3. Alternating currents

* Introduction to Alternating current

Alternating Current (AC) :-

If the direction of current changes alternatively (periodically) and its magnitude changes continuously with respect to time, then the current is called alternating current. It is sinusoidal (i.e represented by sine or cosine angles) in nature.

Alternating current can be defined as the current whose magnitude and direction changes with time and attains the same magnitude and direction after a definite time interval. It changes continuously between zero and a maximum value and flows in one direction in the first half cycle and in the opposite direction in the next half cycle.

The instantaneous value of AC is given by

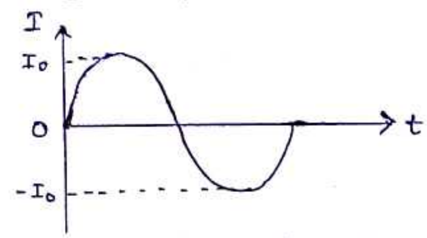
$$I = I_0 \sin\omega t \qquad \left[\omega = \frac{2\pi}{T} = 2\pi f\right]$$

where, I = current at any instant t

I_0 = maximum / peak value of AC

f = frequency and

ω = angular frequency.



Current vs time graph of an AC

Note :- Current whose direction does not change with time through a load is known as direct circuit (DC).

* **Advantages of AC over DC :-**

(1) AC generation is easy and economical.

(2) It can be easily converted into DC with the help of rectifier.

(3) In AC, energy loss is minimum, so it can be transmitted over large distances.

* **Disadvantages of AC over DC :-**

(1) AC shock is attractive, while DC shock is repulsive, so 220V AC is more dangerous than 220V DC.

(2) AC cannot be used in electroplating process because here constant current with constant polarity is needed which is given by DC.

* **Alternating emf or voltage :-**

It can be defined as the voltage whose magnitude and direction changes with time and attains the same magnitude and direction after a definite time interval. The instantaneous value of alternating emf or voltage is given by

$$V = V_0 \sin \omega t$$

where V = voltage at any time t,

V_0 = maximum / peak value of alternating voltage

and ω = angular frequency.

Note :- Alternating current, alternating emf, flux etc... all are sinusoidal waves.

* **Mean or Average value of AC :-**

It is defined as the value of AC which would send same amount of charge through a circuit in half cycle (i.e $T/2$) that is

Sent by Steady current in the Same time. It is denoted by I_m or I_{av}

Let the instantaneous value of alternating current is represented by $I = I_0 \sin\omega t \longrightarrow (1)$

The AC changes continuously with time. Suppose current is kept constant for small time (dt). Then, small amount of charge (dq) in small time (dt) is given by $dq = I\,dt = I_0 \sin\omega t \cdot dt$ (from -(1))

To calculate total charge send by AC over half cycle is given by

$$\int dq = \int_0^{T/2} I_0 \sin\omega t \cdot dt$$

$$q = I_0 \int_0^{T/2} \sin\omega t \cdot dt$$

Here q is steady charge over half cycle.

$$q = I_0 \left[\frac{-\cos\omega t}{\omega}\right]_0^{T/2}$$

$$q = -\frac{I_0}{\omega}\left[\cos\frac{\omega T}{2} - \cos 0^\circ\right]$$

$$q = -\frac{I_0}{\omega}\left[\cos\frac{\omega T}{2} - 1\right] \qquad \left[\omega = \frac{2\pi}{T} \Rightarrow \omega T = 2\pi\right]$$

$$q = -\frac{I_0}{\omega}\left[\cos\frac{2\pi}{2} - 1\right]$$

$$q = -\frac{I_0}{\omega}\left[-1-1\right]$$

$$\Longrightarrow q = \frac{2I_0}{\omega}$$

Also, the charge sent by AC in positive half cycle is $q_{AC} = I_m \times \frac{T}{2}$

where, I_m is mean value of AC over half cycle.

According to the definition

$$q = q_{AC} \quad \text{(over any half cycle)}$$

$$\frac{2I_0}{\omega} = I_m \times \frac{T}{2}$$

$$\Rightarrow I_m = \frac{4I_0}{\omega T} = \frac{4I_0}{2\pi} \quad [\omega T = 2\pi]$$

$$\Rightarrow I_m = \frac{2I_0}{\pi}$$

$$\boxed{I_m = 0.637\ I_0}$$

Mean value of AC (I_m) is 63.7% of the peak value of AC (I_0) over positive half cycle. For negative half cycle, the mean value of AC will be $-2I_0/\pi$. Therefore, in a complete cycle, the mean value of AC will be zero.

In the same way, mean value of alternating emf (V_m) is $\boxed{V_m = \frac{2V_0}{\pi} = 0.637\ V_0}$

Root Mean Square (RMS) value of AC :-

It is defined as that value of AC over a complete cycle which would generate same amount of heat in a given resistor that is generated by steady current in the same resistor and in the same time during a complete cycle.

It is also called virtual value or effective value of AC. It is represented by I_{rms} or I_{eff} or I_v. Suppose I is the current which flows in the resistor having resistance (R) in the time (T) produces heat (Q).

Instantaneous value of AC

$$I = I_0 \sin\omega t$$

If dQ is small amount of heat produced in time dt in resistor R, then

$$dQ = I^2 R\, dt \quad (\because Q = I^2RT) \rightarrow ①$$

In complete cycle ($0 \rightarrow T$) the total heat produced is Q.

After integrating Eq-(1), we get

$$\int dQ = \int_0^T I^2 R\,dt \Rightarrow Q = \int_0^T I^2 R\,dt.$$

Put the value of I in the above equation, we get

$$Q = \int_0^T (I_0 \sin\omega t)^2 R\,dt$$

$$= I_0^2 R \int_0^T \sin^2\omega t\,dt$$

$$= I_0^2 R \int_0^T \left[\frac{1-\cos 2\omega t}{2}\right] dt$$

$$= \frac{I_0^2 R}{2} \int_0^T (1-\cos 2\omega t)\,dt$$

$$= \frac{I_0^2 R}{2}\left[\int_0^T dt - \int_0^T \cos 2\omega t\,dt\right]$$

$$= \frac{I_0^2 R}{2}\left([t]_0^T - \left[\frac{\sin 2\omega t}{2\omega}\right]_0^T\right)$$

$$= \frac{I_0^2 R}{2}\left[(T-0) - \frac{1}{2\omega}[\sin 2\omega T - \sin 0]\right]$$

$$= \frac{I_0^2 R}{2}\left[T - \frac{1}{2\omega}(\sin 2\times 2\pi - 0)\right] \quad (\because \omega T = 2\pi)$$

$$= \frac{I_0^2 R}{2}\left[T - \frac{1}{2\omega}(0)\right] \quad (\because \sin 4\pi = 0)$$

$$\therefore Q = \frac{I_0^2 RT}{2} \longrightarrow ②$$

If I_{rms} is rms value of AC and Q is the heat produced by rms current (I_{rms}), then

$$Q = I_{rms}^2 RT \longrightarrow ③$$

On comparing Eqs (2) and (3), we get

$$I_{rms}^2 RT = \frac{I_0^2 RT}{2} \Rightarrow I_{rms}^2 = \frac{I_0^2}{2}$$

$$\Rightarrow \boxed{I_{rms} = \sqrt{\frac{I_0^2}{2}} = \frac{I_0}{\sqrt{2}} = 0.707\,I_0}$$

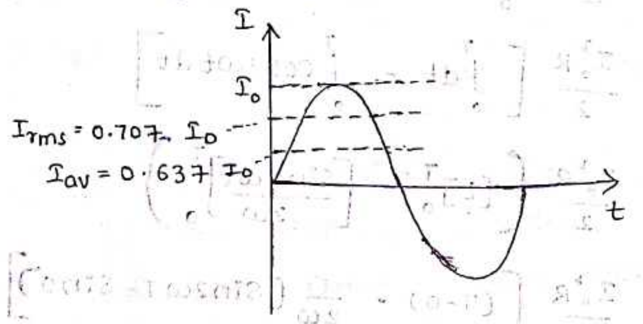
$$\Rightarrow \boxed{I_{rms} = 70.7\% \text{ of } I_0}$$

From the above equation, we conclude that rms value of current is 70.7% of the peak value of current.

In the same way, the rms value of alternating emf (V_{rms} or V_{eff}) is

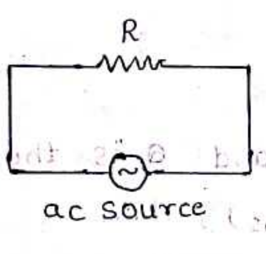
$$\boxed{V_{rms} = \frac{V_0}{\sqrt{2}} = 0.707 V_0 = 70.7\% \text{ of } V_0}$$

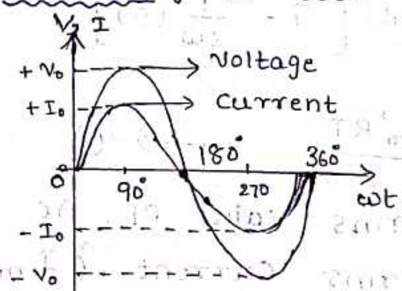
The different values I_0, I_{av} and I_{rms} are shown in figure given below.

