# Electronic Communications

Electronic Communications

Website: www.skyfox.co

Email:skyfox@skyfox.org.in

ISBN 978-93-90357-01-7

9 789390 357017

sky fox

Electronic Communications

Dr. A. Ajay Praveenkumar

Dr. R. Ravibaskar

# Electronic Communications

**AUTHOR INFORMATION:**

**Dr. A. Ajay Praveenkumar., M.Sc., B.Ed., M.Phil., Ph.D.,**
Senior Lecturer
Department of Physics
DMI – St. Eugene University
P.O.Box: 330081, Chibombo
Zambia.

**Dr. R. Ravibaskar., M.Sc., M.Phil., Ph.D.,**
Associate Professor
Department of Physics
DMI – St. Eugene University
P.O.Box: 330081, Chibombo
Zambia.

| | |
|---|---|
| **First Edition** | **: August 2020** |
| **Cover Art and Design** | **: Authors** |
| **ISBN** | **: 978-93-90357-01-7** |
| **DOI** | **: https://doi.org/10.22573/spg.020.BK/S/014** |
| **Copyright** | **: © 2020 by Authors** |

**Publisher Contact &Support**

**Skyfox Publishing Group**
#987, Medical College Road
Thanjavur-613004
Tamil Nadu, India.
Phone: +918300123232
Email: skyfoxpublishing@gmail.com / skyfox@skyfox.org.in
Website: www.skyfox.co

**Headquarters &Marketing Office**

**Skyfox Publishing Group**
333 Cedar Street, PO Box 208002,
New Haven, United States.
CT 06520-8002.
Tel: 203.737.5603 / Fax: 203.785.7446
Email: skyfoxpublishing@gmail.com / skyfox@skyfox.org.in
Website: www.skyfox.co

# Electronic Communications

**Dr. A. Ajay Praveenkumar**
**Dr. R. Ravibaskar**

From the desk of
**Dr. T. X. A. ANANTH, BBA, MSW, MBA, MPhil, PhD,**
President – University Council

Dear Learner,
Welcome to DMI – St. Eugene University!
I am sure you are expert in using the PC Tablets distributed by us. Now your world is open to Internet and using the tablet for your educational learning purposes. The very same book you are holding in your hand now is available in your V-Campus portal. All the teaching and learning materials are available in your portal.
As our Chancellor, Rev.Fr. Dr. J. E. Arulraj , mentioned, it is not just the success for DMI-St. Eugene University alone, it is success for the technology, it is success for the great nation of Zambia and it is success for the continent of Africa.
You can feel the improvement in the quality of the learning materials in the printed format. Improvement is done in quality of the content and the printing. Use it extensively and preserve it for your future references. This will help you to understand the subjects in a better way. The theories of the subjects have been explained thoroughly and the problems have been identified at learners' level. This book is made up of five units and every unit is mapped to the syllabus and discussed in detail.
I am happy at the efforts taken by the University in publishing this book not only in printed format, but also in PDF format in the Internet.

With warm regards

Dr. T. X. A. ANANTH
President – University Council

**CONTENTS**

# Chapter I
# OSCILLATION & WAVES

## Introduction

A motion which repeats itself over and over again after a regular interval of time is called a periodic motion. Oscillatory or vibratory motion is that motion in which a body moves to and fro or back and forth repeatedly about a fixed point in a definite interval of time. Simple harmonic motion is a specific type of oscillatory motion, in which

- particle moves in one dimension,
- particle moves to and fro about a fixed mean position (where $F_{net} = 0$),
- net force on the particle is always directed towards means position, and
- magnitude of net force is always proportional to the displacement of particle from the mean position at that instant.

So, $F_{net} = -kx$

where, k is known as force constant

$ma = -kx$

## Some Important terms

### Amplitude

The amplitude of particle executing S.H.M. is its maximum displacement on either side of the mean position. A is the amplitude of the particle.

### Time Period

Time period of a particle executing S.H.M. is the time taken to complete one cycle and is denoted by T.

### Frequency

The frequency of a particle executing S.H.M. is equal to the number of oscillations completed in one second.

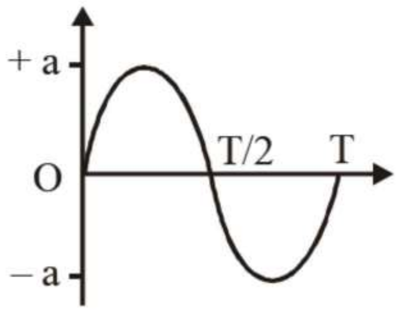
### Phase

The phase of particle executing S.H.M. at any instant is its state as regard to its position and direction of motion at that instant. It is measured as argument (angle) of sine in the equation of S.H.M.

$$\text{Phase} = (\omega t + \phi)$$

At t=0, phase=$\phi$; the constant $\phi$ is called initial phase of the particle or phase constant.

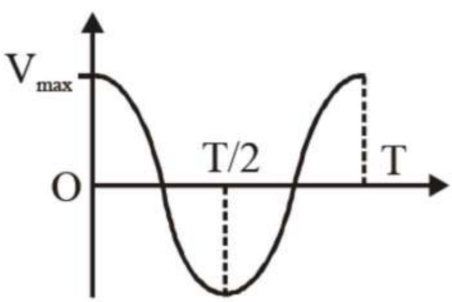
**Position**




If mean position is at origin the position (X coordinate) depends on time in general as:

x (t) = sin ($\omega$t + $\phi$)  At mean position, x=0  At extremes, x=+A,–A

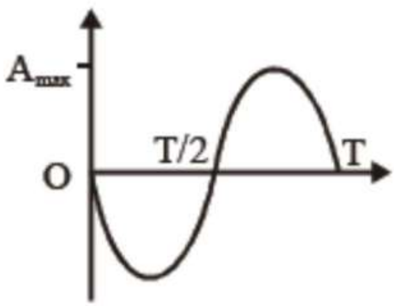
**Velocity**




At any time instant t,v(t) =A$\omega$cos($\omega$t+$\phi$)  At any position x,v(x)=$\pm\omega\sqrt{A^2} - x^2$

Velocity is minimum at extremes because the particle is at rest. i.e., v = 0 at extreme position. Velocity has maximum magnitude at mean position.

**Acceleration**




At any instant t, a (t) = – $\omega^2$ A sin ($\omega$t+$\phi$)

At any position x,a(x)=–$\omega^2$x

Acceleration is always directed towards mean position. The magnitude of acceleration is minimum at mean position and maximum at extremes.

**Displacement as a function of time and Periodic function**

To understand this idea of displacement as a function of time, we will have to derive an expression for displacement, assume a body travelling at an initial velocity of v1 at time t1 and then the body accelerates at a constant acceleration of 'a' for some time and a final velocity of v2 at time t2, keeping these things in assumption let's derive the following.


Let's write displacement as,

$$d = V_{average} * \Delta t$$

Where Δt is the change in time, assuming that the object is under constant acceleration.

$$d = \frac{v_1 - v_2}{2} * \Delta t$$

Where $V_2$ and $V_1$ are final and initial velocities respectively, let's rewrite final velocity in terms of initial velocity for the sake of simplicity.

$$d = \left(\frac{v_1(v_1 + a * \Delta t)}{2}\right) * \Delta\ t$$

Where a is the constant acceleration the body is moving at, now if we rewrite the above as,

$$d = \left(\frac{2 \times v_1}{2} + \frac{a * \Delta t}{2}\right) * \Delta\ t$$

The above expression is one of the most fundamental expressions in kinematics; it is also sometimes given as

$$d = v_i \mathbf{t} + \frac{1}{2} a t^2$$

Where $v_i$ the initial velocity and t is actually the change in time, all the quantities in this derivation like Velocity, displacement and acceleration are vector quantities.

If we have a look at our final expression, it is very clear that change in displacement depends upon time.

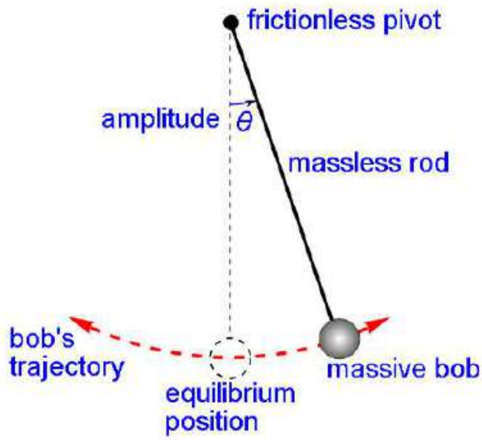
**Example of an oscillating pendulum**

The graph of the motion of the bob shows us how the displacement varies with respect to time, when bob reaches highest point, the potential energy is maximum and the kinetic energy is minimum as the velocity is equal to zero, we can easily deduce that the velocity is equal to zero, by looking at the slope of the displacement-time graph at specific times.

Slope A is positive in the graph stating that the velocity of the body is also positive, or in the forward direction.

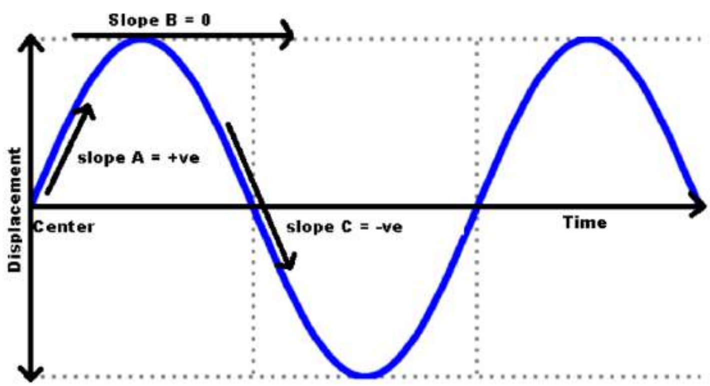
Slope B is equal to zero meaning that the velocity of the body is zero, meaning there is no motion at this point. Slope C is negative in the graph stating that the velocity of the body is also negative, or in the reverse direction.

The magnitude of velocity is always positive; it is the direction that decides if the velocity is positive or negative.




Now if we draw a displacement-time graph of this oscillating pendulum bob, we will get something like this,

The displacement of the bob repeats after certain amount of time, it is a




periodic function hence the displacement at any given time in the future can be predicted if we know the time and the time period of the pendulum, hence we can say that displacement is a function of time.

**Simple harmonic motion**, in physics, repetitive movement back and forth through an equilibrium, or central, position, so that the maximum displacement on one side of this position is equal to the maximum displacement on the other side. The time interval of each complete vibration is the same. The force responsible for the motion is always directed toward the equilibrium position and is directly proportional to the distance from it. That is, $F = -kx$, where $F$ is the force, $x$ is the displacement, and $k$ is a constant. This relation is called Hooke's law.

A specific example of a simple harmonic oscillator is the vibration of a mass attached to a vertical spring, the other end of which is fixed in a ceiling. At the maximum displacement $-x$, the spring is under its greatest tension, which forces the mass upward. At the maximum displacement $+x$, the spring reaches its greatest compression, which forces the mass back downward again. At either position of maximum displacement, the force is greatest and is directed toward the equilibrium position, the velocity ($v$) of the mass is zero, its acceleration is at a maximum, and the mass changes direction. At the equilibrium position, the velocity is at its maximum and the acceleration ($a$) has fallen to zero. Simple harmonic motion is characterized by this changing acceleration that always is directed toward the equilibrium position and is proportional to the displacement from the equilibrium position. Furthermore, the interval of time for each complete vibration is constant and does not depend on the size of the maximum displacement. In some form, therefore, simple harmonic motion is at the heart of timekeeping.

To express how the displacement of the mass changes with time, one can use Newton's second law, $F = ma$, and set $ma = -kx$. The acceleration $a$ is the second derivative of $x$ with respect to time $t$, and one can solve the resulting differential equation with $x = A \cos \omega t$, where $A$ is the maximum displacement and $\omega$ is the angular frequency in radians per second. The time it takes the mass to move from $A$ to $-A$ and back again is the time it takes for $\omega t$ to advance by $2\pi$. Therefore, the period $T$ it takes for the mass to move from $A$ to $-A$ and back again is $\omega T = 2\pi$, or $T = 2\pi/\omega$. The frequency of the vibration in cycles per second is $1/T$ or $\omega/2\pi$.

Many physical systems exhibit simple harmonic motion (assuming no energy loss): an oscillating pendulum, the electrons in a wire carrying alternating current, the vibrating particles of the medium in a sound wave, and other assemblages involving relatively small oscillations about a position of stable equilibrium.

The motion is called harmonic because musical instruments make such vibrations that in turn cause corresponding sound waves in air. Musical sounds are actually a combination of many simple harmonic waves corresponding to the many ways in which the vibrating parts of a musical instrument oscillate in sets of superimposed simple harmonic motions, the frequencies of which are multiples of a lowest fundamental frequency. In fact, any regularly repetitive motion and any wave, no matter how complicated its form, can be treated as the sum of a series of simple harmonic motions or waves, a discovery first published in 1822 by the French mathematician Joseph Fourier.

When you pluck a guitar string, the resulting sound has a steady tone and lasts a long time (Figure 1). The string vibrates around an equilibrium position, and one oscillation is

Figure 1 When a guitar string is plucked, the string oscillates up and down in periodic motion. The vibrating string causes the surrounding air molecules to oscillate, producing sound waves.

completed when the string starts from the initial position, travels to one of the extreme positions, then to the other extreme position, and returns to its initial position. We define **periodic motion** to be any motion that repeats itself at regular time intervals, such as exhibited by the guitar string or by a child swinging on a swing. In this section, we study the basic characteristics of oscillations and their mathematical description.

**Period and Frequency in Oscillations**

In the absence of friction, the time to complete one oscillation remains constant and is called the period (T). Its units are usually seconds, but may be any convenient unit of time. The word 'period' refers to the time for some event whether repetitive or not, but in this chapter, we shall deal primarily in periodic motion, which is by definition repetitive.

A concept closely related to period is the frequency of an event. Frequency (f) is defined to be the number of events per unit time. For periodic motion, frequency is the number of oscillations per unit time. The relationship between frequency and period is

$$f = 1/T$$

The SI unit for frequency is the *hertz* (Hz) and is defined as one *cycle per second.*

A cycle is one complete **oscillation.**

**Characteristics of Simple Harmonic Motion**

A very common type of periodic motion is called simple harmonic motion (SHM). A system that oscillates with SHM is called a simple harmonic oscillator.

In simple harmonic motion, the acceleration of the system, and therefore the net force, is proportional to the displacement and acts in the opposite direction of the displacement.

A good example of SHM is an object with mass mm attached to a spring on a frictionless surface, as shown in Figure 15.2.215.2.2. The object oscillates around the equilibrium position, and the net force on the object is equal to the force provided by the spring. This force obeys Hooke's law Fs = – kx, as discussed in a previous chapter.

If the net force can be described by Hooke's law and there is no damping (slowing down due to friction or other non-conservative forces), then a simple harmonic oscillator oscillates with equal displacement on either side of the equilibrium position, as shown for an object on a spring in Figure 2. The maximum displacement from equilibrium is called the amplitude (A). The units for amplitude and displacement are the same but depend on the type of oscillation. For the object on the spring, the units of amplitude and displacement are meters.

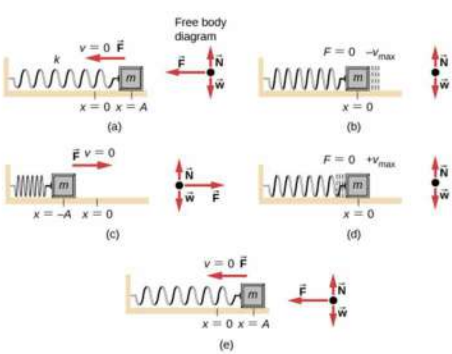


Fig.2 An object attached to a spring sliding on a frictionless surface is an uncomplicated simple harmonic oscillator. In the above set of figures, a mass is attached to a spring and placed on a frictionless table.

The other end of the spring is attached to the wall. The position of the mass, when the spring is neither stretched nor compressed, is marked as $x = 0$ and is the equilibrium position. (a) The mass is displaced to a position $x = A$ and released from rest. (b) The mass accelerates as it moves in the negative x-direction, reaching a maximum negative velocity at $x = 0$. (c) The mass continues to move in the negative x-direction, slowing until it comes to a stop at $x = -A$. (d) The mass now begins to accelerate in the positive x direction, reaching a positive maximum velocity at $x = 0$. (e) The mass then continues to move in the positive direction until it stops at $x = A$. The mass continues in SHM that has an amplitude A and a period T. The object's maximum speed occurs as it passes through equilibrium. The stiffer the spring is, the smaller the period T. The greater the mass of the object is, the greater the period T.

What is so significant about SHM? For one thing, the period T and frequency f of a simple harmonic oscillator are independent of amplitude. The string of a guitar, for example, oscillates with the same frequency whether plucked gently or hard.

Two important factors do affect the period of a simple harmonic oscillator. The period is related to how stiff the system is. A very stiff object has a large **force constant (k)**, which causes the system to have a smaller period. For example, you can adjust a diving board's stiffness—the stiffer it is, the faster it vibrates, and the shorter its period. Period also depends on the mass of the oscillating system. The more massive the system is, the longer the period. For example, a heavy person on a diving board bounces up and down more slowly than a light one. In fact, the mass m and the force constant k are the only factors that affect the period and frequency of SHM. To derive an equation for the period and the frequency, we must first define and analyze the equations of motion. Note that the force constant is sometimes referred to as the *spring constant*.

## Equations of SHM

Consider a block attached to a spring on a frictionless table (Figure 3).The equilibrium position (the position where the spring is neither stretched nor compressed) is marked as x = 0. At the equilibrium position, the net force is zero.

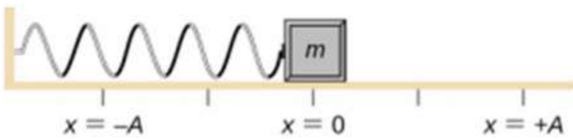



Figure 3 A block is attached to a spring and placed on a frictionless table. The equilibrium position, where the spring is neither extended nor compressed, is marked as x = 0.

Work is done on the block to pull it out to a position of x = + A, and it is then released from rest. The maximum x-position (A) is called the amplitude of the motion. The block begins to oscillate in SHM between x = + A and x = –A, where A is the amplitude of the motion and T is the period of the oscillation. The period is the time for one oscillation. Figure 4 shows the motion of the block as it completes one and a half oscillations after release.

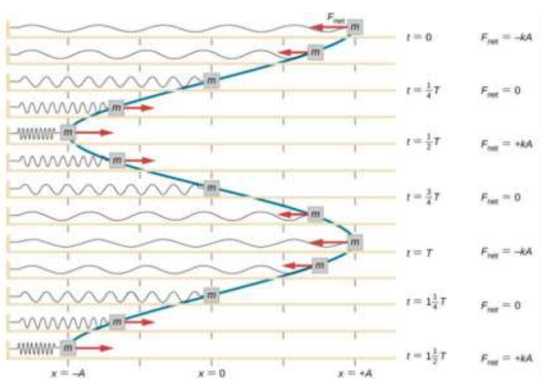


Figure 4 shows a plot of the position of the block versus time. When the position is plotted versus time, it is clear that the data can be modeled by a cosine function with an amplitude A and a period T. The cosine function cosθ repeats every multiple of 2π, whereas the motion of the block repeats every period T. However, the function $\cos(\frac{2\pi}{T}t)$repeats every integer multiple of the period. The maximum of the cosine function is one, so it is necessary to multiply the cosine function by the amplitude A.

$$\boldsymbol{x(t) = A\omega s\left(\frac{2\pi}{T}t\right) = Acos(\omega t)}$$

Recall from the chapter on rotation that the angular frequency equals $\omega=\frac{d\theta}{dt}$. In this case, the period is constant, so the angular frequency is defined as 2π divided by the period, $\omega=\frac{2\pi}{T}$.

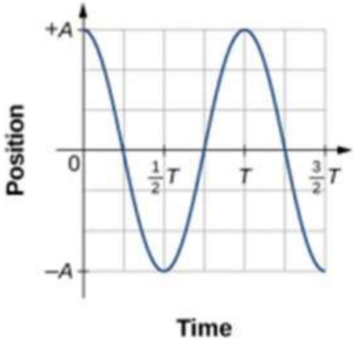



Figure 5: A graph of the position of the block shown in Figure 4 as a function of time. The position can be modeled as a periodic function, such as a cosine or sine function.

The equation for the position as a function of time x(t)=Acos(ωt)x(t)=Acos (ωt) is good for modeling data, where the position of the block at the initial time t = 0.00 s is at the amplitude A and the initial velocity is zero. Often when taking experimental data, the position of the mass at the initial time t = 0.00 s is not equal to the amplitude and the initial velocity is not zero. Consider 10 seconds of data collected by a student in lab, shown in Figure 6.

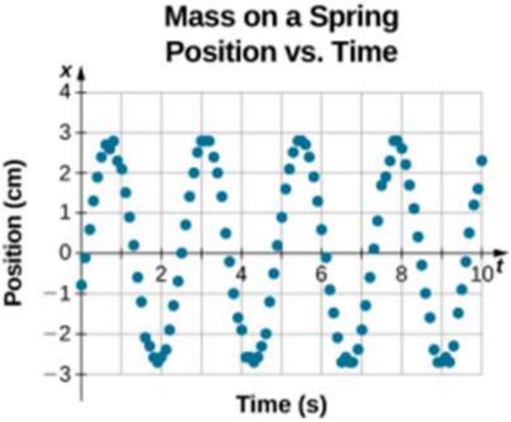



Figure 6: Data collected by a student in lab indicate the position of a block attached to a spring, measured with a sonic range finder. The data are collected starting at time t = 0.00s, but the initial position is near position $x \approx -0.80$ cm $\neq 3$

The data in Figure 6 can still be modeled with a periodic function, like a cosine function, but the function is shifted to the right. This shift is known as a **phase shift** and is usually represented by the Greek letter phi (φ). The equation of the position as a function of time for a block on a spring becomes

$$x(t) = A\cos(\omega t + \phi)$$

This is the generalized equation for SHM where t is the time measured in seconds, ω is the angular frequency with units of inverse seconds, A is the amplitude measured in meters or centimeters, and φ is the phase shift measured in radians (Figure 7). It should be noted that because sine and cosine functions differ only by a phase shift, this motion could be modeled using either the cosine or sine function.

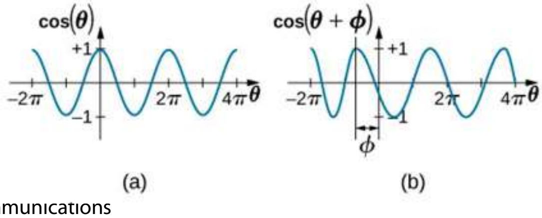


Figure 7 (a) A cosine function. (b) A cosine function shifted to the right by an angle $\phi$ . The angle $\phi$ is known as the phase shift of the function.

The velocity of the mass on a spring, oscillating in SHM, can be found by taking the derivative of the position equation:

$$v(t) = \frac{dx}{dt} = \frac{d}{dt}\left(Acos(\omega t + \phi)\right) = -A\omega sin(\omega t + \phi)$$
$$= -v_{max}\, sin(\omega t + \phi)$$

Because the sine function oscillates between −1 and +1, the maximum velocity is the amplitude times the angular frequency, $v_{max} = A\omega$. The maximum velocity occurs at the equilibrium position (x = 0) when the mass is moving toward x = + A. The maximum velocity in the negative direction is attained at the equilibrium position (x = 0) when the mass is moving toward x = −A and is equal to $-v_{max}$ .

The acceleration of the mass on the spring can be found by taking the time derivative of the velocity:

$$a(t) = \frac{dv}{dt} = \frac{d}{dt}\left(-A\omega sin(\omega t + \phi)\right) = -A\omega^2 cos(\omega t + \phi)$$
$$= -a_{max} cos(\omega t + \phi)$$

The maximum acceleration is $a_{max}$. = $A\omega^2$. The maximum acceleration occurs at the position (x = −A), and the acceleration at the position (x = −A) and is equal to $-a_{max}$.

**Energy in Simple Harmonic Motion**

The total energy that a particle possesses while performing simple harmonic motion is energy in simple harmonic motion. Take a pendulum for example. When it is at its mean position, it is at rest. When it moves towards its extreme position, it is in motion and as soon as it reaches its extreme position, it comes to rest again. Therefore, in order to calculate the energy in simple harmonic motion, we need to calculate the kinetic and potential energy that the particle possesses.

**Kinetic Energy (K.E.) in S.H.M**

Kinetic energy is the energy possessed by an object when it is in motion. Let's learn how to calculate the kinetic energy of an object. Consider a particle with mass m performing simple harmonic motion along a path AB. Let O be its mean position. Therefore, OA = OB = a.

The instantaneous velocity of the particle performing S.H.M. at a distance x from the mean position is given by

$v= \pm\omega \sqrt{a^2 - x^2}$

$\therefore v^2 = \omega^2 (a^2 - x^2)$

$\therefore$ Kinetic energy= $1/2\ mv^2 = 1/2\ m\ \omega^2 (a^2 - x^2)$

As, $k/m = \omega^2$

$\therefore k = m\ \omega^2$

Kinetic energy= $1/2\ k (a^2 - x^2)$ . The equations Ia and Ib can both be used for calculating the kinetic energy of the particle.

**Potential Energy(P.E.) of Particle Performing S.H.M.**

Potential energy is the energy possessed by the particle when it is at rest. Let's learn how to calculate the potential energy of a particle performing S.H.M. Consider a particle of mass m performing simple harmonic motion at a distance x from its mean position. You know the restoring force acting on the particle is F= -kx where k is the force constant.

Now, the particle is given further infinitesimal displacement dx against the restoring force F. Let the work done to displace the particle be dw. Therefore, The work done dw during the displacement is

$dw = -fdx = -(-kx)dx = kxdx$

Therefore, the total work done to displace the particle now from 0 to x is

$\int dw= \int kxdx = k \int x\ dx$

Hence Total work done = $1/2\ Kx^2 = 1/2\ m\ \omega^2x^2$

The total work done here is stored in the form of potential energy.

Therefore Potential energy = $1/2\ kx^2 = 1/2\ m\ \omega^2x^2$

Equations IIa and IIb are equations of potential energy of the particle. Thus, potential energy is directly proportional to the square of the displacement, that is P.E. α x2.

**Total Energy in Simple Harmonic Motion (T.E.)**

The total energy in simple harmonic motion is the sum of its potential energy and kinetic energy.

Thus, T.E. = K.E. + P.E. = $1/2\ k (a^2- x^2) + 1/2\ Kx^2 = 1/2\ ka^2$

Hence, T.E.= E = $1/2\ m\ \omega^2a^2$

Equation III is the equation of total energy in a simple harmonic motion of a particle performing the simple harmonic motion. As $\omega^2$, $a^2$ are constants, the total energy in the simple harmonic motion of a particle performing simple harmonic motion remains constant. Therefore, it is independent of displacement x.

As $\omega = 2\pi f$ , $E = 1/2\ m\ (2\pi f)^2 a^2$

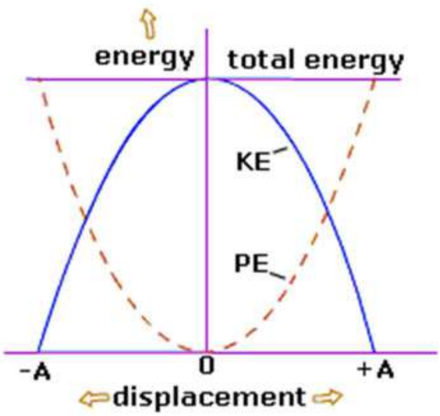
$\therefore E = 2m\pi^2 f^2 a^2$

As 2 and $\pi^2$ constants, we have T.E. ~ m, T.E. ~ f 2, and T.E. ~ $a^2$

Thus, the total energy in the simple harmonic motion of a particle is:

- Directly proportional to its mass
- Directly proportional to the square of the frequency of oscillations and
- Directly proportional to the square of the amplitude of oscillation.

The law of conservation of energy states that energy can neither be created nor destroyed. Therefore, the total energy in simple harmonic motion will always be constant. However, kinetic energy and potential energy are interchangeable. Given below is the graph of kinetic and




potential energy vs. instantaneous displacement.

In the graph, we can see that,

- At the mean position, the total energy in simple harmonic motion is purely kinetic and at the extreme position, the total energy in simple harmonic motion is purely potential energy.
- At other positions, kinetic and potential energies are

interconvertible and their sum is equal to 1/2 k $a^2$.

- The nature of the graph is parabolic.

**Damped Simple Harmonic Motion**

When the motion of an oscillator reduces due to an external force, the oscillator and its motion are **damped**. These periodic motions of gradually decreasing amplitude are damped simple harmonic motion. An example of a damped simple harmonic motion is a simple pendulum.

In the damped simple harmonic motion, the energy of the oscillator dissipates continuously. But for a small damping, the oscillations remain approximately periodic. The forces which dissipate the energy are generally **frictional forces**.

Expression of damped simple harmonic motion

Let's take an example to understand what a damped simple harmonic motion is. Consider a block of mass m connected to an elastic string of spring constant k. In an ideal situation, if we push the block down a little and then release it, its angular frequency of oscillation is $\omega = \sqrt{k}/ m$.

However, in practice, an external force (air in this case) will exert a damping force on the motion of the block and the mechanical energy of the block-string system will decrease. This energy that is lost will appear as the heat of the surrounding medium.

The damping force depends on the nature of the surrounding medium. When we immerse the block in a liquid, the magnitude of damping will be much greater and the dissipation energy is much faster. Thus, the damping force is proportional to the velocity of the bob and acts opposite to the direction of the velocity. If the damping force is Fd, we have,

$$Fd = -b\upsilon \qquad (I)$$

where the constant b depends on the properties of the medium(viscosity, for example) and size and shape of the block. Let's say O is the equilibrium position where the block settles after releasing it. Now, if we pull down or push the block a little, the restoring force on the block due to spring is $Fs = -kx$, where x is the displacement of the mass from its equilibrium position. Therefore, the total force acting on the mass at any time t is, $F = -kx - b\upsilon$.

Now, if a(t) is the acceleration of mass m at time t, then by Newton's Law of Motion along the direction of motion, we have

$$ma(t) = -kx(t) - bυ(t) \qquad (II)$$

Here, we are not considering vector notation because we are only considering the one-dimensional motion. Therefore, using first and second derivatives of s(t), v(t) and a(t), we have,

$$m(d^2x/dt^2) + b(dx/dt) + kx = 0 \qquad (III)$$

This equation describes the motion of the block under the influence of a damping force which is proportional to velocity. Therefore, this is the expression of damped simple harmonic motion. The solution of this expression is of the form

$$x(t) = Ae^{-bt/2m} \cos(ω't + ø) \qquad (IV)$$

where A is the amplitude and ω' is the angular frequency of damped simple harmonic motion given by,

$$ω' = \sqrt{(k/m - b^2/4m^2)} \qquad (V)$$

The function x(t) is not strictly periodic because of the factor e-bt/2m which decreases continuously with time. However, if the decrease is small in one-time period T, the motion is then approximately periodic. In a damped oscillator, the amplitude is not constant but depends on time. But for small damping, we may use the same expression but take amplitude as Ae-bt/2m

$$\therefore E(t) = 1/2\ kAe^{-bt/2m} \qquad (VI)$$

This expression shows that the damping decreases exponentially with time. For a small damping, the dimensionless ratio (b/√km) is much less than 1. Obviously, if we put b = 0, all equations of damped simple harmonic motion will turn into the corresponding equations of un damped motion.

**Forced Simple Harmonic Motion**

When we displace a system, say a simple pendulum, from its equilibrium position and then release it, it oscillates with a natural frequency ω and these oscillations are free oscillations. But all free oscillations eventually die out due to the ever present damping forces in the surrounding.

However, an external agency can maintain these oscillations. These oscillations are known as forced or driven oscillations. The motion that the system performs under this external agency is known as Forced Simple Harmonic Motion. The external force is itself periodic with a frequency ωd which is known as the drive frequency.

A very important point to note is that the system oscillates with the driven frequency and not its natural frequency in Forced Simple Harmonic Motion. If it oscillates with its natural frequency, the motion will die out. A good example of forced oscillations is when a child uses his feet to move the swing or when someone else pushes the swing to maintain the oscillations.

Expression of Forced Simple Harmonic Motion

Consider an external force F(t) of amplitude F0 that varies periodically with time. This force is applied to a damped oscillator. Therefore, we can represent it as,

$$F(t) = F_0 \cos \omega_d t \qquad \text{(I)}$$

Thus, at this time, the forces acting on the oscillator are its restoring force, the external force and a time-dependent driving force. Therefore,

$$ma(t) = -kx(t) - b\upsilon(t) + F_0 \cos \omega_d t \qquad \text{(II)}$$

We know that acceleration = $d^2x/dt^2$. Substituting this value of acceleration in equation II , we get,

$$m(d^2x/dt^2) + b(dx/dt) + kx = F_0 \cos \omega_d t \qquad \text{(III)}$$

Equation III is the equation of an oscillator of mass m on which a periodic force of frequency ωd is applied. Obviously, the oscillator first oscillates with its natural frequency. When we apply the external periodic force, the oscillations with natural frequency die out and the body then oscillates with the driven frequency. Therefore, its displacement after the natural oscillations die out is given by:

$$x(t) = A\cos(\omega_d + \text{ø}) \qquad \text{(IV)}$$

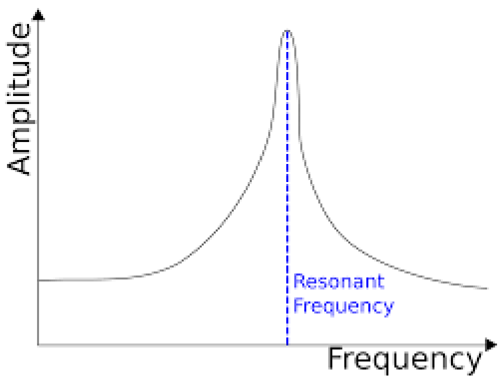
where t is the time from the moment we apply external periodic force.

**Resonance**

The phenomenon of increase in amplitude when the driving force is close to the natural frequency of the oscillator is known as resonance. To understand the phenomenon of resonance, let us consider two pendulums of nearly equal (but not equal) lengths (therefore, different amplitudes) suspended from the same rigid support.

When we swing the first pendulum which is greater in length, it oscillates with its natural frequency. The energy of this pendulum transfers through the rigid support to the second pendulum which is slightly smaller in length. Therefore, the second pendulum starts oscillating with its natural frequency first.

At one point, the frequency with the second pendulum vibrates becomes nearly equal to the first one. Therefore, the second pendulum now starts with the frequency of the first one, which is the driven frequency. When this happens, the amplitude of the oscillations is maximum. Thus, resonance takes place.

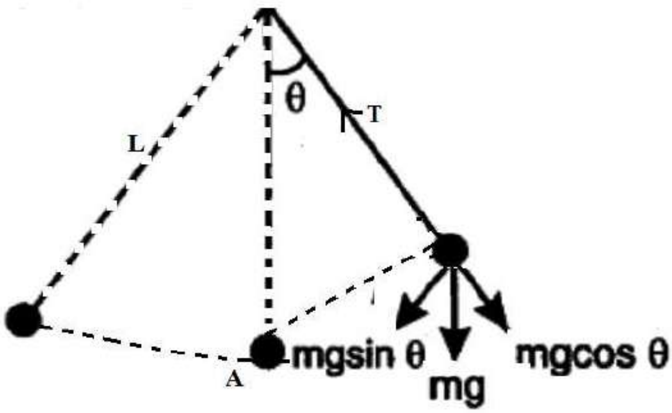



**Time Period of Simple Pendulum**

A point mass M suspended from the end of a light inextensible string whose upper end is fixed to a rigid support. The mass displaced from its mean position.

Assumptions:

- There is negligible friction from the air and the system
- The arm of the pendulum does not bend or compress and is massless
- The pendulum swings in a perfect plane
- Gravity remains constant




**Time Period of Simple Pendulum Derivation**

Using the equation of motion, $T - mg\cos\theta = mv^2L$

The torque tending to bring the mass to its equilibrium position,

$\tau = mgL \times \sin\theta = mg\sin\theta \times L = I \times \alpha$

For small angles of oscillations $\sin \approx \theta$,

Therefore, $I\alpha = -mgL\theta$

$\alpha = -(mgL\theta)/I$

$-\omega_0^2\,\theta = -(mgL\theta)/I$

$\omega_0^2 = (mgL)/I$

$\omega^2{}_0 = \sqrt{(mgL/I)}$

Using $I = ML^2$, [where I denote the moment of inertia of bob]

we get, $\omega_0 = \sqrt{(g/L)}$

Therefore, the time period of a simple pendulum is given by,

$T = 2\pi/\omega_0 = 2\pi \times \sqrt{(L/g)}$

**Wave Motion**

Wave motion is a mode of transmission of energy through a medium in the form of a disturbance. It is due to the repeated periodic motion of the particles of the medium about an equilibrium position transferring the energy from one particle to another.

The waves are of three types - mechanical, electromagnetic and matter waves. Mechanical waves can be produced only in media which possess elasticity and inertia. Water waves, sound waves and seismic waves are common examples of this type. Electromagnetic waves do not require any material medium for propagation. Radio waves, microwaves, infrared rays, visible light, the ultraviolet rays, X rays and Y rays are electromagnetic waves. The waves associated with particles like electrons, protons and fundamental particles in motion are matter waves.

**Waves on surface of water**

In order to understand the concept of wave motion, let us drop a stone in a trough of water. We find that small circular waves seem to originate from the point where the stone touches the surface of water. These waves spread out in all directions. It appears as if water moves away from that point. If a piece of paper is placed on the water surface, it will be observed that the piece of paper moves up and down, when the waves pass through it. This shows that the waves are formed due to the vibratory motion of the water particles, about their mean position.

Wave motion is a form of disturbance which travels through a medium due to the repeated periodic motion of the particles of the medium about their mean position. The motion is transferred continuously from one particle to its neighboring particle.

**Characteristics of wave motion**

Wave motion is a form of disturbance travelling in the medium due to the periodic motion of the particles about their mean position.

- It is necessary that the medium should possess elasticity and inertia.
- All the particles of the medium do not receive the disturbance at the same instant (i.e) each particle begins to vibrate a little later than its predecessor.
- The wave velocity is different from the particle velocity. The velocity of a wave is constant for a given medium, whereas the velocity of the particles goes on changing and it becomes maximum in their mean position and zero in their extreme positions.
- During the propagation of wave motion, there is transfer of energy from one particle to another without any actual transfer of the particles of the medium.

The waves undergo reflection, refraction, diffraction and interference.

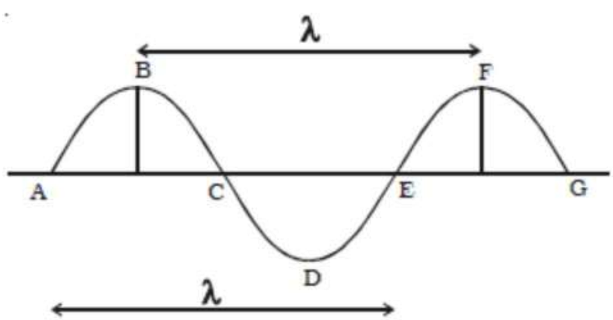
Mechanical wave motion

The two types of mechanical wave motion are (i) transverse wave motion and (ii) longitudinal wave motion

**Transverse wave motion**

Transverse wave motion is that wave motion in which particles of the medium execute SHM about their mean positions in a direction perpendicular to the direction of propagation of the wave. Such waves are called transverse waves. Examples of transverse waves are waves produced by plucked strings of veena, sitar or violin and electromagnetic waves. Transverse waves travel in the form of crests and troughs. The maximum displacement of the particle in the positive direction i.e. above its mean position is called crest and maximum displacement of the particle in the negative direction i.e below its mean position is called trough. Thus if ABCDEFG is a transverse Wave, the points B and F are crests while D is trough (Fig.1) For the propagation of transverse cannot be produced in gases and liquids.

Transverse waves can be produced in solids and surfaces of liquids only.




**Longitudinal wave motion**

‘Longitudinal wave motion is that wave motion in which each particle of the medium executes simple harmonic motion about its mean position along the direction of propagation of the wave.’

Sound waves in fluids (liquids and gases) are examples of longitudinal wave. When a longitudinal wave travels through a medium, it produces compressions and rarefactions.

**Important terms used in wave motion**

(i) Wavelength (λ)

The distance travelled by a wave during which a particle of the medium completes one vibration is called wavelength. It is also defined as the distance between any two nearest particles on the wave having same phase.

Wavelength may also be defined as the distance between two successive crests or troughs in transverse waves, or the distance between two successive compressions or rarefactions in longitudinal waves.

(ii) Time period (T)

The time period of a wave is the time taken by the wave to travel a distance equal to its wavelength.

(iii) Frequency (n)

This is defined as the number of waves produced in one second. If T represents the time required by a particle to complete one vibration, then it makes 1/T waves in one second.

Therefore frequency is the reciprocal of the time period (i.e) n = 1/T.

Relationship between velocity, frequency and wavelength of a wave

The distance travelled by a wave in a medium in one second is called the velocity of propagation of the wave in that medium. If v represents the velocity of propagation of the wave, it is given by

$$\text{Velocity} = \frac{\text{Distance travelled}}{\text{Time taken}}$$

$$v = \frac{\lambda}{T} = n\lambda \qquad \left[\because n = \frac{1}{T}\right]$$

The velocity of a wave (v) is given by the product of the frequency and wavelength. During day time, the upper layers of air are cooler than the layers of air near the surface of the Earth. During night, the layers of air near the Earth are cooler than the upper layers of air. As sound travels faster in hot air, during day-time, the sound waves will be refracted upwards and travel a short distance on the surface of the Earth. On the other hand, during night the sound waves are refracted downwards to the Earth and will travel a long distance.

**Superposition principle**

When two waves travel in a medium simultaneously in such a way that each wave represents its separate motion, then the resultant displacement at any point at any time is equal to the vector sum of the individual displacements of the waves.

This principle is illustrated by means of a slinky in the Fig.10 (a).

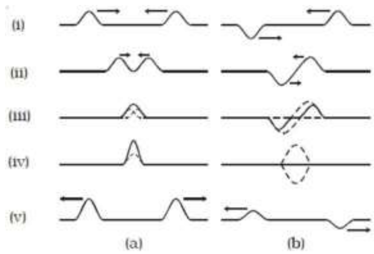
In the figure,(i) shows that the two pulses pass each other,

In the figure, (ii) shows that they are at some distance apart

In the figure, (iii) shows that they overlap partly

In the figure, (iv) shows that resultant is maximum

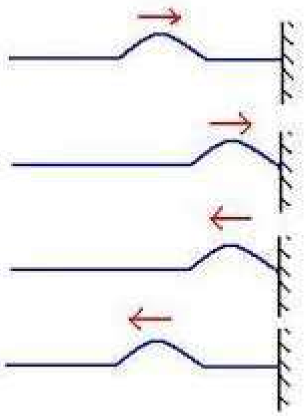
Fig. 10 b illustrates the same events but with pulses that are equal and opposite.




**Reflection of Waves**

We encounter situations involving the reflection of waves all around us, for example, in the phenomenon of echo, the sound reflected from a distant object reaches the listener with a little delay. In this section, we will learn more about the reflection of a wave from a fixed and a free end.

In the image shown below, we can see what happens when a pulse or a travelling wave encounters a rigid boundary. We see how under such a situation the pulse or the wave gets reflected.

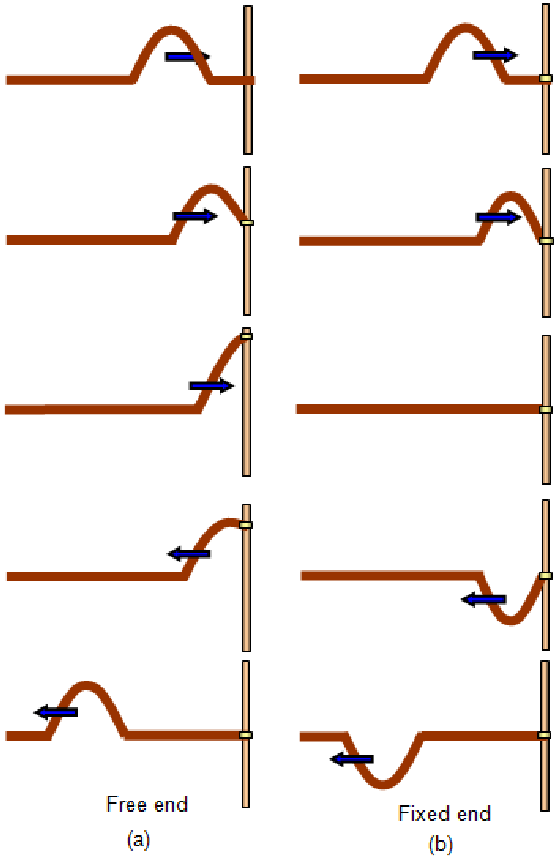


When a wave is an incident on any surface, a part of the incident wave is reflected and a part is transmitted into the second medium. If the wave is incident obliquely on the boundary, the transmitted wave can also be termed as a reflected wave.

