due to the wave properties of light when it propagates in a medium with sharply expressed in homogeneity, like holes in opaque screens, edges of opaque bodies, etc.

Sound waves are diffracted through wide opening while light waves are diffracted through very narrow openings, in line with Huygens's principle. When a beam of light approaches a small hole in an opaque screen, all the secondary wavelets are stopped except those arriving directly at the hole. These wavelets will be transmitted into the geometrical shadow where they will spread on their own.

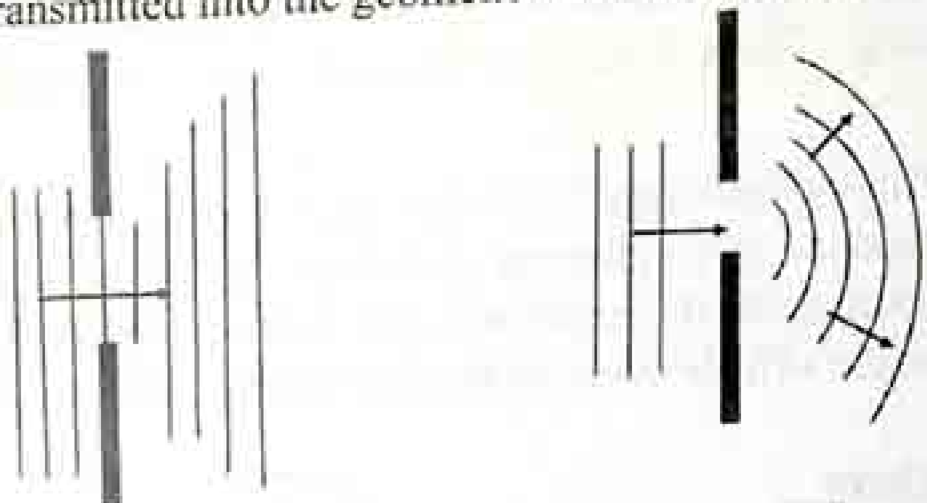


Fig. 22.3: Diffraction of waves by a wide opening and a narrow aperture

Put in a slightly different way, diffraction effect occurs when incident ray interacts with obstacles or apertures with finite size. The waves bend round the obstacles or spread out at the other side of the apertures. Diffraction is observed if the size of the obstacle is small compared to the wavelength of the wave. This condition makes diffraction to be easily compared in sound and water waves. The smallness of light wavelength makes its diffraction hard to observe in effect which made Newton to reject the Huygens wave theory. Nevertheless light wave can be diffracted.
The Italian physicist Francesco Grimaldi first recorded the broadening of a beam of light as it passes through a narrow slit in 1665. The small water droplets that condense on a car window show beautiful holes round the street lights as the car passes by due to diffraction. In diffraction the incident wavelength is altered by the obstacle to produce a new secondary wavefront according to Huygens theory.

The most conspicuous feature of diffraction is the deviation from rectilinear propagation when a wave is obstructed by an opaque obstacle. The incident waves appear to bend round the obstacle and form fringe patterns in the region of geometrical shadow. Similarly, waves going through a transparent hole in an opaque object are spread out in the region of the geometrical shadow. The obstacle in a transparent medium and the aperture in an opaque object are called diffraction centres. The phenomenon is like that of interference, each resulting from the superposition of several wavelets. The observed diffraction events are due to forced vibration of parts of the diffraction centers at the wave frequency, the production of new waves by these shaken parts, and the superposition of the new waves and incident wave to give rise to the diffracted wave. Diffraction also confirmed the wave nature of light and of electrons and other material particles.
Huygens'-Fresnel principle is used to analyze diffraction problems and it states: Every point on a wavefront serves as a source of spherical secondary wavelets of the same frequency as the primary wave. The optical field at any point beyond an obstruction is the superposition of all such wavelets reaching that point.

The diffraction effects are large when the dimensions of the diffraction centers are comparable with the wavelength of the waves involved. Consequently, diffraction sets a limit to the precision with which light or other waves can be used to record the fine detail in a physical object being examined. Thus the dimensions of an object can be very precisely measured in terms of the wavelength of the diffracted waves such as light, X-rays, or de Broglie waves of electrons. On the other hand, diffraction effects are not observable when the wavelength of the waves is much smaller than the dimension of the diffraction centers. Diffraction limits the sharpness of the image produced by optical instruments. However, the shorter the wavelength of the waves and the larger the diameter of a lens or mirror in an optical instrument, the less significant the diffraction effect.

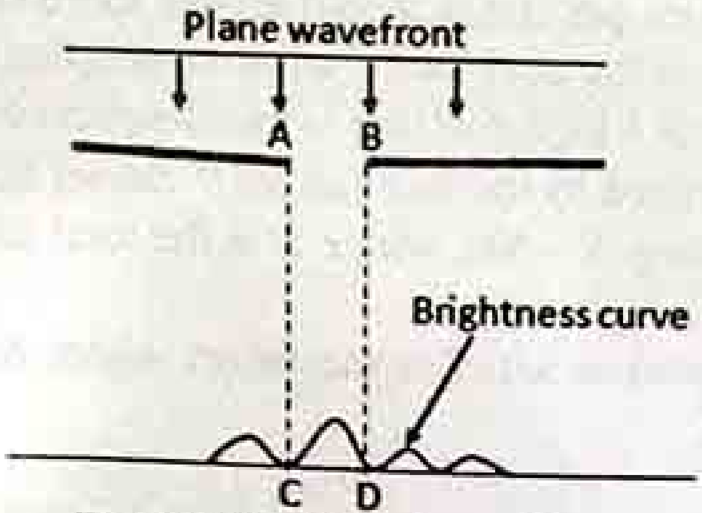


Fig. 22.4: Diffraction of light

Consider two points on the same wavefront, for example, the two points A and B on a plane wavefront of light incident on a narrow slit in a screen (Figure. 22.4). According to Huygen's principle A and B are secondary sources of light and since they are on the same wavefront, they have identical amplitudes and frequencies and are in phase. A and B are therefore coherent sources which produces an interference pattern on a screen in front of the slit if the dimension of the slit is small compared with the wavelength of the incident light. Alternate bright and dark bands are actually observed in the geometrical shadow, i.e., a short distance beyond the edges C and D of the projection of A and B. This deviation from rectilinear propagation is the phenomenon of diffraction.

Two kinds of diffraction phenomena can be observed with light waves one called Fresnel diffraction which occurs with spherical wavefronts and the other called Fraunhofer diffraction, involving plane wavefronts; hence the latter is easier to interprete mathematically.

22.9 Diffraction Grating

Diffraction grating is an instrument where diffraction phemomena is applied

A diffraction grating is an optical device consisting of a large number of identical concentrated in a small space. A typical example is a certain type of Louvre glasses which is made up of a regular pattern of parallel fine lines. Specially designed grating glasses are used in the laboratories for light experiments. There may be as many as 10^4 lines per *cm*. This size of the line or slit is called grating spacing, and can be as small as 10^{-5}.

Light falling on the grating interferes to produce an interference pattern given by the equation:

$$d\sin\theta = m\lambda \qquad (22.12)$$

where $m = 0, 1, 2 \ldots$ etc., d is the grating spacing, θ is the diffraction angle and λ is the wavelength of the light used.