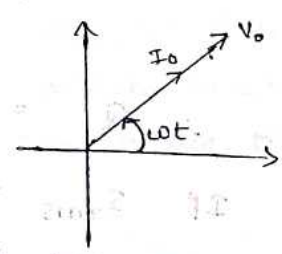


RMS and average value of current on the same graph.

* AC circuit containing a pure resistor :-







Pure resistive load

Variation of V and I with ωt

Phasor diagram.

Consider a pure resistor of resistance R connected to an ac source, which gives an ac voltage

$$V = V_0 \sin\omega t \longrightarrow (1)$$

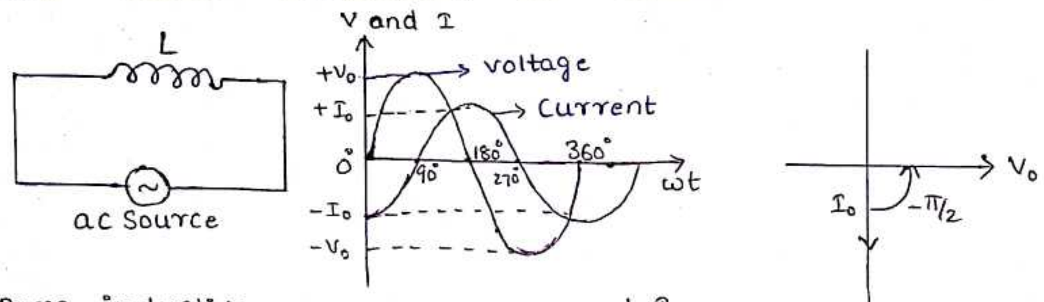
The current through the resistor at any instant t is $I = \frac{V}{R} = \left(\frac{V_0}{R}\right) \sin\omega t$

i.e $I = I_0 \sin\omega t \longrightarrow (2)$

where $I_0 = \frac{V_0}{R}$, is the peak current.

From Eqs, (1) and (2) it follows that I and V are zero at the same time, reach their maximum or minimum values at the same time. Hence, we say that I is in phase with V when the load is purely resistive. This is shown in the phasor diagram.

* AC Circuit Containing an inductor :-



Pure inductive load

Variation of V and I versus ωt

Phasor diagram

The circuit shows an inductor of self inductance L connected to an ac source of voltage $V = V_0 \sin\omega t$. In practice, an inductor will have a certain resistance in their windings. Let us assume that the resistance of the inductor coil is negligibly small.

Let the ac voltage across the inductor at an instant of time t be $V = V_0 \sin\omega t \longrightarrow (1)$

where V_0 = peak ac voltage

ω = angular frequency of ac

The induced emf at the instant t is given by

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

where $\frac{dI}{dt}$ = rate of change of current at that instant.

According to Kirchhoff's law, $v + e = 0$ (since there is no resistance in the circuit).

$$V_0 \sin\omega t - L \frac{dI}{dt} = 0$$

$$L \frac{dI}{dt} = V_0 \sin\omega t$$

$$dI = \frac{V_0}{L} \sin\omega t \cdot dt.$$

The instantaneous current I is obtained by integrating the above expression.

$$I = \int dI = \int \frac{V_0}{L} \sin\omega t \, dt.$$

$$I = \frac{V_0}{L}\left(-\frac{\cos\omega t}{\omega}\right) + \text{Constant of integration}.$$

It can be shown that the constant of integration is zero. (The current in the circuit oscillates symmetrically about zero. There is no time independent component. But the integration constant has dimensions of current and is independent of time. Hence, the integration constant is zero).

We know that $-\cos\omega t = \sin\left(\omega t - \frac{\pi}{2}\right)$

$$I = \frac{V_0}{\omega L} \sin\left(\omega t - \frac{\pi}{2}\right)$$

$$I = I_0 \sin(\omega t - \pi/2) \longrightarrow (2)$$

where $I_0 = \frac{V_0}{\omega L}$ = Peak value of ac

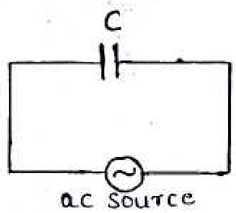
The quantity $\omega L = X_L$ represents the effective opposition of the inductance coil to ac and it is called the inductive reactance.

$$\therefore \quad X_L = \omega L = 2\pi f L = \frac{V_0}{I_0} = \frac{V_{rms}}{I_{rms}}, \quad f = \text{frequency of ac}$$

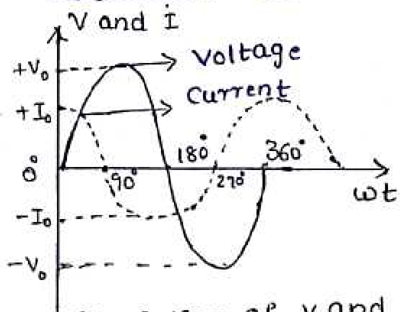
The inductive reactance of an inductor is the effective opposition offered by the inductor to the flow of ac and is defined as the ratio of rms value of current through the inductor.

From Eqs (1) and (2) we find that the current in an ac circuit containing pure inductor lags behind the applied voltage by 90° or $\pi/2$ radian. The phase relationship is shown in the phasor diagram.

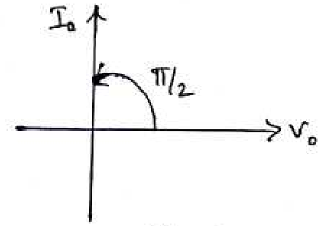
* **Ac Circuit Containing a Capacitor :-**



pure capacitance load



Variation of v and I versus ωt.



Phasor diagram.

The circuit shows a capacitor of capacitance C. Connected to an ac source. The instantaneous voltage is given by $\quad v = V_0 \sin\omega t \longrightarrow (1)$

where V_0 = peak ac voltage, ω = angular frequency of ac

If q be the charge on the capacitor at an instant of time t, then the potential difference across capacitor is $\quad v = \frac{q}{C} \Rightarrow q = vC$.

The instantaneous current I is given by

$$I = \frac{dq}{dt} = \frac{d}{dt}(vC) = \frac{d}{dt}(V_0 \sin\omega t).C$$

$$I = \omega C V_0 \cos\omega t$$

we can write, $\quad \cos\omega t = \sin(\omega t + \pi/2)$

$$I = \frac{V_0}{\left(\frac{1}{\omega C}\right)} \sin(\omega t + \pi/2)$$

$$I = I_0 \sin(\omega t + \pi/2) \longrightarrow (2)$$

where $I_0 = \frac{V_0}{\left(\frac{1}{\omega C}\right)} = \frac{V_0}{X_C} =$ peak value of ac

The quantity $\frac{1}{\omega C} = X_C$ represents the effective opposition of the capacitor to the ac and is called capacitive reactance. $X_C = \frac{V_0}{I_0} = \frac{V_{rms}}{I_{rms}}$

Also, $X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C}$, where f = frequency of ac.

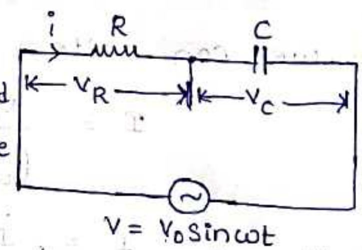
The capacitive reactance of a capacitor is the effective opposition offered by a capacitor to the flow of ac and is defined as the ratio of rms value of voltage to the rms value of current through the capacitor.

From Eqs (1) and (2) we find that current in an ~~arre~~ ac circuit containing capacitor leads the applied ac voltage by $\pi/2$ radian or 90°. The phase relationship is shown in the phasor diagram.

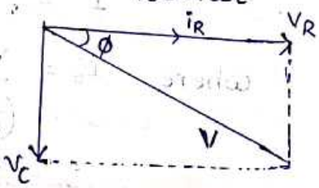
* Phase and amplitude relations for ac and voltage :-

Circuit element	Symbol	Impedance	Phase of the Current	Phase angle ϕ	Amplitude relation
Resistor	R	R	In phase with V_R	0°	$V_R = i_R R$
Capacitor	C	X_C	lead V_C by $\pi/2$	$+\pi/2$	$V_C = i_C X_C$
Inductor	L	X_L	lags V_L by $\pi/2$	$-\pi/2$	$V_L = i_L X_L$

* **RC - Circuit :-**



Consider a circuit with resistor R and capacitor C, connected in series with an alternating source of potential V. If V_R and V_C are the p.d across resistor and capacitor respectively, then



(i) Supply voltage $V = \sqrt{V_R^2 + V_C^2} = \sqrt{(IR)^2 + (IX_C)^2}$

(ii) Impedance $Z = \frac{V}{I} = \sqrt{R^2 x X_C^2} = \sqrt{R^2 + \left(\frac{1}{\omega C}\right)^2}$

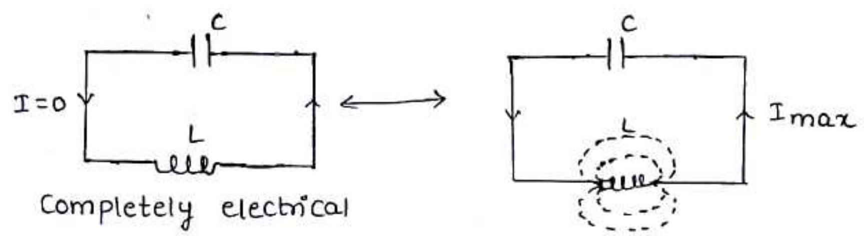
(iii) Current $I = I_0 \sin(\omega t + \phi)$

where $\tan\phi = \frac{X_C}{R} = \frac{1}{\omega CR}$

(iv) Peak Current $I_0 = \frac{V_0}{Z} = \frac{V_0}{\sqrt{R^2 + X_C^2}}$

* Free Oscillations of LC circuit :-

when a capacitor is supplied with an AC current, it gets charged.



