as i increases, r_1 increases as $\frac{\sin i}{\sin r_1} = n$ but r_2 and e decrease. However, variation in δ with increasing i is different. It is as plotted in the Fig. 9.19.

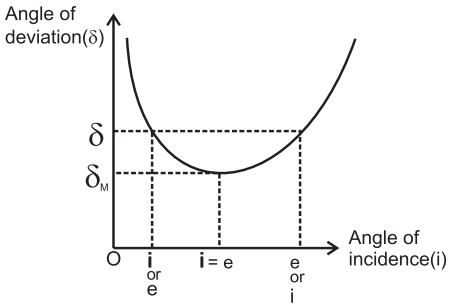


Fig. 9.19: Deviation curve for a prism.

It shows that, with increasing values of i, the angle of deviation δ decreases initially to a certain minimum (δ_m) and then increases. It should also be noted that the curve is not a symmetric parabola, but the slope in the part after is less. It is clear that except at $\delta = \delta_m$, (Angle of minimum deviation) there are two values of i for any given δ. By applying the principle of reversibility of light to path PQRS it is obvious that if one of these values is i, the other must be e and vice versa. Thus at $\delta = \delta_m$, we have $i = e$. Also, in this case, $r_1 = r_2$ and $A = r_1 + r_2 = 2r \ \therefore r = \frac{A}{2}$

Only in this case QR is parallel to base BC and the figure is symmetric.

Using these in Eq. (9.11), we get,

$$i + i = A + \delta_m \therefore i = \frac{(A + \delta_m)}{2}$$

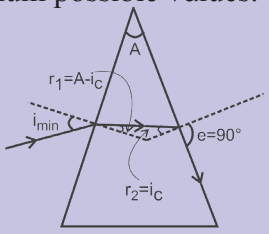
According to Snell's law,

$$\therefore n = \frac{\sin\left(\frac{A + \delta_m}{2}\right)}{\sin\left(\frac{A}{2}\right)} \qquad \text{--- (9.12)}$$

Equation (9.12) is called prism formula.

Example 9.9: For a glass (n =1.5) prism having refracting angle 60^0, determine the range of angle of incidence for which emergent ray is possible from the opposite surface and the corresponding angles of emergence. Also calculate the angle of incidence for which $i = e$. How much is the corresponding angle of minimum deviation?

(I) Grazing emergence and minimum angle of incidence: At the point of emergence, the ray travels form a denser medium into rarer (popular prisms are of denser material, kept in rarer). Thus if $r_2 = \sin^{-1}\left(\frac{1}{n}\right)$ is the critical angle, the angle of emergence $e = 90^0$. This is called grazing emergence or we say that the ray just emerges. Angle of prism A is constant for a given prism and $A = r_1 + r_2$. Hence the corresponding r_1 and i will have their minimum possible values.



Grazing emergence
or the ray just emerges

(II) For commonly used glass prisms,

$$n = 1.5, \sin^{-1}\left(\frac{1}{n}\right) = \sin^{-1}\left(\frac{1}{1.5}\right)$$

$$= 41^0 49' = (r_2)_{max}$$

If prism is symmetric (equilateral),

$$A = 60^0 \therefore r_1 = 60^0 - 41^0 49' = 18^0 11'$$

$$\therefore n = 1.5 = \frac{\sin(i_{min})}{\sin 18^0 11'} \therefore i_{min} = 27^0 55' \cong 28^0.$$

(III) For a symmetric (equilateral) prism, the prism formula can be written as

$$n = \frac{\sin\left(\frac{60 + \delta_m}{2}\right)}{\sin\left(\frac{60}{2}\right)} = \frac{\sin\left(30 + \frac{\delta_m}{2}\right)}{\sin(30)}$$

$$= 2\sin\left(30 + \frac{\delta_m}{2}\right)$$

(IV) For a prism of denser material, kept in a rarer medium, the incident ray deviates towards the normal during the first refraction and away from the normal during second refraction. However, during both the refractions it deviates towards the base only.

Solution: As shown in the box above, $i_{min} = 27^0 55'$. Angle of emergence for this is $e_{max} = 90^0$.

From the principle of reversibility of light, $i_{max} = 90^0$ *and* $e_{min} = 27^0 55'$

Also, from the box above,

$$n = \frac{\sin\left(\frac{60+\delta_m}{2}\right)}{\sin\left(\frac{60}{2}\right)}$$

$$= \frac{\sin\left(30+\frac{\delta_m}{2}\right)}{\sin(30)} = 2\sin\left(30+\frac{\delta_m}{2}\right)$$

$$\therefore 1.5 = 2\sin\left(30+\frac{\delta_m}{2}\right) \quad \therefore 0.75 = \sin\left(30+\frac{\delta_m}{2}\right)$$

$$\therefore \left(30+\frac{\delta_m}{2}\right) = 48^0 35'$$

$$\therefore \frac{\delta_m}{2} = 18^0 35' \quad \therefore \delta_m = 37^0 10'$$

$i + e = A + \delta$ and $i = e$ for $\delta = \delta_m$

$\therefore\ i + i = 60 + 37°\ 10' = 97°10' \ \therefore\ i = 48°35'$

Thin prisms: Prisms having refracting angle less than 10^0 $\left(A < 10^0\right)$ are called thin prisms. For such prisms we can comfortably use $\sin\theta \cong \theta$. For such prisms to deviate the incident ray towards the base during both refractions, it is essential that i should also be less than 10^0 so that all the other angles will also be small.

Thus

$$n = \frac{\sin i}{\sin r_1} \cong \frac{i}{r_1} \text{ and } n = \frac{\sin e}{\sin r_2} \cong \frac{e}{r_2}$$

$\therefore i \cong nr_1$ and $e \cong nr_2$

Using these in Eq. (9.11), we get,

$$i + e = nr_1 + nr_2 = n(r_1 + r_2) = nA = A + \delta$$

$$\therefore \delta = A(n-1) \qquad \text{--- (9.13)}$$

A and n are constant for a given prism. Thus, for a thin prism, for small angles of incidences, angle of deviation is constant (independent of angle of incidence).

Angular dispersion and mean deviation: As discussed earlier, if a polychromatic beam is incident upon a prism, the emergent beam consists of all the individual colours angularly separated. This is angular dispersion (Fig. 9.20).

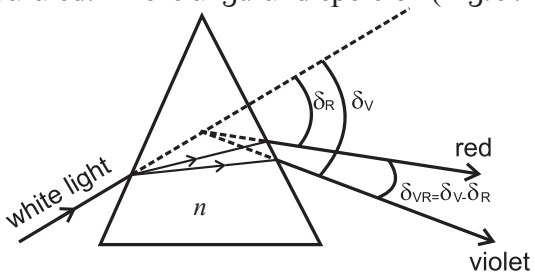


Fig. 9.20: Angular dispersion through a prism.

It is measured for any two component colours.

$$\therefore \delta_{21} = \delta_2 - \delta_1$$

Normally we do it for extreme colours.

For white light, violet and red are the extreme colours.

$$\therefore \delta_{VR} = \delta_V - \delta_R$$

Using deviation for thin prism (Eq. 9.13), we can write

$$\therefore \delta_{21} = \delta_2 - \delta_1 = A(n_2 - 1) - A(n_1 - 1)$$

$$= A(n_2 - n_1)$$

where n_1 and n_2 are refractive indices for the two colours.

Also,

$$\delta_{VR} = \delta_V - \delta_R = A(n_V - 1) - A(n_R - 1)$$

$$= A(n_V - n_R) \qquad \text{--- (9.14)}$$

Yellow is practically chosen to be the mean colour for violet and red.

This gives mean deviation

$$\delta_{VR} = \frac{\delta_V + \delta_R}{2} \cong \delta_Y = A(n_Y - 1) \qquad \text{--- (9.15)}$$

Do you know ?

(i) If you see a rainbow widthwise, yellow appears to be centrally located. Hence angular deviation of yellow is average for the entire colour span. This may be the reason for choosing yellow as the mean colour. Remember, red band is widest and violet is much thinner than blue.

(ii) While obtaining the expression for ω, we have used thin prism formula for δ. However, the expression for ω (equation 9.16) is valid as well for equilateral prisms or right-angled prisms.

Dispersive power: Ability of an optical material to disperse constituent colours is its dispersive power. It is measured for any two colours as the ratio of angular dispersion to the mean deviation for those two colours. Thus, for the extreme colours of white light, dispersive power is given by

$$\omega = \frac{\left[\delta_V - \delta_R\right]}{\left[\frac{\delta_V + \delta_R}{2}\right]} \cong \frac{\delta_V - \delta_R}{\delta_Y}$$

$$= \frac{A\left(n_V - n_R\right)}{A\left(n_Y - 1\right)} = \frac{n_V - n_R}{n_Y - 1} \qquad \text{--- (9.16)}$$

As ω is the ratio of same physical quantities, it is unitless and dimensionless quantity. From the expression in terms of refractive indices it should be understood that dispersive power depends only upon refractive index (hence material only) and not upon the dimensions of prism. For commonly used glasses it is around 0.03.

Example 10: For a dense flint glass prism of refracting angle 10^0, obtain angular deviation for extreme colours and dispersive power of dense flint glass. ($n_{red} = 1.712$, $n_{violet} = 1.792$)

$$\delta_V = A(n_V - 1) = 10(1.792 - 1) = (7.92)°$$

$$\delta_R = A(n_R - 1) = 10(1.712 - 1) = (7.12)°$$

$\therefore$ Angular dispersion, $\delta_{VR} = \delta_V - \delta_R = \left(0.8\right)^0$

dispersive power, ω =

$$= \frac{\delta_V - \delta_R}{\left(\frac{\delta_V + \delta_R}{2}\right)}$$

$$= 2\left(\frac{7.92 - 7.12}{7.92 + 7.12}\right)$$

$$= \frac{2 \times 0.8}{15.04} = 0.1064$$

(This is much higher than popular crown glass)

9.9 Some natural phenomena due to Sunlight:

Mirage: On a hot clear Sunny day, along a level road, a pond of water appears to be there ahead. However, if we physically reach the spot, there is nothing but the dry road and water pond again appears ahead. This illusion is called a mirage (Fig. 9.21).

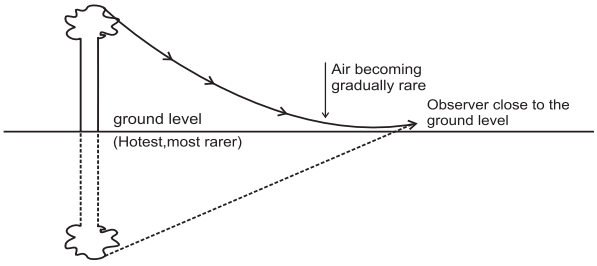


Fig. 9.21: The Mirage.

On a hot day the air in contact with the road is hottest and as we go up, it gets gradually cooler. The refractive index of air thus increases with height. As shown in the figure, due to this gradual change in the refractive index, the ray of light coming from the top of an object becomes more and more horizontal as it almost touches the road. For some reason (mentioned later) it bends above. Then onwards, upward bending continues due to denser air. As a result, for an observer, it appears to be coming from below thereby giving an illusion of reflection from an (imaginary) water surface.

Rainbow: Undoubtedly, rainbow is an eye-catching phenomenon occurring due to rains and Sunlight. It is most popular because it is observable from anywhere on the Earth. A few lucky persons might have observed two rainbows simultaneously one above the other. Some might have seen a complete circular rainbow from an aeroplane (Of course, this time it's not a bow!). Optical phenomena discussed till now are sufficient to explain the formation of a rainbow.

The facts to be explained are:

(i) It is seen during rains and on the opposite side of the Sun.

(ii) It is seen only during mornings and evenings and not throughout the day.

(iii) In the commonly seen rainbow red arch is outside and violet is inside.

(iv) In the rarely occurring concentric secondary rainbow, violet arch is outside and red is inside.

(v) It is in the form of arc of a circle.

(vi) Complete circle can be seen from a higher altitude, i.e., from an aeroplane.

(vii) Total internal reflection is not possible in this case.

Conditions necessary for formation of a rainbow: Light shower with relatively large raindrops, morning or evening time and enough Sunlight.

Optical phenomena involved: During the formation of a rainbow, the rays of Sunlight incident on water drops, deviate and disperse during refraction, internally (NOT total internally) reflect once (for primary rainbow) or twice (for secondary rainbow) and finally refract again into air. At all stages there is angular dispersion which leads to clear separation of the colours.

Primary rainbow: Figure 9.22 (a) shows the optical phenomena involved in the formation of a primary rainbow due to a spherical water drop.

Do you know ?

Possible reasons for the upward bending at the road during mirage could be:

(i) Angle of incidence at the road is glancing. At glancing incidence, the reflection coefficient is very large which causes reflection.

(ii) Air almost in contact with the road is not steady. The non-uniform motion of the air bends the ray upwards and once it has bent upwards, it continues to do so.

(iii) Using Maxwell's equations for EM waves, correct explanation is possible for the reflection.

It may be pointed out that total internal reflection is NEVER possible here because the relative refractive index is just less than 1 and hence the critical angle (discussed in the article 9.6) is also approaching 90^0.

White ray AB from the Sun strikes from upper portion of a water drop at an incident angle *i*. On entering into water, it deviates and disperses into constituent colours. Extreme colours violet(V) and red(R) are shown. Refracted rays BV and BR strike the opposite inner surface of water drop and suffer internal (NOT total internal) reflection. These reflected rays finally emerge from V′ and R′ and can be seen by an observer on the ground. For the observer they appear to be coming from opposite side of the Sun. Minimum deviation rays of red and violet colour are inclined to the ground level at $\theta_R = 42.8° \cong 43°$ and $\theta_V = 40.8 \cong 41°$ respectively. As a result, in the 'bow' or arch, the red is above or outer and violet is lower or inner.

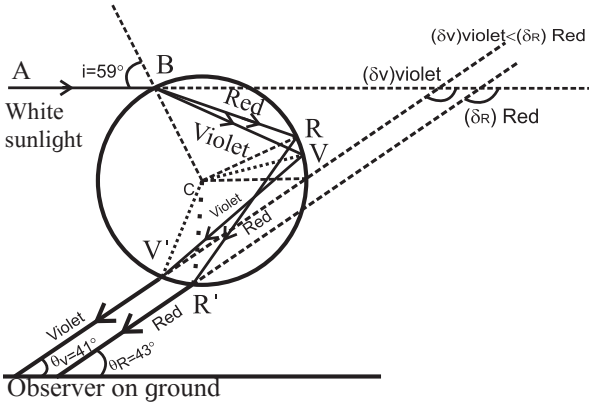


Fig. 9.22 (a): Formation of primary rainbow.

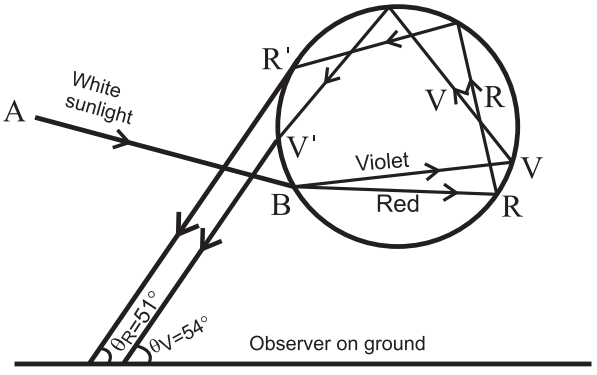


Fig. 9.22 (b): Formation of secondary rainbow.

Secondary rainbow: Figure 9.22 (b) shows some optical phenomena involved in the formation of a secondary rainbow due to a spherical water drop. White ray AB from the Sun strikes from lower portion of a water drop at an incident angle *i*. On entering into water, it deviates and disperses into constituent colours. Extreme colours violet(V) and red(R) are shown. Refracted rays BV and BR finally emerge the drop from V' and R' after suffering two internal reflections and can be seen by an observer on the ground. Minimum deviation rays of red and violet colour are inclined to the ground level at $\theta_R \cong 51°$ and $\theta_V \cong 53°$ respectively. As a result, in the 'bow' or arch, the violet is above or outer and red is lower or inner.

Do you know ?

(I) Why total internal reflection is not possible during formation of a rainbow?

Angle of incidence i in air, at the water drop, can't be greater than 90°. As a result, angle of refraction r in water will *always less than the critical angle*. From Fig a and b and by simple geometry, it is clear that this r itself is the angle of incidence at any point for one or more internal reflections. Obviously, total internal reflection is not possible.

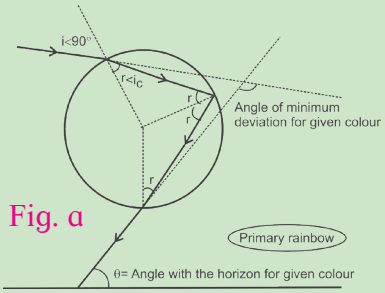


Fig. a

(II) Why is rainbow seen only for a definite angle range with respect to the ground?

For clear visibility we must have a beam of enough intensity. From the deviation curve (Fig 9.19) it is clear that near minimum deviation the curve is almost parallel to x-axis, i.e., for majority of angles of incidence in this range, the angle of deviation is nearly the same and those are almost parallel forming a beam of enough intensity. Thus, the rays in the near vicinity of minimum deviation are almost parallel to each other. Rays beyond this range suffer wide angular dispersion and thus will not have enough intensity for visibility.

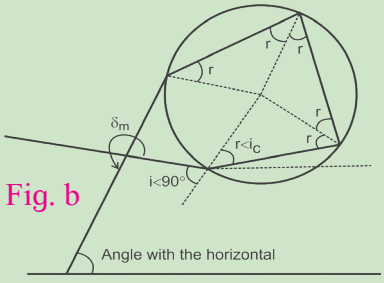
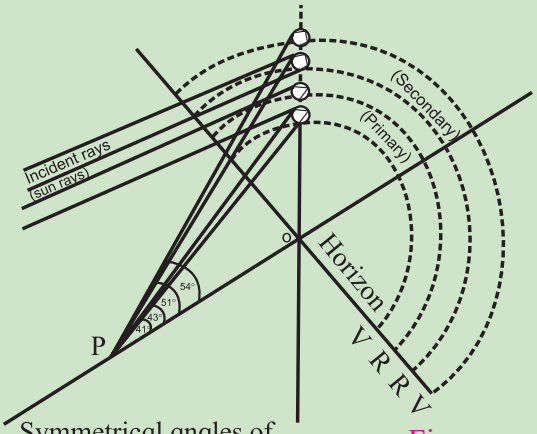


Fig. b

By using simple geometry for Figs. a and b it can be shown that the angle of deviation between final emergent ray and the incident ray is $\delta = \pi + 2i - 4r$ during primary rainbow, and $\delta = 2\pi + 2i - 6r$ during secondary rainbow. Using these relations and Snell's law $\sin i = n \sin r$, we can obtain derivatives of δ. Second derivative of δ comes out to be negative, which shows that it is the minima condition. Equating first derivative to zero we can obtain corresponding values of i and r. Again, by using Figs. a and b, we can obtain the corresponding angles θ_R and θ_V at the horizontal, which is the visible angular position for the rainbow.

(III) Why is the rainbow a bow or an arch? Can we see a complete circular rainbow?

Figure c illustrates formation of primary and secondary rainbows with their common centre O is the point where the line joining the sun and the observer meets the Earth when extended. P is location of the observer. Different colours of rainbows are seen on arches of cones of respective angles described earlier.



Symmetrical angles of respective curves of colours

Fig. c

Smallest half angle refers to the cone of violet colour of primary rainbow, which is 41^0. As the Sun rises, the common centre of the rainbows moves down. Hence as the Sun comes up, smaller and smaller part of the rainbows will be seen. If the Sun is above 41^0, violet arch of primary rainbow cannot be seen. Obviously beyond 53^0, nothing will be seen. That is why rainbows are visible only during mornings and evenings.

However, if observer moves up (may be in an aeroplane), the line PO itself moves up making lower part of the arches visible. After a certain minimum elevation, entire circle for all the cones can be visible.

(IV) Size of water drops convenient for rainbow: Water drops responsible for the formation of a rainbow should not be too small. For too small drops the phenomenon of diffraction (redistribution of energy due to obstacles, discussed in XIIth standard) dominates and clear rainbow can't be seen.

9.10 Defects of lenses (aberrations of optical images):

As mentioned in the section 9.4 for aberration for curved mirrors, while deriving various relations, we assume most of the rays to be paraxial by using lenses of small aperture. In reality, we have objects of finite sizes. Also, we need optical devices of large apertures (lenses and/or mirrors of size few meters for telescopes, etc.). In such cases the beam of rays is no more paraxial, quite often not parallel also. As a result, the spherical oberration discussed for spherical mirrors can occur for lenses also. Only one defect is mentioned corresponding to monochromatic beam of light.

Chromatic aberration: In case of mirrors there is no dispersion of light due to refractive index. However, lenses are prepared by using a transparent material medium having different refractive index for different colours. Hence angular dispersion is present. A convex lens can be approximated to two thin prisms connected base to base and for a concave lens those are vertex to vertex. (Fig. 9.23 (a) and 9.23 (b))

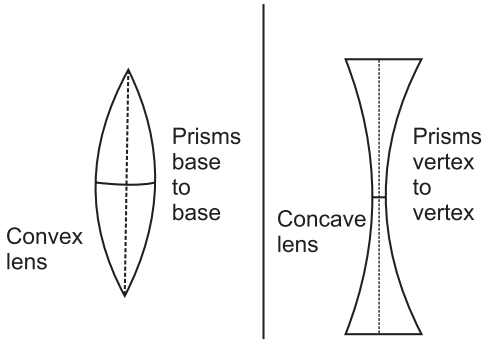


Fig. 9.23: (a) Convex lens (b) Concave lens

If the lens is thick, this will result into notably different foci corresponding to each colour for a polychromatic beam, like a white light. This defect is called chromatic aberration, violet being focused closest to pole as it has maximum deviation. (Fig 9.24 (a) and 9.24 (b)) Longitudinal chromatic aberration, transverse chromatic aberration and circle of least confusion are defined in the same manner as that of spherical aberration for spherical mirrors.

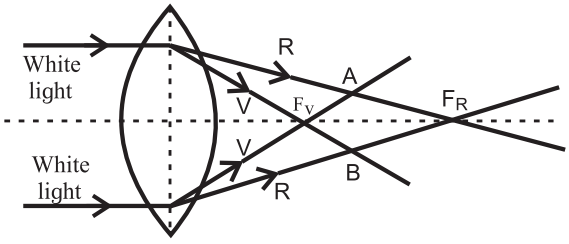


Fig. 9.24: Chromatic aberration: (a) Convex lens.

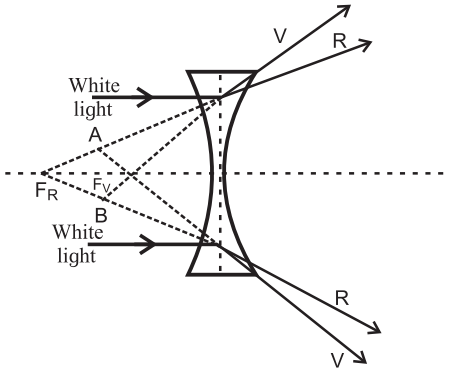


Fig. 9.24: Chromatic aberration: (b) Concave lens

Reducing/eliminating chromatic aberration:

Eliminating chromatic aberration simultaneously for all the colours is impossible. We try to eliminate it for extreme colours which reduces it for other colours. Convenient methods to do it use either a convex and a concave lens in contact or two thin convex lenses with proper separation. Such a combination is called achromatic combination.

Achromatic combination of two lenses in contact: Let ω_1 and ω_2 be the dispersive powers of materials of the two component lenses used in contact for an achromatic combination. Their focal lengths f for violet, red and yellow (assumed to be the mean colour) are suffixed by respective letters V, R and Y.

Also, let $K_1 = \left(\frac{1}{R_1} - \frac{1}{R_2}\right)_1$ for lens 1 and

$K_2 = \left(\frac{1}{R_1} - \frac{1}{R_2}\right)_2$ for lens 2.

For two thin lenses in contact,

$$\frac{1}{f} = \frac{1}{f_1} + \frac{1}{f_2} \quad \ldots\ldots$$

To be used separately for respective colours.

For the combination to be achromatic, the

resultant focal length of the combination must be the same for both the colours, i.e.,

$$f_V = f_R \text{ or } \frac{1}{f_V} = \frac{1}{f_R}$$

$$\therefore \frac{1}{f_{1V}} + \frac{1}{f_{2V}} = \frac{1}{f_{1R}} + \frac{1}{f_{2R}}$$

$(n_{1V}\text{-}1)\,K_1 + (n_{2V}\text{-}1)\,K_2 = (n_{1R}\text{-}1)\,K_1 + (n_{2R}\text{-}1)\,K_2$

…… using lens makers' Eq. (9.7)

$$\therefore \frac{K_1}{K_2} = \frac{n_{2V} - n_{2R}}{n_{1V} - n_{1R}} \qquad \text{--- (9.17)}$$

For mean colour yellow,

$$\frac{1}{f_Y} = \frac{1}{f_{1Y}} + \frac{1}{f_{2Y}}$$

$$\text{with } \frac{1}{f_{1Y}} = (n_{1Y} - 1)K_1$$

$$\text{and } \frac{1}{f_{2Y}} = (n_{2Y} - 1)K_2$$

$$\therefore \frac{K_1}{K_2} = \left(\frac{n_{2Y} - 1}{n_{1Y} - 1}\right)\left(\frac{f_{2Y}}{f_{1Y}}\right) \qquad \text{--- (9.18)}$$

Equating R.H.S. of (9.17) and (9.18) and rearranging, we can write

$$\frac{f_{2Y}}{f_{1Y}} = -\left(\frac{n_{2V} - n_{2R}}{n_{2Y} - 1}\right) \div \left(\frac{n_{1R} - n_{1R}}{n_{1Y} - 1}\right)$$

$$= -\frac{\omega_2}{\omega_1} \qquad \text{--- (9.19)}$$

Equation (9.19) is the condition for achromatic combination of two lenses, in contact.

Dispersive power ω is always positive. Thus, one of the lenses must be convex and the other concave.

If second lens is concave, f_{2Y} is negative.

$$\therefore \frac{1}{f_Y} = \frac{1}{f_{1Y}} - \frac{1}{f_{2Y}}$$

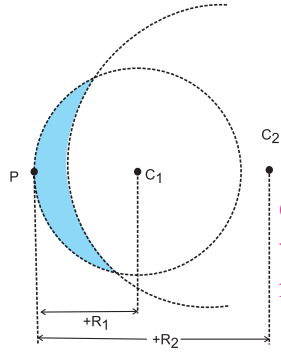
For this combination to be converging, f_Y should be positive.

Hence, $f_{1Y} < f_{2Y}$ **and** $\omega_1 < \omega_2$

Thus, for an achromatic combination if there is a choice between flint glass ($n = 1.655$) and crown glass ($n = 1.517$), the convergent (convex) lens must be of crown glass and the divergent (concave) lens of flint glass.

Example 9.11: After Cataract operation, a person is recommended with concavo-convex spectacles of curvatures 10 cm and 50 cm. Crown glass of refractive indices 1.51 for red and 1.53 for violet colours is used for this. Calculate the lateral chromatic aberration occurring due to these glasses.

Solution: For a concavo-concave lens, both the radii of curvature are either positive or both negative. If convex shape faces object, both will be positive. See the accompanying figure.



Concavo-convex lens with convex face receiving incident rays

$\therefore R_1 = 10\,\text{cm}$ and $R_2 = 50\,\text{cm}$

$$\therefore \left(\frac{1}{R_1} - \frac{1}{R_2}\right) = \left(\frac{1}{+10} - \frac{1}{+50}\right) = 0.08 \text{ cm}^{-1}$$

$$\therefore \frac{1}{f_R} = (n_R - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$= (1.51 - 1) \times 0.08 = 0.0408$$

$$\therefore f_R = 25.51\,\text{cm}$$

$$\text{and } \frac{1}{f_V} = (n_V - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$= (1.53 - 1) \times 0.08 = 0.0424$$

$$\therefore f_V = 23.58\,\text{cm}$$

$\therefore$ Longitudinal chromatic aberration

$= f_V - f_R = 25.51 - 23.58$

= 1.93 cm,... (quite appreciable!)

Verify that you get the same answer even if you consider the concave surface facing the incident rays.

Spherical aberration: Longitudinal spherical aberration, transverse spherical aberration and circle of least confusion are defined in the same manner as that for spherical mirrors. (Fig 9.25 (a) and 9.25 (b))

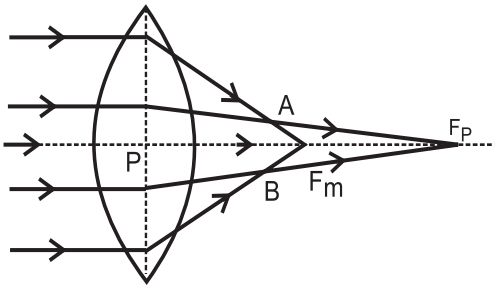


Fig. 9.25 (a): Spherical aberration, Convex lens.

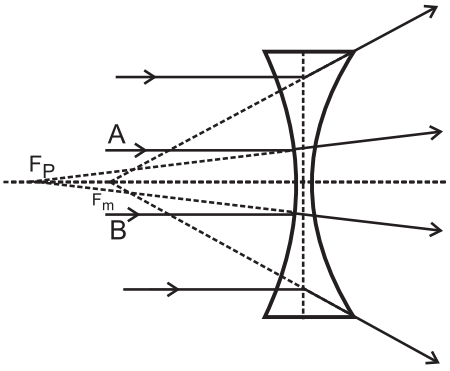


Fig. 9.25 (b): Spherical aberration, Concave lens

Methods to reduce/eliminate spherical aberration of lenses:

(i) Cheapest method to reduce the spherical aberration is to use a planoconvex or planoconcave lens with curved side facing the incident rays (real object). Reversing it increases the aberration appreciably.

(ii) Certain ratio of radii of curvature for a given refractive index almost eliminates the spherical aberration. For $n = 1.5$, the ratio is $\frac{R_1}{R_2} = \frac{1}{6}$ and for $n = 2$, it is $\frac{1}{5}$

(iii) Use of two thin converging lenses separated by distance equal to difference between their focal lengths with lens of larger focal length facing the incident rays considerably reduces spherical aberration.

(iv) Spherical aberration of a convex lens is positive (for real image), while that of a concave lens is negative. Thus, a suitable combination of them (preferably a double convex lens of smaller focal length and a planoconcave lens of greater focal length) can completely eliminate spherical aberration.

9.11 Optical instruments:

Introduction: Whether an object appears bigger or not does not necessarily depend upon its own size. Huge mountains far off may appear smaller than a small tree close to us. This is because the angle subtended by the mountain at the eye from that distance (called the visual angle) is smaller than that subtended by the tree from its position. Hence, apparent size of an object depends upon the visual angle subtended by the object from its position. Obviously, for an object to appear bigger, we must bring it closer to us or we should go closer to it.

However, due to the limitation for focusing the eye lens it is not possible to take an object closer than a certain distance. This distance is called least distance of distance vision D. For a normal, unaided human eye $D = 25$cm. If an object is brought closer than this, we cannot see it clearly. If an object is too small (like the legs of an ant), the corresponding visual angle from 25 cm is not enough to see it and if we bring it closer than that, its image on the retina is blurred. Also, the visual angle made by cosmic objects far away from us (such as stars) is too small to make out minor details and we cannot bring those closer. In such cases we need optical instruments such as a microscope in the former case and a telescope in the latter. It means that microscopes and telescopes help us in increasing the visual angle. This is called angular magnification or magnifying power.

Magnifying power: Angular magnification or magnifying power of an optical instrument is defined as the ratio of the visual angle made by the image formed by that optical instrument (β) to the visual angle subtended by the object when kept at the least distance of distinct vision (α). (Figure 9.26 (a) and 9.26 (b)) In the case of telescopes, α is the angle subtended by the object from its own position as it is not possible to get it closer.

Simple microscope or a reading glass: In order to read very small letters in a newspaper, sometimes we use a convex lens. You might have seen watch-makers using a special type of small convex lens while looking at very tiny

parts of a wrist watch. Convex lens used for this purpose is a simple microscope.

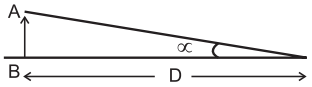


Fig. 9.26: (a) Visual Angle α.

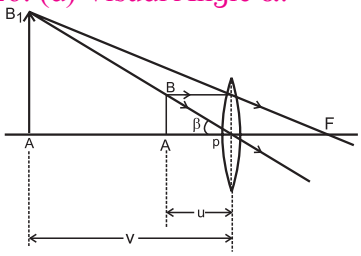


Fig. 9.26: (b) Visual Angle β.

Figure 9.26 (a) shows visual angle α made by an object, when kept at the least distance of distinct vision D. Without an optical instrument this is the greatest possible visual angle as we cannot get the object closer than this. Figure 9.26 (b) shows a convex lens forming erect, virtual and magnified image of the same object, when placed within the focus. The visual angle β of the object and the image in this case are the same. However, this time the viewer is looking at the image which is not closer than D. Hence the same object is now at a distance smaller than D. It makes β greater then α and the same object appears bigger.

Angular magnification or magnifying power, in this case, is given by

$$M = \frac{\text{Visual angle of the image}}{\text{Visual angle of the object, when at } D} = \frac{\beta}{\alpha}$$

For small angles α and β, we can write,

$$M = \frac{\beta}{\alpha} \cong \frac{\tan(\beta)}{\tan(\alpha)} = \frac{\left(\frac{BA}{PA}\right)}{\left(\frac{BA}{D}\right)} = \frac{D}{u}$$ - (numerically)

Limiting cases:

(i) For maximum magnifying power, the image should be nearest possible, i.e., at D.

For a thin lens, $\frac{1}{f} = \frac{1}{v} - \frac{1}{u}$

In this case, $f = +f, v = v_{min} = -D, u = -u$ and $M = M_{max}$

$$\therefore \frac{1}{f} = \frac{1}{-D} - \frac{1}{-u} \quad \therefore \frac{D}{f} = \frac{D}{-D} + \frac{D}{u}$$

$$\therefore M_{max} = \frac{D}{u} = 1 + \frac{D}{f}$$

(ii) For minimum magnifying power, $v = \infty$, i.e., $u = f$ (numerically)

$$\therefore M_{min} = \frac{D}{u} = \frac{D}{f}$$

Thus the angular magnification by a lens of focal length f is between $\left(\frac{D}{f}\right)$ and $\left(1 + \frac{D}{f}\right)$ only.

For common human eyesight, D = 25 cm. Thus, if f = 5 cm,

$$M_{min} = \left(\frac{D}{f}\right) = 5 \text{ and } M_{max} = \left(1 + \frac{D}{f}\right) = 6.$$

Hence image appears to be only 5 to 6 times bigger for a lens of focal length 5 cm.

For $M_{min} = \left(\frac{D}{f}\right) = 5$, $v = \infty$. $\therefore \quad m = \frac{v}{u} = \infty$.

Thus, the image size is infinite times that of the object, but appears only 5 times bigger.

For

$$M_{max} = 1 + \left(\frac{D}{f}\right) = 6,$$

$v = -25$ cm. Corresponding $u = \frac{-25}{6}$ cm

$\therefore \quad m = \frac{v}{u} = 6$. Thus, image size is 6 times that of the object, and appears also 6 times larger.

Example 9.12: A magnifying glass of focal length 10 cm is used to read letters of thickness 0.5 mm held 8 cm away from the lens. Calculate the image size. How big will the letters appear? Can you read the letters if held 5 cm away from the lens? If yes, of what size would the letters appear? If no, why not?

$f = +10$ cm, $u = -8$ cm, $v = ?$

$$\frac{1}{f} = \frac{1}{v} - \frac{1}{u} \quad \therefore \frac{1}{10} = \frac{1}{v} - \frac{1}{-8} \quad \therefore v = -40 \text{ cm}$$

$$m = \frac{v}{u} = \frac{\text{Image size } h_i}{\text{Object size } h_o} \therefore \frac{40}{8} = \frac{h_i}{0.5}$$

$\therefore h_i = 2.5$ cm (5 times that of the object)

$$M = \frac{D}{u} = \frac{25}{8} = 3.125$$

$\therefore$ Image will appear to be 3.125 times bigger. i.e., $3.125 \times 0.5 = 1.5625$ cm

For $\mu = -5$ cm, v will be - 10 cm.

For an average human being to see clearly, the image must be at or beyond 25 cm. Thus, it will not possible to read the letters if held 5 cm away from the lens.

Compound microscope: As seen above, the magnifying power of a simple microscope is inversely proportional to its focal length. However, if we need focal length to be smaller and smaller, the corresponding lens becomes thicker and thicker. For such a lens both spherical as well as chromatic aberrations are dominant. Thus, if higher magnifying power is needed, we go for using more than one lenses. The instrument is then called a compound microscope. It is used to view very small objects (sizes $\sim 10^{-1}$ mm to 10^{-3} mm). Also, whether the image is erect or inverted is immaterial.

A compound microscope essentially uses two convex lenses of suitable focal lengths fit into a cylindrical tube with some adjustment possible for its length. The smaller lens (~ 4 mm to 6 mm aperture) facing the object is called the objective. Other lens with which the observer jams her/his eye is litter larger and called as the eye lens. (Fig 9.27) During this discussion we consider the eye lens to be a single lens, but in practice it is an eyepiece, itself consisting of two planoconvex lenses.

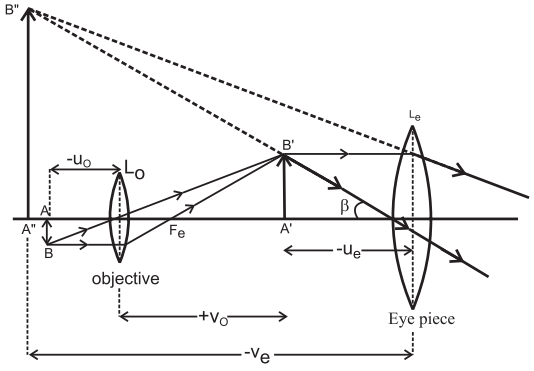


Fig. 9.27: Compound Microscope.