Note the grating spacing d, and the number of line per *cm*, N are related by:

$$d = 1/N \qquad (22.13)$$

Put in another way, obstacles with a large number of slits are called *diffraction gratings*. Such gratings are made of a large number of narrow lines on glass plate, which varies from grating to grating. They can contain up to as much as 3000 lines per millimeter. The grating constant N is defined as the number of lines per unit length. The grating spacing d is the inverse of N.

That is $d = 1/N$

Therefore the fringe width y is inversely proportional to N.

For a parallel beam of monochromatic light incident normally on a diffraction grating with a grating spacing e, the path difference δ:

$$\delta = m\lambda = d\sin\theta$$

$\therefore$ For a maximum $\lambda = d\sin\theta, = 0,1,2,3,4$

m is the order of the spectrum; $m = 0$, corresponds to the central maximum. The first order spectrum is formed for $m = 1$, second order $m = 2$ e.t.c.

Factors that influence the choice and application of gratings:

The number of orders that can be viewed with a given grating depends on

(i) The wavelength range to be studied,

(ii) The slit spacing; fewer lines per meter would enable more order to be seen, which could be confusing (useful though, in the analysis of infrared radiation). With more lines per meter the first order is expanded but at a high cost due to the closeness of the rulings.

(iii) The geometry, which is limited by the fact that $\sin\theta$ cannot be greater than unity, and

(iv) Resolving power of a grating $R-Nm$, where N is the total number of lines in the grating and m is the order.

The condition of the principal minima satisfies the points where destructive interference of waves occur i.e. $d\sin\theta = m\lambda$ (minima).

Example 22.5

A diffraction grating with 4.0×10^4 lines per meter is set up on a spectrometer table at what angle will the second other maxima be formed if the incident wave is of $600nm$ wavelength?

Solution

$m=2,\ N=4.0\times10^4 m^{-1}$

Grating spacing $d=1/N=2.5\times10^{-5}$

$$m\lambda = d\sin\theta,\quad \sin\theta = 2\times600\times10^{-9}/2\times10^{-5}m$$

$$\therefore\quad \theta = 3.44°$$

Example 22.6

In the diffraction pattern of a single slit, the separation between the first minimum on one side and the first minimum on the other side is $5.2mm$. The distance of the screen from the slit is $80.0cm$ and the wavelength of light used is $546nm$. What is the width of the slit?

Solution

Distance between first minimum and the central maximum

$$x=\frac{1}{2}(5.2)mm = 2.6\times10^{-3}m$$

Using the equation, $\theta=\lambda/a$ since θ is small and $\theta = x/D$

$$\therefore\ \frac{\lambda}{a}=\frac{x}{D},$$

$$\therefore\ \frac{\lambda}{a}=\frac{x}{D},\ \therefore \text{Slit width, } a=\frac{\lambda D}{x}=\frac{(546\times10^{-9})(0.80)}{(2.6\times10^{-3})}=1.68\times10^{-4}m$$

Example 22.7

What is the angular separation between the first minimum and the central maximum for a diffraction pattern of a single slit if the slit width is equal to,

(a) one wavelength, (b) 5 wavelengths and (c) 10 wavelengths?

Solution

Since the slit width a is compatible to λ, the equation $\sin\theta=\lambda/a$ is used rather than $\theta=\lambda/a$. The equation $\theta=\lambda/a$ is used when the slit width a is large compared to λ.

(a) when $a=\lambda$, $\sin\theta=\frac{\lambda}{a}=\frac{\lambda}{\lambda}=1$

$\therefore$ Angular separation, $\theta=90°$

(b) When $a=5\lambda$, $\sin\theta=\frac{\lambda}{a}=\frac{\lambda}{5\lambda}=0.2$

$\therefore$ Angular separation, $\theta=11.54°$

(c) When $a=10\lambda, \sin\theta=\frac{\lambda}{a}=\frac{\lambda}{10\lambda}=0.1$

$\therefore$ Angular separation, $\theta=5.74°$

Example 22.8

Find the angle of the fourth bright fringe if two slits $0.3mm$ apart are illuminate by red light of wavelength $700nm$.

Solution

Using the equation, $d\sin\theta = m\lambda$

$\sin\theta = m\lambda/d = 4\left(700\times10^{-9}m/0.3\times10^{-3}m\right) = 9.3\times10^{-3}$

$\theta = 0.534^{\circ}$

22.10 Polarization of Waves

If the vibration of particles of the medium in a transverse wave remains parallel to a fixed line in space, the wave is said to be linearly polarized. In the case of plane polarized light, the electric field vector E, always remains in a given plane of polarization. The name is given to the plane of the electric field, rather than the magnetic field because it is electric field which is responsible for photographic effect. In contrast to plane – polarized light, ordinary light produced by incandescent sources, for example a tungsten filament lamp is unpolarized, that is electric field although still transverse to the direction of travel; are otherwise oriented at random and have no common direction. We can visualize polarization most easily by considering mechanical waves on a string. If one end is moved up and down, the resulting waves on the string are linearly polarized with each element of the string vibrating in the vertical direction. Similarly if one end of the string is moved along a horizontal line (perpendicular to the string) the displacements of the string are linearly polarized in the horizontal direction. If one end is moved with a constant speed in a circle the resulting wave is circularly polarized. Un-polarized waves can be produced by moving the end of the string vertically and horizontally in a random way.

Most waves produced by a single source are polarized. For example, string waves produced by regular vibration of one end of a string or electromagnetic waves produced by a single atom or by many sources are usually un-polarized.

Interference and diffraction can occur with all kinds of waves. Polarization occurs only with transverse waves; the phenomenon has therefore been used as such to establish the transverse nature of light and electromagnetic waves in general. Electromagnetic waves such as light waves consists of electric (E) and magnetic (B) fields which are perpendicular to the direction of the wave and to each other. The direction and plane of polarization of an electromagnetic wave are respectively the direction and plane of its electric field. For plane polarized light wave, for example, its electric field oscillates within a single plane and traces a sinusoidal curve along the direction of propagation (Figure. 22.5).

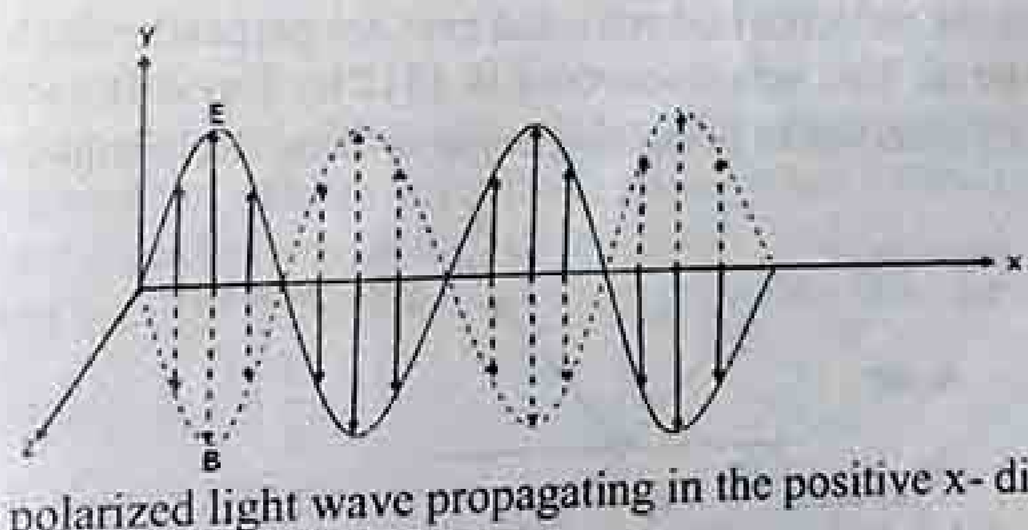


Fig. 22.5: Plane polarized light wave propagating in the positive x- direction

There are four phenomena that can produce polarized light from unpolarized light: Absorption, scattering, reflection and birefringence (also called double refraction).