Here, the incident and the refracted waves obey Snell's Law of refraction and the incident and the reflected waves obey the laws of reflection. The reflection of wave or a pulse can happen from two types of surfaces, it can either be a fixed wall or a ring, as shown in the image below.

Fixed End Reflection

Let us consider the situation where a string is fixed to a rigid wall at its right end. When we allow a pulse to propagate through these strings, the pulse reaches the right end, gets reflected as shown in the figure above. When the pulse arrives at the fixed end, it exerts a force on the wall and according to Newton's third law, the wall exerts an equal and



opposite force on the string. This second force generates a pulse at the support, which travels back along the string in the direction opposite to that of the incident pulse. In a reflection of this kind, there is no displacement at the support as the string is fixed there. The reflected and incident pulses have opposite signs and they cancel each other at that point. Thus, in the case of a travelling wave, the reflection at a rigid boundary takes place with a phase reversal or with a phase difference of $\pi$.

Free End Reflection

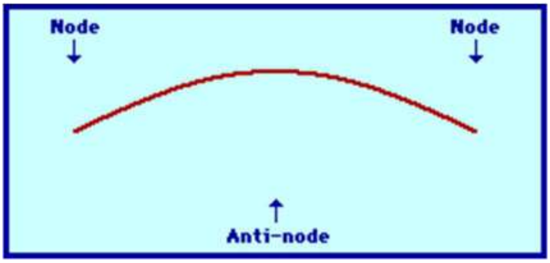
When the right end of the string is tied to a ring, which slides up and down without any friction on a rod, we term it as a free end. In this case, when the pulse arrives at the right end, the ring moves up the rod and as it moves, it pulls on the string, stretching the string and producing a reflected pulse with the same sign and amplitude as the incident pulse. Thus, in such a reflection, the incident and reflected pulses reinforce each other, creating the maximum displacement at the end of the string: the maximum displacement of the ring is twice the amplitude of either of the pulses. Thus, the reflection occurs without any additional phase shift. In case of a travelling wave the reflection at an open boundary the reflection takes place without any phase change.

Summarizing the above result, we can say that the reflection of waves at a boundary between two media takes place accordingly. A travelling wave, at a rigid boundary or a closed end, is reflected with a phase reversal but the reflection at an open boundary takes place without any phase change.

**Standing Waves on a String**

A standing wave pattern is a pattern which results from the interference of two or more waves along the same medium. All standing wave patterns are characterized by positions along the medium which are standing still. Such positions are referred to as nodal positions or nodes. Nodes occur at locations where two waves interfere such that one wave is displaced upward the same amount that a second wave is displaced downward. This form of interference is known as destructive interference and leads to a point of "no displacement." A node is a point of no displacement. Standing wave patterns are also characterized by antinodal positions - positions along the medium that vibrate back and

forth between a maximum upward displacement to a maximum downward displacement. Antinodes are located at positions along the medium where the two interfering waves are always undergoing constructive interference. Standing wave patterns are always characterized by an alternating pattern of nodes and antinodes.
There are a variety of patterns which could be produced by vibrations within a string, slinky, or rope. Each pattern corresponds to vibrations which occur at a particular frequency and is known as a harmonic. The lowest possible frequency at which a string could vibrate to form a standing wave pattern is known as the fundamental frequency or the



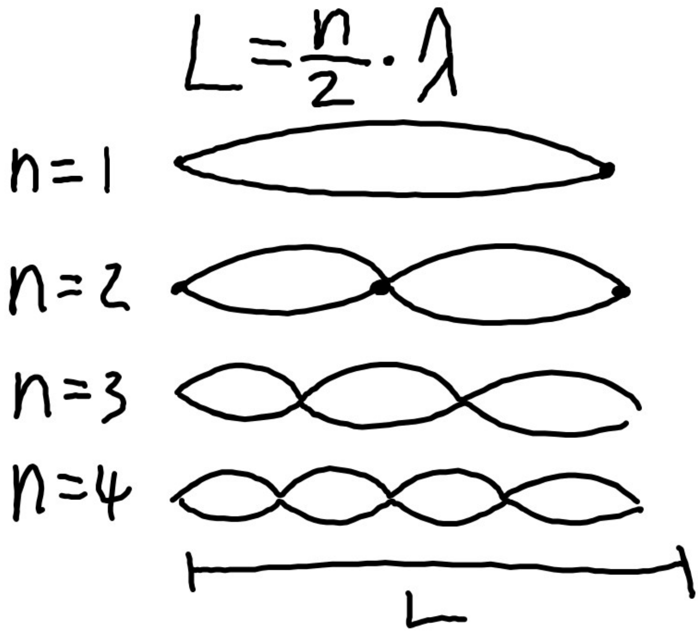

first harmonic. An animation of a string vibrating with the first harmonic is shown below.

The frequency associated with each harmonic is dependent upon the speed at which waves move through the medium and the wavelength of the medium. The speed at which waves move through a medium is dependent upon the properties of the medium (tension of the string, thickness of the string, material composition of the string, etc.). The wavelength of the harmonic is dependent upon the length of the string and the harmonic number (first, second, third, etc.). Variations in either the properties of the medium or the length of the medium will result in variations in the frequency at which the string will vibrate.

**Standing Waves on a String and pipes**

The important idea is that when a traveling wave hits a barrier there will be a reflection. If those reflections are timed exactly, then they will result in constructive and destructive interference in stationary locations. The locations where destructive interference occurs are called nodal points. At these nodal points, there is no movement of the string. The locations where constructive interference occurs are called anti-nodal points. At these points, the string would oscillate between Super crests and super troughs.

We also found for that a given string of a length L, with fixed end points, there can exist only half-integer number of wavelengths for standing waves on the string. λ is wavelength and n is an integer which happens

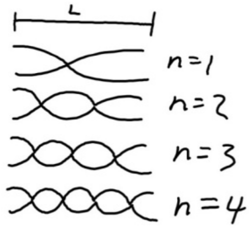



to be equal to the number of anti-nodes.

We can further, using the wave equation, express this as the possible frequencies that a string will resonate at.

$$f = nv/2L$$

where f is the frequencies, and v is the speed of sound in the string. So actually presents us with an equation for the harmonic frequencies that the string will vibrate at. As the week progresses we will connect this to actual musical notes. This of course, related to the different kinds of sounds that we can here on string instruments. The pitch or frequency that we hear depends on the length of the string and on the speed of sound in the string. And the speed of sound depends on the bulk density and the tension in the string.



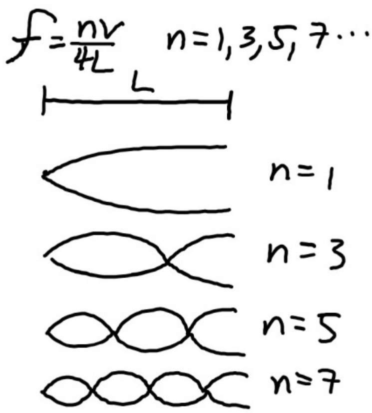

We looked at a case for a solid rod vibrating with loose ends. This sketch has the 1st, 2nd, 3rd and 4th harmonics.

harmonics for an open wind pipe would be identical to this as would be the equation. Even though there is a 1/4 wavelength at the ends, two quarters add up to a half. However, in a pipe, it is air that is vibrating so in the equation

$$f = nv/2L$$

v is the speed of sound in air (about 340 m/s) and not the speed in the material.

Whereas for a closed pipe (such as s Coke Bottle), with one end open and the other end closed, there will be an odd integer number of 1/4 wavelengths.




Also, when thinking about a Coke bottle making noise, remember the Length L, is not the length of the bottle, but how much air is in the bottle. So if you fill it with water, you decrease the amount of air.

## Introduction: vibrations, strings, pipes

To make a sound, we need something that vibrates. If we want to make musical notes you usually need the vibration to have an almost constant frequency: that means stable pitch. We also want a frequency that can be easily controlled by the player. In electronic instruments this is done with electric circuits or with clocks and memories. In non-electronic instruments, the stable, controlled vibration is produced by a standing wave. Here we discuss the way strings work. This also a useful introduction for studying wind instruments, because vibrating strings are easier to visualise than the vibration of the air in wind instruments. Both are less complicated than the vibrations of the bars and skins of the percussion family. For the physics of standing waves, there is a multimedia tutorial.

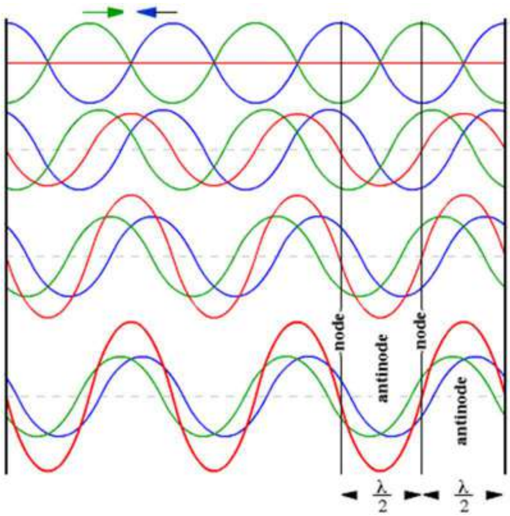
## Travelling waves in strings

The strings in the violin, piano and so on are stretched tightly and vibrate so fast that it is impossible to see what is going on. If you can find a long spring (a toy known as a 'slinky' works well) or several metres of flexible rubber hose you can try a few fun experiments which will make it easy to understand how strings work. (Soft rubber is good for this, garden hoses are not really flexible enough.) First hold or clamp one end and then, holding the other end still in one hand, stretch it a little (not too much, a little sag won't hurt). Now pull it aside with the other hand to make a kink, and then let it go. (This, in slow motion, is what happens when you pluck a string.) You will probably see that the kink travels down the "string", and then it comes back to you. It will suddenly tug your hand sideways but, if you are holding it firmly, it will reflect again.

First you will notice that the speed of the wave in the string increases if you stretch it more tightly. This is useful for tuning instruments - but we're getting ahead of ourselves. It also depends on the "weight" of the string - it travels more slowly in a thick, heavy string than in a light string of the same length under the same tension. (Strictly, it is the ratio of tension to mass per unit length that determines speed, as we'll see below.)

Next let's have a close look at the reflection at the fixed end. You'll notice that if you initially pull the string to the left, the kink that travels away from you is to the left, but that it comes back as a kink to the right - the reflection is inverted. This effect is important not only in string instruments, but in winds and percussion as well. When a wave encounters a boundary with something that won't move or change (or that doesn't change easily), the reflection is inverted. (The fact that it is inverted gives zero displacement at the end. However, reflection with any phase change will give a standing wave).

**Travelling waves and standing waves**




An interesting effect occurs if you try to send a simple wave along the string by repeatedly waving one end up and down. If you have found a suitable spring or rubber hose, try it out. Otherwise, look at these diagrams.

The figure shows the interaction of two waves, with equal frequency and magnitude, travelling in opposite directions: blue to the right, green to the left. The red line is their sum: the red wave is what happens when the two travelling waves add together (superpose is the technical term). By stopping the animation, you can check that the red wave really is the sum of the two interacting travelling waves.

The figure at right is the same diagram represented as a time sequence - time increases from top to bottom. You could think of it as representing a series of photographs of the waves, taken very quickly. The red wave is what we would actually see in a such photographs.

Suppose that the right hand limit is an immoveable wall. As discussed above, the wave is inverted on reflection so, in each "photograph", the blue plus green adds up to zero on the right hand boundary. The reflected (green) wave has the same frequency and amplitude but is travelling in the opposite direction.

At the fixed end they add to give no motion - zero displacement: after all it is this condition of immobility which causes the inverted reflection. But if you look at the red line in the animation or the diagram (the sum of the two waves) you'll see that there are other points where the string never moves! They occur half a wavelength apart. These motionless points are called nodes of the vibration, and they play an important role in nearly all of the instrument families. Halfway between the nodes are antinodes: points of maximum motion. But note that these peaks are not travelling along the string: the combination of two waves travelling in opposite directions produces a standing wave.

This is shown in the animation and the figure. Note the positions (nodes) where the two travelling waves always cancel out and the others (antinodes where they add to give an oscillation with maximum amplitude.

You could think of this diagram as a representation (not to scale) of the fifth harmonic on a string whose length is the width of the diagram. This brings us to the next topic.

**Harmonics and modes**

The string on a musical instrument is (almost) fixed at both ends, so any vibration of the string must have nodes at each end. Now that limits the possible vibrations. For instance the string with length L could have a standing wave with wavelength twice as long as the string (wavelength $\lambda = 2L$) as shown in the first sketch in the next series. This gives a node at either end or an antinode in the middle.

This is one of the modes of vibration of the string ("mode of vibration" just means style or way of vibrating). Several standing waves are shown in the next sketch.

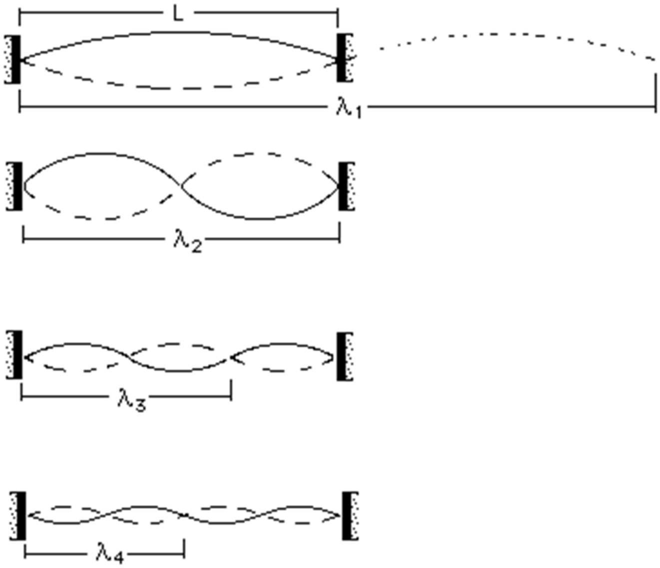
A sketch of the first four modes of vibration of an idealised stretched string with a fixed length. The vertical axis has been exaggerated.

Let's work out the relationships among the frequencies of these modes. For a wave, the frequency is the ratio of the speed of the wave to the length of the wave: $f = v/\lambda$. Compared to the string length L, you can see that these waves have lengths 2L, L, 2L/3, L/2. We could write this as 2L/n, where n is the number of the harmonic.

The fundamental or first mode has frequency $f_1 = v/\lambda_1 = v/2L$, The second harmonic has frequency $f_2 = v/\lambda_2 = 2v/2L = 2f_1$ The third harmonic has frequency $f_3 = v/\lambda_3 = 3v/2L = 3f_1$, The fourth harmonic has frequency $f_4 = v/\lambda_4 = 4v/2L = 4f_1$, and, to generalise,

The nth harmonic has frequency $f_n = v/\lambda_n = nv/2L = nf_1$.

All waves in a string travel with the same speed, so these waves with




different wavelengths have different frequencies as shown. The mode with the lowest frequency ($f_1$) is called the fundamental. Note that the nth mode has frequency n times that of the fundamental. All of the modes (and the sounds they produce) are called the harmonics of the string. The frequencies f, 2f, 3f, 4f etc are called the harmonic series. This series will be familiar to most musicians, particularly to buglers and players of natural horns. If for example the fundamental is the note C3 or viola C (a nominal frequency of 131 Hz: see this link for a table), then the harmonics would have the pitches shown in the next figure. These pitches have been approximated to the nearest quarter tone. The octaves are exactly octaves, but all other intervals are slightly different from the intervals in the equal tempered scale.

**Beats**

When two waves of nearly equal frequencies travelling in a medium along the same direction superimpose upon each other, beats are produced. The amptitude of the resultant sound at a point rises and falls regularly.

The intensity of the resultant sound at a point rises and falls regularly with time. When the intensity rises to maximum we call it as waxing of sound, when it falls to minimum we call it as waning of sound.

The phenomenon of waxing and waning of sound due to interference of two sound waves of nearly equal frequencies are called beats. The number of beats produced per second is called beat frequency, which is equal to the difference in frequencies of two waves.

**Uses of beats**

(i) The phenomenon of beats is useful in tuning two vibrating bodies in unison. For example, a sonometer wire can be tuned in unison with a tuning fork by observing the beats. When an excited tuning fork is kept on the sonometer and if the sonometer wire is also excited, beats are heard, when the frequencies are nearly equal. If the length of the wire is adjusted carefully so that the number of beats gradually decreases to zero, then the two are said to be in unison. Most of the musical instruments are made to be in unison based on this method.

(ii) The frequency of a tuning fork can be found using beats. A standard tuning fork of frequency N is excited along with the experimental fork. If the number of beats per second is n, then the frequency of experimental tuning fork is N+n. The experimental tuning fork is then loaded with a little bees' wax, thereby decreasing its frequency. Now the observations are repeated. If the number of beats increases, then the frequency of the experimental tuning fork is N-n, and if the number of beats decreases its frequency is N + n.

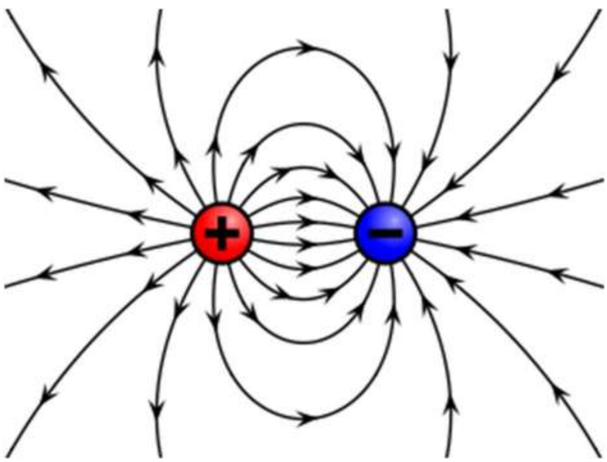
## Chapter II
## ELECTROSTATICS

Electrostatics is the branch of Physics, which deals with static electric charges or charges at rest. In this chapter, we shall study the basic phenomena about static electric charges. The charges in a electrostatic field are analogous to masses in a gravitational field. These charges have forces acting on them and hence possess potential energy. The ideas are widely used in many branches of electricity and in the theory of atom.

**Electrostatics – frictional electricity**

In 600 B.C., Thales, a Greek Philosopher observed that, when a piece of amber is rubbed with fur, it acquires the property of attracting light objects like bits of paper. In the 17$^{th}$ century, William Gilbert discovered that, glass, ebonite etc, also exhibit this property, when rubbed with suitable materials. The substances which acquire charges on rubbing are said to be 'electrified' or charged. These terms are derived from the Greek word electron, meaning amber. The electricity produced by friction is called frictional electricity. If the charges in a body do not move, then, the frictional electricity is also known as Static Electricity.

**Introduction to Protons & Electrons**



Majority of the electric charge is contained with the protons and electrons present within the atom. The negative charge is carried by electrons, whereas protons carry the positive charge. It is vital to know that, electrons and protons attract each other; the standard notion of “opposites attract” as framed by Coulomb.

Furthermore, protons and electrons are responsible for the development of electric fields, which apply a force termed as Coulomb force. This force is known to be outward radiating in all directions. Since protons are usually limited to the nuclei implanted inside atoms, their movement isn’t that free as compared to electrons.

Hence, whenever there is a question related to electric charge, it always points out to surplus or shortage of electrons. In case some imbalance happens, and electrons are allowed to flow, the generation of electric current can be experienced. After understanding the data mentioned above, this is the point when the question: what is charge? grows a bit clear to the readers.

**Electric charge**

Electric charge is a characteristic of certain subatomic particles. Both the proton and the electron carry electric charge while the neutron contains no charge. Scientists arbitrarily assigned a negative value to the electron and a positive value to the proton as a result of the fact that they attract. Therefore "like charges repel" and "unlike charges attract". In the same way that energy is conserved and cannot be destroyed nor created, electric charge also cannot be destroyed or created. In other words, electric charge is conserved and the total electric charge in our universe remains constant.

Electric Charge is nothing but the amount of energy or electrons that pass from one body to another by different modes like conduction, induction or other specific methods. This is a basic electric charge definition. There are two types of electric charges. They are positive charges and negative charges.

Charges are present in almost every type of body. All those bodies having no charges are the neutrally charged ones. We denote a charge y the symbol ‘q’ and its standard unit is Coulomb. Mathematically, we can say that a charge is the number of electrons multiplied by the charge on 1 electron. Symbolically, it is

$$Q = ne$$

Where q is a charge, n is a number of electrons and e is a charge on 1 electron ($1.6 \times 10^{-19}$C). The two very basic natures of electric charges are

Like charges repel each other.

Unlike charges attract each other.

This means that while protons repel protons, they attract electrons. The nature of charges is responsible for the forces acting on them and coordinating the direction of the flow of them. The charge on electron and proton is the same in magnitude which is $1.6 \times 10^{-19}$ C. The difference is only the sign that we use to denote them, + and -.

**Conservation of Charges**

A charge is a property associated with the matter due to which it produces and experiences electrical and magnetic effects. The basic idea behind the conservation of charge is that the total charge of the system is conserved. We can define it as:

Conservation of Charge is the principle that the total electric charge in an isolated system never changes. The net quantity of electric charge, the amount of positive charge minus the amount of negative charge in the universe, is always conserved.

As we know, the system is the group of objects and its interaction with charges is similar to conservation of energy and momentum, but this conservation law is more intuitive because the net charge of an object depends on the number of electrons and protons. The protons and electron cannot just appear or disappear out of nowhere; the total charge has to be the same. That's the reason there is always the same number of electrons and protons in a body.

It is known that every atom is electrically neutral, containing as many electrons as the number of protons in the nucleus. Bodies can also have any whole multiples of the elementary charge:

Electrical charge resides in electrons and protons, the smallest charge that a body can have is the charge of one electron or proton. [ie. $-1.6 \times 10^{-19}$ C or $+1.6 \times 10^{-19}$ C]

Explanation:

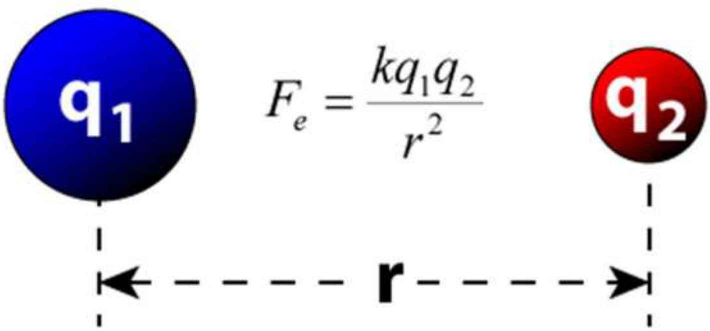
Law of conservation of charge says that the net charge of an isolated system will always remain constant. Let's try to understand it in more depth. There is a list of basically two ideal states for a system for multiple objects.

The first one is that the entire object has a net neutral charge. So in the whole system, there is the same number of protons and electrons, for each proton, there is an electron to balance it.

Another ideal state would be the net charge of the system being distributed uniformly in the objects. So rather than concentrating negative charge in a few bodies, the charge on the body is evenly distributed throughout by the transfer of the electron, and this can be achieved by the transfer of electrons from higher to lower polarity. Only electrons can be involved in the transfer charges, not proton.

**Coulomb's Law**

Coulomb's Law gives an idea about the force between two point charges. By the word point charge, we mean that in physics, the size of linear charged bodies is very small as against the distance between them. Therefore, we consider them as point charges as it becomes easy for us to calculate the force of attraction/ repulsion between them.




Charles-Augustin de Coulomb, a French physicist in 1784, measured the force between two point charges and he came up with the theory that the force is inversely proportional to the square of the distance between the charges. He also found that this force is directly proportional to the product of charges (magnitudes only).

We can show it with the following explanation. Let's say that there are two charges q1 and q2. The distance between the charges is 'r', and the force of attraction/repulsion between them is 'F'. Then

$$F \propto q_1 q_2$$
$$\text{Or, } F \propto 1/r^2$$

$$F = k\, q_1 q_2 / r^2$$

where k is proportionality constant and equals to $1/4\pi\varepsilon_0$. Here, $\varepsilon_0$ is the epsilon naught and it signifies permittivity of a vacuum. The value of k comes $9 \times 10^9$ Nm$^2$/ C$^2$ when we take the S.I unit of value of $\varepsilon_0$ is $8.854 \times 10^{-12}$ C$^2$ N$^{-1}$ m$^{-2}$.

According to this theory, like charges repel each other and unlike charges attract each other. This means charges of same sign will push each other with repulsive forces while charges with opposite signs will pull each other with attractive force.

**The force between two charged bodies**

The force between two charged bodies was studied by Coulomb in 1785. Coulomb's law states that the force of attraction or repulsion between two point charges is directly proportional to the product of the charges and inversely proportional to the square of the distance between them. The direction of forces is along the line joining the two point charges.

Let q1 and q2 be two point charges placed in air or vacuum at a distance r apart (Fig. 1.3a). Then, according to Coulomb's law,

$$F \alpha \frac{q_1 q_2}{r^2} \text{ or } F = k\frac{q_1 q_2}{r^2}$$

where k is a constant of proportionality. In air or vacuum, k = $\frac{1}{4\pi\varepsilon_0}$, where $\varepsilon_0$ is the permittivity of free space (i.e., vacuum) and the value of $\varepsilon_0$ is $8.854 \times 10^{-12}$ C$^2$ N$^{-1}$ m$^{-2}$.

$$F = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2} \qquad (1)$$

$$\frac{1}{4\pi\varepsilon_0} = 9 \times 10^9 \text{ N m}^2 \text{ C}^{-2}$$

In the above equation, if $q_1 = q_2 = 1$C and r = 1m then,

$$F = (9 \times 10^9) \times \frac{1 \times 1}{1^2} = 9 \times 10^9 \text{ N}$$

One Coulomb is defined as the quantity of charge, which when placed at a distance of 1 meter in air or vacuum from an equal and similar charge, experiences a repulsive force of $9 \times 10^9$ N.

If the charges are situated in a medium of permittivity ε, then the magnitude of the force between them will be,

$$F_m = \frac{1}{4\pi\varepsilon}\frac{q_1 q_2}{r^2} \qquad (2)$$

Dividing equation (1) by (2)

$$\frac{F}{F_m} = \frac{\varepsilon}{\varepsilon_0} = \varepsilon_r$$

The ratio $\frac{\varepsilon}{\varepsilon_0} = \varepsilon_r$, is called the relative permittivity or dielectric constant of the medium. The value of $\varepsilon_r$ for air or vacuum is 1.

$$\therefore \varepsilon = \varepsilon_0 \varepsilon_r$$

Since $F_m = \frac{F}{\varepsilon_r}$, the force between two point charges depends on the nature of the medium in which the two charges are situated.

**Vector Form of Coulomb's Law**

The physical quantities are of two types namely scalars (with the only magnitude) and vectors (those quantities with magnitude and direction). Force is a vector quantity as it has both magnitude and direction. The Coulomb's law can be re-written in the form of vectors. Remember we denote the vector "F" as F, vector r as r and so on.

Let there be two charges $q_1$ and $q_2$, with position vectors $r_1$ and $r_2$ respectively. Now, since both the charges are of the same sign, there will be a repulsive force between them. Let the force on the q1 charge due to $q_2$ be $F_{12}$ and force on $q_2$ charge due to $q_1$ charge be $F_{21}$. The corresponding vector from $q_1$ to $q_2$ is $r_{21}$ vector.

$r_{21} = r_2 - r_1$

To denote the direction of a vector from position vector r1 to $r_2$, and from $r_2$ to $r_1$ as:

$$\hat{r}_{21} = \frac{r_{21}}{r_{21}} . \hat{r}_{12} = \frac{r_{12}}{r_{12}} . \hat{r}_{21} = \hat{r}_{12}$$

Now, the force on charge $q_2$ due to $q_1$, in vector form is:

$$F_{21} = \frac{1}{4\pi\varepsilon_0} \frac{q_1 q_2}{r_{21}^2} \hat{r}_{21}$$

The above equation is the vector form of Coulomb's Law.

**Limitations of Coulomb's Law**

Coulomb's Law is derived under certain assumptions and can't be used freely like other general formulas. The law is limited to following points:

- We can use the formula if the charges are static ( in rest position)

- The formula is easy to use while dealing with charges of regular and smooth shape, and it becomes too complex to deal with charges having irregular shapes
- The formula is only valid when the solvent molecules between the particle are sufficiently larger than both the charges

**Principle of Superposition**

The principle of superposition is to calculate the electric force experienced by a charge $q_1$ due to other charges $q_2$, $q_3$ ……. $q_n$. The total force on a given charge is the vector sum of the forces exerted on it due to all other charges.

The force on $q_1$ due to $q_2$

$$\overrightarrow{F_{12}} = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r_{21}^2}\widehat{r_{21}}$$

Similarly, force on $q_1$ due to $q_3$

$$\overrightarrow{F_{13}} = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_3}{r_{31}^2}\widehat{r_{31}}$$

The total force $F_1$ on the charge $q_1$ by all other charges is,

$$\overrightarrow{F_1} = \overrightarrow{F_{12}} + \overrightarrow{F_{13}} + \overrightarrow{F_{14}} \ldots\ldots + \overrightarrow{F_{1n}}$$

Therefore,

$$\overrightarrow{F_{13}} = \frac{1}{4\pi\varepsilon_0}\left(\frac{q_1 q_2}{r_{21}^2}\widehat{r_{21}} + \frac{q_1 q_3}{r_{31}^2}\widehat{r_{31}} + \cdots \frac{q_1 q_n}{r_{n1}^2}\widehat{r_{n1}}\right)$$

**Electric Field Intensity (E)**

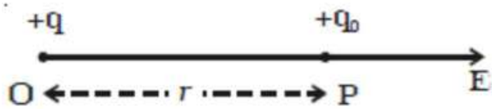
Electric field at a point is measured in terms of electric field intensity. Electric field intensity at a point, in an electric field is defined as the force experienced by a unit positive charge kept at that point.

It is a vector quantity. $|\vec{E}| = \frac{|\vec{F}|}{q_0}$ The unit of electric field intensity is N $C^{-1}$.

The electric field intensity is also referred as electric field strength or simply electric field. So, the force exerted by an electric field on a charge is $F = q_o E$.

**Electric field due to a point charge**

Let q be the point charge placed at O in air. A test charge $q_0$ is placed at P at a distance r from q. According to Coulomb's law, the force acting on $q_0$ due to q is



$\frac{1}{4\pi\varepsilon}\frac{qq_0}{r^2}$

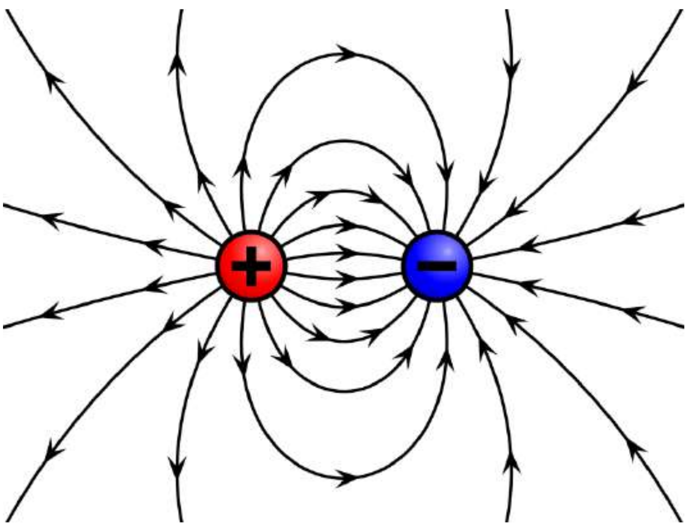
The electric field at a point P is, by definition, the force per unit test charge.

$E = \frac{F}{q_0} = \frac{1}{4\pi\varepsilon}\frac{q}{r^2}$

The direction of E is along the line joining O and P, pointing away from q, if q is positive and towards q, if q is negative.In vector notation $\hat{E} = \frac{1}{4\pi\varepsilon}\frac{q}{r^2}\hat{r}$, where $\hat{r}$ is a unit vector pointing away from q.

**Electric Field Lines**

An electric field can be used in the pictorial form to describe the overall intensity of the field around it. This pictorial representation is called the electric field lines. There are certain properties, rules, and applications of electric field lines. Electric Field Lines can be easily defined as a curve which shows the direction of an electric field when we draw a tangent at its point.




The concept of electric field was first proposed by Michael Faraday, in the 19th century. Faraday always thought of electric field lines as ones which can be used to describe and interpret the invisible electric field.

Instead of using complex vector diagram each time, electric field lines can be used to describe the electric field around a system of charges in an easier way.

The strength of electric fields is usually directly proportional to the lengths of electric field lines. Also, since the electric field is inversely proportional to the square of the distance, the electric field strength decreases, as we move away from the charge. The direction of arrows of field lines depicts the direction of the electric field, which is pointing outwards in case of positive charge and pointing inwards in case of a negative charge.

Further, the magnitude of an electric field is well described by the density of charges. The lines closer to the charge represent a strong electric field and the lines away from charge correspond to the weak electric field. This is because the strength of the electric field decreases as we move away from the charge.

**Properties of Electric Field Lines**

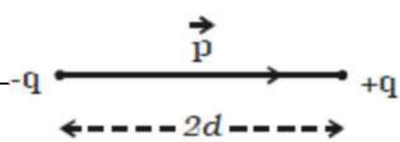
Electric field lines generally show the properties to account for nature of electric fields. Some general properties of these lines are as follows:

- Electric field lines start from a positive charge and end at a negative charge, in case of a single charge, electric field lines end at infinity
- In a charge-free region, electric field lines are continuous and smooth
- Two electric field lines never intersect or cross each other, as if they do, there will be two vectors depicting two directions of the same electric field, which is not possible
- These lines never form a closed loop. This is because an electric field is conservative in nature and hence the lines don’t form a closed loop.

**Electric dipole and electric dipole moment**

Two equal and opposite charges separated by a very small distance constitute an electric dipole.

Water, ammonia, carbon-dioxide and chloroform molecules are some examples of permanent electric dipoles. These molecules behave like

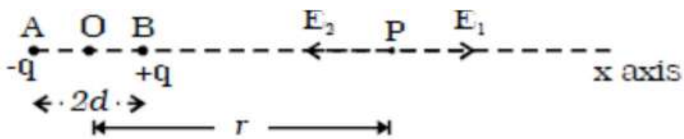


electric dipole, because the centers of positive and negative charge do not coincide and are separated by a small distance.

Two point charges +q and –q are kept at a distance 2d apart. The magnitude of the dipole moment is given by the product of the magnitude of the one of the charges and the distance between them.
∴ Electric dipole moment, p = q2d or 2qd.
It is a vector quantity and acts from –q to +q. The unit of dipole moment is C m.

**Electric field due to an electric dipole at a point on its axial line.**

AB is an electric dipole of two point charges –q and +q separated by a




small distance 2d. P is a point along the axial line of the dipole at a distance r from the midpoint O of the electric dipole.
The electric field at the point P due to +q placed at B is,

$E_1 = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r-d)^2}$ (along BP)

The electric field at the point P due to –q placed at A is,

$E_1 = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r+d)^2}$ (along PA)

$E_1$ and $E_2$ act in opposite directions.
Therefore, the magnitude of resultant electric field (E) acts in the direction of the vector with a greater magnitude. The resultant electric field at P is,

$E = E_1 + (-E_2)$

$E = \frac{1}{4\pi\varepsilon_0}\left[\frac{q}{(r-d)^2} - \frac{q}{(r+d)^2}\right]$ along BP

$E = \frac{q}{4\pi\varepsilon_0}\left[\frac{1}{(r-d)^2} - \frac{1}{(r+d)^2}\right]$ along BP

$$E = \frac{q}{4\pi\varepsilon_0}\left[\frac{4rd}{(r^2-d^2)^2}\right]$$

If the point P is far away from the dipole, then d <<r

$E = \frac{q}{4\pi\varepsilon_0}\frac{4rd}{r^4} = \frac{q}{4\pi\varepsilon_0}\frac{4d}{r^3}$

$\frac{q}{4\pi\varepsilon_0}\frac{4p}{r^3}$along BP [∵ Electric dipole moment p = q x 2d]

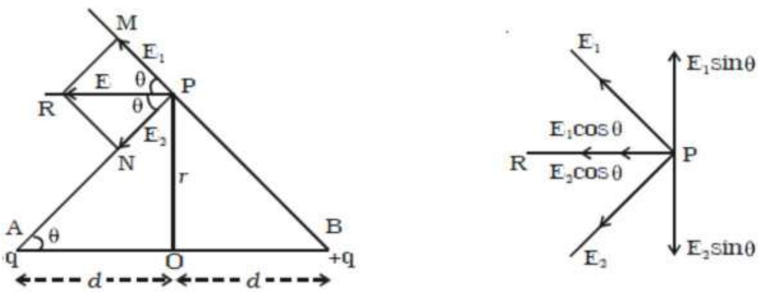
E acts in the direction of dipole moment.

**Electric field due to an electric dipole at a point on the equatorial line.**

Consider an electric dipole AB. Let 2d be the dipole distance and p be the dipole moment. P is a point on the equatorial line at a distance r from the midpoint O of the dipole (Fig a).

Electric field at a point P due to the charge +q of the dipole,

$E_1 = \frac{1}{4\pi\varepsilon_0}\frac{q}{BP^2}$along BP




*(a) Electric field at a point on equatorial line*

*(b) The components of the electric field*

$= \frac{1}{4\pi\varepsilon_0}\frac{q}{(r+d)^2}$along BP (∵$BP^2 = OP^2 + OB^2$)

Electric field ($E_2$) at a point P due to the charge –q of the dipole

$E_1 = \frac{1}{4\pi\varepsilon_0}\frac{q}{AP^2}$along PA

$= \frac{1}{4\pi\varepsilon_0}\frac{q}{(r+d)^2}$along PA

The magnitudes of $E_1$ and $E_2$ are equal. Resolving $E_1$ and $E_2$ into their horizontal and vertical components (Fig b), the vertical components $E_1$ sinθ and $E_2$ sinθ are equal and opposite, therefore they cancel each other.

The horizontal components $E_1$ cosθ and $E_2$ cosθ will get added along PR.

Resultant electric field at the point P due to the dipole is

E = $E_1$ cosθ + $E_2$ cosθ (along PR)

$= 2\ E_1\cos\theta\ (\because E1 = E2)$

$$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r+d)^2} \text{x } 2\cos\theta$$

$$\text{Cos}\theta = \frac{d}{\sqrt{r^2+d^2}}$$

$$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r+d)^2} \times \frac{2d}{(r^2+d^2)^{1/2}} = \frac{1}{4\pi\varepsilon_0}\frac{q2d}{(r^2+d^2)^{3/2}}$$

$$= \frac{1}{4\pi\varepsilon_0}\frac{P}{(r^2+d^2)^{3/2}} \quad (\because p = q2d)$$

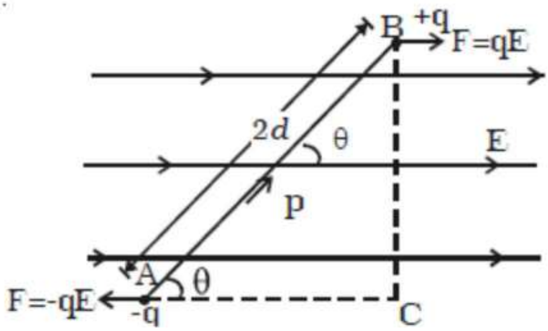
For a dipole, d is very small when compared to r

$$\therefore \frac{1}{4\pi\varepsilon_0}\frac{P}{r^3}$$

The direction of E is along PR, parallel to the axis of the dipole and directed opposite to the direction of dipole moment.

**Electric dipole in a uniform electric field**

Consider a dipole AB of dipole moment p placed at an angle θ in an uniform electric field E. The charge +q experience a force qE in the direction of the field. The charge –q experiences an equal force in the opposite direction. Thus the net force on the dipole is zero. The two equal and unlike




parallel forces are not passing through the same point, resulting in a torque on the dipole, which tends to set the dipole in the direction of the electric field.

The magnitude of torque is,

$\tau$ = One of the forces x perpendicular distance between the forces

= F x 2d sin $\theta$

= qE x 2d sin $\theta$ = pE sin $\theta$ ($\because$ q × 2d = P)

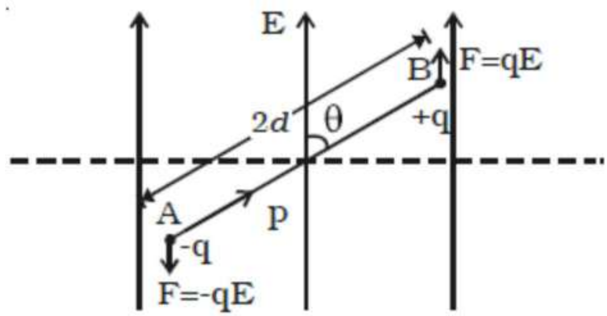
In vector notation, $\vec{\tau} = \vec{P} \times \vec{E}$

Note : If the dipole is placed in a non–uniform electric field at an angle θ, in addition to a torque, it also experiences a force.

**Electric potential energy of an electric dipole in an electric field.**

Electric potential energy of an electric dipole in an electrostatic field is the work done in rotating the dipole to the desired position in the field. When an electric dipole of dipole moment p is at an angle $\theta$ with the electric field E, the torque on the dipole is

$\tau$ = pE sin $\theta$




Work done in rotating the dipole through d θ,

dw = $\tau$ .d$\theta$

= pE sin $\theta$. d$\theta$

The total work done in rotating the dipole through an angle θ is

W = $\int dw$

W = pE $\int \sin\theta \, d\theta$ = –pE cos $\theta$

This work done is the potential energy (U) of the dipole.

∴ U = – pE cos $\theta$

When the dipole is aligned parallel to the field, $\theta$ = 0°

∴U = –pE

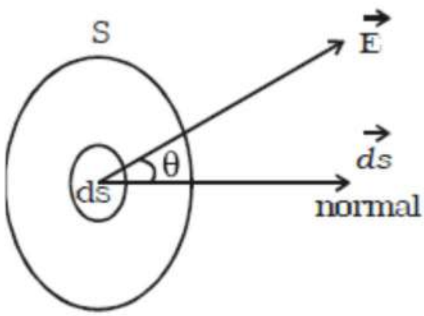
This shows that the dipole has a minimum potential energy when it is aligned with the field. A dipole in the electric field experiences a torque ($\vec{\tau} = \vec{P} \times \vec{E}$) which tends to align the dipole in the field direction, dissipating its potential energy in the form of heat to the surroundings.

**Gauss's law and its applications**

**Electric flux**

Consider a closed surface S in a non–uniform electric field (Fig 1.16). Consider a very small area $\overrightarrow{ds}$ on this surface. The direction of ds is drawn

normal to the surface outward. The electric field over ds is supposed to be a constant $\vec{E}$. $\vec{E}$ and $\overrightarrow{ds}$ make an angle $\theta$ with each other.
The electric flux is defined as the total number of electric lines of force, crossing through the given area. The electric flux dϕ through the area ds is,
$d\phi = E\, ds = E\, ds \cos\theta$




The total flux through the closed surface S is obtained by integrating the above equation over the surface.
$\phi = \oint d\varphi = \oint \vec{E}\,\overrightarrow{ds}$
The circle on the integral indicates that, the integration is to be taken over the closed surface. The electric flux is a scalar quantity. Its unit is N $m^2 C^{-1}$.

**Gauss's law**

The law relates the flux through any closed surface and the net charge enclosed within the surface. The law states that the total flux of the electric field E over any closed surface is equal to
$\frac{1}{\varepsilon_0}$ times the net charge enclosed by the surface.
$\phi = \frac{q}{\varepsilon_0}$
This closed imaginary surface is called Gaussian surface. Gauss's law tells us that the flux of E through a closed surface S depends only on the value of net charge inside the surface and not on the location of the charges. Charges outside the surface will not contribute to flux.