Completely electrical Energy

Completely magnetic Energy.

when this charged capacitor is connected with an inductor, current flows through inductor, giving rise to magnetic flux. Hence, induced emf is produced in the circuit. Due to this, the charge (or energy) on the capacitor decreases and an equivalent amount of energy is stored in the inductor in the form of magnetic field. when the discharging of the capacitor completes, current and magnetic flux linked with L starts decreasing.

Therefore, an induced emf is produced which recharges the capacitor in opposite direction. This process of charging and discharging capacitor is repeated and energy taken once from source keeps on oscillating between C and L.

According to Kirchhoff's loop rule, we have

$$\frac{q}{c} - L\frac{dI}{dt} = 0 \longrightarrow (1)$$

But $I = -\frac{dq}{dt} \Rightarrow \frac{dI}{dt} = -\frac{d^2q}{dt^2}$

Negative Sign indicates that as q decreases, I increases.

Putting $\frac{dI}{dt}$ in Eq (1), we get

$$\frac{d^2q}{dt^2} + \frac{1}{LC}q = 0$$

Compare this equation with equation of Simple harmonic oscillator, we get

$$\frac{d^2x}{dt^2} + \omega_0^2 x = 0 \quad \left(\because \omega_0 = \frac{1}{\sqrt{LC}}\right)$$

Therefore, the charge oscillates with a frequency

$$\boxed{f = \frac{\omega}{2\pi} = \frac{1}{2\pi\sqrt{LC}}}$$

The L-C oscillations discussed above are not realistic for the two reasons.

(i) Every inductor has some resistance. The effect of this resistance will introduce a damping effect on the charge and current in the circuit. Thus, the oscillations finally die away.

(ii) Even, if the resistance is zero, the total energy of the system would not remain constant. It is radiated away from the system in the form of electromagnetic waves. In fact, radio and TV transmitters depend on this radiation.

For L-C oscillators, the energy oscillates in between the capacitor and inductor as electrostatic energy and magnetic energy.

It is given as,

$$U = \frac{1}{2} L I^2 = \frac{1}{2} \frac{q^2}{C}$$

The table below gives the analogy between some important quantities of mechanical and electrical systems.

Mechanical System	Electrical System
Mass, m	Inductance, L
Force constant, k	Reciprocal capacitance, $1/C$
Displacement, x	Charge, q
Velocity, $v = dx/dt$	Current, $I = dq/dt$
Mechanical energy, $E = \frac{1}{2} kx^2 + \frac{1}{2} mv^2$	Electromagnetic energy, $U = \frac{1}{2}\frac{q^2}{C} + \frac{1}{2} L I^2$

* Freely oscillating RLC Circuit :-

Here we add a resistor to an LC circuit without any external alternating emf.

Apply KVL

$$-L\frac{dI}{dt} - \frac{Q}{C} - IR = 0$$

Here $I = \frac{dQ}{dt}$

$$\Rightarrow \quad L\frac{d^2Q}{dt^2} + R\frac{dQ}{dt} + \frac{1}{C}Q = 0$$

$$\frac{d^2Q}{dt^2} + \left(\frac{R}{L}\right)\frac{dQ}{dt} + \frac{1}{LC}Q = 0$$

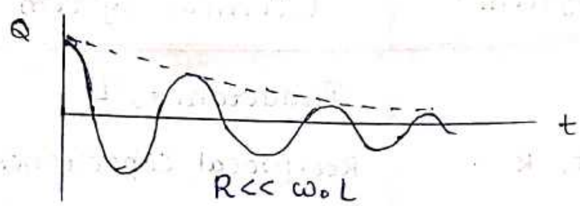
→ This equation describes a damped harmonic oscillator

→ Here Amplitude of the oscillations decrease with time.

Scanned by CamScanner

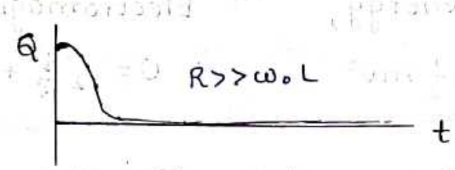
→ Current flows through resistor, resulting Joul Heating represents some energy loss or damping during each cycle of oscillation.

→ If the resistance is small, these loses are small and oscillation amplitude decreases slowly as shown in figure.



The circuit is said to be underdamped.

→ If the resistance is too large, the current drops to zero before any oscillation can occur in figure.



The Circuit is said to be overdamped.

→ From the theory of linear differential equation, in under damped case

Charge $Q = Q_{max} e^{-\left(\frac{Rt}{2L}\right)} \cos \omega_d t$

Here 'ω_d' is damped oscillations.

and $\omega_d = \sqrt{\omega_0^2 - \frac{R^2}{4L^2}}$

Here ω_0 is natural undamped frequency.

for small value of 'R' $\omega_d \approx \omega_0$

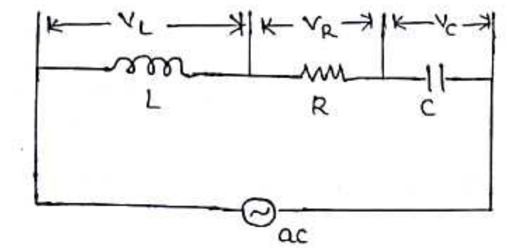
for large value of 'R' $\omega_d = 0$

$$\therefore \quad 0 = \sqrt{\omega_0^2 - \frac{R^2}{4L^2}}$$

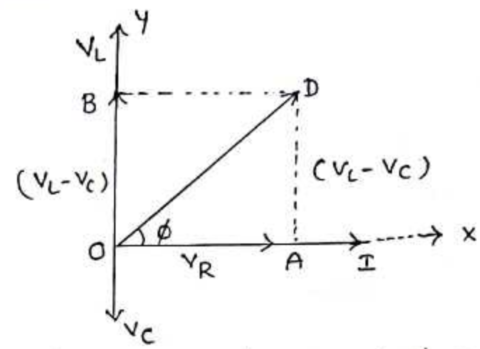
$$\Rightarrow \omega_0^2 = \frac{R^2}{4L^2} \Rightarrow \left(\frac{1}{\sqrt{LC}}\right)^2 = \frac{R^2}{4L^2}$$

$$\Rightarrow \frac{1}{LC} = \frac{R^2}{4L^2} \Rightarrow \boxed{R = 2\sqrt{L/C}}$$

* AC Circuit Containing Resistance, Inductance and Capacitance (Series LCR circuit) - Phasor Diagram analysis :-



(a) Series LCR Circuit



(b) phasor diagram (when $X_L > X_C$)

Fig (a) shows a circuit containing a resistance R, inductance L and a capacitance C in series with an ac source. Let I be the rms value of current in the circuit.

The rms values voltage across the combination is made up of three parts, namely

(1) The voltage V_R across the resistance, which is in phase with the current in it.

(2) The voltage V_L across the inductance leading the current by 90° and

(3) The voltage V_C across the capacitor lagging the current by 90°.

They are represented in a phasor diagram as shown in fig (b). V_L is along the +ve y-axis and V_C is along the -ve y-axis. If V_L is greater than V_C their net voltage $V_L - V_C$ is represented by OB.

V_R is represented by OA along the x-axis. The diagonal OD of the rectangle OADB gives the resultant voltage across the combination.