As shown in the Fig. 9.27, a tiny object AB is placed between f and $2f$ of the objective which produces its real, inverted and magnified image A′ B′ in front of the eye lens. Position of the eye lens is so adjusted that the (intermediate) image A′ B′ is within its focus. Hence, for this object A′ B′, the eye lens behaves as a simple microscope and produces its virtual and magnified image A″ B″, which is inverted with respect to original object AB.

Magnifying power of a compound microscope with two lenses: From its position, the final image A″ B″ makes a visual angle β at the eye (jammed at the eye lens). Visual angle made by the object from distance D is α.

$$\therefore \tan\beta = \frac{A''B''}{v_e} = \frac{A'B'}{u_e}$$

$$\tan\alpha = \frac{AB}{D} \quad \text{(Fig. 9.29 (a))}$$

$\therefore$ Angular magnification or magnifying power,

$$M = \frac{\beta}{\alpha} \cong \frac{\tan\beta}{\tan\alpha} = \left(\frac{A'B'}{u_e}\right) \times \left(\frac{D}{AB}\right)$$

$$= \left(\frac{A'B'}{AB}\right) \times \left(\frac{D}{u_e}\right)$$

$$\therefore M = m_o \times M_e$$

Where, $\left(\frac{A'B'}{AB}\right) = m_o = \frac{v_o}{u_o}$ is the linear (lateral) magnification of the objective and

$\left(\frac{D}{u_e}\right) = M_e$ is the angular magnification or magnifying power of the eye lens. Length of the compound microscope then becomes L = distance between the two lenses $v_0 + u_e$.

Remarks:

(i) In order to increase m_o, we need to decrease u_o. Thereby, the object comes closer and closer to the focus of the objective. This increases v_0 and hence length of the microscope. Thus m_o can be increased only within the limitation of length of the microscope.

(ii) Minimum value of M_e is $\left(\frac{D}{f_e}\right)$ for final image at infinity and maximum value of M_e is $\left(1+\frac{D}{f_e}\right)$ for final image at D respectively. M_e and m_o together decide the minimum and maximum magnifying power of the microscope.

Example 9.13: The pocket microscope used by a student consists of eye lens of focal length 6.25 cm and objective of focal length 2 cm. At microscope length 15 cm, the final image appears biggest. Estimate distance of the object from the objective and magnifying power of the microscope.

Solution:

$f_e = 6.25\,\text{cm},\ f_o = 2\,\text{cm},\ L = |v_o| + |u_e|$

$= 15\,\text{cm},\ v_e = 25\,\text{cm}\ (\text{Image appears largest})$

$$\frac{1}{f_e} = \frac{1}{v_e} - \frac{1}{u_e} \quad \therefore \frac{1}{6.25} = \frac{4}{25}$$

$$= \frac{1}{-25} - \frac{1}{u_e} \quad \therefore |u_e| = 5\,\text{cm}$$

$$\therefore |v_o| = L - |u_e| = 15 - 5 = 10\,\text{cm}$$

$$\frac{1}{f_o} = \frac{1}{v_o} - \frac{1}{u_o} \quad \therefore \frac{1}{2} = \frac{1}{+10} - \frac{1}{u_o} \quad \therefore |u_o| = 2.5\,\text{cm}$$

$$M = m_o \times M_e = \left(\frac{v_0}{u_o}\right)\left(\frac{D}{u_e}\right)$$

$$= \left(\frac{10}{2.5}\right)\left(\frac{25}{5}\right) = 4 \times 5 = 20$$

Telescope: Telescopes are used to see terrestrial or astronomical bodies. A telescope essentially uses two lenses (or one large parabolic mirror and a lens). The lens facing the object (called objective) is of aperture as large as possible. For Newtonian telescopes, a large parabolic mirror faces the object.

For terrestrial telescopes the objects to be seen are on the Earth , like mountains, trees, players playing a match in a stadium, etc. In such case, the final image must be erect. Eye lens used for this purpose must be concave and such a telescope is popularly called a binocular. A variety of binoculars use three convex lenses with proper separation. The third lens again inverts the second intermediate image and makes final image erect with respect to the object. In this text we shall be discussing astronomical telescope.

For an astronomical telescope, the objects to be seen are planets, stars, galaxies, etc. In this case there is no necessity of erect image. Such telescopes use convex lens as eye lens. (Fig. 9.27).

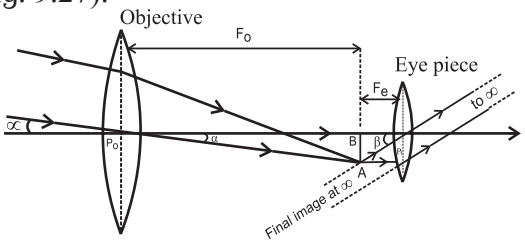


Fig. 9.28: Telescope.

Magnifying power of a telescope: Objects to be seen through a telescope cannot be brought to distance D from the objective, like in microscopes. Hence, for telescopes, α is the visual angle of the object from its own position, which is practically at infinity. Visual angle of the final image is β and its position can be adjusted to be at D. However, under normal adjustments, the final image is also at infinity but making a greater visual angle than that of the object. (If the image is really at infinity, there will not be any parallax at the cross wires). Beam of incident rays is now inclined at an angle α with the principal axis while emergent beam is inclined at a greater angle β with the principal axis causing angular magnification. (Fig. 9.28)

Objective of focal length f_o focusses the parallel incident beam at a distance f_o from the objective giving an inverted image AB. For normal adjustment, the eye lens is so adjusted that the intermediate image AB happens to be at the focus of the eye lens. Rays refracted beyond the eye lens form a parallel beam inclined at an angle β with the principal axis resulting into the image also at infinity.

$\therefore$ Angular magnification or magnifying power,

$$M = \frac{\beta}{\alpha} \cong \frac{\tan\beta}{\tan\alpha} = \frac{\left(\frac{BA}{P_eB}\right)}{\left(\frac{BA}{P_oB}\right)} = \frac{\left(\frac{BA}{f_e}\right)}{\left(\frac{BA}{f_o}\right)}$$

$$\therefore M = \frac{f_o}{f_e}$$

Length of the telescope for normal adjustment is $L = f_o + f_e$

Under the allowed limit of length objective of

maximum possible focal length f_o and eye lens of minimum possible focal length f_e can be chosen for maximum magnifying power.

Example 14: Focal length of the objective of an astronomical telescope is 1 m. Under normal adjustment, length of the telescope is 1.05 m. Calculate focal length of the eyepiece and magnifying power under normal adjustment.

Solution: For astronomical telescope,

$L = f_0 + f_e \therefore 1.05 = 1 + f_0 \therefore f_e = 0.05$ m = 5 cm

Under normal adjustments,

$$M = \frac{f_o}{f_e} = \frac{1}{0.05} = 20$$

Exercises

1. Choose the correct option

i. As per recent understanding *light* consists of

(A) rays

(B) waves

(C) corpuscles

(D) photons obeying the rules of waves

ii. Consider optically denser lenses P, Q, R and S drawn below. According to Cartesian sign convention which of these have positive focal length?

(A) Only P

(B) Only P and Q

(C) Only P and R

(D) Only Q and S

iii. Two plane mirrors are inclined at angle 40^0 between them. Number of images seen of a tiny object kept between them is

(A) Only 8 (B) Only 9

(C) 8 or 9 (D) 9 or 10

iv. A concave mirror of curvature 40 cm, used for *shaving purpose* produces image of double size as that of the object. Object distance must be

(A) 10 cm only

(B) 20 cm only

(C) 30 cm only

(D) 10 cm or 30 cm

v. Which of the following aberrations will NOT occur for spherical mirrors?

(A) Chromatic aberration

(B) Coma

(C) Distortion

(D) Spherical aberration

vi. There are different fish, monkeys and water on the habitable planet of the star Proxima b. A fish swimming underwater feels that there is a monkey at 2.5 m on the top of a tree. The same monkey feels that the fish is 1.6 m below the water surface. Interestingly, height of the tree and the depth at which the fish is swimming are exactly same. Refractive index of that water must be

(A) 6/5 (B) 5/4

(C) 4/3 (D) 7/5

vii. Consider following phenomena/ applications: P) Mirage, Q) rainbow, R) Optical fibre and S) glittering of a diamond. Total internal reflection is involved in

(A) Only R and S (B) Only R

(C) Only P, R and S (D) all the four

viii. A student uses spectacles of number -2 for seeing distant objects. Commonly used lenses for her/his spectacles are

(A) bi-concave

(B) double concave

(C) concavo-convex

(D) convexo-concave

ix. A spherical marble of refractive index 1.5 and curvature 1.5 cm, contains a tiny air bubble at its centre. Where will it appear when seen from outside?

(A) 1 cm inside (B) at the centre

(C) 5/3 cm inside (D) 2 cm inside

x. Select the WRONG statement.

(A) Smaller angle of prism is recommended for greater angular dispersion.

(B) Right angled isosceles glass prism is commonly used for total internal reflection.

(C) Angle of deviation is practically constant for thin prisms.

(D) For emergent ray to be possible from the second refracting surface, certain minimum angle of incidence is necessary from the first surface.

xi. Angles of deviation for extreme colours are given for different prisms. Select the one having maximum dispersive power of its material.

(A) 7°, 10° (B) 8°, 11°

(C) 12°, 16° (D) 10°, 14°

xii. Which of the following is not involved in formation of a rainbow?

(A) refraction

(B) angular dispersion

(C) angular deviation

(D) total internal reflection

xiii. Consider following statements regarding a simple microscope:

(P) It allows us to keep the object within the least distance of distant vision.

(Q) Image appears to be biggest if the object is at the focus.

(R) It is simply a convex lens.

(A) Only (P) is correct

(B) Only (P) and (Q) are correct

(C) Only (Q) and (R) are correct

(D) Only (P) and (R) are correct

2. Answer the following questions.

i) As per recent development, what is the nature of light? Wave optics and particle nature of light are used to explain which phenomena of light, respectively?

ii) Which phenomena can be satisfactorily explained using ray optics? State the assumptions on which ray optics is based.

iii) What is focal power of a spherical mirror or of a lens? What may be the reason for using $P = \frac{1}{f}$ as its expression?

iv) At which positions of the objects do spherical mirrors produce (i) diminished image, (ii) magnified image?

v) State the restrictions for having images produced by spherical mirrors to be appreciably clear.

vi) Explain spherical aberration for spherical mirrors. How can it be minimized? Can it be eliminated by some curved mirrors?

vii) Define absolute refractive index and relative refractive index. Explain in brief, with an illustration for each.

viii) Explain 'mirage' as an illustration of refraction.

ix) Under what conditions is total internal reflection possible? Explain it with a suitable example. Define critical angle of incidence and obtain an expression for it.

x) Describe construction and working of an optical fibre. What are the advantages of optical fibre communication over electronic communication?

xi) Why is a prism binoculars preferred over traditional binoculars? Describe its working in brief.

xii) A spherical surface separates two transparent media. Derive an expression that relates object and image distances with the radius of curvature for a point object. Clearly state the assumptions, if any.

xiii) Derive lens makers' equation. Why is it called so? Under which conditions focal length f and radii of curvature R are numerically equal for a lens?

2. Answer the following questions in detail.

i) What are different types of dispersions of light? Why do they occur?

ii) Define angular dispersion for a prism. Obtain its expression for a thin prism. Relate it with the refractive indices of the material of the prism for corresponding colours.

iii) Explain and define dispersive power of a transparent material. Obtain its expressions in terms of angles of deviation and refractive indices.

iv) (i) State the conditions under which a rainbow can be seen.

(ii) Explain the formation of a primary rainbow. For which angular range with the horizontal is it visible?

(iii) Explain the formation of a secondary rainbow. For which angular range with the horizontal is it visible?

(iv) Is it possible to see primary and secondary rainbow simultaneously? Under what conditions?

v) (i) Explain chromatic aberration for spherical lenses. State a method to minimize or eliminate it.

(ii) What is achromatism? Derive a condition to achieve achromatism for a lens combination. State the conditions for it to be converging.

vi) Describe spherical aberration for spherical lenses. What are different ways to minimize or eliminate it?

vii) Define and describe magnifying power of an optical instrument. How does it differ from linear or lateral magnification?

viii) Derive an expression for magnifying power of a simple microscope. Obtain its minimum and maximum values in terms of its focal length.

ix) Derive the expressions for the magnifying power and the length of a compound microscope using two convex lenses.

x) What is a terrestrial telescope and an astronomical telescope?

xi) Obtain the expressions for magnifying power and the length of an astronomical telescope under normal adjustments.

xii) What are the limitations in increasing the magnifying powers of (i) simple microscope (ii) compound microscope (iii) astronomical telescope?

3. Solve the following numerical examples

i) A monochromatic ray of light strike the water (n = 4/3) surface in a cylindrical vessel at angle of incidence 53^0. Depth of water is 36 cm. After striking the water surface, how long will the light take to reach the bottom of the vessel? [Angles of the most popular Pythagorean triangle of sides in the ratio 3:4:5 are nearly 37^0, 53^0 and 90^0]

[Ans: 2 ns]

ii) Estimate the number of images produced if a tiny object is kept in between two plane mirrors inclined at 35^0, 36^0, 40^0 and 45^0.

[Ans: 10, 9, 9 or 8, 7 respectively]

iii) A rectangular sheet of length 30 cm and breadth 3 cm is kept on the principal axis of a concave mirror of focal length 30 cm. Draw the image formed by the mirror on the same ray diagram, as far as possible on scale.

[Ans: Inverted image starts from 50 cm and ends at 90 cm. Its height in the beginning is 2 cm and at the end it is 6 cm. At 60 cm, image height is 3 cm. Thus, outer boundary if the image is a curve]

iv) A car uses a convex mirror of curvature 1.2 m as its rear-view mirror. A minibus of cross section 2.4 m × 2.4 m is 6.6 m away from the mirror. Estimate the image size.

[Ans: A square of edge 0.2 m]

v) A glass slab of thickness 2.5 cm having refractive index 5/3 is kept on an ink spot. A transparent beaker of very thin bottom, containing water of refractive index 4/3 up to 8 cm, is kept on the glass block. Calculate apparent depth of the ink spot when seen from the outside air.

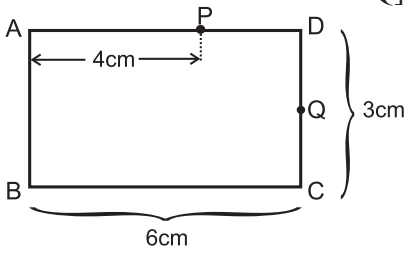
[Ans: 7.5 cm]

vi) A convex lens held some distance above a 6 cm long pencil produces its image of SOME size. On shifting the lens by a distance equal to its focal length, it again produces the image of the SAME size as earlier. Determine the image size.

[Ans: 12 cm]

vii) Figure below shows the section ABCD of a transparent slab. There is a tiny green LED light source at the bottom left corner B. A certain ray of light from B suffers total internal reflection at nearest point P on the surface AD and strikes the surface CD at point Q. Determine refractive index of the material of the slab and distance DQ. At Q, the ray PQ will suffer partial or total internal reflection? [You may use the approximation given in Q 1 above].

[Ans: n = 5/4, DQ = 1.5 cm, Partial internal reflection at Q]



viii) A point object is kept 10 cm away from one of the surfaces of a thick double convex lens of refractive index 1.5 and radii of curvature 10 cm and 8 cm. Central thickness of the lens is 2 cm. Determine location of the final image considering paraxial rays only.

Hint : Single spherical surface formula to be used twice.

[Ans: 64 cm away from the other surface]

ix) A monochromatic ray of light is incident at 37^0 on an equilateral prism of refractive index 3/2. Determine angle of emergence and angle of deviation. If angle of prism is adjustable, what should its value be for emergent ray to be just possible for the same angle of incidence.

[Ans: $e = 63°$, $\delta = 40°$, $A = 65°\ 24'$ for $e = 90°$ (just emerges)]

x) From the given data set, determine angular dispersion by the prism and dispersive power of its material for extreme colours. n_R = 1.62 n_V = 1.66, $\delta_R = 3.1°$

[Ans: $\delta_{VR} = 0.2°$, $\omega_{VR} = \frac{1}{16} = 0.0625$]

xi) Refractive index of a flint glass varies from 1.60 to 1.66 for visible range. Radii of curvature of a thin convex lens are 10 cm and 15 cm. Calculate the chromatic aberration between extreme colours.

[Ans: 10/11 cm]

xii) A person uses spectacles of 'number' 2.00 for reading. Determine the range of magnifying power (angular magnification) possible. It is a concavo-convex lens (n = 1.5) having curvature of one of its surfaces to be 10 cm. Estimate that of the other.

[Ans: $M_{min} = 0.5$, $M_{max} = 1.5$ $R_2 = 50/3$ cm]

xiii) Focal power of the eye lens of a compound microscope is 6 dioptre. The microscope is to be used for maximum magnifying power (angular magnification) of at least 12.5. The packing instructions demand that length of the microscope should be 25 cm. Determine minimum focal power of the objective. How much will its radius of curvature be if it is a biconvex lens of n = 1.5.

[Ans: 40 dioptre, 2.5 cm]

***

10. Electrostatics

Can you recall?

1. Have you experienced a shock while getting up from a plastic chair and shaking hand with your friend?
2. Ever heard a crackling sound while taking out your sweater in winter?
3. Have you seen the lightning striking during pre-monsoon weather?

10.1 Introduction:

Electrostatics deals with static electric charges, the forces between them and the effects produced in the form of electric fields and electric potentials. We have already studied some aspects of electrostatics in earlier standards. In this Chapter we will review some of them and then go on to study some aspects in details.

Current electricity, which plays a major role in our day to day life, is produced by moving charges. Charges are present everywhere around us though their presence can only be felt under special circumstances. For example, when we remove our sweater in winter on a dry day, we hear some crackling sound and the sweater appears to stick to our body. This is because of the electric charges produced due to friction between our body and the sweater. Similarly, the lightening that we see in the sky is also due to the flow of large amount of electric charges that develop on the clouds due to friction.

10.2 Electric Charges:

Historically, opposite electric charges were known to the Greeks in the 600 BC. They realized that equal and opposite charges develop on amber and fur when rubbed against each other. Now we know that electric charge is a basic property of elementary particles of which matter is made of. These elementary particles are proton, neutron, and electron. Atoms are made of these particles and matter is made from atoms. A proton is considered to be positively charged and electron to be negatively charged. Neutron is electrically neutral, i.e., it has no charge. An atomic nucleus is made up of protons and neutrons and hence is positively charged. Negatively charged electrons surround the nucleus so as to make an atom electrically neutral. Thus, most matter around us is electrically neutral.

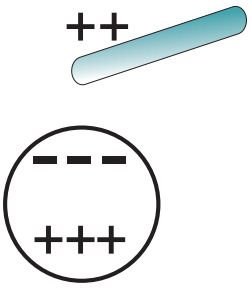


Fig. 10.1 a

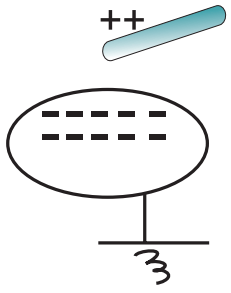


Fig. 10.1 b

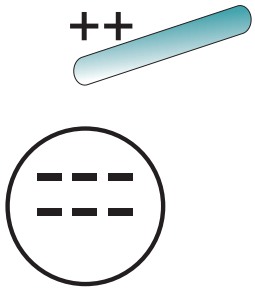


Fig. 10.1 c

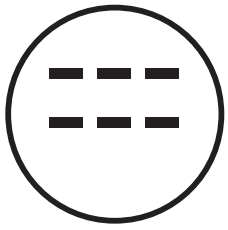


Fig. 10.1 d

Fig. 10.1 (a): Insulated conductor

Fig. 10.1 (b): +ve charge is neutralized by electron from Earth

Fig. 10.1 (c): Earthing is removed -ve charge still stays on the conductor due to +ve charged rod

Fig. 10.1 (d): Rod removed -ve charge is distributed over the surface of the conductor

When certain dissimilar substances, like fur and amber or comb and dry hair are rubbed against each other, electrons get transferred to the other substance making them charged. The substance receiving electrons develops a negative charge while the other is left with an equal amount of positive charge. This can be called charging by conduction as charges are transfered from one body to another. Charges can be separated by other means as well, like

chemical reactions (in cells), convection (in clouds), diffusion (in living cells) etc.

If an uncharged conductor is brought near a charged body, (not in physical contact) the nearer side of the conductor develops opposite charge to that on the charged body and the far side of the conductor develops charge similar to that on the charged body. This is called induction. This happens because the electrons in a conductor are free and can move easily in presence of a charged body. This can be seen from Fig. 10.1.

A charged body attracts or repels electrons in a conductor depending on whether the charge on the body is positive or negative respectively. Positive and negative charges are redistributed and are accumulated at the ends of the conductor near and away from the changed body. From the above discussion it can be inferred that there are only two types of charges found in nature, namely, positive and negative charges. In induction, there is no transfer of charges between the charged body and the conductor. So when the charged body is moved away from the conductor, the charges in the conductor are free again.

Can you tell?

1. When a petrol or a diesel tanker is emptied in a tank, it is grounded.
2. A thick chain hangs from a petrol or a diesel tanker and it is in contact with ground when the tanker is moving.

10.3 Basic Properties of Electric Charge:

10.3.1 Additive Nature of Charge:

Electric charge is additive, similar to mass. The total electric charge on an object is equal to the algebraic sum of all the electric charges distributed on different parts of the object.

It may be pointed out that while taking the algebraic sum, the sign (positive or negative) of the electric charges must be taken into account. Thus if two bodies have equal and opposite charges, the net charge on the system of the two bodies is zero. This is similar to that in case of atoms where the nucleus is positively charged and this charge is equal to the negative charge

Gold Leaf Electroscope:

This is a classic instrument for detecting presences of electric charge. A metal disc is connected to one end of a narrow metal rod and a thin piece of gold leaf is fixed to the other end. The whole of this part of the electroscope is insulted from the body of the instrument. A glass front prevents air draughts but allows to observe the effect of charge on the leaf.

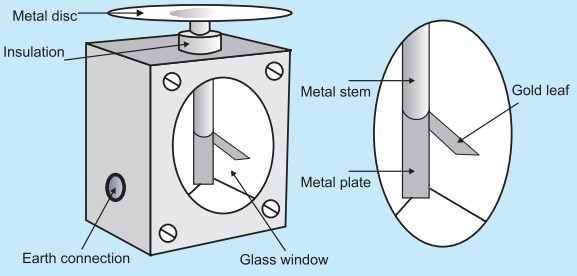
When a charge is put on the disc at the top it spreads down to the plate and leaf moves away from the plate. This happens because similar charges repel. The more the charge on the disc, more is the separation of the leaf from the plate.

The leaf can be made to fall again by touching the disc. This is done by earthing the electroscope. An earth terminal prevents the case from accumulating any stray charge. The electroscope can be charged in two ways.

(a) by contact- a charged rod is brought in contract with the disc and charge is transferred to the electroscope. This method gives the gold leaf the same charge as that on the conductor. This is not a very effective method of charging the electroscope.

(b) by induction- a charged rod is brought close to the disc (not touching it) and the electroscope is earthed. The rod is then removed. This method give the gold leaf opposite charges.

The following diagrams show how the charges spread to the gold leaf and lift it.



of the electrons making the atoms electrically neutral .

It is interesting to compare the additive property of charge with that of mass.

1) The masses of the particles constituting an object are always positive, whereas the charges distributed on different parts of the object may be positive or negative.
2) The total mass of an object is always positive whereas, the total charge on the object may be positive, zero or negative.

10.3.2 Quantization of Charge:

The minimum value of the charge on an electron as determined by the Milikan's oil drop experiment is $e = 1.6\times10^{-19}$ C. This is called the elementary charge. Here, C stands for coulomb which is the unit of charge in SI system. Unit of charge is defind in article 10.4.3. Since protons (+ve) and electrons (-ve) are the charged particles constituting matter, the charge on an object must be an integral multiple of $\pm e$. $q = \pm ne$, where n is an integer.

Further, charge on an object can be increased or decreased in multiples of e. It is because, during the charging process an integral number of electrons can be transferred from one body to the other body. This is known as **quantization of charge** or discrete nature of charge.

The discrete nature of electric charge is usually not observable in practice. It is because the magnitude of the elementary electric charge, e, is extremely small. Due to this, the number of elementary charges involved in charging an object becomes extremely large. Suppose, for example, when a glass rod is rubbed with silk, a charge of the order of one μC (10^{-6} C) appears on the glass rod or silk. Since elementary charge $e = 1.6\times10^{-19}$ C, the number of elementary charges on the glass rod (or silk) is given by

$$n = \frac{10^{-6}\ \text{C}}{1.6\times10^{-19}\ \text{C}} = 6.25\times10^{12}$$

Since it is a tremendously large number, the quantization of charge is not observed and one usually observes a continuous variation of charge.

Example 10.1: How much positive and negative charge is present in 1gm of water? How many electrons are present in it? Given, molecular mass of water is 18.0 g.

Solution: Molecular mass of water is 18.0 gm, that means the number of molecules in 18.0 gm of water is 6.02×10^{23}.

$\therefore$ Number of molecules in 1gm of water $= 6.02\times10^{23}/18$. One molecule of water (H_2O) contains two hydrogen atoms and one oxygen atom. Thus the number of electrons in H_2O is sum of the number of electrons in H_2 and oxygen. There are 2 electrons in H_2 and 8 electrons oxygen.

$\therefore$ Number of electrons in $H_2O = 2+8 = 10$.

Total number of protons / electrons in 1.0 gm of water $= \frac{6.02\times10^{23}}{18}\times10 = 3.34\times10^{23}$

Total positive charge $= 3.34\times10^{23}\times$ charge on a proton

$= 3.34\times10^{23}\times1.6\times10^{-19}\ \text{C} = 5.35\times10^{4}$ C

This positive charge is balanced by equal amount of negative charge so that the water molecule is electrically neutral.

Do you know ?

According to recent advancement in physics, it is now believed that protons and neutrons are themselves built out of more elementary units called quarks. They are of six types, having fractional charge (-1/3)e or +(2/3)e. A proton or a neutron consists of a combination of three quarks. It may be clearly understood that even in the quark model, quantization of charge is not affected. It is only the step size of the charge that decreases from e to e/3. Quarks are always present in bound states and no free quarks are known to exist.
In modern day experiments it is possible to observe the discrete nature of charge in very sensitive divides such as single electron transistor

10.3.3 Conservation of Charge:

We know that when a glass rod is rubbed with silk, it becomes positively charged and silk becomes negatively charged. The amount

of positive charge on glass rod is found to be exactly the same as negative charge on silk. Thus, the systems of glass rod and silk together possesses zero net charge after rubbing.

Result and conclusion of this experiment can be generalized and we can say that "in any given physical process, charge may get transferred from one part of the system to another, but the total charge in the system remains constant" or, **for an isolated system total charge cannot be created nor destroyed.** In simple words, the total charge of an isolated system is always conserved.

10.3.4 Forces between Charges:

It was observed in carefully conducted experiments with charged objects that they experience force when brought close (not touching) to each other. This force can be attractive or repulsive. **Like charges repel each other and unlike charges attract each other.** Figure 10.2 describes this schematically. This is the reason for charging by induction as described in section 10.2 and Fig. 10.1.

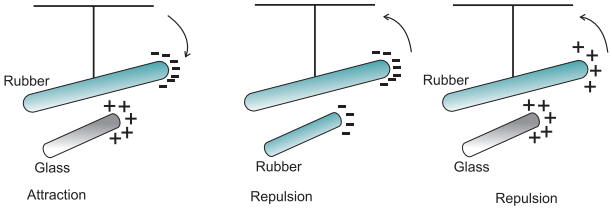


Fig. 10.2: Attractive and repulsive force.

10.4 Coulomb's Law:

The electric interaction between two charged bodies can be expressed in terms of the forces they exert on each other. Coulomb (1736-1806) made the first quantitative investigation of the force between electric charges. He used point charges at rest to study the interaction. A point charge is a charge whose dimensions are negligibly small compared to its distance from another bodies. **Coulomb's law is a fundamental law governing interaction between charges at rest.**

10.4.1 Scalar form of Coulomb's Law:

Statement : *The force of attraction or repulsion between two point charges at rest is directly proportional to the product of the magnitude of the charges and inversely proportional to the square of the distance between them.* This force acts along the line joining the two charges.

Let q_1 and q_2 be two point charges at rest with respect to each other and separated by a distance r. The magnitude F of the force between them is given by,

$$F \alpha \frac{q_1 q_2}{r^2}$$

$$F = K \frac{q_1 q_2}{r^2} \qquad \text{--- (10.1)}$$

where K is the constant of proportionality. Its magnitude depends on the units in which F, q_1, q_2 and r are expressed and also on the properties of the medium around the charges.

The force between the two charges will be attractive if they are unlike (one positive and one negative). The force will be repulsive if charges are similar (both positive or both negative). Figure 10.3 describes this schematically.

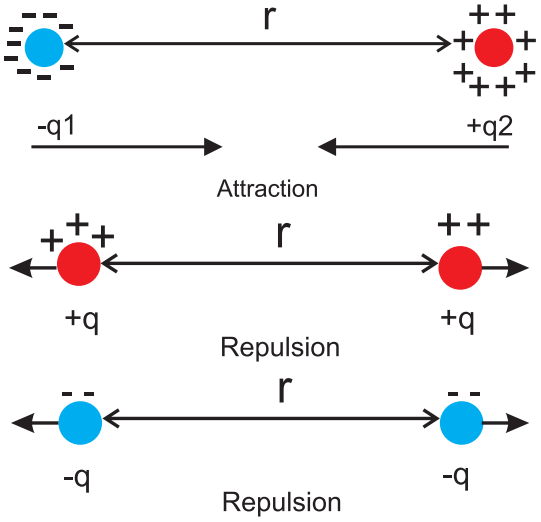


Fig. 10.3: Coulomb's law.

10.4.2 Relative Permittivity or Dielectric Constant:

While discussing the coulomb's law it was assumed that the charges are held stationery in vacuum. When the charges are kept in a material medium, such as water, mica or parafined paper, the medium affects the force between the charges. The force between the two charges placed in a medium may be written as,

$$F_{med} = \frac{1}{4\pi\varepsilon}\left(\frac{q_1 q_2}{r^2}\right) \qquad \text{--- (10.2)}$$

where ε is called the absolute permittivity of the medium. The force between the same two charges placed in free space or vacuum at distance r is given by,

$$F_{vac} = \frac{1}{4\pi\varepsilon_0}\left(\frac{q_1 q_2}{r^2}\right) \quad \text{--- (10.3)}$$

Dividing Eq. (10.3) by (10.2)

$$\frac{F_{vac}}{F_{med}} = \frac{\frac{1}{4\pi\varepsilon_0}\left(\frac{q_1 q_2}{r^2}\right)}{\frac{1}{4\pi\varepsilon}\left(\frac{q_1 q_2}{r^2}\right)} = \frac{\varepsilon}{\varepsilon_0}$$

The ratio $\frac{\varepsilon}{\varepsilon_0}$ is the relative permittivity or dielectric constant of the medium and is denoted by ε_r or K.

$$K \text{ or } \varepsilon_r = \frac{\varepsilon}{\varepsilon_0} = \frac{F_{vac}}{F_{med}} \quad \text{--- (10.4)}$$

Thus,

(i) ε_r is the ratio of absolute permittivity of a medium to the permittivity of free space.

(ii) ε_r is the ratio of the force between two point charges placed a certain distance apart in free space or vacuum to the force between the same two point charges when placed at the same distance in the given medium.

ε_r is a dimensionless quantity.

(iii) ε_r is also called specific inductive capacity.

The force between two point charges q_1 and q_2 placed at a distance r in a medium of relative permittivity ε_r, is given by

$$F = \frac{1}{4\pi\varepsilon_0\varepsilon_r}\frac{q_1 q_2}{r^2} \quad \text{--- (10.5)}$$

For water, $\varepsilon_r = 80$ then from Eq. (10.4)

$$\frac{F_{vac}}{F_{water}} = \varepsilon_r = 80$$

$$F_{water} = \frac{F_{vac}}{80}$$

This means that when two point charges are placed some distance apart in water, the force between them is reduced to $\left(\frac{1}{80}\right)^{th}$ of the force between the same two charges placed at the same distance in vacuum.

Thus, a material medium reduces the force between charges by a factor of ε_r, its relative permittivity.

While using Eq. (10.5) we assume that the medium is homogeneous, isotropic and infinitely large.

10.4.3 Definition of Unit Charge from the Coulomb's Law:

The force between two point charges q_1 and q_2, separated by a distance r in free space, is written by using Eq. (10.2),

$$F = \frac{1}{4\pi\varepsilon_0} \times \frac{q_1 q_2}{r^2} = 9\times10^9 \times \frac{q_1 q_2}{r^2}$$

$\varepsilon_0 = 8.85 \times 10^{-12}\ C^2\ N^{-1}\ m^{-2}$

If $q_1 = q_2 = 1C$ and $r = 1.0$ m

Then $F = 9.0\times10^9$ N

From this, we define, coulomb (C) the unit of charge in SI units.

One coulomb is the amount of charge which, when placed at a distance of one metre from another charge of the same magnitude in vacuum, experiences a force of 9.0×10^9 N. This force is a tremendously large force realisable in practical situations. It is, therefore, necessary to express the charge in smaller units for practical purpose. Subunits of coulomb are used in electrostatics. For example, micro-coulomb (10^{-6} C, μC), nano-coulomb (10^{-9} C, nC) or pico-coulomb. (10^{-12} C, pC) are normally used units.

Do you know ?

Force between two charges of 1.0 C each, separated by a distance of 1.0 m is 9.0×10^9 N or, about 10 million metric tonne. A normal truck-load is about 10 metric tonne. So, this force is equivalent to about one million truck-loads. A tremendously large force indeed !

Example 10.2: Charge on an electron is 1.6×10^{-19} C. How many electrons are required to accumulate a charge of one coulomb?

Solution: 1.6×10^{-19} C = 1 electron

$$\therefore 1C = \frac{1}{1.6\times10^{-19}} \text{ electrons}$$

$$= 0.625\times10^{19} = 6.25\times10^{18} \text{ electrons}$$

6.25×10^{18} electrons are required to accumulate a charge of one coulomb.

It is now possible to measure a very small amount of current in otto-amperes which measures flow of single electron.

10.4.4 Coulomb's Law in Vector Form:

As shown in Fig 10.4, q_1 and q_2 are two similar point charges situated at points A and B. r_{12} is the distance of separation between them. $\vec{F}_{21}$ denotes the force exerted on q_2 by q_1

$$\vec{F}_{21} = \frac{1}{4\pi\varepsilon_0} \times \frac{q_1 q_2}{|\hat{r}_{21}|^2} \times \hat{r}_{21} \qquad \text{--- (10.6)}$$

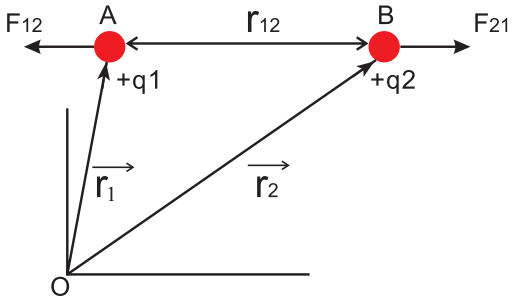


Fig. 10.4: Coulomb's law in vector form

$\hat{r}_{21}$ is the unit vector along $\overline{AB}$, away from B. Similarly, the force $\vec{F}_{12}$ exerted on q_1 by q_2 is given by

$$\vec{F}_{12} = \frac{1}{4\pi\varepsilon_0} \times \frac{q_1 q_2}{|r_{12}|^2} \times \hat{r}_{12} \qquad \text{--- (10.7)}$$

$\hat{r}_{12}$ is the unit vector along $\overline{BA}$, away from A. $\vec{F}_{12}$ acts on q_1 at A and is directed along BA, away from A. The unit vectors $\hat{r}_{12}$ and $\hat{r}_{21}$ are oppositely directed i.e., $\hat{r}_{12} = -\hat{r}_{21}$ hence,

$$\vec{F}_{21} = -\vec{F}_{12}$$

Thus, the two charges experience force of equal magnitude and opposite in direction. These two forces form an action- reaction pair. **As $\vec{F}_{21}$ and $\vec{F}_{12}$ act along the line joining the two charges, the electrostatic force is a central force.**

Example 10.3: Calculate and compare the electrostatic and gravitational forces between two protons which are 10^{-15} m apart. Value of G = 6.674×10^{-11} m^3 kg^{-1} s^{-2} and mass of the proton is 1.67×10^{-27} kg

Solution: The electrostatic force between the protons is given by $F_e = \frac{1}{4\pi\varepsilon_0}\frac{q_1 q_2}{r^2}$

Here, $q_1 = q_2 = +1.6\times10^{-19}$ C, r = 10^{-15}m

$$\therefore F_e = 9\times10^9 \frac{(1.6\times10^{-19})(1.6\times10^{-19})}{(10^{-15})^2}$$

$$= 9\times1.6\times1.6\times10^{1}\,\text{N}$$

$$F_e = 2.3\times10^{2}\,\text{N} \qquad \text{--- (10.8.a)}$$

The gravitational force between the protons is given by

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

$$= \frac{6.674\times10^{-11}\times1.67\times10^{-27}\times1.67\times10^{-27}}{(10^{-15})^2}$$

$$F_g = 1.86\times10^{-34}\,\text{N} \qquad \text{--- (10.8.b)}$$

Comparing 10.8. (a) and 10.8.(b)

$$\frac{F_e}{F_g} = \frac{2.30\times10^{-2}\,\text{N}}{1.86\times10^{-34}\,\text{N}} = 1.23\times10^{36}$$

Thus, the electrostatic force is about 36 orders of magnitude stronger than the gravitational force.

Comparison of gravitational and electrostatic forces:

Similarities

1. Both forces obey inverse square law : $F \propto \frac{1}{r^2}$
2. Both are central forces : act along the line joining the two objects.