**Applications of Gauss's Law**

i) Field due to an infinite long straight charged wire

Consider a uniformly charged wire of infinite length having a constant linear charge density $\lambda$ (charge per unit length). Let P be a point at a distance r from the wire (Fig.) and E is the electric field at the point P. A cylinder of length l, radius r, closed at each end by plane caps normal to the axis is chosen as Gaussian surface. Consider very small area ds on the Gaussian surface.

By symmetry, the magnitude of the electric field will be the same at all points on the curved surface of the cylinder and directed radially outward. $\vec{E}$ $and$ $\overrightarrow{ds}$ are along the same direction.

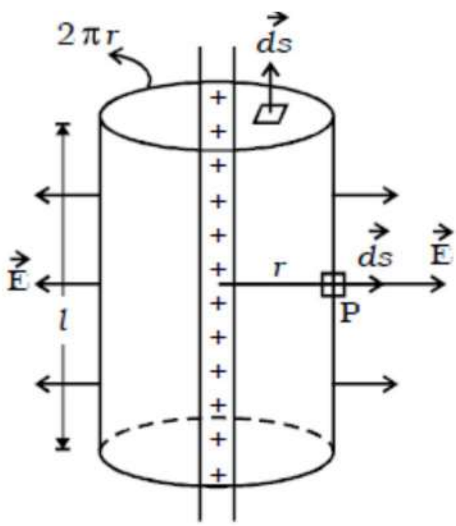
The electric flux (ϕ) through curved surface $= \oint \mathrm{E\ ds} \cos\theta$

$\phi = \oint \mathrm{E\ ds}$ [$\because \theta = 0; \cos\theta$]

= E ($2\pi$rl)

($\because$The surface area of the curved part is $2\pi$rl)

Since $\vec{E}$ $and$ $\overrightarrow{ds}$ are right angles to each other, the electric flux through the plane caps = 0




$\therefore$ Total flux through the Gaussian surface, ϕ = E. ($2\pi$rl)

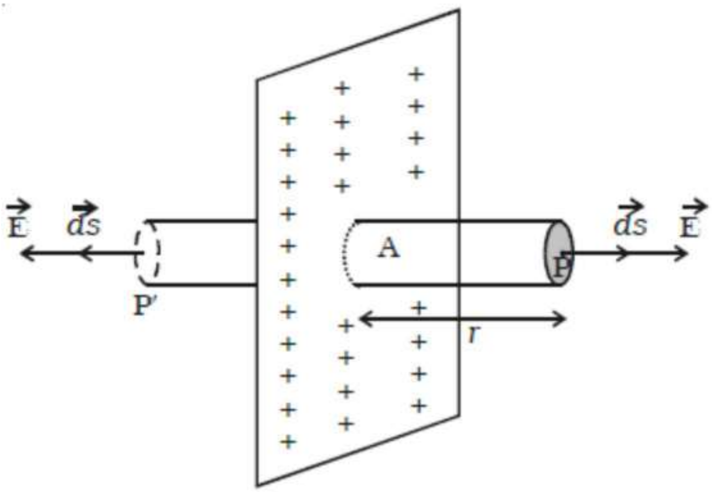
The net charge enclosed by Gaussian surface is, q = $\lambda l$

E. ($2\pi$rl) = $\frac{\lambda}{\varepsilon_0}$ or $\frac{\lambda}{2\pi\varepsilon_0 r}$

The direction of electric field E is radially outward, if line charge is positive and inward, if the line charge is negative.

**Electric field due to an infinite charged plane sheet**

Consider an infinite plane sheet of charge with surface charge density $\sigma$. Let P be a point at a distance r from the sheet (Fig) and E be the electric field at P. Consider a Gaussian surface in the form of cylinder of cross sectional area A and length 2r perpendicular to the sheet of charge.




By symmetry, the electric field is at right angles to the end caps and away from the plane. Its magnitude is the same at P and at the other cap at P′.

Therefore, the total flux through the closed surface is given by

$\phi = [\oint Eds]_P + [\oint Eds]_{P'}$ ($\because \theta = 0, \cos\theta = 1$)

$= EA + EA = 2EA$

If $\sigma$ is the charge per unit area in the plane sheet, then the net positive charge q within the Gaussian surface is, $q = \sigma A$

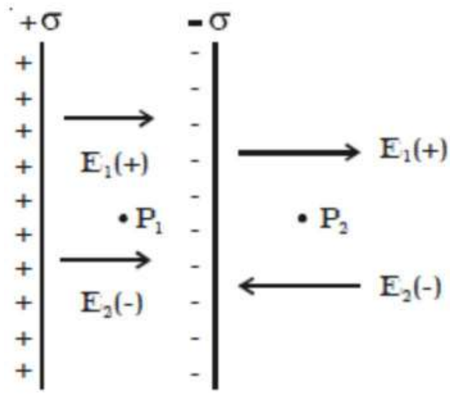
Using Gauss's law,

$2EA = \frac{\sigma A}{\varepsilon_0}$

$\because E = \frac{\sigma}{2\varepsilon_0}$

**Electric field due to two parallel charged sheets**

Consider two plane parallel infinite sheets with equal and opposite charge densities + $\sigma$and $-\sigma$ as shown in Fig. The magnitude of electric field on either side of a plane sheet of charge is E = $\frac{\sigma}{2\varepsilon_0}$and acts perpendicular to the sheet, directed outward (if the charge is positive) or inward (if the charge is negative).




(i) When the point $P_1$ is in between the sheets, the field due to two sheets will be equal in magnitude and in the same direction. The resultant field at $P_1$ is,

$$E = E_1 + E_2 = \frac{\sigma}{2\varepsilon_0} + \frac{\sigma}{2\varepsilon_0} = \frac{\sigma}{\varepsilon_0} \text{(towards the right)}$$

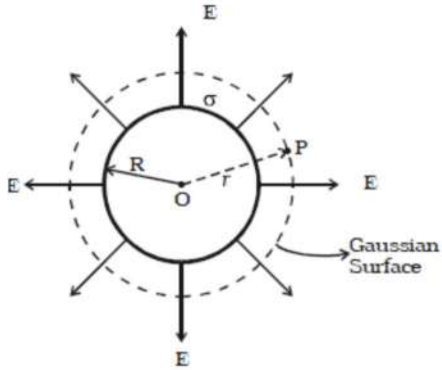
(ii) At a point $P_2$ outside the sheets, the electric field will be equal in magnitude and opposite in direction. The resultant field at $P_2$ is,

$$E = E_1 - E_2 = \frac{\sigma}{2\varepsilon_0} - \frac{\sigma}{2\varepsilon_0} = 0$$

**Electric field due to uniformly charged spherical shell**

Case (i) At a point outside the shell.

Consider a charged shell of radius R (Fig a). Let P be a point outside the shell, at a distance r from the center O. Let us construct a Gaussian surface with r as radius. The electric field E is normal to the surface.




The flux crossing the Gaussian sphere normally in an outward direction is,

$\phi = \int_s \vec{E} \cdot \overrightarrow{ds} = \int_s E.\, ds$ = E $(4\pi r^2)$

(since angle between E and ds is zero)

By Gauss's law, E. $(4\pi r^2) = \frac{q}{\varepsilon_0}$ or E = $\frac{1}{4\pi\varepsilon_0}\frac{q}{r^2}$

It can be seen from the equation that, the electric field at a point outside the shell will be the same as if the total charge on the shell is concentrated at its center.

Case (ii) At a point on the surface.

The electric field E for the points on the surface of charged spherical shell is,

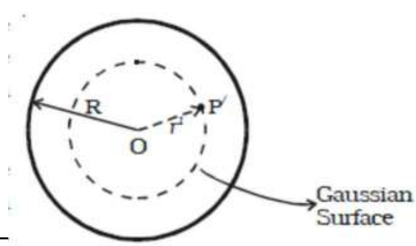
E = $\frac{1}{4\pi\varepsilon_0}\frac{q}{R^2}$ ($\because$ r = R)

Case (iii) At a point inside the shell.

Consider a point P′ inside the shell at a distance r′ from the center of the shell. Let us construct a Gaussian surface with radius r′. The total flux crossing the Gaussian sphere normally in an outward direction is

$\phi = \int_s \vec{E} \cdot \overrightarrow{ds} = \int_s E.\, ds$ = E $(4\pi r'^2)$

since there is no charge enclosed by the gaussian surface, according to Gauss's Law



E. $(4\pi r'^2) = \frac{q}{\varepsilon_0} = 0$

(i.e) the field due to a uniformly charged thin shell is zero at all points inside the shell.

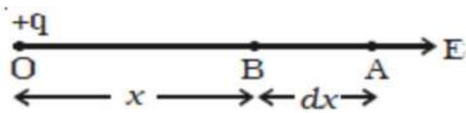
**Electric potential**

Let a charge +q be placed at a point O (Fig 1.11). A and B are two points, in the electric field. When a unit positive charge is moved from A to B against the electric force, work is done. This work is the potential difference between these two points. i.e., $dV = W_{A \rightarrow B}$.

The potential difference between two points in an electric field is defined as the amount of work done in moving a unit positive charge from one point to the other against the electric force.

The unit of potential difference is volt.

The potential difference between two points is 1 volt if 1 joule of work is




done in moving 1 Coulomb of charge from one point to another against the electric force. The electric potential in an electric field at a point is defined as the amount of work done in moving a unit positive charge from infinity to that point against the electric forces.

**Relation between electric field and potential**

Let the small distance between A and B be dx. Work done in moving a unit positive charge from A to B is dV = E. dx.

The work has to be done against the force of repulsion in moving a unit positive charge towards the charge +q. Hence,

dV = – E. dx

$E = -\frac{dV}{dX}$

The change of potential with distance is known as potential gradient; hence the electric field is equal to the negative gradient of potential.

The negative sign indicates that the potential decreases in the direction of electric field. The unit of electric intensity can also be expressed as $Vm^{-1}$.

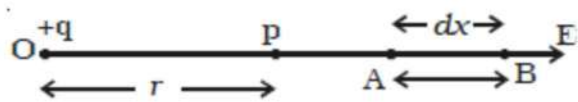
**Electric potential at a point due to a point charge**

Let +q be an isolated point charge situated in air at O. P is a point at a distance r from +q. Consider two points A and B at distances x and x + dx from the point O.

The potential difference between A and B is,

dV = –E dx

The force experienced by a unit positive charge placed at A is




$E = \frac{1}{4\pi\varepsilon_0}\frac{q}{X^2}$

$\therefore dV = \mathrm{E} = \frac{1}{4\pi\varepsilon_0}\frac{q}{X^2}dx$

The negative sign indicates that the work is done against the electric force. The electric potential at the point P due to the charge +q is the total work done in moving a unit positive charge from infinity to that point.

$v = -\int_{\infty}^{r}\frac{q}{4\pi\varepsilon_0 X^2}.dx = \frac{q}{4\pi\varepsilon_0 r}$

**Electric potential at a point due to an electric dipole**

Two charges –q at A and +q at B separated by a small distance 2d constitute an electric dipole and its dipole moment is p. Let P be the point at a distance r from the midpoint of the dipole O and θ be the angle between PO and the axis of the dipole OB. Let r1 and r2 be the distances of the point P from +q and –q charges respectively.

Potential at P due to charge (+q) $= \frac{1}{4\pi\varepsilon_0}\frac{q}{r_1}$

Potential at P due to charge (–q) $= \frac{1}{4\pi\varepsilon_0}\left(-\frac{q}{r_2}\right)$

Total potential at P due to dipole is, $V = \frac{1}{4\pi\varepsilon_0}\frac{q}{r_1} - \frac{1}{4\pi\varepsilon_0}\frac{q}{r_2}$

$V = \frac{1}{4\pi\varepsilon_0}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$ …. (1)

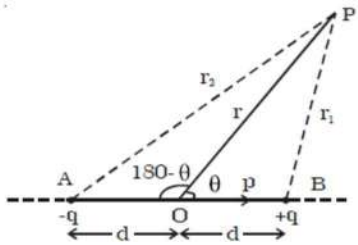
Applying cosine law,

$r_1^2 = r^2 + d^2 - 2rd\cos\theta$

$r_1^2 = r^2\left(1 - \frac{2d}{r}\cos\theta\right)^{\frac{1}{2}}$

$\frac{1}{r_1} = \frac{1}{r}\left(1 - \frac{2d}{r}\cos\theta\right)^{\frac{1}{2}}$

Using the Binomial theorem and neglecting higher powers,




$\frac{1}{r_1}=\frac{1}{r}\left(1+\frac{2d}{r}\cos\theta\right)$ .... (2)

Similarly,

$r_1^2 = r^2 + d^2 - 2rd\cos(180-\theta)$

Or $r_1^2 = r^2 + d^2 + 2rd\cos\theta$

$r_2 = r\left(1+\frac{2d}{r}\cos\theta\right)^{\frac{1}{2}}$ ($\because \frac{d^2}{r^2}$ is negligible)

$or\ \frac{1}{r_2}=\frac{1}{r}\left(1+\frac{2d}{r}\cos\theta\right)^{-\frac{1}{2}}$

Using the Binomial theorem and neglecting higher powers,

$\frac{1}{r_2}=\frac{1}{r}\left(1-\frac{d}{r}\cos\theta\right)$ .... (3)

Substituting equation (2) and (3) in equation (1) and simplifying

$V=\frac{q}{4\pi\varepsilon_0}\frac{1}{r}\left(1+\frac{d}{r}\cos\theta-1+\frac{d}{r}\cos\theta\right)$

$\therefore V=\frac{q2d\cos\theta}{4\pi\varepsilon_0 r^2}=\frac{q}{4\pi\varepsilon_0}\frac{P.\cos\theta}{r^2}$

Special cases:

1. When the point P lies on the axial line of the dipole on the side of +q, then $\theta = 0$

$\therefore V=\frac{P}{4\pi\varepsilon_0 r^2}$

2. When the point P lies on the axial line of the dipole on the side of –q, then $\theta = 180$

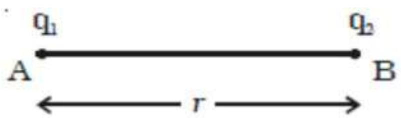
$\therefore V=-\frac{P}{4\pi\varepsilon_0 r^2}$

3. When the point P lies on the equatorial line of the dipole, then, $\theta = 90°$,

$\therefore V = 0$

**Electric potential energy**

The electric potential energy of two point charges is equal to the work done to assemble the charges or work done in bringing each charge or work done in bringing a charge from infinite distance. Let us consider a




point charge $q_1$, placed at A.

The potential at a point B at a distance r from the charge $q_1$ is

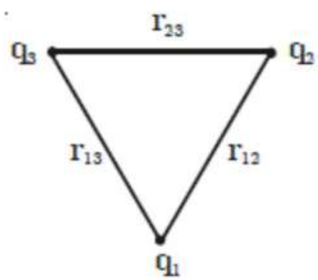
$V = \frac{q_1}{4\pi\varepsilon_0 r}$

Another point charge $q_2$ is brought from infinity to the point B. Now the work done on the charge $q_2$ is stored as electrostatic potential energy (U) in the system of charges $q_1$ and $q_2$.

$\therefore$ Work done, $w = Vq_2$

Potential energy (U) $= \frac{q_1 q_2}{4\pi\varepsilon_0 r}$

Keeping $q_2$ at B, if the charge $q_1$ is imagined to be brought from infinity to the point A, the same amount of work is done.




Also, if both the charges $q_1$ and $q_2$ are brought from infinity, to points A and B respectively, separated by a distance r, then potential energy of the system is the same as the previous cases.

For a system containing more than two charges, the potential energy (U) is given by

$U = \frac{1}{4\pi\varepsilon_0}\left[\frac{q_1 q_2}{r_{12}} + \frac{q_1 q_3}{r_{13}} + \frac{q_2 q_3}{r_{23}}\right]$

**Equipotential Surface**

If all the points of a surface are at the same electric potential, then the surface is called an equipotential surface.

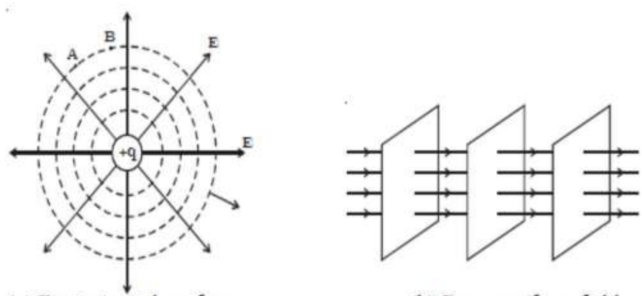
(i) In case of an isolated point charge, all points equidistant from the charge are at same potential. Thus, equipotential surfaces in this case will be a series of concentric spheres with the point charge as their center (Fig a). The potential will however be different for different spheres.

If the charge is to be moved between any two points on an equipotential surface through any path, the work done is zero. This is because the potential difference between two points A and B is defined as $V_B - V_A = \frac{W_{AB}}{q}$ . If $V_A = V_B$ then $W_{AB} = 0$. Hence the electric field lines must be normal to an equipotential surface.

(ii) In case of uniform field, equipotential surfaces are the parallel planes with their surfaces perpendicular to the lines of force as shown in Fig b.

**Electrostatic shielding**

It is the process of isolating a certain region of space from external field. It is based on the fact that electric field inside a conductor is zero.

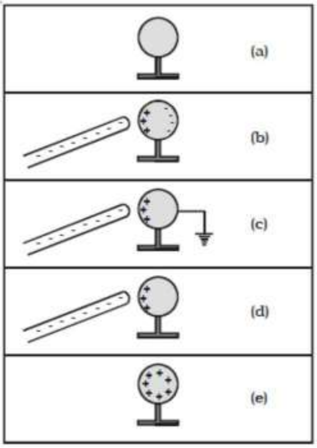



(a) Equipotential surface (spherical) Fig 15 (b) For a uniform field (plane)

During a thunder accompanied by lightning, it is safer to sit inside a bus than in open ground or under a tree. The metal body of the bus provides electrostatic shielding, where the electric field is zero. During lightning the electric discharge passes through the body of the bus.

**Electrostatic induction**

It is possible to obtain charges without any contact with another charge. They are known as induced charges and the phenomenon of producing




induced charges is known as electrostatic induction. It is used in electrostatic machines like Van de Graaff generator and capacitors.

Fig. shows the steps involved in charging a metal sphere by induction.

(a) There is an uncharged metallic sphere on an insulating stand.

(b) When a negatively charged plastic rod is brought close to the sphere, the free electrons move away due to repulsion and start pulling up at the farther end. The near end becomes positively charged due to deficit of electrons. This process of charge distribution stops when the net force on the free electron inside the metal is zero (this process happens very fast).

(c) When the sphere is grounded, the negative charge flows to the ground. The positive charge at the near end remains held due to attractive forces.

(d) When the sphere is removed from the ground, the positive charge continues to be held at the near end.

(e) When the plastic rod is removed, the positive charge spreads uniformly over the sphere.

**Capacitance of a conductor**

When a charge q is given to an isolated conductor, its potential will change. The change in potential depends on the size and shape of the conductor. The potential of a conductor changes by V, due to the charge q given to the conductor.

q α V or q = CV

i.e. C = q/V

Here C is called as capacitance of the conductor. The capacitance of a conductor is defined as the ratio of the charge given to the conductor to the potential developed in the conductor.

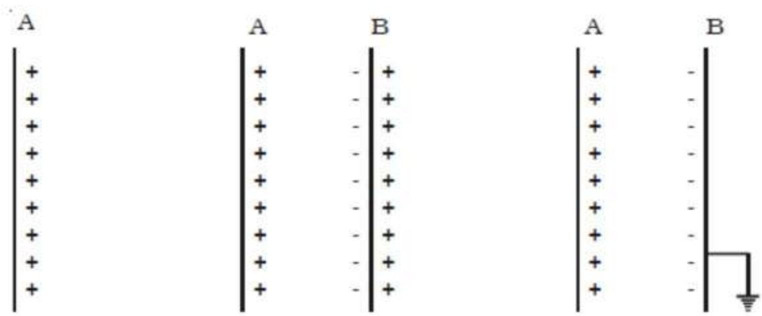
The unit of capacitance is farad. A conductor has a capacitance of one farad, if a charge of 1 coulomb given to it, raises its potential by 1 volt.

The practical units of capacitance are μF and pF.

**Principle of a capacitor**

Consider an insulated conductor (Plate A) with a positive charge 'q' having potential V (Fig a). The capacitance of A is C = q/V.

When another insulated metal plate B is brought near A, negative charges are induced on the side of B near A. An equal amount of positive charge is induced on the other side of B (Fig b). The negative

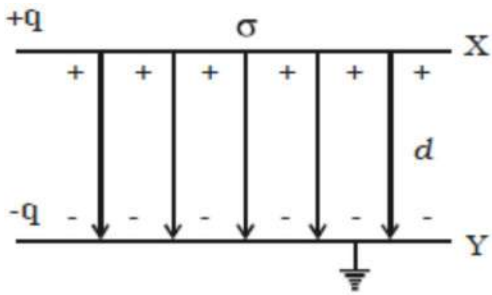


(a) (b) (c)

charge in B decreases the potential of A. The positive charge in B increases the potential of A. But the negative charge on B is nearer to A than the positive charge on B. So the net effect is that, the potential of A decreases. Thus the capacitance of A is increased. If the plate B is earthed, positive charges get neutralized (Fig c). Then the potential of A decreases further. Thus the capacitance of A is considerably increased. The capacitance depends on the geometry of the conductors and nature of the medium. A capacitor is a device for storing electric charges.

**Capacitance of a parallel plate capacitor**

The parallel plate capacitor consists of two parallel metal plates X and Y each of area A, separated by a distance d, having a surface charge




density$\sigma$. The medium between the plates is air. A charge +q is given to the plate X. It induces a charge –q on the upper surface of earthed

plate Y. When the plates are very close to each other, the field is confined to the region between them. The electric lines of force starting from plate X and ending at the plate Y are parallel to each other and perpendicular to the plates. By the application of Gauss's law, electric field at a point between the two plates is,

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

Potential difference between the plates X and Y is

$V = \int_d^0 -E\,dr = \int_d^0 -\frac{\sigma}{\varepsilon_0}\,dr = \frac{\sigma d}{\varepsilon_0}$

The capacitance (C) of the parallel plate capacitor

$C = \frac{q}{V} = \frac{\sigma A}{\sigma d/\varepsilon_0} = \frac{\varepsilon_0 A}{d}$ [since, $\sigma = \frac{q}{A}$]

$\therefore C = \frac{\varepsilon_0 A}{d}$

**Dielectrics and polarisation**

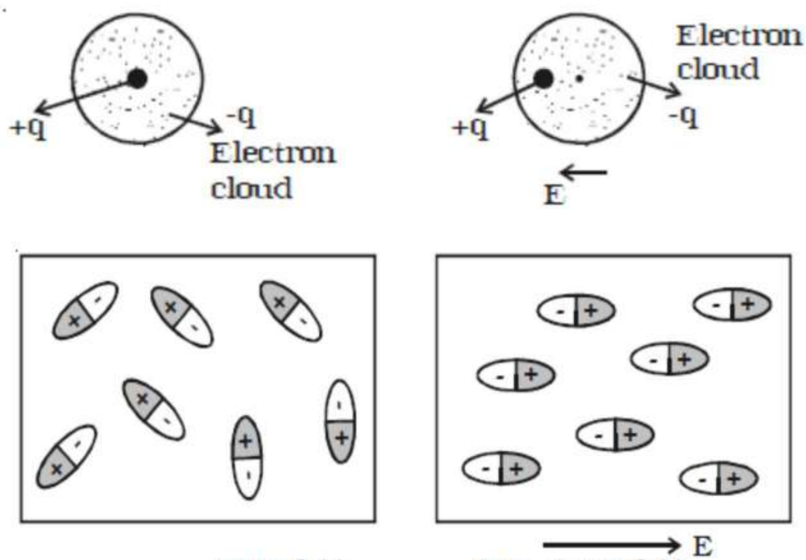
Dielectrics

A dielectric is an insulating material in which all the electrons are tightly bound to the nucleus of the atom. There are no free electrons to carry current. Ebonite, mica and oil are few examples of dielectrics. The electrons are not free to move under the influence of an external field.

**Polarisation**

A nonpolar molecule is one in which the centre of gravity of the positive charges (protons) coincide with the center of gravity of the negative charges (electrons). Example: $O_2$, $N_2$, $H_2$. The nonpolar molecules do not have a permanent dipole moment. If a non-polar dielectric is placed in an electric field, the center of charges gets displaced. The molecules are then said to be polarized and are called induced dipoles. They acquire induced dipole moment p in the direction of electric field (Fig).

A polar molecule is one in which the centre of gravity of the positive charges is separated from the centre of gravity of the negative charges




(a) No field (b) In electric field

by a finite distance. Examples: $N_2O$, $H_2O$, HCl, $NH_3$. They have a permanent dipole moment. In the absence of an external field, the dipole moments of polar molecules orient themselves in random directions. Hence no net dipole moment is observed in the dielectric. When an electric field is applied, the dipoles orient themselves in the direction of electric field. Hence a net dipole moment is produced.

The alignment of the dipole moments of the permanent or induced dipoles in the direction of applied electric field is called polarisation or electric polarisation. The magnitude of the induced dipole moment p is directly proportional to the external electric field E.

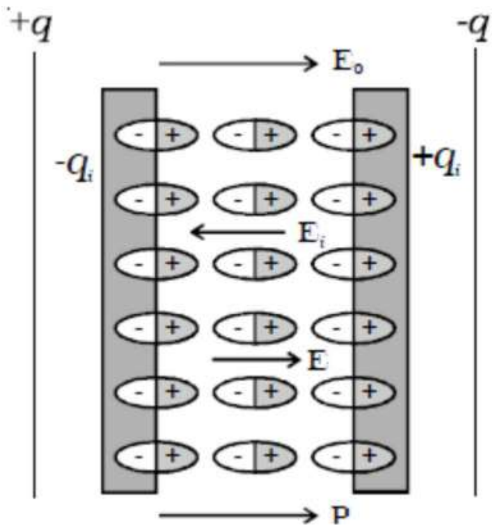
$\therefore p \propto E$ or $p = \alpha E$, where α is the constant of proportionality and is called molecular polarisability.

**Polarisation of dielectric material**

Consider a parallel plate capacitor with +q and –q charges. Let $E_0$ be the electric field between the plates in air. If a dielectric slab is introduced in the space between them, the dielectric slab gets polarised. Suppose $+q_i$ and $-q_i$ be the induced surface charges on the face of dielectric opposite to the plates of capacitor (Fig). These induced charges produce their own field $E_i$ which opposes the electric field $E_o$. So, the resultant field, $E < E_o$. But the direction of E is in the direction of $E_o$.

$\therefore E = Eo + (-Ei)$

($\because E_i$ is opposite to the direction of $E_o$)




**Capacitance of a parallel plate capacitor with a dielectric medium.**

Consider a parallel plate capacitor having two conducting plates X and Y each of area A, separated by a distance d apart. X is given a positive charge so that the surface charge density on it is $\sigma$ and Y is earthed. Let a dielectric slab of thick-ness t and relative permittivity $\varepsilon_r$ be introduced between the plates (Fig).

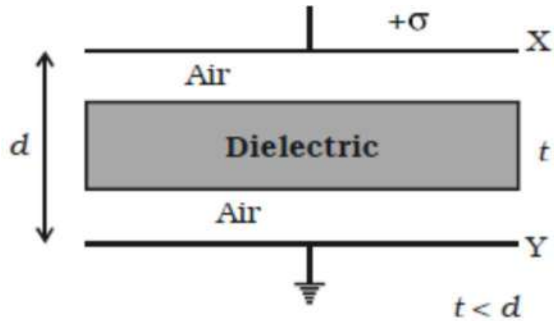
Thickness of dielectric slab = t

Thickness of air gap = (d−t)

Electric field at any point in the air between the plates,

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

Electric field at any point, in the dielectric slab E′ = $\frac{\sigma}{\varepsilon_0 \varepsilon_r}$




The total potential difference between the plates, is the work done in crossing unit positive charge from one plate to another in the field E over a distance (d-t) and in the field E′ over a distance t, then

V = E (d-t) + E′t

$= \frac{\sigma}{\varepsilon_0}(d-t) + \frac{\sigma t}{\varepsilon_0 \varepsilon_r}$

$= \frac{\sigma}{\varepsilon_0}\left[(d-t) + \frac{t}{\varepsilon_r}\right]$

The charge on the plate X, q = $\sigma$A

Hence the capacitance of the capacitor is,

$C = \frac{q}{V} = \frac{\sigma A}{\frac{\sigma}{\varepsilon_0}\left[(d-t) + \frac{t}{\varepsilon_r}\right]} = \frac{\varepsilon_0 A}{(d-t) + \frac{t}{\varepsilon_r}}$

**Effect of dielectric**

In capacitors, the region between the two plates is filled with dielectric like mica or oil.

The capacitance of the air filled capacitor, $C = \frac{\varepsilon_0 A}{d}$

The capacitance of the dielectric filled capacitor, $C' = \frac{\varepsilon_r \varepsilon_0 A}{d}$

$\therefore \frac{C'}{C} = \varepsilon_r$ or $C' = \varepsilon_r C$

since, $\varepsilon_r > 1$ for any dielectric medium other than air, the capacitance increases, when dielectric is placed.

**Applications of capacitors.**

(i) They are used in the ignition system of automobile engines to eliminate sparking.

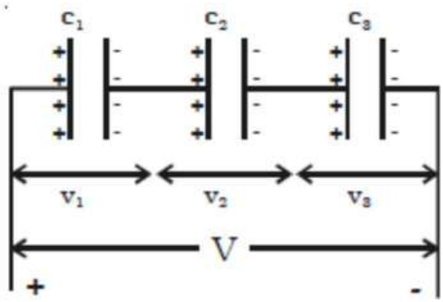
(ii) They are used to reduce voltage fluctuations in power supplies and to increase the efficiency of power transmission.

(iii) Capacitors are used to generate electromagnetic oscillations and in tuning the radio circuits.

**Capacitors in series and parallel**

(i) Capacitors in series

Consider three capacitors of capacitance $C_1$, $C_2$ and $C_3$ connected in series (Fig). Let V be the potential difference applied across the series combination. Each capacitor carries the same amount of charge q. Let $V_1$, $V_2$, $V_3$ be the potential difference across the capacitors $C_1$, $C_2$, $C_3$




respectively. Thus $V = V_1 + V_2 + V_3$

The potential difference across each capacitor is,

$V_1 = \frac{q}{c_1}$; $V_2 = \frac{q}{c_2}$; $V_3 = \frac{q}{c_3}$

$V = \frac{q}{c_1} + \frac{q}{c_2} + \frac{q}{c_3} = q\left[\frac{1}{c_1} + \frac{1}{c_2} + \frac{1}{c_3}\right]$

If $C_S$ be the effective capacitance of the series combination, it should acquire a charge q when a voltage V is applied across it.

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

$$\frac{q}{c_s} = \frac{q}{c_1} + \frac{q}{c_2} + \frac{q}{c_3}$$

$$\therefore \frac{1}{c_s} = \frac{1}{c_1} + \frac{1}{c_2} + \frac{1}{c_3}$$

when a number of capacitors are connected in series, the reciprocal of the effective capacitance is equal to the sum of reciprocal of the capacitance of the individual capacitors.

(ii) Capacitors in parallel

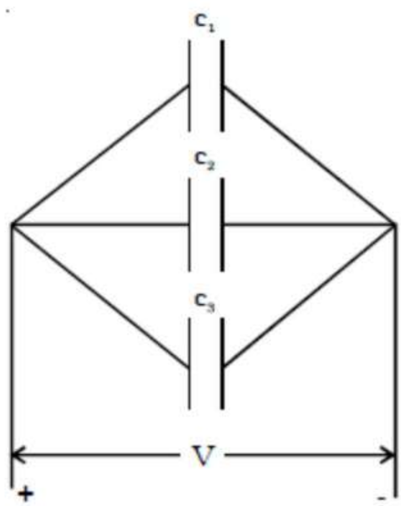
Consider three capacitors of capacitances $C_1$, $C_2$ and $C_3$ connected in parallel (Fig). Let this parallel combination be connected to a potential difference V. The potential difference across each capacitor is the same. The charges on the three capacitors are,

$q_1 = C_1V$, $q_2 = C_2 V$, $q_3 = C_3V$.

The total charge in the system of capacitors is

$q = q_1 + q_2 + q_3$

$q = C_1V + C_2 V + C_3V$




But $q = C_p.V$ where $C_p$ is the effective capacitance of the system

$\therefore C_pV = V (C_1 + C_2 + C_3)$

$\therefore C_P = C_1 + C_2 + C_3$

Hence the effective capacitance of the capacitors connected in parallel is the sum of the capacitances of the individual capacitors.

**Energy stored in a capacitor**

The capacitor is a charge storage device. Work has to be done to store the charges in a capacitor. This work done is stored as electrostatic potential energy in the capacitor.

Let q be the charge and V be the potential difference between the plates of the capacitor. If dq is the additional charge given to the plate, then work done is, $dw = Vdq$

$dw = \frac{q}{c}dq \qquad \left(\because V = \frac{q}{c}\right)$

Total work done to charge a capacitor is w = $\int dw = \int_0^q \frac{q}{c} dq = \frac{1}{2}\frac{q^2}{C}$

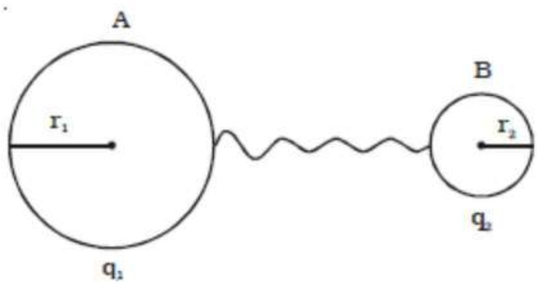
This work done is stored as electrostatic potential energy (U) in the capacitor.

U = $\frac{1}{2}\frac{q^2}{C} = \frac{1}{2}CV^2$ ($\therefore$ q = CV )

This energy is recovered if the capacitor is allowed to discharge.

**Distribution of charges on a conductor and action of points**

Let us consider two conducting spheres A and B of radii $r_1$ and $r_2$respectively connected to each other by a conducting wire (Fig). Let r1 be greater than $r_2$. A charge given to the system is distributed as $q_1$ and $q_2$ on the surface of the spheres A and B. Let $\sigma_1$, $\sigma_2$ be the charge




densities on the sphere A and B.

The potential at A,

$V_1 = \frac{q_1}{4\pi\varepsilon_0 r_1}$

The potential at B,

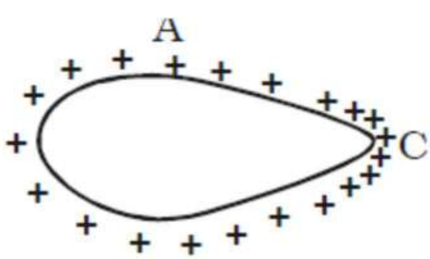
$V_2 = \frac{q_1}{4\pi\varepsilon_0 r_2}$

Since they are connected, their potentials are equal

$\frac{q_1}{4\pi\varepsilon_0 r_1} = \frac{q_1}{4\pi\varepsilon_0 r_2}$ $[\because q_1 = 4\pi r_1^2 \sigma_1 and\ q_2 = 4\pi r_2^2 \sigma_2]$

$\sigma_1 r_1 = \sigma_2 r_2$

i.e., $\sigma r$ is a constant. From the above equation it is seen that, smaller the radius, larger is the charge density.



In case of conductor, shaped as in Fig. the distribution is not uniform. The charges accumulate to a maximum at the pointed end where the curvature is maximum or the radius is minimum. It is found experimentally that a charged conductor with sharp points on its surface, loses its charge rapidly. The reason is that the air molecules which come in contact with the sharp points become ionized. The positive ions are repelled and the negative ions are attracted by the sharp points and the charge in them is therefore reduced. Thus, the leakage of electric charges from the sharp points on the charged conductor is known as action of points or corona discharge. This principle is made use of in the electrostatic machines for collecting charges and in lightning arresters (conductors).

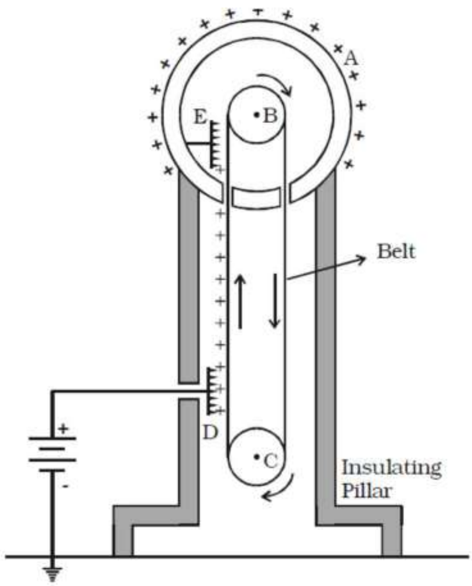
**Lightning conductor**

This is a simple device used to protect tall buildings from the lightning. It consists of a long thick copper rod passing through the building to ground. The lower end of the rod is connected to a copper plate buried deeply into the ground. A metal plate with number of spikes is connected to the top end of the copper rod and kept at the top of the building. When a negatively charged cloud passes over the building, positive charge will be induced on the pointed conductor. The positively charged sharp points will ionize the air in the vicinity. This will partly neutralize the negative charge of the cloud, thereby lowering the potential of the cloud. The negative charges that are attracted to the conductor travels down to the earth. Thereby preventing the lightning stroke from the damage of the building.

**Van de Graaff Generator**

In 1929, Robert J. Van de Graaff designed an electrostatic machine which produces large electrostatic potential difference of the order of 107 V. The working of Van de Graaff generator is based on the principle of electrostatic induction and action of points. A hollow metallic sphere A is mounted on insulating pillars as shown in the Fig. A pulley B is mounted at the centre of the sphere and another pulley C is mounted near the bottom. A belt made of silk moves over the pulleys. The pulley C is driven continuously by an electric motor. Two comb–shaped conductors D and E having number of needles are mounted near the pulleys. The comb D is maintained at a positive potential of the order of 104 volt by a power supply. The upper comb E is connected to the inner side of the hollow metal sphere. Because of the high electric field near the comb D, the air gets ionised due to action of points, the negative charges in air move towards the needles and positive charges are repelled on towards the belt. This positive charge stick to the belt moves up and reaches near the comb E.

Because of the high electric field near the comb D, the air gets ionised due to action of points, the negative charges in air move towards the




needles and positive charges are repelled on towards the belt. This positive charge stick to the belt, move up and reaches near the comb E. As a result of electrostatic induction, the comb E acquires negative charge and the sphere acquires positive charge. The acquired positive charge is distributed on the outer surface of the sphere. The high electric field at the comb E ionises the air. Hence, negative charges are repelled to the belt, neutralises the positive charge on the belt before the belt passes over the pulley. Hence the descending belt will be left uncharged. Thus the machine, continuously transfers the positive charge to the sphere. As a result, the potential of the sphere keeps increasing till it attains a limiting value (maximum). After this stage no more charge can be placed on the sphere, it starts leaking to the surrounding due to ionisation of the air. The leakage of charge from the sphere can be reduced by enclosing it in a gas filled steel chamber at a very high pressure. The high voltage produced in this generator can be used to accelerate positive ions (protons, deuterons) for the purpose of nuclear disintegration.

## Unit III
## CURRENT ELECTRICITY

The branch of Physics which deals with the study of motion of electric charges is called current electricity. In an uncharged metallic conductor at rest, some (not all) electrons are continually moving randomly through the conductor because they are very loosely attached to the nuclei. The thermodynamic internal energy of the material is sufficient to liberate the outer electrons from individual atoms, enabling the electrons to travel through the material. But the net flow
of charge at any point is zero. Hence, there is zero current. These are termed as free electrons. The external energy necessary to drive the free electrons in a definite direction is called electromotive force (emf). The emf is not a force, but it is the work done in moving a unit charge from one end to the other. The flow of free electrons in a conductor constitutes electric current.

**Electric current**

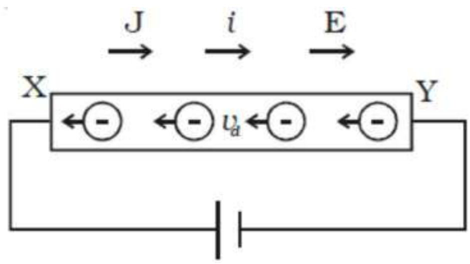
The current is defined as the rate of flow of charges across any cross sectional area of a conductor. If a net charge $q$ passes through any cross section of a conductor in time $t$, then the current $I = q / t$, where $q$ is in coulomb and $t$ is in second. The current I is expressed in ampere. If the rate of flow of charge is not uniform, the current varies with time and the instantaneous value of current $i$ is given by,

$i = dq/dt$

Current is a scalar quantity. The direction of conventional current is taken as the direction of flow of positive charges or opposite to the direction of flow of electrons.

**Drift velocity and mobility**

Consider a conductor XY connected to a battery. A steady electric field E is established in the conductor in the direction X to Y. In the absence of an electric field, the free electrons in the conductor move randomly in all possible directions.




They do not produce current. But, as soon as an electric field is applied, the free electrons at the end Y experience a force F = eE in a direction opposite to the electric field. The electrons are accelerated and in the process they collide with each other and with the positive ions in the conductor.

Thus due to collisions, a backward force acts on the electrons and they are slowly drifted with a constant average drift velocity $v_d$ in a direction opposite to electric field. Drift velocity is defined as the velocity with which free electrons get drifted towards the positive terminal, when an electric field is applied.

If $\tau$ is the average time between two successive collisions and the acceleration experienced by the electron be a, then the drift velocity is given by,

$$v_d = a\tau$$

The force experienced by the electron of mass m is

F = ma

Hence a = $\frac{eE}{m}$

$\therefore v_d = \frac{eE}{m}\tau = \mu E$

Where, $\mu = \frac{e\tau}{m}$ is the mobility and is defined as the drift velocity acquired per unit electric field. It takes the unit $m^2V^{-1}s^{-1}$. The drift velocity of electrons is proportional to the electric field intensity. It is very small and is of the order of 0.1 cm $s^{-1}$.

**Current density**

Current density at a point is defined as the quantity of charge passing per unit time through unit area, taken perpendicular to the direction of flow of charge at that point.

The current density J for a current I flowing across a conductor having an area of cross section A is

$$J = \frac{(q/t)}{A} = \frac{I}{A}$$

Current density is a vector quantity. It is expressed in A $m^{-2}$

**Relation between current and drift velocity**

Consider a conductor XY of length L and area of cross section A. An electric field E is applied between its ends. Let n be the number of free electrons per unit volume. The free electrons move towards the left with a constant drift velocity $v_d$.

The number of conduction electrons in the conductor = nAL

The charge of an electron = e

The total charge passing through the conductor q = (nAL) e

The time in which the charges pass through the conductor,t =$\frac{L}{v_d}$

The current flowing through the conductor, I = $\frac{q}{t} = \frac{(nAL)_e}{(L/v_d)}$

$I = NAev_d$ $\quad\quad\quad\quad\quad\quad\quad\quad$ … (1)

The current flowing through a conductor is directly proportional to the drift velocity.

From equation (1), $\frac{I}{A} = nev_d$

J = $nev_d$ $\quad\quad$ $\left[\because J = \frac{I}{A}, current\ density\right]$

**Ohm's law**

George Simon Ohm established the relationship between potential difference and current, which is known as Ohm's law. The current flowing through a conductor is,

$I = NAev_d$

But $v_d = \frac{eE}{m}.\tau$

$\therefore I = nAe\frac{eE}{m}.\tau$

$I = \frac{nAe^2}{mL}$ $\quad\quad$ $\left[\because E = \frac{V}{L}\right]$

where V is the potential difference. The quantity $\frac{mL}{nAe^2}$ is a constant for a given conductor, called electrical resistance (R).

$$\therefore I\ \alpha\ V$$

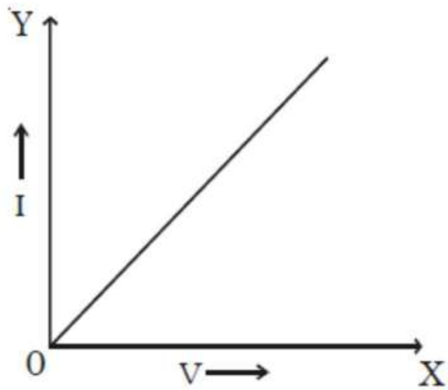
The law states that, at a constant temperature, the steady current flowing through a conductor is directly proportional to the potential difference between the two ends of the conductor.