22.11 Polarization by absorption

Several naturally occurring crystals absorb and transmit light differently depending on the polarization of light. These crystals can be used to produce linearly, polarized light. Example is a dichroic material such as Polaroid which was invented by E. H. Land in 1938. Dichroic molecules have two coefficients of absorption one for one direction of the electric field and one for another direction.

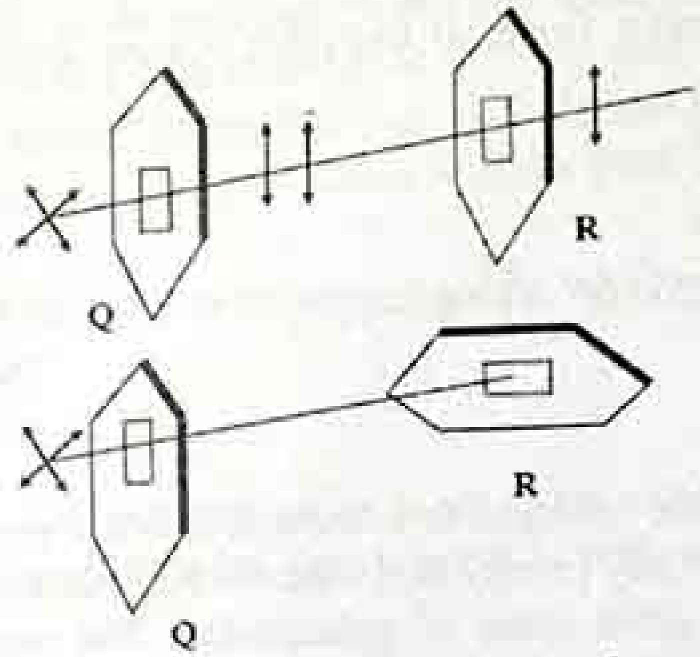


Fig: 22.6: Polarization of light through parallel and perpendicular crystals

The material contains long-chain hydrocarbon molecules which absorb light strongly when the plane of polarization is parallel to the chain axis, but little when the electric field is perpendicular to the axis. If unpolarized light is incident on the material, light with predominately one plane of polarization is transmitted since the component with its electric field along the molecular axis is absorbed.

22.12 Polarization by Scattering

The phenomenon of absorption and radiation is called scattering. An example of light scattering is that from clusters of air molecules (due to random fluctuations in the density of air), which tend to scatter short wavelengths more than long wavelengths, thereby giving the sky its blue color. We can understand polarization by scattering if we think of an absorbing molecule as an electric- dipole antenna that radiates waves with maximum intensity in the direction perpendicular to the antenna and zero intensity in the direction perpendicular to the antenna and zero intensity in the direction along the antenna a scattering center at the origin. The electric field in the light beam has components in both the x and y directions perpendicular to the direction of motion of the light beam. The light scattered in the x direction is polarized in the y direction and that scattered in y direction is polarized in the x direction.

22.13 Polarization by Reflection

When unpolarized light is reflected from a plane surface boundary between two transparent media such as air and glass, the reflected light is partially polarized. The degree of polarization is found to depend on the angle of incidence and the refractive index of the media. Experiment shows that when the angle of incidence is such that the reflected and refracted rays are perpendicular to each other, the reflected light is completely polarized. This was discovered in 1812 by David Brewster. The angle of incidence θ_p for which this occurs is known as polarizing angle or Brewster's angle.

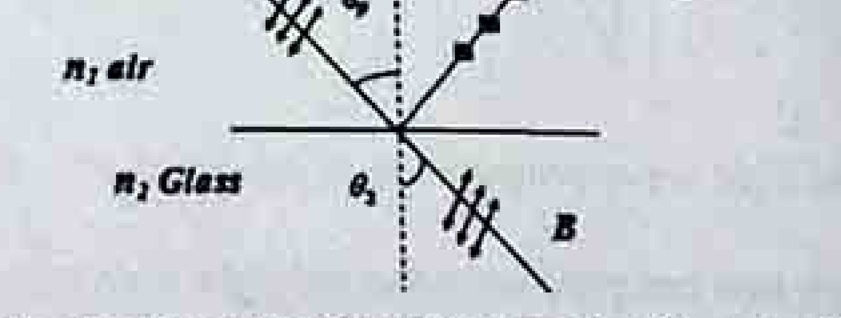


Fig. 22.7: Polarization by reflection showing the Brewster's angle

Snell's law gives: $n_1 \sin\theta_p = n_2 \sin\theta_2$.

But $\theta_2 + \theta_p = 90$ or $\theta_2 = 90 - \theta_p$,

$$n_1 \sin\theta_p = n_2 \sin(90-\theta_p) = n_2 \cos\theta_p$$

$$\tan\theta_p = n_2/n_1 \qquad (22.14)$$

The equation 22.14 is called the *Brewster's law equation* where n_1 and n_2 are the refractive indices of the incident and refracted media respectively.

Because of the polarization of reflected light, sunglasses made of polarizing material can be very effective in cutting out glare.

Example 22.8

The reflected light from an oily concrete floor of refractive index 1.45 appears to be fully polarized. Find the angle of the incident light.

Solution

We have $n_1 = 1$ for air and $n_2 = 1.45$. So, the reflected beam is polarized horizontal when

$\tan\theta_p = n_1/n_2 = 1.45/1$,

$\therefore\ \theta_p = 55°24'$

Example 22.9

A beam of unpolarized light of intensity I_0 is incident upon a stack of two Polaroid's with their transmission axes rotated 22.5° and 67.5° clockwise, respectively, compute the transmitted intensity, I_0.

Solution

The intensity I_1, after passing through the first Polaroid is

$I_1 = I_0 \cos 22.5° = 0.92 I_0$.

The final linear polarized light emerging from the second Polaroid has intensity

$I_2 = I_1 \cos 67.5° = 0.38 I_1$.

22.14 Polarization by Birefringence (Double refraction)

Birefringence also called double refraction was first observed by Bartholines in 1669. Birefringence occurs in calcite, liquid crystals, and some non-cubic crystals because of their atomic structure, birefringent are anisotropic therefore the speed of light depends on its direction through the material (unlike isotropic materials where the speed is the same in all directions).

Light rays incident on birefringent materials are usually separated in two rays. One is called the ordinary ray because it obeys the normal laws of refraction, and the other the extraordinary ray *E*.

These two rays are totally plane-polarized at right angles to each other and travel with different speeds.

Summary

1. The wave nature of light is demonstrated by interference and diffraction.
2. Huygen's principle is used to predict the propagation of waves. The principle states that every point on a wavefront becomes a new or secondary centre of disturbance.
3. The wavefront concept of light is used to explain reflection and refraction. The absolute refractive index of a medium is $n = \frac{c}{v}$.
4. The speed of wave is affected by the medium in which the wave is traveling through or along.
5. A change in wave speed causes a change in wavelength as the frequency remains constant.
6. Wave like light obeys the laws of reflection, refraction, interference, diffraction.
7. Interference refers to the superposition of two or more isolated portions of a wavefront of the same source or waves from independent coherent sources to produce a combine field effect.
8. Interference can either be constructive or destructive. It is constructive if two superposed waves are in phase and destructive if two superposed waves are in phase and destructive if two are II radians, out of phase and with some amplitude.