Differences

1. Gravitational force between two objects is always attractive while electrostatic force between two charges can be either attractive or repulsive depending on the nature of charges.
2. Gravitational force is about 36 orders of magnitude weaker than the electrostatic force.

10.5 Principle of Superposition:

The principle of superposition states that *when a number of charges are interacting, the resultant force on a particular charge is given by the vector sum of the forces exerted by individual charges.*

Consider a number of point charges q_1, q_2, q_3 ------- kept at points A_1, A_2, A_3--- as shown in Fig. 10.5. The force exerted on the charge q_1 by q_2 is $\vec{F}_{12}$. The value of $\vec{F}_{12}$ is calculated

by ignoring the presence of other charges. Similarly, we find $\vec{F}_{13}$, $\vec{F}_{14}$ etc, one at a time, using the coulomb's law.

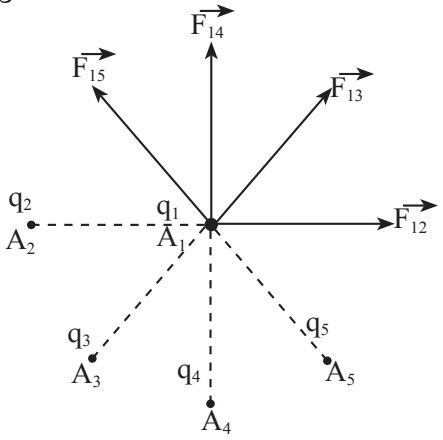


Fig. 10.5: Principle of superposition.

Total force $\vec{F}_1$ on charge q_1 is the vector sum of all such forces.

$$\vec{F}_1 = \vec{F}_{12} + \vec{F}_{13} + \vec{F}_{14} + ...$$

$$= \frac{1}{4\pi\varepsilon_0}\left[\frac{q_1 q_2}{|\vec{r}_{12}|^2}\hat{r}_{12} + \frac{q_1 q_3}{|\vec{r}_{13}|^2}\hat{r}_{13} + ...\right],$$

where $\hat{r}_{12}, \hat{r}_{13}$ etc., are unit vectors directed to q_1 from q_2, q_3 etc., and r_{12}, r_{13}, r_{14},etc., are the distances from q_1 to q_2, q_3 etc respectively.

Let there be N point charges q_1, q_2,q_3 etc., q_N. The force $\vec{F}$ exerted by these charges on a test charge q_0 can be written using the summation notation Σ as follows,

$$\vec{F}_{test} = \vec{F}_1 + \vec{F}_2 + \vec{F}_3 + ---+ \vec{F}_N \quad \text{--- (10.9)}$$

$$= \sum_{n=1}^{N} F_n = \frac{1}{4\pi\varepsilon_0} \sum_{n=1}^{N} \frac{q_0 q_n}{r_{0n}^2}\hat{r}_{0n} \quad \text{--- (10.10)}$$

Where $\hat{r}_{0n}$ is a unit vector directed from the n^{th} charge to the test charge q_o and r_{0n} is the separation between them, $\vec{r}_{0n} = r_{0n}\, r_{0n}$

Can you tell?

Three charges, *q* each, are placed at the vertices of an equilateral triangle. What will be the resultant force on charge *q* placed at the centroid of the triangle?

Example 10.4: Three charges of 2μC, 3μC and 4μC are placed at points A, B and C respectively, as shown in Fig. a. Determine the force on A due to other charges.

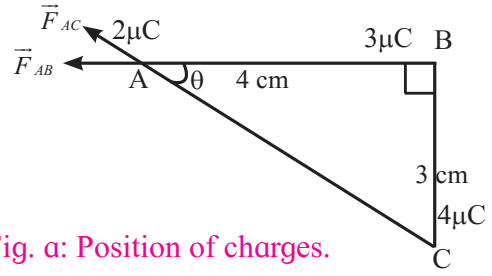


Fig. a: Position of charges.

Solution : Given,

AB = 4.0 cm, BC = 3.0 cm

$\therefore$ AC = $\sqrt{4^2 + 3^2}$ = 5.0 cm

Magnitude of force $\vec{F}_{AB}$ on A due to B is,

$$\vec{F}_{AB} = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{2\times10^{-6}\times3\times10^{-6}}{(4\times10^{-2})^2}$$

$$= \frac{9\times10^9\times6}{16\times10^{-4}}\times10^{-12}$$

$$= 3.37\times10$$

$$= 33.7\ \text{N}$$

This force acts at point A and is directed along $\overrightarrow{BA}$ (Fig. (b)).

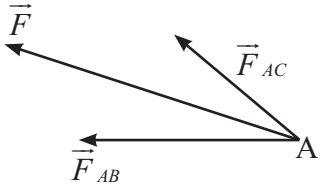


Fig. b: Forces acting at point A.

Magnitude of force $\vec{F}_{AC}$ on A due to C is,

$$\vec{F}_{AC} = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{2\times10^{-6}\times4\times10^{-6}}{(5\times10^{-2})^2}$$

$$= \frac{9\times10^9\times8.0\times10^{-12}}{25\times10^{-4}}$$

$$= \frac{72}{25}\times10 = 28.8\ \text{N}$$

This force acts at point A and is directed along $\overrightarrow{CA}$. (Fig. 10.6.(b))

$$\vec{F} = \vec{F}_{AB} + \vec{F}_{AC}$$

Magnitude of resultant force is,

$$F = \left[F_{AC}^2 + F_{AB}^2 + 2F_{AC}\cdot F_{AB}\cdot\cos\theta\right]^{1/2}$$

$$= 59.3\ \text{N}$$

Using Eq. (2.10)

calculaling tan α, α =16.93^0

Direction of the resultant force is 16.9° north of west. (Fig. c)

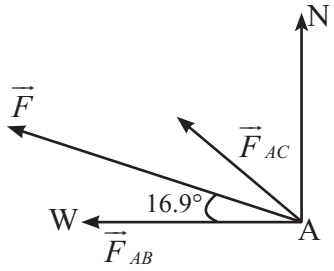


Fig. c: Direction of the resultant force.

10.6 Electric Field:

Space around a charge Q gets modified so that when a test charge is brought in this region, it experiences a coulomb force. *This region around a charged object in which coulomb force is experienced by another charge is called electric field.*

Mathematically, electric field is defined as the force experienced per unit charge. Let Q and q be two charges separated by a distance r.

The coulomb force between them is given by $\vec{F} = \frac{1}{4\pi\varepsilon_0}\frac{Qq}{r^2}\hat{r}$, where, $\hat{r}$ is the unit vector along the line joining Q to q.

Therefore, electric field due to charge Q is given by,

$$\vec{E} = \frac{\vec{F}}{q} = \frac{Q}{4\pi\varepsilon_0 r^2}\hat{r} \qquad \text{--- (10.11)}$$

The coulomb force acts across an empty space (vacuum) and does not need any intervening medium for its transmission.

The electric field exists around a charge irrespective of the presence of other charges.

Since the coulomb force is a vector, the electric field of a charge is also a vector and is directed along the direction of the coulomb force, experienced by a test charge.

A precise definition of electric field is: Electric field is the force experienced by a test charge in presence of the given charge at the given distance from it.

$$E = \lim_{q\to 0}\frac{\vec{F}}{q}$$

Test charge is a positive charge so small in magnitude that it dose not affect the surroundings of the given charge.

The magnitude of electric field at a distance r from a point charge Q is same at all points on the surface of a sphere of radius r as shown in Fig. 10.6. Its direction is along the radius of the sphere, pointing away from its centre if the charge is positive.

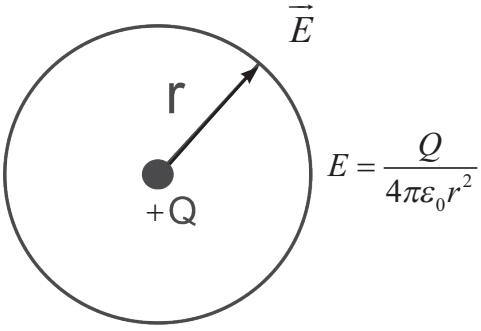


Fig. 10.6: Electric field due to a point charge (+Q).

SI unit of electric intensity is newton per coulomb (NC^{-1}). Practically, electric field is expressed in volt per metre(Vm^{-1}). This is discussed in article 10.6.2.

Dimensional formula of E is,

$$E = \frac{F}{q_0} = \frac{[LMT^{-2}]}{[IT]}$$

$$E = [LMT^{-3}I^{-1}]$$

10.6.1 Electric Field Intensity due to a Point Charge in a Material Medium:

Consider a point charge q placed at point O in a medium of dielectric constant K as shown in Fig. 10. 7.

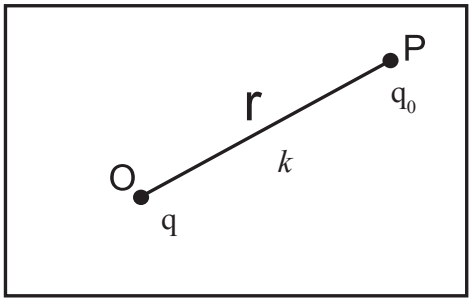


Fig. 10.7: Field in a material medium.

Consider the point P in the electric field of point charge q at distance r from it. A test charge q_0 placed at the point P will experience a force which is given by the Coulomb's law,

$$\vec{F} = \frac{1}{4\pi\varepsilon_0 K}\frac{q\,q_0}{r^2}\hat{r}$$

where $\hat{r}$ is the unit vector in the direction of force i.e., along OP.

By the definition of electric field intensity

$$\vec{E} = \frac{\vec{F}}{q_0} = \frac{1}{4\pi\varepsilon_0 K}\frac{q}{r^2}\hat{r}$$

The direction of $\vec{E}$ will be along OP when q is positive and along PO when q is negative.

The magnitude of electric field intensity in a medium is given by

$$E = \frac{1}{4\pi\varepsilon_0 K}\frac{q}{r^2} \qquad \text{--- (10.12)}$$

For air or vacuum $K = 1$ then

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

The coulomb force between two charges and electric field E of a charge both follow the inverse square law, ($F \propto 1/r^2$, $E \propto 1/r^2$) Fig. 10.8.

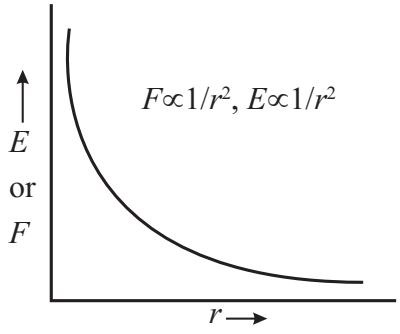


Fig. 10.8: Variation of Coulomb force/ Electric field due to a point charge.

1. **Uniform electric field:** A uniform electric field is a field whose magnitude and direction is same at all points. For example, field between two parallel plates. Fig 10.9.a
2. **Non uniform electric field:** A field whose magnitude and direction is not the same at all points. For example, field due to a point charge. In this case, the magnitude of field is same at distance r from the point charge in any direction but the direction of the field is not same. Fig 10.9.b

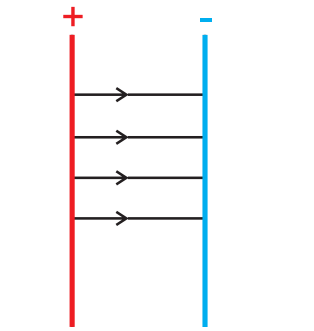


Fig. 10.9 (a): uniform electric field.

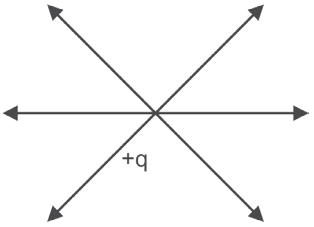


Fig. 10.9 (b): non uniform electric field.

10.6.2 Practical Way of Calculating Electric Field

A pair of charged parallel plates is arranged as shown in Fig. 10.10. The electric field between them is uniform. A potential difference V is applied between two parallel plates separated by a distance 'd'. The electric field between them is directed from plate A to plate B as shown.

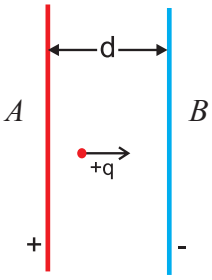


Fig 10.10: Electric field between two parallel plates.

A charge $+q$ placed between the plates experiences a force F due to the electric field. If we have to move the charge against the direction of field, i.e., towards the positive plate, we have to do some work on it. If we move the charge $+q$ from the negative plate B to the positive plate A, the work done against the field is $W = Fd$; where 'd' is the separation between the plates. The potential difference V between the two plates is given by

$W = Vq$, but $W = Fd$

$\therefore Vq = Fd \;\therefore\; F/q = V/d = E$

$\therefore$ Electric field can be defined as

$E = V/d$ --- (10.13)

This is the commonly used definition of electric field.

Example 10.5: Gap between two electrodes of the spark-plug used in an automobile engine is 1.25 mm. If the potential of 20 V is applied across the gap, what will be the magnitude of electric field between the electrodes?

Solution:

$$E = \frac{V}{d}$$

$$E = \frac{20V}{1.25\times10^{-3}m} = 16\times10^{-3} = 1.6\times10^{4}\,\frac{V}{m}$$

This electric field is sufficient to ionize the gaseous mixture of fuel compressed in the cylinder and ignite it.

Can you tell?

Why a small voltage can produce a reasonably large electric field?

Example 10.6: Three point charges are placed at the vertices of a right isosceles triangle as shown in the Fig. a. What is the magnitude and direction of the resultant electric field at point P which is the mid point of its hypotenuse?

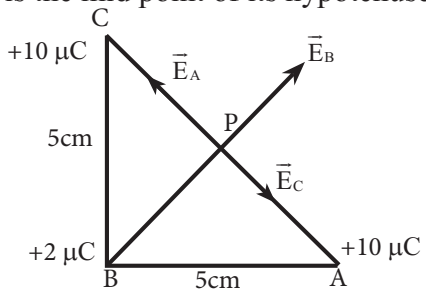


Fig (a): Position of charges.

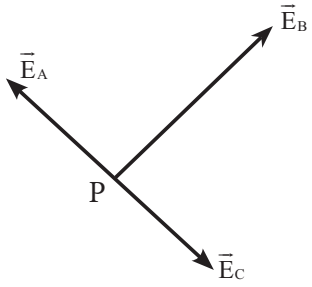


Fig (b): Electric field at point P.

Solution: Electric field is the force an a unit positive charge, the fields at P due to the charges at A, B and C are shown in the Fig. b. $\vec{E}_A$ is the field at P due to charge at A and $\vec{E}_C$ is the field at P due to charge at C. Since P is the midpoint of AC and the fields at A and C are equal in magnitudes and are opposite in direction, $\vec{E}_A = -\vec{E}_C$. $\vec{E}_A + \vec{E}_C = 0$. Thus, the field at P is only to the charge at B and can be written as

$$\vec{E}_p = \vec{E}_B = \frac{2\times10^{-6}}{4\pi\varepsilon_0(BP)^2}$$

$$\vec{E}_p = \frac{2\times10^{-6}\times9\times10^{9}}{(5/\sqrt{2})^2}$$

$$= \frac{2\times9\times10^{3}\times2}{25}$$

$$= \frac{36}{25}\times10^{3}$$

$$= 1.44\times10^{15}\ \text{NC}^{-1}\ \text{along}\ \overrightarrow{BP}$$

To calculate BP

$$\left|\overrightarrow{BP}\right| = (BA)\cos(45)^\circ = \frac{5}{\sqrt{2}}$$

Example 10.7: A simplified model of hydrogen atom consists of an electron revolving about a proton at a distance of 5.3×10^{-11}m. The charge on a proton is $+1.6\times10^{-19}$ C. Calculate the intensity of the electric field due to proton at this distance.

Solution:

$$E = \frac{q}{4\pi\varepsilon_o r^2}$$

$$q = +1.6\times10^{-19}\ C$$

$$r = 5.3\times10^{-11}\,m$$

$$\frac{1}{4\pi\varepsilon_o} = 9.0\times10^{9}\,Nm^2C^{-1}$$

$$E = \frac{1.6\times10^{-19}}{(5.3\times10^{-11})^2}\times9.0\times10^{9} = 5.1\times10^{11}\,NC^{-1}$$

The force between electron and proton in hydrogen atom can be calculated by using the electric field. We have, $E = \frac{F}{q} \therefore F = qE$

$$F = -1.6\times10^{-19}\ \text{C}\times 5.1\times10^{11}\ \text{NC}^{-1}$$

$$= -8.16\times10^{8}\ \text{N}.$$

This force is attractive.

Using the Coulomb's law,

$$F = \frac{1}{4\pi\varepsilon_0}\frac{q_1q_2}{r^2}$$

$$= 9.0\times10^{9}\,Nm^2C^{-1}\times\frac{(-1.6\times10^{-19}C)\times(+1.6\times10^{-19}C)}{(5.3\times10^{-11}m)^2}$$

$$= -8.6\times10^{8}\,N$$

Knowing electric field at a point is useful to estimate the force experienced by a charge at that point.

10.6.3 Electric Lines of Force:

Michael Faraday (1791-1867) introduced the concept of lines of force for visualising electric and magnetic fields. *An electric line of force is an imaginary curve drawn in such a way that the tangent at any given point on this curve gives the direction of the electric field at that point.* See Fig.10.11. If a test charge is placed in an electric field it would be acted upon by a force at every point in the field and will move along a path. **The path along which the unit positive charge moves is called a line of force.**

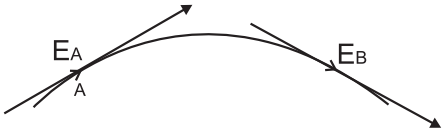


Fig. 10.11: Electric line of force.

A line of force is defined as a curve such that the tangent at any point to this curve gives the direction of the electric field at that point.

The density of field lines indicates the strength of electric fields at the given point in space. Figure 10.12.

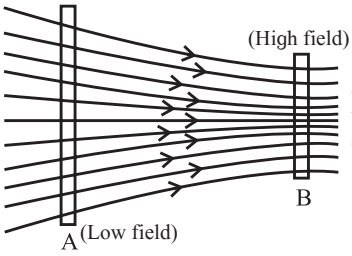


Fig. 10.12: density of field lines and strength of electric field.

Characteristics of electric lines of force

(1) The lines of force originate from a positively charged object and terminate on a negatively charged object.

(2) The lines of force neither intersect nor meet each other, as it will mean that electric field has two directions at a single point.

(3) The lines of force leave or terminate on a conductor normally.

(4) The lines of force do not pass through conductor i.e. electric field inside a conductor is always zero, but they pass through insulators.

(5) Magnitude of the electric field intensity is proportional to the number of lines of force per unit area of the surface held perpendicular to the field.

(6) Electric lines of force are crowded in a region where electric intensity is large.

(7) Electric lines of force are widely separated from each other in a region where electric intensity is small

(8) The lines of force of an uniform electric field are parallel to each other and are equally spaced.

The lines of force are purely a geometric construction which help us visualise the nature of electric field in a region. **The lines of force have no physical existence.**

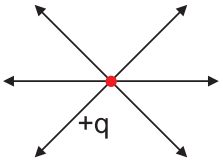


Fig. 10.13 (a): Lines of force due to positive charge.

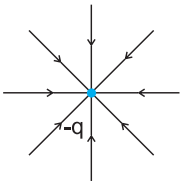


Fig. 10.13 (b): Lines of force due to negative charge.

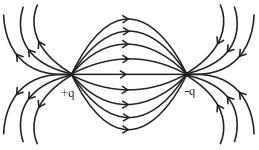


Fig. 10.13 (c): Lines of force due to opposite charge.

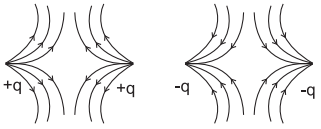


Fig. 10.13; (d): Lines of force due to similar charge.

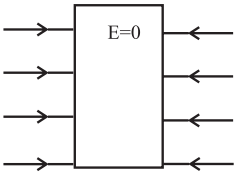


Fig. 10.13 (e): Lines of force terminate on a conductor.

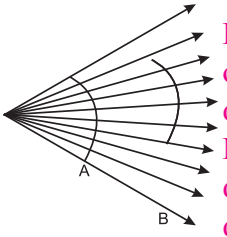


Fig. 10.13 (f): Intensity of a electric field is more at point A and less at B. More lines cross the area at A and less at the same area at B.

Fig. 10.13: The lines of force due to various geometrical arrangement of electrical charges.

Can you tell?

Lines of force are imaginary, can they have any practical use?

10.7 Electric Flux:

As discussed previously, the number of lines of force per unit area is the intensity of the electric field $\vec{E}$.

$$\therefore E = \frac{\text{Number of lines of force}}{\text{Area enclosing the lines of force}} \quad \text{- (10.14)}$$

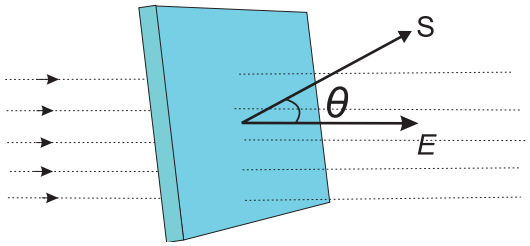


Fig. 10.14: Flux through area S.

Number of lines of force = $(E)\cdot(\text{Area})$ When the area is inclined at an angle θ with the direction of electric field, Fig. 10.14, the electric flux can be calculated as follows.

Let the angle between electric field $\vec{E}$ and area vector $d\vec{S}$ be θ, then the electric flux passing through area dS is given by

$d\phi$ = (component of dS along $\vec{E}$)·(area of $d\vec{S}$)

$d\phi = E\,(dS\cos\theta)$

$d\phi = EdS\cos\theta$

$$d\phi = \vec{E}\cdot d\vec{S} \qquad \text{--- (10.15)}$$

Total flux through the entire surface

$$\Phi = \int d\phi = \int_s \vec{E}.d\vec{S} = \vec{E}.\vec{S} \qquad \text{--- (10.16)}$$

The SI unit of electric flux can be calculated using,

$\Phi = \vec{E}\cdot\vec{S}$ =(V/m) m^2 =Vm

10.8 Gauss' Law:

Karl Friedrich Gauss (1777-1855) one of the greatest mathematician of all times, formulated a law expressing the relationship between the electric charge and its electric field which is called the Gauss' law. Gauss' law is analongous to Coulomb's law in the sense that it too expresses the relationship between electric field and electric charge. Gauss' law provides equivalent method for finding electric intensity. It relates values of field at a closed surface and the total charges enclosed by that surface.

Consider a closed surface of any shape which encloses number of positive electric charges (Fig. 10.15). To prove Gauss' theorem, imagine a small charge +q present at a point O inside closed surface. Imagine an infinitesimal area dA of the given irregular closed surface.

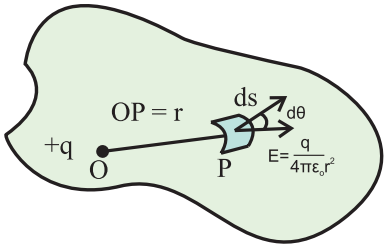


Fig. 10.15: Gauss' law.

The magnitude of electric field intensity at point P on dS due to charge +q at point O is,

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

The direction of E is away from point O. Let θ be the angle subtended by normal drawn to area dS and the direction of E. Electric flux, $d\phi$, passing through area dS, $= E\cos\theta\, dS$

$$= \frac{q}{4\pi\varepsilon_0 r^2}\cos\theta\, dS$$

$$= \left(\frac{q}{4\pi\varepsilon_0}\right)\left(\frac{dS\cos\theta}{r^2}\right)$$

$$d\phi = \left(\frac{q}{4\pi\varepsilon_0}\right)d\omega \qquad \text{--- (10.17)}$$

where, $d\omega = \frac{dS\cos\theta}{r^2}$ is the solid angle subtended by area dS at point O.

Total electric flux, ϕ_E, crossing the given closed surface can be obtained by integrating Eq. (10.17) over its area. Thus,

$$\Phi_E = \int d\phi = \int_s \vec{E}.\vec{ds} = \int \frac{q}{4\pi\varepsilon_0}\,d\omega = \frac{q}{4\pi\varepsilon_0}\int d\omega$$

But $\int d\omega = 4\pi =$ solid angle subtended by entire closed surface at point O

$$\text{Total flux} = \frac{q}{4\pi\varepsilon_0}(4\pi)$$

$$\Phi_E = \int_s \vec{E}\cdot\vec{ds} = +q/\varepsilon_0 \qquad \text{--- (10.18)}$$

This is true for every electric charge enclosed by a given closed surface.

Total flux due to charge q_1, over the given closed surface = $+ q_1/\varepsilon_0$

Total flux due to charge q_2, over the given closed surface = + q_2/ε_0

Total flux due to charge q_n, over the given closed surface = + q_n/ε_0

Positive sign in Eq. (10.18) indicates that the flux is directed outwards, away form the charge. If the charge is negative, the flux will be is directed inwards as shown in Fig 10.16 (b). If a charge is outside the closed surface the net flux through it will be zero Fig 10.16 (c).

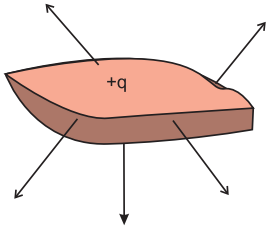


Fig. 10.16 (a): Flux due to positive charge.

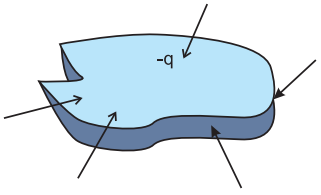


Fig. 10.16 (b): Flux due to negative charge.

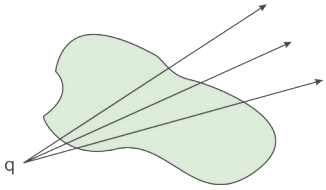


Fig. 10.16 (c): Flux due to charge outside a closed surface is zero.

According to the superposition principle, the total flux ϕ due to all charges enclosed within the given closed surface is

$$\Phi_E = \frac{q_1}{\varepsilon_0} + \frac{q_2}{\varepsilon_0} + \frac{q_3}{\varepsilon_0} + \text{----} + \frac{q_n}{\varepsilon_0} = \sum_{i=1}^{i=n} \frac{q_i}{\varepsilon_0} = \frac{Q}{\varepsilon_0}$$

Statement of Gauss' law

The flux of the net electric field through a closed surface equals the net charge enclosed by the surface divided by ε_0

$$\int \vec{E} \cdot \vec{dS} = \frac{Q}{\varepsilon_0}$$

where Q is the total charge within the surface.

Gauss' law is applicable to any closed surface of regular or irregular shape.

Example 10.8: A charge of 5.0 C is kept at the centre of a sphere of radius 1 m. What is the flux passing through the sphere? How will this value change if the radius of the sphere is doubled?

Solution: Flux per unit area is given by Eq. 10.16.

According to Gauss law, the total flux through the sphere $\Phi = \int \vec{E}.\vec{ds}$, where the integration is over the surface of the sphere. As the electric field is same all over the sphere i.e. $|\bar{E}|$ = constant and the direction of $\bar{E}$ as well as that of $\bar{ds}$ is along the radius, we get

$$\text{flux} = \Phi = |E| 4\pi R^2$$

$$E = \frac{q}{4\pi\varepsilon_0 r^2} = 9\times10^9 \times \frac{5.0\,\text{C}}{(1.0\,\text{m})^2}$$

$$E = 9\times10^9 \times 5 = 4.5\times10^{10}\,\text{NC}^{-1}$$

$$\phi = \vec{E}\cdot\vec{S}$$

$$\therefore \text{flux} = 4.5\times10^{10} \times 4\pi(1)^2$$

$$= 5.65\times10^{11}\,\text{Vm}$$

Thus the total flux is independent of radius.

$E \propto 1/r^2$, and area $\propto r^2$. This can also be seen from Gauss' law, where the net flux crossing a closed surface is equal to q/ε_0 where q is the net charge inside the closed surface. As the charge inside the sphere is unchanged, the flux passing through a sphere of any radius is the same. Thus, if the radius of the sphere is increased by a factor of 2, the net flux passing through its surface remains unchanged. As shown in Fig. 10.17, same number of lines of force cross both the surfaces. The total flux is independent of shape of the closed surface because Eq. 10.18 does not involve any radius.

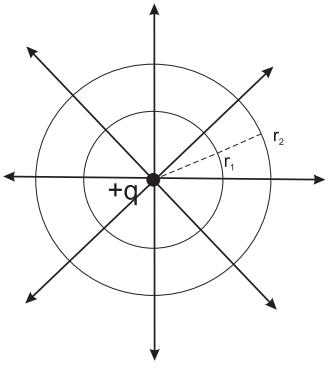


Fig. 10.17: Flux is independent of the shape and size of closed surface.

Do you know ?

Gaussian surface

All the lines of force originating from a point charge penetrate an imaginary three dimensional surface. The total flux $\Phi_E = q/\varepsilon_0$. The same number of lines of force will cross the surface of any shape. The total flux through both the surfaces is the same. Calculating flux involves calculating $\int \vec{E} \cdot \vec{ds}$, hence it is convenient to consider a regular surface surrounding the given charge distribution. A surface enclosing the given charge distribution and symmetric about it is a Gaussian surface.

For example. if we have a point charge the Gaussian surface will be a sphere. If the charge distribution is linear, the Gaussian surface would be a cylinder with the charges distributed along its axis. Gaussian surface offers convenience of calculating the integral $\int \vec{E} \cdot \vec{ds}$.

Remember that a Gaussian surface is purely imaginary and does not exist physically.

10.9 Electric Dipole:

A pair of equal and opposite charges separated by a finite distance is called an electric dipole. It is shown in Fig. (10.18).

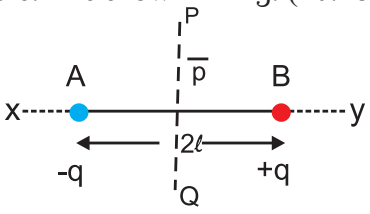


Fig. 10.18: Electric dipole. x-y axial line, P-Q equatorial line.

Line joining the two charges is called the dipole axis. A line passing through the dipole axis is called **axial line**. A line passing through the centre of the dipole and perpendicular to the axial line is called the **equatorial line** as shown in Fig. 10.18.

Strength of a dipole is measured in terms of a quantity called the **dipole moment**. Let q be the magnitude of each charge and $2\vec{l}$ be the distance from negative charge to positive charge. Then the product q $(2\vec{l})$ is called the dipole moment $\vec{p}$.

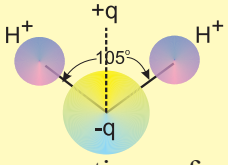
Dipole moment is defined as

$$\vec{p} = q(2\vec{l}) \qquad \text{--- (10.19)}$$

A dipole moment is a vector whose magnitude is q $(2l)$ and the direction is from the negative to the positive charge. The unit of dipole moment is Columb-meter (Cm) or Debye (D). $1D = 3.33\times10^{-30}$ Cm. If two charges +e and -e are separated by $1.0A^0$, the dipole moment is 1.6×10^{-29} Cm or 4.8 D. For example, a water molecule has a permanent dipole moment of

Natural dipole:

The water molecule is non-linear, i.e., the two hydrogen atoms and one oxygen atom are not in a straight line. The two hydrogen-oxygen bonds in water molecule are at an angle of 105°. The positive charge of a water molecule is effectively concentrated on the hydrogen side and the negative charge on the oxygen side of the molecule. Thus, the positive and negative charges of the water molecule are inherently separated by a small distance. This separation of positive and negative charges gives rise to the permanent dipole moment of a water molecule.



Molecules of water, ammonia, sulphur di-oxide, sodium chloride etc. have an inherent separation of centers of positive and negative charges. Such molecules are called polar molecules.

Polar molecules are the molecules in which the center of positive charge and the negative charge is naturally separated.

Molecules such as H_2, Cl_2, CO_2 CH_4 and many others have their positive and negative charges effectively centered at the same point and are called non-polar molecules.

Non-polar molecules are the molecules in which the center of positive charge and the negative charge is one and the same. They do not have a permanent electric dipole. When an external electric field is applied to such molecules the centers of positive and negative charge are displaced and a dipole is induced.

6.172×10^{-30} Cm or 1.85 D. Its direction is from oxygen to hydrogen. See box on Natural dipole.

10.9.1 Couple Acting on an Electric Dipole in a Uniform Electric Field:

Consider an electric dipole placed in a uniform electric field E. The axis of electric dipole makes an angle θ with the direction of electric field as shown in Fig. 10.19 a.

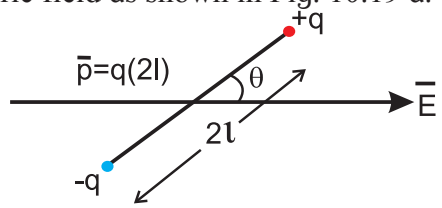


Fig. 10.19 (a): Dipole in uniform electric field.

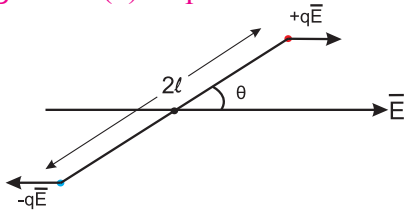


Fig. 10.19 (b): Couple acting on a dipole.

Figure 10.19. b shows the couple acting on an electric dipole in uniform electric field.

The force acting on charge - q at A is

$\vec{F}_A = -q\vec{E}$ in the direction opposite to that of $\vec{E}$ and the force acting on charge +q at B is $\vec{F}_B = +q\vec{E}$ in the direction of $\vec{E}$. Since $\vec{F}_A = -\vec{F}_B$, the two equal and opposite forces separated by a distance form a couple. Moment of the couple is called torque and is defined by $\vec{\tau} = \vec{d} \times \vec{F}$ where, d is the perpendicular distance between the two equal and opposite forces.

∴Magnitude of Torque =
Magnitude of force × Perpendicular distance

∴ Torque on the dipole = $\vec{\tau} = \overrightarrow{BP} \times q\vec{E}$

∴ $\tau = qE2l\sin\theta$ --- (10.20)

but $\vec{p} = q \times 2\vec{l}$

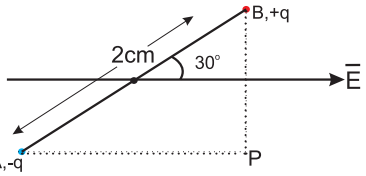
∴ $\tau = pE\sin\theta$ --- (10.21)

In vector form $\vec{\tau} = \vec{p} \times \vec{E}$ --- (10.22)

If θ =90° sin θ =1, then $\tau = pE$

When the axis of electric dipole is perpendicular to uniform electric field, torque of the couple acting on the electric dipole is maximum, i.e., $\tau = pE$. It θ = 0 then τ = 0, this is the minimum torque on the dipole. Torque tends to align the dipole axis along the direction of electric field.

Example 10.9: An electric dipole of length 2.0 cm is placed with its axis making an angle of 30° with a uniform electric field of 10^5 N/C. as shown in figure. If it experiences a torque of $10\sqrt{3}$ Nm, calculate the magnitude of charge on the dipole.



Solution: Given

$\tau = 10\sqrt{3}, Nm,\ E = 10^5\ N/C,$

$2l = 2.0\times10^{-2}\,m,\ \theta = 30^0$

$\tau = qE2l\sin\theta$

$10\sqrt{3}Nm = q10^5\ N/C 2.0\times10^{-2}\,m\left(\frac{1}{2}\right)$

$\therefore q = \frac{10\sqrt{3}}{10^3} = \sqrt{3}\times10^{-2}C$

$= 1.73\times10^{-2}C$

10.9.2 Electric Intensity at a Point due to an Electric Dipole:

Case 1 : At a point on the axis of a dipole.

Consider an electric dipole consisting of two charges -q and +q separated by a distance $2l$ as shown in Fig. 10.20. Let P be a point at a distance r from the centre C of the dipole. The electric intensity $\vec{E}_a$ at P due to the dipole is the vector sum of the field due to the charge - q at A and + q at B.

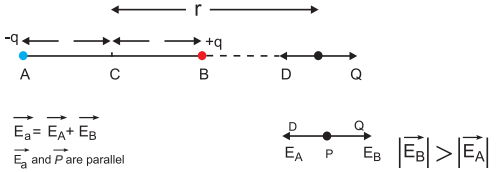


Fig. 10.20: Electric field of a dipole along its axis.

Electric field intensity at P due to the charge -q at A

$$= \vec{E}_A = \frac{1}{4\pi\varepsilon_0}\frac{(-q)}{(r+l)^2}\ \widehat{PD}$$

where $\widehat{PD}$ is unit vector directed along $\overrightarrow{PD}$

Electric intensity at P due to charge +q at B

$$= \vec{E}_B = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r-l)^2}\ \widehat{PQ}$$

where $\widehat{PQ}$ is a unit vector directed along $\overrightarrow{PQ}$.

The magnitude of $\vec{E}_B$ is greater than that of $\vec{E}_A$. (Because BP < AB)

Resultant field $\vec{E}_a$ at P on the axis, due to the dipole is

$$\vec{E}_a = \vec{E}_B + \vec{E}_A$$

The magnitude of $\vec{E}_a$ is given by

$$|\vec{E}_a| = \frac{1}{4\pi\varepsilon_0}\left[\frac{q}{(r-l)^2} - \frac{q}{(r+l)^2}\right]$$

$$|\vec{E}_a| = \frac{q}{4\pi\varepsilon_0}\left[\frac{r^2 + l^2 + 2lr - r^2 + 2lr - l^2}{(r^2 - l^2)^2}\right]$$

$$|\vec{E}_a| = \frac{2(2lq)r}{4\pi\varepsilon_0 (r^2 - l^2)^2}$$

But $2lq$ = p, the dipole moment

$$|\vec{E}_a| = \frac{1}{4\pi\varepsilon_0}\frac{2pr}{(r^2 - l^2)^2} \qquad \text{--- (10.23)}$$

$\vec{E}_a$, is directed along PQ, which is the direction of the dipole moment $\vec{p}$ i.e. from the negative to the positive charge, parallel to the axis. If r>> l, l^2 can be neglected compared to r^2,

$$\left|\vec{E}_a\right| = \frac{1}{4\pi\varepsilon_0}\frac{2p}{r^3} \qquad \text{--- (10.24)}$$

The field will be along the direction of the dipole moment $\vec{p}$.