(i.e) $I\ \alpha$ V or I = $\frac{1}{R}$V

$V = IR$ or $R = \frac{V}{I}$

Resistance of a conductor is defined as the ratio of potential difference across the conductor to the current flowing through it. The unit of resistance is ohm (Ω)

The reciprocal of resistance is conductance. Its unit is mho ($\Omega^{-1}$). Since, potential difference V is proportional to the current I, the graph (Fig) between V and I is a straight line for a conductor. Ohm's law holds well only when a steady current flows through a conductor.




**Electrical Resistivity and Conductivity**

The resistance of a conductor R is directly proportional to the length of the conductor l and is inversely proportional to its area of cross section A.

$$R \; \alpha \; \frac{1}{A} \; or \; R = \frac{\rho l}{A}$$

ρ is called specific resistance or electrical resistivity of the material of the conductor.

If l = l m, A = l $m^2$, then ρ = R

The electrical resistivity of a material is defined as the resistance offered to current flow by a conductor of unit length having unit area of cross section. The unit of $\rho$ is ohm-m(Ω m). It is a constant for a particular material.

The reciprocal of electrical resistivity, is called electrical conductivity, $\sigma = \frac{1}{\rho}$

The unit of conductivity is mho $m^{-1}$ ($\Omega^{-1}$ $m^{-1}$)

**Classification of materials in terms of resistivity**

The resistivity of a material is the characteristic of that particular material. The materials can be broadly classified into conductors and insulators. The metals and alloys which have low resistivity of the order of $10^{-6}$ - $10^{-8}\Omega$ m are good conductors of electricity. They carry current without appreciable loss of energy. Example: silver, aluminium, copper, iron, tungsten, nichrome, manganin, constantan. The resistivity of metals increase with increase in temperature.

Insulators are substances which have very high resistivity of the order of $10^{8}$ - $10^{14}\Omega$ m. They offer very high resistance to the flow of current and are termed non–conductors. Example: glass, mica, amber, quartz, wood, teflon, bakelite. In between these two classes of materials lie the semiconductors (Table). They are partially conducting. The resistivity of semiconductor is $10^{-2}$ - $10^{4}\Omega$ m. Example: germanium, silicon.

**Superconductivity**

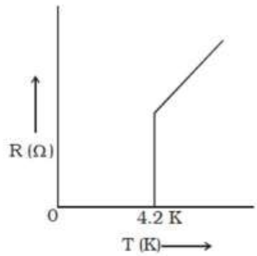
Ordinary conductors of electricity become better conductors at lower

| Classification | Material | ρ (Ω m) |
|---|---|---|
| conductors | silver | $1.6 \times 10^{-8}$ |
| | copper | $1.7 \times 10^{-8}$ |
| | aluminium | $2.7 \times 10^{-8}$ |
| | iron | $10 \times 10^{-8}$ |
| Semiconductors | germanium | 0.46 |
| | silicon | 2300 |
| Insulators | glass | $10^{10} - 10^{14}$ |
| | wood | $10^{8} - 10^{11}$ |
| | quartz | $10^{13}$ |
| | rubber | $10^{13} - 10^{16}$ |

temperatures. The ability of certain metals, their compounds and alloys to conduct electricity with zero resistance at very low temperatures is called superconductivity. The materials which exhibit this property are called superconductors. The phenomenon of superconductivity was first observed by Kammerlingh Onnes in 1911. He found that mercury suddenly showed zero resistance at 4.2 K (Fig 2.3).

The first theoretical explanation of superconductivity was given by Bardeen, Cooper and Schrieffer in 1957 and it is called the BCS theory.

The temperature at which electrical resistivity of the material suddenly drops to zero and the material changes from normal conductor to a superconductor is called the transition temperature or critical temperature TC. At the transition temperature the following changes




are observed:

- The electrical resistivity drops to zero.
- The conductivity becomes infinity
- The magnetic flux lines are excluded from the material.

**Applications of superconductors**

- Superconductors form the basis of energy saving power systems, namely the superconducting generators, which are smaller in size and weight, in comparison with conventional generators.
- Superconducting magnets have been used to levitate trains above its rails. They can be driven at high speed with minimal expenditure of energy.
- Superconducting magnetic propulsion systems may be used to launch satellites into orbits directly from the earth without the use of rockets.
- High efficiency ore–separating machines may be built using superconducting magnets which can be used to separate tumor cells from healthy cells by high gradient magnetic separation method.
- Since the current in a superconducting wire can flow without any change in magnitude, it can be used for transmission lines.

- Superconductors can be used as memory or storage elements in computers.

**Carbon resistors**

The wire wound resistors are expensive and huge in size. Hence, carbon resistors are used. Carbon resistor consists of a ceramic core, on which a thin layer of crystalline carbon is deposited. These resistors are cheaper, stable and small in size. The resistance of a carbon resistor is indicated by the colour code drawn on it (Table). A three colour code carbon resistor is discussed here. The silver or gold ring at one end corresponds to the tolerance. It is a tolerable range (+) of the resistance.

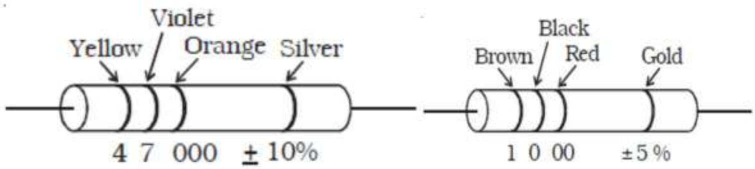
| Colour | Number |
|---|---|
| Black | 0 |
| Brown | 1 |
| Red | 2 |
| Orange | 3 |
| Yellow | 4 |
| Green | 5 |
| Blue | 6 |
| Violet | 7 |
| Grey | 8 |
| White | 9 |

The tolerance of silver, gold, red and brown rings is 10%, 5%, 2% and 1% respectively. If there is no coloured ring at this end, the tolerance is 20%. The first two rings at the other end of tolerance ring are significant figures of resistance in ohm. The third ring indicates the powers of 10 to be multiplied or number of zeroes following the significant figure.

Example:

The first yellow ring in Fig.a corresponds to 4. The next violet ring corresponds to 7. The third orange ring corresponds to 103. The silver ring represents 10% tolerance. The total resistance is $47 \times 10^3 + 10\%$ i.e.

47 k Ω, 10%. Fig.b shows 1 k Ω, 5% carbon resistor.
Presently four colour code carbon resistors are also used. For certain critical applications 1% and 2% tolerance resistors are used.




**Combination of resistors**

In simple circuits with resistors, Ohm's law can be applied to find the effective resistance. The resistors can be connected in series and parallel.

**Resistors in series**

Let us consider the resistors of resistances $R_1$, $R_2$, $R_3$ and $R_4$ connected in series as shown in Fig.

When resistors are connected in series, the current flowing through each resistor is the same. If the potential difference applied between the ends of the combination of resistors is V, then the potential difference across each resistor $R_1$, $R_2$, $R_3$ and $R_4$ is $V_1$, $V_2$, $V_3$ and $V_4$ respectively.

The net potential difference $V = V_1 + V_2 + V_3 + V_4$

By Ohm's law

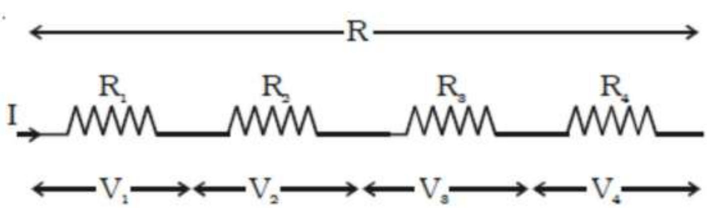
$V_1 = IR_1$, $V_2 = IR_2$, V3 = $IR_3$, $V_4 = IR_4$ and $V = IR_s$

where $R_S$ is the equivalent or effective resistance of the series combination.

Hence, $IR_S = IR_1 + IR_2 + IR_3 + IR_4$ or $R_S = R_1 + R_2 + R_3 + R_4$

Thus, the equivalent resistance of a number of resistors in series connection is equal to the sum of the resistance of individual resistors.

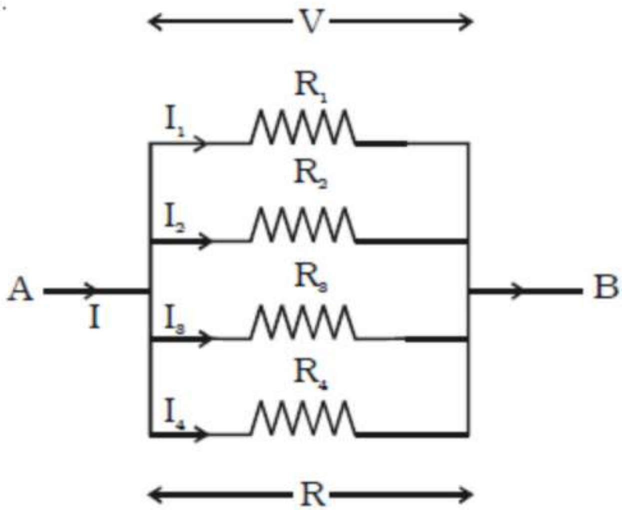
**Resistors in parallel**



Consider four resistors of resistances $R_1$, $R_2$, $R_3$ and $R_4$ are connected in parallel as shown in Fig. A source of emf V is connected to the parallel combination. When resistors are in parallel, the potential difference (V) across each resistor is the same.

A current I entering the combination gets divided into $I_1$, $I_2$, $I_3$ and $I_4$ through $R_1$, $R_2$, $R_3$ and $R_4$ respectively,

such that $I = I_1 + I_2 + I_3 + I_4$.




By Ohm's law

$I_1 = \frac{V}{R_1}, I_2 = \frac{V}{R_2}, I_3 = \frac{V}{R_3}, I_4 = \frac{V}{R_4}$ and $I = \frac{V}{R_P}$

where $R_P$ is the equivalent or effective resistance of the parallel combination.

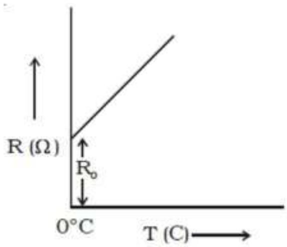
$$\therefore \frac{V}{R_P} = \frac{V}{R_1} + \frac{V}{R_2} + \frac{V}{R_3} + \frac{V}{R_4}$$

$$\frac{1}{R_P} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4}$$

Thus, when a number of resistors are connected in parallel, the sum of the reciprocal of the resistance of the individual resistors is equal to the reciprocal of the effective resistance of the combination.

**Temperature dependence of resistance**

The resistivity of substances varies with temperature. For conductors the resistance increases with increase in temperature. If $R_0$ is the resistance of a conductor at 0° C and $R_t$ is the resistance of same conductor at to C, then




$R_t = R_0 (1 + \alpha t)$

where $\alpha$ is called the temperature coefficient of resistance.

$$\alpha = \frac{R_t - R_0}{R_0 t}$$

The temperature coefficient of resistance is defined as the ratio of increase in resistance per degree rise in temperature to its resistance at 0° C. Its unit is per °C.

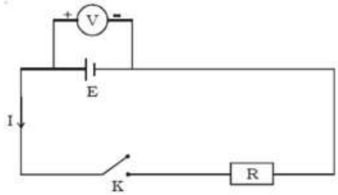
The variation of resistance with temperature is shown in Fig. Metals have positive temperature coefficient of resistance, i.e., their resistance increases with increase in temperature. Insulators and semiconductors have negative temperature coefficient of resistance, i.e., their resistance decreases with increase in temperature. A material with a negative temperature coefficient is called a thermistor. The temperature coefficient is low for alloys.

**Internal resistance of a cell**

The electric current in an external circuit flows from the positive terminal to the negative terminal of the cell, through different circuit elements. In order to maintain continuity, the current has to flow through the electrolyte of the cell, from its negative terminal to positive terminal. During this process of flow of current inside the cell, a resistance is offered to current flow by the electrolyte of the cell. This is termed as the internal resistance of the cell. A freshly prepared cell has low internal resistance and this increases with ageing.

**Determination of internal resistance of a cell using voltmeter**

The circuit connections are made as shown in Fig. With key K open, the emf of cell E is found by connecting a high resistance voltmeter across it. Since the high resistance voltmeter draws only a very feeble




current for deflection, the circuit may be considered as an open circuit. Hence the voltmeter reading gives the emf of the cell. A small value of resistance R is included in the external circuit and key K is closed. The potential difference across R is equal to the potential difference across cell (V).

The potential drop across R, V = IR ... (1)

Due to internal resistance r of the cell, the voltmeter reads a value V, less than the emf of cell.

Then V = E – Ir or Ir = E–V ... (2)

Dividing equation (2) by equation (1)

$\frac{Ir}{IR} = \frac{E-V}{V}$ or $r = \left(\frac{E-V}{V}\right) R$

Since E, V and R are known, the internal resistance r of the cell can be determined.

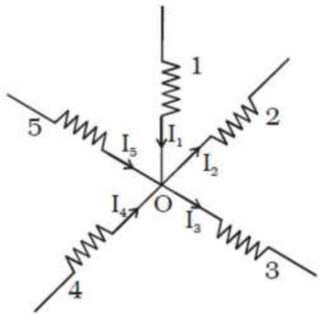
**Kirchoff's law**

Ohm's law is applicable only for simple circuits. For complicated circuits, Kirchoff's laws can be used to find current or voltage. There are two generalised laws:

(i) Kirchoff's current law (ii) Kirchoff's voltage law

**Kirchoff's first law (current law)**

Kirchoff's current law states that the algebraic sum of the currents meeting at any junction in a circuit is zero. The convention is that, the current flowing towards a junction is positive and the current flowing




away from the junction is negative. Let 1,2,3,4 and 5 be the conductors meeting at a junction O in an electrical circuit (Fig). Let $I_1$, $I_2$, $I_3$, $I_4$ and $I_5$ be the currents passing through the conductors respectively. According to Kirchoff's first law.

$I_1 + (-I_2) + (-I_3) + I_4 + I_5 = 0$ or $I_1 + I_4 + I_5 = I_2 + I_3$.

The sum of the currents entering the junction is equal to the sum of the currents leaving the junction. This law is a consequence of conservation of charges.

**Kirchoff's second law (voltage law)**

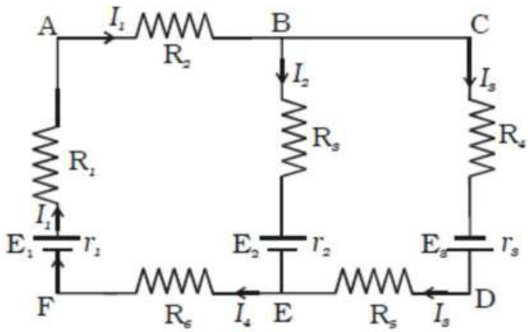
Kirchoff's voltage law states that the algebraic sum of the products of resistance and current in each part of any closed circuit is equal to the algebraic sum of the emf's in that closed circuit. This law is a consequence of conservation of energy.

In applying Kirchoff's laws to electrical networks, the direction of current flow may be assumed either clockwise or anticlockwise. If the assumed direction of current is not the actual direction, then on solving the problems, the current will be found to have negative sign. If the result is positive, then the assumed direction is the same as actual direction. It should be noted that, once the particular direction has been assumed, the same should be used throughout the problem. However, in the application of Kirchoff's second law, we follow that the current in clockwise direction is taken as positive and the current in anticlockwise direction is taken as negative.

Let us consider the electric circuit given in Fig.

Considering the closed loop ABCDEFA,

$I_1R_2 + I_3R_4 + I_3r_3 + I_3R_5 + I_4R_6 + I_1r_1 + I_1R_1 = E_1 + E_3$




Both cells $E_1$ and $E_3$ send currents in clockwise direction.

For the closed loop ABEFA

$I_1R_2 + I_2R_3 + I_2r_2 + I_4R_6 + I_1r_1 + I_1R_1 = E_1 - E_2$

Negative sign in $E_2$ indicates that it sends current in the anticlockwise direction.

As an illustration of application of Kirchoff's second law, let us calculate the current in the following networks.

Illustration I

Applying first law to the Junction B, (Fig)

$I_1 - I_2 - I_3 = 0$

$\therefore I_3 = I_1 - I_2$ ... (1)

For the closed loop ABEFA,

$132\ I_3 + 20I_1 = 200$ ... (2)

Substituting equation (1) in equation (2)

$132\ (I_1 - I_2) + 20I_1 = 200$

$152I_1 - 132I_2 = 200$ ... (3)

For the closed loop BCDEB,

$60I_2 - 132I_3 = 100$

substituting for $I_3$,

$\therefore 60I_2 - 132\ (I_1 - I_2) = 100$

$-132I_1 + 192I_2 = 100$ ... (4)

Solving equations (3) and (4), we obtain

$I_l = 4.39$ A and $I_2 = 3.54$ A

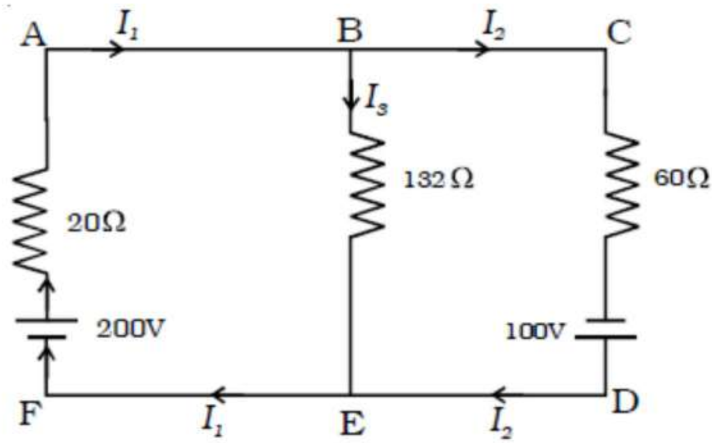



Illustration 2

Taking the current in the clockwise direction along ABCDA as positive (Fig)

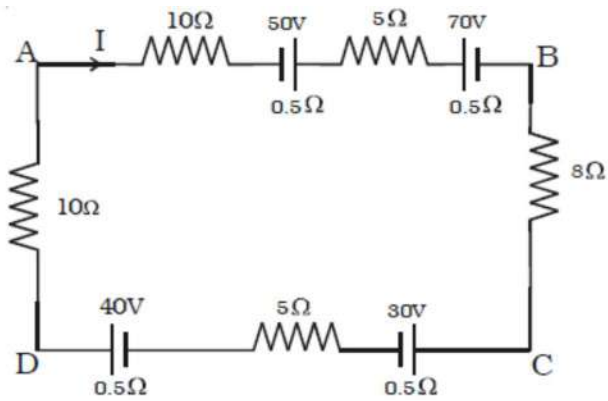
$10\,I + 0.5\,I + 5\,I + 0.5\,I + 8\,I + 0.5\,I + 5\,I + 0.5\,I + 10\,I = 50 - 70 - 30 + 40$

$I\,(10 + 0.5 + 5 + 0.5 + 8 + 0.5 + 5 + 0.5 + 10) = -10$

$40\,I = -10$

$\therefore I = \frac{-10}{40} = -0.25$ A

The negative sign indicates that the current flows in the anticlockwise direction.




**Wheatstone's bridge**

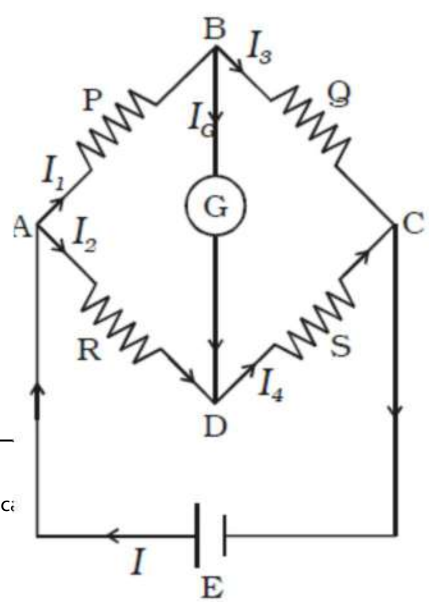
An important application of Kirchoff's law is the Wheatstone's bridge (Fig). Wheatstone's network consists of resistances P, Q, R and S connected to form a closed path. A cell of emf E is connected between points A and C. The current I from the cell is divided into $I_1$, $I_2$, $I_3$ and $I_4$ across the four branches. The current through the galvanometer is $I_g$. The resistance of
galvanometer is G.

Applying Kirchoff's current law to junction B,

$I_1 - I_g - I_3 = 0$ ... (1)

Applying Kirchoff's current law to junction D

$I_2 + I_g - I_4 = 0$ ... (2)



Applying Kirchoff's voltage law to closed path ABDA

$I_1 P + I_g G - I_2 R = 0$ ... (3)

Applying Kirchoff's voltage law to closed path ABCDA

$I_1P + I_3Q - I_4S - I_2R = 0$ ... (4)

When the galvanometer shows zero deflection, the points B and D are at same potential and $I_g = 0$. Substituting $I_g = 0$ in equation (1), (2) and (3)

$I_1 = I_3$ ... (5)

$I_2 = I_4$ ... (6)

$I_1P = I_2R$ ... (7)

Substituting the values of (5) and (6) in equation (4)

$I_1P + I_1Q - I_2S - I_2R = 0$

$I_1 (P + Q) = I_2 (R+S)$ ... (8)

Dividing (8) by (7)

$$\frac{I_1(P+Q)}{I_1P} = \frac{I_2(R+S)}{I_2R}$$

$$\therefore \frac{P+Q}{P} = \frac{R+S}{R}$$

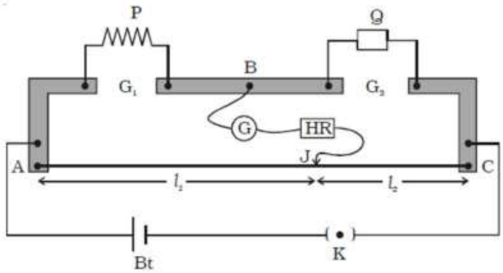
$$1 + \frac{Q}{P} = 1 + \frac{S}{R}$$

$$\therefore \frac{Q}{P} = \frac{S}{R} \text{ or } \frac{P}{Q} = \frac{R}{S}$$

This is the condition for bridge balance. If P, Q and R are known, the resistance S can be calculated.

**Metre Bridge**

Metre bridge is one form of Wheatstone's bridge. It consists of thick strips of copper, of negligible resistance, fixed to a wooden board. There are two gaps G1 and G2 between these strips. A uniform manganin wire AC of length one metre whose temperature coefficient is low, is stretched along a metre scale and its ends are soldered to two copper strips. An unknown resistance P is connected in the gap G1 and a standard resistance Q is connected in the gap G2 (Fig 2.13). A metal jockey J is connected to B through a galvanometer (G) and a high resistance (HR) and it can make contact at any point on the wire AC. Across the two ends of the wire, a Leclanche cell and a key are connected. Adjust the position of metal jockey on metre bridge wire so




that the galvanometer shows zero deflection. Let the point be J. The portions AJ and JC of the wire now replace the resistances R and S of Wheatstone's bridge. Then

$$\frac{P}{Q} = \frac{R}{S}\frac{r.AJ}{r.JC}$$

where r is the resistance per unit length of the wire.

$$\therefore \frac{P}{Q} = \frac{AJ}{JC} = \frac{l_1}{l_2}$$

where $AJ = l_1$ and $JC = l_2$

$$\therefore P = Q\,\frac{l_1}{l_2}$$

Though the connections between the resistances are made by thick copper strips of negligible resistance, and the wire AC is also soldered to such strips a small error will occur in the value of $\frac{l_1}{l_2}$due to the end resistance. This error can be eliminated, if another set of readings are taken with P and Q interchanged and the average value of P is found, provided the balance point J is near the mid-point of the wire AC.

Determination of specific resistance

The specific resistance of the material of a wire is determined by knowing the resistance (P), radius (r) and length (L) of the wire using the expression

$$\rho = \frac{P\pi r^2}{L}$$

Determination of temperature coefficient of resistance

If $R_1$ and $R_2$ are the resistances of a given coil of wire at the temperatures $t_1$ and $t_2$, then the temperature coefficient of resistance of the material of the coil is determined using the relation,

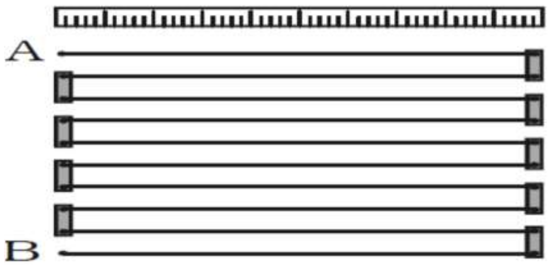
$$\alpha = \frac{R_2 - R_1}{R_1 t_1 - R_2 t_1}$$

**Potentiometer**

The Potentiometer is an instrument used for the measurement of potential difference (Fig). It consists of a ten metre long uniform wire of manganin or constantan stretched in ten segments, each of one metre length. The segments are stretched parallel to each other on a horizontal wooden board. The ends of the wire are fixed to copper strips with binding screws. A metre scale is fixed on the board, parallel to the wire. Electrical contact with wires is established by pressing the jockey J.

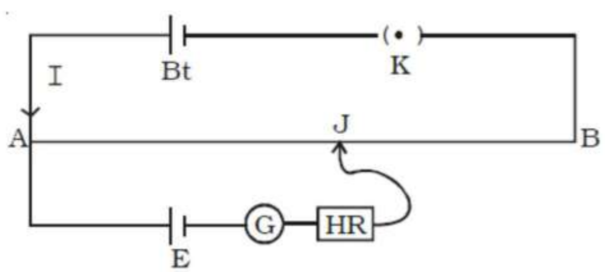
**Principle of potentiometer**

A battery Bt is connected between the ends A and B of a potentiometer




wire through a key K. A steady current I flow through the potentiometer wire (Fig). This forms the primary circuit. A primary cell is connected in series with the positive terminal A of the potentiometer, a galvanometer, high resistance and jockey. This forms the secondary circuit.

If the potential difference between A and J is equal to the emf of the cell, no current flows through the galvanometer. It shows zero deflection. AJ is called the balancing length. If the balancing length is l, the potential difference across AJ = Irl where r is the resistance per unit length of the potentiometer wire and I the current in the primary circuit.




$\therefore E = Irl$,

since I and r are constants, E $\alpha$ l

Hence emf of the cell is directly proportional to its balancing length. This is the principle of a potentiometer.

**Comparison of emfs of two given cells using potentiometer**

The potentiometer wire AB is connected in series with a battery (Bt), Key (K), rheostat (Rh) as shown in Fig. This forms the primary circuit. The end A of potentiometer is connected to the terminal C of a DPDT switch (six ways key–double pole double throw). The terminal D is connected to the jockey (J) through a galvanometer (G) and high resistance (HR). The cell of emf $E_1$ is connected between terminals $C_1$ and $D_1$ and the cell of emf $E_2$ is connected between $C_2$ and $D_2$ of the DPDT switch. Let I be the current flowing through the primary circuit and r be the resistance of the potentiometer wire per metre length. The DPDT switch is pressed towards $C_1$, $D_1$ so that cell $E_1$ is included in the secondary circuit. The jockey is moved on the wire and adjusted for zero deflection in galvanometer. The balancing length is $l_1$. The potential difference across the balancing length $l_1 = Irl_1$. Then, by the principle of potentiometer,

$E_1 = Irl_1$ ... (1)

The DPDT switch is pressed towards $E_2$. The balancing length $l_2$ for zero deflection in galvanometer is determined. The potential difference across the balancing length is $l_2 = Irl_2$, then

$E_2 = Irl_2$ ... (2)

Dividing (1) and (2) we get

$\frac{E_1}{E_2} = \frac{l_1}{l_2}$

If emf of one cell ($E_1$) is known, the emf of the other cell ($E_2$) can be calculated using the relation.

$$E_2 = E_1 \frac{l_2}{l_1}$$

**Comparison of emf and potential difference**

- The difference of potentials between the two terminals of a cell in an open circuit is called the electromotive force (emf) of a cell. The difference in potentials between any two points in a closed circuit is called potential difference.
- The emf is independent of external resistance of the circuit, whereas potential difference is proportional to the resistance between any two points.

**Electric energy and electric power**

If I is the current flowing through a conductor of resistance R in time t, then the quantity of charge flowing is, q = It. If the charge q, flows between two points having a potential difference V, then the work done in moving the charge is = V. q = VIt. Then, electric power is defined as the rate of doing electric work.

Power $= \frac{Work\ done}{time} = \frac{VIt}{t} = VI$

Electric power is the product of potential difference and current strength.

Since V = IR, Power = $I^2R$

Electric energy is defined as the capacity to do work. Its unit is joule. In practice, the electrical energy is measured by watt hour (Wh) or kilowatt hour (kWh). 1 kWh is known as one unit of electric energy.

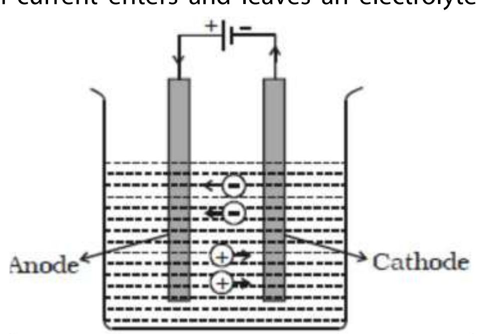
(1 kWh = 1000 Wh = $1000 \times 3600$ J = $36 \times 10^5$ J)

**Wattmeter**

A wattmeter is an instrument used to measure electrical power consumed i.e energy absorbed in unit time by a circuit. The wattmeter consists of a movable coil arranged between a pair of fixed coils in the form of a solenoid. A pointer is attached to the movable coil. The free end of the pointer moves over a circular scale. When current flows through the coils, the deflection of the pointer is directly proportional to the power.

**Chemical effect of current**

The passage of an electric current through a liquid causes chemical changes and this process is called electrolysis. The conduction is possible, only in liquids wherein charged ions can be dissociated in opposite directions (Fig). Such liquids are called electrolytes. The plates through which current enters and leaves an electrolyte are known as



electrodes. The electrode towards which positive ions travel is called the cathode and the other, towards which negative ions travel is called anode. The positive ions are called cations and are mostly formed from metals or hydrogen. The negative ions are called anions.

**Faraday's laws of electrolysis**

The factors affecting the quantities of matter liberated during the process of electrolysis were investigated by Faraday.

**First Law:** The mass of a substance liberated at an electrode is directly proportional to the charge passing through the electrolyte.

If an electric current I is passed through an electrolyte for a time t, the amount of charge (q) passed is I t. According to the law, mass of substance liberated (m) is

m $\alpha$ q or m = zIt

where Z is a constant for the substance being liberated called as electrochemical equivalent. Its unit is kg $C^{-1}$.

The electrochemical equivalent of a substance is defined as the mass of substance liberated in electrolysis when one coulomb charge is passed through the electrolyte.

**Second Law:** The mass of a substance liberated at an electrode by a given amount of charge is proportional to the chemical equivalent of the substance. If E is the chemical equivalent of a substance, from the second law

m $\alpha$ E

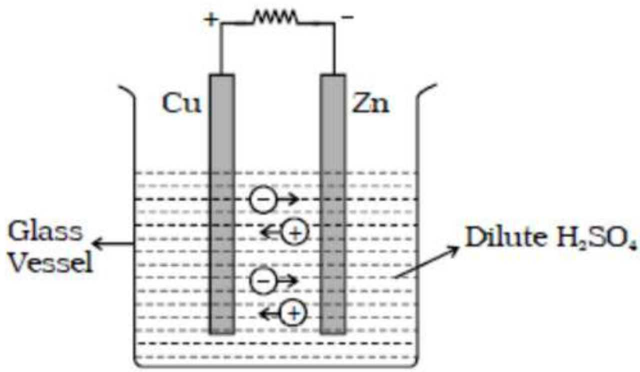
**Electric cells**

The starting point to the development of electric cells is the classic experiment by Luige Galvani and his wife Lucia on a dissected frog hung from iron railings with brass hooks. It was observed that, whenever the leg of the frog touched the iron railings, it jumped and this led to the introduction of animal electricity. Later, Italian scientist and genius professor Alessandro Volta came up with an electrochemical battery. The battery Volta named after him consisted of a pile of copper and zinc discs placed alternately separated by paper and introduced in salt solution. When the end plates were connected to an electric bell, it continued to ring, opening a new world of electrochemical cells. His experiment established that, a cell could be made by using two dissimilar metals and a salt solution which reacts with at least one of the metals as electrolyte.

**Voltaic cell**

The simple cell or voltaic cell consists of two electrodes, one of copper and the other of zinc dipped in a solution of dilute sulphuric acid in a glass vessel (Fig). On connecting the two electrodes externally, with a piece of wire, current flows from copper to zinc outside the cell and from zinc to copper inside it. The copper electrode is the positive pole or copper rod of the cell and zinc is the negative pole or zinc rod of the cell. The electrolyte is dilute sulphuric acid.

The action of the cell is explained in terms of the motion of the charged ions. At the zinc rod, the zinc atoms get ionized and pass into solution as $Zn^{++}$ ions. This leaves the zinc rod with two electrons more, making it negative. At the same time, two hydrogen ions ($2H^+$) are discharged at the copper rod, by taking these two electrons. This makes the copper



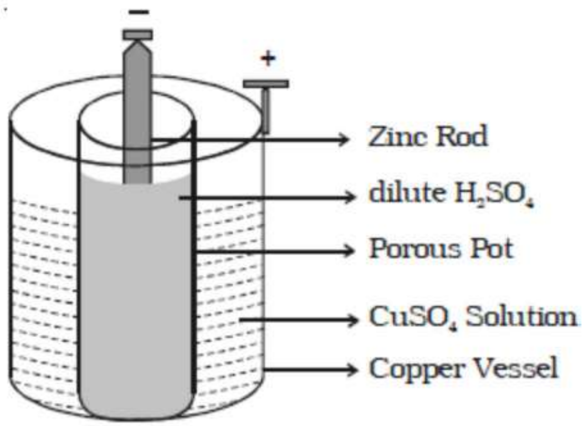

rod positive. As long as excess electrons are available on the zinc electrode, this process goes on and a current flows continuously in external circuit. This simple cell is thus seen as a device which converts chemical energy into electrical energy. Due to opposite charges on the two plates, a potential difference is set up between copper and zinc, copper being at a higher potential than zinc. The difference of potential between the two electrodes is 1.08V.

**Primary Cell**

The cells from which the electric energy is derived by irreversible chemical actions are called primary cells. The primary cell is capable of giving an emf, when its constituents, two electrodes and a suitable electrolyte, are assembled together. The three main primary cells, namely Daniel Cell and Leclanche cell are discussed here. These cells cannot be recharged electrically.

**Daniel cell**

Daniel cell is a primary cell which cannot supply steady current for a long time. It consists of a copper vessel containing a strong solution of copper sulphate (Fig). A zinc rod is dipped in dilute sulphuric acid contained in a porous pot. The porous pot is placed inside the copper sulphate solution. The zinc rod reacting with dilute sulphuric acid




produces $Zn^{++}$ ions and 2 electrons.

$Zn^{++}$ ions pass through the pores of the porous pot and reacts with copper sulphate solution, producing $Cu^{++}$ ions. The $Cu^{++}$ ions deposit on the copper vessel. When Daniel cell is connected in a circuit, the two electrons on the zinc rod pass through the external circuit and reach the copper vessel thus neutralizing the copper ions. This constitutes an electric current from copper to zinc. Daniel cell produces an emf of 1.08 volt.

**Leclanche cell**

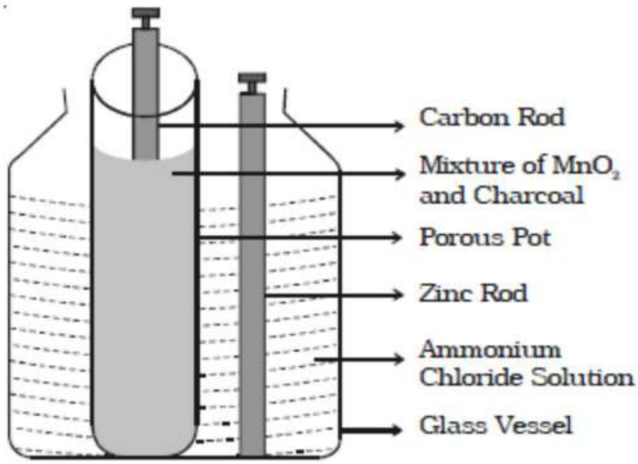
A Leclanche cell consists of a carbon electrode packed in a porous pot containing manganese

dioxide and charcoal powder (Fig). The porous pot is immersed in a saturated solution of ammonium chloride (electrolyte) contained in an outer glass vessel. A zinc rod is immersed in electrolytic solution.

At the zinc rod, due to oxidation reaction Zn atom is converted into $Zn^{++}$ ions and 2 electrons. $Zn^{++}$ ions reacting with ammonium chloride produces zinc chloride and ammonia gas.

i.e $Zn^{++} + 2\,NH_4Cl \rightarrow 2NH_3 + ZnCl_2 + 2\,H^+ + 2e^-$

The ammonia gas escapes. The hydrogen ions diffuse through the pores of the porous pot and react with manganese dioxide. In this process the positive charge of hydrogen ion is transferred to carbon rod. When zinc rod and carbon rod are connected externally, the two electrons from the zinc rod move towards carbon and neutralizes the positive charge. Thus current flows from carbon to zinc. Leclanche cell is useful for supplying




intermittent current. The emf of the cell is about 1.5 V, and it can supply a current of 0.25 A.

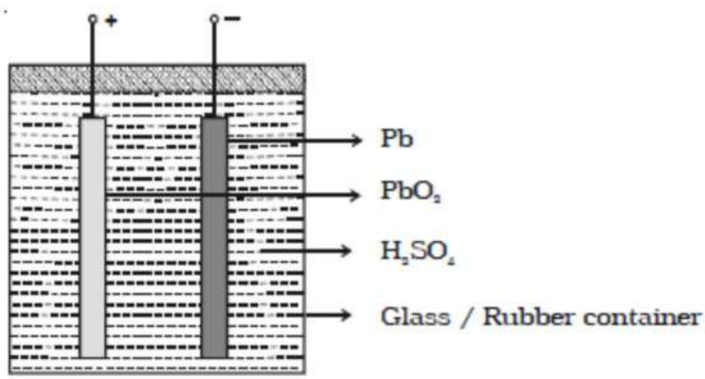
**Secondary Cells**

The advantage of secondary cells is that they are rechargeable. The chemical reactions that take place in secondary cells are reversible. The active materials that are used up when the cell delivers current can be reproduced by passing current through the cell in opposite direction. The chemical process of obtaining current from a secondary cell is called discharge. The process of reproducing active materials is called charging. The most common secondary cells are lead acid accumulator and alkali accumulator.

**Lead – Acid accumulator**

The lead acid accumulator consists of a container made up of hard rubber or glass or celluloid. The container contains dilute sulphuric acid which acts as the electrolyte. Spongy lead (Pb) acts as the negative electrode and lead oxide ($PbO_2$) acts as the positive electrode (Fig). The electrodes are separated by suitable insulating materials and assembled in a way to give low internal resistance. When the cell is connected in a circuit, due to the oxidation reaction that takes place at the negative electrode, spongy lead reacting with dilute sulphuric acid produces lead sulphate and two electrons. The electrons flow in the external circuit from negative electrode to positive electrode where the reduction action takes place. At the positive electrode, lead oxide on reaction with sulphuric acid produces lead sulphate and the two electrons are neutralized in this process. This makes the conventional current to flow from positive electrode to negative electrode in the external circuit.

The emf of a freshly charged cell is 2.2 Volt and the specific gravity of the electrolyte is 1.28. The cell has low internal resistance and hence can deliver high current. As the cell is discharged by drawing current from it, the emf falls to about 2 volts. In the process of charging, the chemical reactions are reversed.




**Applications of secondary cells**

The secondary cells are rechargeable. They have very low internal resistance. Hence they can deliver a high current if required. They can be recharged a very large number of times without any deterioration in properties. These cells are huge in size. They are used in all automobiles like cars, two wheelers, trucks etc. The state of charging these cells is, simply monitoring the specific gravity of the electrolyte. It should lie between 1.28 to 1.12 during charging and discharging respectively.

## Unit IV

## Electromagnetic Induction and Alternating Current

In the year 1820, Hans Christian Oersted demonstrated that a current carrying conductor is associated with a magnetic field. Thereafter, attempts were made by many to verify the reverse effect of producing an induced emf by the effect of magnetic field.

### Electromagnetic induction

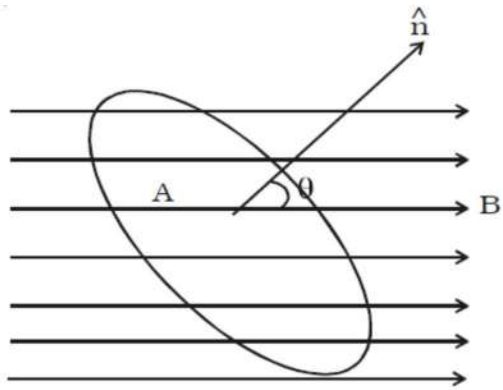
Michael Faraday demonstrated the reverse effect of Oersted experiment. He explained the possibility of producing emf across the ends of a conductor when the magnetic flux linked with the conductor changes. This was termed as electromagnetic induction. The discovery of this phenomenon brought about a revolution in the field of power generation.

### Magnetic flux

The magnetic flux (φ) linked with a surface held in a magnetic field (B) is defined as the number of magnetic lines of force crossing a closed area (A) (Fig ). If θ is the angle between the direction of the field and normal to the area, then

$\varphi = B . A$

$\varphi = BA \cos\theta$

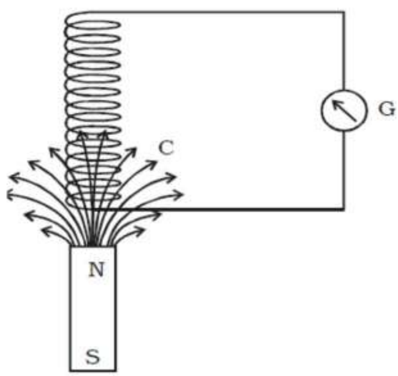



### Induced emf and current – Electromagnetic induction.

Whenever there is a change in the magnetic flux linked with a closed circuit an emf is produced. This emf is known as the induced emf and the current that flows in the closed circuit is called induced current. The phenomenon of producing an induced emf due to the changes in the magnetic flux associated with a closed circuit is known as electromagnetic induction.

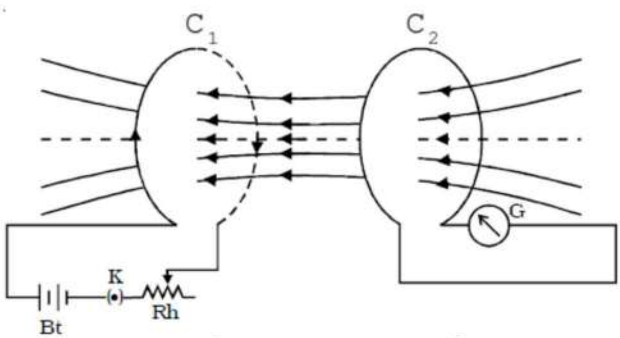
Faraday discovered the electromagnetic induction by conducting several experiments. Fig. consists of a cylindrical coil C made up of several turns of insulated copper wire connected in series to a sensitive galvanometer G. A strong bar magnet NS with its North Pole pointing towards the coil is moved up and down. The following inferences were made by Faraday.

- Whenever there is a relative motion between the coil and the magnet, the galvanometer shows deflection indicating the flow of induced current.




- The deflection is momentary. It lasts so long as there is relative motion between the coil and the magnet.
- The direction of the flow of current changes if the magnet is moved towards and withdrawn from it.
- The deflection is more when the magnet is moved faster, and less when the magnet is moved slowly.
- However, on reversing the magnet (i.e) South Pole pointing towards the coil, same results are obtained, but current flows in the opposite direction.

Faraday demonstrated the electromagnetic induction by another experiment also. This Fig. shows two coils $C_1$ and $C_2$ placed close to each other. The coil $C_1$ is connected to a battery Bt through a key K and a rheostat. Coil $C_2$ is connected to a sensitive galvanometer G and kept




close to $C_1$. When the key K is pressed, the galvanometer connected

with the coil $C_2$ shows a sudden momentary deflection. This indicates that a current is induced in coil $C_2$. This is because when the current in $C_1$ increases from zero to a certain steady value, the magnetic flux linked with the coil $C_1$ increases. Hence, the magnetic flux linked with the coil $C_2$ also increases. This causes the deflection in the galvanometer. On releasing K, the galvanometer shows deflection in the opposite direction. This indicates that a current is again induced in the coil $C_2$. This is because when the current in $C_1$ decreases from maximum to zero value, the magnetic flux linked with the coil $C_1$ decreases. Hence, the magnetic flux linked with the coil $C_2$ also decreases. This causes the deflection in the galvanometer in the opposite direction.