Now, $V^2 = V_R^2 + (V_L - V_C)^2$

V, V_L, V_C and V_R represent either peak values or rms values. I is the peak current or rms current accordingly

But $V_R = IR$, $V_L = IX_L$ and $V_C = IX_C$

where I is the current in the circuit

$$\therefore V^2 = I^2 [R^2 + (X_L - X_C)^2]$$

The current $I = \dfrac{V}{\sqrt{R^2 + (X_L - X_C)^2}} = \dfrac{V}{Z}$

where $Z = \sqrt{R^2 + (X_L - X_C)^2}$

Z is called the impedance of the circuit

Impedance :- In a series LCR circuit with ac source, the combined opposition offered by resistance and reactance to the flow of ac is called the impedance of the circuit. It is denoted by Z. $I = \dfrac{V}{Z}$

$$\text{Impedance } Z = \frac{\text{Peak value of voltage } (V_0)}{\text{Peak value of current } (I_0)}$$

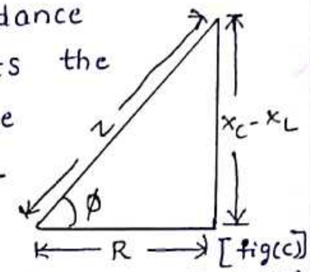
$$= \frac{\text{rms value of voltage } (V_{rms})}{\text{rms value of current } (I_{rms})}$$

→ Impedance of a series LCR circuit is the effective opposition offered by the circuit to the flow of ac and is defined as the ratio of rms value of voltage across the circuit to the rms value of current in the circuit

the expression of current $I = I_0 \sin(\omega t + \phi)$.

where ϕ is the phase angle b/w V and I

The Fig (c) shows an impedance triangle or impedance diagram. It is a right angled triangle with its sides representing ohmic resistance, net reactance and impedance of the circuit. The base represents the ohmic resistance R in the circuit. The Perpendicular represents the net reactance $X_L - X_C$. The hypotenuse represents the impedance Z of the LCR circuit.



(Impedance diagram)

The angle ϕ represents the phase angle by which voltage and current differ.

From fig (c) $\tan\phi = \dfrac{V_L - V_C}{V_R} = \dfrac{IX_L - IX_C}{IR}$

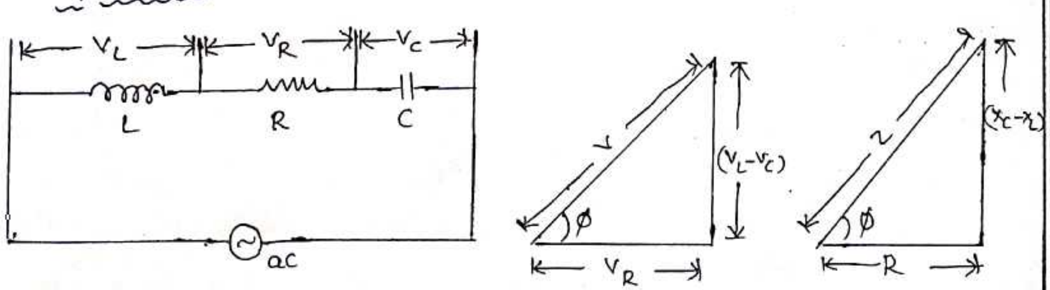
Hence, we get $\tan\phi = \dfrac{X_L - X_C}{R}$

→ If $X_L > X_C$, the current in the circuit lags behind the applied voltage by an angle ϕ and the circuit is inductive.

→ If $X_L < X_C$, the current in the circuit leads the applied voltage by an angle ϕ and the circuit is capacitive.

→ If $X_L = X_C$ the current in the circuit is in phase with the applied voltage and the circuit is resistive or resonant.

* **AC Circuit Containing Resistance, Inductance and Capacitance (Series LCR circuit) - Analytical Solution :-**



(a) Series LCR Circuit. (b) Voltage diagram (c) Impedance diagram

Fig (a) Shows a Series LCR Circuit having a Sinusoidal ac Source of angular frequency ω.

Scanned by CamScanner

At any arbitary instant of time t,
let V be the applied voltage and I be the current in the circuit.

The voltage V can be written as

$$V = V_0 \sin\omega t \longrightarrow (1)$$

where $V_0 \rightarrow$ Peak value of applied voltage.

The voltage V divides into V_L, V_R and V_C across the inductor (L), the resistor (R) and the capacitor (C) respectively.

Then $V_L = L\frac{dI}{dt}$, $V_R = IR$ and $V_C = q/C$

where q is the charge on the capacitor plates at time t.

According to the law of conservation of energy

$$V_L + V_R + V_C = V$$

(or) $$L\frac{dI}{dt} + IR + q/C = V_0 \sin\omega t \longrightarrow (2)$$

Here, $I = \frac{dq}{dt} \Rightarrow \frac{dI}{dt} = \frac{d^2q}{dt^2}$

$$L\frac{d^2q}{dt^2} + R\frac{dq}{dt} + \frac{q}{C} = V_0 \sin(\omega t) \longrightarrow (3)$$

This equation represents an equation for damped simple harmonic motion. The solution to Eq (3) is of the form

$$q = q_0 \sin(\omega t + \theta)$$

Hence, $\frac{dq}{dt} = \omega q_0 \cos(\omega t + \theta)$ and

$$\frac{d^2q}{dt^2} = -\omega^2 q_0 \sin(\omega t + \theta)$$

Substituting these into Eq (3), we get.

$$\Rightarrow V_0 \sin\omega t = L[-q_0\omega^2 \sin(\omega t + \theta)] + R[q_0\omega \cos(\omega t + \theta)] + \frac{\omega q_0 \sin(\omega t + \theta)}{\omega C}$$

$$\Rightarrow \omega q_0 \left\{\left[-\omega L \sin(\omega t+\theta) + \frac{\sin(\omega t+\theta)}{\omega C}\right] + R\cos(\omega t+\theta)\right\} = V_0 \sin \omega t$$

Let inductive reactance $x_L = \omega L$ and capacitive reactance $x_C = \frac{1}{\omega C}$

$$\Rightarrow \omega q_0 \left\{ R\cos(\omega t+\theta) - (x_L - x_C)\sin(\omega t+\theta)\right\} = V_0 \sin \omega t$$

Multiplying and dividing the LHS of the above equation by Z, the impedance of the circuit, we get

$$\Rightarrow \omega q_0 Z \left\{ \frac{R}{Z}\cos(\omega t+\theta) - \frac{(x_L - x_C)}{Z}\sin(\omega t+\theta)\right\} = V_0 \sin \omega t$$

let $R/Z = \cos\phi \rightarrow (4)$ and $\frac{x_L - x_C}{Z} = \sin\phi \rightarrow (5)$

The above equation becomes

$$\Rightarrow \omega q_0 Z \left\{ \cos\phi . \cos(\omega t+\theta) - \sin\phi . \sin(\omega t+\theta)\right\} = V_0 \sin \omega t$$

But $\left\{ \cos\phi . \cos(\omega t+\theta) - \sin\phi . \sin(\omega t+\theta)\right\} = \cos(\omega t+\theta-\phi)$

Hence $\omega q_0 Z \cos(\omega t+\theta+\phi) = V_0 \sin \omega t$

let $\theta + \phi = \pi/2 \Rightarrow \theta = \pi/2 - \phi$

$$\Rightarrow \omega q_0 Z \cos(\omega t + \pi/2) = V_0 \sin \omega t$$

$$\Rightarrow \omega q_0 Z \sin(\omega t) = V_0 \sin(\omega t)$$ (Ignoring -ve sign)

Comparing the two sides of this equation, we get.

$$V_0 = \omega q_0 Z$$

let $I_0 = \omega q_0$

Therefore $V_0 = I_0 Z$ or $\boxed{Z = \frac{V_0}{I_0}}$

Squaring and adding the eq^n (4) and (5), we get

$$\frac{R^2}{Z^2} + \frac{(x_L - x_C)^2}{Z^2} = \cos^2\phi + \sin^2\phi = 1$$

Hence $Z^2 = R^2 + (x_L - x_C)^2$

The impedance of the circuit, $Z = \sqrt{R^2 + (x_L - x_C)^2}$

again from eq^n(4) & (5), we get $\tan\phi = \frac{x_L - x_C}{R}$

Thus, the analytical solution for the amplitude and phase of the current in the circuit agrees with that obtained by the method of phasors.

* Resonance :-

In a Series L-C-R Circuit, when phase (ϕ) between Current and voltage is zero, then the circuit is said to be a resonant circuit.

As applied frequency increases, then $x_L = \omega L$, x_L increases and $x_c = \frac{1}{\omega c}$, x_c decreases.

At some frequency (angular frequency - (ω_r)),

$$x_L = x_c$$

where $x_L = \omega_r L$, $x_c = \frac{1}{\omega_r C}$

The frequency at which x_c and x_L become equal, is called resonant frequency.

$$\Rightarrow \omega_r L = \frac{1}{\omega_r C} \quad (\text{or}) \quad \omega_r^2 = \frac{1}{LC}$$

$$(2\pi \nu_r)^2 = \frac{1}{LC} \quad [\because \omega_r = 2\pi\nu_r, \text{ where } \nu_r \text{ is resonating frequency}$$

$$\Rightarrow 2\pi \nu_r = \frac{1}{\sqrt{LC}}$$

$$\boxed{\therefore \nu_r = \frac{1}{2\pi\sqrt{LC}}}$$

At resonating frequency,

$$Z = R = \text{minimum}$$

$$\therefore I = \frac{E}{Z} = \text{maximum}$$

Since, Z is minimum, therefore I will be maximum.