Case 2: At a point on the equatorial line. As shown in Fig. 10.21 (a)

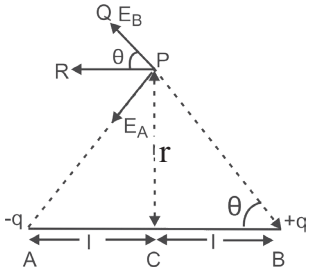


Fig. 10.21 (a): Electric field of a dipole at a point on the equatorial line.

Electric field at point P due to charge -q at A is: $\vec{E}_A = \frac{1}{4\pi\varepsilon_0}\frac{(-q)}{(AP)^2}\widehat{PA}$

where PA is the unit vector direction along $\overrightarrow{PA}$.

Similarly, Electric field at P due to charge + q at B is: $\vec{E}_B = \frac{1}{4\pi\varepsilon_0}\frac{(+q)}{(BP)^2}\widehat{PQ}$

where $\widehat{PQ}$ is the unit vector directed along $\overrightarrow{PQ}$ or $\overrightarrow{BP}$

Electric field at $\vec{P}$ is the, sum of $\vec{E}_A$ and $\vec{E}_B$

$$\therefore \vec{E}_{eq} = \vec{E}_A + \vec{E}_B$$

Consider Δ ACP

$$(AP)^2 = (PC)^2 + (AC)^2 = r^2 + l^2 = (BP)^2$$

$$\therefore \left|\vec{E}_A\right| = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r^2 + l^2)} \qquad \text{--- (10.25)}$$

$$\left|\vec{E}_B\right| = \frac{1}{4\pi\varepsilon_0}\frac{q}{(r^2 + l^2)} \qquad \text{--- (10.26)}$$

$$\left|\vec{E}_A\right| = \left|\vec{E}_B\right|$$

The resultant of fields $\vec{E}_A$ and $\vec{E}_B$ acting at point P can be calculated by resolving these vectors $\vec{E}_A$ and $\vec{E}_B$ along the equatorial line and along a direction perpendicular to it.

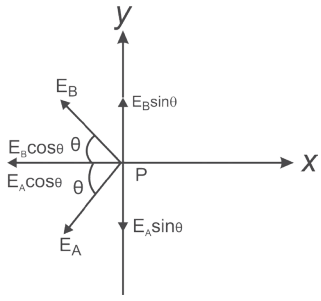


Fig. 10.21 (b): Components of the field at point P.

Consider Fig. 10.21 (b). Let the y-axis coincide with the equator of the dipole x-axis will be parallel to dipole axis, as shown. The origin is at point P.

The y-components of E_A and E_B are $E_A\sin\theta$ and $E_B\sin\theta$ respectively. They are equal in magnitude but opposite in direction and cancel each other. There is no contribution from them towards the resultant.

The x-components of E_A and E_B are $E_A\cos\theta$ and $E_B\cos\theta$ respectively. They are of equal magnitude and are in the same direction

$$\therefore |\vec{E}_{eq}| = E_A\cos\theta + E_B\cos\theta \qquad \text{--- (10.27)}$$

By using Eq. 10.25 and 10.26

$$\left|\vec{E}_{eq}\right| = 2E_A\cos\theta$$

$$= 2\left(\frac{q}{4\pi\varepsilon_0(r^2 + l^2)}\right)\frac{l}{\sqrt{r^2 + l^2}}$$

$$= \frac{2ql}{4\pi\varepsilon_0(r^2 + l^2)^{3/2}}$$

If r>>l then l^2 is very small compared to r^2

$$\left|\vec{E}_{eq}\right| = \frac{1}{4\pi\varepsilon_0}\frac{p}{(r^2)^{3/2}} = \frac{1}{4\pi\varepsilon_0}\frac{p}{r^3} \quad \text{--- (10.28)}$$

The direction of this field is along - $\vec{p}$ (anti-parallel to $\vec{p}$) as shown in Fig. 10.21 (c).

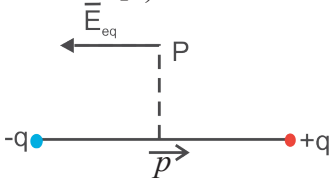


Fig. 10.21 (c): Electric field at point P is anti-parallel to $\vec{p}$.

Comparing Eq. 10.28 and 10.24 we find that the electric intensity at an axial point is twice that at a point on the equatorial position, lying at the same distance from the centre of the dipole.

10.10 Continuous Charge Distribution:

A system of charges can be considered as a continuous charge distribution, if the charges are located very close together, compared to their distances from the point where the intensity of electric field is to be found out.

The charge distribution is continuous in the sense that, a system of closely spaced charges is equivalent to a total charge which is continuously distributed along a line or a surface or a volume. To find the electric field due to continuous charge distribution, we define following terms for different types of charge distribution.

(a) Linear charge density (λ).

As shown in Fig. 10.22 charge q is uniformly distributed along a liner conductor of length l. The linear charge density λ is defined as,

$$\lambda = \frac{q}{l} \quad \text{--- (10.29)}$$

SI unit of λ is (C/m).

For example, charge distributed uniformly on a straight thin rod or a thin nylon thread. If the charge is not distributed uniformly over the length of thin conductor then charge dq on small element of length dl can be written as dq = λdl

Fig. 10.22: Linear charge.

(b) Surface charge density (σ)

Suppose a charge q is uniformly distributed over a surface of area A. As shown in Fig. 10.23, then the surface charge density σ is defined as

$$\sigma = \frac{q}{A} \quad \text{--- (10.30)}$$

SI unit of σ is (C/m²)

For example, charge distributed uniformly on a thin disc or a synthetic cloth. If the charge is not distributed uniformally over the surface of a conductor, then charge dq on small area element dA can be written as $dq = \sigma\, dA$.

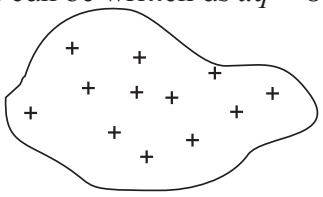


Fig. 10.23: Surface charge.

(c) Volume charge density (ρ)

Suppose a charge q is uniformly distributed throughout a volume V, then the volume charge density ρ is defined as the charge per unit volume.

$$\rho = \frac{q}{V} \quad \text{--- (10.31)}$$

S.I. unit of ρ is (C/m³)

For example, charge on a solid plastic sphere or a solid plastic cube.

If the charge is not distributed uniformaly over the volume of a material, then charge dq over small volume element dV can be written as $dq = \rho\, dV$.

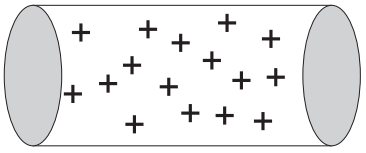


Fig. 10.24:Volume charge.

Electric field due to a continuous charge distribution can be calculated by adding electric fields due to all these small charges.

Can you tell?

The surface charge density of Earth is σ = -1.33 nC/m². That is about 8.3×10⁹ electrons per square meter. If that is the case why don't we feel it?

Do you know ?

Static charge can be useful

Static charges can be created whenever there is a friction between an insulator and other object. For example, when an insulator like rubber or ebonite is rubbed against a cloth, the friction between them causes electrons to be transferred from one to the other. This property of insulators is used in many applications such as Photocopier, Inkjet printer, Panting metal panels, Electrostatic precipitation/separators etc.

Static charge can be harmful

i. When charge transferred from one body to other is very large sparking can take place. For example lightning in sky.
ii. Sparking can be dangerous while refuelling your vehicle.
iii. One can get static shock if charge transferred is large.
iv. Dust or dirt particles gathered on computer or TV screens can catch static charges and can be troublesome.

Precautions against static charge

i. Home appliances should be grounded.
ii. Avoid using rubber soled footwear.
iii. Keep your surroundings humid. (dry air can retain static charges).

Internet my friend

1. https://www.physicsclassroom.com>class
2. https://courses.lumenlearing.com>elect
3. https://www.khanacademy.org>science.
4. https://www.topper.com>guides>physics

Exercises

1. Choose the correct option.

i. A positively charged glass rod is brought close to a metallic rod isolated from ground. The charge on the side of the metallic rod away from the glass rod will be
(A) same as that on the glass rod and equal in quantity
(B) opposite to that on the glass of and equal in quantity
(C) same as that on the glass rod but lesser in quantity
(D) same as that on the glass rod but more in quantity

ii. An electron is placed between two parallel plates connected to a battery. If the battery is switched on, the electron will
(A) be attracted to the +ve plate
(B) be attracted to the -ve plate
(C) remain stationary
(D) will move parallel to the plates

iii. A charge of + 7 μC is placed at the centre of two concentric spheres with radius 2.0 cm and 4.0 cm respectively. The ratio of the flux through them will be
(A) 1:4 (B) 1:2
(C) 1:1 (D) 1:16

iv. Two charges of 1.0 C each are placed one meter apart in free space. The force between them will be
(A) 1.0 N (B) 9×10^{9} N
(C) 9×10^{-9} N (D) 10 N

v. Two point charges of +5 μC are so placed that they experience a force of 80×10^{-3} N. They are then moved apart, so that the force is now 2.0×10^{-3} N. The distance between them is now
(A) 1/4 the previous distance
(B) double the previous distance
(C) four times the previous distance
(D) half the previous distance

vi. A metallic sphere A isolated from ground is charged to +50 μC. This sphere is brought in contact with other isolated metallic sphere B of half the radius of sphere A. The charge on the two sphere will be now in the ratio
(A) 1:2 (B) 2:1
(C) 4:1 (D) 1:1

vii. Which of the following produces uniform electric field?

(A) point charge
(B) linear charge
(C) two parallel plates
(D) charge distributed an circular any

viii. Two point charges of A = +5.0 μC and B = -5.0 μC are separated by 5.0 cm. A point charge C = 1.0 μC is placed at 3.0 cm away from the centre on the perpendicular bisector of the line joining the two point charges. The charge at C will experience a force directed towards
(A) point A
(B) point B
(C) a direction parallel to line AB
(D) a direction along the perpendicular bisector

2. Answer the following questions.

i. What is the magnitude of charge on an electron?
ii. State the law of conservation of charge.
iii. Define a unit charge.
iv. Two parallel plates have a potential difference of 10V between them. If the plates are 0.5 mm apart, what will be the strength of electric charge.
v. What is uniform electric field?
vi. If two lines of force intersect of one point. What does it mean?
vii. State the units of linear charge density.
viii. What is the unit of dipole moment?
ix. What is relative permittivity?

3. Solve numerical examples.

i. Two small spheres 18 cm apart have equal negative charges and repel each other with the force of 6×10^{-3} N. Find the total charge on both spheres.
[Ans: $q = 2.94\times10^{-10}$ C]

ii. A charge +q exerts a force of magnitude -0.2 N on another charge -2q. If they are separated by 25.0 cm, determine the value of q.
[Ans: q = 0.8333 μC]

iii. Four charges of + 6×10^{-8} C each are placed at the corners of a square whose sides a are 3 cm each. Calculate the resultant force on each charge and show its direction on a digram drawn to scale.
[Ans: 6.89×10^{-2} N]

iv. The electric field in a region is given by $\vec{E}$ = 5.0 kN/C. Calculate the electric flux Through a square of side 10.0 cm in the following cases
(a) the square is along the XY plane
[Ans: = 5.0×10^{-2} Vm]
(b) The square is along XZ plane
[Ans: Zero]
(c) The normal to the square makes an angle of 45^0 with the Z axis.
[Ans: 3.5×10^{-2} Vm]

v. Three equal charges of 10×10^{-8} C respectively, each located at the corners of a right triangle whose sides are 15 cm, 20 cm and 25 cm respectively. Find the force exerted on the charge located at the 90° angle.
[Ans: $4.59.\times10^{-3}$ N]

vi. A potential difference of 5000 volt is applied between two parallel plates 5 cm a part a small oil drop having a charge of 9.6×10^{-19} C falls between the plates. Find (a) electric field intensity between the plates and (b) the force on the oil drop.
[Ans: (a) $1.0.\times10^{5}$ N/C
(b) $9.6.\times10^{-14}$ N]

vii. Calculate the electric field due to a charge of -8.0×10^{-8} C at a distance of 5.0 cm from it.
[Ans: 2.88×10^{-2} N/C]

***

11. Electric Current Through Conductors

Can you recall?

1. Do you recall that the flow of charged particles in a conductor constitutes a current?
2. An electric current in a metallic conductor such as a wire is due to flow of electrons, the negatively charged particles in the wire.
3. What is the role of the valence electrons which are the outermost electrons of an atom?

11.1 Introduction:

The valence electrons become de-localized when a large number of atoms come together in a metal. These are the conduction electrons or free electrons constituting an electric current when a potential difference is applied across the conductor.

11.2 Electric current :

Consider an imaginary gas of both negatively and positively charged particles. Fig. 11.1 shows the negatively and positively charged particles flowing randomly in various directions across a plane P. In a time interval t, let the amount of positive charge flowing in the forward direction be q^+ and the amount of negative charge flowing in the forward direction be q^-.

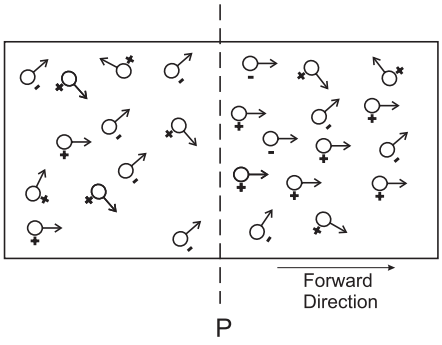


Fig. 11.1: Flow of charged particles.

Thus the net charge flowing in the forward direction is $q = q^+ - q^-$. For a steady flow, this quantity is proportional to the time t. The ratio $\frac{q}{t}$ is defined as the current I.

$$I = \frac{q}{t} \qquad \text{--- (11.1)}$$

SI unit of the current is ampere (A), that of the charge and time is coulomb (C) and second (s) respectivly.

Let I be the current varying with time. Let Δq be the amount of net charge flowing across the plane P from time t to $t + \Delta t$, i.e. during the time interval Δt. Then the current is given by

$$I(l) = \lim_{\Delta t \to o} \frac{\Delta q}{\Delta t} \qquad \text{--- (11.2)}$$

Here, the current is expressed as the limit of the ratio $\Delta q / \Delta t$ as Δt tends to zero.

The current during lightening could be as high as 10,000 A, while the current in the house hold circuit could be of the order of a few amperes. Currents of the order of a milliampere (mA), a microampere (μA) or a nanoampere (nA) are common in semiconductor devices.

11.3 Flow of current through a conductor :

A current can be generated by positively or negatively charged particles. In an electrolyte, both positively and negatively charged particles take part in the conduction. In a metal, the free electrons are responsible for conduction. These electrons flow and generate a net current under the action of an applied electric field. As long as a steady field exists, the electrons continue to flow in the form of a steady current. Such steady electric fields are generated by cells and batteries.

Do you know ?

Sign convention : The direction of the current in a circuit is drawn in the direction in which positively charged particles would move, even if the current is constituted by the negatively charged particles, electrons, which move in the direction opposite to that the electric field. We use this as a convention.

11.4 Drift speed :

Imagine a copper rod with no current flowing through it. Fig 11.2 shows the schematic of a conductor with the free electrons

in random motion. There is no net motion of these electrons in any direction. If electric field is applied along the length of the copper rod, and a current is set up in the rod, these electrons still move randomly, but tend to 'drift' in a particular direction. Their direction is opposite to that of the applied electric field.

Direction of electric field : Direction of an electric field at a point is the direction of the force on the test charge placed at that point.

The electrons under the action of the applied electric field drift with a drift speed V_d. The drift speed in a copper conductor is of the order of 10^{-4} m/s-10^{-5} m/s, whereas the electron random speed is of the order of 10^6 m/s.

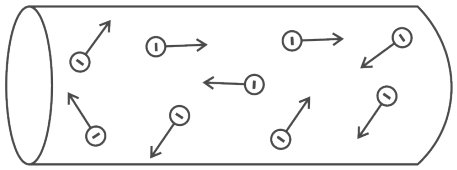


Fig. 11.2: Free electrons in random motion inside the conductor.

How is the current through a conductor related to the drift speed of electrons? Figure 11.3 shows a part of conducting wire with its free electrons having the drift speed V_d in the direction opposite to the electric field $\vec{E}$.

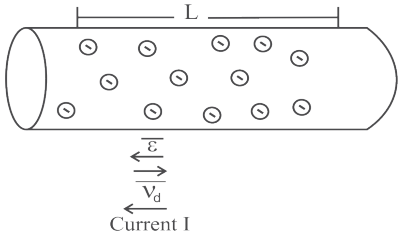


Fig. 11.3: Conducting wire with the applied electric field.

It is assumed that all the electron move with the same drift speed V_d and that, the current I is the same throughout the cross section (A) of the wire. Consider the length L of the wire. Let n be the number of free electrons per unit volume of the wire. Then the total number of electrons in the length L of the conducting wire is nAL. The total charge in the length L is,

$$q = n\,A\,L\,e \qquad \text{--- (11.3)}$$

where e is the electron charge.

This is total charge that moves through any cross section of the wire in a certain time interval t,

$$t = \frac{L}{V_d} \qquad \text{--- (11.4)}$$

From the Eq. (11.1), and Eq. (11.3), the current

$$\text{I} = \frac{q}{t} = \frac{n\,A\,L\,e}{L/V_d} = n\,A V_d\,e \qquad \text{--- (11.5)}$$

Hence

$$V_d = \frac{I}{nAe} = \frac{J}{ne} \qquad \text{--- (11.6)}$$

where $J = I/A$ is current density. J is uniform over the cross sectional area A of the wire. Its unit is A/m^2

Here, $J = \frac{I}{A}$ --- (11.7)

From Eq. (11.6),

$$\vec{J} = (ne)\vec{V}_d \qquad \text{--- (11.8)}$$

For electrons, ne is negative and $\vec{J}$ and $\vec{V}_d$ have opposite directions, $\vec{V}_d$ is the drift velocity.

Example 11.1: A metallic wire of diameter 0.02m contains 10^{28} free electrons per cubic meter. Find the drift velocity for free electrons, having an electric current of 100 amperes flowing through the wire.

(Given : charge on electron = 1.6×10^{-19}C)

Solution: Given

$e = 1.6 \times 10^{-19}$ C

$n = 10^{28}$ electrons/m^3

$D = 0.02$m $\quad r = D/2 = 0.01$m

$I = 100$ A

$$V_d = \frac{J}{ne} = \frac{I}{nAe}$$

where A is the cross sectional area of the wire.

$A = \pi r^2 = 3.142 \times (0.01)^2$

$= 3.142 \times 10^{-4}$m^2

$$V_d = \frac{100}{3.142 \times 10^{-4} \times 10^{28} 1.6 \times 10^{-19}}$$

$$= \frac{10^{2+4-9}}{5.027}$$

$V_d = 10^{-3} \times 0.1989 = 1.9 \times 10^{-4}$ m/s

Example 11.2: A copper wire of radius 0.6 mm carries a current of 1A. Assuming the current to be uniformly distributed over a cross sectional area, find the magnitude of current density.

Solution: Given

$r = 0.6 \text{ mm} = 0.6 \times 10^{-3} \text{ m}$

$I = 1\text{A}$

$J = ?$

Area of copper wire $= \pi r^2$

$= 3.142 \times (0.6)^2 \times 10^{-6}$

$= 3.142 \times 0.36 \times 10^{-6}$

$= 1.1311 \times 10^{-6} \text{ m}^2$

$$J = \frac{I}{A} = \frac{1}{1.1311 \times 10^{-6}}$$

$J = 0.884 \times 10^6 \text{ A/m}^2$

11.5 Ohm's law :

The relationship between the current through a conductor and applied potential difference was first discovered by German scientist George Simon Ohm in 1828 AD. This relationship is known as Ohm's law.

It states that "The current I through a conductor is directly proportional to the potential difference V applied across its two ends provided the physical state of the conductor is unchanged".

The graph of current versus potential difference across the conductor is a straight line as shown in Fig. 11.4

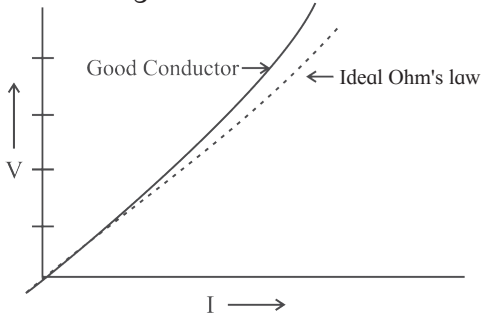


Fig. 11.4: I-V curve for a conductor.

In general, $I \propto V$

or $V = IR$ or $R = \frac{V}{I}$, --- (11.9)

where R is a proportionality constant and is called the resistance of the conductor. The unit of resistance is ohm (Ω),

$$1\Omega = \frac{1\text{volt}}{1\text{ampere}}$$

If potential difference of 1volt across a conductor produces a current of 1ampere through it, then the resistance of the conductor is 1Ω.

Reciprocal of resistance is called conductance.

$$C = \frac{1}{R} \qquad \text{--- (11.10)}$$

The unit of conductance is siemens or $(\Omega)^{-1}$

Example 11.3: A Flashlight uses two 1.5V batteries to provide a steady current of 0.5A in the filament. Determine the resistance of the glowing filament.

Solution:

$$R = \frac{V}{I} = \frac{3}{0.5} = 6.0\Omega$$

∴ Resistance of the glowing filament is 6.0 Ω.

Physical origin of Ohm's law :

We know that electrical conduction in a conductor is due to mobile charge carriers, the electrons. It is assumed that these conduction electrons are free to move inside the volume of the conductor. During their random motion, electrons collide with the ion cores within the conductor. It is assumed that electrons do not collide with each other. These random motions average to zero. On the application of an electric field $\vec{E}$, the motion of the electrons is a combination of the random motion of electrons due to collisions and that due to the electric field $\vec{E}$. The electrons drift under the action of the field $\vec{E}$ and move in a direction opposite to the direction of the field $\vec{E}$.

Consider an electron of mass m subjected to an electric field $\vec{E}$. The force experienced by the electron will be $\vec{F} = e\vec{E}$. The acceleration experienced by the electron will then be

$$\vec{a} = \frac{e\vec{E}}{m} \qquad \text{--- (11.11)}$$

The type of collision the conduction electrons undergo is such that the drift velocity attained before the collision has nothing to do with the drift velocity after the collision. After the collision, the electron will move in random direction, but will still drift in the direction opposite to $\vec{E}$.

Let τ be the average time between two successive collisions. Thus on an average, the

electrons will acquire a drift speed $V_d = a\tau$, where a is the acceleration given by Eq (11.11). Also, at any given instant of time, the average drift speed of the electron will also be $V_d = a\tau$. From Eq. (11.11),

$$V_d = a\tau = \frac{eE\tau}{m} \qquad \text{--- (11.12)}$$

From the Eq. (11.6) and Eq. (11.12),

$$V_d = \frac{J}{ne} = \frac{eE\tau}{m} \qquad \text{--- (11.13)}$$

which gives

$$E = \left(\frac{m}{e^2 n\tau}\right) J \qquad \text{--- (11.14)}$$

or, $E = \rho J$, where ρ is the resistivity of the material and

$$\rho = \frac{m}{ne^2\tau} \qquad \text{--- (11.15)}$$

For a given material, m, n, e^2 and τ will be constant and ρ will also be constant, ρ is independent of $\vec{E}$, the externally applied electric field.

11.6 Limitations of the Ohm's law:

Ohm's law is obeyed by various materials and devices. The devices for which potential difference (V) versus current (I) curve is a straight line passing through origin, inclined to V-axis, are called linear devices or ohmic devices (Fig. 11.4). Resistance of these devices is constant. Several conductors obey the Ohms law. They follow the linear I-V characteristic.

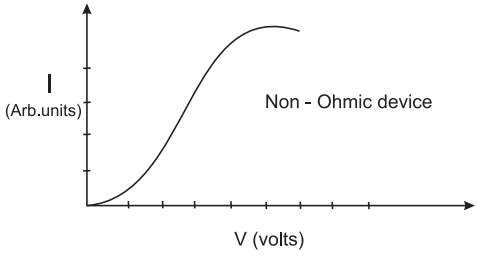


Fig. 11.5: I-V curve for non-Ohmic devices.

The devices for which the I-V curve is not a straight line as shown in Fig. 11.5 are called non-ohmic devices. They do not obey the Ohm's law and the resistance of these devices is a function of V or I; e.g. liquid electrolytes, vacuum tubes, junction diodes, thermistors etc. Resistance R for such non-linear devices at a particular value of the potential difference V is given by,

$$R = \lim_{\Delta I \to 0} \frac{\Delta V}{\Delta I} = \frac{dV}{dI} \qquad \text{--- (11.16)}$$

where ΔV is the potential difference between the two values of potential

$$V - \frac{\Delta V}{2} \quad \text{to} \quad V + \frac{\Delta V}{2},$$

and ΔI is the corresponding change in the current.

11.7 Electrical Energy and Power:

Consider a resistor AB connected to a cell in a circuit shown in Fig. 11.6 with current flowing from A to B. The cell maintains a potential difference V between the two terminals of the resistor, higher potential at A and lower at B. Let Q be the charge flowing in time Δt through the resistor from A to B. The potential difference V between the two points A and B, is equal to the amount of work W, done to carry a unit positive charge from A to B. It is given by

$$V = \frac{W}{Q}, \qquad W = VQ \qquad \text{--- (11.17)}$$

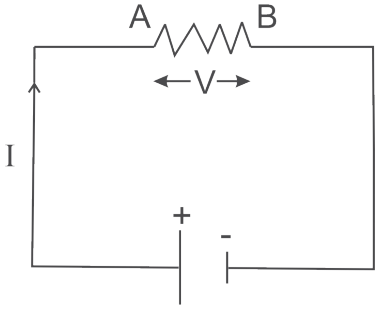


Fig. 11.6: A simple circuit with a cell and a resistor.

The cell provides this energy through the charge Q, to the resistor AB where the work is performed. When the charge Q flows from the higher potential point A to the lower potential point B, i.e. through a decrease in potential of value V, its potential energy decreases by an amount

$$\Delta U = QV = I\,\Delta t V \qquad \text{--- (11.18)}$$

where I is current due to the charge Q flowing in time Δt. Where will this energy go? By the principle of conservation of energy, it is

converted into some other form of energy.

In the limit as $\Delta t \longrightarrow 0$,

$$\frac{dU}{dt} = I.V \qquad \text{--- (11.19)}$$

Here, $\frac{dU}{dt}$ is power, the time rate of transfer of energy and is given by,

$$P = \frac{dU}{dt} = I.V \qquad \text{--- (11.20)}$$

We can also say that this power is transferred by the cell to the resistor or any other device in place of the resistor, such as a motor, a rechargeable battery etc.

Because of the presence of an electric field, the free electrons move across a resistor and there would be an increase in their kinetic energy as they move. When the electrons collide with the ion cores the energy gained by them is shared among the ion cores. Consequently, vibrations of the ions increase, resulting in heating up of the resistor. Thus, some amount of energy is dissipated in the form of heat in a resistor. The energy dissipated in time interval Δt is given by Eq. (11.18). The energy dissipated per unit time is actually the power dissipated and is given by Eq. (11.20).

Using Eq. (11.20), and using Ohm's law, $V=IR$,

$$\therefore P = \frac{V^2}{R} = I^2R \qquad \text{--- (11.21)}$$

It is the power dissipation across a resistor which is responsible for heating it up. For example, the filament of an electric bulb heats up to incandescence, radiating out heat and light.

Example 11.4 : An electric heater takes 6A current from a 230V supply line, calculate the power of the heater and electric energy consumed by it in 5 hours.

Solution : Given

$I = 6A$, $V = 230V$

We know that,

$P = I \times V = (6A)\ (230V) = 1380$ W

$P = 1.38$ kW

Energy consumed = Power × time

= (1.38 kW) × (5 h)

= 6.90kWh (1.0 Kwh = 1 unit of power)

= 6.9 units of electrical energy.

11.8 Resistors:

Resistors are used to limit the current following through a particular path of a circuit. Commercially available resistors are mainly of two types :

Carbon resistors and Wire wound resistors. High value resistors are mostly carbon resistors. They are small and inexpensive. The values of these resistors are colour coded to mark their values in ohms. The colour coding is standardized by Electronic Industries Association (EIA). One such resistor is shown in Fig. 11.7.

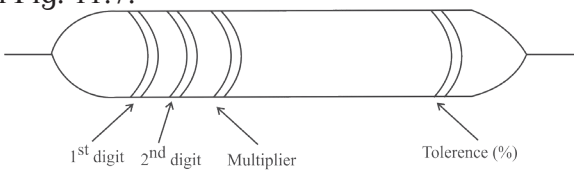


Fig. 11.7: Carbon composition resistor.

Colour code:

Colours	1st digit	2nd digit	Multiplier	Tolerance
Black	0	0	$\times 10^0$	
Brown	1	1	$\times 10^1$	±1%
Red	2	2	$\times 10^2$	±2%
Orange	3	3	$\times 10^3$	
Yellow	4	4	$\times 10^4$	
Green	5	5	$\times 10^5$	
Blue	6	6	$\times 10^6$	
Violet	7	7	$\times 10^7$	
Gray	8	8	$\times 10^8$	
White	9	9	$\times 10^9$	
For Gold			$\times 10^{-1}$	±5%
For Silver			$\times 10^{-2}$	±10%
No colour			-	±20%

Easy Bytes:

Finding it difficult to memorize the colour code sequence? No need to worry, we have a one liner which will help you out "B. B. Roy in Great Britain has Very Good Wife"

B B R O Y G B V G W

This funny one liner makes it easy to recall the sequence of digits and multipliers.

In the four band resistor colour code illustrated in the above table, the first three bands (closest together) indicate the value in ohms. The first two bands indicate two numbers and third band often called decimal multiplier. The fourth band separated by a space from the three value bands, (so that you know which end to start reading from), indicates tolerance of the resistor.

Example

i. Colour code of resistor is

	Yellow	Violet	Orange	Gold
Value :	4	7	10^3	±5%

i.e. $47\times10^3 = 47000\Omega = 47k\Omega\pm5\%$

The value of the resistor is 47kΩ ±5%

ii. From given values of resistor; find the colour bands of this resistor

$330\Omega = 33\times10$

3	3	10^1	
Orange	Orange	Brown	tolerance band

11.8.1 Rheostat:

A rheostat shown in Fig. 11.8 is an adjustable resistor used in applications that require adjustment of current or resistance in an electric circuit. The rheostat can be used to adjust potential difference between two points in a circuit, change the intensity of lights and control the speed of motors, etc. Its resistive element can be a metal wire or a ribbon, carbon films or a conducting liquid, depending upon the application. In hi-fi equipment, rheostats are used for volume control.

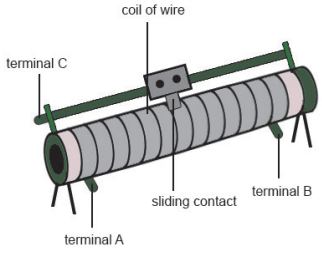


Fig . 11.8: Rheostat.

11.8.2 Combination of Resistors:

I. Series combination of Resistors:

In series combination of resistors, these are connected in single electrical path as shown in Fig 11.9. Hence the same electric current flows through each resistor in a series combination. Because of series combination, the supply voltage between two resistors R_1 and R_2 is V_1 and V_2, respectively and the same current I flows through the resistor R_1 and the resistor R_2. i.e. in series combination, supply voltage is divided and the current remains the same in all the resistors.

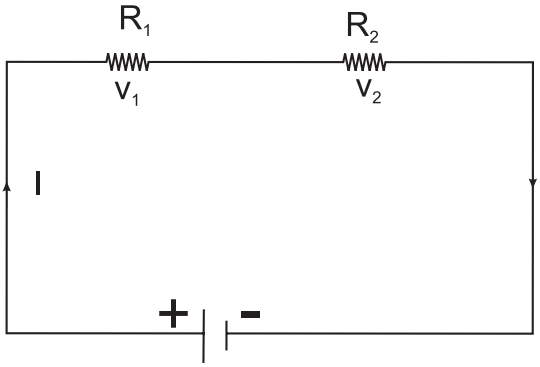


Fig. 11.9: Series combination of two resistors R_1 and R_2.

According to Ohm's law,

$$R_1 = \frac{V_1}{I}, \quad R_2 = \frac{V_2}{I} \qquad \text{--- (11.22)}$$

Total voltage $V = V_1 + V_2$ --- (11.23)

From equation.... (11.22) and (11.23)

we write

$$V = I(R_1 + R_2) \qquad \text{--- (11.24)}$$

$$\therefore V = I R_s \qquad \text{--- (11.25)}$$

Thus the equivalent resistance of the series circuit $R_s = R_1 + R_2$

When a number of resistors are connected in series, the equivalent resistance is equal to the sum of individual resistances.

For n number of resistors,

$$R_s = R_1 + R_2 + R_2 + \ldots\ldots\ldots + R_n = \sum_{i=1}^{i=n} R_i \quad \text{-- (11.26)}$$

II. Parallel Combination of Resistors:

In the parallel combination, the resistors are connected in such a way that the same voltage is applied across each resistor.

A number of resistors are said to be connected in parallel if all of them are connected between the same two electrical points each having individual path as shown in Fig. 11.10.

In parallel combination the total current I is divided into I_1 and I_2 as shown in the circuit

diagram Fig.11.10, whereas voltage V across them remains the same,

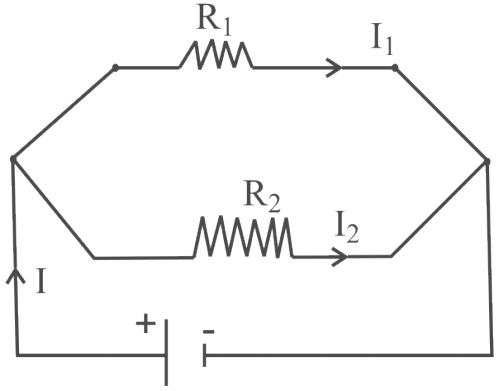


Fig. 11.10 : Two resistors in parallel combination.

$$I = I_1 + I_2 \qquad \text{-- (11.27)}$$

where I_1 is current flowing through R_1 and I_2 is current flowing through R_2.

When Ohm's law is applied to R_1

$$V = I_1 R_1 \quad \text{i.e. } I_1 = \frac{V}{R_1} \qquad \text{--- (11.28a)}$$

Ohm's law applied to R_2

$$V = I_2 R_2 \quad \text{i.e. } I_2 = \frac{V}{R_2} \qquad \text{--- (11.28b)}$$

From Eq. (11.27) and Eq. (11.28),

$$\therefore \quad I = \frac{V}{R_1} + \frac{V}{R_2},$$

If, $$I = \frac{V}{R_p},$$

$$\frac{V}{R_p} = \frac{V}{R_1} + \frac{V}{R_2},$$

$$\therefore \quad \frac{1}{R_p} = \frac{1}{R_1} + \frac{1}{R_2}, \qquad \text{--- (11.29)}$$

where R_p is the equivalent resistance in parallel combination.

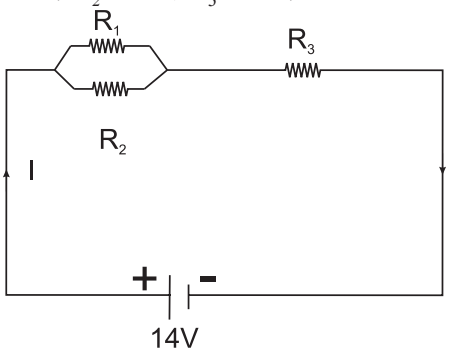
If n resistors R_1, R_2, R_3....., R_n are connected in parallel, the equivalent resistance of the combination is given by

$$\frac{1}{R_p} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots\ldots\ldots\ldots \frac{1}{R_n} = \sum_{i=1}^{n} \frac{1}{R} \qquad \text{--(11.30)}$$

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

Example 11.5: Calculate i) total resistance and ii) total current in the following circuit.

$R_1 = 3\Omega$, $R_2 = 6\Omega$, $R_3 = 5\Omega$, $V = 14V$



Circuit diagram

Solution:

i) Total resistance = $R_T = R_P + R_3$

$$R_P = \frac{R_1 R_2}{R_1 + R_2} = \frac{3 \times 6}{9} = 2\Omega$$

$R_T = 2 + 5 = 7\Omega$

Total Resistance = 7Ω

ii) Total current :

$$I = \frac{V}{R_T} = \frac{14V}{7\Omega}$$

$I = 2A$

11.9 Specific Resistance (Resistivity):

At a particular temperature, the resistance of a given conductor is observed to depend on the nature of material of conductor, the area of its cross-section, and its length.

It is found that resistance R of a conductor of uniform cross section is

i. directly proportional to its length l,
i.e. $R \propto l$

ii. inversely proportional to its area of cross section A,

i.e. $R \propto \frac{l}{A}$

From **i** and **ii**

$$R = \rho \frac{l}{A} \qquad \text{--- (11.31)}$$

where ρ is a constant of proportionality and it is called specific resistance or resistivity of the material of the conductor at a given temperature.

From Eq. (11.31), we write

$$\rho = \frac{RA}{l} \qquad \text{--- (11.32)}$$

SI unit of resistivity is ohm-meter. Resistivity of a conductor is numerically the resistance per unit length, and per unit area of cross-section of material of the conductor.

i.e. when, $R = 1\Omega$, $A = 1m^2$ and $l = 1m$,

then, $\rho = 1\Omega m$

Conductivity : Reciprocal of resistivity is called conductivity of a material, $\sigma = (\rho)^{-1}$.