**Faraday's laws of electromagnetic induction**

Based on his studies on the phenomenon of electromagnetic induction, Faraday proposed the following two laws.

First law

Whenever the amount of magnetic flux linked with a closed circuit changes, an emf is induced in the circuit. The induced emf lasts so long as the change in magnetic flux continues.

Second law

The magnitude of emf induced in a closed circuit is directly proportional to the rate of change of magnetic flux linked with the circuit.

Let $\phi_1$ be the magnetic flux linked with the coil initially and $\phi_2$ be the magnetic flux linked with the coil after a time t. Then

Rate of change of magnetic flux = $\frac{\phi_2-\phi_1}{t}$

According to Faraday's second law, the magnitude of induced emf is, e$\frac{\phi_2-\phi_1}{t}$. If d$\phi$ is the change in magnetic flux in a time dt, then the above equation can be written as e $\alpha$ $\frac{d\phi}{dt}$

**Lenz's law**

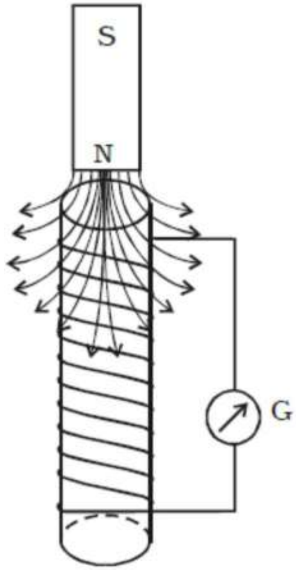
The Russian scientist H.F. Lenz in 1835 discovered a simple law giving the direction of the induced current produced in a circuit. Lenz's law states that the induced current produced in a circuit always flows in such a direction that it opposes the change or cause that produces it.

If the coil has N number of turns and $\phi$ is the magnetic flux linked with each turn of the coil then, the total magnetic flux linked with the coil at any time is N$\phi$.

$\therefore e = -\frac{d}{dt}(N\phi)$ $e = \frac{-Nd\phi}{dt}$ $= \frac{-N(\phi_2-\phi_1)}{t}$

**Lenz's law - a consequence of conservation of energy**

Copper coils are wound on a cylindrical cardboard and the two ends of the coil are connected to a sensitive galvanometer. A magnet is moved towards the coil (Fig). The upper face of the coil acquires north polarity.




Consequently work has to be done to move the magnet further against the force of repulsion. When we withdraw the magnet away from the coil, its upper face acquires south polarity. Now the workdone is against the force of attraction. When the magnet is moved, the number of magnetic lines of force linking the coil changes, which causes an induced current to flow through the coil. The direction of the induced current, according to Lenz's law is always to oppose the motion of the magnet. The workdone in moving the magnet is converted into electrical energy. This energy is dissipated as heat energy in the coil. If on the contrary, the direction of the current were to help the motion of the magnet, it would start moving faster increasing the change of magnetic flux linking the coil. This results in the increase of induced current. Hence kinetic energy and electrical energy would be produced without any external work being done, but this is impossible. Therefore, the induced current always flows in such a direction to oppose the cause. Thus it is proved that Lenz's law is the consequence of conservation of energy.

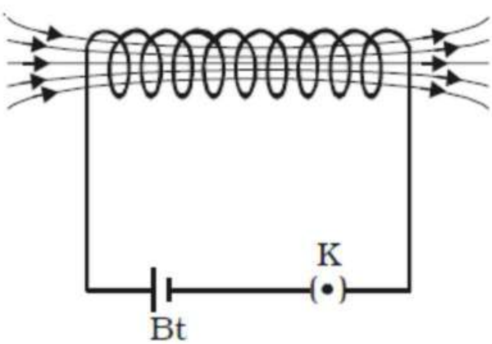
**Fleming's right hand rule**

The forefinger, the middle finger and the thumb of the right hand are held in the three mutually perpendicular directions. If the forefinger points along the direction of the magnetic field and the thumb are along the direction of motion of the conductor, then the middle finger points in the direction of the induced current. This rule is also called generator rule.

**Self-Induction**

The property of a coil which enables to produce an opposing induced emf in it when the current in the coil changes is called self-induction. A coil is connected in series with a battery and a key (K) (Fig.).

On pressing the key, the current through the coil increases to a



maximum value and correspondingly the magnetic flux linked with the coil also increases. An induced current flows through the coil which according to Lenz's law opposes the further growth of current in the coil. On releasing the key, the current through the coil decreases to a zero value and the magnetic flux linked with the coil also decreases. According to Lenz's law, the induced current will oppose the decay of current in the coil.

**Coefficient of self-induction**

When a current I flows through a coil, the magnetic flux ($\phi$) linked with the coil is proportional to the current.

$\phi \; \alpha$ I or $\phi$ = LI

where L is a constant of proportionality and is called coefficient of self-induction or self-inductance.

If I = 1A, $\phi$ = L × 1, then L = $\phi$ Therefore, coefficient of self-induction of a coil is numerically equal to the magnetic flux linked with a coil when unit current flows through it. According to laws of electromagnetic induction.

$e = -\frac{d\phi}{dt} = -\frac{d}{dt}(LI)$ or $e = -L\frac{dI}{dt}$

If $\frac{dI}{dt} = 1$ A $s^{-1}$, then L = −e

The coefficient of self-induction of a coil is numerically equal to the opposing emf induced in the coil when the rate of change of current through the coil is unity. The unit of self-inductance is henry (H). One henry is defined as the self-inductance of a coil in which a change in current of one ampere per second produces an opposing emf of one volt.

**Energy associated with an inductor**

Whenever current flows through a coil, the self–inductance opposes the growth of the current. Hence, some work has to be done by external agencies in establishing the current. If e is the induced emf then,

$$e = -L\frac{dI}{dt}$$

The small amount of work dw done in a time interval dt is

dw = e.I dt

$= -L\frac{dI}{dt}I.\,dt$

The total work done when the current increases from 0 to maximum value ($I_0$) is

$$w = \int dw = \int_0^{I_0} -L\,I\,dI$$

This work done is stored as magnetic potential energy in the coil.

∴ Energy stored in the coil

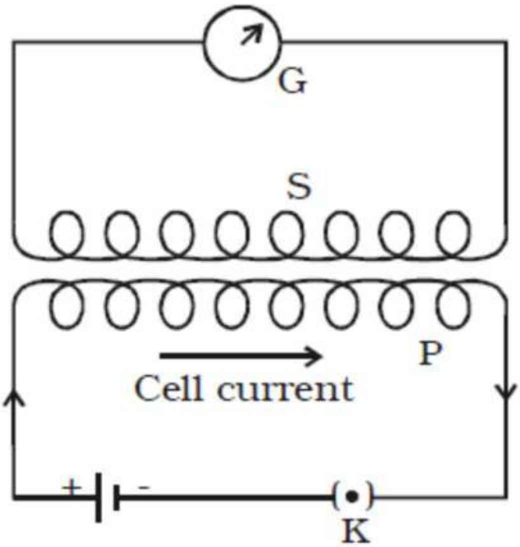
$= -L \int_0^{I_0} I\, dI = -\frac{1}{2} LI_0^2$

Negative sign is consequence of Lenz's Law. Hence, quantitatively, the energy stored in an inductor is $\frac{1}{2} LI_0^2$.

**Mutual induction**

Whenever there is a change in the magnetic flux linked with a coil, there is also a change of flux linked with the neighbouring coil, producing an induced emf in the second coil. This phenomenon of producing an induced emf in a coil due to the change in current in the other coil is known as mutual induction.

P and S are two coils placed close to each other (Fig.). P is connected to a battery through a key K. S is connected to a galvanometer G. On




pressing K, current in P starts increasing from zero to a maximum value. As the flow of current increases, the magnetic flux linked with P increases. Therefore, magnetic flux linked with S also increases producing an induced emf in S. Now, the galvanometer shows the deflection. According to Lenz's law the induced current in S would oppose the increase in current in P by flowing in a direction opposite to the current in P, thus delaying the growth of current to the maximum value. When the key 'K' is released, current starts decreasing from maximum to zero value, consequently magnetic flux linked with P decreases. Therefore magnetic flux linked with S also decreases and hence, an emf is induced in S. According to Lenz's law, the induced current in S flows in such a direction so as to oppose the decrease in current in P thus prolonging the decay of current.

**Coefficient of mutual induction**

IP is the current in coil P and $\phi_s$ is the magnetic flux linked with coil S due to the current in coil P.

$\therefore \phi_s \; \alpha$ IP or $\phi_s = M I_P$

where M is a constant of proportionality and is called the coefficient of mutual induction or mutual inductance between the two coils.

If $I_P$ = 1A, then, M = $\phi_s$

Thus, coefficient of mutual induction of two coils is numerically equal to the magnetic flux linked with one coil when unit current flows through the neighboring coil. If $e_s$ is the induced emf in the coil (S) at any instant of time, then from the laws of electromagnetic induction,

$$e_S = -\frac{d\phi_S}{dt} = -\frac{d}{dt}(MI_P) = M\frac{dI_P}{dt}$$

$$\therefore M = \frac{e_S}{-\left(\frac{dI_P}{dt}\right)}$$

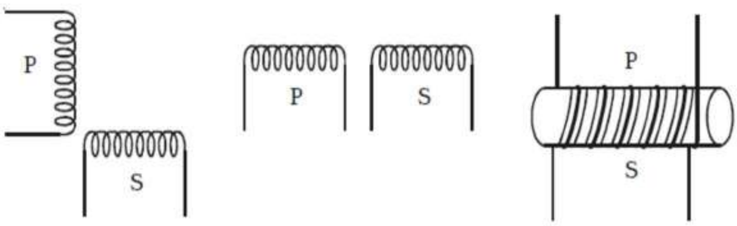
If $\frac{dI_P}{dt} = 1$ A s$^{-1}$, then, M = $-e_s$

Thus, the coefficient of mutual induction of two coils is numerically equal to the emf induced in one coil when the rate of change of current through the other coil is unity. The unit of coefficient of mutual induction is henry. One henry is defined as the coefficient of mutual induction between a pair of coils when a change of current of one ampere per second in one coil produces an induced emf of one volt in the other coil. The coefficient of mutual induction between a pair of coils depends on the following factors

(i) Size and shape of the coils, number of turns and permeability of material on which the coils are wound.

(ii) proximity of the coils Two coils P and S have their axes perpendicular to each other (Fig.a). When a current is passed through coil P, the magnetic flux linked with S is small and hence, the coefficient of mutual induction between the two coils is small. The two coils are placed in such a way that they have a common axis (Fig. b). When current is passed through the coil P the magnetic flux linked with coil S is large and hence, the coefficient of mutual induction between the two coils is large.

If the two coils are wound on a soft iron core (Fig. c) the mutual induction is very large.




(a) (b) (c)

**Mutual induction of two long solenoids.**

$S_1$ and $S_2$ are two long solenoids each of length l. The solenoid $S_2$ is wound closely over the solenoid $S_1$ (Fig.). $N_1$ and $N_2$ are the number of turns in the solenoids $S_1$ and $S_2$ respectively. Both the solenoids are considered to have the same area of cross section A as they are closely wound together. $I_1$ is the current flowing through the solenoid $S_1$. The magnetic field $B_1$ produced at any point inside the solenoid $S_1$ due to the current $I_1$ is

$B_1 = \mu_0 \frac{N_1}{l} I_1$ … (1)

The magnetic flux linked with each turn of $S_2$ is equal to $B_1A$.

Total magnetic flux linked with solenoid $S_2$ having $N_2$ turns is

$\phi_2 = B_1AN_2$

Substituting for $B_1$ from equation (1)

$\phi_2 = \left(\mu_0 \frac{N_1}{l} I_1\right)AN_2$

$\phi_2 = \frac{\mu_0 N_1 N_2 A I_1}{l}$ … (2)

But $\phi_2 = MI_1$ … (3)

where M is the coefficient of mutual induction between $S_1$ and $S_2$.

From equations (2) and (3)

$MI_1 = \frac{\mu_0 N_1 N_2 A I_1}{l}$

$M = \frac{\mu_0 N_1 N_2 A}{l}$

If the core is filled with a magnetic material of permeability $\mu$,

$M = \frac{\mu N_1 N_2 A}{l}$

**Methods of producing induced emf**

We know that the induced emf is given by the expression

$e_s = -\frac{d\phi}{dt} = -\frac{d}{dt}$ (NBA cos θ)

Hence, the induced emf can be produced by changing

- the magnetic induction (B)
- area enclosed by the coil (A) and
- the orientation of the coil (θ) with respect to the magnetic field.

**Emf induced by changing the magnetic induction.**

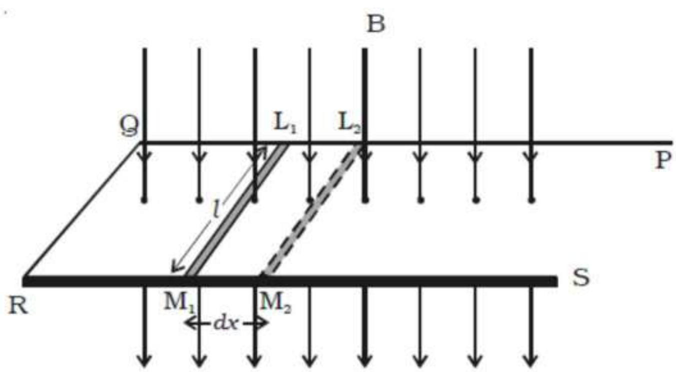
The magnetic induction can be changed by moving a magnet either towards or away from a coil and thus an induced emf is produced in the coil. The magnetic induction can also be changed in one coil by changing the current in the neighbouring coil thus producing an induced emf.

$$\therefore e = -NA\ \cos\theta \left(\frac{dB}{dt}\right)$$

**Emf induced by changing the area enclosed by the coil**

PQRS is a conductor bent in the shape as shown in the Fig. $L_1M$= is a sliding conductor of length l resting on the arms PQ and RS. A uniform magnetic field 'B' acts perpendicular to the plane of the conductor. The closed area of the conductor is $L_1QRM_1$. When $L_1M_1$ is moved through a distance dx in time dt, the new area is $L_2QRM_2$. Due to the change in area $L_2L_1M_1M_2$, there is a change in the flux linked with the conductor. Therefore, an induced emf is produced.

Change in area dA = Area $L_2L_1M_1M_2$




∴ dA = $l$dx

Change in the magnetic flux, d$\phi$ = B.dA = B $l$dx

e = $-\frac{d\phi}{dt}$

$\therefore$ e = $-\frac{B\,l dx}{dt}$ =B $l\ v$

where $v$ is the velocity with which the sliding conductor is moved.

**Emf induced by changing the orientation of the coil**

PQRS is a rectangular coil of N turns and area A placed in a uniform magnetic field B (Fig.). The coil is rotated with an angular velocity $\omega$ in the clockwise direction about an axis perpendicular to the direction of the magnetic field. Suppose, initially the coil is in vertical position, so that the angle between normal to the plane of the coil and magnetic field is zero. After a time t, let $\theta$ $(=\omega t)$ be the angle through which the coil is rotated. If $\phi$ is the flux linked with the coil at this instant, then

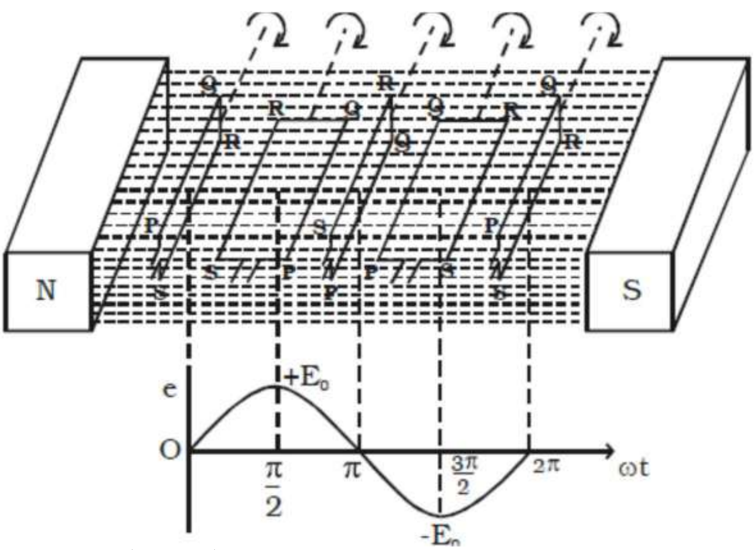
$$\phi = NBA\cos\theta$$

The induced emf is, e = $-\frac{d\phi}{dt}$ = -NBA$\frac{d}{dt}\cos\omega t$

$\therefore\ e = NBA\omega\sin\omega t$ ... (1)

The maximum value of the induced emf is, $E_0$ = NAB$\omega$

Hence, the induced emf can be represented as e = $E_0$ sin$\omega$t. The induced emf e varies sinusoidally with time t and the frequency being v cycles




per second $\left(v = \frac{\omega}{2\pi}\right)$

(i) When $\omega$t = 0, the plane of the coil is perpendicular to the field B and hence e = 0.

(ii) When $\omega$t = $\pi/2$, the plane of the coil is parallel to B and hence e = $E_0$

(iii) When $\omega$t = $\pi$, the plane of the coil is at right angle to B and hence e = 0.

(iv) When $\omega t = 3\pi/2$, the plane of the coil is again parallel to B and the induced emf is $e = -E_0$.
(v) When $\omega t = 2\pi$, the plane of the coil is again perpendicular to B and hence $e = 0$.
If the ends of the coil are connected to an external circuit through a resistance R, current flows through the circuit, which is also sinusoidal in nature.

**AC generator (Dynamo) - Single phase**

The ac generator is a device used for converting mechanical energy into electrical energy. The generator was originally designed by a Yugoslav scientist Nikola Tesla.

**Principle**

It is based on the principle of electromagnetic induction, according to which an emf is induced in a coil when it is rotated in a uniform magnetic field.

Essential parts of an AC generator

(i) Armature

Armature is a rectangular coil consisting of a large number of loops or turns of insulated copper wire wound over a laminated soft iron core or ring. The soft iron core not only increases the magnetic flux but also serves as a support for the coil

(ii) Field magnets

The necessary magnetic field is provided by permanent magnets in the case of low power dynamos. For high power dynamos, field is provided by electro magnet. Armature rotates between the magnetic poles such that the axis of rotation is perpendicular to the magnetic field.

(iii) Slip rings

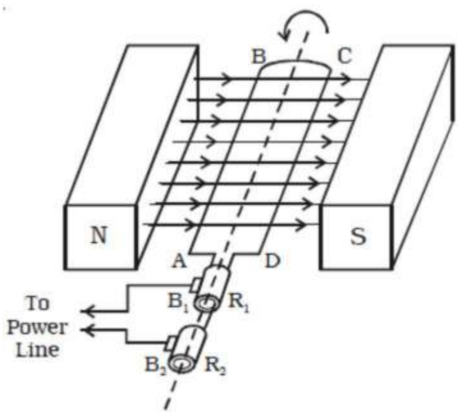
The ends of the armature coil are connected to two hollow metallic rings R1 and R2 called slip rings. These rings are fixed to a shaft, to which the armature is also fixed. When the shaft rotates, the slip rings along with the armature also rotate.

(iv) Brushes

B1 and B2 are two flexible metallic plates or carbon brushes. They provide contact with the slip rings by keeping themselves pressed against the ring. They are used to pass on the current from the armature to the external power line through the slip rings.

Working

Whenever, there is a change in orientation of the coil, the magnetic flux linked with the coil changes, producing an induced emf in the coil. The direction of the induced current is given by Fleming's right hand rule. Suppose the armature ABCD is initially in the vertical position. It is rotated in the anticlockwise direction. The side AB of the coil moves downwards and the side DC moves upwards (Fig.). Then according to Flemings right hand rule the current induced in arm AB flows from B to A and in CD it flows from D to C. Thus the current flows along DCBA in the coil. In the external circuit the current flows from $B_1$ to $B_2$.




On further rotation, the arm AB of the coil moves upwards and DC moves downwards. Now the current in the coil flows along ABCD. In the external circuit the current flows from $B_2$ to $B_1$. As the rotation of the coil continues, the induced current in the external circuit keeps changing its direction for every half a rotation of the coil. Hence the induced current is alternating in nature (Fig.). As the armature completes ν rotations in one second, alternating current of frequency ν cycles per second is produced. The induced emf at any instant is given by $e= E_0 \sin \omega t$

The peak value of the emf, $E_0 = NBA$

where N is the number of turns of the coil,

A is the area enclosed by the coil,

B is the magnetic field and

$\omega$ is the angular velocity of the coil

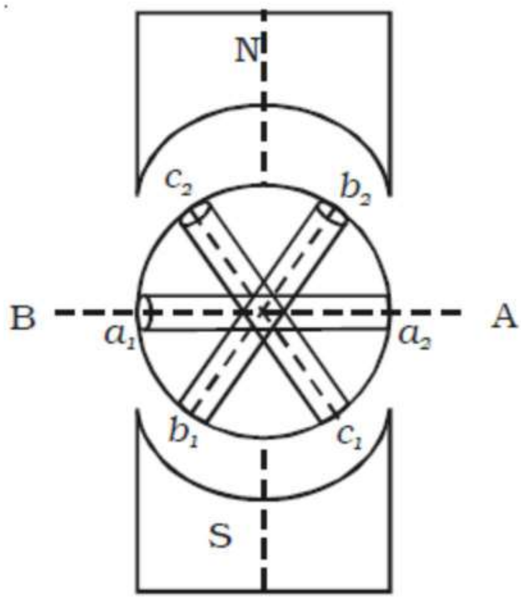
**AC generator (Alternator) – Three phase**

A single phase a.c. generator or alternator has only one armature winding. If a number of armature windings are used in the alternator it is known as polyphase alternator. It produces voltage waves equal to the number of windings or phases. Thus a polyphase system consists of a numerous windings which are placed on the same axis but displaced from one another by equal angle which depends on the number of phases. Three phase alternators are widely preferred for transmitting large amount of power with less cost and high efficiency.

**Generation of three phase emf**

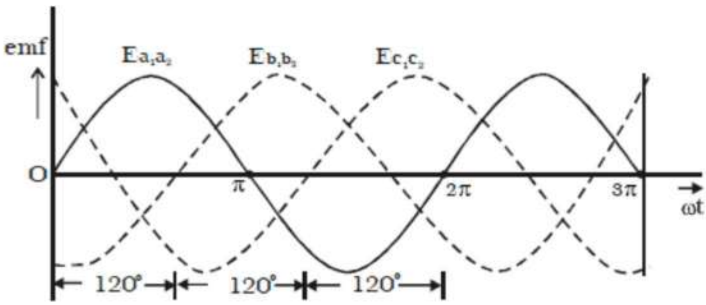
In a three - phase a.c. generator three coils are fastened rigidly together and displaced from each other by 120°. It is made to rotate about a fixed axis in a uniform magnetic field. Each coil is provided with a separate set of slip rings and brushes.

An emf is induced in each of the coils with a phase difference of 120°. Three coils $a_1$ $a_2$, $b_1$ $b_2$ and $c_1$ $c_2$ are mounted on the same axis but

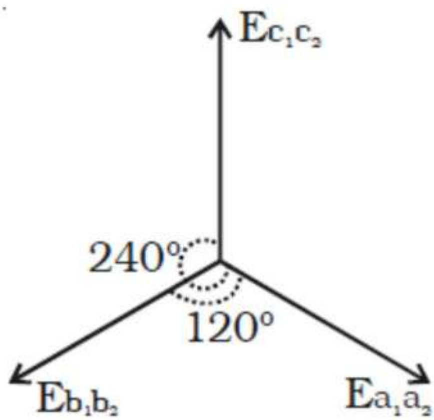



displaced from each other by 120°, and the coils rotate in the anticlockwise direction in a magnetic field (Fig.a).

When the coil $a_1a_2$ is in position AB, emf induced in this coil is zero and starts increasing in the positive direction. At the same instant the coil $b_1b_2$ is 120° behind coil $a_1$ $a_2$, so that emf induced in this coil is approaching its maximum negative value and the coil $c_1$ $c_2$ is 240° behind the coil $a_1$ $a_2$, so the emf induced in this coil has passed its positive maximum value and is decreasing. Thus the emfs induced in all the three coils are equal in magnitude and of same frequency. The emfs induced in the three coils are;

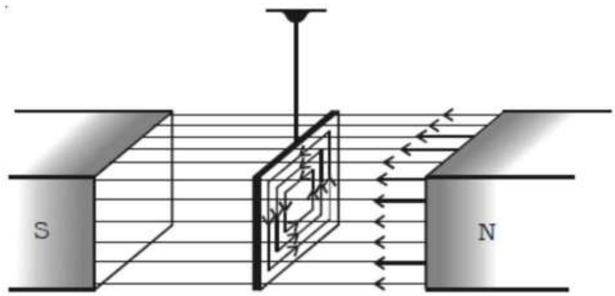



$e_{a_1a_2}$ = $E_0$ sin $\omega$t
$e_{b_1b_2}$= $E_0$ sin ($\omega$t – 2$\pi$/3)
$e_{c_1c_2}$= Eo sin ($\omega$t – 4$\pi$/3)
The emfs induced and phase difference in the three coils $a_1$ $a_2$, $b_1$ $b_2$ and $c_1$ $c_2$ are shown in Fig b & Fig c.



**Eddy currents**

Foucault in the year 1895 observed that when a mass of metal moves in a magnetic field or when the magnetic field through a stationary mass of metal is altered, induced current is produced in the metal. This




induced current flows in the metal in the form of closed loops resembling 'eddies' or whirl pool. Hence this current is called eddy current. The direction of the eddy current is given by Lenz's law. When a conductor in the form of a disc or a metallic plate as shown in Fig. , swings between the poles of a magnet, eddy currents are set up inside the plate. This current acts in a direction so as to oppose the motion of the conductor with a strong retarding force, that the conductor almost comes to rest. If the metallic plate with holes drilled in it is made to swing inside the magnetic field, the effect of eddy current is greatly reduced consequently the plate swings freely inside the field. Eddy current can be minimised by using thin laminated sheets instead of solid metal.

**Applications of Eddy current**

(i) Dead beat galvanometer when current is passed through a galvanometer, the coil oscillates about its mean position before it comes to rest. To bring the coil to rest immediately, the coil is wound on a metallic frame. Now, when the coil oscillates, eddy currents are set up in the metallic frame, which opposes further oscillations of the coil. This in turn enables the coil to attain its equilibrium position almost instantly.

Since the oscillations of the coil die out instantaneously, the galvanometer is called dead beat galvanometer.

(ii) Induction furnace

In an induction furnace, high temperature is produced by generating eddy currents. The material to be melted is placed in a varying magnetic field of high frequency. Hence a strong eddy current is developed inside the metal. Due to the heating effect of the current, the metal melts.

(iii) Induction motors

Eddy currents are produced in a metallic cylinder called rotor, when it is placed in a rotating magnetic field. The eddy current initially tries to decrease the relative motion between the cylinder and the rotating magnetic field. As the magnetic field continues to rotate, the metallic cylinder is set into rotation. These motors are used in fans.

(iv) Electromagnetic brakes

A metallic drum is coupled to the wheels of a train. The drum rotates along with the wheel when the train is in motion. When the brake is applied, a strong magnetic field is developed and hence, eddy currents are produced in the drum which opposes the motion of the drum. Hence, the train comes to rest.

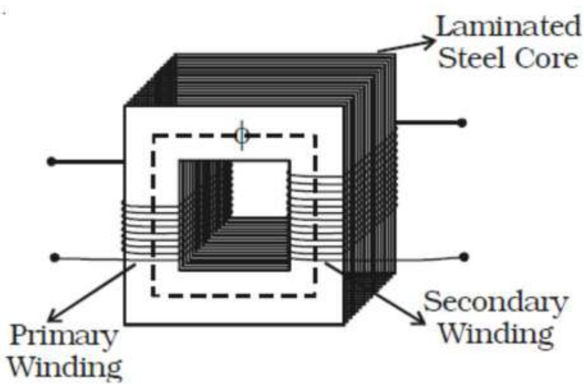
(v) Speedometer

In a speedometer, a magnet rotates according to the speed of the vehicle. The magnet rotates inside an aluminium cylinder (drum) which is held in position with the help of hair springs. Eddy currents are produced in the drum due to the rotation of the magnet and it opposes the motion of the rotating magnet. The drum inturn experiences a torque and gets deflected through a certain angle depending on the speed of the vehicle. A pointer attached to the drum moves over a calibrated scale which indicates the speed of the vehicle.

**Transformer**

Transformer is an electrical device used for converting low alternating voltage into high alternating voltage and vice versa. It transfers electric power from one circuit to another. The transformer is based on the principle of electromagnetic induction.

A transformer consists of primary and secondary coils insulated from




each other, wound on a soft iron core (Fig.). To minimise eddy currents a laminated iron core is used. The a.c. input is applied across the primary coil. The continuously varying current in the primary coil produces a varying magnetic flux in the primary coil, which in turn produces a varying magnetic flux in the secondary. Hence, an induced emf is produced across the secondary. Let $E_P$ and $E_S$ be the induced emf in the primary and secondary coils and $N_P$ and $N_S$ be the number of turns in the primary and secondary coils respectively. Since same flux links with the primary and secondary, the emf induced per turn of the two coils must be the same

(i.e) $\frac{E_P}{N_P} = \frac{E_S}{N_S}$

$\frac{E_S}{E_P} = \frac{N_S}{N_P}$ … (1)

For an ideal transformer, input power = output power

$E_p I_p = E_S I_S$

where $I_p$ and $I_s$ are currents in the primary and secondary coils.

(i.e) $\frac{E_S}{E_P} = \frac{I_P}{I_S}$ … (2)

From equations (1) and (2)

$\frac{E_S}{E_P} = \frac{N_S}{N_P} = \frac{I_P}{I_S} = K$

where k is called transformer ratio.

(for step up transformer k > 1 and for step down transformer k < 1)

In a step up transformer $E_s > E_p$ implying that $I_s < I_p$. Thus a step up transformer increases the voltage by decreasing the current, which is in accordance with the law of conservation of energy. Similarly a step down transformer decreases the voltage by increasing the current.

**Efficiency of a transformer**

Efficiency of a transformer is defined as the ratio of output power to the input power.

$$\eta = \frac{output\ power}{input\ power} = \frac{E_S I_S}{E_P I_P}$$

The efficiency $\eta$ = 1(ie. 100%), only for an ideal transformer where there is no power loss. But practically there are numerous factors leading to energy loss in a transformer and hence the efficiency is always less than one.

**Energy losses in a transformer**

(1) Hysteresis loss

The repeated magnetisation and demagnetisation of the iron core caused by the alternating input current, produces loss in energy called hysterisis loss. This loss can be minimised by using a core with a material having the least hysterisis loss. Alloys like mumetal and silicon steel are used to reduce hysterisis loss.

(2) Copper loss

The current flowing through the primary and secondary windings lead to Joule heating effect. Hence some energy is lost in the form of heat. Thick wires with considerably low resistance are used to minimize this loss.

(3) Eddy current loss (Iron loss)

The varying magnetic flux produces eddy current in the core. This leads to the wastage of energy in the form of heat. This loss is minimised by using a laminated core made of stelloy, an alloy of steel.

(4) Flux loss

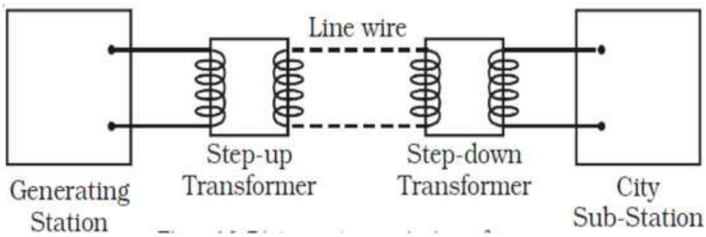
The flux produced in the primary coil is not completely linked with the secondary coil due to leakage. This results in the loss of energy. This loss can be minimised by using a shell type core.

In addition to the above losses, due to the vibration of the core, sound is produced, which causes a loss in the energy.

**Long distance power transmission**

The electric power generated in a power station situated in a remote place is transmitted to different regions for domestic and industrial use. For long distance transmission, power lines are made of conducting material like aluminium. There is always some power loss associated with these lines.

If I is the current through the wire and R the resistance, a considerable amount of electric power $I_2R$ is dissipated as heat. Hence, the power at the receiving end will be much lesser than the actual power generated. However, by transmitting the electrical energy at a higher voltage, the




power loss can be controlled as is evident from the following two cases.

Case (i) A power of 11,000 W is transmitted at 220 V.

Power P = VI

$\therefore I = \frac{P}{V} = \frac{11000}{220} = 50A$

If R is the resistance of line wires,

Power loss = $I^2R = 50^2R = 2500(R)$ watts

Case (ii) 11,000 W power is transmitted at 22,000 V

$\therefore I = \frac{P}{V} = \frac{11000}{220} = 50A$

Power loss = $I^2R = (0.5)^2R = 0.25(R)$ watts

Hence it is evident that if power is transmitted at a higher voltage the loss of energy in the form of heat can be considerably reduced. For transmitting electric power at 11,000 W at 220 V the current capacity of line wires has to be 50 A and if transmission is done at 22,000 V, it is only 0.5 A. Thus, for carrying larger current (50A) thick wires have to be used. This increases the cost of transmission. To support these thick wires, stronger poles have to be erected which further adds on to the cost. On the other hand if transmission is done at high voltages, the wires required are of lower current carrying capacity. So thicker wires can be replaced by thin wires, thus reducing the cost of transmission considerably.

For example, 400MW power produced at 15,000 V in the power station at Neyveli, is stepped up by a step-up transformer to 230,000 V before transmission. The power is then transmitted through the transmission lines which forms a part of the grid. The grid connects different parts of the country. Outside the city, the power is stepped down to 110,000 V by a step-down transformer. Again the power is stepped down to 11,000 V by a transformer. Before distribution to the user, the power is stepped down to 230 V or 440 V depending upon the need of the user.

**Alternating current**

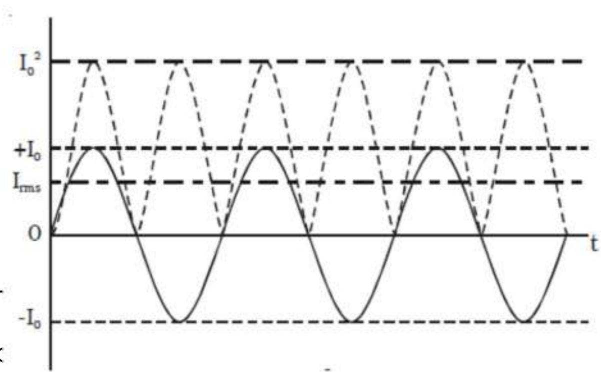
As we have seen earlier a rotating coil in a magnetic field, induces an alternating emf and hence an alternating current. Since the emf induced in the coil varies in magnitude and direction periodically, it is called an alternating emf. The significance of an alternating emf is that it can be changed to lower or higher voltages conveniently and efficiently using a transformer. Also the frequency of the induced emf can be altered by changing the speed of the coil. This enables us to utilize the whole range of electromagnetic spectrum for one purpose or the other. For example domestic power in India is supplied at a frequency of 50 Hz. For transmission of audio and video signals, the required frequency range of radio waves is between 100 KHz and 100 MHz. Thus owing to its wide applicability most of the countries in the world use alternating current.

**Measurement of AC**

Since alternating current varies continuously with time, its average value over one complete cycle is zero. Hence its effect is measured by rms value of a.c.

**RMS value of a.c.**

The rms value of alternating current is defined as that value of the steady current, which when passed through a resistor for a given time, will generate the same amount of heat as generated by an alternating current when passed through the same resistor for the same time. The rms value is also called effective value of an a.c. and is denoted by $I_{rms}$ or $I_{eff}$. when an alternating current $i=I_0 \sin \omega t$ flows through a resistor of resistance R, the amount of heat produced in the resistor in a small time



dt is

dH = $i^2$ R dt

The total amount of heat produced in the resistance in one complete cycle is

$$H = \int_0^T i^2 R dt = \int_0^T (I_0^2 sin^2 \omega t)\, R dt$$

$$= I_0^2 R \int_0^T \left(\frac{1 - \cos 2\omega t}{2}\right) dt = \frac{I_0^2 R}{2}\left[\int_0^T dt - \int_0^T cos 2\omega t . dt\right]$$

$$= \frac{I_0^2 R}{2}\left[t - \frac{sin 2\omega t}{2\omega}\right]_0^T = \frac{I_0^2 R}{2}\left[\int_0^T dt - \int_0^T \cos 2\omega t . dt\right] \qquad \because T = \frac{2\pi}{\omega}$$

$$H = \frac{I_0^2 RT}{2}$$

But this heat is also equal to the heat produced by rms value of AC in the same resistor (R) and in the same time (T),

(i.e) H = $I^2_{rms}$ RT

$$\therefore I_{rms}^2 RT = \frac{I_0^2}{2}$$

$$I_{rms} = \frac{I_0}{\sqrt{2}} = 0.707 I_0$$

Similarly, it can be calculated that

$$E_{rms} = \frac{E_0}{\sqrt{2}}$$

Thus, the rms value of an a.c is 0.707 times the peak value of the a.c. In other words it is 70.7 % of the peak value.

**AC Circuit with resistor**

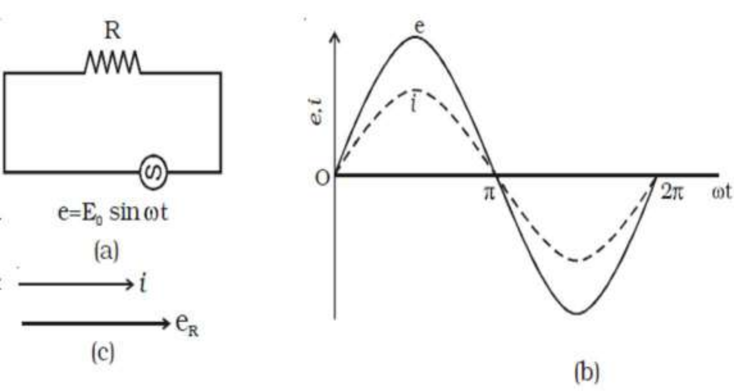
Let an alternating source of emf be connected across a resistor of resistance R.

The instantaneous value of the applied emf is

$e = E_o \sin \omega t$ ... (1)

If i is the current through the circuit at the instant t, the potential drop across R is, e = i R

Potential drop must be equal to the applied emf.



Hence, $iR = E_0 \sin \omega t$

$i\frac{E_0}{R} \sin \omega t$; $i = I_0 \sin \omega t$ ...(2)

where $I_0 = \frac{E_0}{R}$, is the peak value of a.c in the circuit. Equation (2) gives the instantaneous value of current in the circuit containing R. From the expressions of voltage and current given by equations (1) and (2) it is evident that in a resistive circuit, the applied voltage and current are in phase with each other (Fig.b). Fig.c is the phasor diagram representing the phase relationship between the current and the voltage.

**AC Circuit with an inductor**

Let an alternating source of emf be applied to a pure inductor of inductance L. The inductor has a negligible resistance and is wound on a laminated iron core. Due to an alternating emf that is applied to the inductive coil, a self-induced emf is generated which opposes the applied voltage. (eg) Choke coil.

The instantaneous value of applied emf is given by

$e = E_0 \sin\omega t$ ... (1)

Induced emf $e' = -L.\frac{di}{dt}$

where L is the self-inductance of the coil. In an ideal inductor circuit induced emf is equal and opposite to the applied voltage.

Therefore $e = -e'$

$E_0 \sin\omega t = -\left(-L.\frac{di}{dt}\right)$

$\therefore E_0 \sin\omega t = L.\frac{di}{dt}$

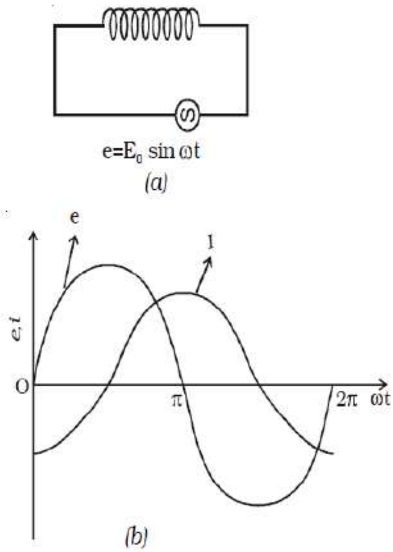
$di = \frac{E_0}{L} \int sin\omega t \; dt$

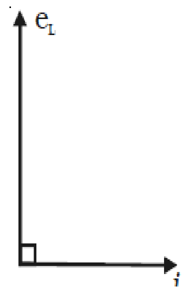
$= \frac{E_0}{L}\left[-\frac{\cos \omega t}{\omega}\right] = \frac{E\cos \omega t}{\omega L}$

$i = \frac{E_0}{\omega L} \sin \omega t \;\; \frac{\pi}{t}$ ... (2)

where $I_0 = \frac{E_0}{\omega L}$. Here, ωL is the resistance offered by the coil. It is called inductive reactance. Its unit is ohm.

From equations (1) and (2) it is clear that in an a.c. circuit containing a pure inductor the current i lags behind the voltage e by the phase angle of $\pi/2$. . Conversely the voltage across L leads the current by the phase angle of $\pi/2$. This fact is presented graphically in Fig. b. Fig. c represents the phasor diagram of a.c. circuit containing only L.




**Inductive reactance**

XL = $\omega L = 2\pi v L$, where $v$ is the frequency of the a.c. supply For d.c. $v$ = 0; ∴ XL = 0

Thus a pure inductor offers zero resistance to d.c. But in an a.c. circuit the reactance of the coil increases with increase in frequency.

**AC Circuit with a capacitor**

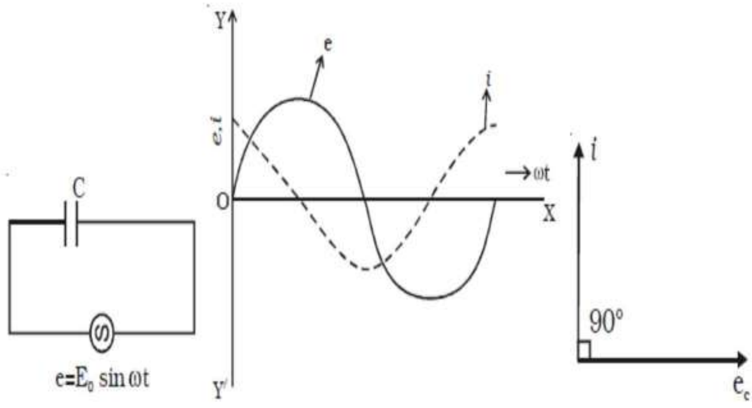
An alternating source of emf is connected across a capacitor of capacitance C (Fig. a). It is charged first in one direction and then in the other direction.

The instantaneous value of the applied emf is given by

$e = E_0 \sin \omega t$ ... (1)

At any instant the potential difference across the capacitor will be equal to the applied emf

∴ e = q/C, where q is the charge in the capacitor




(a) (b) (c)

But, i = $\frac{dq}{dt} = \frac{d}{dt}(Ce)$

i = $\frac{d}{dt}(C\ E_0\ \sin\ \omega t\ ) = \omega\ C\ E_0\ \sin\ \omega t$

i = $\frac{E_0}{(1/_{\omega C})}\sin\left(\omega t + \frac{\pi}{2}\right)$

i = $I_0\ \sin\left(\omega t + \frac{\pi}{2}\right)$ … (2)

$$I_0 = \frac{E_0}{(1/_{\omega C})}$$

$\left(\frac{1}{\omega C}\right) = X_C$ is is the resistance offered by the capacitor. It is called capacitive reactance. Its unit is ohm. From equations (1) and (2), it follows that in an a.c. circuit with a capacitor, the current leads the voltage by a phase angle of $\pi/2$. In otherwords the emf lags behind the current by a phase angle of $\pi/2$. This is represented graphically in Fig. b. Fig. c represents the phasor diagram of a.c. circuit containing only C.

$$X_C = \left(\frac{1}{\omega C}\right) = \frac{1}{2\pi v C}$$

where $v$ is the frequency of the a.c. supply. In a d.c. circuit $v = 0$

∴ $X_C = \infty$

Thus a capacitor offers infinite resistance to d.c. For an a.c. the capacitive reactance varies inversely as the frequency of a.c. and also inversely as the capacitance of the capacitor.