* Quality Factor (Q- Factor) :-

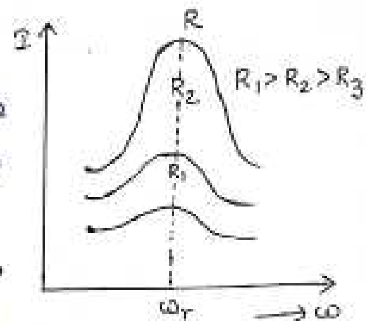
It is the measure of sharpness of the resonance of an L-C-R circuit. It is defined as the ratio of voltage developed across the inductance or capacitance at resonance to the impressed voltage, which is the voltage applied across R.

$$Q\text{-factor} = \frac{\text{Voltage across } L \text{ (or } C)}{\text{Voltage across } R}$$

$$Q\text{-factor} = \frac{V_L \text{ or } V_C}{V_R} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

$$\boxed{Q\text{-factor} = \frac{\omega_0 L}{R} = \frac{1}{\omega_0 RC}}$$

Q is just a number having no dimensions, it can also be called voltage multiplication factor of the circuit.



The electronic circuit with high Q values would respond to a very narrow range of frequencies and vice-versa. Higher the value of Q, the narrower and sharper is the resonance.

Q-factor can also be defined as the ratio of the resonant frequency to the difference in two frequencies taken on both sides of the resonant frequency such that at each frequency, the current amplitude becomes $\frac{1}{\sqrt{2}}$ times the value at resonant frequency.

Mathematically,

$$Q\text{-factor or } Q = \frac{\omega_r}{\omega_1 - \omega_2}$$

where ω_1 and ω_2 are frequencies when current decreases to 0.707 ($1/\sqrt{2}$) times the peak value of current. We can also write, $\omega_1 = \omega_r + \Delta\omega$

$$\omega_2 = \omega_r - \Delta\omega$$

The difference $\omega_1 - \omega_2 = 2\Delta\omega$ is often called the bandwith of the circuit.

Thus, from the above, Q-factor can also be defined as the ratio of resonant angular frequency to bandwith of the circuit.

The smaller the bandwith ($\Delta\omega$), the sharper and narrower is the resonance.

* **Significance of Q-factor :-**

(1) Q-factor denotes the Sharpness of turning.

(2) High Q-factor indicates lower rate of energy loss.

(3) Higher value of Q-factor indicates Sharper peak in the current.

(4) For R = 0, Q-factor = infinity.

* **Average Power Associated in AC Circuit :-**

Power is defined as the rate of doing work

$$p = \frac{dw}{dt} \longrightarrow (1)$$

(or)

Power is defined as the product of voltage and current. In AC circuit, both emf and current change continuously with respect to time. So in it we have to calculate average power in complete cycle.

Instantaneous power $p = VI \longrightarrow (2)$

$[\because V = V_0 \sin\omega t \ , \ I = I_0 \sin(\omega t + \phi)]$

Here, V and I are instantaneous voltage and current, respectively. If the instantaneous power remains constant for a small time dt, then small amount of work done in maintaining the current for a small time dt is

$$\frac{dw}{dt} = VI$$

$$dw = VI\,dt \longrightarrow (3)$$

Integrating on both sides, we get

$$\int dw = \int_0^{T} VI\,dt$$

Total work done or energy spent in maintaining current over one full cycle,

$$w = \int_0^T V_0 \sin\omega t \cdot I_0 \sin(\omega t + \phi)\, dt$$

$$= V_0 I \int_0^T \sin\omega t\,(\sin\omega t \cdot \cos\phi + \cos\omega t \cdot \sin\phi)\, dt$$

$$= V_0 I \left[\cos\phi \int_0^T \sin^2\omega t\, dt + \sin\phi \int_0^T \sin\omega t \cdot \cos\omega t\, dt\right]$$

$$= V_0 I \left[\cos\phi \int_0^T \left(\frac{1-\cos 2\omega t}{2}\right) dt + \frac{\sin\phi}{2} \int_0^T 2\sin\omega t \cdot \cos\omega t\, dt\right]$$

$$= \frac{V_0 I}{2}\left[\left(\cos\phi\, [t]_0^T - \int_0^T \cos 2\omega t\, dt\right) + \sin\phi \int_0^T \sin 2\omega t\, dt\right]$$

But $\int_0^T \cos 2\omega t\, dt = 0$ or and $\int_0^T \sin 2\omega t\, dt = 0$

$$\therefore\ w = \frac{V_0 I_0 T}{2} \cos\phi$$

Average power associated in AC circuit,

$$P_{av} = \frac{w}{T} = \frac{V_0 I_0 T \cos\phi}{2T} = \frac{V_0 I_0}{2} \cos\phi.$$

$$\boxed{P_{av} = \frac{V_0}{\sqrt{2}} \cdot \frac{I_0}{\sqrt{2}} \cos\phi}$$

or $P_{av} = V_{rms} I_{rms} \cos\phi$

$= V_v I_v \cos\phi$

Here $\cos\phi$ is power factor, which is defined as the cosine of the angle of lag or lead.

If P_{av} is true power or average power, then power factor is given by

$$\cos\phi = \frac{P_{av}}{V_{rms} I_{rms}} = \frac{\text{True power}}{\text{Apparent Power}} \quad \cos\phi = \frac{R}{Z}$$

Here, ϕ is the phase difference b/w I_{rms} and V_{rms}

* Special cases :-

(i) AC circuit Containing R

when $\phi = 0°$, then $P_{av} = V_v I_v \cos 0 = V_v I_v$

So, average power in R is maximum.

(ii) AC circuit containing L

when $\phi = \pi/2$ then $P_{av} = V_v I_v \cos \pi/2 = 0$

So, average power in L is zero.

(iii). AC circuit containing C

when $\phi = \pi/2$, then $P_{av} = V_v I_v \cos \pi/2 = 0$

So, average power in C is zero.

(iv) AC circuit containing L and R

when $\tan\phi = \frac{\omega L}{R} \Rightarrow \cos\phi = \frac{R}{\sqrt{R^2+\omega^2L^2}}$

then $P_{av} = V_v I_v \cdot \frac{R}{\sqrt{R^2+\omega^2L^2}}$

(V) AC circuit containing C and R

when $\tan\phi = \frac{1/\omega C}{R} \Rightarrow \cos\phi = \frac{R}{\sqrt{R^2+1/\omega^2C^2}}$

then $P_{av} = V_v I_v \cdot \frac{R}{\sqrt{R^2+1/\omega^2C^2}}$

(vi) AC circuit containing L, C and R

when $\tan\phi = \frac{\omega L - 1/\omega L}{R}$

$\Rightarrow \cos\phi = \frac{R}{\sqrt{R^2+(\omega L - 1/\omega C)^2}}$

then $P_{av} = V_v I_v \cdot \frac{R}{\sqrt{R^2+\left(\omega L - \frac{1}{\omega C}\right)^2}}$

* **Wattless current :-**

The current which consumes no power for its maintenance in the circuit is called wattless current or idle current

If the resistance in an AC circuit is zero, although current flows in the circuit, then the average power remains zero, i.e there is no energy dissipation in the circuit, Such a circuit is called wattless ~~current~~ circuit and the current flowing is called wattless current.

If the circuit contains either inductance or capacitance only, then phase difference between current and voltage is 90°, i.e $\phi = 90°$. The average power in such a circuit is

$$P_{av} = V_{rms} \times I_{rms} \times \cos\phi = V_{rms} \times I_{rms} \times \cos 90° = 0.$$

* **Choke coil :-**

Choke coil is an electrical device used for controlling current in an AC circuit without wasting electrical energy in the form of heat.

(i) To reduce low frequency alternating currents, choke coils with laminated soft iron cores are used. These are called af choke coils

(ii) To reduce high frequency alternating currents, choke coils with air cores are used. These are called rf choke coils.

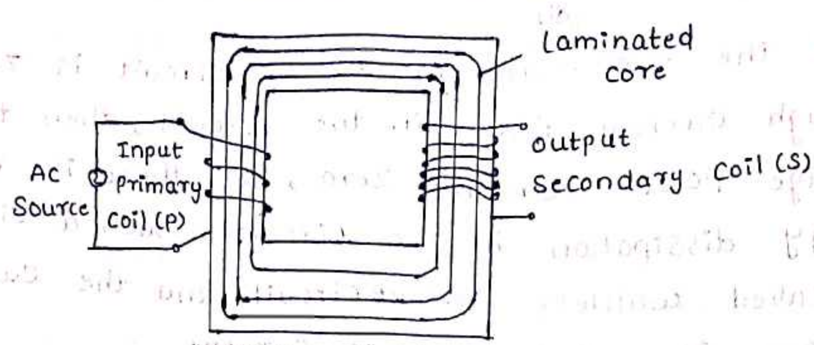
* **Transformer :-**

It is a device which is used to increase or decrease the alternating voltage.

The transformers are of the following types

1. Step-up transformer
2. Step-down transformer.

Scanned by CamScanner



Step - down transformer

Principle :- Transformer is based upon the principle of mutual induction.