9. Necessary conditions for interference are:
   - The light sources must have the same frequency.
   - The light sources must have a constant phase difference between them.
   - The wave trains must overlap in space.
10. Young's double – slit experiment was the first experiment that confirmed the wave theory of light first proposed by Huygens.
11. At constructive interference, bright fringes are observed at path fringes difference, $m = d\sin\theta = yd/Dm = 0,1,2,3,\ldots$
12. At destructive interference dark fringes are observed. Path difference $= \left(m+\frac{1}{2}\right)d, (m = 0,1,2,3,\ldots)$
13. Diffraction effects occur when incident waves interact with obstacles or apertures with finite size. The dimension of the obstacles or apertures must be compared to the wavelength of the wave.
14. A linearly polarized wave is a transverse wave whose vibration remains parallel to a fixed line in space.
15. In plane polarized light, the electric field vector always remains in a given plane.
16. Polarized light can be produced from unpolarized light by absorption, scattering, reflection and birefringence.
17. Brewster's Law is given by $\tan\theta_p = n_2/n_1$.

Exercise 22

22.1 Which of the following statement is not correct, necessary condition for interference

A. The light sources must have the same frequency

B. The light source must have constant phase difference.

C. The sources must have large amplitude

D. The wave train must overlap in space.

22.2 Which statement is correct for interference

A. Two or more waves must interfere

B. If the waves are in phase it is constructive interference

C. It is destructive interference if they are out of phase

D. Only mechanical waves undergo interference

22.3 Two sources are said to be coherent sources if they have

A. Same frequency B. Same amplitude

C. Same wavelength D. are in phase

22.4 In young double – slit experiment which of the following is wrong:

A. Any bright source such as tungsten filament or an arc could be used

B. An illuminated narrow slits are used S_1 and S_2.

C. S_1 and S_2 have equal amplitude

D. S_1 and S_2 have different phase

22.5 In young double-slit experiment which of the following is wrong

A. D is the distance between the sources and the screen.

B. Separation between two bright fringes $= D/d$.

C. d = distance between the slits. D. $m = 0, 1, 2....$ is the order of inference.

22.6 In young slit experiment separation between two slit is $2mm$, fring separation is $3mm$ and distance between slit and plane of observer is $20m$. What is the wave length of light?

A. 3×10^{-7} B. 10^{-6} C. 3×10^{-3} D. 10^{-5}

22.7 Which of the following statement is correct?

If a beam of unpolarized light is reflected from glass or water the reflected the reflected ray is fully polarized

A. If the angle between the reflected and refracted rays is 90°

B. That angle of incidence is also 90°

C. That angle of incidence is related to refractive index by $n = \tan i$

D. The refracted and reflected rays are partially polarized

22.8 Which of the following is not correct? Polarized light can be produced from unpolarized light by A. Absorption B. Scattering C. Reflection D. Diffusion

22.9 At destructive interference

A. Bright fringes are observed B. The wave amplitude is zero

C. Intensity at the point is maximum D. The frequency of the wave is maximum

22.10 Which of the following statement is not correct

A. Ordinary source like the sun is polarized B. Flourescent lamp is unpolarized

C. Unpolarized means that the planes of vibration are randomly oriented

D. Electromagnetic waves can be elliptically polarized

22.11 (a) State Brewster's law/equation? (b) State Snell's law

22.12 From Snell's law obtain Brewster's law and discuss its physical application

22.13 A parallel beam of white light of wavelength $4.5\times10^{-7}\,m$ to $7.5\times10^{-7}\,m$ is incident normally on a diffraction grating. The wavelength which is deviated most in the second order spectrum is diffracted through an angle 60^o from the direction of incidence. How many lines per metre are there on the gravity?

22.14 Light from a sodium vapour lamp is incident normally on a diffraction grating having 6000 lines per cm. The spectrum obtained consists of two strong lines of $\lambda = 577\,nm$ and $\lambda = 579\,nm$. What is the angular separation between these two lines for the second order diffracted beam?

22.15 In a young's double slit arrangement the distance between the centers of the slits is $0.25mm$ and $\lambda = 6.0\times10^{-4}\,nm$. Calculate the angle θ subtended at the double slit by the neighboring maxima of the interference pattern.

22.16 In a young double slit experiment, the slit separation is $0.05cm$, and the distance between the double slit and the screen is $200cm$. When blue light is used the distance of the first bright fringe from the centre of interference pattern is $0.13cm$. Calculate the wavelength of blue light used and the distance of the fourth dark fringe from the center of the pattern.

22.17 A parallel beam of white light of wavelength $4.5\times10^{-7}\,m$ to $7.5\times10^{-7}\,m$ is incident normally on a diffraction grating. The wavelength which is deviated most in the second order spectrum is diffracted through an angle 60^o from the direction of incidence. How many lines per metre are there on the gravity?

22.18 Light from a source is incident normally on a diffraction grating which has 4000 lines per cm. If the light consists of two lines of wavelength $656nm$ and $410nm$ respectively, determine the angular separation between the two lines in the second order spectrum produced by the grating.

22.19 In a Young's experiment a monochromatic light passed through two narrow parallel slits that are separated by $0.40mm$, the bright fringes formed on a screen $1.6m$ away are observed to have a spacing of $2.99mm$. What is the wavelength of the incident light?

22.20 If a monochromatic light of wavelength $5860\,\overset{o}{A}$ falls normally onto a slit $0.02mm$ wide, find the angles at which the first two diffraction minima will form on the screen.?

Appendix I
Relevant Units in Physics

Quantity	Unit	Symbol
Length	metre	m
Mass	kilogram	kg
Atomic mass	atomic mass unit	u
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	Kelvin	K
Luminous intensity	candela	cd
Amount of substance	mole	mol
Frequency	hertz	Hz
Force	newton	N
Pressure and stress	pascal	Pa
Work, energy, heat	joule	J
Power	watt	W
Electric charge	coulomb	C
Electric potential difference	volt	V
Electromotive force	volt	V
Electric resistance	ohm	Ω
Electric conductance	Siemens	S
Electric capacitance	farad	F
Magnetic flux	weber	W
Magnetic flux density	(magnetic induction)	tesla T
Inductance	henry	H
Momentum	N s	P
Moment of a force	N m	M
Torque	N m	T
Electrical resistivity	Ω m	ρ
Electrical conductivity	S m^{-1}	σ
Current density	A m^{-2}	j
Permittivity	F m^{-1}	ε
Electric field strength	NC^{-1} or V m^{-1}	E
Permeability	H m	μ
Moment of inertia	kg.m	I
Angular momentum	J s	L