**Resistor, inductor and capacitor in series**

Let an alternating source of emf e be connected to a series combination of a resistor of resistance R, inductor of inductance L and a capacitor of capacitance C (Fig. a).

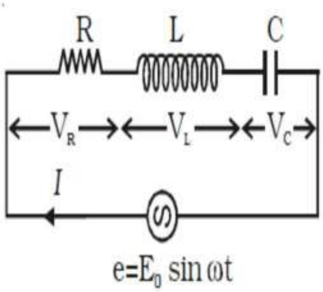
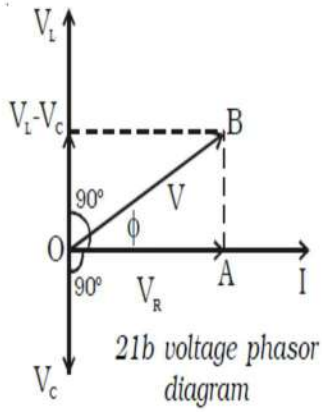


Fig 21a RLC sereis circuit



21b voltage phasor diagram

Let the current flowing through the circuit be I. The voltage drop across the resistor is, $V_R = I R$ (This is in phase with I)

The voltage across the inductor coil is $V_L = I X_L$

($V_L$ leads I by $\pi/2$)

The voltage across the capacitor is, $V_C = IX_C$

($V_C$ lags behind I by $\pi/2$)

The voltages across the different components are represented in the voltage phasor diagram (Fig. b).

$V_L$ and $V_C$ are 180° out of phase with each other and the resultant of $V_L$ and $V_C$ is ($V_L - V_C$), assuming the circuit to be predominantly inductive. The applied voltage 'V' equals the vector sum of $V_R$, $V_L$ and $V_C$.

$OB_2 = OA_2 + AB_2$;

$V^2 = V_R^2 + (V_L - V_C)^2$

$V = \sqrt{V_R^2 + (V_L - V_C)^2}$

$V = \sqrt{(IR)^2 + (IX_L - IX_C)^2}$

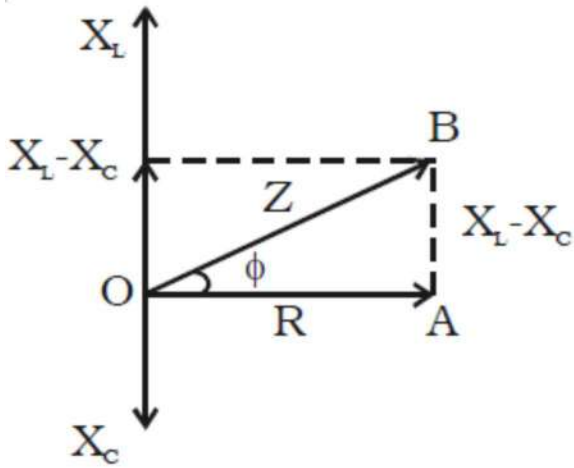
$= I\sqrt{(R)^2 + (X_L - X_C)^2}$

$\frac{V}{I} = Z = \sqrt{(R)^2 + (X_L - X_C)^2}$

The expression $\sqrt{(R)^2 + (X_L - X_C)^2}$ is the net effective opposition offered by the combination of resistor, inductor and capacitor known as the impedance of the circuit and is represented by Z. Its unit is ohm. The values are represented in the impedance diagram (Fig.). Phase angle $\phi$ between the voltage and current is given by

$\tan\phi = \frac{V_L - V_C}{V_R} = \frac{IX_L - IX_C}{IR}$

$\tan\phi = \frac{X_L - X_C}{R} = \frac{net\ reactance}{resistance}$




$\therefore \phi = \tan^{-1}\left(\frac{X_L - X_C}{R}\right)$

∴ Io sin ($\omega$t +$\phi$) is the instantaneous current flowing in the circuit.

**Series resonance or voltage resonance in RLC circuit**

The value of current at any instant in a series RLC circuit is given by

$I = \frac{I}{Z} = \frac{V}{\sqrt{(R)^2 + (X_L - X_C)^2}} = \frac{V}{\sqrt{(R)^2 + \left(\omega_L - \frac{1}{\omega_L}\right)^2}}$

At a particular value of the angular frequency, the inductive reactance and the capacitive reactance will be equal to each other (i.e.) $\omega$L = $\frac{1}{\omega_L}$, so that the impedance becomes minimum and it is given by Z = R

i.e. I is in phase with V

The particular frequency $v_0$ at which the impedance of the circuit becomes minimum and therefore the current becomes maximum is called Resonant frequency of the circuit. Such a circuit which admits maximum current is called series resonant circuit or acceptor circuit. Thus the maximum current through the circuit at resonance is

$I_0 = \frac{V}{R}$

Maximum current flows through the circuit, since the impedance of the circuit is merely equal to the ohmic resistance of the circuit. i.e Z = R

$\omega_L = \frac{1}{\omega C}$

$$\omega = 2\pi v_0 = \frac{1}{\sqrt{LC}}$$

$$v_0 = \frac{1}{2\pi\sqrt{LC}}$$

**Acceptor circuit**

The series resonant circuit is often called an 'acceptor' circuit. By offering minimum impedance to current at the resonant frequency it is able to select or accept most readily this particular frequency among many frequencies. In radio receivers the resonant frequency of the circuit is tuned to the frequency of the signal desired to be detected. This is usually done by varying the capacitance of a capacitor.

**Q-factor**

The selectivity or sharpness of a resonant circuit is measured by the quality factor or Q factor. In other words it refers to the sharpness of tuning at resonance. The Q factor of a series resonant circuit is defined as the ratio of the voltage across a coil or capacitor to the applied voltage.

$Q = \frac{\text{voltage across L or C}}{\text{applied voltage}}$ ... (1)

Voltage across L = $I\omega_0 L$ ... (2)

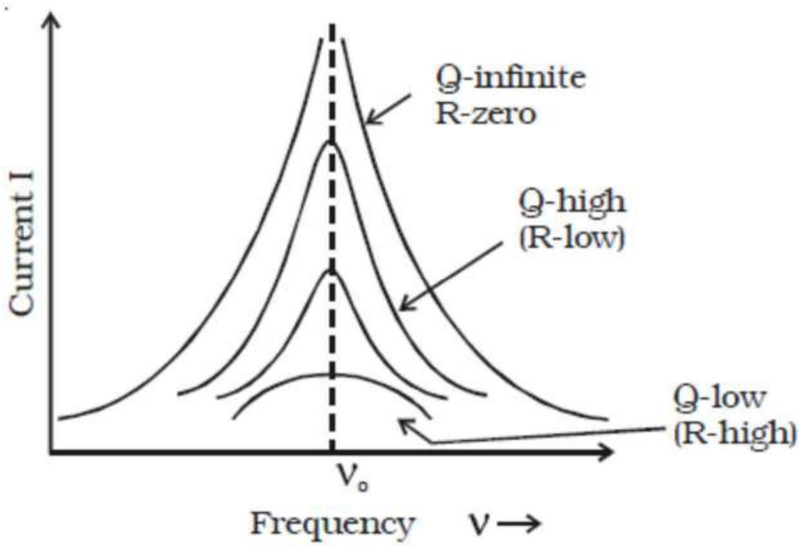
where $\omega_0$ is the angular frequency of the a.c. at resonance. The applied voltage at resonance is the potential drop across R, because the potential drop across L is equal to the drop across C and they are 180° out of phase. Therefore they cancel out and only potential drop across R will exist.

Applied Voltage = IR ... (3)

Substituting equations (2) and (3) in equation (1)

$$Q = \frac{I\omega_0 L}{IR} = \frac{\omega_0 L}{R}$$

$$Q = \frac{1}{\sqrt{LC}}\frac{L}{R} = \frac{1}{R}\sqrt{\frac{L}{C}} \qquad \left\{\because \omega_0 = \frac{1}{\sqrt{LC}}\right\}$$




Q is just a number having values between 10 to 100 for normal frequencies. Circuit with high Q values would respond to a very narrow frequency range and vice versa. Thus a circuit with a high Q value is sharply tuned while one with a low Q has a flat resonance. Q-factor can be increased by having a coil of large inductance but of small ohmic resistance.

Current frequency curve is quite flat for large values of resistance and becomes more sharp as the value of resistance decreases. The curve shown in Fig. is also called the frequency response curve.

**Power in an ac circuit**

In an a.c circuit the current and emf vary continuously with time. Therefore power at a given instant of time is calculated and then its mean is taken over a complete cycle. Thus, we define instantaneous power of an a.c. circuit as the product of the instantaneous emf and the instantaneous current flowing through it.

The instantaneous value of emf and current is given by

e = $E_0 \sin \omega t$

i = $I_0 \sin(\omega t + \phi)$

where$\phi$ is the phase difference between the emf and current in an a.c circuit

The average power consumed over one complete cycle is

$$P_{av}=\frac{\int_0^T iedt}{\int_0^T dt}=\frac{\int_0^T [I_0 \sin(\omega t+\phi)E_0 \sin \omega t]dt}{T}$$

On simplification, we obtain

$$P_{av} = \frac{E_0 I_0}{2}\cos\phi$$

$$P_{av} = \frac{E_0}{\sqrt{2}}.\frac{I_0}{\sqrt{2}}\cos\phi = E_{rms}I_{rms}\cos\phi$$

$P_{av}$ = apparent power × power factor

where Apparent power = $E_{rms}I_{rms}$ and power factor = $\cos\phi$

The average power of an ac circuit is also called the true power of the circuit.

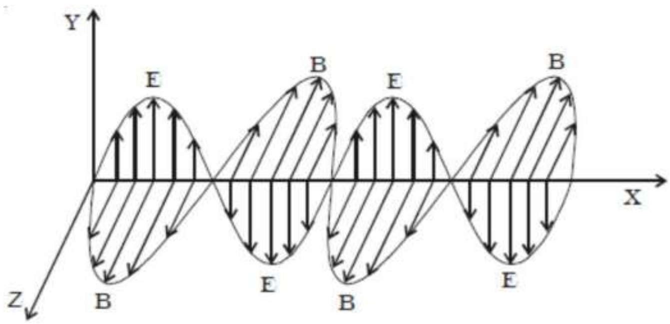
**Electromagnetic Waves**

The phenomenon of Faraday's electromagnetic induction concludes that a changing magnetic field at a point with time produces an electric field at that point. Maxwell in 1865, pointed out that there is symmetry in nature (i.e) changing electric field with time at a point produces a magnetic field at that point. It means that a change in one field with time (either electric or magnetic) produces another field. This idea led Maxwell to conclude that the variation in electric and magnetic fields perpendicular to each other produces electromagnetic disturbances in space. These disturbances have the properties of a wave and propagate through space without any material medium. These waves are called electromagnetic waves.

Electromagnetic waves

According to Maxwell, an accelerated charge is a source of electromagnetic radiation. In an electromagnetic wave, electric and magnetic field vectors are at right angles to each other and both are at right angles to the direction of propagation. They possess the wave character and propagate through free space without any material medium. These waves are transverse in nature. Fig. shows the variation of electric field $\vec{E}$ along Y direction and magnetic field $\vec{B}$along Z direction and wave propagation in +X direction.

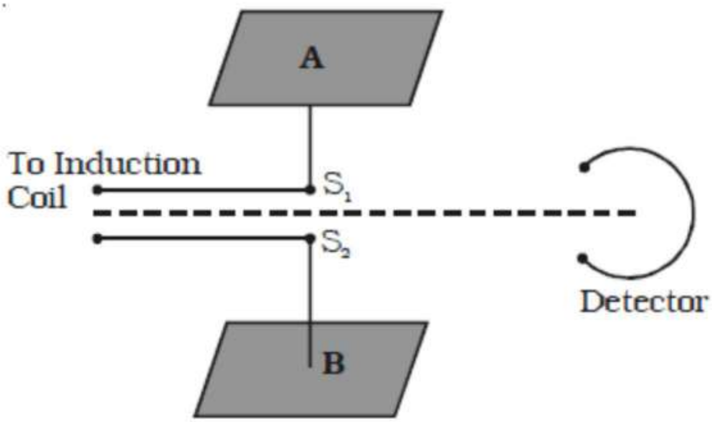
Characteristics of electromagnetic waves




- Electromagnetic waves are produced by accelerated charges.
- They do not require any material medium for propagation.
- In an electromagnetic wave, the electric ($\vec{E}$) and magnetic ($\vec{B}$) field vectors are at right angles to each other and to the direction of propagation. Hence electromagnetic waves are transverse in nature.
- Variation of maxima and minima in both $\vec{E}$ and $\vec{B}$ occur simultaneously.
- They travel in vacuum or free space with a velocity $3 \times 10^8$ $m^{-1}$ given by the relation $C = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$.
- ($\mu_0$-permeability of free space and $\varepsilon_0$ - permittivity of free space)
- The energy in an electromagnetic wave is equally divided between electric and magnetic field vectors.
- The electromagnetic waves being chargeless, are not deflected by electric and magnetic fields.

**Hertz experiment**

The existence of electromagnetic waves was confirmed experimentally by Hertz in 1888. This experiment is based on the fact that an oscillating electric charge radiates electromagnetic waves. The energy of these




waves is due to the kinetic energy of the oscillating charge.

The experimental arrangement is as shown in Fig. It consists of two metal plates A and B placed at a distance of 60 cm from each other. The metal plates are connected to two polished metal spheres $S_1$ and $S_2$ by means of thick copper wires. Using an induction coil a high potential difference is applied across the small gap between the spheres. Due to high potential difference across $S_1$ and $S_2$, the air in the small gap between the spheres gets ionized and provides a path for the discharge of the plates. A spark is produced between Hertz was able to produce electromagnetic waves of frequency about $5 \times 10^7$ Hz.

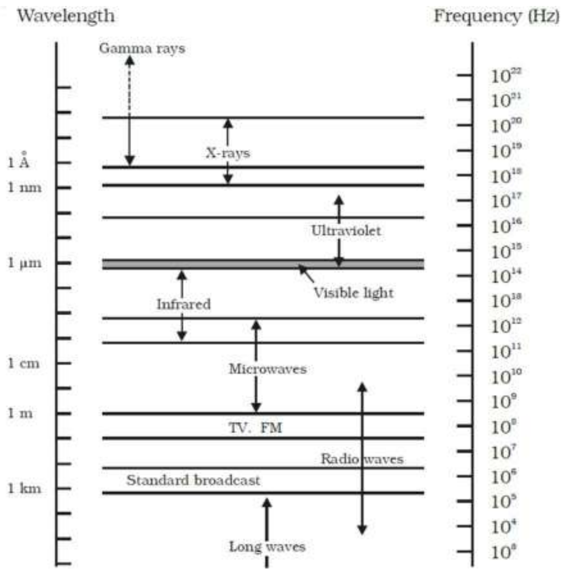
Here the plates A and B act as a capacitor having small capacitance value C and the connecting wires provide low inductance L. The high frequency oscillation of charges between the plates is given by $V = \frac{1}{2\pi\sqrt{LC}}$

**Electromagnetic Spectrum**

After the demonstration of electromagnetic waves by Hertz, electromagnetic waves in different regions of wavelength were produced by different ways of excitation.

The orderly distribution of electromagnetic waves according to their wavelength or frequency is called the electromagnetic spectrum. Electromagnetic spectrum covers a wide range of wavelengths (or) frequencies. The whole electromagnetic spectrum has been classified into different parts and sub parts, in order of increasing wavelength and type of excitation. All electromagnetic waves travel with the velocity of light. The physical properties of electromagnetic waves are determined by their wavelength and not by their method of excitation.

The overlapping in certain parts of the spectrum shows that the particular wave can be produced by different methods. Table shows




various regions of electromagnetic spectrum with source, wavelength and frequency ranges of different electromagnetic waves.

**Uses of electromagnetic spectrum**

The following are some of the uses of electromagnetic waves.

1. Radio waves: These waves are used in radio and television communication systems. AM band is from 530 kHz to 1710 kHz. Higher frequencies upto 54 MHz are used for short waves bands.

Television waves range from 54 MHz to 890 MHz. FM band is from 88 MHz to 108 MHz. Cellular phones use radio waves in ultra-high frequency (UHF) band.

2. Microwaves: Due to their short wavelengths, they are used in radar communication system. Microwave ovens are an interesting domestic application of these waves.

3. Infrared waves:

- Infrared lamps are used in physiotherapy.
- Infrared photographs are used in weather forecasting.
- As infrared radiations are not absorbed by air, thick fog, mist etc, they are used to take photograph of long distance objects.
- Infra-red absorption spectrum is used to study the molecular structure.

4. Visible light: Visible light emitted or reflected from objects around us provides information about the world. The wavelength range of visible light is 4000 Å to 8000 Å.

5. Ultra-violet radiations

- They are used to destroy the bacteria and for sterilizing surgical instruments.
- These radiations are used in detection of forged documents, finger prints in forensic laboratories.
- They are used to preserve the food items.
- They help to find the structure of atoms.

6. X rays:

- X rays are used as a diagonistic tool in medicine.
- It is used to study the crystal structure in solids.

7. $\gamma$-rays: Study of $\gamma$ rays gives useful information about the nuclear structure and it is used for treatment of cancer.

| Sl.No. | Name | Source | Wavelength range (m) | Frequency range (Hz) |
|---|---|---|---|---|
| 1. | $\gamma$ – rays | Radioactive nuclei, nuclear reactions | $10^{-14}$ – $10^{-10}$ | $3 \times 10^{22}$ – $3 \times 10^{18}$ |
| 2. | x – rays | High energy electrons suddenly stopped by a metal target | $1 \times 10^{-10}$–$3 \times 10^{-8}$ | $3 \times 10^{18}$ – $1 \times 10^{16}$ |
| 3. | Ultra-violet (UV) | Atoms and molecules in an electrical discharge | $6 \times 10^{-10}$–$4 \times 10^{-7}$ | $5 \times 10^{17}$ – $8 \times 10^{14}$ |
| 4. | Visible light | incandescent solids Fluorescent lamps | $4 \times 10^{-7}$ – $8 \times 10^{-7}$ | $8 \times 10^{14}$ – $4 \times 10^{14}$ |
| 5. | Infra-red (IR) | molecules of hot bodies | $8 \times 10^{-7}$ – $3 \times 10^{-5}$ | $4 \times 10^{14}$ – $1 \times 10^{13}$ |
| 6. | Microwaves | Electronic device (Vacuum tube) | $10^{-3}$ – 0.3 | $3 \times 10^{11}$ – $1 \times 10^{9}$ |
| 7. | Radio frequency waves | charges accelerated through conducting wires | $10$–$10^{4}$ | $3 \times 10^{7}$ – $3 \times 10^{4}$ |

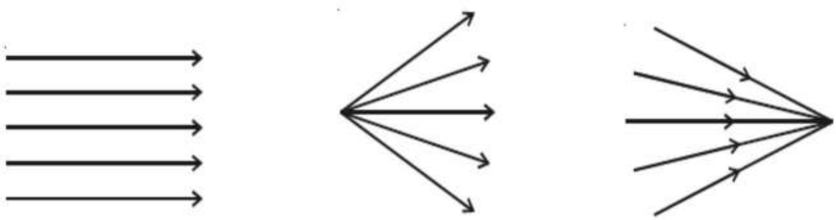
## Unit V
## OPTICS

**Light rays and beams**

A ray of light is the direction along which the light energy travels. In practice a ray has a finite width and is represented in diagrams as straight lines. A beam of light is a collection of rays. A search light emits a parallel beam of light (Fig.a). Light from a lamp travels in all directions which is a divergent beam. (Fig.b). A convex lens produces a convergent beam of light, when a parallel beam falls on it (Fig.c).

**Reflection of light**

Highly polished metal surfaces reflect about 80% to 90% of the light



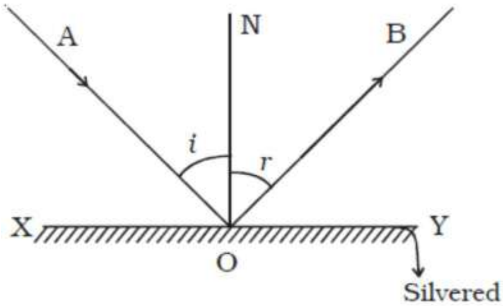
(a) Parallel beam (b) Divergent Beam (c) Convergent Beam

incident on them. Mirrors in everyday use are therefore usually made of depositing silver on the backside of the glass. The largest reflector in the world is a curved mirrors nearly 5 meters across, whose front surface is coated with aluminium. It is the hale Telescope on the top of Mount Palomar, California, U.S.A. Glass by itself, will also reflect light, but the percentage is small when compared with the case of silvered surface. It is about 5% for an air-glass surface.

**Laws of reflection**

Consider a ray of light, AO, incident on a plane mirror XY at O. It is reflected along OB. Let the normal ON is drawn at the point of incidence. The angle AON between the incident ray and the normal is called angle of incidence, i (Fig) the angle BON between the reflected ray and the normal is called angle of reflection, r. Experiments

show that : (i) The incident ray, the reflected ray and the normal drawn to the reflecting surface at the point of incidence, all lie in the same plane.




(ii) The angle of incidence is equal to the angle of reflection. (i.e) i = r.
These are called the laws of reflection.

**Deviation of light by plane mirror**

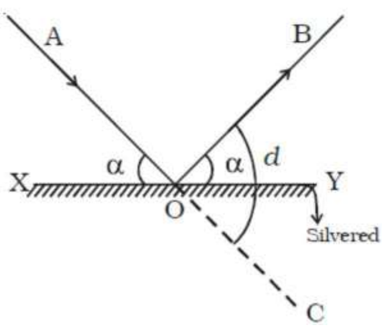
Consider a ray of light, AO, incident on a plane mirror XY (Fig) at O. It is reflected along B. The angle AOX made by AO with XY is known as the glancing angle $\alpha$ with the mirror. Since the angle of reflection is equal to the angle of incidence, the glancing angle BOY made by the reflected ray OB with the mirror is also equal to $\alpha$.

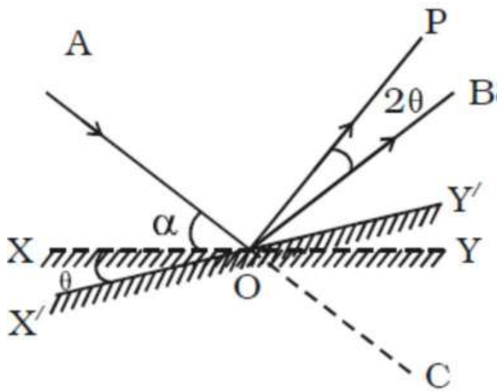
The light has been deviated from a direction AO to a direction OB. Since angle COY = angle AOX, it follows that angle of deviation, d = $2\alpha$

So, in general, the angle of deviation of a ray by a plane mirror or a plane surface is twice the glancing angle.

**Deviation of light due to rotation of a mirror**

Let us consider a ray of light AO incident on a plane mirror XY at O. It






is reflected along OB. Let $\alpha$ be the glancing angle with XY (Fig.). We know that the angle of deviation COB = $2\alpha$.

Suppose the mirror is rotated through an angle $\theta$ to a position X'Y'.

The same incident ray AO is now reflected along OP. Here the glancing angle with X'Y' is $(\alpha + \theta)$. Hence the new angle of deviation COP = 2 $(\alpha + \theta)$. The reflected ray has thus been rotated through an angle BOP when the mirror is rotated through an angle $\theta$.

∟BOP = ∟COP-COB

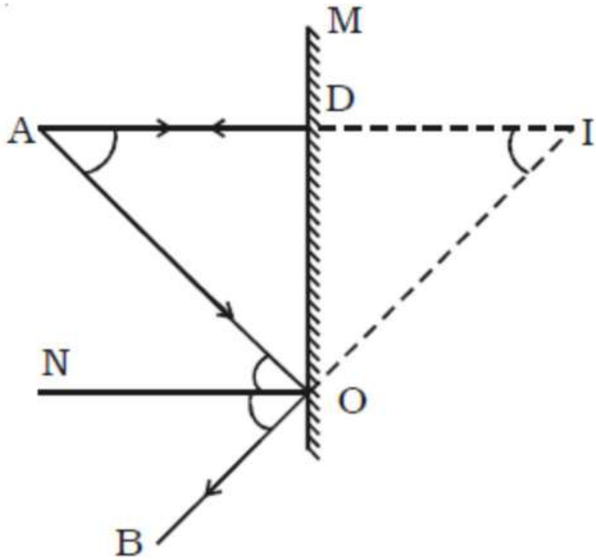
BOP = 2 $(\alpha + \theta) - 2\alpha = 2\theta$

For the same incident ray, when the mirror is rotated through an angle, the reflected ray is rotated through twice the angle.

**Image in a plane mirror**

Let us consider a point object A placed in front of a plane mirror M as shown in the Fig. Consider a ray of light AO from the point object incident on the mirror and reflected along OB. Draw the normal ON to the mirror at O.

The angle of incidence AON = angle of reflection BON Another ray AD incident normally on the mirror at D is reflected back along DA. When BO and AD are produced backwards, they meet at I. Thus the rays

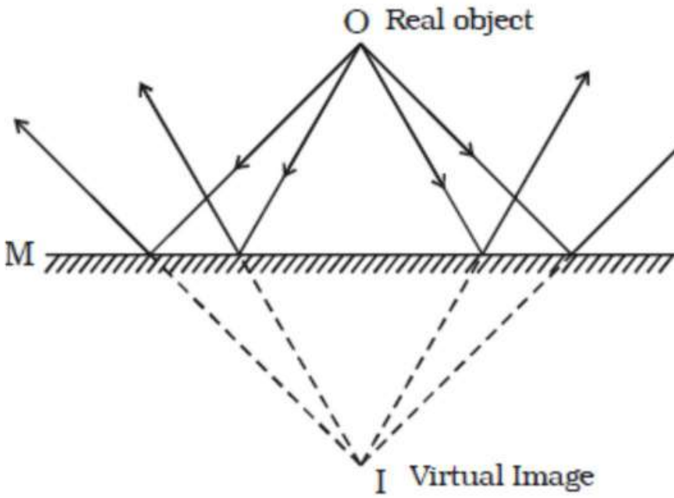



reflected from M appear to come from a point I behind the mirror.

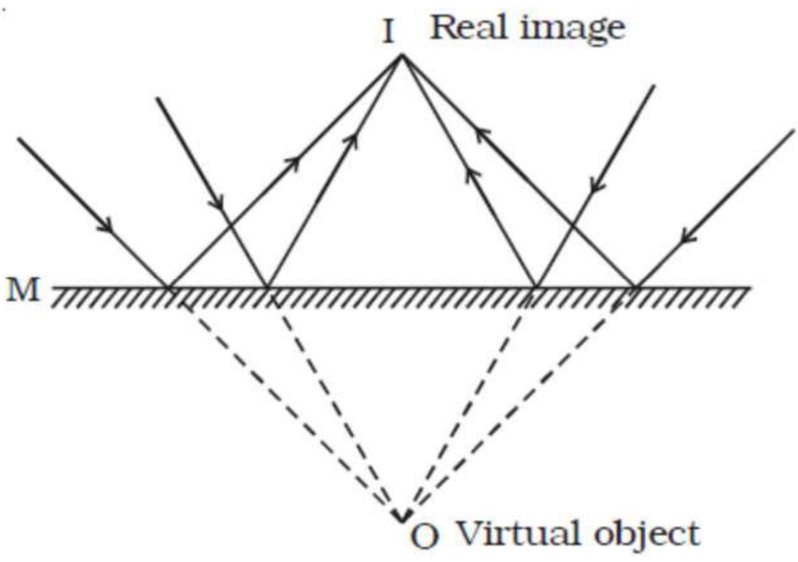
From the figure
∟AON = ∟DAO, alternate angles and ∟BON = ∟DIO, corresponding angles it follows that ∟DAO=∟DIO.
The triangles ODA and ODI are congruent
∴ AD = ID, For a given position of the object, A and D are fixed points. Since AD = ID, the point I is also fixed. It should be noted that AO = OI. So the object and its image in a plane mirror are at equal perpendicular distances from the mirror.

**Virtual and real images**

An object placed in front of a plane mirror has an image behind the mirror. The rays reflected from the mirror do not actually meet through I, but only appear to meet and the image cannot be received on the screen, because the image is behind the mirror. This type of image is called an unreal or virtual image (Fig.a).




If a convergent beam is incident on a plane mirror, the reflected rays pass through a point I in front of M, as shown in the Fig. b. In the Fig. a, a real object (divergent beam) gives rise to a virtual image. In the Fig. b, a virtual object (convergent beam) gives a real image.



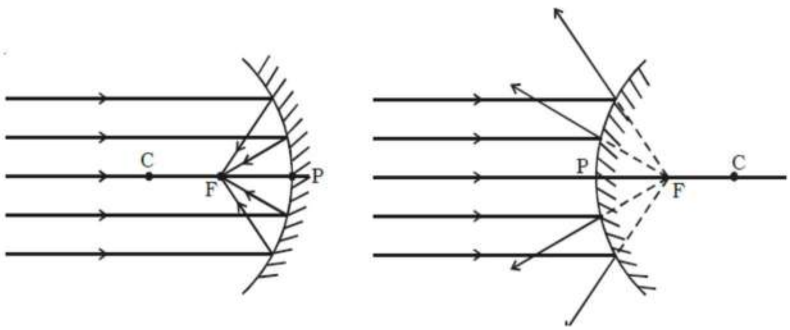

Hence plane mirrors not only produce virtual images for real objects but also produce real images for virtual objects.

**Characteristics of the image formed by a plane mirror**

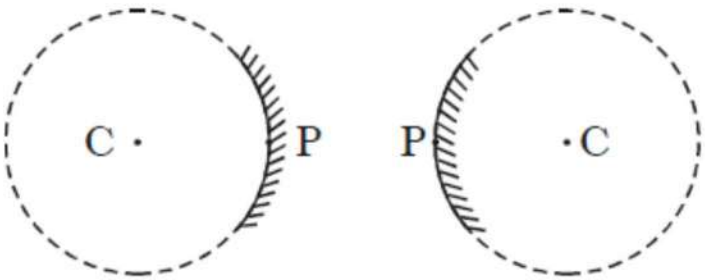
- Image formed by a plane mirror is as far behind the mirror as the object is in front of it and it is always virtual.
- The image produced is laterally inverted.
- The minimum size of the mirror required to see the complete image of the object is half the size of the object.
- If the mirror turns by an angle θ, the reflected ray turns through an angle $2\theta$.
- If an object is placed between two plane mirrors inclined at an angle θ, then the number of images formed is $n = \frac{360°}{\theta} - 1$

**Reflection at curved surfaces**

In optics we are mainly concerned with curved mirrors which are the part of a hollow sphere (Fig.). One surface of the mirror is silvered.



Reflection takes place at the other surface. If the reflection takes place at the concave surface, (which is towards the centre of the sphere) it is called concave mirror. If the reflection takes place at the convex surface,




(which is away from the centre of the sphere) it is called convex mirror. The laws of reflection at a plane mirror are equally true for spherical mirrors also. The centre of the sphere, of which the mirror is a part is called the centre of curvature (C). The geometrical centre of the mirror is called its pole (P). The line joining the pole of the mirror and its centre of curvature is called the principal axis. The distance between the pole and the centre of curvature of the spherical mirror is called the radius of curvature of the mirror and is also equal to the radius of the sphere of which the mirror forms a part. When a parallel beam of light is incident on a spherical mirror, the point where the reflected rays converge (concave mirror) or appear to diverge from the point (convex mirror) on the principal axis is called the principal focus (F) of the mirror. The distance between the pole and the principal focus is called the focal length (f) of the mirror (Fig.).

**Images formed by a spherical mirror**

The images produced by spherical mirrors may be either real or virtual and may be either larger r smaller than the object. The image can be located by graphical construction as shown in Fig. by adopting any two of the following rules.

- A ray parallel to the principal axis after reflection by a concave mirror passes through the principal focus of the concave mirror and appears to come from the principal focus in a convex mirror.

- A ray passing through the centre of curvature retraces its path after reflection.
- A ray passing through the principal focus, after reflection is rendered parallel to the principal axis.
- A ray striking the pole at an angle of incidence i is reflected at the same angle i to the axis.

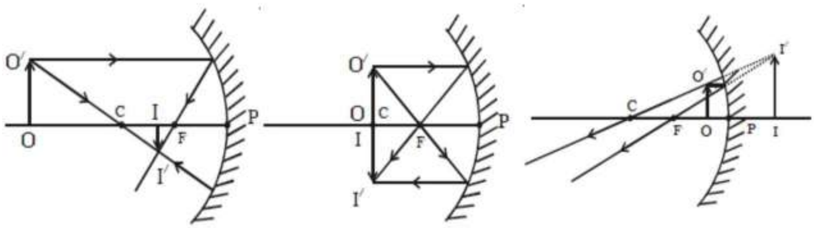
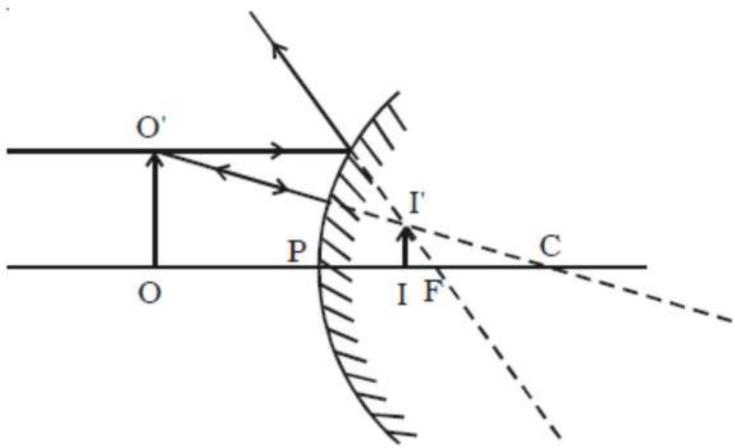



**Image formed by a convex mirror**

In a convex mirror irrespective of the position of the object, the image formed is always virtual, erect but diminished in size. The image lies



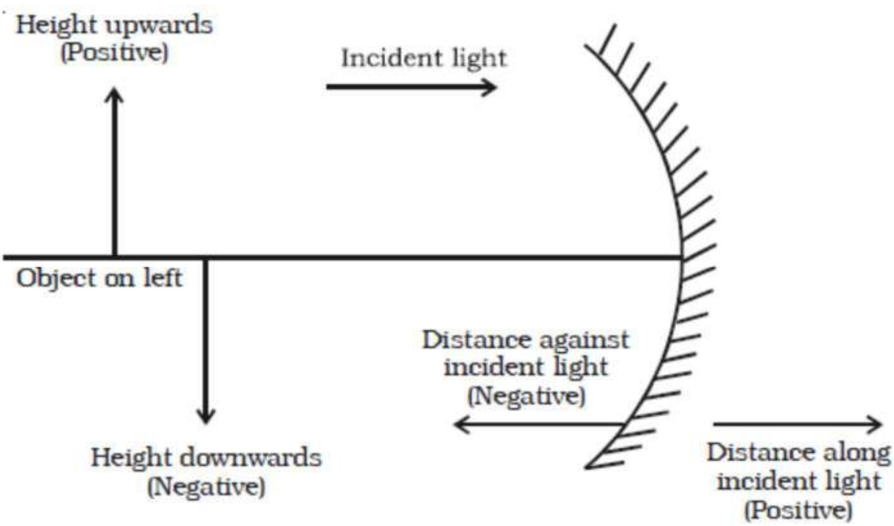

between the pole and the focus (Fig.).

In general, real images are located in front of a mirror while virtual images behind the mirror.

**Cartesian sign convention**

The following sign conventions are used.

- All distances are measured from the pole of the mirror (in the case of lens from the optic centre).
- The distances measured in the same direction as the incident light, are taken as positive.
- The distances measured in the direction opposite to the direction of incident light are taken as negative.
- Heights measured perpendicular to the principal axis, in the upward direction are taken as positive.
- Heights measured perpendicular to the principal axis, in the downward direction are taken as negative.
- The size of the object is always taken as positive, but image size is positive for erect image and negative for an inverted image.
- The magnification is positive for erect (and virtual) image, and negative for an inverted (and real) image.




**Relation between u, v and f for spherical mirrors**

A mathematical relation between object distance u, the image distance v and the focal length f of a spherical mirror is known as mirror formula.

(i) Concave mirror - real image

Let us consider an object OO′ on the principal axis of a concave mirror beyond C. The incident and the reflected rays are shown in the Fig. A ray O′A parallel to principal axis is incident on the concave mirror at A, close to P. After reflections the ray passes through the focus F. Another ray O′C passing through centre of curvature C, falls normally on the mirror and reflected back along the same path. A third ray O′P incident at the pole P is reflected along PI′. The three reflected rays intersect at the point I′. Draw perpendicular I′I to the principal axis. II′ is the real, inverted image of the object OO′. Right angled triangles, II′P and OO′P are similar.

$\therefore \frac{II'}{OO'} = \frac{PI}{PO}$ … (1)

Right angled triangles II′F and APF are also similar (A is close to P; hence AP is a vertical line)

$\therefore \frac{II'}{AP'} = \frac{IF}{PF}$

AP = OO′. Therefore the above equation becomes,

$\therefore \frac{II'}{OO'} = \frac{IF}{PF}$ … (2)

Comparing the equations (1) and (2)

$\frac{PI}{PO} = \frac{IF}{PF}$ … (3)

But, IF = PI – PF

Therefore equation (3) becomes,

$\frac{PI}{PO} = \frac{PI-PF}{PF}$ … (4)

Using sign conventions, we have PO = –u,

PI = $-v$ and PF = -$f$

Substituting the values in the above equation, we get

$\frac{-v}{-u} = \frac{-v-(-f)}{-f}$ (or)

$\frac{v}{u} = \frac{v-f}{f} = \frac{v}{f}$ -1

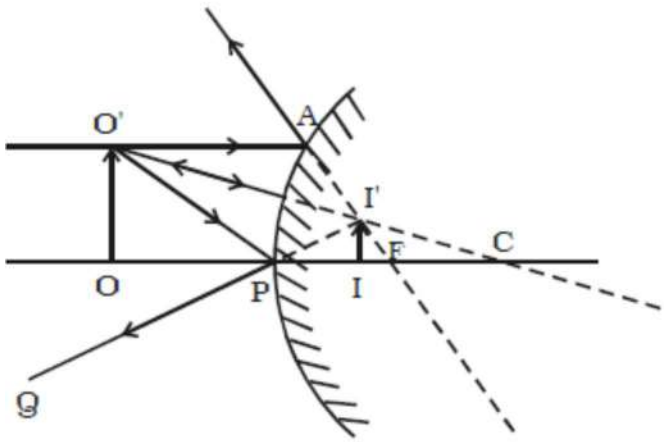
Dividing by v and rearranging, $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

This is called mirror equation. The same equation can be obtained for virtual image also.

**(ii) Convex mirror - virtual image**

Let us consider an object OO′ anywhere on the principal axis of a convex mirror. The incident and the reflected rays are shown in the Fig. A ray O′A parallel to the principal axis incident on the convex mirror at A close to P. After reflection the ray appears to diverge from the focus F. Another ray O′C passing through centre of curvature C, falls normally on

the mirror and is reflected back along the same path. A third ray O'P incident at the pole P is reflected along PQ. The three reflected rays when produced appear to meet at the point I'. Draw perpendicular II' to the principal axis. II' is the virtual image of the object OO'.
Right angled triangles, II'P and OO'P are similar.




$\therefore \frac{II'}{OO'} = \frac{PI}{PO}$ … (1)

Right angled triangles II'F and APF are also similar (A is close to P; hence AP is a vertical line)

$\frac{II'}{AP'} = \frac{IF}{PF}$

AP = OO'. Therefore the above equation becomes,

$\frac{II'}{OO'} = \frac{IF}{PF}$ … (2)

Comparing the equations (1) and (2)

$\frac{PI}{PO} = \frac{IF}{PF}$ … (3)

But, IF = PF – PI. Therefore equation (3) becomes,

$\frac{PI}{PO} = \frac{PF-PI}{PF}$

Using sign conventions, we have PO = -u, PI = +v and PF = +f. Substituting the values in the above equation, we get

$\frac{+v}{-u} = \frac{+f-(+v)}{+f}$ or $-\frac{v}{v} = \frac{f-v}{f} = 1 - \frac{v}{f}$

Dividing by v and rearranging we get, $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

This is called mirror equation for convex mirror producing virtual image.

**Magnification**

The linear or transverse magnification is defined as the ratio of the size of the image to that of the object.

∴ Magnification $=\frac{Sive\ of\ the\ image}{size\ of\ the\ object}=\frac{h_2}{h_1}$

where $h_1$ and $h_2$ represent the size of the object and image respectively.

From Fig. it is known that $\frac{II'}{OO'}=\frac{PI}{PO}$

Applying the sign conventions,

II' = $-h_2$ (height of the image measured downwards)

OO' = $+h_1$ (height of the object measured upwards)

PI = $-v$ (image distance against the incident light)

PO = $-u$ (object distance against the incident light)

Substituting the values in the above equation, we get

magnification m = $\frac{-h_2}{+h_1}=\frac{-v}{-u}$ (or) m= $\frac{h_2}{h_1}=\frac{-v}{u}$

For an erect image m is positive and for an inverted image m is negative. This can be checked by substituting values for convex mirror also. Using mirror formula, the equation for magnification can also be obtained as

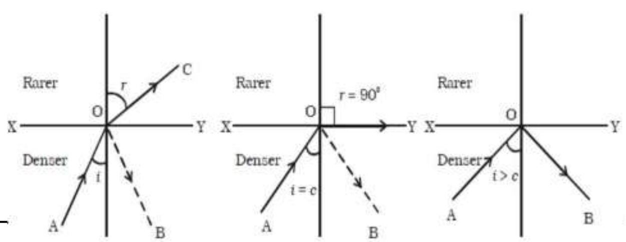
m= $\frac{h_2}{h_1}=\frac{-v}{u}=\frac{f-v}{f}=\frac{f}{f-u}$

This equation is valid for both convex and concave mirrors.

**Total internal reflection**

When a ray of light AO passes from an optically denser medium to a rarer medium, at the interface XY, it is partly reflected back into the same medium along OB and partly refracted into the rarer medium along OC (Fig.).

If the angle of incidence is gradually increased, the angle of refraction r will also gradually increase and at a certain stage r becomes 90°. Now the refracted ray OC is bent so much away from the normal and it grazes the surface of separation of two media. The angle of incidence in the denser medium at which the refracted ray just grazes the surface of separation is called the critical angle c of the denser medium. If i is



increased further, refraction is not possible and the incident ray is totally reflected into the same medium itself. This is called total internal reflection.

If $\mu_d$ is the refractive index of the denser medium then, from Snell's Law, the refractive index of air with respect to the denser medium is given by,

$${}_d\mu_a = \frac{\sin i}{\sin r}$$

$$\frac{\mu_a}{\mu_d} = \frac{\sin i}{\sin r}$$

$$\frac{1}{\mu_d} = \frac{\sin i}{\sin r} (\because \mu_a = 1 \text{ for air})$$

If r = 90°, i = c

$$\frac{\sin c}{\sin 90°} = \frac{1}{\mu_d} \text{ (or) } \sin c = \frac{1}{\mu_d} \text{ or } c = \sin^{-1}\left(\frac{1}{\mu_d}\right)$$

If the denser medium is glass, $c = \sin^{-1}\left(\frac{1}{\mu_g}\right)$

Hence for total internal reflection to take place (i) light must travel from a denser medium to a rarer medium and (ii) the angle of incidence inside the denser medium must be greater than the critical angle i.e. i > c.

**Applications**

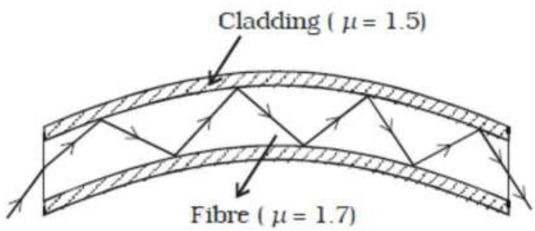
| Medium | Refractive index | Critical angle |
|---|---|---|
| Water | 1.33 | 48.75° |
| Crown glass | 1.52 | 41.14° |
| Dense flint glass | 1.62 | 37.31° |
| Diamond | 2.42 | 24.41° |

(i) Diamond

Total internal reflection is the main cause of the brilliance of diamonds. The refractive index of diamond with respect to air is 2.42. Its critical angle is 24.41°. When light enters diamond from any face at an angle greater than 24.41° it undergoes total internal reflection. By cutting the diamond suitably, multiple internal reflections can be made to occur.