Construction :-

It consists of two coils, primary coil (P) and secondary coil (S) insulated from each other wounded on soft iron core. Often the primary coil is the input coil and secondary coil is the output coil. These soft iron cores are laminated to minimise eddy current loss.

Working and Theory :-

The value of the emf induced in secondary coil due to alternating voltage applied to primary coil depends on the number of turns in the secondary coil. We consider an ideal transformer in which the primary coil has negligible resistance and all the flux in the core links both primary and secondary windings. Let ϕ be the flux in each turn in the core at time t due to current in the primary when a voltage V_p is applied to it.

Then, the induced emf or voltage E_s, in the secondary with N_s turns is

$$E_s = -N_s \frac{d\phi}{dt} \longrightarrow (1)$$

The alternating flux ϕ also induces an emf, called back emf in the primary. This is

$$E_p = -N_p \, d\phi/dt \longrightarrow (2)$$

But $E_p = V_p$. If this was not, so the primary current would be infinite, since the primary has zero resistance (as considered). If the secondary is an open circuit or the current taken from it is small, then to a good approximation.

$$E_S = V_S$$

where V_S is the voltage across the secondary.

Therefore, Eqs (1) and (2) can be written as

$$V_S = -N_S \frac{d\phi}{dt} \longrightarrow (3)$$

and $$V_P = -N_P \frac{d\phi}{dt} \longrightarrow (4)$$

From eqs (3) and (4) we have

$$\frac{V_S}{V_P} = \frac{N_S}{N_P} \longrightarrow (5)$$

The above relation has been obtained using three assumptions.

(i) The primary resistance and current are small.

(ii) The same flux links both the primary and the secondary as very little flux escapes from the core.

(iii) The secondary current is small.

If the transformer is assumed to be 100% efficient (no energy losses), the power input is equal to the power output. Since $P = IV$, we get

$$I_P V_P = I_S V_S \longrightarrow (6)$$

Although, some energy is always lost, still this is a good approximation, since a well designed transformer may have an efficiency of more than 95%.

Combining Eqs (5) and (6), we have

$$\boxed{\frac{I_P}{I_S} = \frac{V_S}{V_P} = \frac{N_S}{N_P}} \longrightarrow (7).$$

Since I and V both oscillate with the same frequency as the AC source, Eq (7) also gives the ratio of the amplitudes or rms values of corresponding quantities.

Now, we can observe how a transformer affects the voltage and current, we have

$$V_S = \left(\frac{N_S}{N_P}\right) V_P \quad \text{and} \quad I_S = \left(\frac{N_P}{N_S}\right) I_P \longrightarrow (8)$$

That is, if the secondary coil has a greater number of turns than the primary (i.e $N_S > N_P$), the voltage is stepped up ($V_S > V_P$). This type of arrangement is called a step-up transformer. However, in this arrangement, there is less current in the secondary than in the primary (i.e. $N_P/N_S < 1$ and $I_S < I_P$).

⟹ If the secondary coil has less number of turns than the primary (i.e $N_S < N_P$) we have a step-down transformer.

⟹ In this case $V_S < V_P$ and $I_S > I_P$. That is, the voltage is stepped-down, (or reduced) and the current is increased. The equations obtained above apply to ideal transformers (without any energy losses)

* **Energy Loss in Transformers :-**

In actual transformers, small energy losses do occur due to the following reasons.

(i) Flux leakage :- There is always some leakage of flux that is not all of the flux due to primary passes through the secondary. This is due to poor design of the core or the air gaps in the core.

It can be reduced by winding the primary and Secondary coils one over the other.

(ii) Resistance of the windings :- The wire used for the windings has some resistance and so, energy is also lost due to heat produced in the wire (I^2R). In high current, low voltage windings, energy losses are minimised by using thick wire.

(iii) Eddy currents :- The alternating magnetic flux induces eddy currents in the iron core and causes heating. The effect is reduced by having a laminated core.

(iv) Hysteresis :- The magnetisation of the core is repeatedly reversed by an alternating magnetic field. The resulting expenditure of energy in the core appears as heat and is kept to a minimum by using a magnetic material which has a low hysteresis loss.

* Uses of Transformers :-

Transformers are used in almost all AC operations. Some of the following are given below.

(1) In the induction furnaces

(2) In voltage regulator for T.V, refrigerator, computer, air conditioner, etc..

(3) Small transformers are used in Radio sets, telephones, loud speakers and electric bells, etc..

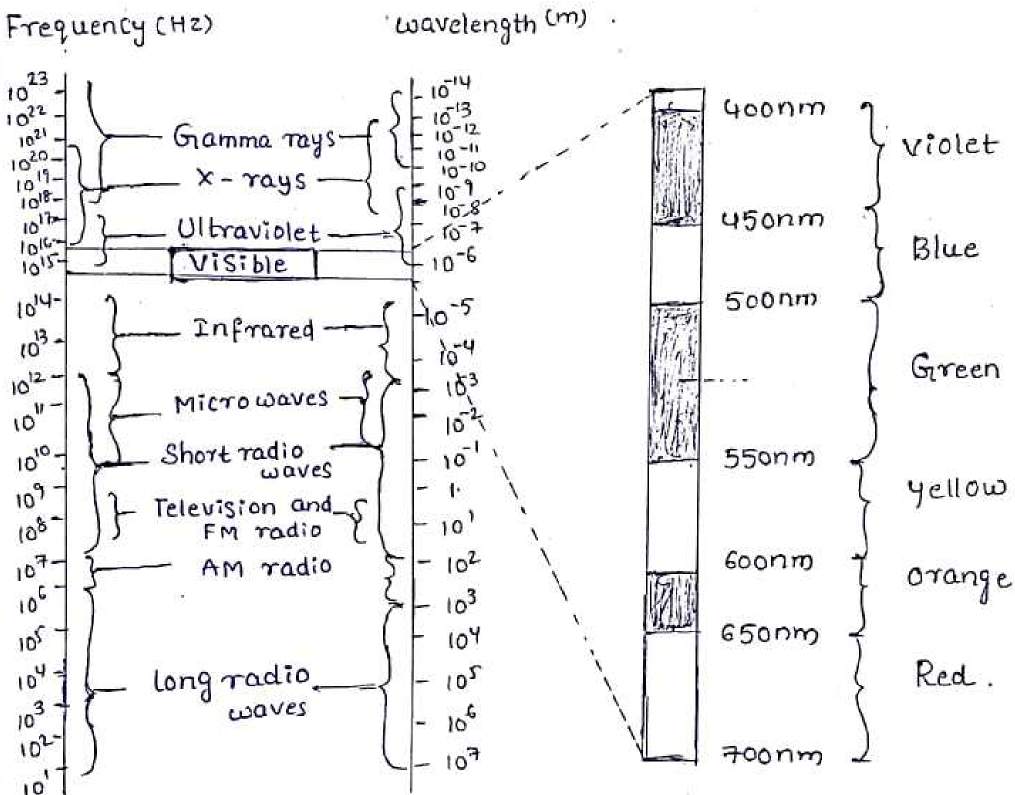
(4) A step down transformer is used for welding purposes.

(5) A step down transformer is used for obtaining large current.

4. Electromagnetic Waves

* **Electromagnetic Spectrum :-**

The orderly arrangement of EM waves in increasing or decreasing order of wavelength λ or frequency ν is called electromagnetic Spectrum. The range varies from 10^{-12} m to 10^{4} m, i.e from γ- rays to radio waves.



Electromagnetic Spectrum with Common names for various parts of it.

The wavelength ranges, frequency ranges and use of various regions of electromagnetic Spectrum are summarised below.

* **Radio waves :-**

These are produced due to oscillating charge particles. The frequency varies from 500 kHz to 1000 MHz.

Uses of radio waves are given below.

(i) These are used in AM (Amplitude Modulation) from 530 KHz to 1710 KHz. These are also used in ground wave propagation.

(ii) These are used in TV waves ranging from 54 MHz to 890 MHz.

(iii) These are used in FM (Frequency Modulation) ranging from 88 MHz to 108 MHz.

(iv) UHF (Ultra High Frequency) waves are used in Cellular phones.

* **Microwaves :-**

These waves are called short wavelength radio waves which are produced by vaccum tubes. Their frequency lies in the range of 1GHz to 300GHz (gigahertz).

Uses of microwaves are given below.

(i) These are used in RADAR systems for aircraft navigation.

(ii) These are used in microwave oven for cooking purpose.

(iii) These are used in study of atomic and molecular structures.

(iv) These are used to measure the speed of vehicle, speed of cricket ball, etc ...

* **Infrared waves :-**

These waves were discovered by Herschell. These waves are called heat waves. These waves are produced from the ~~heatin~~ heat radiating bodies and molecules.

They have high penetration power. Its frequency range is from 3×10^{11} Hz to 4×10^{14} Hz

Uses of infrared waves are given below.

(i) These are used in physical therapy.
(ii) These are used in Satellite for army purpose.
(iii) These are used in weather forecasting.
(iv) These are used for producing dehydrated fruits.
(v) These are used in solar water heater, solar cells and cooker.