Appendix II
Table of Fundamental Physical Constants

Constant	Symbol	Value
Atomic mass unit	u	1.660556×10^{-27} kg
Avogadro constant	N_0	6.022169×10^{23} mole^{-1}
Bohr magneton	M_B	9.27410×10^{-24} joule m^{2}Wb^{-1}
Boltzmann's constant	k	1.380623×10^{-23} joule k^{-1}
Electron charge	e	1.602192×10^{-19} coul
Electron mass	m_e	9.10956×10^{-31} kg
Faraday constant	F	9.64867×10^{4} coul.mole^{-1}
Gas constant	R	8.3143 joule mole^{-1}k^{-1}
Gravitational constant	G	6.6732×10^{-11} N.m.^{2}kg^{-2}
Mean radius of sun	r_{sun}	6.96×10^{5} km
Neutron mass	m_n	1.67492×10^{-27} kg = 1.00865u
Nuclear magneton	M_n	5.05095×10^{-27} joule m^{2}Wb^{-1}
Permeability of vacuum	μ_0	1.256637×10^{-6} H.m^{-1}
Permittivity of vacuum	ε_0	8.854185×10^{-12} F.m^{-1}
Planck's constant	h	6.62620×10^{-34} joule sec.
	$\hbar(=h/2\pi)$	1.054×10^{-34} J.s.
Proton mass	m_p	1.67261×10^{-27} kg=1.00727u
Speed of light in vacuum	c	2.997925×10^{8} m.s.$^{-1}$
Stefan-Boltzmann constant	σ	5.6696×10^{-8} W.m.$^{-2}$ K^{-4}
Sun-Earth distance	d_s	1.49×10^{8} km

Appendix III

Basic Mathematics Tools

A. Trigonometric Functions

The six trigonometric functions of angle θ can be defined in terms of the sides of a right angled triangle as shown in Figure A.1.

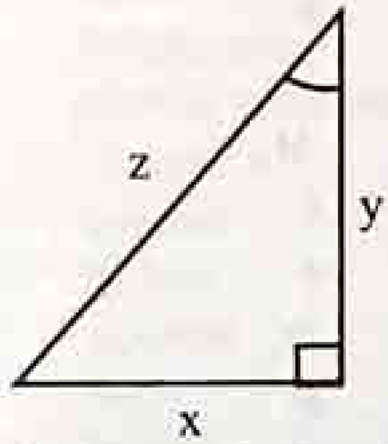


Fig. A.1 Trigonometric Functions

$$\text{cosine or } \cos\theta = \frac{\text{adjacent side}}{\text{hypotenuse side}} = \frac{y}{z}$$

$$\text{sine or } \sin\theta = \frac{\text{opposite side}}{\text{hypotenuse side}} = \frac{x}{z}$$

$$\text{tangent or } \tan\theta = \frac{\text{opposite side}}{\text{adjacent side}} = \frac{x}{y}$$

$$\text{cosecant or } \csc\theta = \frac{\text{hypotenuse side}}{\text{opposite side}} = \frac{z}{x}$$

$$\text{secant or } \sec\theta = \frac{\text{hypotenuse side}}{\text{adjacent side}} = \frac{z}{y}$$

$$\text{cotangent or } \cot\theta = \frac{\text{adjacent side}}{\text{opposite side}} = \frac{y}{x}$$

$$\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{1}{\cot\theta}$$

$$\csc\theta = \frac{1}{\sin\theta}$$

$$\sec\theta = \frac{1}{\cos\theta}$$

Other formulae are

$$\cos^2\theta + \sin^2\theta = 1$$

$$\cos(\theta \pm \Phi) = \cos\theta\cos\Phi \pm \sin\theta\sin\Phi$$

$$\sin(\theta \pm \Phi) = \sin\theta\cos\Phi \pm \cos\theta\sin\Phi$$

$$\sin 2\theta = 2\sin\theta\cos\theta$$

$$\cos 2\theta = \cos^2\theta - \sin^2\theta$$

$$= 1 - 2\sin^2\theta = 2\cos^2\theta - 1$$

$$\sin\theta = \sin(180^\circ - \theta)$$
$$\cos\theta = \cos(360^\circ - \theta)$$
$$\cos\theta = \sin(90^\circ - \theta)$$

B. Radians and Degrees

In a circle of radius R, an arc of length S subtends an angle of one radian at the centre of the circle (Figure A.2).

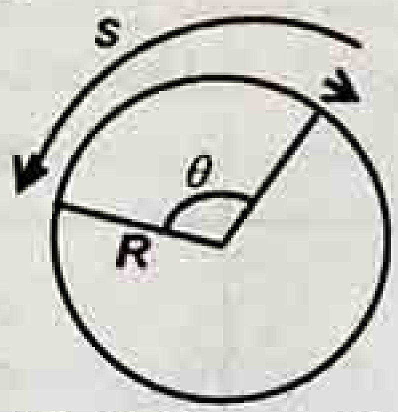


Fig. A.2 Radians and Degrees

$\therefore$ An arc of length S subtends angle $\theta = S/R\, radians$

$\therefore\ S = R\theta$

The circumference of a circle has length $2\pi R$.

$\therefore$ The circle subtends angle $\dfrac{2\pi R}{R} = 2\pi$ radians at the centre.

Also the circumference subtends 360° at the centre.

$\therefore\ 2\pi$ radians = 360°

1 radian $= \dfrac{360}{2\pi} = \dfrac{360}{2\times 3.142} = 57.3^\circ$

Similarly, $1^\circ = \dfrac{2\pi}{360} radians = 0.0175 radian.$

C. Small Angle Approximations

AMN is an arc of a circle of centre O (Figure A.3)

$$\theta = S/R,$$

where arc of length S subtends angle θ radians at the centre of circle. Also,

$\sin\theta = y/R$

$\tan\theta = y/x$

$\cos\theta = x/R$

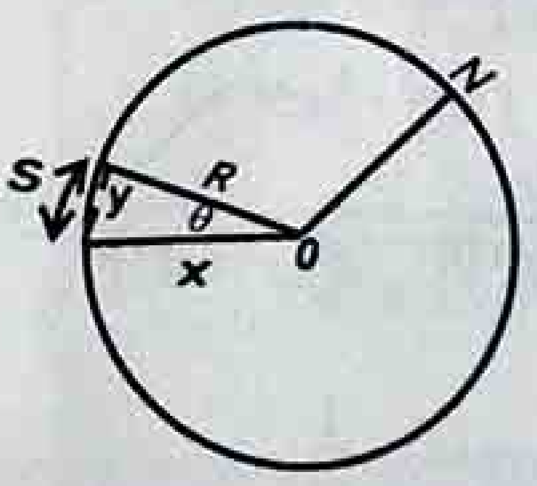


Fig. A.3

As θ gets smaller, $S \approx y$, and $x \approx R$

$\therefore\ \sin\theta \approx \dfrac{S}{R} = \theta \qquad \tan\theta \approx \dfrac{S}{R} = \theta$

That is, as θ gets smaller, the sine and tangent of the angle are approximately equal to the value of the angle in radian. Also, for small θ

$$\cos\theta \approx \frac{R}{R} = 1$$

These approximations are adequate for $\theta \leq 10^{\circ}$

D. Derivatives

The general rules for derivatives of some simple functions are provided. If y is the function, dy/dx is the derivative with respect to x as tabulated below:

y	dy/dx
a (constant)	0
x^n	nx^{n-1}
$\sin ax$	$a \cos ax$
$\cos ax$	$-a \sin ax$
$\tan ax$	$a \sec^2 ax$
e^{ax}	ae^{ax}
$\ln ax$	$1/x$

E. Indefinite Integrals

The integrals of some simple functions are given below:

$\int dx = x + c$, where c is a constant

$$\int X^n dx = \frac{1}{n+1} X^{n+1} + c$$

$$\int \sin ax \, dx = -\frac{1}{a}\cos ax + c$$

$$\int \cos ax \, dx = -\frac{1}{a}\sin ax + c$$

$$\int \frac{1}{x} dx = \ln x + c$$

$$\int e^{ax} = \frac{1}{a} e^{ax} + c$$

$$\int \frac{dx}{\left(x^2 + R^2\right)^{3/2}} = \frac{x}{R^2\left(x^2 + R^2\right)^{1/2}} + c$$

F. Definite Integrals

$$\int_o^\pi \sin \omega t = -\frac{1}{\omega}[\cos \omega t]_o^\pi = \frac{1}{\omega}(\cos\pi - \cos o) = \frac{2}{\omega}$$

$$\int_o^\pi \cos \omega t = \frac{1}{\omega}[\cos \omega t]_{-\pi}^\pi = \frac{1}{\omega}(\sin\pi - \sin -\pi) = 0$$

N.B. sin $\sin\pi = 0, \cos\pi = -1$

G. Vectors

A vector $\underline{a}$, is written as $\underline{a} = |\underline{a}| < \theta$

$|\underline{a}|$ is the magnitude of the vector, θ is its direction of $\underline{a}$ is $\underline{u} = \frac{\underline{a}}{|\underline{a}|}$

A vector $\underline{a}$ in two dimensional spaces can be written as:

$\underline{a} = |\underline{a}|\cos\theta, a_y = |\underline{a}|\sin\theta$ are the components along x and y, i and j are the unit vectors along x and y axis.

and $$|\underline{a}| = \sqrt{a_x^2 + a_y^2}, \tan\theta = a_y / a_x$$

In terms of position vectors, $\underline{a}$ may be written as $\underline{a} = (a_x, a_y)$. (Figure A.4).

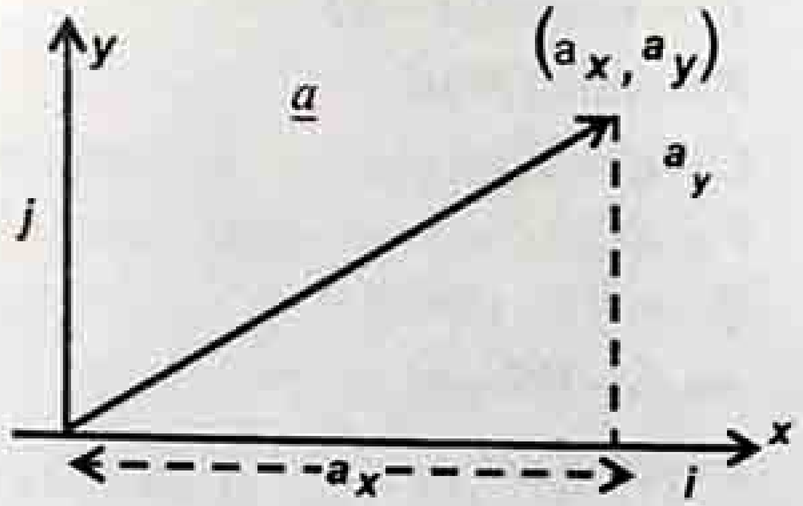


Fig. A.4

We can generalize to 3-dimension $\quad \underline{a} = a_x i + a_y j + a_z k$

If $r = x\,i + y\,j$ is the position vector of a point on a curve (Figure A.5), then the velocity of the point is given by

$$v = \frac{dr}{dt} = \frac{dx}{dt}i + \frac{dy}{dt}j.$$

Acceleration $$a = \frac{d^2x}{dt^2}i + \frac{d^2y}{dt^2}j.$$

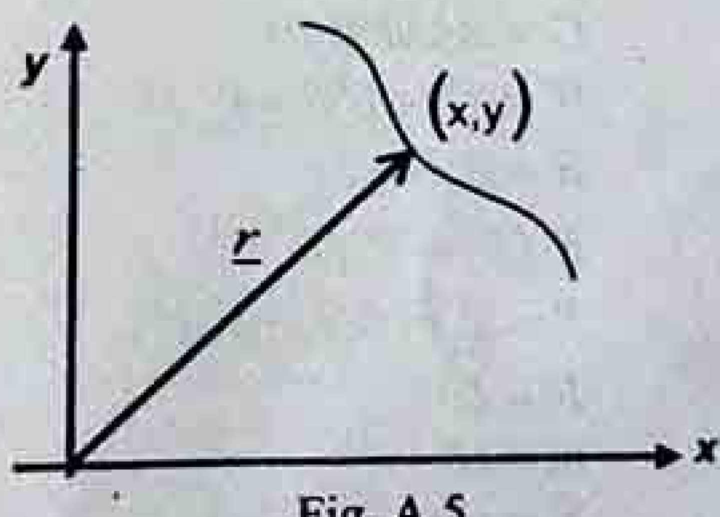


Fig. A.5

H. Summation

The addition such as the following can be written briefly as:

$$m_1 r_1^2 + m_2 r_2^2 + m_3 r_3^2 + \dots m_n r_n^2 = \sum_{i=1}^{n} m_i r_i^2$$

Appendix IV
Formulae List in Electromagnetism and Modern Physics

Charge	$Q = It$
Capacitance of a sphere	$C = 4\pi\varepsilon r$
Current	$I = nAve$
Parallel-plate capacitor	$C = \varepsilon A/t$
Electrical energy	Energy = QV
Concentric spheres	$C = 4\pi\varepsilon ab/(a-b)$
Force on charge	$F = QE = QV/d$
Parallel capacitors	$C = C_1 + C_2$
Ohm's law	$V = IR$
Series capacitors	$1/C = 1/C_1 + 1/C_2$
Internal resistance	$e.m.f. = I(R+r)$
Energy store	$E = \frac{1}{2}QV = \frac{1}{2}CV^2$
Resistivity	$\rho = RA/l$
Capacitance comparison	$C_1/C_2 = \theta_1/\theta_2$
Temperature variation of resistance	$R_\theta = R_0(1+\alpha\theta)$
Capacitor discharge	$V = V_o e^{-t/RC}$
Capacitor charge	$V = V_o(1-e^{-t/RC})$
Series resistance	$R = R_1 + R_2$
Force on current	$F = BIl\sin\theta$
Parallel resistance	$I/R = I/R_1 + 1/R_2$
Couple of coil	$C = BANI\sin\theta$
Power	$W = VI = I^2R = V^2/R$
Field at centre of coil	$B = \mu_0 N1/2r$
Wheatstone bridge	$R_1/R_2 = R_3/R_4$
Field in solenoid	$B = \mu_0 nl$ (n turns m^{-1})
Faraday constant	$F = Le$
Mass liberated in electrolysis	$m = zIt$
Field at end of long solenoid	$B = \mu_0 n1/2$
Electric field strength	$E = -dV/dt$
Helmholtz coils	$B = 8\mu_0 N1/(125)^{1/2} r$
Field near straight wire	$B = \mu_0 1/2\pi r$
Force between charges	$F = \frac{1}{4\pi\varepsilon}\cdot\frac{Q_1Q_2}{d_2}$
Velocity of *e.m.* waves	$c = 1/(\varepsilon_0\mu_0)^{1/2}$
Field due to point charge Q	$E = Q/(4\pi\varepsilon d^2)$
Current sensitivity	$\theta = BAN\,I/c$