(ii) Optical fibres The total internal reflection is the basic principle of optical fibre. An optical fibre is a very thin fibre made of glass or quartz having radius of the order of micrometer ($10^{-6}$ m). A bundle, of such thin fibres forms a 'light pipe' (Fig. a). Fig. b shows the principle of light transmission inside an optical fibre. The refractive index of the material of the core is higher than that of the cladding. When the light is incident at one end of the fibre at a small angle, the light passes inside, undergoes repeated total internal reflections along the fibre and finally comes out. The angle of incidence is always larger than the critical angle of the core material with respect to its cladding. Even if the fibre is bent or twisted, the light can easily travel through the fibre. Light pipes are used in medical and optical examination. They are also used to transmit communication signals.

(a)



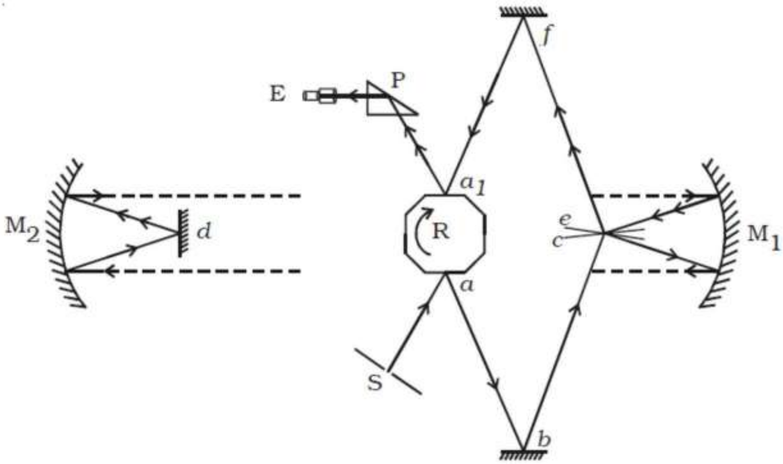

(b)

**Michelson's method**

A.A. Michelson, an American physicist, spent many years of his life in measuring the velocity of light and he devised a method in the year

1926 which is considered as accurate.

The experimental set up is shown in Fig. Light from an arc source after passing through a narrow slit S is reflected from one face $a$ of an octagonal mirror R. The ray after reflections at small fixed mirrors b and c is then rendered parallel by a concave mirror $M_1$ placed in the observing station on Mt. Wilson. This parallel beam of light travels a distance of 35 km and falls on another concave mirror $M_2$ placed at Mt. St Antonio, and it is reflected to a plane mirror d placed at the focus of the concave mirror $M_2$ . The ray of light from d is rendered parallel after getting




reflected by $M_2$ and travels back to the concave mirror $M_1$. After reflections at $M_1$ and the plane mirrors e and f, the ray falls on the opposite face a 1 of the octagonal mirror. The final image which is totally reflected by a total reflecting prism P, is viewed through an eye piece E. When the octagonal mirror is stationary, the image of the slit is seen through the eye piece. When it is rotated the image disappears. The speed of rotation of R is suitably adjusted so that the image is seen again clearly as when R is stationary. The speed of revolution is measured by stroboscope. Let D be the distance travelled by light from face a to face a 1 and n be the number of rotations made by R per second.

The time taken by R to rotate through 45° or $\frac{1}{8}$ of a rotation = $\frac{1}{8n}$

During this time interval, the distance travelled by the light = D

∴The velocity of light c = Distance travelled/Time taken = $\frac{D}{\frac{1}{8}}$= 8nD.

In general, if the number of faces in the rotating mirror is N, the velocity of light = NnD.

The velocity of light determined by him is 2.99797 × $10^8$ m $s^{-1}$.

**Importance of velocity of light**

The value of velocity of light in vacuum is of great importance in science. The following are some of the important fields where the value of velocity of light is used.

(1) Frequency - wavelength relation: From the relation c = $v\lambda$, the frequency of electromagnetic radiations can be calculated if the wavelength is known and vice versa.

(2) Relativistic mass variation with velocity: Theory of relativity has shown that the mass m of a moving particle varies with its velocity v according to the relation m = $\frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}}$

Here $m_0$ is the rest mass of the particle.

(3) Mass - Energy relation: E = $mc^2$ represents conversion of mass into energy and energy into mass. The energy released in nuclear fission and fusion is calculated using this relation.

(4) Measurement of large distance in astronomy: Light year is a unit of distance used in astronomy. A light year is the distance travelled by light in one year. It is equal to 9.46 × $10^{15}$ metre.

(5) Refractive index: The refractive index $\mu$ of a medium is given by

$$\mu = \frac{velocity\ of\ light\ in\ vaccum}{velocity\ of\ light\ in\ medium} = \frac{c}{v}$$

**Refraction of light**

When a ray of light travels from one transparent medium into another medium, it bends while crossing the interface, separating the two media. This phenomenon is called refraction. Image formation by spherical lenses is due to the phenomenon of refraction. The laws of refraction at a plane surface are equally true for refraction at curved surfaces also. While deriving the expressions for refraction at spherical surfaces, we make the following assumptions.

(i) The incident light is assumed to be monochromatic and

(ii) the incident pencil of light rays is very narrow and close to the principal axis.

**Cartesian sign convention**

The sign convention followed in the spherical mirror is also applicable to refraction at spherical surface. In addition to this two more sign conventions to be introduced which are:

(i) The power of a converging lens is positive and that of a diverging lens is negative.

(ii) The refractive index of a medium is always said to be positive. If two refractions are involved, the difference in their refractive index is also taken as positive.

**Refraction at a spherical surface**

Let us consider a portion of a spherical surface AB separating two media having refracting indices $\mu_1$ and$\mu_1$(Fig.). This is symmetrical about an axis passing through the centre C and cuts the surface at P. The point P is called the pole of the surface. Let R be the radius of curvature of the surface. Consider a point object O on the axis in the first medium. Consider two rays OP and OD originating from O. The ray OP falls normally on AB and goes into the second medium, undeviated. The ray OD falls at D very close to P. After refraction, it meets at the point I on the axis, where the image is formed. CE is the normal drawn to the point D. Let i and r be the angle of incidence and refraction respectively.

Let ∟DOP= $\alpha$ , ∟DCP =$\beta$, ∟ DIC = $\gamma$

Since D is close to P, the angles $\alpha$, $\beta$ and $\gamma$ are all small. From the Fig.

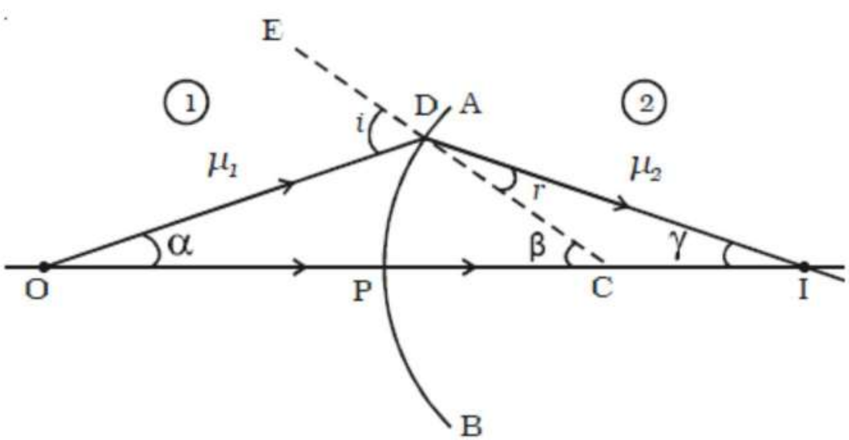
$\tan\alpha = \frac{DP}{PO}$, $\tan\beta = \frac{DP}{PC}$ and $\tan\gamma = \frac{DP}{PI}$

$\therefore\ \alpha = \frac{DP}{PO}$, $\beta = \frac{DP}{PC}$ and $\gamma = \frac{DP}{PI}$

From the ΔODC, i = $\alpha + \beta$

... (1)

From the ΔDCI, $\beta$ = r + $\gamma$ or r = $\beta - \gamma$ ...



(2)

From Snell's Law, $\frac{\mu_2}{\mu_1} = \frac{\sin i}{\sin r}$ and for small angles of i and r, we can write,

$\mu_1 i = \mu_1 r$

$\dots$ (3)

From equations (1), (2) and (3)

we get $\mu_1(\alpha + \beta) = \mu_2(\beta - \gamma)$ or $\mu_1\alpha + \mu_1\gamma = (\mu_2 - \mu_1)\beta$ $\dots$ (4)

Substituting the values of $\alpha, \beta$ and $\gamma$ in equation (4)

$$\mu_1\left(\frac{DP}{PO}\right) + \mu_2\left(\frac{DP}{PI}\right) = (\mu_2 - \mu_1) - \left(\frac{DP}{PC}\right)$$

$$\frac{\mu_1}{PO} + \frac{\mu_2}{PI} = \left(\frac{\mu_2 - \mu_1}{PC}\right) \quad \dots (5)$$

As the incident ray comes from left to right, we choose this direction as the positive direction of the axis. Therefore u is negative, whereas v and R are positive substitute PO = $-u$ PI = $+v$ and PC = +R in equation (5),

$$\frac{\mu_1}{-u} + \frac{\mu_2}{v} = \frac{\mu_2 - \mu_1}{PR}$$

$$\frac{\mu_2}{-v} - \frac{\mu_1}{u} = \frac{\mu_2 - \mu_1}{R} \quad \dots (6)$$

Equation (6) represents the general equation for refraction at a spherical surface. If the first medium is air and the second medium is of refractive index $\mu$, then

$$\frac{\mu}{v} = \frac{1}{u} = \frac{\mu - 1}{R} \quad \dots (7)$$

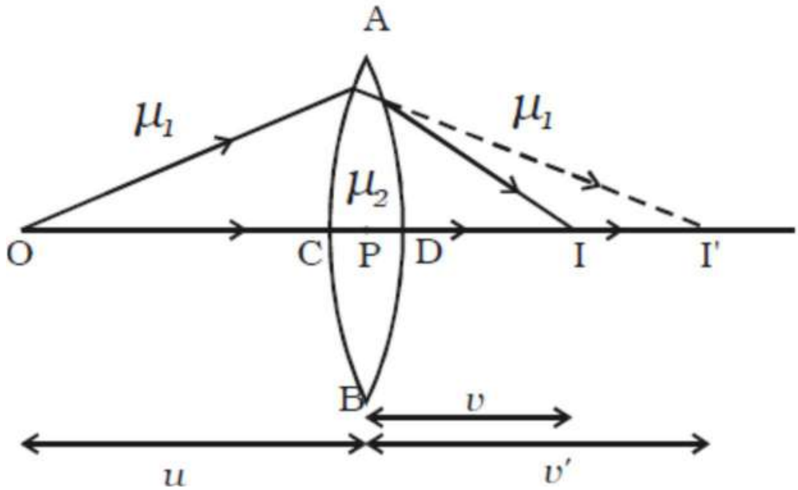
**Refraction through thin lenses**

A lens is one of the most familiar optical devices. A lens is made of a transparent material bounded by two spherical surfaces. If the distance between the surfaces of a lens is very small, then it is a thin lens.

As there are two spherical surfaces, there are two centres of curvature $C_1$ and $C_2$ and correspondingly two radii of curvature $R_1$ and $R_2$. The line joining $C_1$ and $C_2$ is called the principal axis of the lens. The centre P of the thin lens which lies on the principal aixs is called the optic centre.

**Lens maker's formula and lens formula**

Let us consider a thin lens made up of a medium of refractive index $\mu_2$ placed in a medium of refractive index $\mu_1$. Let $R_1$ and $R_2$ be the radii of curvature of two spherical surfaces ACB and ADB respectively and P be the optic centre. Consider a point object O on the principal axis. The ray OP falls normally on the spherical surface and goes through the lens undeviated. The ray OA falls at A very close to P. After refraction at the




surface ACB the image is formed at I'. Before it does so, it is again refracted by the surface ADB. Therefore the final image is formed at I as shown in Fig.

The general equation for the refraction at a spherical surface is given by

$$\frac{\mu_2}{v} - \frac{\mu_1}{u} = \frac{\mu_2 - \mu_1}{R} \qquad \dots (1)$$

For the refracting surface ACB, from equation (1) we write

$$\frac{\mu_2}{v'} - \frac{\mu_1}{u} = \frac{\mu_2 - \mu_1}{R} \qquad \dots (2)$$

The image I' acts as a virtual object for the surface ADB and the final image is formed at I. The second refraction takes place when light travels from the medium of refractive index $\mu_2$ *to* $\mu_1$. For the refracting surface ADB, from equation (1) and applying sign conventions, we have

$$\frac{\mu_1}{v} - \frac{\mu_2}{v'} = \left(\frac{\mu_2 - \mu_1}{-R_2}\right) \qquad \dots (3)$$

Adding equations (2) and (3) $\frac{\mu_1}{v} - \frac{\mu_1}{u} = (\mu_2 - \mu_1)\left[\frac{1}{R_1} - \frac{1}{R_2}\right]$ Dividing the above equation by $\mu_1$

$$\frac{1}{v} - \frac{1}{u} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left[\frac{1}{R_1} - \frac{1}{R_2}\right] \qquad \dots (4)$$

If the object is at infinity, the image is formed at the focus of the lens.

Thus, for u = ∞, $v = f$. Then the equation (4) becomes.

$\frac{1}{f} = \left(\frac{\mu_2}{\mu_1} - 1\right)\left[\frac{1}{R_1} - \frac{1}{R_2}\right]$

If the refractive index of the lens is $\mu$ and it is placed in air, $\mu_2 = \mu$ and $\mu_1$= 1. So the equation (5) becomes

$\frac{1}{f} = (\mu - 1)\left[\frac{1}{R_1} - \frac{1}{R_2}\right]$ … (5)

This is called the lens maker's formula, because it tells what curvature will be needed to make a lens of desired focal length. This formula is true for concave lens also. Comparing equation (4) and (5)

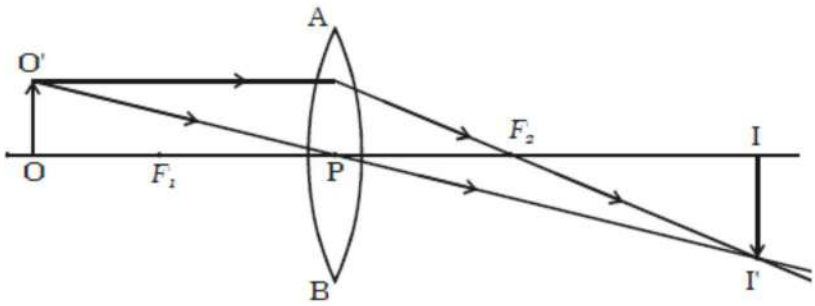
We get $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$ … (6)

which is known as the lens formula.

**Magnification**

Let us consider an object OO' placed on the principal axis with its height perpendicular to the principal axis as shown in Fig. The ray OP passing through the optic centre will go undeviated. The ray O'A parallel to the principal axis must pass through the focus $F_2$. The image is formed where O'PI' and $AF_2I'$ intersect. Draw a perpendicular from I' to the principal axis. This perpendicular II' is the image of OO'.

The linear or transverse magnification is defined as the ratio of the size




of the image to that of the object.

∴ Magnification m = $\frac{Size\ of\ the\ image}{Size\ of\ the\ object} = \frac{II'}{OO'} = \frac{h_2}{h_1}$

where $h_1$ is the height of the object and $h_2$ is the height of the image. From the similar right angled triangles OO'P and II'P, we have

$\frac{II'}{OO'} = \frac{PI}{PO}$

Applying sign convention,

$II' = -h_2;$ $OO' = +h_1;$

$PI = +v;\qquad PO = -u;$

Substituting this in the above equation, we get magnification

$$m = \frac{-h_2}{+h_1} = \frac{+v}{-u}$$

$$\therefore m = +\frac{v}{u}$$

The magnification is negative for real image and positive for virtual image. In the case of a concave lens, it is always positive.

Using lens formula the equation for magnification can also be obtained as m $= \frac{h_2}{h_1} = \frac{v}{u} = \frac{f-v}{f} = \frac{f}{f+u}$

This equation is valid for both convex and concave lenses and for real and virtual images.

**Power of a lens**

Power of a lens is a measure of the degree of convergence or divergence of light falling on it. The power of a lens (P) is defined as the reciprocal of its focal length.

$$P = \frac{1}{f}$$

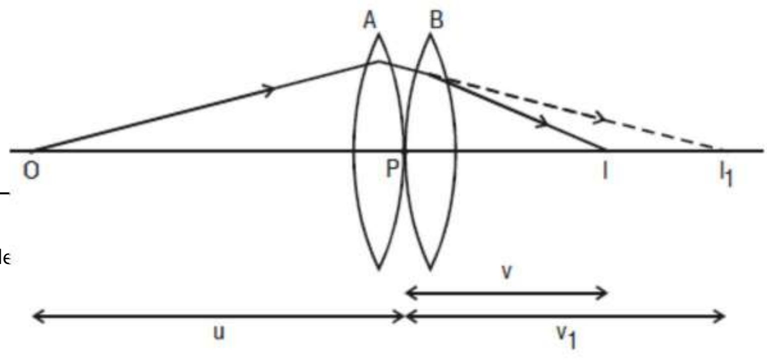
The unit of power is dioptre (D): 1 D = 1 $m^{-1}$. The power of the lens is said to be 1 dioptre if the focal length of the lens is 1 metre. P is positive for converging lens and negative for diverging lens. Thus, when an optician prescribes a corrective lens of power + 0.5 D, the required lens is a convex lens of focal length + 2 m. A power of -2.0 D means a concave lens of focal length -0.5 m.

**Combination of thin lenses in contact**

Let us consider two lenses A and B of focal length $f_1$ and $f_2$ placed in contact with each other.

An object is placed at O beyond the focus of the first lens A on the common principal axis. The lens A produces an image at $I_1$. This image $I_1$ acts as the object for the second lens B. The final image is produced at I as shown in Fig. Since the lenses are thin, a common optical centre P is chosen.

Let PO = u, object distance for the first lens (A), PI = $v$, final image



distance and $PI_1 = v_1$, image distance for the first lens (A) and also object distance for second lens (B).

For the image $I_1$ produced by the first lens A,

$$\frac{1}{v_1} = \frac{1}{u} = \frac{1}{f_1} \qquad \text{... (1)}$$

For the final image I, produced by the second lens B,

$$\frac{1}{v} = \frac{1}{v_1} = \frac{1}{f_2} \qquad \text{... (2)}$$

Adding equations (1) and (2),

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f_1} + \frac{1}{f_2} \qquad \text{... (3)}$$

If the combination is replaced by a single lens of focal length F such that it forms the image of O at the same position I, then

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f} \qquad \text{... (4)}$$

From equations (3) and (4)

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} \qquad \text{... (5)}$$

This F is the focal length of the equivalent lens for the combination. The derivation can be extended for several thin lenses of focal lengths $f_1$, $f_2$, $f_3$... in contact. The effective focal length of the combination is given by

$$\frac{1}{F} = \frac{1}{f_1} + \frac{1}{f_2} + \frac{1}{f_3} + \dots \qquad \text{... (6)}$$

In terms of power, equation (6) can be written as

$$P = P_1 + P_2 + P_3 + \qquad \text{... (7)}$$

Equation (7) may be stated as follows:

The power of a combination of lenses in contact is the algebraic sum of the powers of individual lenses.

The combination of lenses is generally used in the design of objectives of microscopes, cameras, telescopes and other optical instruments.

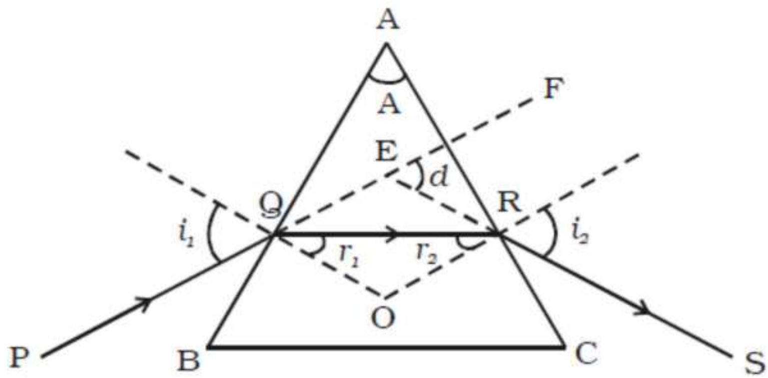
**Prism**

A prism is a transparent medium bounded by the three plane faces. Out of the three faces, one is grounded and the other two are polished. The polished faces are called refracting faces. The angle between the refracting faces is called angle of prism, or the refracting angle. The third face is called base of the prism.

## Refraction of light through a prism

Fig. shows the cross section of a triangular prism ABC, placed in air. Let 'A' be the refracting angle of the prism. A ray of light PQ incident on the refracting face AB, gets refracted along QR and emerges along RS. The angle of incidence and refraction at the two faces are $i_1$, $r_1$, $r_2$ and $i_2$ respectively. The angle between the incident ray PQ and the emergent ray RS is called angle of deviation, d.

In the ΔQER, the exterior angle ∟FER= ∟EQR+∟ERQ




$d = (i_1 - r_1) + (i_2 - r_2)$

$\therefore d = (i_1 + i_2) - (r_1 + r_2)$ ... (1)

In the quadrilateral AQOR, the angles at Q and R are right angles

Q + R = 180°

∴ A + QOR = 180o ... (2)

Also, from the ΔQOR

$r_1 + r_2 + QOR = 180°$ ... (3)

From equation (2) and (3)

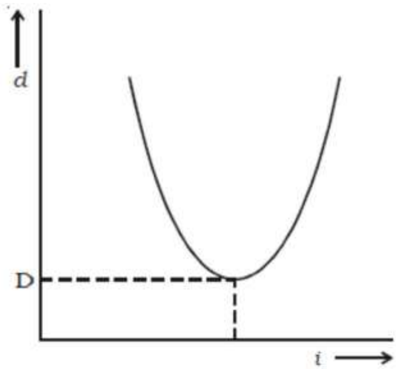
$r_1 + r_2 = A$ ... (4)

Substituting in (1),

$d = i_1 + i_2 - A$

or $A + d = i_1 + i_2$ ... (5)

For a given prism and for a light of given wavelength, the angle of




deviation depends upon the angle of incidence.

As the angle of incidence i gradually increases, the angle of deviation d decreases, reaches a minimum value D and then increases. D is called the angle of minimum deviation. It will be seen from the graph (Fig.) that there is only one angle of incidence for which the deviation is a minimum. At minimum deviation position the incident ray and emergent ray are symmetric with respect to the base of the prism. (i.e) the refracted ray QR is parallel to the base of the prism. At the minimum deviation $i_1 = i_2 = i$ and $r_1 = r_2 = r$

∴from equation (4) 2r = A or $r = \frac{A}{2}$ and from equation (5) 2i = A + D or $i = \frac{A+D}{2}$

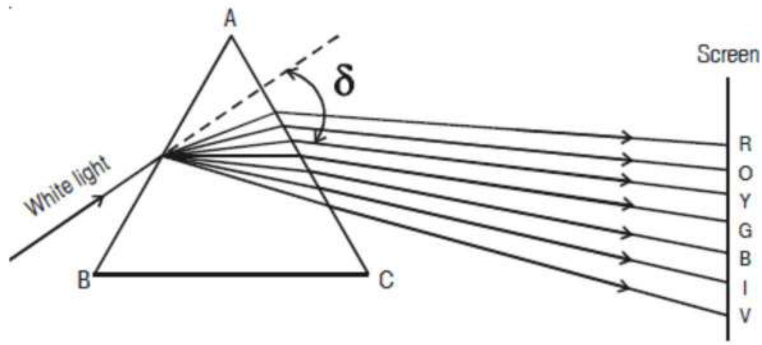
The refractive index is $\mu = \frac{\sin i}{\sin r}$

$$\therefore \mu = \frac{\sin\left(\frac{A+D}{2}\right)}{\sin\left(\frac{A}{2}\right)}$$

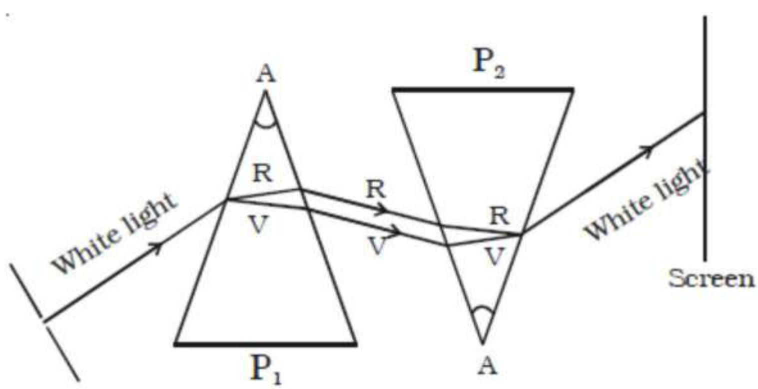
**Dispersion of light**

Dispersion is the splitting of white light into its constituent colours. This band of colours of light is called its spectrum.  n the visible region of spectrum, the spectral lines are seen in the order from violet to red. The colours are given by the word VIBGYOR (Violet, Indigo, Blue, Green, Yellow, Orange and Red) (Fig.)

The origin of colour after passing through a prism was a matter of much debate in physics. Does the prism itself create colour in some way or




does it only separate the colours already present in white light? Sir Isaac Newton gave an explanation for this. He placed another similar prism in an inverted position. The emergent beam from the first prism was made to fall on the second prism (Fig.). The resulting emergent beam was found to be white light. The first prism separated the white light into its constituent colours, which were then recombined by the inverted prism to give white light. Thus it can be concluded that the prism does not create any colour but it only separates the white light into its




constituent colours. Dispersion takes place because the refractive index of the material of the prism is different for different colours (wavelengths). The deviation and hence the refractive index is more for violet rays of light than the corresponding values for red rays of light. Therefore the violet ray travels with a smaller velocity in glass prism than red ray. The deviation and the refractive index of the yellow ray are taken as the mean values. Table gives the refractive indices for different wavelength for crown glass and flint glass.

| Colour | Wave length (nm) | Crown glass | Flint glass |
|---|---|---|---|
| Violet | 396.9 | 1.533 | 1.663 |
| Blue | 486.1 | 1.523 | 1.639 |
| Yellow | 589.3 | 1.517 | 1.627 |
| Red | 656.3 | 1.515 | 1.622 |

The speed of light is independent of wavelength in vacuum. Therefore vacuum is a non-dispersive medium in which all colours travel with the same speed.

**Dispersive power**

The refractive index of the material of a prism is given by the relation $\mu = \frac{\sin\left(\frac{A+D}{2}\right)}{\sin\left(\frac{A}{2}\right)}$.

Here A is the angle of the prism and D is the angle of minimum deviation.

If the angle of prism is small of the order of 10°, the prism is said to be small angled prism. When rays of light pass through such prisms the angle of deviation also becomes small.

If A be the refracting angle of a small angled prism and $\delta$ the angle of deviation, then the prism formula becomes $\mu = \frac{\sin\left(\frac{A+\delta}{2}\right)}{\sin\left(\frac{A}{2}\right)}$

For small angles A and $\delta$, $\sin\frac{A+\delta}{2} = \frac{A+\delta}{2}$ and $\sin\frac{A}{2} = \frac{A}{2}$

$$\therefore \mu = \frac{\frac{A+\delta}{2}}{\frac{A}{2}}$$

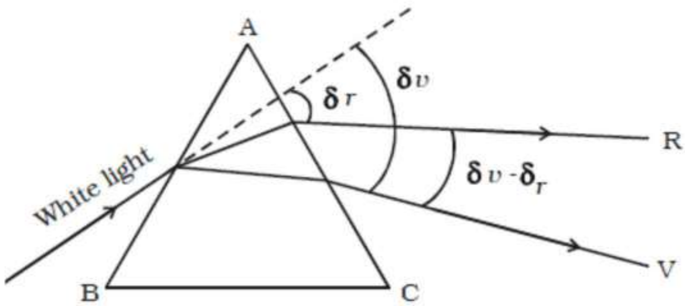
$$\mu A = A + \delta$$

$$\delta = (\mu - 1)A \qquad \ldots (1)$$

If $\delta_v$ and $\delta_r$ are the deviations produced for the violet and red rays and $\mu_v$ and $\mu_r$ are the corresponding refractive indices of the material of the small angled prism then,

for violet light, $\delta_v = (\mu_v - \mu_r)A$ … (2)

for red light, $\delta_r = (\mu_r - 1)A$ … (3)




From equations (2) and (3)

$\delta_v - \delta_r = (\mu_v - \mu_r)A$ … (4)

$\delta_v - \delta_r$ is called the angular dispersion which is the difference in deviation between the extreme colours (Fig.).

If $\delta_y$ and $\mu_y$ are the deviation and refractive index respectively for yellow ray (mean wavelength) then, for yellow light, $\delta_y = (\mu_y - 1)A$ ... (5)

Dividing equation (4) by (5) we get $\frac{\delta_v - \delta_r}{\delta_y} = \frac{(\mu_v - \mu_r)A}{(\mu_y - 1)A}$

$\frac{\delta_v - \delta_r}{\delta_y} = \frac{\mu_v - \mu_r}{\mu_y - 1}$

The expression $\frac{\delta_v - \delta_r}{\delta_y}$ is known as the dispersive power of the material of the prism and is denoted by $\omega$.

$\therefore \omega = \frac{\mu_v - \mu_r}{\mu_y - 1}$

The dispersive power of the material of a prism is defined as the ratio of angular dispersion for any two wavelengths (colours) to the deviation of mean wavelength.

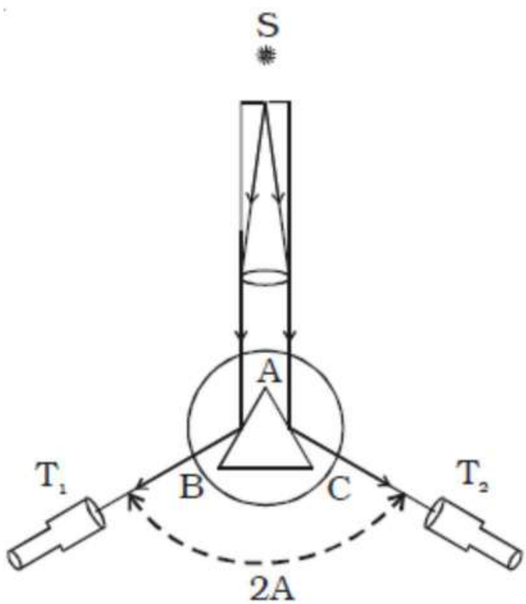
**Determination of the refractive index of the material of the prism**

The preliminary adjustments of the telescope, collimator and the prism table of the spectrometer are made. The refractive index of the prism can be determined by knowing the angle of the prism and the angle of minimum deviation.

**(i) Angle of the prism (A)**

The prism is placed on the prism table with its refracting edge facing the collimator as shown in Fig. The slit is illuminated by a sodium vapour lamp.

The parallel rays coming from the collimator fall on the two faces AB and AC. The telescope is rotated to the position $T_1$ until the image of the slit, formed by the reflection at the face AB is made to coincide with the vertical cross wire of the telescope. The readings of the verniers are

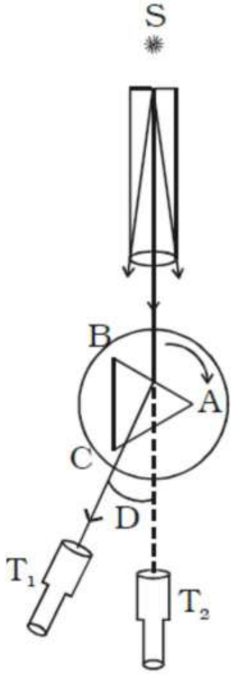



noted. The telescope is then rotated to the position $T_2$ where the image of the slit formed by the reflection at the face AC coincides with the vertical cross wire. The readings are again noted.

The difference between these two readings gives the angle rotated by the telescope. This angle is equal to twice the angle of the prism. Half of this value gives the angle of the prism A. (ii) Angle of minimum deviation (D) The prism is placed on the prism table so that the light from the collimator falls on a refracting face, and the refracted image is observed through the telescope (Fig.). The prism table is now rotated so that the angle of deviation decreases. A stage comes when the image stops for a moment and if we rotate the prism table further in the same direction, the image is seen to recede and the angle of deviation increases. The vertical cross wire of the telescope is made to coincide with the image of the slit where it turns back. This gives the minimum deviation position. The readings of the verniers are noted. Now the prism is removed and the telescope is turned to receive the direct ray and the vertical cross wire is made to coincide with the image. The readings of the verniers are noted. The difference between the two readings gives the angle of minimum deviation D.

The refractive index of the material of the prism $\mu$ is calculated using the formula $\mu = \frac{\sin\left(\frac{A+D}{2}\right)}{\sin\left(\frac{A}{2}\right)}$.

The refractive index of a liquid may be determined in the same way




using a hollow glass prism filled with the given liquid.

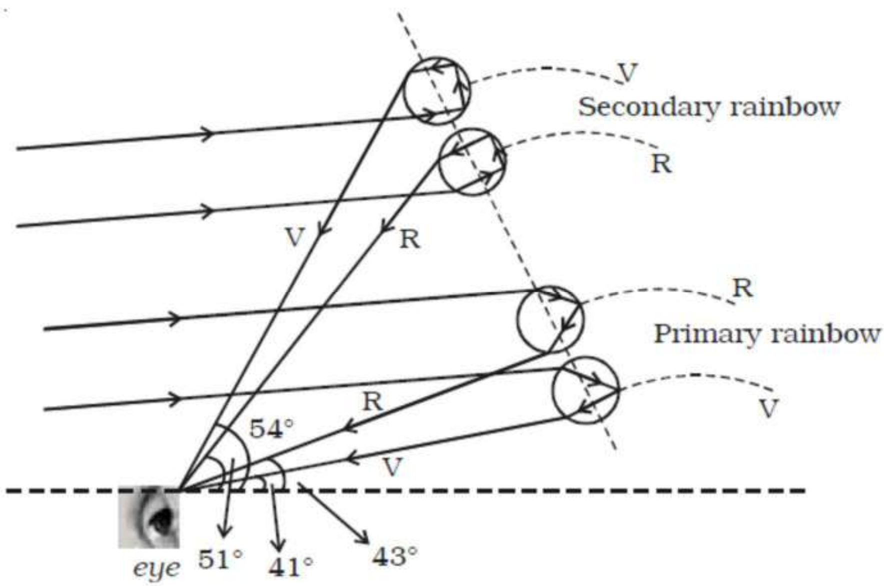
**Rainbow**

One of the spectacular atmospheric phenomena is the formation of rainbow during rainy days. The rainbow is also an example of dispersion of sunlight by the water drops in the atmosphere.

When sunlight falls on small water drops suspended in air during or after a rain, it suffers refraction, internal reflection and dispersion.

If the Sun is behind an observer and the water drops infront, the observer may observe two rainbows, one inside the other. The inner one is called primary rainbow having red on the outer side and violet on the inner side and the outer rainbow is called secondary rainbow, for which violet on the outer side and red on the inner side.

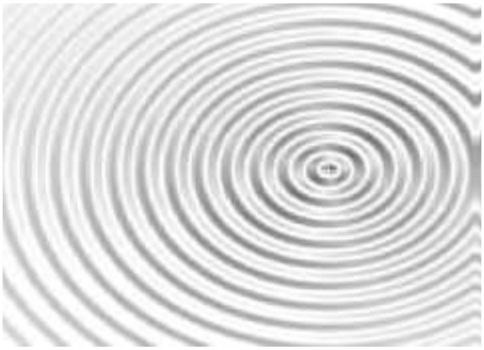
Fig. shows the formation of primary rainbow. It is formed by the light from the Sun undergoing one internal reflection and two refractions and emerging at minimum deviation. It is however, found that the intensity of the red light is maximum at an angle of 43° and that of the violet rays at 41°. The other coloured arcs occur in between violet and red (due to other rain drops). The formation of secondary rainbow is also shown in Fig. It is formed by the light from the Sun undergoing two internal



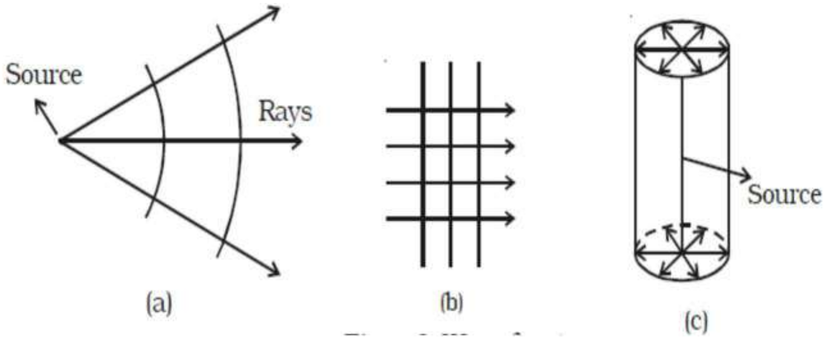

reflections and two refractions and also emerging at minimum deviation. In this case the inner red edge subtends an angle of 51° and the outer violet edge subtends an angle of 54°. This rainbow is less brighter and narrower than the primary rainbow. Both primary and secondary rainbows exhibit all the colours of the solar spectrum. From the ground level an arc of the rainbow is usually visible. A complete circular rainbow may be seen from an elevated position such as from an aeroplane.

**Wave front**

When a stone is dropped in a still water, waves spread out along the surface of water in all directions with same velocity. Every particle on the surface vibrates. At any instant, a photograph of the surface of water would show circular rings on which the disturbance is maximum (Fig.). It is clear that all the particles on such a circle are vibrating in phase,



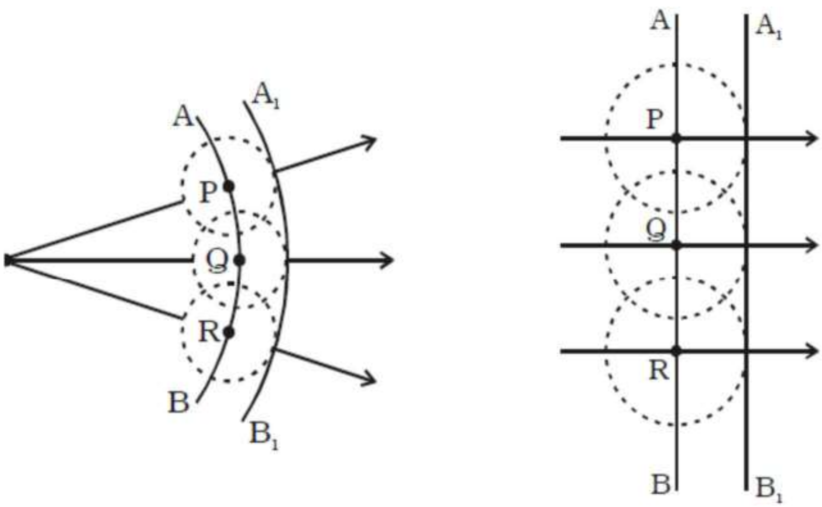
because these particles are at the same distance from the source. Such a surface which envelopes the particles that are in the same state of vibration is known as a wave front. The wave front at any instant is defined as the locus of all the particles of the medium which are in the same state of vibration. A point source of light at a finite distance in an isotropic medium emits a spherical wave front (Fig. a). A point source of light in an isotropic medium at infinite distance will give rise to plane wavefront (Fig. b). A linear source of light such as a slit illuminated by a lamp, will give rise to cylindrical wavefront (Fig. c).



**Huygen's principle**

Huygen's principle helps us to locate the new position and shape of the wavefront at any instant, knowing its position and shape at any previous instant. In other words, it describes the progress of a wave front in a medium. Huygen's principle states that, (i) every point on a given wave front

may be considered as a source of secondary wavelets which spread out with the speed of light in that medium and (ii) the new wavefront is the forward envelope of the secondary wavelets at that instant. Huygen's construction for a spherical and plane wavefront is shown in Fig.a. Let AB represent a given wavefront at a time t = 0. According to Huygen's principle, every point on AB acts as a source of secondary wavelets which travel with the speed of light c. To find the position of the wave front after a time t, circles are drawn with points P, Q, R ... etc as centres on AB and radii equal to ct. These are the traces of secondary wavelets. The arc A1B1 drawn as a forward envelope of the small circles is the new wavefront at that instant. If the source of light is at a large distance, we




(a) (b)

obtain a plane wave front $A_1$ $B_1$ as shown in Fig. b.

**Reflection of a plane wave front at a plane surface**

Let XY be a plane reflecting surface and AB be a plane wavefront incident on the surface at A. PA and QBC are perpendiculars drawn to AB at A and B respectively. Hence they represent incident rays. AN is the normal drawn to the surface. The wave front and the surface are perpendicular to the plane of the paper (Fig.). According to Huygen's principle each point on the wavefront acts as the source of secondary wavelet. By the time, the secondary wavelets from B travel a distance BC, the secondary wavelets from A on the reflecting surface would travel the same distance BC after reflection. Taking A as centre and BC as radius an arc is drawn. From C a tangent CD is drawn to this arc. This tangent CD not only envelopes the wavelets from C and A but also the wavelets from all the points between C and A. Therefore CD is the reflected plane wavefront and AD is the reflected ray.

**Laws of reflection**

(i) The incident wavefront AB, the reflected wavefront CD and the reflecting surface XY all lie in the same plane.

(ii) Angle of incidence i = ∠ PAN = 900 − ∠ NAB = ∠BAC Angle of reflection r = ∠ NAD = 900 − ∠ DAC = ∠DCA In right angled triangles ABC and ADC

∠B = ∠ D = 900

BC=AD and AC is common

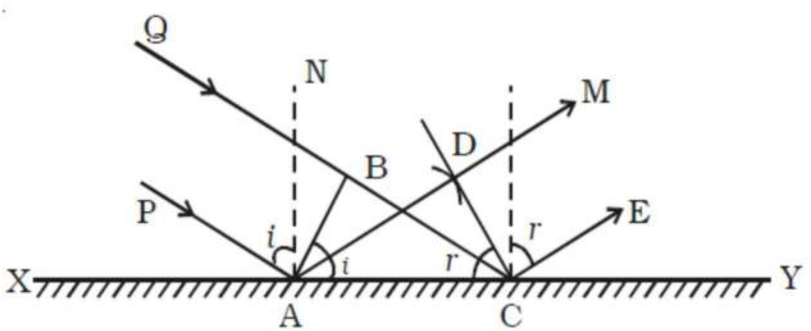
∴The two triangles are congruent

∠ BAC = ∠DCA

i.e. i= r

Thus the angle of incidence is equal to angle of reflection.

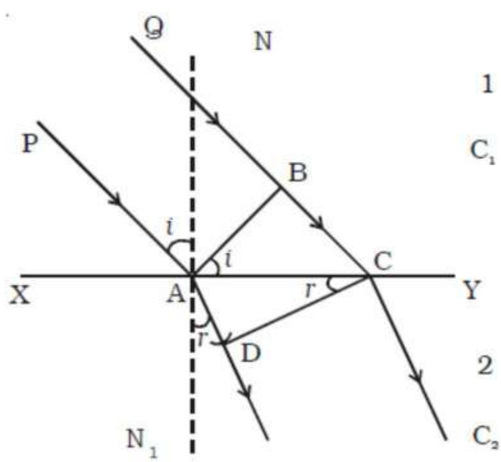
**Refraction of a plane wavefront at a plane surface**



Let XY be a plane refracting surface separating two media 1 and 2 of refractive indices μ1 and μ2 (Fig 5.12). The velocities of light in these two media are respectively c1 and c2. Consider a plane wave front AB incident on the refracting surface at A. PA and QBC are perpendiculars drawn to AB at A and B respectively. Hence they represent incident rays. NAN1 is the normal drawn to the surface. The wave front and the surface are perpendicular to the plane of the paper.

According to Huygen's principle each point on the wave front act as the source of secondary wavelet. By the time, the secondary wavelets from B, reaches C, the secondary wavelets from the point A would travel a distance AD = C2t, where t is the time taken by the wavelets to travel the distance BC.

∴ BC = C1t and AD = $C_2$t = C2 $\frac{BC}{C_1}$ . Taking A as centre and C2 $\frac{BC}{C_1}$ as radius an arc is drawn in the second medium. From C a tangent CD is drawn to this arc. This tangent CD not only envelopes the wavelets from C and A but also the wavelets from all the points between C and A. Therefore CD is the refracted plane wavefront and AD is the refracted ray.




**Laws of refraction**

(i) The incident wave front AB, the refracted wave front CD and the refracting surface XY all lie in the same plane.