* **Visible Rays :-**

It is that part of spectrum which is visible by human eye and its frequency range is from 4×10^{14} Hz to 7×10^{14} Hz.

Uses of visible rays is given below :-

Visible rays are used by the optical organs of humans and animals for three primary purposes given below.

(i) To see things, avoid bumping from them and escape danger.
(ii) To find stuff to eat.
(iii) To find other living things with which to consort so as to prolong the species.

* **Ultraviolet Rays :-**

These rays were discovered by Ritter in 1801. These rays are produced by special lamps and very hot bodies. The Sun is an important source of UV - rays but fortunately absorbed by ozone layer at an altitude of about 40-50 km. Its frequency range is from 10^{14} Hz to 10^{16} Hz.

Uses of ultraviolet rays are given below.

(i) These are used in burglar alaram.
(ii) These are used in checking mineral sample.
(iii) These are used to study molecular structure.
(iv) To kill germs in minerals.

(V) To Sterilise Surgical instruments.

(Vi) These rays can be focussed into very narrow beams for high precision applications such as LASIK eye surgery.

* **X - Rays :-**

These rays were discovered by German professor Roentgen. Its frequency range is from 3×10^{16} Hz to 3×10^{21} Hz.

Uses of x-rays are given below :

(i) These are used in surgery to detect the fracture, diseased organs, stones in the body, etc...

(ii) These are used in engineering to detect fault, crack on bridges, testing of welds.

(iii) These are used at metro station to detect metal or explosive material.

(iv) These are used in scientific research.

* **Gamma (γ) Rays :-**

These rays were discovered by Rutherford. They travel with the speed of light and having high penetration power. The frequency ranges from 3×10^{18} Hz to 5×10^{22} Hz.

Uses of gamma (γ) rays are given below :

(i) These are used to produce nuclear reaction.

(ii) These are used in radio therapy for the treatment of tumour and cancer.

(iii) These are used in food industry to kill pathogenic microorganism.

(iv) These are used to provide valuable information about the structure of atomic nucleus.

Different types of Electromagnetic waves.

Type	Wavelength range	Production	Detection.
Radio wave	> 0.1 m	Rapid acceleration and decelerations of electrons in aerials	Receiver's aerials.
Micro wave	0.1 m to 1 mm	klystron valve or magnetron valve	point contact diodes.
Infrared wave	1 mm to 700 nm	vibration of atoms and molecules	Thermopiles bolometer, Infrared photographic film.
Light	700 nm to 400 nm	Electrons in atoms emit light when they move from higher energy level to a lower energy level	The eye, photocells, photographic film
Ultraviolet rays	400 nm to 1 nm	Inner shell electrons in atoms moving from higher energy level to a lower energy level	photocells, photographic film.
X-rays	1 nm to 10^{-3} nm	x-ray tubes or inner shell electrons	Photographic film, Geiger tubes, Ionisation chamber
Gamma rays	$< 10^{-3}$ nm	Radioactive decay of the nucleus	photographic film, ionisation chamber.

Note :- This EM spectrum and its properties have been frequently asked in previous years 2014, 2013, 2012, 2011, 2010.

Displacement Current :-

Ampere's circuital law states that, the line integral of magnetic field B around any closed path is equal to μ_0 times the total current threading the closed path.

$$\oint B \cdot dl = \mu_0 I \longrightarrow (i)$$

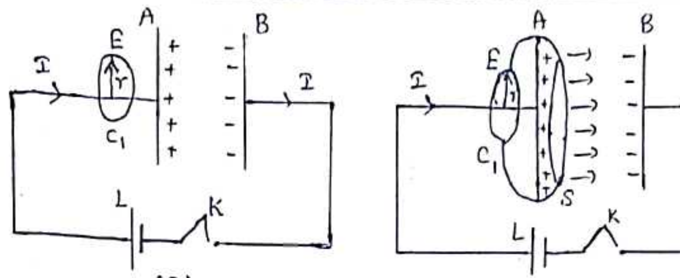
where, I is the net current threading the surface bounded by a closed path C.

Origin of Displacement Current :-

According to Maxwell, the eq (i) is logically inconsistent. with the help of following observations, it explained the same. He considered a parallel plate capacitor having plates A and B connected to a battery L, through a tapping key K. After pressing the key K, the conduction current flows through the connecting wires and the capacitor starts storing charge. As the charge on the capacitor grows, the conduction current in the wire decreases.

When the capacitor is fully charged, the conduction current stops flowing in the wire. But during the charging of capacitor, there is no conduction current between the plates of capacitor. Let at an instant during charging, I be the conduction current in the wires. This current will produce magnetic field around the wires which can be detected by using a compass needle.

After this, the magnetic field was found out at point E, which is at a perpendicular distant r from connecting wire, in a region outside the parallel plate capacitor.



(a)

Circuit diagrams showing the incosisteney of Ampere's circuital law.

For this, a plane circular loop C_1 of radius r is considered. Its centre lies on wire and its plane is perpendicular to the direction of current carrying wire [See fig(a)]. The magnitude of magnetic field is same at all points on the loop and is acting tangentially along the circumference of the loop. If B is the magnitude of magnetic field at E, then by using Ampere's circuital law for loop C_1, we get

$$\oint_{C_1} B \cdot dl = \oint_{C_1} B\,dl \cos 0 = B \times 2\pi r = \mu_0 I$$

$$\Rightarrow \quad B = \frac{\mu_0 I}{2\pi r} \quad \text{----}\!\!\rightarrow \text{(ii)}$$

Now, a different surface, i.e a tiffin box surface is considered. This surface is without lid with its circular rim, which has the same boundary as that of loop C_1 [See fig (b)]

On applying Ampere's circuital law to loop C_1 of this tiffin surface, we get

$$\oint B \cdot dl = B \cdot 2\pi r = \mu_0 \times 0 = 0 \; \longrightarrow \text{(iii)}$$

From Eqs (ii) and (iii), it has been noticed that there is a magnetic field at E calculated through one way and no magnetic field at E, calculated through another way. As this contradiction arises from the use of Ampere's circuital law, hence, Ampere's circuital law is logically inconsistent.

Basic Idea of Displacement Current :-

Since, Ampere's Circuital law for Conduction Current during Charging of a capacitor was found inconsistent. Maxwell argued that the above inconsistency of Ampere's Circuital Law is because of some missing term. The term must be related to a Changing electric field which passes through Surface S between the plates of Capacitor during Charging. So, maxwell introduced this missing term, i.e. displacement current, In order to make Ampere's Circuital law logically Consistent.

Displacement current is that current which comes into play in the region in which the electric field and the electric flux is Changing with time.

i.e Displacement current, $\boxed{I_d = \epsilon_0 \frac{d\phi_E}{dt}}$

Ampere's circuital law $\left(\oint B.dl = \mu_0 I\right)$ was modified to $\boxed{\oint B\cdot dl = \mu_0 (I_c + I_d)}$

where, I_c = Conduction Current and

I_d = displacement current.

→ It is called modified Ampere's Circuital law or Ampere Maxwell's - Circuital law.

Therefore, modified Ampere's Circuital Law may also be expressed as

$$\boxed{\oint B.dl = \mu_0 \quad I_c + \epsilon_0 \frac{d\phi_E}{dt}}$$

The ~~interference~~ inferences Can be drawn from the above discussions as given below.

(i) The conduction and displacement currents are individually discontinuous ; but the currents together possess the property of continuity through any closed electric circuit.

(ii) The displacement current is precisely equal to the conduction current, when the two present in different parts of the circuit.

(iii) The displacement current arises due to rate of change of electric flux (or electric field) between the two plates of the capacitor.

(iv) Just as the conduction current, the displacement current is also the source of varying magnetic field.

* Maxwell's Equations :-

Maxwell's equations are the basic laws of electricity and magnetism. These equations give complete description of all electromagnetic interactions. Maxwell on the basis of his equations, predicted the existence of electromagnetic waves.

There are four Maxwell's equations which are explained as given below.

Gauss's Law of Electrostatics :-

This law states that, the total electric flux through any closed surface is always equal to $\frac{1}{\epsilon_0}$ times the net charge enclosed by that surface. It is given by

$$\boxed{\oint E.ds = \frac{q}{\epsilon_0}}$$

This equation is called Maxwell's first equation.

* **Gauss's Law in Magnetostatics :-**

This law states that, the net magnetic flux through any closed surface is always zero. It is given by

$$\boxed{\oint B \cdot ds = 0}$$

This equation is called Maxwell's Second equation.

* **Faraday's Law of Electromagnetic Induction :-**

This law states that, the induced emf produced in a circuit is numerically equal to rate of change of magnetic flux through it. It is given by

$$\boxed{\oint E \cdot dl = -\frac{d\phi_B}{dt}}$$

This equation is called Maxwell's third equation.

* **Ampere - Maxwell's Circuital Law :-**

This law states that, the line integral of the magnetic field along a closed path is equal to μ_0 times the total current (i.e Sum of Conduction Current and displacement Current) threading the surface bounded by that closed path. It is given by

$$\boxed{\oint B \cdot dl = \mu_0 \left(I_C + \epsilon_0 \frac{d\phi_E}{dt} \right)}$$

This equation is called Maxwell's fourth equation.