SI unit of σ is: $(\Omega m)^{-1}$ i.e. siemens/meter (Sm^{-1})

Table 11.1 : Resistivity of various materials

Material	Resistivity ρ (Ω.m)	Material	Resistivity ρ (Ω.m)
Conductors		**Semiconductors**	
Silver	1.59×10^{-8}	Carbon	3.5×10^{-5}
Copper	1.72×10^{-8}	Germanium	0.5
Gold	2.44×10^{-8}	Silicon	3×10^{4}
Aluminium	2.82×10^{-8}	**Insulators**	
Tungsten	5.6×10^{-8}	Glass	10^{11}-10^{13}
Iron	9.7×10^{-8}	Mica	10^{11}-10^{15}
Mercury	95.8×10^{-8}	Rubber (hard)	10^{13}-10^{16}
Nichrome (alloy)	100×10^{-8}	Teflon	10^{16}
		Wood (maple)	3×10^{8}

Example 11.6: Calculate the resistance per metre, at room temperature, of a constantan (alloy) wire of diameter 1.25mm. The resistivity of constantan at room temperature is 5.0×10^{-7} Ωm.

Solution: $\rho = 5.0 \times 10^{-7}\,\Omega m$

$d = 1.25 \times 10^{-3}$ m

$r = .625 \times 10^{-3}$ m

Cross-sectional Area $= \pi r^2$

$$\text{Resistivity } \rho = \frac{RA}{l}$$

$$\text{Resistance per meter} = \frac{R}{l}$$

i.e. $$\frac{R}{l} = \frac{\rho}{A} = \frac{5\times10^{-7}}{(0.625\times10^{-3})^2 \times 3.142}$$

$$\frac{R}{l} = 0.41\,\Omega m^{-1}$$

$\therefore$ Resistance per metre= 0.41 Ωm^{-1}

Resistivity ρ is a property of a material, while the resistance R refers to a particular object. Similarly, the electric field $\vec{E}$ at a point is specified in a material with the potential difference across the resistance, and the current density $\vec{J}$ in a material instead of the current I in the resistor. Then for an isotropic material,

$$\rho = \frac{E}{J} \quad \text{or} \quad \vec{E} = \rho\vec{J} \qquad \text{---- (11.33)}$$

Again, the SI unit of ρ is

$$\frac{\text{unit}(E)}{\text{unit}(J)} = \frac{\text{V/m}}{\text{A/m}^2} = \frac{\text{V}}{\text{A}}\,\text{m} = \Omega.\text{m}$$

In terms of conductivity σ of a material, from (11.33),

$$\vec{J} = \frac{1}{\rho}\vec{E} = \sigma\vec{E} \qquad \text{--- (11.34)}$$

For a particular resistor, we had (Eq. 11.9) the resistance R given by

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

Compare this with the above Eq (11.33).

11.10 Variation of Resistance with Temperature:

Resistivity of a material varies with temperature. It is a property of material. Fig. 11.11 shows the variation of resistivity of copper as a function of temperature (K). It can be seen that the variation is linear over a certain range of temperatures. Such a linear relation can be expressed as,

$$\rho = \rho_0[1 + \alpha(T - T_0)], \qquad \text{--- (11.35)}$$

where T_0 is the chosen reference temperature and ρ_0 in the resistivity at the chosen temperature, for example, T_0 can be 0 °C.

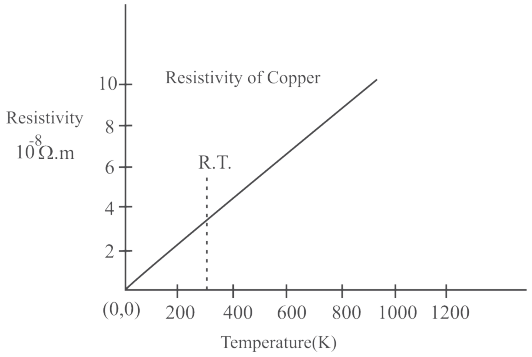


Fig. 11.11: Resistivity as a function of temperature (K).

In the above Eq. (11.35),

$$\alpha = \frac{\rho - \rho_0}{\rho_0 (T - T_0)} = \frac{R - R_0}{R_0 (T - T_0)} \quad \text{--- (11.36)}$$

Here, α is called the temperature coefficient of resistivity. Table (11.1) shows the resistivity of some of the metals. The temperature coefficient of resistance is defined as the increase in resistance per unit original resistance at the chosen reference temperature, per degree rise in temperature. The unit of α is $^oC^{-1}$ or $^oK^{-1}$ (per degree celcius or per degree kelvin).

From Eq. (11.36)

$$R = R_0 [1 + \alpha (T - T_0)] \quad \text{--- (11.37)}$$

For small difference in temperatures,

$$\alpha = \frac{1}{R_0} \cdot \frac{dR}{dT} \quad \text{--- (11.38)}$$

Do you know ?

Here, the temperature difference is more important than the temperature alone. Therefore, as the sizes of degrees on the Celsius scale and the Absolute scale are identical, any scale can be used.

Example 11.7: A piece of platinum wire has resistance of 2.5 Ω at 0° C. If its temperature coefficient of resistance is $4 \times 10^{-3}/^oC$. Find the resistance of the wire at 80° C.

Solution:

$R_0 = 2.5\ \Omega$

$\alpha = 0.004/^oC$

$T - 0 = T = 80^o C$

$R_T = R_0(1 + \alpha T)$

$R_T = 2.5\ (1 + 0.004 \times 80) = 2.5\ (1 + 0.32)$

$R_T = 2.5 \times 1.32$

$R_T = 3.3\Omega$

Superconductivity :

We know that the resistivity of a metal decreases as the temperature decreases. In case of some metals and metal alloys, the resistivity suddenly drops to zero at a particular temperature (T_c). This temperature is called critical temperature, for example, mercury loses its resistance completely to zero at 4.2K.

Superconductivity can be harnessed so as to be useful for mankind. It is already in use in obtaining very high magnetic field (a few Tesla) in superconducting magnet. These magnets are used in research quality NMR spectrometers. For its operation, the current carrying coils are required to be kept at a temperature less than the critical temperature of the coil material.

11.11 Electromotive Force (emf):

When charges flow through a conductor, a potential difference has to be established between the two ends of the conductor. For a steady flow of charges, this potential difference is required to be maintained across the two ends of the conductor, the terminals. There is a device that does so by doing work on the charges, thereby maintaining the potential difference. Such a device is called an emf device and it provides the emf ε. The charges move in the conductor owing to the energy provided by the emf device. The device supplies this energy through the work it does.

You must have used some of these emf devices. Power cells, batteries,Solar cells, fuel cells, and even generators, are some examples of emf devices familiar to you.

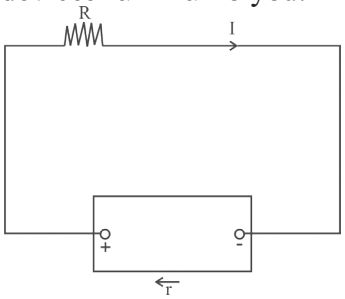


Fig. 11.12: Circuit with emf device.

Fig. 11.12 shows a circuit with an emf device and a resistor R. Here, the emf device keeps the positive terminal (+) at a higher electric potential than the negative terminal (-).

The emf is represented by an arrow from the negative terminal to the positive terminal of a device such as a Voltaic cell. When the circuit is open, there is no net flow of charge carriers within the device. When connected in a circuit, there is a flow of carriers from one terminal to the other terminal inside the emf device. The positive charge carriers move towards the positive terminal which acts as cathode inside the emf device. Thus the positive charge carriers move from the region of lower potential energy, to the region of higher potential energy which is cathode inside the emf device. Here, the energy source is chemical in nature. In a Solar cell, it is the photon energy in the Solar radiation.

Now suppose that a charge dq flows through the cross section of the circuit (Fig. 11.12), in time dt.

It is clear that the same amount of charge dq flows throughout the circuit, including the emf device. It enters the negative terminal (low potential terminal) and leaves the positive terminal (higher potential terminal). Hence, the device must do work dw on the charge dq, so that it moves in the above manner. Thus we define the emf of the emf device.

$$\varepsilon = \frac{dw}{dq} \qquad \text{--- (11.39)}$$

The SI unit of emf is joule/coulomb (J/C).

In an ideal device, there is no internal resistance to the motion of charge carriers. The emf of the device is then equal to the potential difference across the two terminals of the device. In a real emf device, there is an internal resistance to the motion of charge carriers. If such a device is not connected in a circuit, there is no current through it. In that case the emf is equal to the potential difference across the two terminals of the emf device connected in a circuit, there is no current through it. If a current (I) flows through an emf device, there is an internal resistance (r) and the emf (ε) differs from the potential difference across its two terminals (V).

$$V = \varepsilon - (I)\,(r) \qquad \text{--- (11.40)}$$

The negative sign is due to the fact that the current I flows through the emf device from the negative terminal to the positive terminal.

By the application of Ohm's law Eq. (11.9),

$V = IR$

Hence $IR = \varepsilon - Ir$ --- (11.41)

Or

$$I = \frac{\varepsilon}{R + r} \qquad \text{--- (11.42)}$$

Thus, the maximum current that can be drawn from the emf device is when $R = 0$, i.e.

$$I_{max} = \frac{\varepsilon}{r} \qquad \text{--- (11.43)}$$

This is the maximum allowed current from an emf device (or a cell). This decides the maximum current rating of a cell or a battery.

11.12 Cells in Series:

In a series combination, cells are connected in single electrical path, such that the positive terminal of one cell is connected to the negative terminal of the next cell, and so on. The terminal voltage of battery/cell is equal to the sum of voltages of individual cells in series, as shown in Fig 11.13 a.

Figure shows two 1.5V cells in series. This combination provides total voltage of 3.0V (1.5×2).

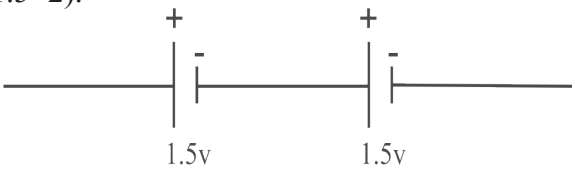


Fig. 11.13 (a): Cells in parallel.

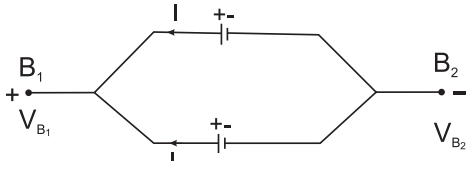


Fig. 11.13 (b): Cells in parallel.

The equivalent emf of n number of cells in series combination is the algebraic sum of their individual emf. The equivalent internal resistance of n cells in a series combination is the sum of their individual internal resistance.

$$V = \sum_i \varepsilon_i - I.\sum_i r_i \qquad \text{--- (11.44)}$$

- **Advantages of cells in series.**

(i) The cells connected in series produce a larger resultant voltage.

(ii) Cells which are damaged can be easily identified, hence can be easily replaced.

11.13 Cells in parallel:

Consider two cells which are connected in parallel. Here, positive terminals of all the cells are connected together and the negative terminals of all the cells are connected together. In parallel connection, the current is divided among the branches i.e. I_1 and I_2 as shown in Fig. 11.13b. Consider points B_1 and B_2 having potentials V_{B_1} and V_{B_2}, respectively.

For the first cell the potential difference across its terminals is,

$$V = V_{B_1} - V_{B_2} = \varepsilon_1 - I_1 r_1 \qquad \text{--- (11.45)}$$

$$\therefore \quad I_1 = \frac{\varepsilon_1 - V}{r_1} \qquad \text{--- (11.46)}$$

Point B_1 and B_2 are connected exactly similarly to the second cell.

Hence, considering the second cell we write,

$$V = V_{B_1} - V_{B_2} = \varepsilon_2 - I_2 r_2 \; ; \; I_2 = \frac{\varepsilon_2 - V}{r_2} \qquad \text{--- (11.47)}$$

We know that $I = I_1 + I_2$

Combining the last three equations,

$$\therefore \quad I = \frac{\varepsilon_1}{r_1} - \frac{V}{r_1} + \frac{\varepsilon_2}{r_2} - \frac{V}{r_2} \qquad \text{--- (11.48)}$$

$$= \left(\frac{\varepsilon_1}{r_1} + \frac{\varepsilon_2}{r_2}\right) - V\left(\frac{1}{r_1} + \frac{1}{r_2}\right)$$

$$\text{Thus,} \quad V\left(\frac{1}{r_1} + \frac{1}{r_2}\right) = \left(\frac{\varepsilon_1}{r_1} + \frac{\varepsilon_2}{r_2}\right) - I$$

$$\therefore \quad V\left(\frac{r_1 + r_2}{r_1 r_2}\right) = \frac{\varepsilon_1 r_2 + \varepsilon_2 r_1}{r_1 r_2} - I$$

$$V = \frac{\varepsilon_1 r_2 + \varepsilon_2 r_1}{r_1 + r_2} - I\frac{r_1 r_2}{r_1 + r_2} \qquad \text{---(11.49)}$$

If we replace the cells by a single cell connected between points B_1 and B_2 with the emf ε_{eq} and the internal resistance r_{eq} as in Fig. (11.13b),

then,

$$V = \varepsilon_{eq} - I r_{eq} \qquad \text{--- (11.50)}$$

Considering Eq. (11.49) and Eq. (11.50) we can write,

$$\varepsilon_{eq} = \frac{\varepsilon_1 r_2 + \varepsilon_2 r_1}{r_1 + r_2}$$

$$r_{eq} = \frac{r_1 r_2}{r_1 + r_2}$$

$$\text{i.e.} \quad \frac{1}{r_{eq}} = \frac{1}{r_1} + \frac{1}{r_2}$$

$$\frac{\varepsilon_{eq}}{r_{eq}} = \frac{\varepsilon_1}{r_1} + \frac{\varepsilon_2}{r_2} \qquad \text{--- (11.51)}$$

For n number of cells connected in parallel with emf $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots\ldots, \varepsilon_n$ and internal resistance $r_1, r_2, r_3, \ldots\ldots, r_n$

$$\frac{1}{r_{eq}} = \frac{1}{r_1} + \frac{1}{r_2} + \frac{1}{r_3} + \ldots\ldots\ldots + \frac{1}{r_n} \qquad \text{--- (11.52)}$$

and

$$\frac{\varepsilon_{eq}}{r_{eq}} = \frac{\varepsilon_1}{r_1} + \frac{\varepsilon_2}{r_2} + \ldots\ldots\ldots\ldots + \frac{\varepsilon_n}{r_n}$$

--- (11.53)

Substitution of emfs should be done algebraically by considering proper ± signs according to polarity.

- **Advantages of cells in parallel :** For cells connected in parallel in a circuit, the circuit will not break open even if a cell gets damaged or open.
- **Disadvantages of cells in parallel :** The voltage developed by the cells in parallel connection cannot be increased by increasing number of cells present in circuit.

11.14 Types of Cells:

Electrical cells can be divided into several categories like primary cell, secondary cell, fuel cell, etc.

A primary cell cannot be charged again. It can be used only once. Dry cells, alkaline cells are different examples of primary cells. Primary cells are low cost and can be used easily. But these are not suitable for heavy loads. Secondary cells are used for such applications. The secondary cell are rechargeable and can be reused. The chemical reaction in a secondary cells is reversible. Lead acid cell, and fuel cell

are some examples of secondary cells. Lead acid battery is used widely in vehicles and other applications which require high load currents. Solar cells are secondary cells that convert Solar energy into electrical energy.

Fuel cells vehicles (FCVs) are electric vehicles that use fuel cells instead of lead acid batteries to power the vehicles. Hydrogen is used as a fuel in fuel cells. The by- product after its burning is water. This is important in terms of reducing emission of greenhouse gases produced by traditional gasoline fueled vehicles. The hydrogen fuel cell vehicles are thus more environment friendly.

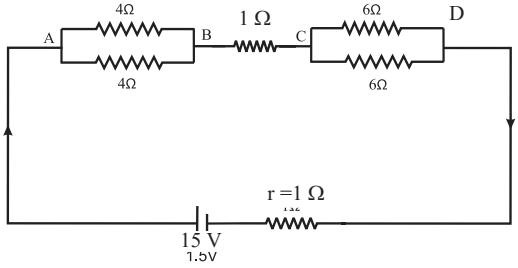
Example 11.8: A network of resistors is connected to a 15 V battery with internal resistance 1 Ω as shown in the circuit diagram.

Calculate

(i) The equivalent resistance,

(ii) Current in each resistor,

(iii) Voltage drops V_{AB}, V_{BC} and V_{DC}.



Solution :

i) Equivalent Resistance $(R_{eq}) = R_{AB} + R_{BC} + R_{DC}$

$$R_{AB} = \frac{4\times4}{4+4} = 2\Omega, \qquad R_{CD} = \frac{6\times6}{6+6} = 3\Omega$$

$$R_{BC} = 1\Omega$$

$$R_T = R_{eq} = 2 + 1 + 3 = 6\ \Omega$$

$\therefore$ Equivalent Resistance is 6 Ω

ii. Current in each resistor :

Total current I in the circuit is,

$$I = \frac{\varepsilon}{R_T + r} = \frac{15}{6+1} = 2.1\text{A}$$

Consider resistors between A and B.

Let I_1 be the current through one of the 4Ω resistors and I_2 be the current in the other resistor

$I_1 \times 4 = I_2 \times 4$

that is, $I_1 = I_2$ from symmetry of the two arms.

But $I_1 + I_2 = I = 2.1\text{A}$

$\therefore \quad I_1 = I_2 = 1.05\text{A}$

that is, the current in each 4Ω resistor is 1.05A, the current in 1Ω resistor between B and C would be 2.1A.

Now, consider the resistances between C and D

Let I_3 be the current through one of the 6 Ω resistors and I_4 be the current in the other resistor.

$$I_3 \times 6 = I_4 \times 6$$

$\therefore \quad I_3 = I_4 = 1.05\text{A}$

That is, current in each 6 Ω resistor is 1.05A

iii. Voltage drop across BC is V_{BC}

$$V_{BC} = I \times 1 = 2.1 \times 2.1 = 2\ \text{V}$$

Voltage drop across CD is V_{CD}

$$V_{CD} = I \times R_{CD} = 2.1 \times 3 = 6.3\text{V}$$

[Note : Total voltage drop across AD is (4.2 V+2.1V+6.3 V) =12.6 V, while its emf is 15 V. The loss of the voltage is 2.4 V].

Internet my friend

https://www.britannica.com/science/superconductivityphysics

Exercises

1. Choose correct alternative

i) You are given four bulbs of 25 W, 40 W, 60 W and 100 W of power, all operating at 230 V. Which of them has the lowest resistance?

(A) 25 W (C) 40 W

(C) 60 W (D) 100 W

ii) Which of the following is an ohmic conductor?

(A) transistor (B)vacuum tube

(C) electrolyte (D) nichrome wire

iii) A rheostat is used

(A) to bring on a known change of resistance in the circuit to alter the current

(B) to continuously change the resistance in any arbitrary manner and there by alter the current

(C) to make and break the circuit at any instant

(D) neither to alter the resistance nor the current

iv) The wire of length L and resistance R is stretched so that its radius of cross-section is halved. What is its new resistance?

(A) 5R (B) 8R

(C)4R (D) 16R

v) Masses of three pieces of wires made of the same metal are in the ratio 1:3:5 and their lengths are in the ratio 5:3:1. The ratios of their resistances are

(A) 1:3:5 (B) 5:3:1

(C) 1:15:125 (D) 125:15:1

vi) The internal resistance of a cell of emf 2V is 0.1Ω it is connected to a resistance of 0.9Ω. The voltage across the cell will be

(A) 0.5 V (B) 1.8 V

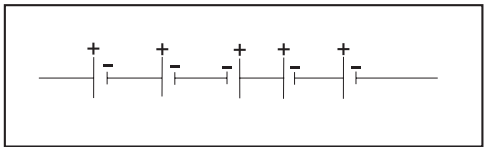
(C) 1.95 V (D) 3V

vii) 100 cells each of emf 5V and internal resistance 1Ω are to be arranged so as to produce maximum current in a 25Ω resistance. Each row contains equal number of cells. The number of rows should be

(A) 2 (B) 4

(C) 5 (D) 100

viii) Five dry cells each of voltage 1.5 V are connected as shown in diagram



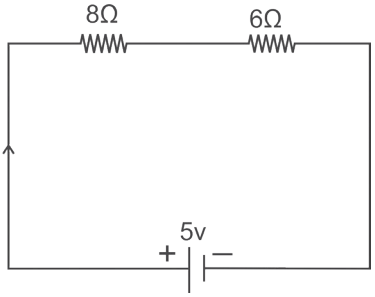
What is the overall voltage with this arrangement?

(A) 0V (B) 4.5V

(C) 6.0V (D) 7.5V

2. Give reasons / short answers

i) In given circuit diagram two resistors are connected to a 5V supply.



a] Calculate potential difference across the 8Ω resistor.

b] A third resistor is now connected in parallel with 6Ω resistor. Will the potential difference across the 8Ω resistor the larger, smaller or the same as before? Explain the reason for your answer.

ii) Prove that the current density of a metallic conductor is directly proportional to the drift speed of electrons.

3. Answer the following questions.

i) Distinguish between Ohmic and non-ohmic substances; explain with the help of example.

ii) DC current flows in a metal piece of non-uniform cross-section. Which of these quantities remains constant along the conductor: current, current density or drift speed?

4. Solve the following problems.

i) What is the resistance of one of the rails of a railway track 20 km long at 20° C? The cross section area of rail is 25 cm^2 and the rail is made of steel having resistivity at 20° C as 6×10^{-8} Ω m.

[Ans: 0.48 Ω]

ii) A battery after a long use has an emf 24 V and an internal resistance 380 Ω. Calculate the maximum current drawn from the battery? Can this battery drive starting motor of car?

[Ans: 0.063 A]

iii) A battery of emf 12 V and internal resistance 3 Ω is connected to a resistor. If the current in the circuit is 0.5 A,

a] Calculate resistance of resistor.

b] Calculate terminal voltage of the battery when the circuit is closed.

[Ans: a) 21 Ω, b) 10.5 V]

iv) The magnitude of current density in a copper wire is 500 A/cm^2. If the number of free electrons per cm^3 of copper is 8.47×10^{22} calculate the drift velocity of the electrons through the copper wire (charge on an e = 1.6×10^{-19} C)

[Ans: 3.69×10^{-4} m/s]

v) Three resistors 10 Ω, 20 Ω and 30 Ω are connected in series combination.

i] Find equivalent resistance of series combination.

ii] When this series combination is connected to 12V supply, by neglecting the value of internal resistance, obtain potential difference across each resistor.

[Ans: i) 60 Ω, ii) 2 V, 4 V, 6 V]

vi) Two resistors 1k Ω and 2k Ω are connected in parallel combination.

i] Find equivalent resistance of parallel combination

ii] When this parallel combination is connected to 9 V supply, by neglecting internal resistance calculate current through each resistor.

[Ans: i) 0.66 kΩ, ii) 9 mA, 4.5 mA]

vii) A silver wire has a resistance of 4.2 Ω at 27° C and resistance 5.4 Ω at 100° C. Determine the temperature coefficient of resistance.

[Ans: 3.91×10^{-3}/°C]

viii) A 6m long wire has diameter 0.5 mm. Its resistance is 50 Ω. Find the resistivity and conductivity.

[Ans: 1.636×10^{-6}Ω/m, 6.112×10^{5}m/Ω]

ix) Find the value of resistances for the following colour code.

1. Blue Green Red Gold

[Ans: 6.5 kΩ ± 5%]

2. Brown Black Red Silver

[Ans: 1.0 kΩ ± 10%]

3. Red Red Orange Gold

[Ans: 2.2 kΩ ± 5%]

4. Orange White Red Gold

[Ans: 3.9 kΩ ± 5%]

5. Yellow Violet Brown Silver

[Ans: 4.70 kΩ ± 10%]

x) Find the colour code for the following value of resistor having tolerance ± 10%

a) 330Ω b) 100Ω c) 47kΩ
d) 160Ω e) 1kΩ

xi) A current 4A flows through an automobile headlight. How many electrons flow through the headlight in a time 2hrs.

[Ans : 1.8×10^{23}]

xii) The heating element connected to 230V draws a current of 5A. Determine the amount of heat dissipated in 1 hour (J = 4.2 J/cal.).

[Ans : 985.7 kcal]

***

12. Magnetism

Can you recall?

1. What is a bar magnet?
2. What are the magnetic lines of force?
3. What are the rules concerning the lines of force?
4. If you freely hang a bar magnetic horizontally, in which direction will it become stable?

12.1 Introduction:

The history of magnetism dates back to earlier than 600 B.C., but it is only in the twentieth century that scientists began to understand it and developed technologies based on this understanding. William Gilbert (1544-1603) was the first to systematically investigate the phenomenon of magnetism using scientific method. He also discovered that Earth is a weak magnet. Danish physicist Hans Oersted (1777-1851) suggested a link between electricity and magnetism. James Clerk Maxwell (1831-1879) proved that electricity and magnetism represent different aspects of the same fundamental force field.

In electrostatics you have learnt about the relationship between the electric field and force due to electric charges and electric dipoles. Analogous concepts exist in magnetism except that magnetic poles do not exist in isolation, and we always have a magnetic dipole or a quadrupole. In this Chapter the main focus will be on elementary aspects of magnetism and terrestrial magnetism.

Do you know ?

Some commonly known facts about magnetism.

(i) Every magnet regardless of its size and shape has two poles called north pole and south pole.

(ii) If a magnet is broken into two or more pieces then each piece behaves like an independent magnet with somewhat weaker magnetic field.
Thus isolated magnetic monopoles do not exist. The search for magnetic monopoles is still going on.

(iii) Like magnetic poles repel each other, whereas unlike poles attract each other.

(iv) When a bar magnet/ magnetic needle is suspended freely or is pivoted, it aligns itself in geographically North-South direction.

12.2 Magnetic Lines of Force and Magnetic Field:

You have studied properties of electric lines of force earlier in the Chapter on electrostatics. In a similar manner, magnetic lines of force originate from the north pole and end at the south pole of a bar magnet. The magnetic lines of force of a magnet have the following properties:

i) The magnetic lines of force of a magnet or a solenoid form closed loops. This is in contrast to the case of an electric dipole, where the electric lines of force originate from the positive charge and end on the negative charge, without forming a complete loop (see Fig. 12.4).

ii) The direction of the net magnetic field $\vec{B}$ at a point is given by the tangent to the magnetic line of force at that point in the direction of line of force.

iii) The number of lines of force crossing per unit area decides the magnitude of the magnetic field $\vec{B}$.

iv) The magnetic lines of force do not intersect. This is because had they intersected, the direction of magnetic field would not be unique at that point.

Try this

You can take a bar magnet and a small compass needle. Place the bar magnet at a fixed position on a paper and place the needle at various positions. Noting the orientation of the needle, the magnetic field direction at various locations can be traced.

Density of lines of force i.e., the number of lines of force per unit area normal to the surface

around a particular point determines the strength of the magnetic field at that point. The number of lines of force is called magnetic flux (ϕ). SI unit of magnetic flux (ϕ) is weber (Wb). For a specific case of uniform magnetic field which is normal to the finite area A, the magnitude of magnetic field strength B at a point in the area A is given by

$$\text{Magnetic Field} = \frac{\text{magnetic flux}}{\text{area}}$$

$$\text{i.e.} \quad B = \frac{\phi}{A} \qquad \text{--- (12.1)}$$

SI unit of magnetic field (B) is expressed as weber/m^2 or Tesla.

1 Tesla = 10^4 Gauss.

However, magnetic lines are only a crude way of representing magnetic field. It is a pictorial representation of the strength of the magnetic field (B). It is better defined in terms of Lorentz force law which you will learn in std XII.

12.3 The Bar magnet:

A bar magnet is said to have magnetic pole strength $+q_m$ and $-q_m$ at the north and south poles, respectively. The separation of magnetic poles inside the magnet is $2l$. As the bar magnet has two poles, with equal and opposite pole strength, it is called a magnetic dipole. This is analogous to an electric dipole. The magnetic dipole moment, therefore, becomes $\vec{m} = q_m.2\vec{l}$ ($2\vec{l}$ is a vector from south pole to north pole) in analogy with the electric dipole moment.

SI unit of pole strength (q_m) is A m.

SI unit of magnetic dipole moment m is A m^2.

Axis:- It is the line passing through both the poles of a bar magnet. Obviously, there is only one axis for a given bar magnet.

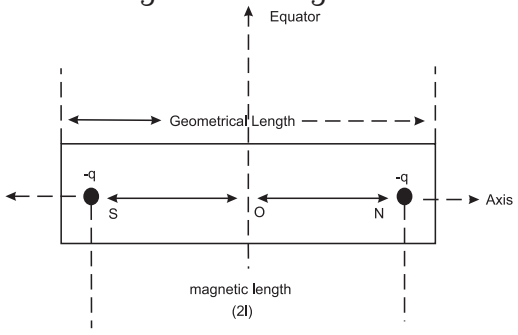


Fig. 12.1: Bar magnet

Equator:- A line passing through the centre of a magnet and perpendicular to its axis is called magnetic equator. The plane containing all equators is called the equatorial plane. The locus of points, on the equatorial plane, which are equidistant from the centre of the magnet is called the equatorial circle. The popularly known 'equator' in Geography is actually an 'equatorial circle'. Such a circle with any diameter is an equator.

Magnetic length ($2l$):- It is the distance between the two poles of a magnet.

$$\text{Magnetic length } (2l) = \frac{5}{6} \times \text{Geometric length} \qquad \text{--- (12.2)}$$

12.3.1 Magnetic field due to a bar magnet at a point along its axis and at a point along its equator:

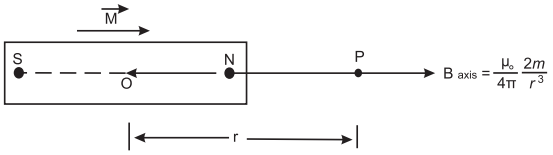


Fig. 12.2 (a): Magnetic field at a point along the axis of the magnet.

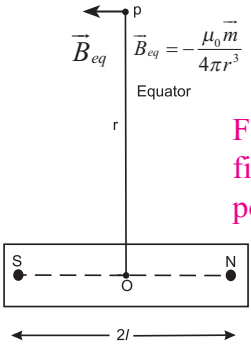


Fig. 12.2 (b): Magnetic field along the equatorial point.

Consider a bar magnet of dipole length $2l$ and magnetic dipole moment $\vec{m}$ as shown in Fig. 12.2 (a). We will now find magnetic field at a point P along the axis of the bar magnet.

Let r be the distance of point P from the centre O of the magnetic dipole.

$OS = ON = l$

$$\therefore \; NP = SP = \sqrt{\left(r^2 + l^2\right)}$$

We now use the electrostatic analogy to obtain the magnetic field due to a bar magnet at a large distance $r >> l$. Consider the electric field due to an electric dipole with a dipole moment p.

The Electrostatic Analogue:

As suggested by Maxwell, electricity and magnetism could be studied analogously. The pole strength (q_m) in magnetism can be compared with charge q in electrostatics. Accordingly, we can write the equivalent physical quantities in electrostatics and magnetism as shown in table 12.1.

Table 12.1: The Electrostatic Analogue

Quantity	Electrostatics	Magnetism
Basic physical quantity	Electrostatic charge	Magnetic pole
Field	Electric Field $\vec{E}$	Magnetic Field $\vec{B}$
Constant	$\frac{1}{4\pi\varepsilon_0}$	$\frac{\mu_0}{4\pi}$
Dipole moment	$\vec{p} = q\,(2\vec{l}\,)$ along (-ve) ⟶ (+ve) charge	$\vec{m} = q_m\,(2\vec{l}\,)$ (bar magnet) along S ⟶ N pole
Force	$\vec{F} = q\vec{E}$	$\vec{F} = q_m\vec{B}$
Energy (In external field) of a dipole	$U = -\,\vec{p}.\vec{E}$	$U = -\vec{m}.\vec{B}$
Coulomb's law	$F = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{q_1 q_2}{r^2}$	No analogous law as magnetic monopoles do not exist
Axial field for a short dipole	$\frac{2p}{4\pi\varepsilon_0 r^3}$ along $\vec{p}$	$\frac{\mu_0 2\vec{m}}{4\pi r^3}$
Equatorial field for a short dipole	$\frac{p}{4\pi\varepsilon_0 r^3}$ opposite to $\vec{p}$	$\frac{-\mu_0\vec{m}}{4\pi r^3}$

You have studied the electric field due to an electric dipole of length $2l$ (p=2ql) at a distance r along the dipolar axis (Eq. 10.24) which is given by,

$$\left|\vec{E}_a\right| = \frac{1}{4\pi\varepsilon_0}\frac{2p}{r^3}, \qquad r >> l$$

The electric field on the equator (Eq. 10.28) is antiparallel to $\vec{p}$ and is given by

$$\left|\vec{E}_{eq}\right| = \frac{1}{4\pi\varepsilon_0}\frac{p}{r^3}, \qquad r >> l$$

Using the analogy given in Table 12.1, we can thus write the axial magnetic field of a bar magnet at a distance r, $r >> l$, $2l$ being the length of bar magnet,

$$\vec{B}_a = +\frac{\mu_0}{4\pi}\frac{2\vec{m}}{r^3} \qquad \text{--- (12.3)}$$

Similarly, the equatorial magnetic field

$$\vec{B}_{eq} = -\frac{\mu_0\vec{m}}{4\pi r^3} \qquad \text{--- (12.4)}$$

Negative sign shows that the direction of $\vec{B}_{eq}$ is opposite to $\vec{m}$.

For the same distance from centre O of a bar magnet,

$$B_{axis} = 2B_{eq} \qquad \text{--- (12.5)}$$

12.3.2 Magnetic field due to a bar magnet at an arbitrary point:

Fig. 12.3 Shows a bar magnet of magnetic moment $\vec{m}$ with centre at O. P is any point in its magnetic field. Magnetic moment $\vec{m}$ is resolved (*about the centre of the magnet*) into components along $\vec{r}$ and perpendicular to $\vec{r}$. For the component $m\cos\theta$ along $\vec{r}$, the point P is an axial point.

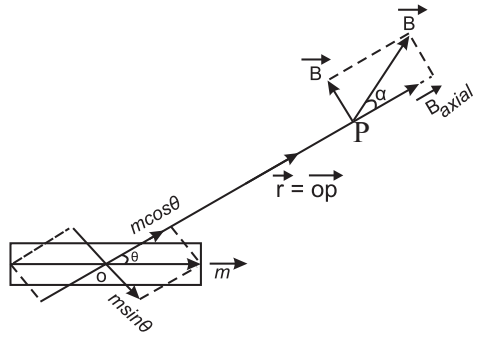


Fig. 12.3: Magnetic field at an arbitrary point.

Also, for the component $m\sin\theta$ perpendicular to $\vec{r}$, the point P is an equatorial point at the same distance $\vec{r}$. Using the results of axial and equatorial fields, we get

$$B_a = \frac{\mu_o}{4\pi}\frac{2m\cos\theta}{r^3} \qquad \text{--- (12.6)}$$

directed along $m\cos\theta$ and

$$B_{eq} = \frac{\mu_o}{4\pi}\frac{m\sin\theta}{r^3} \qquad \text{--- (12.7)}$$

directed opposite to $m\sin\theta$

Thus, the magnitude of the resultant magnetic field B, at point P is given by

$$B = \sqrt{B_a^{\,2} + B_{eq}^{\,2}}$$

$$\therefore\ B = \frac{\mu_o}{4\pi}\frac{m}{r^3}\sqrt{\left[2\cos\theta\right]^2 + \left[\sin\theta\right]^2}$$

$$\therefore B = \frac{\mu_o}{4\pi}\frac{m}{r^3}\sqrt{3\cos^2\theta + 1} \qquad \text{--- (12.8)}$$

Let α be the angle made by the direction of $\vec{B}$ with $\vec{r}$. Then, by using eq (12.6) and eq (12.7),

$$\tan\alpha = \frac{B_{eq}}{B_a} = \frac{1}{2}\left(\tan\theta\right) \qquad \text{--- (12.9)}$$

The angle between directions of $\vec{B}$ and $\vec{m}$ is then $\left(\theta + \alpha\right)$.

Example 12.1: A short magnetic dipole has magnetic moment 0.5 A m^2. Calculate its magnetic field at a distance of 20 cm from the centre of magnetic dipole on (i) the axis (ii) the equatorial line (Given $\mu_0 = 4\pi \times 10^{-7}$ SI units)

Solution :

$m = 0.5\ \text{Am}^2,\ r = 20\ \text{cm} = 0.2\ \text{m}$

$$B_a = \frac{\mu_0}{4\pi}\frac{2m}{r^3} = \frac{10^{-7}\times 2\times 0.5}{(0.2)^3} = \frac{1\times 10^{-7}}{8\times 10^{-3}}$$

$$= \frac{1}{8}\times 10^{-4} = 1.25\times 10^{-5}\ \text{Wb/m}^2$$

$$B_{eq} = \frac{\mu_0}{4\pi}\frac{m}{r^3}$$

$$= \frac{10^{-7}\times 0.5}{(0.2)^3} = \frac{5\times 10^{-8}}{8\times 10^{-3}} = 0.625\times 10^{-5}\ \text{Wb/m}^2$$

12.4 Gauss' Law of Magnetism:

The Gauss' law for electric field is known to you. It states that the net electric flux through a closed Gaussian surface is proportional to the net electric charge enclosed by the surface (Eq. (10.18)). The Gauss' law for magnetic fields states that the net magnetic flux Φ_B through a closed Gaussian surface is zero, i.e.,

$\Phi_B = \int \vec{B}\cdot d\vec{S} = 0$ (Gauss' law for magnetic fields)

The magnetic force lines of (a) bar magnet, (b) current carrying finite solenoid, and (c) electric dipole are shown in Fig.12.4(a), 12.4(b) and 12.4(c), respectively. The curves labelled (i) and (ii) are cross sections of three dimensional closed Gaussian surfaces.

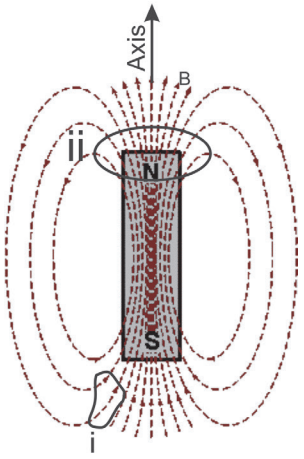


Fig. 12.4 (a): Bar magnet.