Potential	$V = W/Q_0$
Voltage sensitivity	$\theta/V = BAN/cr$
Potential due to charge Q	$V = Q/4\pi\varepsilon d$
Self-inductance	$L = N\phi/I$
Capacitance	$C = Q/V$
Mutual inductance	$M = N_s\phi_s/I_p$
Capacitance of a sphere	$C = 4\pi\varepsilon r$
Induced *e.m.f.* ε	$\varepsilon = -L\,dI/dt$
Induced *e.m.f.* ε_s	$\varepsilon_s = -M\,dI_p/dt$
Induced *e.m.f.* in a rotating coil	$\varepsilon = BAN\omega\sin\theta$
Induced *e.m.f.* (Neumann's law)	$\varepsilon = -N\,d\phi/dt$
Ballistic galvanometer	$Q\alpha\theta$
Transformer	$n_p/n_s = V_p/V_s$, $I_p/I_s = n_s/n_p$
Root mean square current (I)	$I = i_o/\sqrt{2}$
Alternating current	$i = i_o\sin(\omega t)$
Capacitative reactance	$X_c = 1/\omega C$
Inductive reactance	$X_L = \omega L$
Impedance (series RCL)	$Z = \left(R^2 + (X_L - X_c)^2\right)^{1/2}$
Resonance condition for I	$X_L = X_C$
Electrostatic force on electron	$F = eE$
Electromagnetic force on electron	$F = Bev$
Crossed fields	$eE = Bev$
Energy gain	$E = eV$
Kinetic energy	$eV = \frac{1}{2}mv^2$
Circular orbit	$BeV = mv^2/r$
Mass spectrometer	$q/M = v/Br$
X-ray diffraction	$m\lambda = 2d\sin\theta$
Quantum energy	$E = hf$
Relativistic mass energy relation	$E = mc^2$
de Broglie equation	$\lambda = h/mv$
Moseley's law	$f = k(Z-b)^2$
Work function	$W = hf_0$
Einstein's p.e. equation	$hf = hf_o + \frac{1}{2}mv^2$
Photoelectric effect	$hf = eV$
Radioactive decay	$N = N_0e^{-\lambda t}$
Half-life	$T = 0.693/\lambda$
Serial relation	$\lambda_1N_1 = \lambda_2N_2$
Hall voltage	$V = BI/net$

ANSWERS TO EXERCISES

Answers to Exercise 1

1.1 C, **1.2** A, **1.3** C, **1.4** B, **1.5** B, **1.6** D, **1.7** A, **1.8** C, **1.9** D, **1.10** D, **1.15** 10^{12} electrons; **1.16** 1.2×10^{36}; **1.17** 53N, 16N, 37; **1.18** 0.4Q, ½, 73; **1.19.** 3.45×10^{17} N/C; **1.20** $-3qa^2/4\pi\varepsilon x^4$, $6q/4\pi\varepsilon y^4$

Answers to Exercise 2

2.1 C, **2.2** B, **2.3** B, **2.4** A, **2.5** A, **2.6** A, **2.8** 5.48cos 60° N-m²/C; **2.9** 3.7×10^6 N-m²/C; **2.10** (i)2.67 $\times 10^6$ N-m²/C (ii) 0.00 (iii) 3.57×10^6 N-m²/C; **2.11** 0.00, 3.72×10^7Cos15 N-m²; **2.12** 2.36×10^7 V; **2.13**4.99 $\times 10^3$ N-m²; **2.14** 5.6×10^2 C; **2.15** 1.87×10^3 N-m²/C , 8.3×10^9 C;**2.17**0.443 mm;**2.18** - 1.18×10^{-12} C/m³; **2.19**-2.8×10^3 C/m³, r < 60cm and $q/4\pi\varepsilon_0 r^2$ for r > 90cm

Answers to Exercise 3

3.1 C, **3.2** A, **3.3** B, **3.4** B, **3.5** B, **3.6** B, **3.7** B, **3.8** A, **3.9** A, **3.10** D, **3.11** (i) 28.3 pF (ii) 141.5 pF; **3.13** μC; **3.15** 0.25 μF, 0.125 J; **3.17** (i) 4 μF (ii) 666.7μF, 333.3V, (iii) 0.004C; **3.19** 4.15μF **3.20** 1.5μF

Answers to Exercise 4

4.1 A, **4.2** C, **4.3** B ,**4.4** D, **4.5** A, **4.6** A, **4.7** A, **4.8** A, **4.9** B, **4.10** D, **4.11** $1.19\times10^{16}S^{-1}$, 2.42×10^3A/m²; **4.12** (i) 6.79×10^6A/m² (i) 4.24×10^{-4}m/s (iii) 1.90×10^{-2}V/m; **4.13** (i) 0.654μA (i) 83.4mA; **4.14** 0.63 $I_0\tau$ (ii) $I_0\tau$; **4.15** (i) $\frac{V}{2\pi r}$ (ii) $\frac{Vq}{2\pi r}$; **4.16** 5.34Ω, **4.17**l = 1.818m, d = 0.28mm; **4.18** (i) 4.62×10^{-4}Ω-m (ii) 42.86Ω; **4.19** 6.67×10^{-2}S

Answers to Exercise 5

5.1 A, **5.2** B, **5.3** C, **5.4** D, **5.5** D, **5.6** A, **5.7** B, **5.8** B, **5.9** B, **5.10** C, **5.11** 0.55R; **5.12** (i) 75V (ii) 37.5W; **5.13** 3.41kΩ; **5.14** R_x = 1.8kΩ; **5.15** 5.75Ω; **5.16** 73.5Ω, 882V; **5.17** (i) 6.67Ω (ii) 6.67Ω; **5.18** 0, 2.0, 2.40, 2.86, 3.0, 3.60, 3.75, 3.94A

Answers to Exercise 6

6.1 D, **6.2** A, **6.3** A, **6.4** D, **6.5** B, **6.6** C, **6.7** C, **6.8** D, **6.9** A, **6.10** D, **6.11** 1.6N, **6.12** -2.88N j, **6.13** 0.042T k, **6.14** 14.0N **k**, **6.15** -346.67N **k**, **6.16** {- 12.3**i** – 4.48**j** + 1.6**k**} $\times 10^{-18}$ N, **6.17** 13 T North, **6.18** (i) East (ii) 6.28×10^{14} m/s (iii) 2.98mm; **6.19** 1.98 cm;**6.20** B ≥ $\{2mK/(e^2l^2)\}^{1/2}$

Answers to Exercise 7

7.1 C, **7.2** D, **7.3** A, **7.4** B, **7.5** C, **7.6** A, **7.7** D, **7.8** C, **7.9** C, **7.10** B, **7.11** 0.001T; **7.12** 0.1m; **7.14** 10A, attractive; **7.16** 7.2×10^{-4} N; **7.17** 4A;

7.18 (i) 6.3×10^{-6} T (ii) 2.3×10^{-6} T **7.20** $B_p = \mu_o I\left(\frac{1}{4}+\frac{1}{2\pi}\right)/R$ BQ = $\mu_o I/\pi R$

Answers to Exercise 8

8.1 B, **8.2** B, **8.3** B, **8.4** D, **8.5** C, **8.6** A, **8.7** C, **8.8** C, **8.9** True, **8.10** C, **8.16** step-down, $i_{p=}$ 0.48A, $i_{s=}$2.4A; **8.17** 1.9×10^{-6}Wb; **8.18** i) 8.0V ii) 0.53A; **8.19** 7150 rpm; **8.20** A 136V B 11.3A