(ii) Angle of incidence i = ∠PAN = 90° − ∠ NAB = ∠BAC

Angle of refraction r = ∠$N_1$AD = 90°− ∠ DAC = ∠ACD

$\frac{\sin i}{\sin r} = \frac{BC/_{AC}}{AD/_{AC}} = \frac{BC}{AD} = \frac{BC}{BC.\frac{C_2}{C_1}} = \frac{C_1}{C_2}$= a constant=$_1\mu_2$

$_1\mu_2$ is called the refractive index of second medium with respect to first medium. This is Snell's law of refraction. If $_1\mu_2 > 1$, the first medium is rarer and the second medium is denser. Then

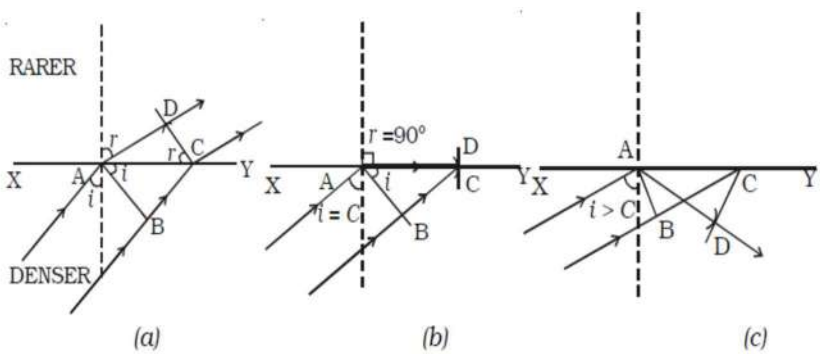
$\frac{C_1}{C_2} > 1$. This means that the velocity of light in rarer medium is greater than that in a denser medium. This conclusion from wave theory is in agreement with the result of Foucault's experiment. It is clear from above discussions that the refractive index of a medium $\mu_m$ is given by

$\mu_m = \frac{\text{velocity of light in vacuum}}{\text{velocity of light in the medium}} = \frac{C_a}{C_m}$

The frequency of a wave does not change when a wave is reflected or refracted from a surface, but wavelength changes on refraction.

i.e. $\mu_m = \frac{C_a}{C_m} = \frac{v\lambda_a}{v\lambda_m} = \frac{\lambda_a}{\lambda_m}$

$\therefore\ \lambda_m = \frac{\lambda_a}{\mu_m}$




where$\lambda_a$ and $\lambda_m$ are the wavelengths in air and medium respectively.

**Total internal reflection by wave theory**

Let XY be a plane surface which separates a rarer medium (air) and a denser medium. Let the velocity of the wavefront in these media be Ca and Cm respectively. A plane wavefront AB passes from denser medium to rarer medium. It is incident on the surface with angle of incidence i. Let r be the angle of refraction.

$\frac{\sin i}{\sin r} = \frac{BC/_{AC}}{AD/_{AC}} = \frac{BC}{AD} = \frac{C_m t}{C_a t} = \frac{C_m}{C_a}$

Since $\frac{C_m}{C_a}$<1 , i is less than r. This means that the refracted wavefront is deflected away from the surface XY. In right angled triangle ADC, there are three possibilities

(i) AD < AC (ii) AD = AC and (iii) AD > AC

(i) AD < AC: For small values of i, BC will be small and so AD > BC but less than AC (Fig.a)

sin r = $\frac{AD}{AC}$ , which is less than unity

i.e r < 90°

For each value of i, for which r < 900, a refracted wavefront is possible

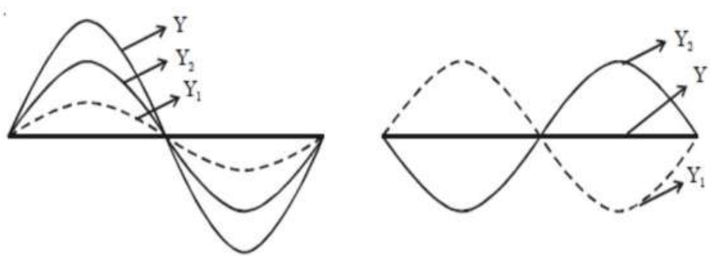
(ii) AD = AC: As i increases r also increases. When AD = AC, sin r = 1 (or) r = 90°. i.e a refracted wavefront is just possible (Fig. b). Now the refracted ray grazes the surface of separation of the two media. The angle of incidence at which the angle of refraction is 90° is called the critical angle C.

(iii) AD > AC: When AD > AC, sin r > 1. This is not possible (Fig. c). Therefore no refracted wave front is possible, when the angle of incidence increases beyond the critical angle. The incident wavefront is totally reflected into the denser medium itself. This is called total internal reflection. Hence for total internal reflection to take place (i) light must travel from a denser medium to a rarer medium and (ii) the angle of incidence inside the denser medium must be greater than the critical angle. i.e i > C.

**Superposition principle**

When two or more waves simultaneously pass through the same medium, each wave acts on every particle of the medium, as if the other waves are not present. The resultant displacement of any particle is the vector addition of the displacements due to the individual waves.

This is known as principle of superposition. If $\overrightarrow{Y_1}$and $\overrightarrow{Y_2}$ represent the



individual displacement then the resultant displacement is given by

$\vec{Y} = \vec{Y_1} + \vec{Y_2}$

**Coherent sources**

Two sources are said to be coherent if they emit light waves of the same wave length and start with same phase or have a constant phase difference. Two independent monochromatic sources, emit waves of same wave length. But the waves are not in phase. So they are not coherent. This is because, atoms cannot emit light waves in same phase and these sources are said to be incoherent sources.

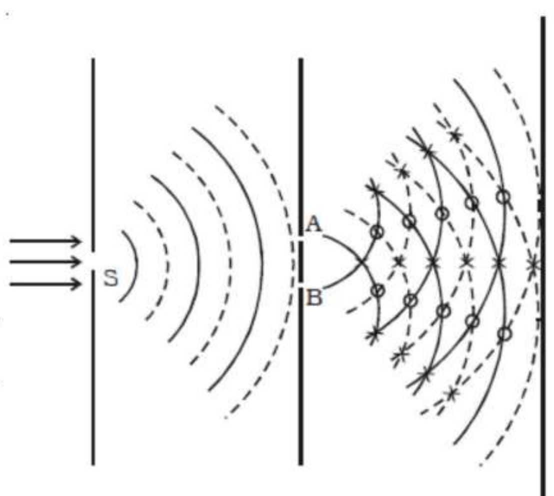
**Phase difference and path difference**

A wave of length $\lambda$ corresponds to a phase of $2\pi$. A distance of $\delta$ corresponds to a phase of $\phi = \frac{2\pi}{\lambda} \times \delta$

**Interference of light**

Two slits A and B illuminated by a single monochromatic source S act as coherent sources. The waves from these two coherent sources travel in the same medium and superpose at various points as hown in Fig. The crest of the wavetrains are shown by thick continuous lines and

troughs are shown by broken lines. At points where the crest of one wave meets the crest of the other wave or the trough of one wave meets the trough of the other wave, the waves are in phase, the displacement is maximum and these points appear bright. These points are marked by crosses (x). This type of interference is said to be constructive interference. At points where the crest of one wave meets the trough of the other wave, the waves are in opposite phase, the displacement is minimum and these points appear dark. These points are marked by circles (O). This type of interference is said to be destructive interference. Therefore, on a screen XY the intensity of light



will be alternatively maximum and minimum i.e. bright and dark bands which are referred as interference fringes. The redistribution of intensity of light on account of the superposition of two waves is called interference.

The intensity of light (I) at a point due to a wave of amplitude (a) is given by $I \propto a^2$. If $a_1$ and $a_2$ are the amplitude of the two interfering waves, then $I_1 \propto a_1^2$ and $I_2 \propto a_2^2$

$\therefore \frac{I_1}{I_2} = \frac{a_1^2}{a_2^2}$

For constructive interference, $I_{max} \propto (a_1 + a_2)^2$ and for destructive interference, $I_{min} \propto (a_1 - a_2)^2$

$\therefore \frac{I_{max}}{I_{min}} = \frac{(a_1+a_2)^2}{(a_1-a_2)^2}$

**Polarisation**

The phenomena of reflection, refraction, interference, diffraction are common to both transverse waves and longitudinal waves. But the transverse nature of light waves is demonstrated only by the phenomenon of polarisation.

**Polarisation of transverse waves.**

Let a rope AB be passed through two parallel vertical slits $S_1$ and $S_2$ placed close to each other. The rope is fixed at the end B. If the free end A of the rope is moved up and down perpendicular to its length, transverse waves are generated with vibrations parallel to the slit. These waves pass through both $S_1$ and $S_2$ without any change in their amplitude. But if $S_2$ is made horizontal, the

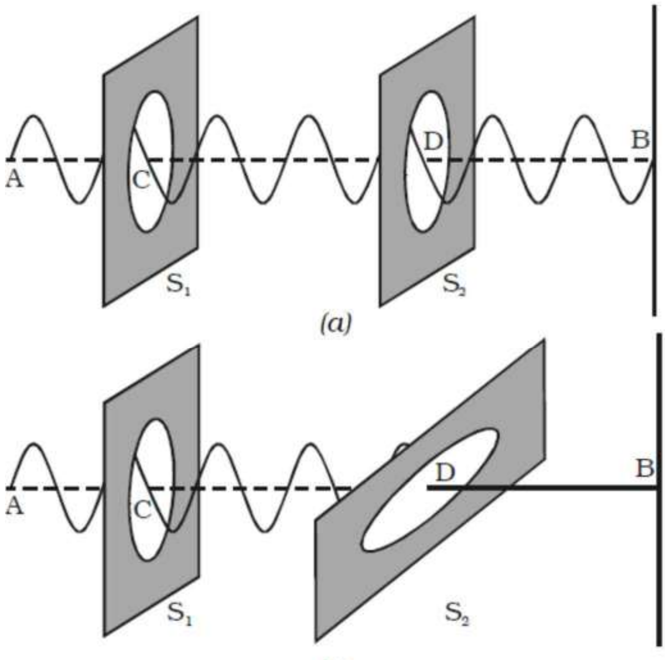
two slits are perpendicular to each other. Now, no vibrations will pass through $S_2$ and amplitude of vibrations will become zero. i.e the portion $S_2B$ is without wave motion as shown in fig .

On the otherhand, if longitudinal waves are generated in the rope by moving the rope along forward and backward, the vibrations will pass through $S_1$ and $S_2$ irrespective of their positions.

This implies that the orientation of the slits has no effect on the propagation of the longitudinal waves, but the propagation of the transverse waves, is affected if the slits are not parallel to each other. A similar phenomenon has been observed in light, when light passes through a tourmaline crystal.

Light from the source is allowed to fall on a tourmaline crystal which is cut parallel to its optic axis (Fig.a).

The emergent light will be slightly coloured due to natural colour of the

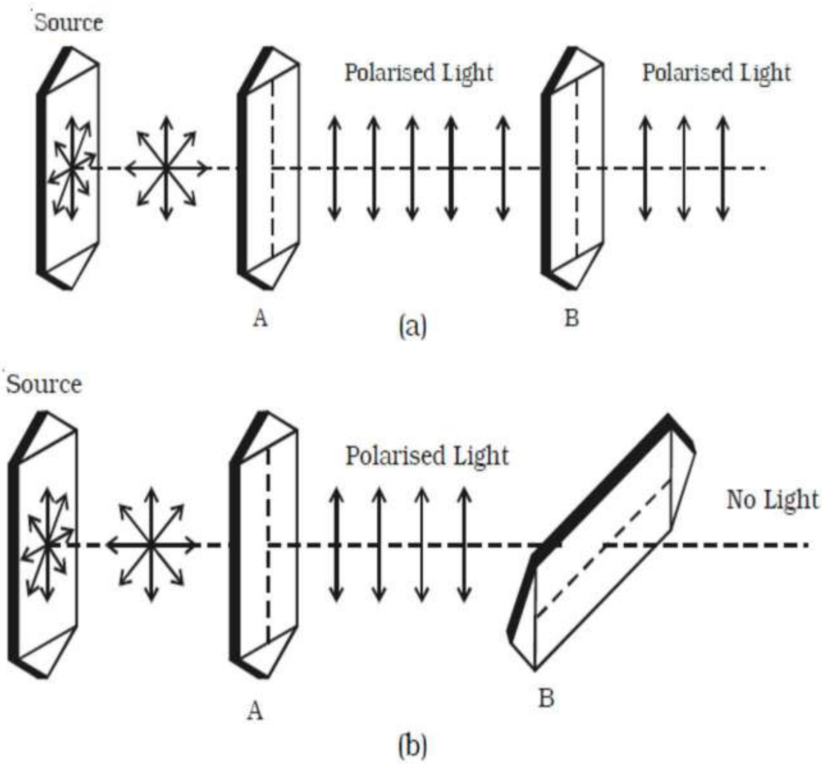



(b)

crystal. When the crystal A is rotated, there is no change in the intensity of the emergent light. Place another crystal B parallel to A in the path of the light. When both the crystals are rotated together, so that their axes are parallel, the intensity of light coming out of B does not change. When the crystal B alone is rotated, the intensity of the emergent light from B gradually decreases. When the axis of B is at right angles to the axis of A, no light emerges from B (Fig.b).

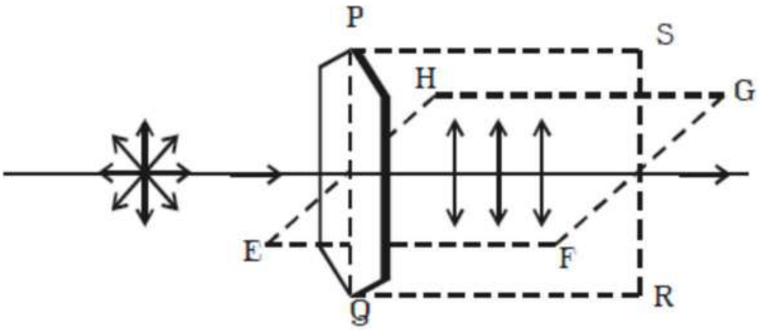
If the crystal B is further rotated, the intensity of the light coming out of B gradually increases and is maximum again when their axis are parallel.

Comparing these observations with the mechanical analogue discussed earlier, it is concluded that the light waves are transverse in nature. Light waves coming out of tourmaline crystal A have their vibrations in only one direction, perpendicular to the direction of propagation. These waves are said to be polarised. Since the vibrations are restricted to only one plane parallel to the axis of the crystal, the light is said to be plane polarised. The phenomenon of restricting the vibrations into a particular plane is known as polarisation.



**Plane of vibration and plane of polarisation**

The plane containing the optic axis in which the vibrations occur is known as plane of vibration. The plane which is at right angles to the plane of vibration and which contains the direction of propagation of the polarised light is known as the plane of polarisation. Plane of polarisation does not contain vibrations in it. In the Fig. PQRS represents

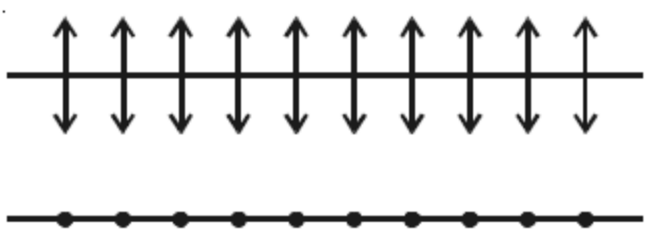



the plane of vibration and EFGH represents the plane of polarisation.

**Representation of light vibrations**

In an unpolarised light, the vibrations in all directions may be supposed to be made up of two mutually perpendicular vibrations. These are represented by double arrows and dots (Fig.).

The vibrations in the plane of the paper are represented by double arrows, while the vibrations



perpendicular to the plane of the paper are represented by dots.

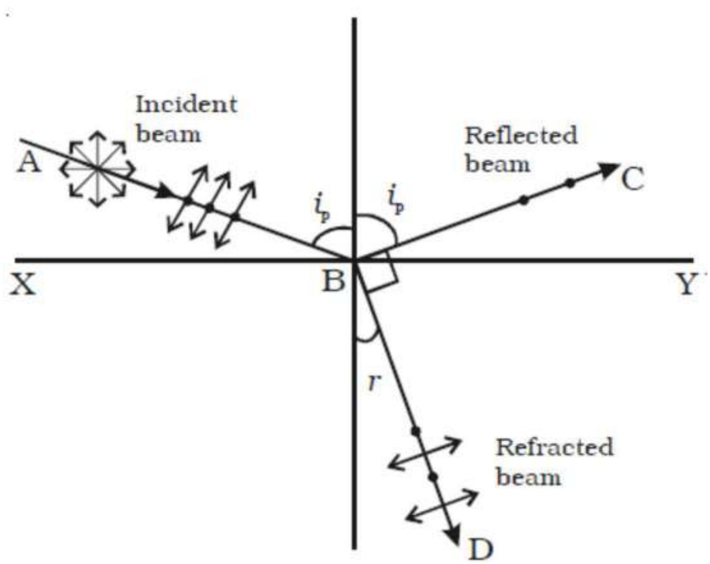
**Polariser and Analyser**

A device which produces plane polarised light is called a polariser. A device which is used to examine, whether light is plane polarised or not is an analyser. A polariser can serve as an analyser and vice versa. A ray of light is allowed to pass through an analyser. If the intensity of the emergent light does not vary, when the analyser is rotated, then the incident light is unpolarised; If the intensity of light varies between maximum and zero, when the analyser is rotated through 90o, then the incident light is plane polarised; If the intensity of light varies between maximum and minimum (not zero), when the analyser is rotated through 90o, then the incident light is partially plane polarised.

**Polarisation by reflection**

The simplest method of producing plane polarised light is by reflection. Malus, discovered that when a beam of ordinary light is reflected from the surface of transparent medium like glass or water, it gets polarised. The degree of polarisation varies with angle of incidence. Consider a beam of unpolarised light AB, incident at any angle on the reflecting glass surface XY.

Vibrations in AB which are parallel to the plane of the diagram are shown by arrows. The vibrations which are perpendicular to the plane of the diagram and parallel to the reflecting surface, shown by dots (Fig.). A part of the light is reflected along BC, and the rest is refracted along BD.




On examining the reflected beam with an analyser, it is found that the ray is partially plane polarised. When the light is allowed to be incident at a particular angle, (for glass it is 57.5°) the reflected beam is completely plane polarised. The angle of incidence at which the reflected beam is completely plane polarised is called the polarising angle ($i_p$).

**Brewster's law**

Sir David Brewster conducted a series of experiments with different reflectors and found a simple relation between the angle of polarization and the refractive index of the medium. It has been observed experimentally that the reflected and refracted rays are at right angles to each other, when the light is incident at polarising angle.

From the above Fig., $i_p + 90° + r = 180°$

$r = 90° - i_p$

From Snell's law,

$$\frac{\sin i_p}{\sin r} = \mu$$

where $\mu$ is the refractive index of the medium (glass)

Substituting for r, we get

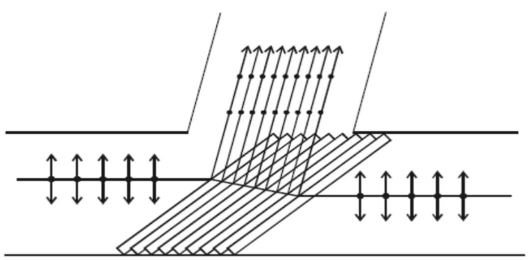
$$\frac{\sin i_p}{\sin(90 - i_p)} = \mu; \frac{\sin i_p}{\cos i_p} = \mu$$

$$\therefore \tan i_p = \mu$$

The tangent of the polarising angle is numerically equal to the refractive index of the medium.

**Pile of plates**

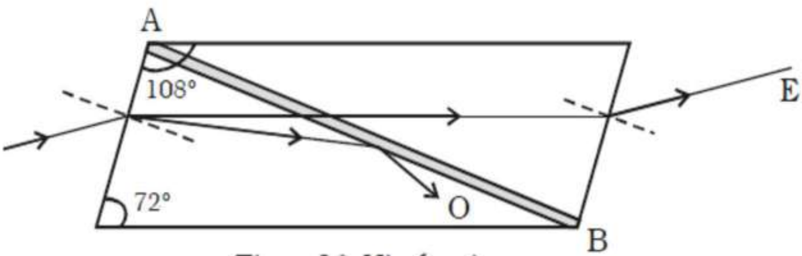
The phenomenon of polarisation by reflection is used in the construction of pile of plates. It consists of a number of glass plates placed one over the other as shown in Fig. in a tube of suitable size. The plates are inclined at an angle of 32.5° to the axis of the tube. A beam of monochromatic light is allowed to fall on the pile of plates along the



axis of the tube. So, the angle of incidence will be 57.5° which is the polarising angle for glass. The vibrations perpendicular to the plane of incidence are reflected at each surface and those parallel to it are transmitted. The larger the number of surfaces, the greater is the intensity of the reflected plane polarised light. The pile of plates is used as a polarizer and an analyser.

**Nicol prism**

Nicol prism was designed by William Nicol. One of the most common forms of the Nicol prism is made by taking a calcite crystal whose length is three times its breadth. It is cut into two halves along the diagonal so that their face angles are 72° and 108°. And the two halves are joined




together by a layer of Canada balsam, a transparent cement as shown in Fig. For sodium light, the refractive index for ordinary light is 1.658 and for extra ordinary light is 1.486. Therefractive index for Canada balsam is 1.550 for both rays, hence Canada balsam does not polarise light. A monochromatic beam of unpolarised light is incident on the face of the nicol prism. It splits up into two rays as ordinary ray (O) and extraordinary ray (E) inside the nicol prism (i.e) double refraction takes place. The ordinary ray is totally internally reflected at the layer of Canada balsam and is prevented from emerging from the other face. The extraordinary ray alone is transmitted through the crystal which is plane polarised. The nicol prism serves as a polariser and also an analyser.

**Polaroids**

A Polaroid is a material which polarises light. The phenomenon of selective absorption is made use of in the construction of polariods.

There are different types of polaroids. A Polaroid consists of micro crystals of herapathite (an iodosulphate of quinine). Each crystal is a doubly refracting medium, which absorbs the ordinary ray and transmits only the extra ordinary ray. The modern polaroid consists of a large number of ultra microscopic crystals of herapathite embedded with their optic axes, parallel, in a matrix of nitro –cellulose. Recently, new types of polariod are prepared in which thin film of polyvinyl alcohol is used. These are colourless crystals which transmit more light, and give better polarisation.

**Uses of Polaroid**

- Polaroids are used in the laboratory to produce and analyse plane polarised light.
- Polaroids are widely used as polarising sun glasses.
- They are used to eliminate the head light glare in motor cars.
- They are used to improve colour contrasts in old oil paintings.
- Polaroid films are used to produce three – dimensional moving pictures.
- They are used as glass windows in trains and aeroplanes to control the intensity of light. In aeroplane one polaroid is fixed outside the window while the other is fitted inside which can be rotated. The intensity of light can be adjusted by rotating the inner polaroid.
- Aerial pictures may be taken from slightly different angles and when viewed through polaroids give a better perception of depth.
- In calculators and watches, letters and numbers are formed by liquid crystal display (LCD) through polarisation of light.
- Polarisation is also used to study size and shape of molecules.

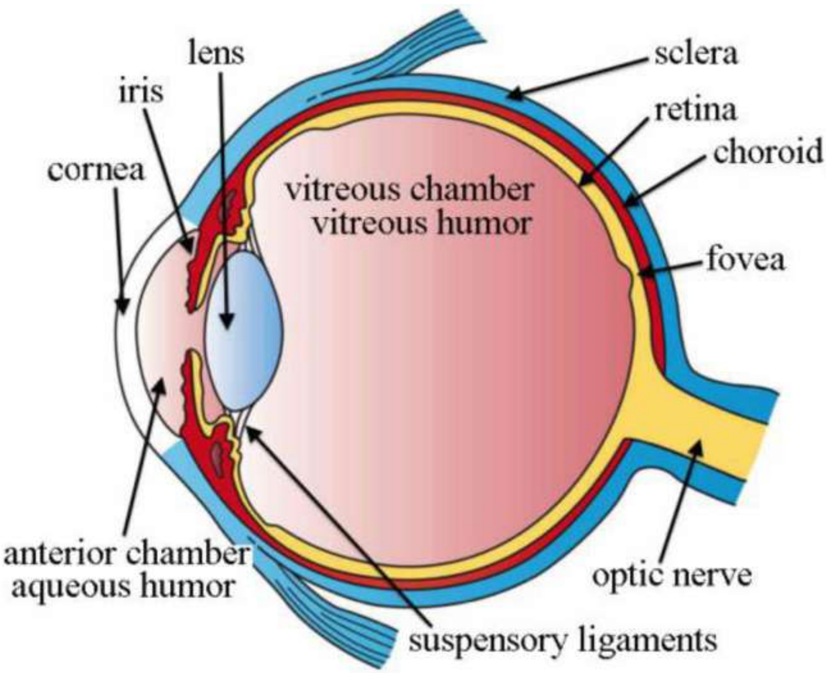
**Optical instruments**

The Human Eye

The human eye is an optical instrument that enables us to view all the objects around us is a very complex organ. Let us study the structure of the human eye. The white protective membrane seen when looked into the eye directly is a Sclera. It is tuff, opaque and fibrous outer layer of the eyeball.

The circular part is the Iris. The color of the eye is determined by the color of the iris. The center transparent area of the iris is the Pupil. The iris works like the shutter of the camera. It absorbs most of the light falling on it and allows it to pass through the pupil.

The amount of light that enters the inner part of the eye depends on the size of the pupil. In bright light, the iris contracts the pupil to restrict the light, whereas in low light it widens the pupil to emit more light into the eye. The eyeball is spherical in shape. The retina of the eye is able to




detect the light and its color because of the presence of senses known as rods and cones.

Light entering the human eye is first refracted by the cornea. The refracted light is then incident on an iris. The lens is just behind the iris and light after refracted through the pupil falls on it and forms a sharp image. Image formation exactly on the retina enables us to see the object clearly.

**Image Formation**

The cornea and lens are at the interface between the physical world of light and the neural encoding of the visual pathways. The cornea and lens bring light into focus at the light sensitive receptors in our retina and initiate a series of visual events that result in our visual experience.

The initial encoding of light at the retina is but the first in a series of visual transformations: The stimulus incident at the cornea is transformed into an image at the retina. The retinal image is transformed into a neural response by the light sensitive elements of the eye, the photoreceptors. The photoreceptor responses are transformed to a neural response on the optic nerve. The optic nerve representation is transformed into a cortical representation, and so forth. We can describe most of our understanding of these transformations, and thus most of our understanding of the early encoding of light by the visual pathways by using linear systems theory. Because all of our visual experience is limited by the image formation within our eye, we begin by describing this transformation of the light signal and we will use this analysis as an introduction to linear methods.

**Accommodation** is the process by which the vertebrate eye changes optical power to maintain a clear image or focus on an object as its distance varies. In this, distances vary for individuals from the far point the maximum distance from the eye for which a clear image of an object can be seen, to the near point the minimum distance for a clear image. Accommodation usually acts like a reflex, including as part of the accommodation-vergence reflex, but it can also be consciously controlled. Mammals, birds and reptiles vary the optical power by changing the form of the elastic lens using the ciliary body (in humans up to 15 dioptres). Fish and amphibians vary the power by changing the distance between a rigid lens and the retina with muscles.

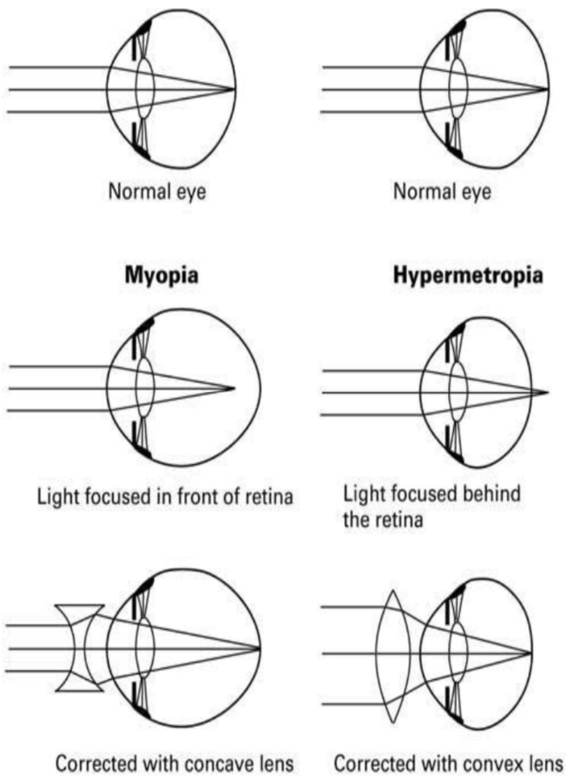
The young human eye can change focus from distance (infinity) to as near as 6.5 cm from the eye. This dramatic change in focal power of the eye of approximately 15 dioptres (the reciprocal of focal length in metres) occurs as a consequence of a reduction in zonular tension induced by ciliary muscle contraction. This process can occur in as little as 224 ± 30 milliseconds in bright light. The amplitude of accommodation declines with age. By the fifth decade of life the accommodative amplitude can decline so that the near point of the eye is more remote than the reading distance. When this occurs the patient is presbyopic. Once presbyopia occurs, those who are emmetropic (do not require optical correction for distance vision) will need an optical aid for near vision; those who are myopic (nearsighted and require an optical correction for distance vision), will find that they see better at near without their distance correction; and those who are hyperopic (farsighted) will find that they may need a correction for both distance and near vision. Note that these effects are most noticeable when the pupil is large; i.e. in dim light. The age-related decline in accommodation occurs almost universally to less than 2 dioptres by the time a person reaches 45 to 50 years, by which time most of the population will have noticed a decrease in their ability to focus on close objects and hence require glasses for reading or bifocal lenses. Accommodation decreases to about 1 dioptre at the age of 70 years. The dependency of accommodation amplitude on age is graphically summarized by Duane's classical curves.

**Eye defects**

Myopia: (nearsightedness) This is a defect of vision in which far objects appear blurred but near objects are seen clearly. The image is focused in front of the retina rather than on it usually because the eyeball is too long or the refractive power of the eye’s lens too strong. Myopia can be corrected by wearing glasses/contacts with concave lenses these help to focus the image on the retina.

Hyperopia: (farsightedness) This is a defect of vision in which there is difficulty with near vision but far objects can be seen easily. The image is focused behind the retina rather than upon it. This occurs when the eyeball is too short or the refractive power of the lens is too weak. Hyperopia can be corrected by wearing glasses/contacts that contain convex lenses.

Astigmatism: This defect is when the light rays do not all come to a single focal point on the retina, instead some focus on the retina and some focus in front of or behind it. This is usually caused by a non-uniform curvature of the cornea. A typical symptom of astigmatism is if you are looking at a pattern of lines placed at various angles and the lines running in one direction appear sharp whilst those in other directions appear blurred. Astigmatism can usually be corrected by using a special spherical cylindrical lens; this is placed in the out-of-focus axis.




Presbyopia

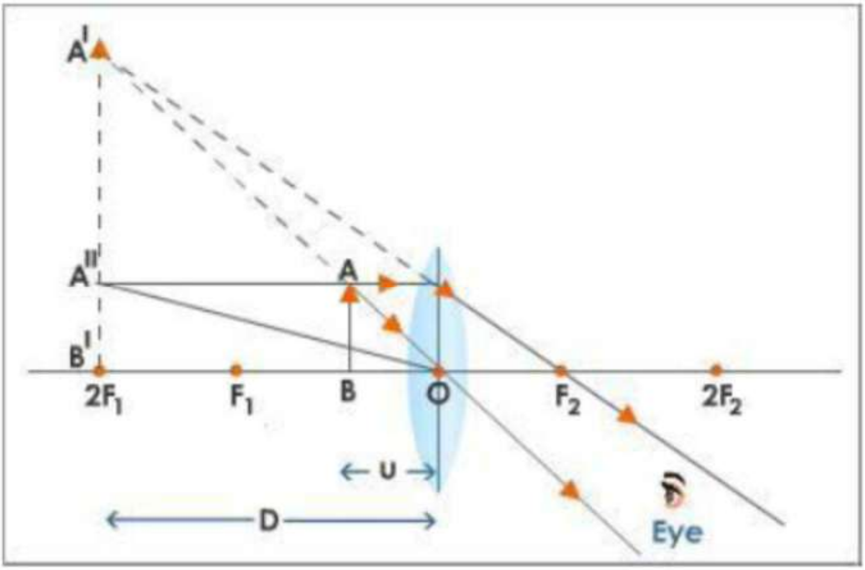
Presbyopia is an age-related process. It is a gradual thickening and loss of flexibility of the natural lens inside your eye.

These age-related changes occur within the proteins in the lens, making the lens harder and less elastic over time. Age-related changes also take place in the muscle fibers surrounding the lens. With less elasticity, it gets difficult for the eyes to focus on close objects. This defect is common in old-people and is caused by the weakening of ciliary muscles. In this defect, both the near-point and the far-point are affected. It is corrected with the help of a bifocal length which is a combination of convex lens and a concave lens.

**The Microscope**

As we all know Microscope is an optical instrument used to view small object. Let us first talk about the simple microscope.

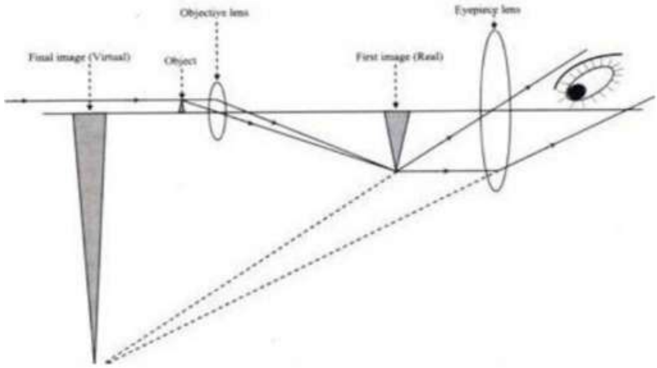
A simple microscope is an optical instrument, we use for the magnification of small objects to get a clear image or vision. It is a convex lens having a short focal length. This microscope is at a small distance from the object for the magnification and hence this forms a virtual image. The simple microscope enables us to view very small




letters and figures. Watchmakers also make use of these. Now let us see what compound microscope is.

With a compound microscope, we get very large values of magnification. We use this microscope to see microscopic objects like microorganisms. It comprises of two convex lenses and magnification occurs in both of these lenses. The components of a compound microscope are eyepiece, objective lens, fine and rough adjustment screw.

Shaping light into an image: first principles




A telescope must bend or reflect light rays to make them converge to a small (ideally smaller that the atmospheric seeing size for ground telescopes) zone in the focal plane.
At optical/nearIR wavelengths, this is done with mirrors or lenses

- choice of materials is important for lenses and mirrors
- coatings (especially for mirrors) are essential to the telescope performance
- optical surfaces of mirrors and lenses must be accurately controlled

At longer wavelength (radio), metal panels or grid can be used
At shorter wavelength (X ray, Gamma ray), materials are poorly reflecive (see next slide)
The telescope must satisfy the previous requirement over a finite field of view with high throughput
Field of view + good image quality → telescope designs with multiple elements (this will be covered in the next lecture)
High throughput over large field of view requires good coating and an optical design which can transmit the full size of the beam for any point in the field of view (no beam clipping)
Refracting telescopes (lenses)

Lenses are easy to manufacture when small First telescopes were refractors (Galileo)
Refractors suffer from serious limitations:
Chromaticity

- refraction index is chromatic: a simple lens has a focal length which changes with wavelength
- achromatic lense designs use combination of several materials to reduce chromaticity

Difficult to implement for large telescopes:

- lens thickness increases with diameter, and needs to be held at its edges
- lens is located at the front of the telescope: center of mass is close to the top of the telescope (top-heavy)

Refracting telescopes (lenses)
Chromaticity problem can be mitigated by adopting long focal length - Refracting telescopes used to be very long and narrow field of view More recently, developments in lens design and manufacturing technology have led to high quality short refractors - Refractors are still used in astronomy for wide field small diameter systems, and the same technology is used to correct for aberrations in wide field reflecting telescopes.
Reflecting telescopes (mirrors) Challenges:
Light bounces back toward the object: focal plane in front of the telescope or secondary mirror needs to be used to send light to instrument / viewer.
Mirrors have ~4x tighter optical surface tolerances than lenses For surface defect h, wavefront error is 2h in reflection, h(n-1) in refraction → mirrors often need to be made non-spherical with 100nm surface accuracy.
Mirrors need to be reflective at first, mirrors were made of metal, then glass
Advantages:

- Achromatic by design (reflection is achromatic)
- Ideally suited for large telescopes:

Mirror is supported from the bottom, and can be thin with active control

Mirror is located at the back of the telescope: center of mass is low

Telescope tube can be relatively short (F ratio of modern large telescopes is ~1 to 2)

**Telescope**

The telescope is of two types. One is the reflecting type and another one is the refracting type. Reflecting telescopes are the ones which do not use lenses at all. They use mirrors to focus the light together. The type of mirror used is a concave mirror.

Mirrors also bend the light together, except that they do it by reflecting the light instead of bending it. Refracting telescopes work by using two lenses to focus the light and make it look like the object is closer to you than it really is. Both the lenses are in a shape of 'convex'. Convex lenses work by bending light inwards.

**The Astronomical Telescope**

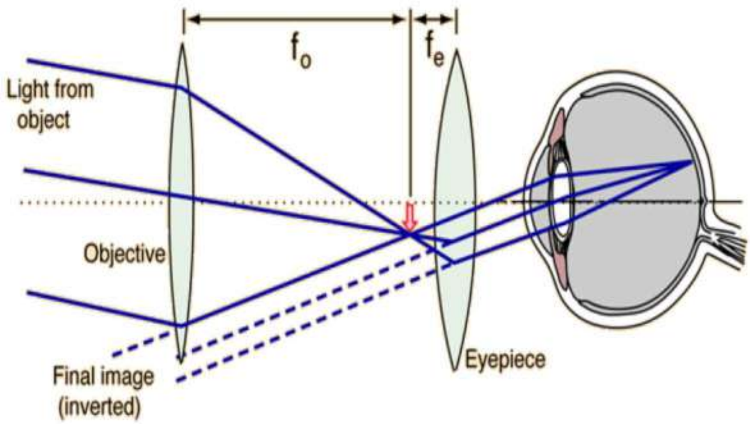
The astronomical telescope makes use of two positive lenses: the objective, which forms the image of a distant object at its focal length, and the eyepiece, which acts as a simple magnifier with which to view the image formed by the objective. Its length is equal to the sum of the focal lengths of the objective and eyepiece, and its angular magnification is -fo/fe , giving an inverted image.

The astronomical telescope can be used for terrestrial viewing, but seeing the image upside down is a definite inconvenience. Viewing stars upside down is no problem. Another inconvenience for terrestrial viewing is the length of the astronomical telescope, equal to the sum of the focal lengths of the objective and eyepiece lenses. A shorter telescope with upright viewing is the Galilean telescope.

Magnifying Power

Two or more lenses may be combined, to form an optical instrument with greater utility than a single lens. For example, a telescope provides an enlarged image of a distant object and a microscope magnifies an object close at hand. The design and magnifying power of telescopes will be studied in this laboratory exercise. In both the telescope and microscope, an objective lens first forms a real image of the object.

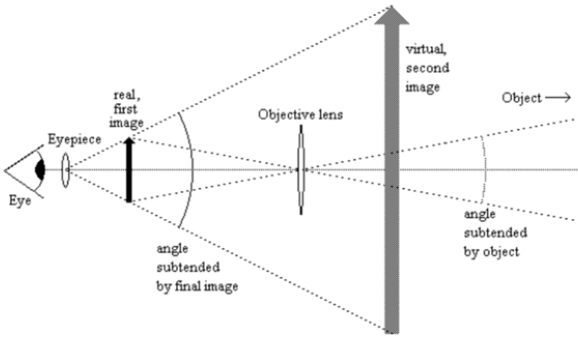
This first image then becomes the object for the second lens, called the eyepiece. Proper adjustment of the instrument causes the eyepiece to form the final image, an imaginary image located at a comfortable viewing distance from the eye. If the object is very far away, the first image is located at the focal point of the objective lens. If the eyepiece is positioned so that its primary focal point is at the location of the first




image, then the final, virtual image is an infinite distance away, where a relaxed eye may focus on it.

The angular magnification is defined as the angle subtended by the final image, divided by the angle subtended when the object is viewed by the naked eye. This ratio can be shown to be equal to the distance from the first image to the objective lens divided by the distance from the first image to the eyepiece. This is the same as the ratio of the two focal lengths in the ideal case where the two focal points coincide.

For small angles, the subtended angle is proportional to the height divided by the distance from the observer. Thus, the angular magnification may be determined by measuring the apparent image height and object height at some common distance from the observer.




The common distance from the observer would cancel out so that the angular magnification is also equal to the ratio of the apparent image height to the apparent object height.

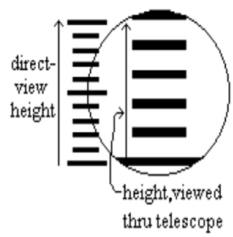
Procedure: First, we must measure the focal lengths of the lenses. The focal point of a lens is defined as the point where rays which come into the lens parallel to the axis will converge. Light rays coming into the lens from a point far away will be approximately parallel to the axis. So the image of a faraway object will be located at the focal point of the lens. Place a converging lens $10^{-20}$ cm from a screen, on the meter stick. Aim the axis of the lens at the lighted object, across the room. Adjust the position of the lens until a sharp image of the object is formed on the screen. Subtract the two positions to get the focal length of the lens. Record the longer focal length (about 20 cm) as $f_A$ and the shorter one (about 10 cm) as $f_B$.

$f_A$ = ____________ cm $f_B$ = ____________ cm $f_C$ = -15 cm

**Magnifying power of Astronomical Telescope**

Place the optical bench at the opposite end of the lab table, aimed at the light source. Position lens B (short positive focal length) about 5 cm from the near end, and lens A (long positive focal length) about 30 cm away. Place the screen at the focal point of lens B (about 10 cm away). Adjust the position of lens A until a real image of the light is formed on the screen. Remove the screen and view the magnified virtual image of the real image of the light source through lens B. Adjust the position of lens B as necessary to obtain a sharp image.

Aim your telescope at a scale mounted on the wall, across the room. Adjust the objective lens until the largest, well focused image that you can get is observed. With one eye, view the image of the scale through the telescope and with the other eye, look directly at the scale. Your brain should eventually superimpose the two images. Align your views so that the lower ends of the two images coincide. Determine where the top mark of the non- telescopic image coincides with the telescopic image. (Note that the scale markings are two cm apart.) Record the two corresponding heights (image height is negative if inverted) and the magnification, equal to the ratio of the two heights. Notice that the full height on the directly viewed scale is actually the apparent size of the image, but the reading on the scale seen through the telescope is the


unmagnified height of the object being observed.

**Resolving Power:**

It is defined as the inverse of the distance or angular separation between two objects which can be just resolved when viewed through the optical instrument.

**Resolving Power of Telescope:**

In telescopes, very close objects such as binary stars or individual stars of galaxies subtend very small angles on the telescope. To resolve them we need very large apertures. We can use Rayleigh's to determine the resolving power. The angular separation between two objects must be

$\triangle\theta = 1.22\ \lambda d$

Resolving power = $1\triangle\theta = d1.22\ \lambda$

Thus higher the diameter d, better the resolution. The best astronomical optical telescopes have mirror diameters as large as 10m to achieve the best resolution. Also, larger wavelengths reduce the resolving power and consequently radio and microwave telescopes need larger mirrors.

**Resolving Power of Microscope:**

For microscopes, the resolving power is the inverse of the distance between two objects that can be just resolved. This is given by the famous Abbe's criterion given by Ernst Abbe in 1873 as

$\triangle$ d = λ2 n sin θ

**Resolving power** = 1$\triangle$ d = 2n sin θλ

Where n is the refractive index of the medium separating object and aperture. Note that to achieve high-resolution n sin θ must be large. This is known as the Numerical aperture.

Thus, for good resolution:

sin θ must be large. To achieve this, the objective lens is kept as close to the specimen as possible.

A higher refractive index (n) medium must be used. Oil immersion microscopes use oil to increase the refractive index. Typically for use in biology studies, this is limited to 1.6 to match the refractive index of glass slides used. (This limits reflection from slides). Thus the numerical aperture is limited to just 1.4-1.6. Thus, optical microscopes (if you do the math) can only image to about 0.1 microns. This means that usually organelles, viruses and proteins cannot be imaged.

Decreasing the wavelength by using X-rays and gamma rays. While these techniques are used to study inorganic crystals, biological samples are usually damaged by x-rays and hence are not used.