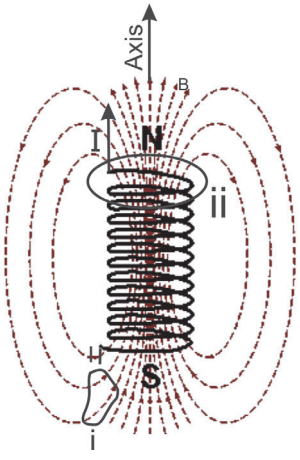


Fig. 12.4 (b): Current (I) carrying solenoid.

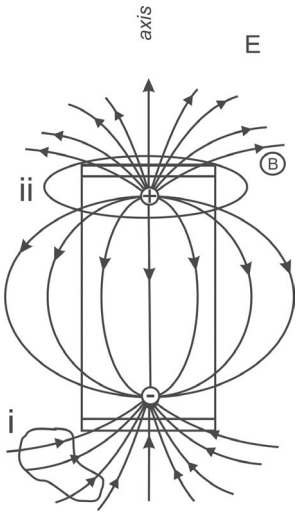


Fig. 12.4 (c): Electric dipole.

If we compare the number of lines of force entering in and leaving out of the surface (i), it is clearly seen that they are equal. The Gaussian surface does not include poles. It means that the flux associated with any closed surface is equal to zero. When we consider surface (ii), in Fig. 12.4 (b), we are enclosing the North pole. As even a thin slice of a bar magnet will have North and South poles associated with it, the closed Gaussian surface will also include a South pole. However in Fig. 12.4(c), for an electric dipole, the field lines begin from positive charge and end on negative charge. For a closed surface (ii), there is a net outward flux since it does include a net (positive) charge. According to the Gauss' law of electrostatics as studied earlier, $\Phi_E = \int \vec{E} \cdot \vec{dS} = \frac{q}{\varepsilon_0}$, where q is the positive charge enclosed. Thus, situation is entirely different from magnetic lines of force, which are shown in Fig. 12.4(a) and Fig. 12.4(b). Thus, Gauss' law of magnetism can be written as $\Phi_B = \int \vec{B} \cdot \vec{dS} = 0$.

From the above we conclude that for electrostatics, an isolated electric charge exists but an isolated magnetic pole does not exist. In short, only dipoles exist in case of magnetism.

12.5 Earth's Magnetism:

It is common experience that a bar magnet or a magnetic needle suspended freely in air always aligns itself along geographic N-S direction. If it has a freedom to rotate about horizontal axis, it inclines with some angle with the horizontal in the vertical N-S plane.

This fact clearly indicates that there is some magnetic field present everywhere on the Earth . This is called Terrestrial Magnetism. It is extremely useful during navigation.

Magnetic parameters of the Earth are described below. The magnetic lines of force enter the Earth's surface at the north pole and emerge from the south pole.

Unless and otherwise stated, the directions mentioned (South, North, etc.) are always, Geographic.

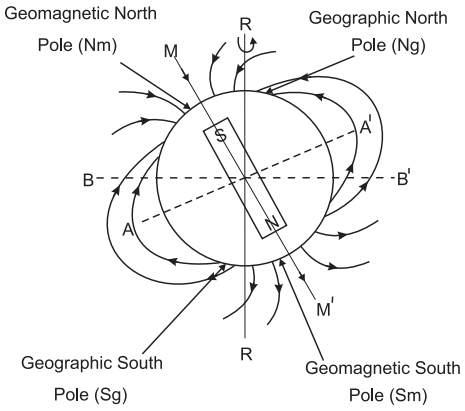


Fig. 12.5: Earth's magnetism.

Magnetic Axis :- The Earth is considered to be a huge magnet. Magnetic north pole (N) of the Earth is located below Antarctica while the south pole (S) is below north Canada. The straight line NS joining these two poles is called the magnetic axis, MM'.

Magnetic equator :- A great circle in the plane perpendicular to magnetic axis is magnetic equatorial circle, AA'. It happens to pass through India near Thiruvananthapuram.

Geographic Meridian:- A plane perpendicular to the surface of the Earth (vertical plane) perpendicular to geographic axis is geographic meridian. (Fig.12.6)

Magnetic Meridian:- A plane perpendicular to surface of the Earth (Vertical plane) and passing through the magnetic axis is magnetic meridian. Direction of resultant magnetic field of the Earth is always along or parallel to magnetic meridian. (Fig.12.6)

Magnetic declination:- Angle between the geographic and the magnetic meridian at a place is called 'magnetic declination' (α). The declination is small in India. It is 0° 58′ west at Mumbai and $0^0$41′ east at Delhi. Thus, at both these places, magnetic needle shows true North

accurately (Fig.12.6).

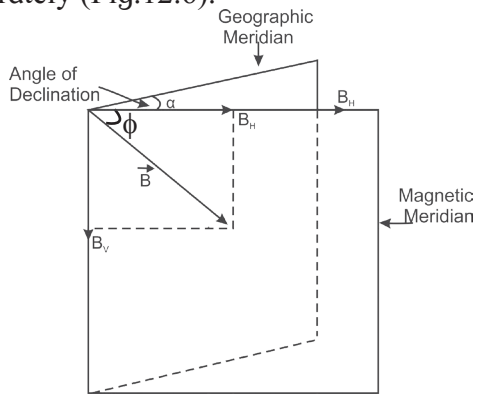


Fig. 12.6: Magnetic declination.

Magnetic inclination or angle of dip (ϕ):- Angle made by the direction of resultant magnetic field with the horizontal at a place is inclination or angle of dip at the place (Fig. 12.7).

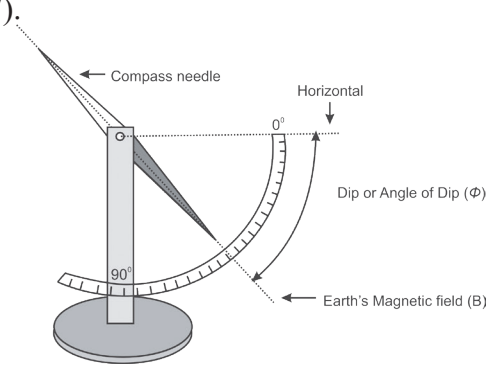


Fig. 12.7: Magnetic inclination.

Earth's magnetic field:- Magnetic force experienced per unit pole strength is magnetic field $\vec{B}$ at that place. It can be resolved in components along the horizontal, $\vec{B}_H$ and along vertical, $\vec{B}_V$. The vertical component can be conveniently determined. The two components can be related with the angle of dip (ϕ) as,

$B_H = B\cos\phi$, $B_V = B\sin\phi$

$$\frac{B_V}{B_H} = \tan\phi \qquad \text{--- (12.10)}$$

$$B^2 = B_V^{\,2} + B_H^{\,2}$$

$$\therefore B = \sqrt{B_V^{\,2} + B_H^{\,2}} \qquad \text{--- (12.11)}$$

Special cases

1) At the magnetic North pole, $\vec{B} = \vec{B}_V$, directed upward, $\vec{B}_H = 0$ and $\phi = 90^0$.
2) At the magnetic south pole, $\vec{B} = \vec{B}_V$,- directed downward, $\vec{B}_H = 0$ and $\phi = 270^0$.
3) Anywhere on the magnetic great circle (magnetic equator) $B = B_H$ along South to North, $B_V = 0$ and $\phi = 0$

Magnetic maps of the Earth:-

Magnetic elements of the Earth (B_H, α and ϕ) vary from place to place and also with time. The maps providing these values at different locations are called magnetic maps. These are extremely useful for navigation. Magnetic maps drawn by joining places with the same value of a particular element are called Iso-magnetic charts.

Lines joining the places of equal horizontal components (B_H) are known as 'Isodynamic lines'

Lines joining the places of equal declination (α) are called Isogonic lines.

Lines joining the places of equal inclination or dip (ϕ) are called Aclinic lines.

Example 12.2: Earth's magnetic field at the equator is approximately 4×10^{-5} T. Calculate Earth's dipole moment. (Radius of Earth = 6.4×10^6 m, $\mu_0 = 4\pi\times10^{-7}$ SI units)

Solution: Given

$B_{eq} = 4\times10^{-5}$ T

$r = 6.4\times10^6$ m

Assume that Earth is a bar magnet with N and S poles being the geographical South and North poles, respectively. The equatorial magnetic field due to Earth's dipole can be written as

$$B_{eq} = \frac{\mu_0 m}{4\pi r^3}$$

$$m = 4\pi B_{eq}\times r^3 / \mu_0$$

$$= 4\times10^{-5}\times(6.4\times10^6)^3\times10^7$$

$$= 1.05\times10^{20}\ \text{A m}^2$$

Example 12.3: At a given place on the Earth, a bar magnet of magnetic moment $\vec{m}$ is kept horizontal in the East-West direction. P and Q are the two neutral points due to magnetic field of this magnet and $\vec{B}_H$ is the horizontal component of the Earth's magnetic field.

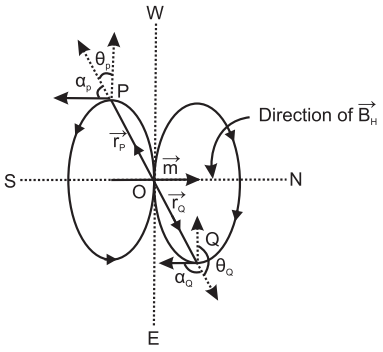
(A) Calculate the angles between position vectors of P and Q with the direction of $\vec{m}$.

(B) Points P and Q are 1 m from the centre of the bar magnet and $B_H = 3.5\times10^{-5}$ T. Calculate

magnetic dipole moment of the bar magnet.

Neutral point is that point where the resultant magnetic field is zero.

Solution: (A) As seen from the figure, the direction of magnetic field $\vec{B}$ due to the bar magnet is opposite to $\vec{B}_H$ at the points P and Q. Also, $(\theta+\alpha) = 90^0$ at P and it is 270^0 at Q.



$$\tan\alpha = \frac{1}{2}\tan\theta$$

$$\therefore \tan\theta = 2\tan\alpha$$

$$= 2\tan(90-\theta) \text{ and } 2\tan(270-\theta)$$

$$\therefore \tan\theta = \pm 2\cot\theta$$

$$\therefore \tan^2\theta = 2$$

$$\therefore \tan\theta = \pm\sqrt{2}$$

$$\therefore \theta = \tan^{-1}(\pm\sqrt{2})$$

$$\therefore \theta = 54^0 44' \text{ and } 180^0 - 54^0 44^0 = 116^0 4'$$

(B)

$$\tan^2\theta = 2$$

$$\therefore \sec^2\theta = 1+\tan^2\theta = 1+2 = 3$$

$$\therefore \cos^2\theta = \frac{1}{3}$$

$r = 1\text{m}$ and $B = B_H = 3.5\times10^{-5}$ T (Given)

we have,

$$B = \frac{\mu_0}{4\pi}\frac{m}{r^3}\sqrt{3\cos^2\theta+1}$$

$$\therefore m = \frac{B_H \times r^3}{\left(\frac{\mu_0}{4\pi}\right)\sqrt{3\cos^2\theta+1}}$$

$$= \frac{3.5\times10^{-5}\times1^3}{10^{-7}\times\sqrt{3\frac{1}{3}+1}}$$

$$\therefore m = \frac{350}{\sqrt{2}} = 247.5 \text{ A m}^2$$

Always remember:
In this Chapter we have used B as a symbol for magnetic field. Calling it magnetic induction is unreasonable. We have used the words **magnetic field** which are used in spoken language.

Exercises

1. Choose the correct option.

i) Let r be the distance of a point on the axis of a bar magnet from its center. The magnetic field at r is always proportional to

(A) $1/r^2$ (B) $1/r^3$ (C) $1/r$

(D) not necessarily $1/r^3$ at all points

ii) Magnetic meridian is the plane

(A) perpendicular to the magnetic axis of Earth

(B) perpendicular to geographic axis of Earth

(C) passing through the magnetic axis of Earth

(D) passing through the geographic axis Earth

iii) The horizontal and vertical component of magnetic field of Earth are same at some place on the surface of Earth. The magnetic dip angle at this place will be

(A) 30^o (B) 45^o

(C) 0^o (D) 90^o

iv) Inside a bar magnet, the magnetic field lines

(A) are not present

(B) are parallel to the cross sectional area of the magnet

(C) are in the direction from N pole to S pole

(D) are in the direction from S pole to N pole

v) A place where the vertical components of Earth's magnetic field is zero has the angle of dip equal to

(A) 0^o (B) 45^o

(C) 60^o (D) 90^o

vi) A place where the horizontal component of Earth's magnetic field is zero lies at

(A) geographic equator

(B) geomagnetic equator

(C) one of the geographic poles

(D) one of the geomagnetic poles

vii) A magnetic needle kept nonparallel to the magnetic field in a nonuniform magnetic field experiences

(A) a force but not a torque

(B) a torque but not a force

(C) both a force and a torque

(D) neither force nor a torque

2. Answer the following questions in brief.

i) What happens if a bar magnet is cut into two pieces transverse to its length/ along its length?

ii) What could be the equation for Gauss' law of magnetism, if a monopole of pole strength p is enclosed by a surface?

3. Answer the following questions in detail.

i) Explain the Gauss' law for magnetic fields.

ii) What is a geographic meridian. How does the declination vary with latitude? Where is it minimum?

iii) Define the Angle of Dip. What happens to angle of dip as we move towards magnetic pole from magnetic equator?

4. Solve the following Problems.

i) A magnetic pole of bar magnet with pole strength of 100 A m is 20 cm away from the centre of a bar magnet. Bar magnet has pole strength of 200 A m and has a length 5 cm. If the magnetic pole is on the axis of the bar magnet, find the force on the magnetic pole.

[Ans: 2.5×10^{-2}N]

ii) A magnet makes an angle of 45^o with the horizontal in a plane making an angle of 30^o with the magnetic meridian. Find the true value of the dip angle at the place.

[Ans: $\tan^{-1}(0.866)$]

iii) Two small and similar bar magnets have magnetic dipole moment of 1.0 Am^2 each. They are kept in a plane in such a way that their axes are perpendicular to each other. A line drawn through the axis of one magnet passes through the center of other magnet. If the distance between their centers is 2 m, find the magnitude of magnetic field at the mid point of the line joining their centers.

[Ans: $\sqrt{5}\times10^{-7}$T]

iv) A circular magnet is made with its north pole at the centre, separated from the surrounding circular south pole by an air a gap. Draw the magnetic field lines in the gap. [The magnet is hypothetical magnet].

Draw a diagram to illustrate the magnetic lines of force between the south poles of two such magnets.

v) Two bar magnets are placed on a straight line with their north poles facing each other on a horizontal surface. Draw magnetic lines around them. Mark the position of any neutral points (points where there is no resultant magnetic field) on your diagram.

***

Internet my friend

https://www.ngdc.noaa.gov

13. Electromagnetic Waves and Communication System

Can you recall?

1. What is a wave?
2. What is the difference between longitudinal and transverse waves?
3. What are electric and magnetic fields and what are their sources?
4. What are Lenz's law, Ampere's law and Faraday's law?
5. By which mechanism heat is lost by hot bodies ?

13.1 Introduction :

The information age in which we live is based almost entirely on the physics of electromagnetic (EM) waves. We are now globally connected by TV, cellphone and internet. All these gadgets use EM waves as carriers for transmission of signals. Energy from the Sun, an essential requirement for life on Earth, reaches us by means of EM waves that travel through nearly 150 million km of empty space. There are EM waves from light bulbs, heated engine blocks of automobiles, x-ray machines, lightning flashes, and some radioactive materials. Stars, other objects in our milky way galaxy and other galaxies are known to emit EM waves. Hence, it is important for us to make a careful study of the properties of EM waves.

13.2 EM wave:

There are four basic laws which describe the behaviour of electric and magnetic fields, the relation between them and their generation by charges and currents. These laws are as follows.

(1) Gauss' law for electrostatics, which is essentially the Coulomb's law, describes the relationship between static electric charges and the electric field produced by them.

(2) Gauss' law for magnetism, which is similar to the Gauss' law for electrostatics mentioned above, states that "magnetic monopoles which are thought to be magnetic charges equivalent to the electric charges, do not exist". Magnetic poles always occur in pairs.

(3) Faraday's law which gives the relation between electromotive force (emf) induced in a circuit when the magnetic flux linked with the circuit changes.

(4) Ampere's law gives the relation between the induced magnetic field associated with a loop and the current flowing through the loop. Maxwell (1831-1879) noticed a major flaw in the Ampere's law for time dependant fields. He noticed that the magnetic field can be generated not only by electric current but also by changing electric field. Therefore in the year 1861, he added one more term to the equation describing this law. This term is called the displacement current. This term is extremely important and the EM waves which are an outcome of these equations would not have been possible in absence of this term.

As a result, the set of four equations describing the above four laws is called Maxwell's equations.

In 1888, H. Hertz (1857-1894) succeeded in producing and detecting the existence of EM waves. He also demonstrated their properties namely reflection, refraction and interference.

In 1895, an Indian physicist Sir Jagdish Chandra Bose (1858-1937) produced EM waves ranging in wavelengths from 5 mm to 25 nm. His work, however, remained confined to laboratory only.

In 1896, an Italian physicist G. Marconi (1874-1937) became pioneer in establishing wireless communication. He was awarded the Nobel prize in physics in 1909 for his work in developing wireless telegraphy, telephony and broadcasting.

13.2.1 Sources of EM waves:

According to Maxwell's theory, "accelerated charges radiate EM waves". Consider a charge oscillating with some frequency. This produces

an oscillating electric field in space, which produces an oscillating magnetic field which in turn is a source of oscillating electric field. Thus varying electric and magnetic fields regenerate each other.

Waves that are caused by the acceleration of charged particles and consist of electric and magnetic fields vibrating sinusoidally at right angles to each other and to the direction of propagation are called EM waves or EM radiation. Figure 13.1 shows an EM wave propagating along *z*-axis. The time varying electric field is along the *x*-axis and time varying magnetic field is along the *y*-axis.

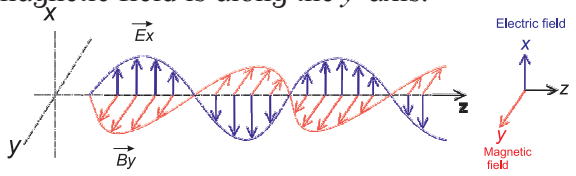


Fig. 13.1: EM wave propagating along *z*-axis.

Do you know ?

In 1865, Maxwell proposed that an oscillating electric charge radiates energy in the form of EM wave. EM waves are periodic changes in electric and magnetic fields, which propagate through space. Thus, energy can be transported in the form of EM waves.

Maxwell's Equations for Charges and Currents in Vacuum

1) $\int \vec{E} \cdot \vec{dS} = \frac{Q_{in}}{\varepsilon_0}$ (Gauss' law)

Here $\vec{E}$ is the electric field and ε_0 is the permittivity of vacuum. The integral is over a closed surface *S*. The law states that electric flux through any closed surface *S* is equal to the total electric charge Q_{in} enclosed by the surface divided by ε_0. Gauss' law describes the relation between an electric charge and electric field it produces.

2) $\int \vec{B} \cdot \vec{dS} = 0$ (Gauss' law for magnetism).

Here $\vec{B}$ is the magnetic field. The integral is over a closed surface *S*. The law states that magnetic flux through a closed surface is always zero, i.e., the magnetic field lines are continuous closed curves, having neither beginning nor end.

3) $\int \vec{E} \cdot \vec{dl} = -\frac{d\phi_m}{dt}$

(Faraday's law with Lenz's law)

Here ϕ_m is the magnetic flux and the integral is over a closed loop. Time varying magnetic field induces an electromotive force (emf) and hence, an electric field. The direction of the induced emf is such that the change is opposed.

4) $\int \vec{B} \cdot \vec{dl} = \mu_0 I + \varepsilon_0 \mu_0 \frac{d\phi_E}{dt}$

(Ampere-Maxwell law)

Here μ_0 is the permeability of vacuum and the integral is over a closed loop, *I* is the current flowing through the loop. ϕ_E is the electric flux linked with the circuit. Magnetic field is generated by moving charges and also by varying electric fields.

13.2.2 Characteristics of EM waves:

1) The electric and magnetic fields, $\vec{E}$ and $\vec{B}$ are always perpendicular to each other and also to the direction of propagation of the EM wave. Thus the EM waves are transverse waves.
2) The cross product $\vec{E} \times \vec{B}$ gives the direction in which the EM wave travels. $\vec{E} \times \vec{B}$ also gives the energy carried by EM wave.
3) The $\vec{E}$ and $\vec{B}$ fields vary sinusoidally and are in phase.
4) EM waves are produced by accelerated electric charges.
5) EM waves can travel through free space as well as through solids, liquids and gases.
6) In free space, EM waves travel with velocity *c*, equal to that of light in free space.

$$c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3 \times 10^8 \text{ m/s},$$

where μ_0 ($4\pi \times 10^{-7}$ Tm/A) is permeability and ε_0 (8.85×10^{-12} C^2/Nm^2) is permittivity of free space.

7) In a given material medium, the velocity (v_m) of EM waves is given by $v_m = \frac{1}{\sqrt{\mu\varepsilon}}$ where μ is the permeability and ε is the

permittivity of the given medium.

8) The EM waves obey the principle of superposition.

9) The ratio of the amplitudes of electric and magnetic fields is constant at any point and is equal to the velocity of the EM wave.

$$|\vec{E_0}| = c\,|\vec{B_0}| \text{ or } \frac{|\vec{E_0}|}{|\vec{B_0}|} = \frac{1}{\sqrt{\mu_0\varepsilon_0}} \quad \text{--- (13.1)}$$

$|\vec{E_0}|$ and $|\vec{B_0}|$ are the amplitudes of $\vec{E}$ and $\vec{B}$ respectively.

10) As the electric field vector ($\vec{E}_0$) is more prominent than the magnetic field vector ($\vec{B}_0$), it is responsible for optical effects due to EM waves. For this reason, electric vector is called light vector.

11) The intensity of a wave is proportional to the square of its amplitude and is given by the equations

$$I_E = \frac{1}{2}\varepsilon_0 E_0^{\,2},\ I_B = \frac{1}{2}\frac{B_0^{\,2}}{\mu_0} \quad \text{--- (13.2)}$$

12) The energy of EM waves is distributed equally between the electric and magnetic fields. $I_E = I_B$

Example 13.1: Calculate the velocity of EM waves in vacuum.

Solution: The velocity of EM wave in free space is given by

$$c = \frac{1}{\sqrt{\mu_0\varepsilon_0}} = \frac{1}{\sqrt{(8.85\times10^{-12}\,\frac{\text{C}^2}{\text{Nm}^2})(4\pi\times10^{-7}\,\frac{\text{T.m}}{\text{A}})}}$$

$$c = 3.00\times10^{+8}\,\text{m/s}$$

Example 13.2: In free space, an EM wave of frequency 28 MHz travels along the *x*-direction. The amplitude of the electric field is E = 9.6 V/m and its direction is along the *y*-axis. What is amplitude and direction of magnetic field B?

Solution : We have,

$$|B| = \frac{|E|}{c} = \frac{9.6\ \text{V/m}}{3\times10^8\ \text{m/s}}$$

$$B = 3.2\times10^{-8}\,\text{T}$$

It is given that E is along *y*-direction and the wave propagates along *x*-axis. The magnetic field B should be in a direction perpendicular to both *x*- and *y*-axes. As per property (2) of EM waves, $\vec{E}\times\vec{B}$ should be along the direction of propagation which is along the *x*- axis

Since $(+\hat{j})\times(+\hat{k}) = \hat{i}$, B is along the $\hat{k}$, i.e., along the *z*-direction.

Thus, the amplitude of B = 3.2×10^{-8} T and its direction is along the *z*-axis.

Do you know ?

According to quantum theory, an electron, while orbiting around the nucleus in a stable orbit does not emit EM radiation even though it undergoes acceleration. It will emit an EM radiation only when it falls from an orbit of higher energy to one of lower energy.

EM waves (such as X-rays) are produced when fast moving electrons hit a target of high atomic number (such as molybdenum, copper, etc.).

An electric charge at rest has an electric field in the region around it but has no magnetic field. When the charge moves, it produces both electric and magnetic fields. If the charge moves with a constant velocity, the magnetic field will not change with time and as such it cannot produce an EM wave. But if the charge is accelerated, both the magnetic and electric fields change with space and time and an EM wave is produced. Thus an oscillating charge emits an EM wave which has the same frequency as that of the oscillation of the charge.

Example 13.3: A beam of red light has an amplitude 2.5 times the amplitude of second beam of the same colour. Calculate the ratio of the intensities of the two waves.

Solution: Intensity $\propto$ (Amplitude)^2

$I_2 \propto (a)^2$ and $I_1 \propto (2.5a)^2$

$$\therefore \frac{I_1}{I_2} = \frac{(2.5a)^2}{a^2} = (2.5)^2 = 6.25$$

In an EM wave, the magnetic field and electric field both vary sinusoidally with *x*. For a wave travelling along *x*-axis having $\vec{E}$ along *y*-axis and $\vec{B}$ along the z axis, with reference to Chapter 8, we can write E_y and B_z as

$$E_y = E_0 \sin(kx-\omega t) \quad \text{--- (13.2)}$$

and $$B_z = B_0 \sin(kx-\omega t), \quad \text{--- (13.3)}$$

where E_0 is the amplitude of the electric field E_y and B_0 is the amplitude of the magnetic field B_z. $k = \frac{2\pi}{\lambda}$ is the propagation constant and λ is the wavelength of the wave. ω = 2πυ is the angular frequency of oscillations, υ being the frequency of the wave.

Both the electric and magnetic fields attain their maximum (and minimum) values at the same time and at the same point in space, i.e., $\vec{E}$ and $\vec{B}$ oscillate in phase with the same frequency.

Example 13.4: An EM wave of frequency 50 MHz travels in vacuum along the positive *x*-axis and $\vec{E}$ at a particular point, *x* and at a particular instant of time t is 9.6 $\hat{j}$ V/m. Find the magnitude and direction of $\vec{B}$ at this point *x* and at time t.

Solution : $B = \frac{E}{c} = \frac{9.6}{3\times10^8} = 3.2\times10^{-8}$ T

As the wave propagates along +*x* axis and *E* is along +*y* axis, direction of *B* will be along +*z*-axis i.e. $B = 3.2\times10^{-8}\ \hat{k}$ T.

Example 13.5: For an EM wave propagating along *x* direction, the magnetic field oscillates along the *z*-direction at a frequency of 3×10^{10} Hz and has amplitude of 10^{-9} T.

a) What is the wavelength of the wave?

b) Write the expression representing the corresponding electric field.

Solution :

a) $\lambda = \frac{c}{\upsilon} = \frac{3\times10^8 \text{ m/s}}{3\times10^{10}/\text{s}} = 10^{-2}\text{ m}$

b) $E_0 = cB_0 = (3\times10^8 \text{ m/s}) \times (10^{-9} \text{ T}) = 0.3$ V/m. Since *B* acts along *z*-axis, *E* acts along *y*-axis. Expression representing the oscillating electric field is

$$E_y = E_0 \sin(kx - \omega t)$$

$$E_y = E_0 \sin\left[\left(\frac{2\pi}{\lambda}\right)x - (2\pi\nu)t\right]$$

$$E_y = E_0 \sin 2\pi\left[\frac{x}{\lambda} - \nu t\right]$$

$$E_y = E_0 \sin 2\pi\left[\frac{x}{10^{-2}} - 3\times10^{10}t\right]$$

$$E_y = E_0 \sin 2\pi\left[100x - 3\times10^{10}t\right] \text{ V/m}$$

Example 13.6: The magnetic field of an EM wave travelling along *x*-axis is $\vec{B} = \hat{k}\ 4\times10^{-4} \sin(\omega t - kx)$. Here *B* is in tesla, *t* is in second and *x* is in m. Calculate the peak value of electric force acting on a particle of charge 5 μC travelling with a velocity of 5×10^5 m/s along the *y*-axis.

Solution :

$B_0 = 4\times10^{-4}$ T, $q = 5\ \mu\text{C} = 5\times10^{-6}$ C

$v = 5\times10^5$ m/s

$E_0 = cB_0 = (3\times10^8) \times (4\times10^{-4})$

$= 12\times10^4$ N/C

Maximum electric force $= qE_0$

$= (5\times10^{-6})\ (12\times10^4)$

$= 60\times10^{-2}$

$= 0.6$ N

13.3 Electromagnetic Spectrum:

The orderly distribution (sequential arrangement) of EM waves according to their wavelengths (or frequencies) in the form of distinct groups having different properties is called the EM spectrum (Fig. 13.2). The properties of different types of EM waves are given in Table 13.1.

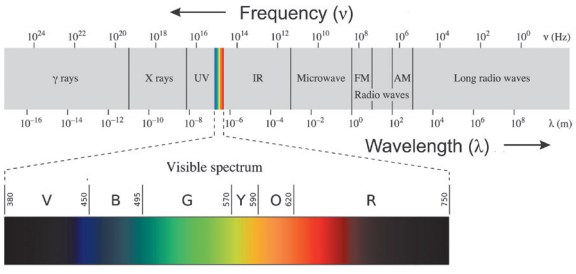


Fig. 13.2: Electromagnetic spectrum.

We briefly describe different types of EM waves in the order of decreasing wavelength (or increasing frequency).

13.3.1 Radio waves :

Radio waves are produced by accelerated motion of charges in a conducting wire. The

frequency of waves produced by the circuit depends upon the magnitudes of the inductance and the capacitance (This will be discussed in XIIth standard). Thus, by choosing suitable values of the inductance and the capacitance, radio waves of desired frequency can be produced.

Properties :

1) They have very long wavelengths ranging from a few centimetres to a few hundreds of kilometres.
2) The frequency range of AM band is 530 kHz to 1710 kHz. Frequency of the waves used for TV-transmission range from 54 MHz to 890 MHz, while those for FM radio band range from 88 MHz to 108MHz.

Notation used for high frequencies
1 kHz = one kilo Hertz =1000 Hz = 10^3 Hz
1 MHz = one mega Hertz = 10^6 Hz
1 GHz = one giga Hertz =10^9 Hz
Notation used for small wavelengths
1 μm = one micrometer = 10^{-6} m
1 Å= one angstrom = 10^{-10} m= 10^{-8} cm
1nm = one nanometer = 10^{-9} m

Uses :

1) Radio waves are used for wireless communication purpose.
2) They are used for radio broadcasting and transmission of TV signals.
3) Cellular phones use radio waves to transmit voice communication in the ultra high frequency (UHF) band.

Table 13.1: Properties of different types of EM waves

Name	Wavelength range in m	Frequency range in Hz	Generated By
Gamma rays	6×10^{-13} to 1×10^{-10}	5×10^{20} to 3×10^{18}	a) Transitions of nuclear energy levels b) Radioactive substances
X-rays	1×10^{-11} to 3×10^{-8}	3×10^{19} to 1×10^{16}	a) Bombardment of high energy electrons (keV) on a high atomic number target (Cu, Mg, Co) b) Energy level transitions of innermost orbital electrons
Ultraviolet (UV waves)	3×10^{-8} to 4×10^{-7}	1×10^{16} to 8×10^{14}	Rearrangement of orbital electrons of atom between energy levels. As in high voltage gas discharge tube, the Sun and mercury vapour lamp, etc.
Visible light	4×10^{-7} to 8×10^{-7}	8×10^{14} to 4×10^{14}	Rearrangement of outer orbital electrons in atoms and molecules e.g., gas discharge tube
Infrared (IR) radiations	8×10^{-7} to 3×10^{-4}	4×10^{14} to 1×10^{12}	Hot objects
Microwaves	3×10^{-4} to 6×10^{-2}	1×10^{12} to 5×10^{9}	Special electronic devices such as klystron tube
Radio waves	6×10^{-4} to 1×10^{5}	5×10^{11} to 8×10^{10}	Acceleration of electrons in circuits

13.3.2 Microwaves :

These waves were discovered of by H. Hertz in 1888. Microwaves are produced by oscillator electric circuits containing a capacitor and an inductor. They can be produced by special vacuum tubes.

Properties

1) They heat certain substances on which they are incident.
2) They can be detected by crystal detectors.

Uses

1) Used for the transmission of TV signals.
2) Used for long distance telephone communication.
3) Microwave ovens are used for cooking.
4) Used in radar systems for the location of

distant objects like ships, aeroplanes etc,

5) They are used in the study of atomic and molecular structure.

13.3.3 Infrared waves

These waves were discovered by William Herschel (1737-1822) in 1800. All hot bodies are sources of infrared rays. About 60% of the solar radiations are infrared in nature. Thermocouples, thermopile and bolometers are used to detect infrared rays.

Properties

1) When infrared rays are incident on any object, the object gets heated.
2) These rays are strongly absorbed by glass.
3) They can penetrate through thick columns of fog, mist and cloud cover.

Uses

1) Used in remote sensing.
2) Used in diagnosis of superficial tumours and varicose veins.
3) Used to cure infantile paralysis and to treat sprains, dislocations and fractures.
4) They are used in Solar water heaters and cookers.
5) Special infrared photographs of the body called thermograms, can reveal diseased organs because these parts radiate less heat than the healthy organs.
6) Infrared binoculars and thermal imaging cameras are used in military applications for night vision.
7) Used to keep green house warm.
8) Used in remote controls of TV, VCR, etc

13.3.4 Visible light :

It is the most familiar form of EM waves. These waves are detected by human eye. Therefore this wavelength range is called the visible light. The visible light is emitted due to atomic excitations.

Properties :

1) Different wavelengths give rise to different colours. These are given in Table 13.2.
2) Visible light emitted or reflected from objects around us provides us information about those objects and hence about the surroundings.

Do you know ?

Stars and galaxies emit different types of waves. Radio waves and visible light can pass through the Earth's atmosphere and reach the ground without getting absorbed significantly. Thus the radio telescopes and optical telescopes can be placed on the ground. All other type of waves get absorbed by the atmospheric gases and dust particles. Hence, the γ-ray, X-ray, ultraviolet, infrared, and microwave telescopes are kept aboard artificial satellites and are operated remotely from the Earth. Even though the visible radiation reaches the surface of the Earth, its intensity decreases to some extent due to absorption and scattering by atmospheric gases and dust particles. Optical telescopes are therefore located at higher altitudes.

The Indian Giant Metrewave Radio Telescope (GMRT) near Pune is an important milestone in the field of Radio-astronomy. Also, Indian Astronomical Observatory houses the Himalayan Chandra Telescope (HCT), the 2 m optical-IR Telescope, which is situated at Hanle, Ladakh, at an altitude of 4500 m.

Table 13.2: Wavelengths of colours in visible light

Colour	Wavelength
violet	380-450 nm
blue	450-495 nm
green	495-570 nm
yellow	570-590 nm
orange	590-620 nm
red	620-750 nm

13.3.5 Ultraviolet rays :

Ultraviolet rays were discovered by J. Ritter (1776-1810) in 1801. They can be produced by the mercury vapour lamp, electric spark and carbon arc lamp. They can also be obtained by striking electrical discharge in hydrogen and xenon gas tubes. The Sun is the most important natural source of ultraviolet rays, most of which are absorbed by the ozone layer in the Earth's atmosphere.

Properties :

1) They produce fluorescence in certain materials, such as 'phosphors'.
2) They cause photoelectric effect.
3) They cannot pass through glass but pass through quartz, fluorite, rock salt etc.
4) They possess the property of synthesizing vitamin D, when skin is exposed to them.

Uses :

1) Ultraviolet rays destroy germs and bacteria and hence they are used for sterilizing surgical instruments and for purification of water.
2) Used in burglar alarms and security systems.
3) Used to distinguish real and fake gems.
4) Used in analysis of chemical compounds.
5) Used to detect forgery.

Do you know ?

1. A fluorescent light bulb is coated from with a powder inside and contains a gas; electricity causes the gas to emit ultraviolet radiation, which then stimulates the tube coating to emit light.
2. The pixels of a television or computer screen fluoresce when electrons from an electron gun strike them.
3. What we call 'visible light' is just the part of the EM spectrum that human eyes see. Many other animals would define 'visible' somewhat differently. For instance, many animals including insects and birds, see in the UV region. Natural world is full of signals that animals see and humans cannot. Many birds including bluebirds, budgies, parrots and even peacocks have ultraviolet patterns that make them even more vivid to each other than they are to us.

13.3.6 *X*-rays:

German physicist W. C. Rontgen (1845-1923) discovered *X*-rays in 1895 while studying cathode rays (which is a stream of electrons emitted by the cathode in a vacuum tube). *X*-rays are also called Rontgen rays. *X*-rays are produced when cathode rays are suddenly stopped by an obstacle.

Properties

1) They are high energy EM waves.
2) They are not deflected by electric and magnetic fields.
3) X-rays ionize the gases through which they pass.
4) They have high penetrating power.
5) Their over dose can kill living plant and animal overdose tissues and hence are harmful.

Uses

1) Useful in the study of the structure of crystals.
2) X-ray photographs are useful to detect bone fracture. X-rays have many other medical uses such as CT scan.
3) X-rays are used to detect flaws or cracks in metals.
4) These are used for detection of explosives, opium etc.

13.3.7 Gamma Rays (γ-rays)

Discovered by P. Villard (1860-1934) in 1900. Gamma rays are emitted from the nuclei of some radioactive elements such as uranium, radium etc.

Properties

1) They are highest energy EM waves. (energy range keV - GeV)
2) They are highly penetrating.
3) They have a small ionising power.
4) They kill living cells.

Uses

1) Used as insecticide disinfection for wheat and flour.
2) Used for food preservation.
3) Used in radiotherapy for the treatment of cancer and tumour.
4) They are used to produce nuclear reactions.

13.4 Propagation of EM Waves:

You must have seen a TV antenna used to receive the TV signals from the transmitting tower or from a satellite. In communication using radio waves, an antenna in the transmitter radiates the EM waves, which travel through space and reach the receiving antenna at the other end. As the EM wave travels away from the transmitter; the strength of the wave keeps

on decreasing. Several factors influence the propagation of EM waves and the path they follow. It is also important to understand the composition of the Earth's atmosphere as it plays a vital role in the propagation of EM waves. Different layers of Earth's atmosphere are shown in Fig. 13.3.

Do you know ?