Answers to Exercise 9

9.1 C, **9.2** A, **9.3** B, **9.4** A, **9.5** D, **9.6** A, **9.7** D, **9.8** C, **9.9** FALSE **9.10** B, **9.12** (i) $0.004J/m^3$ (ii) 0.08A; **9.13** (a) 4.00mJ (b) 0.750mJ 10.14: $\mu_o N_1 N_2 A_2 \frac{\cos\theta}{L}$; **9.17** 1.98pF to 16.7pF; **9.18** (i) 0.689μH (ii) 17.9pJ (iii) 0.11μC; **9.20** 2.456pF

Answer to Exercise 10

10.1 B, **10.2** D, **10.3** D, **10.4** D, **10.5** B, **10.6** C, **10.7** A, **10.8** D, **10.9** B, **10.10** D, **10.11** 50 cm; **10.13** 1.0T; **10.14** 1.000754, 1.266×10^{-6} H/m; **10.18** Paramagnetic: Iron, Ammoium alum, Oxygen (liquid); Diamagnetic: Bismuth, Nitrogen (gas).

Answers to Exercise 11

11.1 B, **11.2**A, **11.3** C, **11.4** B, **11.5** C, **11.6** D, **11.7** A, **11.8** D, **11.9** A, **11.10** B, **11.11** A, **11.15** ($25\sqrt{2}$ V); **11.16** ((i) 20.8A, 29.5A (ii) 41.7A 58.9A; **11.17** 2.1A; **11.18** (i) 7.1 mC (ii) 0.64A; **11.19** 60V 12.20. 7.03H

Answers to Exercise 12

12.1 A, **12.2** C, **12.3** B, **12.4** C, **12.5** B, **12.6** D, **12.7** B, Other questions have hints

Answers to Exercise 13

13.1 A, **13.2** A, **13.3** $F = \dfrac{Ze}{4\pi\varepsilon_0 r^2}$, **13.4** 4.45×10^{-14}, **13.5** C, **13.6** A, **13.7** D, **13.8** B, **13.9** C, **13.10** B, **13.19** 1.503×10^{-10}J, 938MeV; **13.20** 146MeV, 1084MeV

Answers to Exercise 14

14.1 C, **14.2** C, **14.3** C, **14.4** C, **14.5** A, **14.6** A, **14.7** A, **14.8** C, **14.9** A, **14.10** A, **14.16** 3.75×10^{15} V; **14.17** 2.3 eV; **14.18** 0.74 eV; **14.19** 1.8×10^{-7}; **14.20** 0.2439A

Answers to Exercise 15

15.1 D, **15.2** B, **15.3** C, **15.4** C, **15.5** A, **15.6** A, **15.7** B, **15.8** C, **15.9** D, **15.10** A, **15.11** 3.45×10^{6} Bq; **15.12** 200.6 MeV; **15.13** 28.3 MeV; **15.14** 0.27 MeV; **15.15** Cl: p = 17, n = 18; Iron: p = 30, n = 26: Tin: p = 50, n = 70; **15.16** 2.0×10^{3} yr; **15.17** (a) ${}^{0}_{+1}e$ (b) γ (c) ${}^{0}_{-1}e$ (d) ${}^{237}_{94}Pu$; **15.18** 44 yr; **15.19** 3.7×10^{5}; **15.20** 5.37 MeV

Answers to Exercise 16

16.1 B, **16.2** D, **16.3** B, **16.4** B, **16.5** C, **16.6** B, **16.7** B, **16.8** D, **16.9** C, **16.10** B, **16.15** a. 0cm, b. -2cm, c. 2cm, d. -2cm; **16.16** 10^2Hz, 0.007, 14.28 m s^{-1}; **16.17** 0.57m, 6×10^{-7}m; **16.18** b. 0.2s, 5.0Hz, 31.42 rad/s, 2.62; **16.19** a. 34.0cm, b. 170m s^{-1} , c.500Hz; **16.20** 3×10^{8}m s^{-1}, 90.1Hz

Answers to Exercise 17

17.1 A, **17.2** C, **17.3** B, **17.4** B, **17.5** D, **17.6** B, **17.7** C, **17.8** C, **17.9** D, **17.16. a.** $69ms^{-1}$, b. $27.5ms^{-1}$; **17.17** 0.34m, 1.02m; **17.18** $250ms^{-1}$; **17.19** $795.7\times10^{-6}Wm^{-2}$; **17.20** 1032m.

Answers to Exercise 18

18.1 C, **18.2** A, **18.3** A, **18.4** C, **18.5** B, **18.6** D, **18.7** C, **18.8** A, **18.9** D, **18.10** B, **18.16** $500ms^{-1}$, $0.688m$; **18.17** $0.065kgm^{-1}$; **18.18** $51dB$; **18.19** 10^{-8}; **18.20** $1.114\times10^{-3}\,W/m^{-2}$.

Answers to Exercise 19

19.1 C, **19.2** A, **19.3** B, **19.4** B, **19.5** A, **19.6** A, **19.7** C, **19.8** C, **19.9** A, **19.10** A, **19.17** $2.50m$, $6.00m$, $1.25m$, 1.25cm; **19.18** $160.2m$, 1; **19.19**60°; **19.20**30.00cm, 6.00cm, $-30.00cm$, $10.00cm$.

Answers to Exercise 20

20.1 B, **20.2** C, **20.3** C, **20.4** A, **20.5** D, **20.6.** D, **20.7** A, **20.8.** D, **20.9** C, **20.10** A, **20.13** $6.00cm$; **20.14** 74.5°; **20.15** 8°; **20.16** 1.41; **20.17** Virtual image, $-8.6cm$, $^4/_7$, virtual image, $-11.1cm$, $^4/_9$; **20.18** $-100cm$, diverging; **20.19** Diverging lens, $-10.7cm$

Answers to Exercise 21

21.1 B, **21.2.**C, **21.3** A, **21.4** B, **21.5** A, **21.6** C, **21.7** A, **21.8** B, **21.9** C, **21.10** A, **21.11** 37° 10'; **21.12** 54° 46'; **21.13** 30.5cm; **21.14** 0.0155; **21.15** 0.85cm; **21.16** 40cm, 66.7cm; **21.17** 15.34cm

Answers to Exercise 22

22.1C, **22.2** D, **22.3** D, **22.4.** D, **22.5** B, **22.6.** A, **22.7** D, **22.8** D, **22.9** D, **22.10** A, **22.13.** 5.8×10^{5}, (use $d\sin\theta = 2\lambda$, $7.5\times10^{-7}\,m$); **22.14.** 0.19°; **22.15** use: $x = \dfrac{\lambda D}{a}$, $\theta = 2.4\times10^{-3}\,rad$ **22.16** $1.3\times10^{-3}\,m$, $3.25\times10^{-7}\,m$; **22.17** 5.8×10^{5}, (use $d\sin\theta = 2\lambda, 7.5\times10^{-7}\,m$); **22.18** 12.5 degrees; **22.19** $7.46\times10^{-13}\,m$; **22.20** The angles for the minima are 1.68° and 3.36°, respectively.

ISBN: 978-978-953-307-7

Published by Physics Writers Series Creation