Ionizing radiations :

Ultraviolet, *X*-ray and gamma rays have sufficient energy to cause ionization i.e. they strip electrons from atoms and molecules lying along their path. The atoms lose their electrons and are then known as ions. Ionization is harmful to human beings because it can kill or damage living cells, or make them grow abnormally as cancers. Fluorescent lamps are based on ionization of gas. Ionizing radiation is also used in various equipments in laboratory and industry.

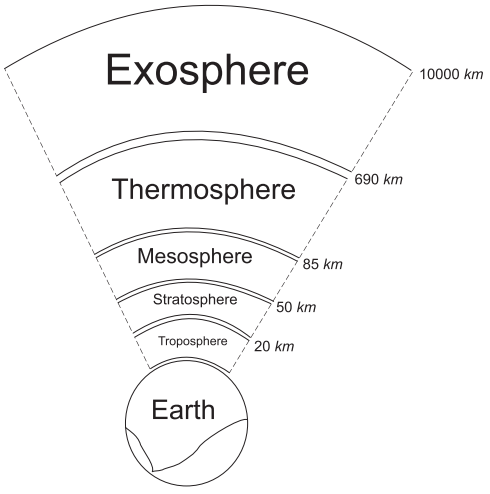


Fig 13.3: Earth and atmospheric layers.

Different modes of propagation of EM waves are described below and are shown in Fig. 13.4.

Do you know ?

X-rays have many practical applications in medicine and industry. Because *X*-ray photons are of such high energy, they can penetrate several centimetres of solid matter and can be used to visualize the interiors of materials that are opaque to ordinary light.

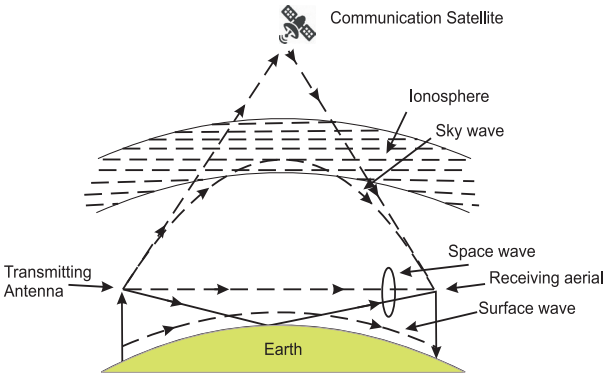


Fig. 13.4: Propagation of EM waves.

13.4.1 Ground (surface) wave:

When a radio wave from a transmitting antenna propagates near surface of the Earth so as to reach the receiving antenna, the wave propagation is called ground wave or surface wave propagation.

In this mode, radio waves travel close to the surface of the Earth and move along its curved surface from transmitter to receiver.

The radio waves induce currents in the ground and lose their energy by absorption. Therefore, the signal cannot be transmitted over large distances. Radio waves having frequency less than 2 MHz (in the medium frequency band) are transmitted by ground wave propagation. This is suitable for local broadcasting only. For TV or FM signals (very high frequency), ground wave propagation cannot be used.

13.4.2 Space wave:

When the radio waves from the transmitting antenna reach the receiving antenna either directly along a straight line (line of sight) or after reflection from the ground or satellite or after reflection from troposphere, the wave propagation is called space wave propagation. The radio waves reflected from troposphere are called tropospheric waves. Radio waves with frequency greater than 30 MHz can pass through the ionosphere (60 km - 1000 km) after suffering a small deviation. Hence, these waves cannot be transmitted by space wave propagation except by using a satellite. Also, for TV signals which have high frequency, transmission over long distance is not possible by means of space wave propagation.

The maximum distance over which a signal can reach is called its range. For larger TV

coverage, the height of the transmitting antenna should be as large as possible. This is the reason why the transmitting and receiving antennas are mounted on top of high rise buildings.

Range is the straight line distance from the point of transmission (the top of the antenna) to the point on Earth where the wave will hit while travelling along a straight line. Range is shown by d in Fig. 13.5. Let the height of the transmitting antenna (AA') situated at A be h. B represents the point on the surface of the Earth at which the space wave hits the Earth. The triangle OA'B is a right angled triangle. From Δ OA' B we can write

$OA'^2 = A'B^2 + OB^2$

$(R+h)^2 = d^2 + R^2$

or $R^2 + h^2 + 2Rh = d^2 + R^2$

As $h << R$, we can ignore h^2 and write

$d \cong \sqrt{2Rh}$

The range can be increased by mounting the receiver at a height h' say at a point C on the surface of the Earth. The range increases to $d + d'$ where d' is $\sqrt{2Rh'}$. Thus

Total range = $d + d' = \sqrt{2Rh} + \sqrt{2Rh'}$

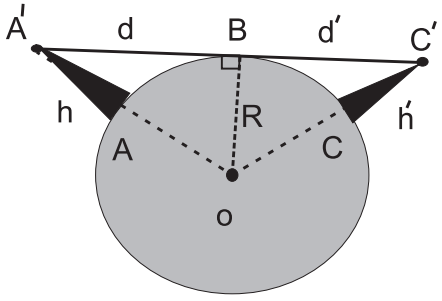


Fig. 13.5: Range of the signal (not to scale).

Example 13.7: A radar has a power of 10 kW and is operating at a frequency of 20 GHz. It is located on the top of a hill of height 500 m. Calculate the maximum distance upto which it can detect object located on the surface of the Earth . (Radius of Earth = 6.4×10^6 m)

Solution:

Maximum distance (range) =

$d = \sqrt{2Rh}$

$= \sqrt{2\times(6.4\times10^6)\times500}$ m

$= 8\times10^4 = 80$ km,

where R is radius of the Earth and h is the height of the radar above Earth's surface.

Example 13.8: If the height of a TV transmitting antenna is 128 m, how much square area can be covered by the transmitted signal if the receiving antenna is at the ground level? (Radius of the Earth = 6400 km)

Solution:

Range = $d = \sqrt{2Rh}$

$= \sqrt{2(6400\times10^3)(128)}$

$= \sqrt{16.384\times10^5\times10^3}$

$= \sqrt{16.384\times10^8}$

$= 4.047\times10^4$

$= 40.470$ km

Area covered = $\pi d^2 = 3.14\times(40)^2$

$= 5144.58$ km^2

Example 13.9: The height of a transmitting antenna is 68 m and the receiving antenna is at the top of a tower of height 34 m. Calculate the maximum distance between them for satisfactory transmission in line of sight mode. (radius of Earth = 6400 km)

$h_t = 68$ m, $h_r = 34$m, $R = 6400$ km $= 6.4\times10^6$m

Solution:

$d_{max} = \sqrt{2Rh_t} + \sqrt{2Rh_r}$

$= \sqrt{2\times6.4\times10^6\times68} + \sqrt{2\times6.4\times10^6\times34}$

$= \sqrt{870}\times10^3 + \sqrt{435}\times10^3$

$= 29.5\times10^3 + 20.9\times10^3$

$= 50.4\times10^3$ m

$= 50.4$ km

13.4.3 Sky wave propagation:

When radio waves from a transmitting antenna reach the receiving antenna after reflection in the ionosphere, the wave propagation is called sky wave propagation.

The sky waves include waves of frequency between 3 MHz and 30 MHz. These waves can suffer multiple reflections between the ionosphere and the Earth. Therefore, they can be transmitted over large distances.

Critical frequency : It is the maximum value of the frequency of radio wave which can be reflected back to the Earth from the ionosphere when the waves are directed normally to ionosphere.

Skip distance (zone) : It is the shortest distance from a transmitter measured along the surface of the Earth at which a sky wave of fixed frequency (if grater than critical frequency) will be returned to the Earth so that no sky waves can be received within the skip distance.

13.5 Introduction to Communication System:

Communication is exchange of information. Since ancient times it is practiced in various ways e.g., through speaking, writing, singing, using body language etc. After the discovery of electricity in the late 19th century, human communication systems changed dramatically. Modern communication is based upon the discoveries and inventions by a number of scientists like J. C. Bose (1858-1937), S. F. B. Morse (1791-1872), G. Marconi (1874-1937) and Alexander Graham Bell (1847-1922) in the 19th and 20th centuries.

In the 20th century we could send messages over large distances using analogue signals, cables and radio waves. With the advancements of digitization technologies, we can now communicate with the entire world almost in real time.

The ability to communicate is an important feature of modern life. We can speak directly to others all around the world and generate vast amount of information every day.

Here we will briefly discuss how communication systems work. A communication system is a device or set up used in transmission and reception of information from one place to another.

13.5.1 Elements of a communication system:

There are three basic (essential) elements of every communication system: a) Transmitter, b) Communication channel and c) Receiver.

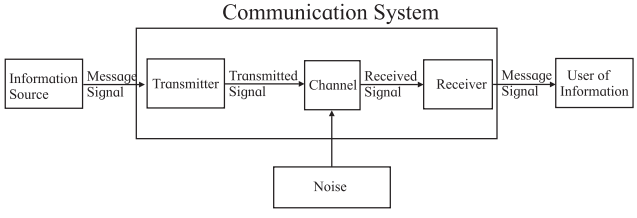


Fig. 13.6: Block diagram of the basic elements of a communication system.

In a communication system, as shown in Fig. 13.6, the transmitter is located at one place and the receiver at another place. The communication channel is a passage through which signals transfer in between a transmitter and a receiver. This channel may be in the form of wires or cables, or may also be wireless, depending on the types of communication system.

There are two basic modes of communication: (i) point to point communication and (ii) broadcast.

In point to point communication mode, communication takes place over a link between a single transmitter and a receiver e.g. Telephony. In the broadcast mode there are large number of receivers corresponding to the single transmitter e.g., Radio and Television transmission.

13.5.2 Commonly used terms in electronic communication system:

Following terms are useful to understand any communication system:

1) Signal :- The information converted into electrical form that is suitable for transmission is called a signal. In a radio station, music and speech are converted into electrical form by a microphone for transmission into space. This electrical form of sound is the signal. A signal can be analog or digital as shown in Fig. 13.7.

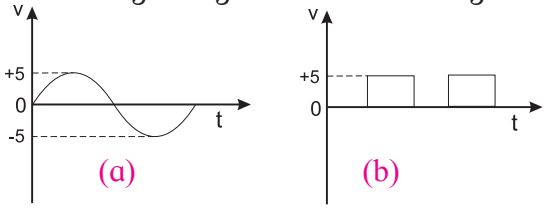


Fig 13.7: (a) Analog signal. (b) Digital signal.

(i) Analog signal: A continuously varying signal (voltage or current) is called an analog signal. Since a wave is a fundamental analog signal, sound and picture signals in TV are analog in nature (Fig 13.7 a)

(ii) Digital signal : A signal (voltage or current) that can have only two discrete values is called a digital signal. For example, a square wave is a digital signal. It has two values viz, +5 V and 0 V. (Fig- 13.7 b)

2) Transmitter :- A transmitter converts the signal produced by a source of information into a form suitable for transmission through a channel and subsequent reception.

3) Transducer :- A device that converts one form of energy into another form of energy is called a transducer. For example, a microphone converts sound energy into electrical energy. Therefore, a microphone is a transducer. Similarly, a loudspeaker is a transducer which converts electrical energy into sound energy.

4) Receiver :- The receiver receives the message signal at the channel output, reconstructs it in recognizable form of the original message for delivering it to the user of information.

5) Noise :- A random unwanted signal is called noise. The source generating the noise may be located inside or outside the system. Efforts should be made to minimise the noise level in a communication system.

6) Attenuation :- The loss of strength of the signal while propagating through the channel is known as attenuation. It occurs because the channel distorts, reflects and refracts the signals as it passes through it.

7) Amplification :- Amplification is the process of raising the strength of a signal, using an electronic circuit called amplifier.

8) Range :- The maximum (largest) distance between a source and a destination up to which the signal can be received with sufficient strength is termed as range.

9) Bandwidth :- The bandwidth of an electronic circuit is the range of frequencies over which it operates efficiently.

10) Modulation :- The signals in communication system (e.g. music, speech etc.) are low frequency signals and cannot be transmitted over large distances. In order to transmit the signal to large distances, it is superimposed on a high frequency wave (called carrier wave). This process is called modulation. Modulation is done at the transmitter and is an important part of a communication system.

11) Demodulation :- The process of regaining signal from a modulated wave is called demodulation. This is the reverse process of modulation.

12) Repeater :- It is a combination of a transmitter and a receiver. The receiver receives the signal from the transmitter, amplifies it and transmits it to the next repeater. Repeaters are used to increase the range of a communication system. These are shown in Fig. 13.8.

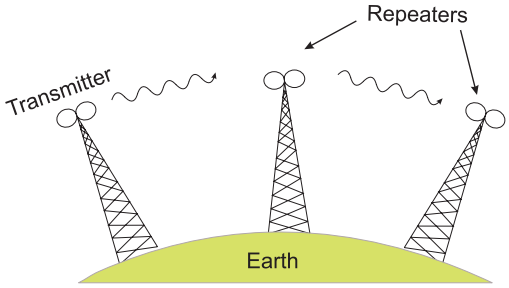


Fig.13.8: Use of repeater station to increase the range of communication.

Do you know ?

To transmit a signal we need an antenna or an aerial. For efficient transmission and reception, the transmitting and receiving antennas must have a length at least $\lambda/4$ where λ is the wavelength of the signal.

For an audio signal of 15kHz, the required length of the antenna is $\lambda/4$ which can be seen to be equal to 5 km.

The highest TV tower in Rameshwaram, Tamilnadu, is the tallest tower in India and is ranked 32nd in the world with pinnacle height of 323 metre. It is used for television broadcast by the Doordarshan.

13.6 Modulation:

As mentioned earlier, an audio signal has low frequency (< 20 KHz). Low frequency signals can not be transmitted over large distances. Because of this, a high frequency wave, called a carrier wave, is used. Some characteristic (e.g. amplitude, frequency or phase) of this wave is changed in accordance with the amplitude of the signal. This process is known as modulation. Modulation also helps avoid mixing up of signals from different transmitters as different carrier wave frequencies can be allotted to different transmitters. Without the use of these waves, the audio signals, if transmitted directly by different transmitters, would have got mixed up.

Modulation can be done by modifying the (i) amplitude (amplitude modulation) (ii) frequency (frequency modulation), and (iii) phase (phase modulation) of the carrier wave in proportion to the amplitude or intensity of the

signal wave keeping the other two properties same. Figure 13.9 (a) shows a carrier wave and (b) shows the signal. The carrier wave is a high frequency wave while the signal is a low frequency wave. Amplitude modulation, frequency modulation and phase modulation of carrier waves are shown in Fig. 13.9 (c), (d) and (e) respectively.

Amplitude modulation (AM) is simple to implement and has large range. It is also cheaper. Its disadvantages are that (i) it is not very efficient as far as power usage is concerned (ii) it is prone to noise and (iii) the reproduced signal may not exactly match the original signal. In spite of this, these are used for commercial broadcasting in the long, medium and short wave bands.

Frequency modulation (FM) is more complex as compared to amplitude modulation and, therefore is more difficult to implement. However, its main advantage is that it reproduces the original signal closely and is less susceptible to noise. This modulation is used for high quality broadcast transmission.

Phase modulation (PM) is easier than frequency modulation. It is used in determining the velocity of a moving target which cannot be done using frequency modulation.

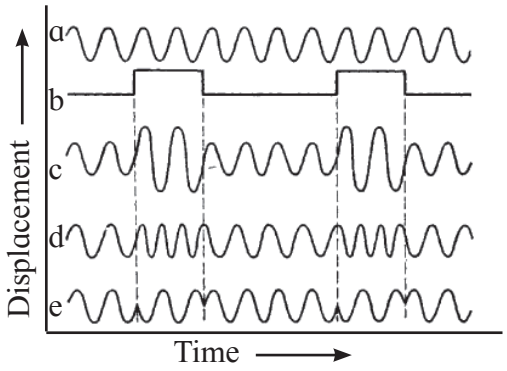


Fig. 13.9: (a) Carrier wave, (b) signal (c) AM (d) FM and (e) PM.

Internet my friend

https://www.iiap.res.in/centers/iao

Exercises

1. Choose the correct option.

i) The EM wave emitted by the Sun and responsible for heating the Earth's atmosphere due to green house effect is
(A) Infra-red radiation (B) X ray
(C) Microwave (D) Visible light

ii) Earth 's atmosphere is richest in
(A) UV (B) IR
(C) X-ray (D) Microwaves

iii) How does the frequency of a beam of ultraviolet light change when it travels from air into glass?
(A) No change (B) increases
(C) decreases (D) remains same

iv) The direction of EM wave is given by
(A) $\vec{E} \times \vec{B}$ (B) $\vec{E} \cdot \vec{B}$
(C) along $\vec{E}$ (D) along $\vec{B}$

v) The maximum distance upto which TV transmission from a TV tower of height h can be received is proportional to
(A) $h^{1/2}$ (B) h
(C) $h^{3/2}$ (D) h^2

vi) The waves used by artificial satellites for communication purposes are
(A) Microwave
(B) AM radio waves
(C) FM radio waves
(D) X-rays

vii) If a TV telecast is to cover a radius of 640 km, what should be the height of transmitting antenna?
(A) 32000 m (B) 53000 m
(C) 42000 m (D) 55000 m

2. Answer briefly.

i) State two characteristics of an EM wave.
ii) Why are microwaves used in radar?
iii) What are EM waves?
iv) How are EM waves produced?
v) Can we produce a pure electric or magnetic wave in space? Why?
vi) Does an ordinary electric lamp emit EM waves?
vii) Why do light waves travel in vacuum whereas sound wave cannot?

viii) What are ultraviolet rays? Give two uses.

ix) What are radio waves? Give its two uses.

x) Name the most harmful radiation entering the Earth's atmosphere from the outer space.

xi) Give reasons for the following:
(i) Long distance radio broadcast uses short wave bands.
(ii) Satellites are used for long distance TV transmission.

xii) Name the three basic units of any communication system.

xiii) What is a carrier wave?

xiv) Why high frequency carrier waves are used for transmission of audio signals?

xv) What is modulation?

xvi) What is meant by amplitude modulation?

xvii) What is meant by noise?

xviii) What is meant by bandwidth?

xix) What is demodulation?

xx) What type of modulation is required for television broadcast?

xxi) How does the effective power radiated by an antenna vary with wavelength?

xxii) Why should broadcasting programs use different frequencies?

xxiii) Explain the necessity of a carrier wave in communication.

xxiv) Why does amplitude modulation give noisy reception?

xxv) Explain why is modulation needed.

2. Solve the numerical problem.

i) Calculate the frequency in MHz of a radio wave of wavelength 250 m. Remember that the speed of all EM waves in vacuum is 3.0×10^{8} m/s.

[Ans: 1.2 MHz]

ii) Calculate the wavelength in nm of an X-ray wave of frequency 2.0×10^{18} Hz.

[Ans: 0.15 nm]

iii) The speed of light is 3×10^{8} m/s. Calculate the frequency of red light of wavelength of 6.5×10^{-7} m.

[Ans: $\upsilon = 4.6\times10^{14}$ Hz]

iv) Calculate the wavelength of a microwave of frequency 8.0 GHz.

[Ans: 3.75 cm]

v) In a EM wave the electric field oscillates sinusoidally at a frequency of 2×10^{10} Hz. What is the wavelength of the wave?

[Ans: 1.5×10^{-2} m]

vi) The amplitude of the magnetic field part of a harmonic EM wave in vacuum is $B_0 = 5\times10^{-7}$ T. What is the amplitude of the electric field part of the wave?

[Ans: 150V/m]

vii) A TV tower has a height of 200 m. How much population is covered by TV transmission if the average population density around the tower is 1000/km^2? (Radius of the Earth = 6.4×10^{6} m)

[Ans: 8×10^{6}]

viii) Height of a TV tower is 600 m at a given place. Calculate its coverage range if the radius of the Earth is 6400 km. What should be the height to get the double coverage area?

[Ans: 87.6 km, 1200 m]

ix) A transmitting antenna at the top of a tower has a height 32 m and that of the receiving antenna is 50 m. What is the maximum distance between them for satisfactory communication in line of sight mode ? Given radius of Earth is 6.4×10^{6} m.

[Ans: 45.537 km]

***

14. Semiconductors

Can you recall?

1. Your mobile handset is very efficient gadget.
2. International Space Station works using solar energy.
3. A LED TV screen produces brighter and vivid colours.
4. Good and bad conductor of electricity.

14.1 Introduction:

Modern life is heavily dependent on many electronic gadgets. It could be a cell phone, a smart watch, a computer or even an LED lamp, they all have one common factor, semiconductor devices that make them work. Semiconductors have made our life very comfortable and easy.

Semiconductors are materials whose electrical properties can be tailored to suit our requirements. Before the discovery of semiconductors, electrical properties of materials could be of two types, conductors or insulators. Conductors such as metals have a very high electrical conductivity, for example, conductivity of silver is 6.25×10^7 Sm^{-1} whereas an insulator or a bad conductor like glass has a very low electrical conductivity of the order of 10^{-10} Sm^{-1}. Electrical conductivity of silicon, a semiconductor, for example is 1.56×10^{-3} Sm^{-1}. It lies between that of a good conductor and a bad conductor. A semiconductor can be customised to have its electrical conductivity as per our requirement. Temperature dependence of electrical conductivity of a semiconductor can also be controlled. Table 14.1 gives electrical conductivity of some materials which are commonly used.

14.2 Electrical conduction in solids:

Electrical conduction in a solid takes place by transport of charge carriers. It depends on its temperature, the number of charge carriers, how easily these carries can move inside a solid (mobility), its crystal structure, types and the nature of defects present in a solid etc. There can be three types of electrical conductors. It could be a good conductor, a semiconductor or a bad conductor.

1. Conductors (Metals): The best example of a conductor is any metal. They have a large number of free electrons available for electrical conduction. (A typical metal will have 10^{28} electrons per m^3). Metals are good conductors of electricity due to the large number of free electrons present in them.

2. Insulators: Glass, wood or rubber are some common examples of insulators. Insulators have very small number (10^{23} per m^3) of free electrons.

3. Semiconductors: Silicon, germanium, gallium arsenide, gallium nitride, cadmium sulphide are some of the commonly used semiconductors. The electrical conductivity of a semiconductor is between the conductivity of a metal and that of an insulator. The number of charge carriers in a semiconductor can be controlled as per our requirement. Their structure can also be designed to suit our requirement. Such materials are very useful in electronic industry and find applications in almost every gadget of daily use such as a cell phone, a solar cell or a complex system such as a satellite or the International Space Station.

Table 14.1: Electrical conductivities of some commonly used materials

Material	Electrical conductivity (S m^{-1})
Silver	6.30×10^7
Copper	5.96×10^7
Aluminium	3.5×10^7
Gold	4.10×10^7
Nichrome	9.09×10^5
Platinum	9.43×10^6
Germanium	2.17
Silicon	1.56×10^{-3}
Air	3×10^{-15} to 8×10^{-15}
Glass	10^{-11} to 10^{-15}
Teflon	10^{-25} to 10^{-23}
Wood	10^{-16} to 10^{-24}

Do you know ?

Electrical conductivity σ of a solid is given by $\sigma = nq\mu$, where,

n = charge carrier density (number of carriers per unit volume)

q = charge on the carriers

μ = mobility of carriers

Mobility of a charge carrier is the measure of the ease with which a carrier can move in a material under the action of an external electric field. It depends upon many factors such as mass of the carrier, whether the material is crystalline or amorphous, the presence of structural defects in a material, the nature of impurities in a material and so on.

Figure 14.1 shows the temperature dependence of the electrical conductivity of a typical metal and a semiconductor. When the temperature of a semiconductor is increased, its electrical conductivity also increases. The electrical conductivity of a metal decreases with increase in its temperature.

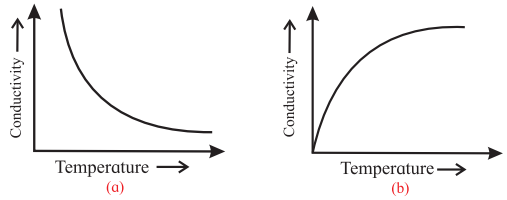


Fig. 14.1: Temperature dependence of electrical conductivity of (a) metals and (b)semiconductors.

Variation of electrical conductivity of semiconductors with change in its temperature is a very useful property and finds applications in a large number of electronic devices. A broad classification of semiconductors can be:

a. Elemental semiconductors: Silicon, germanium

b. Compound Semiconductors: Cadmium sulphide, zinc sulphide, etc.

c. Organic Semiconductors: Anthracene, doped pthalocyanines, polyaniline etc.

Elemental semiconductors and compound semiconductors are widely used in electronic industry. Discovery of organic semiconductors is relatively new and they find lesser applications.

Electrical properties of semiconductors are different from metals and insulators due to their unique conduction mechanism. The electronic configuration of the elemental semiconductors silicon and germanium plays a very important role in their electrical properties. They are from the fourth group of elements in the periodic table. They have a valence of four. Their atoms are bonded by covalent bonds. At absolute zero temperature, all the covalent bonds are completely satisfied in a single crystal of pure silicon or germanium.

The conduction mechanism in a semiconductor can be better understood with the help of the band theory of solids.

14.3 Band theory of solids, a brief introduction:

We begin with the way electron energies in an isolated atom are distributed. An isolated atom has its nucleus at the center which is surrounded by a number of revolving electrons. These electrons are arranged in different and discrete energy levels.

When a solid is formed, a large number of atoms are packed in it. The outermost electronic energy levels in a solid are occupied by electrons from all atoms in a solid. Sharing of the outermost energy levels and resulting formation of energy bands can be easily understood by considering formation of solid sodium.

The electronic configuration of sodium (atomic number 11) is $1s^2$, $2s^2$, $2p^6$, $3s^1$. The outermost level 3s can take one more electron but it is half filled in sodium.

When solid sodium is formed, atoms interact with each other through the electrons in each atom. The energy levels are filled according to the Pauli's exclusion principle. According to this principle, no two electrons can have the same set of quantum numbers, or in simple words, no two electrons with similar spin can occupy the same energy level.

Any energy level can accommodate only two electrons (one with spin up state and the other with spin down state). According to this principle, there can be two states per energy level. Figure 14.2 (a) shows the allowed energy

levels of an isolated sodium atom by horizontal lines. The curved lines represent the potential energy of an electron near the nucleus due to Coulomb interaction.

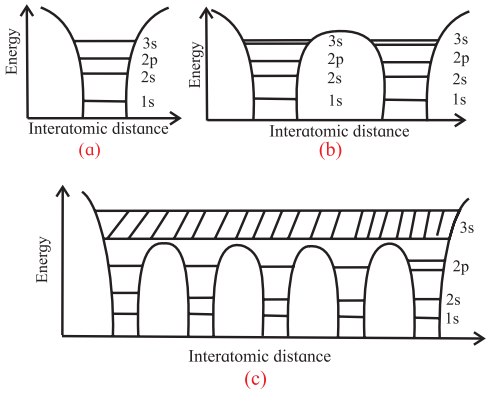


Fig. 14.2: Potential energy diagram, energy levels and bands (a) isolated atom, (b) two atoms, (c) sodium metal.

Consider two sodium atoms close enough so that outer 3s electrons are equally likely to be on any atom. The 3s electrons from both the sodium atoms need to be accommodated in the same level. This is made possible by splitting the 3s level into two sub-levels so that the Pauli's exclusion principle is not violated. Figure 14.2 (b) shows the splitting of the 3s level into two sub levels. When solid sodium is formed, the atoms come close to each other (distance between them ~ 2 - 3Å). Therefore, the electrons from different atoms interact with each other and also with the neighbouring atomic cores. The interaction between the outer most electrons is more due to overlap while the inner core electrons remain mostly unaffected. Each of these energy levels is split into a large number of sub levels, of the order of Avogadro's number. This is because the number of atoms in solid sodium is of the order of this number. The separation between the sublevels is so small that the energy levels appear almost continuous. This continuum of energy levels is called an energy band. The bands are called 1s band, 2s band, 2p band and so on. Figure 14.2 c shows these bands in sodium metal. Broadening of valence and higher bands is more because of stronger interaction of these electrons.

For sodium atom, the topmost occupied energy level is the 3s level. This level is called the valence level. Corresponding energy band is called the **valence band**. Thus, the valence band in solid sodium is the topmost occupied energy band. The valence band is half filled in sodium. Figure 14.3 shows the energy bands in sodium.

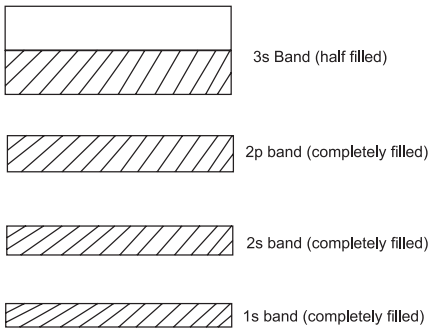


Fig. 14.3: Energy bands in sodium.

When sufficient energy is provided to electrons from the valence band they are raised to higher levels. The immediately next energy level that electrons from valence band can occupy is called conduction level. The band formed by conduction levels is called **conduction band**. In sodium valence and conduction bands overlap.

In a semiconductor or an insulator, there is a gap between the bottom of the conduction band and the top of the valence band. This is called the energy gap or the band gap.

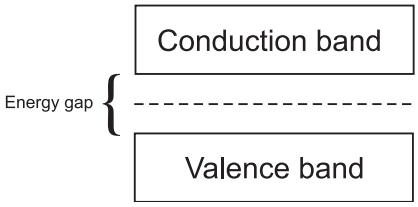


Fig. 14.4: Energy bands for a typical solid.

Figure 14.4 shows the conduction band, the energy gap and the valence band for a typical solid which is not a good conductor. **It is important to remember that this structure is related to the energy of electrons in a solid and it does not represent the physical structure of a solid in any way.**

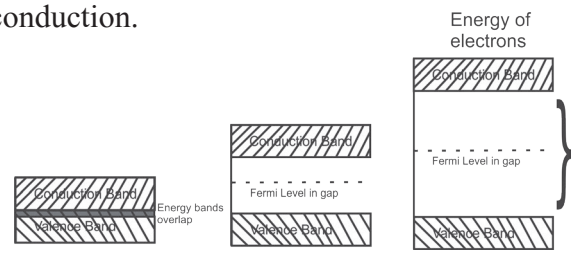
All the energy levels in a band, including the topmost band, in a semiconductor are completely occupied at absolute zero. At some finite temperature T, few electrons gain thermal energy of the order of kT, where k is the Boltzmann constant.

Electrons in the bands below the valence band cannot move to higher band since these are already occupied. Only electrons from the valence band can be excited to the empty

Formation of energy bands in a solid is a result of the small distances between atoms, the resulting interaction amongst electrons and the Pauli's exclusion principle.

conduction band, if the thermal energy gained by these electrons is greater than the band gap. In case of sodium, electrons from the 3s band can gain thermal energy and occupy a slightly higher energy level because the 3s band is only half filled.

Electrons can also gain energy when an external electric field is applied to a solid. Energy gained due to electric field is smaller, hence only electrons at the topmost energy level gain such energy and participate in electrical conduction.



(a) (b) (c)

Fig. 14.5: Band structure of a (a) metal, (b) semiconductor, and an insulator (c).

The difference in electrical conductivities of various solids can be explained on the basis of the band structure of solids. Band structure in a metal, semiconductor and an insulator is different. Figure 14.5 shows a schematic representation of band structure of a metal, a semiconductor and an insulator.

For metals, the valence band and the conduction band overlap and there is no band gap as shown in Fig.14.5 (a). Electrons, therefore, find it easy to gain electrical energy when some external electric field is applied. They are, therefore, easily available for conduction.

In case of semiconductors, the band gap is fairly small, of the order of one electron volt or less as shown in Fig.14.5 (b). When excited, electrons gain energy and occupy energy levels in conduction band easily and can take part in electric conduction.

Insulators, on the contrary, have a wide gap between valence band and conduction band as shown in Fig.14.5 (c). Diamond, for example, has a band gap of about 5.0 eV. In an insulator, therefore, electrons find it very difficult to gain sufficient energy and occupy energy levels in the conduction band.

The magnitude of the band gap plays a very important role in electronic properties of a solid.

Table 14.2: Magnitude of energy gap in silicon, germanium and diamond.

Material	Energy gap (eV) At 300 K
Silicon	1.12
Germanium	0.66
Diamond	5.47

1 eV is the energy gained by an electron while it overcomes a potential difference of one volt. 1 ev = 1.6 × 10^{-19} J.

14.4 Intrinsic Semiconductor:

A pure semiconductor such as pure silicon or pure germanium is called an intrinsic semiconductor. Silicon (Si) has atomic number 14 and its electronic configuration is $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$. Its valence is 4. Each atom of Si forms four covalent bonds with its neighbouring atoms. One Si atom is surrounded by four Si atoms at the corners of a regular tetrahedron Fig. 14.6.

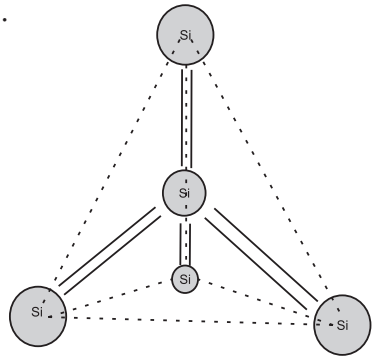


Fig. 14.6: Structure of silicon.

At absolute zero temperature, all valence electrons are tightly bound to respective atoms and the covalent bonds are complete. Electrons are not available to conduct electricity through the crystal because they cannot gain enough energy to get into higher energy levels. At room temperature, however, a few covalent bonds are broken due to thermal agitation and some valence electrons can gain energy. Thus we can say that a valence electron is moved to the conduction band. It creates a vacancy in the valence band as shown in Fig. 14.7.

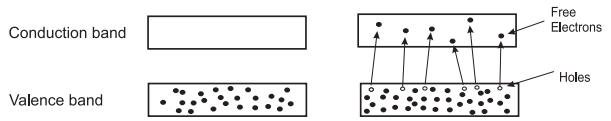


Fig. 14.7: Creation of vacancy in the valence band.

These vacancies of electrons in the valence band are called holes. The holes are thus absence of electrons in the valence band and they carry an *effective* positive charge.

For an intrinsic semiconductor, the number of holes per unit volume, (the number density, n_h) and the number of free electrons per unit volume, (the number density, n_e) is the same.

$$n_h = n_e$$

Electric conduction through an intrinsic semiconductor is quite interesting. There are two different types of charge carriers in a semiconductor. One is the electron and the other is the hole or absence of electron. Electrical conduction takes place by transportation of both carriers or any one of the two carriers in a semiconductor. When a semiconductor is connected in a circuit, electrons, being negatively charged, move towards positive terminal of the battery. Holes have an *effective positive charge*, and move towards negative terminal of the battery. Thus, the current through a semiconductor is carried by two types of charge carriers which move in opposite directions. This conduction mechanism makes semiconductors very useful in designing a large number of electronic devices. Figure 14.8 represents the current through a semiconductor.

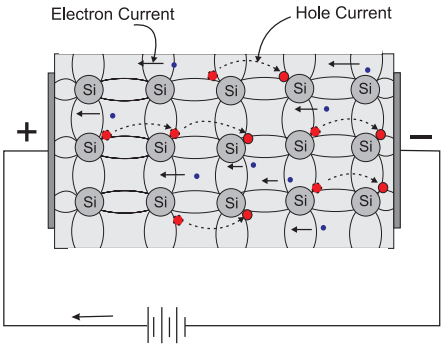


Fig. 14.8: Current through a semiconductor, transport of electrons and holes.

14.5 Extrinsic semiconductors:

The electric conductivity of an intrinsic semiconductor is very low at room temperature; hence no electronic devices can be fabricated using them. Addition of a small amount of a suitable impurity to an intrinsic semiconductor increases its conductivity appreciably. **The process of adding impurities to an intrinsic semiconductor is called doping.** The semiconductor with impurity is called a doped semiconductor or an extrinsic semiconductor. The impurity is called the dopant. The parent atoms are called hosts. The dopant material is so selected that it does not disturb the crystal structure of the host. The size and the electronic configuration of the dopant should be compatible with that of the host. Silicon or germanium can be doped with a pentavalent impurity such as phosphorus (P) arsenic (As) or antimony (Sb) . They can also be doped with a trivalent impurity such as boron (B) aluminium (Al) or indium (In).

Addition of pentavalent or trivalent impurities in intrinsic semiconductors gives rise to different conduction mechanisms. This is very useful in designing many electronic devices. Extrinsic semiconductors can be of two types a) n-type semiconductor or b) p-type semiconductor.

a) n-type semiconductor: When silicon or germanium crystal is doped with a pentavalent impurity such as phosphorus, arsenic, or antimony we get n-type semiconductor. Figure 14.9 shows the schematic electronic structure of antimony.

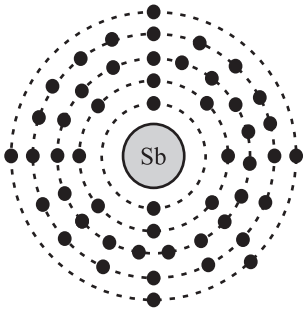


Fig. 14.9: Schematic electronic structure of antimony.

When a dopant atom of 5 valence electrons occupies the position of a Si atom in the crystal lattice, 4 electrons from the dopant form bonds with 4 neighbouring Si atoms and the fifth electron from the dopant remains very weakly bound to its parent atom. Figure 14.10 shows a pentavalent impurity in silicon lattice.

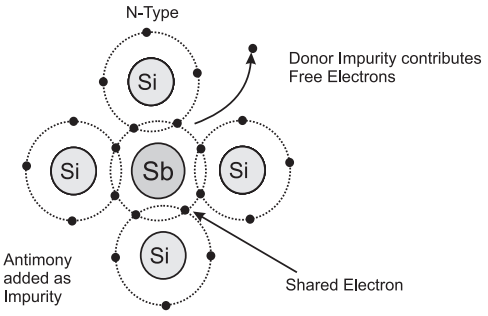


Fig. 14.10: Pentavalent impurity in silicon crystal.

To make this electron free even at room temperature, very small energy is required. It is 0.01 eV for Ge and 0.05 eV for Si.

Do you know ?

One cm^3 specimen of a metal or semiconductor has of the order of 10^{22} atoms. In a metal, every atom donates at least one free electron for conduction, thus 1 cm^3 of metal contains of the order of 10^{22} free electrons, whereas 1 cm^3 of pure germanium at 20 °C contains about 4.2×10^{22} atoms, but only 2.5×10^{13} free electrons and 2.5×10^{13} holes. Addition of 0.001% of arsenic (an impurity) donates 10^{17} extra free electrons in the same volume and the electrical conductivity is increased by a factor of 10,000.

Since every pentavalent dopant atom donates one electron for conduction, it is called a donor impurity. As this semiconductor has large number of electrons in conduction band and its conductivity is due to negatively charged carriers, it is called n-type semiconductor. The n-type semiconductor also has a few electrons and holes produced due to the thermally broken bonds. The density of conduction electrons (n_e) in a doped semiconductor is the sum total of the electrons contributed by donors and the thermally generated electrons from the host. The density of holes (n_h) is only due to the thermal breakdown of some covalent bonds of the host Si atoms. Some electrons and holes recombine continuously because they carry opposite charges. The number of free electrons exceeds the number of holes. Thus, in a semiconductor doped with pentavalent impurity, electrons (negative charge) are the majority carriers and holes are the minority carriers. Therefore, it is called n-type semiconductor. **For n-type semiconductor, $n_e >> n_h$.**

The free electrons donated by the impurity atoms occupy energy levels which are in the band gap and are close to the conduction band. They can be easily available for conduction. Figure 14.11 shows the schematic band structure of an n-type semiconductor.

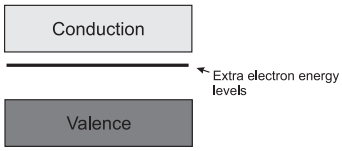


Fig.14.11: Schematic band structure of an n-type semiconductor.

Extrinsic semiconductors are thus far better conductors than intrinsic semiconductors. The conductivity of an extrinsic semiconductor can be controlled by controlling the amount of impurities added. The amount of impurities is expressed as part per million or ppm, that is, one impurity atom per one million atoms of the host.

Features of n-type semiconductors: These are materials doped with pentavalent impurity (donors) atoms . Electrical conduction in these materials is due to electrons as majority charge carriers.

1. The donor atom lose electrons and become positively charged ions.
2. Number of free electrons is very large compared to the number of holes, $n_e >> n_h$. Electrons are majority charge carriers.
3. When energy is supplied externally, negatively charged free electrons (majority charges carries) and positively charged holes (minority charge carriers) are available for conduction.

b) p-type semiconductor: When silicon or germanium crystal is doped with a trivalent impurity such as boron, aluminium or indium, we get a p-type semiconductor. Figure 14.12 shows the schematic electronic structure of boron.

The dopant trivalent atom has one valence electron less than that of a silicon atom. Every trivalent dopant atom shares its three electrons with three neighbouring Si atoms to form

covalent bonds. But the fourth bond between silicon atom and its neighbour is not complete.

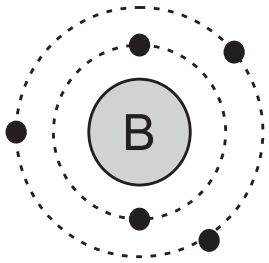


Fig. 14.12: Schematic electronic structure of boron.

Figure 14.13 shows a trivalent impurity in a silicon crystal. The incomplete bond can be completed by another electron in the neighbourhood from Si atom. Since each donar trivalent atom can accept an electron, it is called an acceptor impurity. The shared electron creates a vacancy in its place. This vacancy or the absence of electron is a hole.

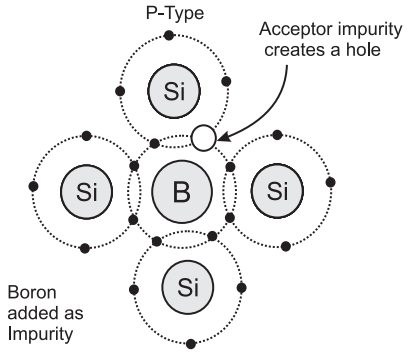


Fig. 14.13: A trivalent impurity in a silicon crystal.

Thus, a hole is available for conduction from each acceptor impurity atom. Holes are majority carriers and electrons are minority carriers in such materials. Acceptor atoms are negatively charged and majority carriers are holes (positively charged). Therefore, extrinsic semiconductor doped with trivalent impurity is called a p-type semiconductor. For a p-type semiconductor, $n_h >> n_e$.

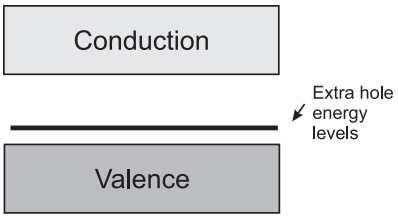


Fig. 14.14: Schematic band structure of a p-type semiconductor.

These vacancies of electrons are created in the valence band; therefore we can say that the holes are created in the valence band. The impurity levels are created just above the valence band in the band gap. Electrons from valence band can easily occupy these levels and conduct electricity. Figure 14.14 shows the schematic band structure of a p-type semiconductor.

Features of p-type semiconductor: These are materials doped with trivalent impurity atoms (acceptors). Electrical conduction in these materials is due to holes as majority charge carriers.

1. The acceptor atoms acquire electron and become negatively charged-ions.
2. Number of holes is very large compared to the number of free electrons. ($n_h \gg n_e$). Holes are majority charge carriers.
3. When energy is supplied externally, positively charged holes (majority charge carriers) and negatively charged free electrons (minority charge carriers) are available for conduction.

c) Charge neutrality of extrinsic semiconductors: The n-type semiconductor has excess of electrons but these extra electrons are supplied by the donor atoms which become positively charged. Since each atom of donor impurity is electrically neutral, the semiconductor as a whole is electrically neutral. Here, excess electron refers to an excess with reference to the number of electrons needed to complete the covalent bonds in a semiconductor crystal. These extra free electrons increase the conductivity of the semiconductor.

Similarly, a p-type semiconductor has holes or absence of electrons in some energy levels. When an electron from a host atom fills this level, the host atom is positively charged and the dopant atom is negatively charged but the semiconductor as a whole is electrically neutral. *Thus, n-type as well as p-type semiconductors are electrically neutral.*

Always remember, for a semiconductor,

$$n_e \cdot n_h = n_i^2$$

Example 14.1: A pure Si crystal has 4×10^{28} atoms m^{-3}. It is doped by 1ppm concentration of antimony. Calculate the number of electrons and holes. Given $n_i = 1.2 \times 10^{16}/m^3$.

Solution: 1 ppm = 1 part per million = $1/10^6$

$\therefore$ no. of Sb atoms $= \frac{4\times10^{28}}{10^6} = 4\times10^{22}$

As one pentavalent impurity atom donates one free electron to the crystal,

Number of free electrons in the crystal $n_e = 4 \times 10^{22}$ m^{-3}

Number of holes,

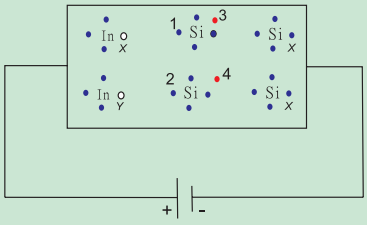
$$n_h = \frac{(n_i)^2}{n_e} = \frac{(1.2\times10^{16})^2}{4\times10^{22}}$$

$n_h = 3.6 \times 10^9$ m^{-3}

Do you know ?

Transportation of holes

Consider a p-type semiconductor connected to terminals of a battery as shown. When the circuit is switched on, electrons at 1 and 2 are attracted to the positive terminal of the battery and occupy nearby holes at x and y. This generates holes at the positions 1 and 2 previously occupied by electrons. Next, electrons at 3 and 4 move towards the positive terminal and create holes in the positions they occupied previously.



Finally, the hole is captured at the negative terminal by the electron supplied by the battery at that end. This keeps the density of holes constant and maintains the current so long as the battery is working.

Thus, physical transportation is of the electrons only. However, we feel that the holes are moving towards the negative terminal of the battery. Positive charge is attracted towards negative terminal. Thus holes, which are not actual charges, behave like a positive charge. In this case, there is an indirect movement of electrons and their drift speed is less than that in the n-type semiconductors. The mobility of holes is, therefore, less than that of the electrons.

Example 14.2: A pure silicon crystal at temperature of 300 K has electron and hole concentration 1.5×10^{16} m^{-3} each. ($n_e = n_h$). Doping by indium increases n_h to 4.5×10^{22} m^{-3}. Calculate n_e for the doped silicon crystal.

Solution: We know,

$n_e n_h = n_i^2$ and $n_e = \frac{(n_i)^2}{n_h}$

Given

$n_i = 1.5 \times 10^{16}$m^{-3} and $n_h = 4.5 \times 10^{22}$ m^{-3}

$n_e = \frac{(1.5\times10^{16})^2}{4.5\times10^{22}} = 5 \times 10^9$ m^{-3}

14.6 p-n junction:

When n-type and p-type semiconductor materials are fused together, a p-n junction is formed. A p-n junction shows many interesting properties and it is the basis of almost all modern electronic devices. Figure 14.15 shows a schematic structure of a p-n junction.

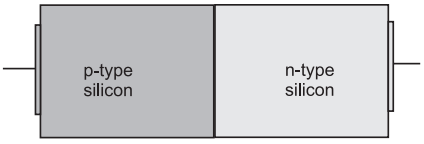


Fig. 14.15: Schematic structure of a p-n junction.

Diffision: When n-type and p-type semiconductor materials are fused together, initially, the number of electrons in the n-side of the junction is very large compared to the number of electrons on the p-side. The same is true for the number of holes on the p-side and on the n-side. Thus, the density of carriers on both sides is different and a large density gradient exists on both sides of the p-n junction. This density gradient causes migration of electrons from the n-side to the p-side of the junction. They fill up the holes in the p-type material and produce negative ions.

When the electrons from the n-side of a junction migrate to the p-side, they leave behind positively charged donor ions on the n-side. Effectively, holes from the p-side migrate into the n-region.

As a result, in the p-type region near the junction there are negatively charged acceptor ions, and in the n-type region near the junction there are positively charged donor ions. The transfer of electrons and holes across the p-n

junction is called diffusion. The extent up to which the electrons and the holes can diffuse across the junction depends on the density of the donor and the acceptor ions on the n-side and the p-side respectively, of the junction. Figure 14.16 shows the diffusion of charge carriers across the junction.

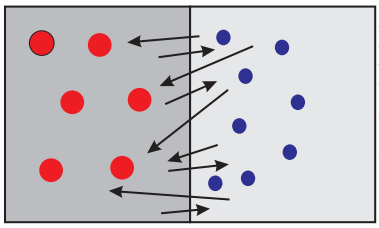


Fig. 14.16: Diffusion of charge carriers across a junction.

Depletion region: The diffusion of carriers across the junction and resultant accumulation of positive and negative charges across the junction builds a potential difference across the junction. This potential difference is called the potential barrier. The magnitude of the potential barrier for silicon is about 0.6 – 0.7 volt and for germanium, it is about 0.3 – 0.35 volt. This potential barrier always exists even if the device is not connected to any external power source. It prevents continuous diffusion of carriers across the junction. A state of electrostatic equilibrium is thus reached across the junction.

Free charge carriers cannot be present in a region where there is a potential barrier. The regions on either side of a junction, therefore, becomes completely devoid of any charge carriers. This region across the p-n junction where there are no charges is called the depletion layer or the depletion region. Figure 14.17 shows the potential barrier and the depletion layer.

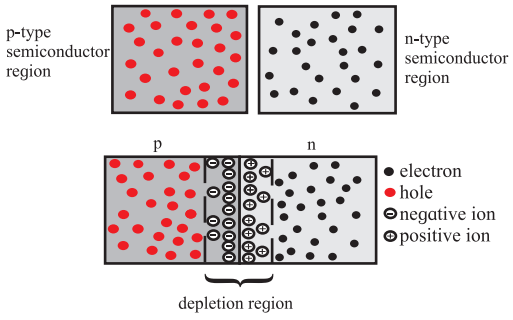


Fig. 14.17: Potential barrier and the depletion layer.

The potential across a junction and width of the potential barrier can be controlled. This is very interesting and useful property of a p-n junction.

The n-side near the boundary of a p-n junction becomes positive with respect to the p-side because it has lost electrons and the p-side has lost holes. Thus the presence of impurity ions on both sides of the junction establishes an electric field across this region such that the n-side is at a positive voltage relative to the p-side. Figure 14.18 shows the electric field thus produced.

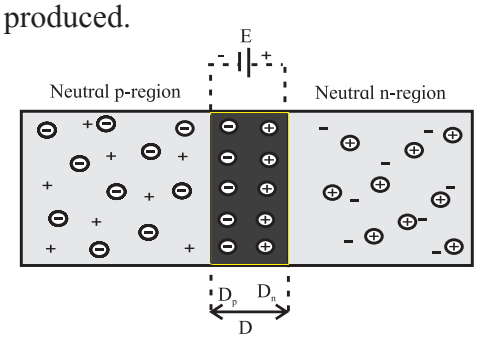


Fig. 14.18: Electric field across a junction.

Biasing a p-n junction: As a result of potential barrier across depletion region, charge carriers require some extra energy to overcome the barrier. A suitable voltage needs to be applied to the junction externally, so that these charge carriers can overcome the potential barrier and move across the junction. Figure 14.19 shows two possibilities of applying this external voltage across the junction.

Figure 14.19 (a) shows a p-n junction connected in an electric circuit where the p-region is connected to the positive terminal and the n-region is connected to the negative terminal of an external voltage source. This external voltage effectively opposes the built-in potential of the junction. The width of potential barrier is thus reduced. Also, negative charge carriers (electrons) from the n-region are pushed towards the junction. A similar effect is experienced by positive charge carriers (holes) in the p-region and they are pushed towards the junction. Both the charge carriers thus find it easy to cross over the barrier and contribute towards the electric current. Such arrangement of a p-n junction in an electric circuit is called **forward bias**.

Figure 14.19 (b) shows the other possibility, where, the p-region is connected to the negative terminal and the n-region is connected to the positive terminal of the external voltage source. This external voltage effectively adds to the

built-in potential of the junction. The width of potential barrier is thus increased. Also, the negative charge carriers (electrons) from the n-region are pulled away from the junction. Similar effect is experienced by the positive charge carriers (holes) in the p-region and they are pulled away from the junction. Both the charge carriers thus find it very difficult to cross over the barrier and thus do not contribute towards the electric current. Such arrangement of a p-n junction in an electric circuit is called **reverse bias**.

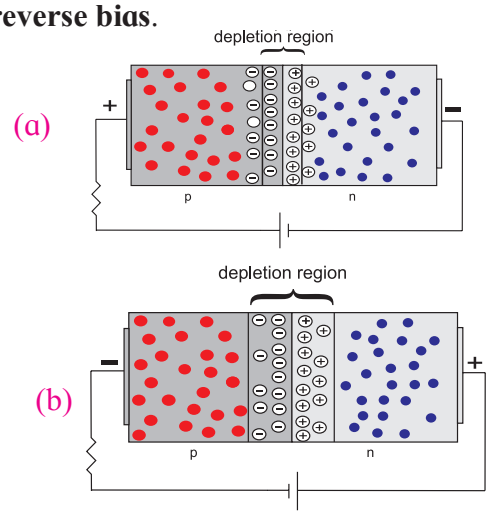


Fig. 14.19: Forward biased (a) and reverse biased (b) junction.

Therefore, when used in forward bias mode, a p-n junction allows a large current to flow across. This current is normally of the order of a few milliamperes, (10^{-3} A). A reverse biased p-n junction on the other hand, carries a very small current that is normally a few microamperes (10^{-6} A).

A p-n junction can be thus used as a one way switch or a gate in an electric circuit. It conducts easily in forward bias and acts as an open switch in reverse bias.

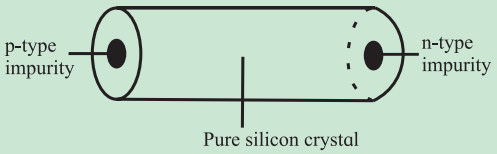
Features of the depletion region:

1. It is formed by diffusion of electrons from n-region to the p-region. This leaves positively charged ions in the n-region.
2. The p-region accumulates electrons (negative charges) and the n-region accumulates the holes (positive charges).
3. The accumulation of charges on either sides of the junction results in forming a potential barrier and prevents flow of charges across it.
4. There are no charges in this region.
5. The depletion region has higher potential on the n-side and lower potential on the p-side of the junction.

Do you know ?

Fabrication of p-n junction diode:

It was mentioned previously, for easy understanding, that a p-n junction is formed by fusing a p-type and a n-type material together. However, in practice, a p-n junction is formed from a crystalline structure of silicon or germanium by adding carefully controlled amounts of donor and acceptor impurities.



The impurities grow on either side of the crystal after heating in a furnace. Electrons and holes combine at the center and the depletion region develops. A junction is thus formed. Electrodes are inserted after cutting transverse sections and hundreds of diodes are prepared. All semiconductor devices, including ICs, are fabricated by 'growing' junctions at the required locations.

Mobility of a hole is less than that of an electron and the hole current is lesser. This imbalance between the two currents is removed by increasing the doping percentage in the p-region. This ensures that the same current flows through the p-region and the n-region of the junction.

14.7 A p-n junction diode:

A p-n junction, when provided with metallic connectors on each side is called a junction diode or simply, a diode. (Diode is a device with two electrodes or di-electrodes). Figure 14.20 shows the circuit symbol for a junction diode.

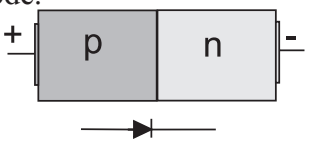


Fig. 14.20: Circuit symbol for a p-n junction diode.

The 'arrow' indicates the direction of the conventional current. The p-side is called the anode and the n-side is called the cathode of the diode. When a diode is connected across a battery, the carriers can gain additional energy to cross the barrier as per biasing.

A diode can be connected across a battery in two different ways, forward bias and reverse bias as shown in the (Fig. 14.21).

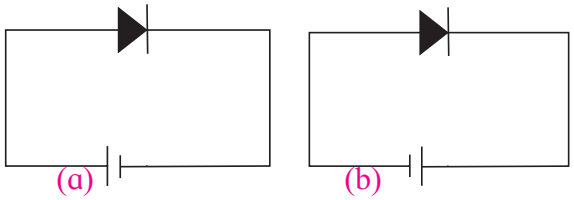


Fig. 14.21: (a) Forward bias, (b) Reverse bias.

The behavior of a diode in both cases is different. This is because the barrier potential is affected differently in the two cases. The barrier potential is reduced in forward biased mode and it is increased in reverse biased mode.

Carriers find it easy to cross the junction in forward bias and contribute towards current for two reasons; first the barrier width is reduced and second, they are pushed towards the junction and gain extra energy to cross the junction. The current through the diode in forward bias is, therefore, large. It is of the order of a few milliamperes (10^{-3} A) for a typical diode.

When connected in reverse bias, width of the potential barrier is increased and the carriers are pushed away from the junction so that very few thermally generated carriers can cross the junction and contribute towards current. This results in a very small current through a reverse biased diode. The current in reverse biased diode is of the order of a few microamperes (10^{-6} A).

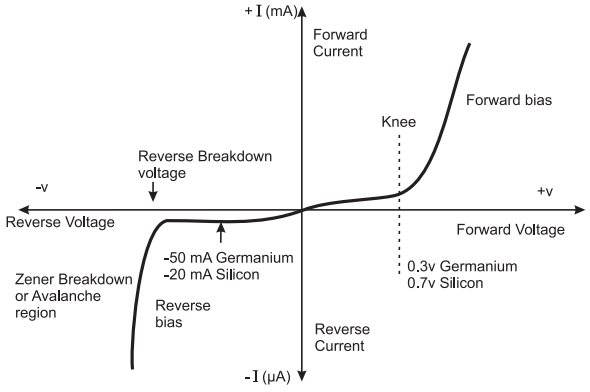


Fig. 14.22: Asymmetrical current flow through a diode.

The width of the depletion layer decreases with an increase in the application of a forward voltage. It increases when a reverse voltage is applied. We have discussed the reasons for this difference earlier. When the polarity of bias voltage is reversed, the width of the depletion layer changes. This results in asymmetrical current flow through a diode as shown in (Fig. 14.22).

A diode can be thus used as a one way switch in a circuit. It is forward biased when its anode is connected to be at a higher potential than that of the cathode. When the anode is at lower potential than that of the cathode, it is reverse biased. A diode can be zero biased if no external voltage is applied across it.

a) **Forward biased:** The positive terminal of the external voltage is connected to the anode (p-side) and negative terminal to the cathode (n-side) across the diode.

In case of forward bias, the width of the depletion region decreases and the p-n junction offers a low resistance path allowing a high current to flow across the junction (Fig. 14.23).

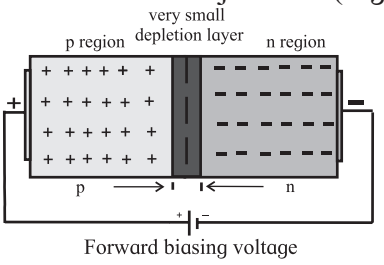


Fig. 14.23: Decrease in width of depletion region.

Figure 14.24 shows the I-V characteristic of a forward biased diode. Initially, the current is very low and then there is a sudden rise in the current. The point at which current rises sharply is shown as the 'knee' point on the I-V characteristic curve. The corresponding voltage is called the 'knee voltage'. It is about 0.7 V for silicon and 0.3 V for germanium.

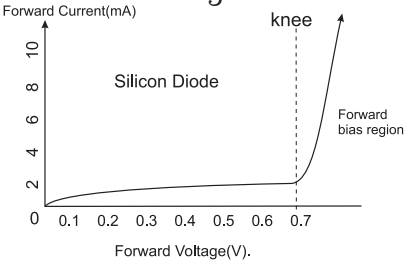


Fig. 14.24: I-V characteristic of a forward biased diode.

A diode effectively becomes a short circuit above this knee point and can conduct a very large current. Resistors are, therefore, used in series with diode to limit its current flow. If the current through a diode exceeds the specified value, it can heat up the diode due to the Joule heating and can result in its physical damage.

b) Reverse biased: The positive terminal of the external voltage is connected to the cathode (n-side) and negative terminal to the anode (p-side) across the diode. In case of reverse bias, the width of the depletion region increases and the p-n junction behaves like a high resistance (Fig. 14.25). Practically, no current flows through it with an increase in the reverse bias voltage. However, a very small **leakage current** does flow through the junction which is of the order of a few micro-amperes, (μA).

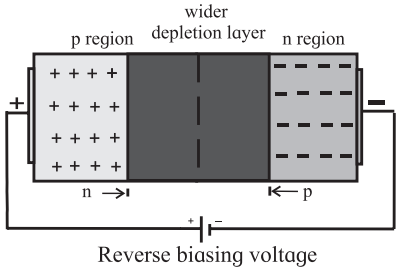


Fig. 14.25: Increase in width of depletion region.

When the reverse bias voltage applied to a diode is increased to sufficiently large value, it causes the p-n junction to overheat. The overheating of the junction results in a sudden rise in the current through the junction. This is because the covalent bonds break and a large number of carriers are available for conduction. The diode, thus, no longer behaves like a diode. This effect is called the avalanche breakdown. The reverse biased characteristic of a diode is shown in Fig 14.26.

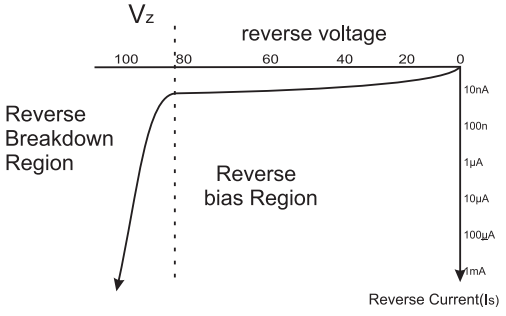
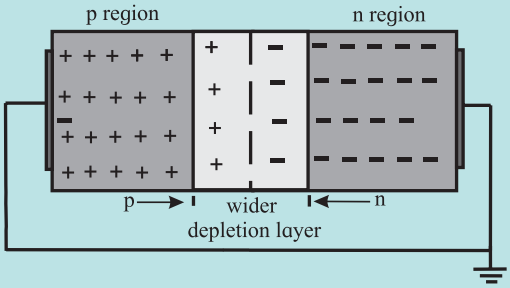


Fig. 14.26: Reverse biased characteristic of a diode.

Zero Biased Junction Diode.

When a diode is connected in a **zero bias** condition, no external potential energy is applied to the p-n junction. When the diode terminals are shorted together, some holes (majority carriers) in the p-side have enough thermal energy to overcome the potential barrier. Such carriers cross the barrier potential and contribute to current. This current is known as the **forward current.** .

Similarly, some holes generated in the n-side (minority carriers), also move across the junction in the opposite direction and contribute to current. This current is known as the **reverse current.** This transfer of electrons and holes back and forth across the p-n junction is known as diffusion, as discussed previously.



Zero biased p-n junction diode

The potential barrier that exists in a junction prevents the diffusion of any more majority carriers across it. However, some minority carriers (few free electrons in the p-region and few holes in the n-region) do drift across the junction.

An equilibrium is established when the majority carriers are equal in number (n_e=n_h) and are moving in opposite directions. The net current flowing across the junction is zero. This is a state of '**dynamic equilibrium**'.

Minority carriers are continuously generated due to thermal energy. When the temperature of the p-n junction is raised, this state of equilibrium is changed. This results in generating more minority carriers and an increase in the leakage current. An electric current, however, cannot flow through the diode because it is not connected in any electric circuit.

c) Static and dynamic resistance of a diode:

One of the most important properties of a diode is its resistance in the forward biased mode and in the reverse biased mode. Figure 14.27 shows the I-V characteristics of an ideal diode.

An ideal diode offers zero resistance in forward biased mode and infinite resistance in reverse biased mode.

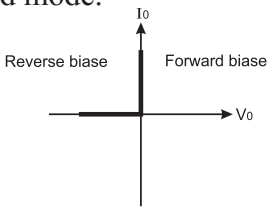


Fig. 14.27: I-V characteristics of an ideal diode.

The I-V characteristics of a forward biased diode (Fig. 14.24) is used to define two of its resistances i) the static (DC) resistance and ii) the dynamic (AC) resistance.

i) Static (DC) resistance: When a p-n junction diode is forward biased, it offers a definite resistance in the circuit. This resistance is called the static or DC resistance (R_g) of a diode. The DC resistance of a diode is the ratio of the DC voltage across the diode to the DC current flowing through it at a particular voltage.

$$R_g = \frac{V}{I}$$

ii) Dynamic (AC) resistance: The dynamic (AC) resistance of a diode, r_g, at a particular applied voltage, is defined as

$$r_g = \frac{\Delta V}{\Delta I}$$

The dynamic resistance of a diode depends on the operating voltage. It is the reciprocal of the slope of the characteristics at that point. Figure 14.28 shows how the DC and the AC resistance of a diode are found out.

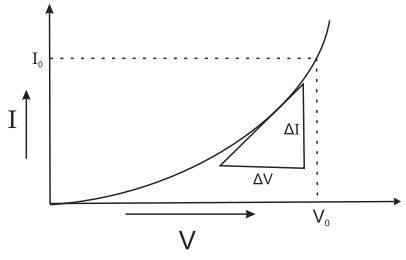
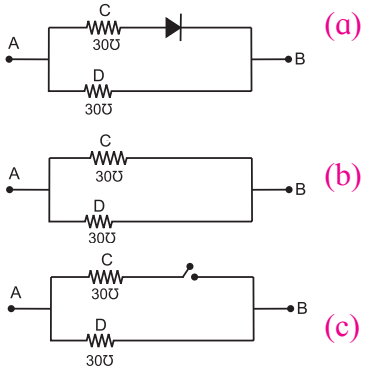


Fig. 14.28: DC and the AC resistance of a diode.

Example 14.3 Refer to the figure a shown below and find the resistance between point A and B when an ideal diode is (1) forward biased and (2) reverse biased.



Solution: We know that for an ideal diode, the resistance is zero when forward biased and infinite when reverse biased.

i) Figure b shows the circuit when the diode is forward biased. An ideal diode behaves as a conductor and the circuit is similar to two resistances in parallel.

R_{AB} = (30 x 30)/(30+30) = 900/60 = 15 Ω

ii) Figure c shows the circuit when the diode is reverse biased. It does not conduct and behaves as an open switch, path ACB. Therefore, R_{AB}= 30 Ω, the only resistance in the circuit along the path ADB.

14.8 Semiconductor devices:

Semiconductor devices find applications in variety of fields. They have many advantages. They also have some disadvantages. Here we discuses some advantages and disadvantages.

14.8.1 Advantages:

1. Electronic properties of semiconductors can be controlled to suit our requirement.
2. They are smaller in size and light weight.
3. They can operate at smaller voltages (of the order of few mV) and require less current (of the order of μA or mA), therefore, consume lesser power.
4. Almost no heating effects occur, therefore these devices are thermally stable.
5. Faster speed of operation due to smaller size.
6. Fabrication of ICs is possible.

14.8.2 Disadvantages:

1. They are sensitive to electrostatic charges.
2. Not vary useful for controlling high power.
3. They are sensitive to radiation.
4. They are sensitive to fluctuations in temperature.
5. They need controlled conditions for their manufacturing.
6. Very few matreials are semiconductors.

14.9 Applications of semiconductors and p-n junction diode:

A p-n junction diode is the basic block of a number of semiconductor devices. A semiconductor device can have more than one junction. Properties of a device can be controlled by controlling the concentration of dopants.

1. **Solar cell**: Converts light energy into electric energy. Useful to produce electricity in remote areas and also for providing electricity for satellites, space probes and space stations.
2. **Photo resistor**: Changes its resistance when light is incident on it.
3. **Bi-polar junction transistor**: These are devices with two junctions and three terminals. A transistor can be a p-n-p or n-p-n transistor. Conduction takes place with holes and electrons. Many other types of transistors are designed and fabricated to suit specific requirements. They are used in almost all semiconductor devices.
4. **Photodiode**: It conducts when illuminated with light.
5. **LED**: Light Emitting Diode: Emits light when current passes through it. House hold LED lamps use similar technology. They consume less power, are smaller in size and have a longer life and are cost effective.
6. **Solid State Laser**: It is a special type of LED. It emits light of specific frequency. It is smaller in size and consumes less power.
7. **Integrated Circuits (ICs):** A small device having hundreds of diodes and transistors performs the work of a large number of electronic circuits.

14.10 Thermistor:

Thermistor is a temperature sensitive resistor. Its resistance changes with change in its temperature. There are two types of thermistors, the Negative Temperature Coefficient (NTC) and the Positive Temperature Coefficient (PTC).

Resistance of a NTC thermistor decreases with increase in its temperature. Its temperature coefficient is negative. They are commonly used as temperature sensors and also in temperature control circuits.

Resistance of a PTC thermistor increases with increase in its temperature. They are commonly used in series with a circuit. They are generally used as a reusable fuse to limit current passing through a circuit to protect against *over current* conditions, as resettable fuses.

Thermistors are made from thermally sensitive metal oxide semiconductors. Thermistors are very sensitive to changes in temperature. A small change in surrounding temperature causes a large change in their resistance. They can measure temperature variations of a small area due to their small size. Both types of thermistors have many applications in industry.

Do you know ?

Electric and electronic devices

Electric devices: These devices convert electrical energy into some other form. Fan, refrigerator, geyser etc. are some examples. Fan converts electrical energy into mechanical energy. A geyser converts it into heat energy. They use good conductors (mostly metals) for conduction of electricity. Common working range of currents for electric circuits is milli ampers (mA) to amperes. Their energy consumption is also moderate to high. A typical geyser consumes about 2.0 to 2.50 kW of power. They are moderate to large in size and are costly.

Electronic devices: Electronic circuits work with control or sequential changes in current through a cell. A calculator, a cell phone a smart watch or the remote control of a TV set are some of the electronic devices. Semiconductors are used to fabricate such devices. Common working range of currents for electronic circuits it is nano-ampere to μA. They consume very low energy. They are very compact, and cost effective.

Internet my friend

1. https://www.electronics-tutorials.ws>diode
2. https://www.hitachi-hightech.com
3. https://ntpel.ac.in>courses
4. https://physics.info>semiconductors
5. https://www.hyperphysics.phy-astr.gsu.edu>semcn

Exercises

1. Choose the correct option.

i) Electric conduction through a semiconductor is due to:
   (A) electrons
   (B) holes
   (C) none of these
   (D) both electrons and holes

ii) The energy levels of holes are:
   (A) in the valence band
   (B) in the conduction band
   (C) in the band gap but close to valence band
   (D) in the band gap but close to conduction band

iii) Current through a reverse biased p-n junction, increases abruptly at:
   (A) breakdown voltage (B) 0.0 V
   (C) 0.3V (D) 0.7V

iv) A reverse biased diode, is equivalent to:
   (A) an off switch
   (B) an on switch
   (C) a low resistance
   (D) none of the above

v) The potential barrier in p-n diode is due to:
   (A) depletion of positive charges near the junction
   (B) accumulation of positive charges near the junction
   (C) depletion of negative charges near the junction,
   (D) accumulation of positive and negative charges near the junction

2. Answer the following questions.

i) What is the importance of energy gap in a semiconductor?

ii) Which element would you use as an impurity to make germanium an n-type semiconductor?

iii) What causes a larger current through a p-n junction diode when forward biased?

iv) On which factors does the electrical conductivity of a pure semiconductor depend at a given temperature?

v) Why is the conductivity of a n-type semiconductor greater than that of p-type semiconductor even when both of these have same level of doping?

3. Answer in detail.

i) Explain how solids are classified on the basis of band theory of solids.

ii) Distinguish between intrinsic semiconductors and extrinsic semiconductors.

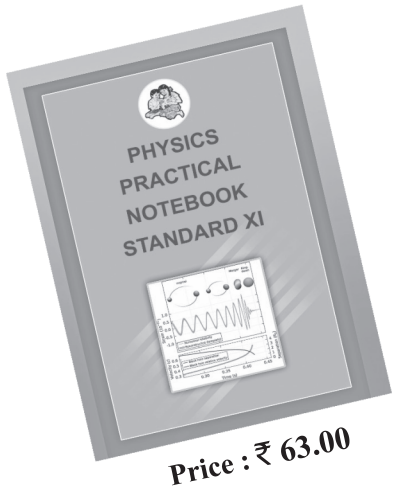
iii) Explain the importance of the depletion region in a p-n junction diode.

iv) Explain the I-V characteristic of a forward biased junction diode.

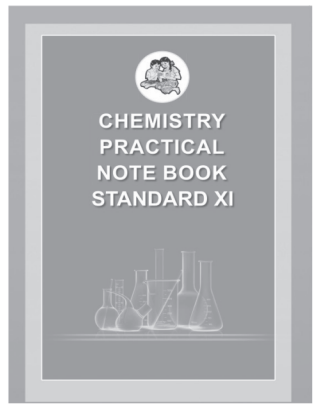
v) Discuss the effect of external voltage on the width of depletion region of a p-n junction

***

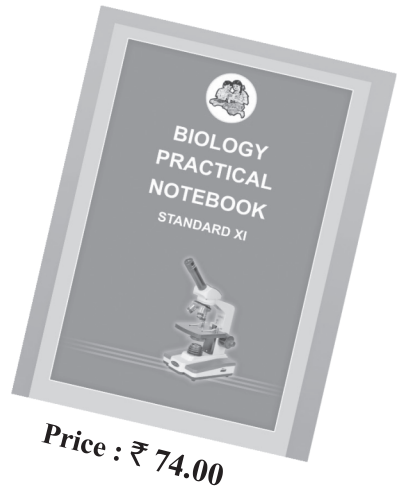
Practical Notebook for Standard XI
Practical Notebook Cum Journal



Price : ₹ 63.00



Price: ₹ 59.00



Price : ₹ 74.00

- Based on Government approved syllabus and textbook
- Inclusion of all practicals as per Evaluation scheme.
- Inclusion of various activities and objective questions
- Inclusion of useful questions for oral examination
- Separate space for writing answers

Practical notebooks are available for sale in the regional depots of the Textbook Bureau.

(1) Maharashtra State Textbook Stores and Distribution Centre, Senapati Bapat Marg, Pune 411004 ☎ 25659465 (2) Maharashtra State Textbook Stores and Distribution Centre, P-41, Industrial Estate, Mumbai - Bengaluru Highway, Opposite Sakal Office, Kolhapur 416122 ☎ 2468576 (3) Maharashtra State Textbook Stores and Distribution Centre, 10, Udyognagar, S. V. Road, Goregaon (West), Mumbai 400062 ☎ 28771842 (4) Maharashtra State Textbook Stores and Distribution Centre, CIDCO, Plot no. 14, W-Sector 12, Wavanja Road, New Panvel, Dist. Rajgad, Panvel 410206 ☎ 274626465 (5) Maharashtra State Textbook Stores and Distribution Centre, Near Lekhanagar, Plot no. 24, 'MAGH' Sector, CIDCO, New Mumbai-Agra Road, Nashik 422009 ☎ 2391511 (6) Maharashtra State Textbook Stores and Distribution Centre, M.I.D.C. Shed no. 2 and 3, Near Railway Station, Aurangabad 431001 ☎ 2332171 (7) Maharashtra State Textbook Stores and Distribution Centre, Opposite Rabindranath Tagore Science College, Maharaj Baug Road, Nagpur 440001 ☎ 2547716/2523078 (8) Maharashtra State Textbook Stores and Distribution Centre, Plot no. F-91, M.I.D.C., Latur 413531 ☎ 220930 (9) Maharashtra State Textbook Stores and Distribution Centre, Shakuntal Colony, Behind V.M.V. College, Amravati 444604 ☎ 2530965

E-learning material (Audio-Visual) for Standards One to Twelve is available through Textbook Bureau, Balbharati...

- Register your demand by scanning the Q.R. Code given alongside.
- Register your demand for E-learning material by using Google play store and downloading ebalbharati app.

www.ebalbharati.in, www.balbharati.in

ebalbharati