* **Electro Magnetic Waves :-**

These waves are produced due to the change in electric field E and magnetic field B sinusoidally and propagating through space such that, the two fields are perpendicular to each other and perpendicular to the direction of wave propagation.

Source of Electromagnetic waves :-

An oscillating charge is an example of accelerating charge. It produces an oscillating electric field in Space, which produces an oscillating magnetic field, which in turn produces an oscillating electric fields and so on. The oscillating electric and magnetic fields regenerate each other as a wave which propagates through space.

The frequency of EM wave is equal to the frequency of oscillation of charge, i.e.

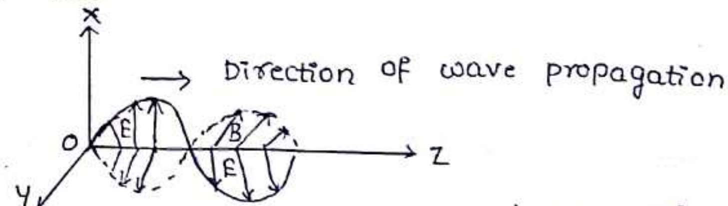
$$v = \frac{1}{2\pi\sqrt{LC}}$$

Electromagnetic waves are also produced when fast moving electrons are suddenly stopped by metal target of high atomic number. These electromagnetic waves are called x-rays.

Transverse Nature of Electromagnetic waves :-

It can be shown from Maxwell's equations that electric and magnetic fields in an electromagnetic wave are perpendicular to each other and to the direction of wave propagation.

It was seen in the discussion of the displacement current also. If we would consider a parallel plate capacitor, the E inside the Parallel-plate capacitor was directed perpendicular to the plates. Also, the B which give rise to the displacement current was parallel to the Capacitor. Thus E and B were perpendicular in that case. But this observation is a general feature.



A plane EM wave travelling along z-axis.

In the above figure, we see that permanent curve shows electric field E which is along x-direction and dotted curve shows Magnetic field B which is along y-direction and the wave propagates along z-direction. Both E and B vary sinusodially and become maximum at same position and time.

Since, in electromagnetic wave, E and B are mutually perpendicular to each other, so they are transverse in nature.

The EM wave propagating in the positive z-direction may be represented by the following equations.

Here $E = E_x = E_0 \sin(kx - \omega t)$

$B = B_y = B_0 \sin(kz - \omega t)$

where $k = 2\pi/\lambda$ [λ = wavelength]

$\omega = 2\pi v$ [v = frequency]

E_0 = amplitude of varying electric field.

and B_0 = amplitude of varying magnetic field.

* Important Characteristics of ElectroMagnetic waves :-

Important characteristics of EM waves are listed below.

(i) The electromagnetic waves are produced by accelerated charge.

(ii) These waves do not require any material medium for propagation.

(iii) These waves travel in free space with the speed of light (3×10^8 m/s) given by the relation $c = 1/\sqrt{\mu_0 \epsilon_0}$. It means that light waves are electromagnetic in nature.

(iv) Speed of electromagnetic wave in a medium is given by, $v = 1/\sqrt{\mu \epsilon}$, where ϵ and μ are the permittivity and magnetic permeability of a material medium, respectively. This means, the speed of EM wave in a medium depends on electric and magnetic properties of a medium.

(v) The direction of variations of electric and magnetic fields are perpendicular to each other and also perpendicular to the direction of wave propagation. Thus electromagnetic waves are transverse in nature.

(vi) In free space, the magnitudes of electric and magnetic fields in electromagnetic waves are related by $E_0 / B_0 = c$.

(vii) The energy in electromagnetic waves is divided, on an average, equally between electric and magnetic fields. $U_e = U_m$

where U_e = energy of electric field

and U_m = energy of magnetic field.

(viii) The energy density (energy per unit volume) in an electric field E in vaccum is $\frac{1}{2}\epsilon_0 E^2$ and that in magnetic field B is $\frac{B^2}{2\mu_0}$

$\therefore$ Energy associated with an electromagnetic wave is given by $\boxed{U = \frac{1}{2}\epsilon_0 E^2 + \frac{1}{2}\cdot\frac{B^2}{\mu_0}}$

Also, average energy density,

$$u_{av} = \frac{1}{4}\epsilon_0 E_0^2 + \frac{1}{4}\frac{B_0^2}{\mu_0}$$

also $$u_{av} = \frac{1}{2}\epsilon_0 E_0^2 = \frac{B_0^2}{2\mu_0}$$

(ix) Electromagnetic waves, being uncharged, are not deflected by electric and magnetic fields.

(x) The electromagnetic wave like other waves carries energy and momentum. Since, it has momentum, an electromagnetic wave also exerts pressure called radiation pressure.

If wave is incident on a completely absorbing surface, then momentum delivered is given by

momentum, $p = \frac{U}{c}$

Note :- Light carries energy from the Sun to the earth, thus making life possible on the earth.

(xi) Electromagnetic waves are polarised and can be easily seen in the response of a portable AM radio to a broadcasting station. If an AM radio has a telescopic antenna, it responds to the electric part of the signal. When the antenna is turned horizontal, the signal will be greatly diminished.

5. Ray Optics and Optical instruments

* ## Ray optics :-

The branch of Study of light is called optics. Broadly optics is divided into three groups.

(i) Geometrical optics (Ray optics)

(ii) Wave optics

(iii) Quantum optics.

Geometrical optics (Ray optics) :-

In this, light is considered as a ray which travels in Straight line. Geometrical optics States that for each and every object, there is an image. It works on the following assumptions:

(i) Rectilinear propagation of light, i.e light ray travels in straight line.

(ii) Laws of reflection.

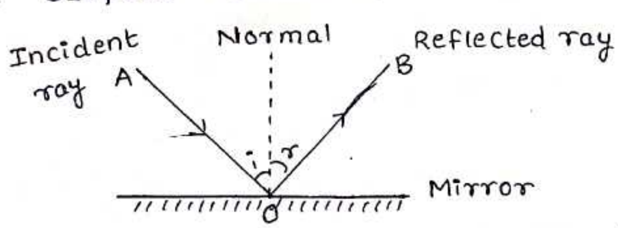
(iii) Laws of refraction.

(iv) Physical independence of light rays i.e two light rays are totally independent of each other.

* ## Reflection of Light :-

Reflection is the phenomenon of change in the path of light without any change in the medium.

The returning back of light in the Same medium from which it has come after Striking a Surface is called reflection of light.



The incident ray reflected ray and the normal to the reflecting Surface lie in Same plane.

Scanned by CamScanner

Laws of Reflection :-

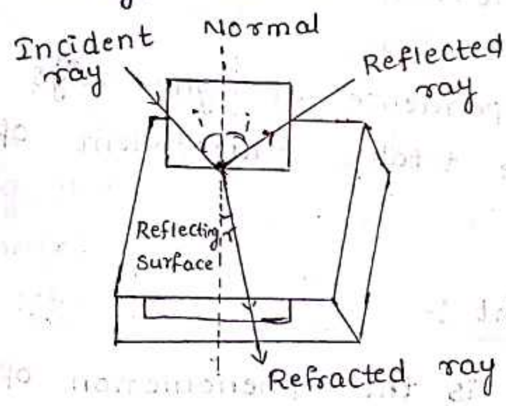
The laws of reflection are as given below

(i) The incident ray, the reflected ray and the normal to the reflecting surface at the point of incidence, all lie in the same plane.

(ii) The angle of reflection (r) is equal to the angle of incidence (i) i.e. $i = r$. For normal incidence $\angle i = 0°$, therefore $\angle r = 0°$. Hence, a ray of light falling normally on a mirror, retraces its path on reflection.

* Refraction :-

Reflection involves change in path of light without any change in the medium, whereas refraction involves change in the path of light due to change in the medium.



When a beam of light encounters another transparent medium, a part of light gets reflected back into the first medium while the rest enters the other. The direction of propagation of an obliquely incident ray of light, that enters the other medium, changes at the interface of two media. This phenomenon is called refraction of light.

Laws of Refraction :-

(i) The incident ray, the refracted ray and the normal to the refracting Surface (or interface) at the point of incidence, all lie in the Same plane.

(ii) The ratio of the Sine angle of incidence to the Sine angle of refraction is Constant for the two given media. This Constant is denoted by ${}^a\mu_b$ and is called the relative refractive index of medium b with respect to medium a.

$$\boxed{\frac{\sin i}{\sin r} = {}^a\mu_b}$$

This law is also called Snell's law of refraction.

* ## Refractive Index :-

The refractive index or index of refraction 'μ' of a material is the ratio of the Speed of light (c) in vaccum to the Speed of light in the medium (v).

Mathematically, refractive index is given by the relation,

$$\mu = \frac{\text{Speed of light in vacuum}}{\text{Speed of light in the material}} = \frac{c}{v}$$

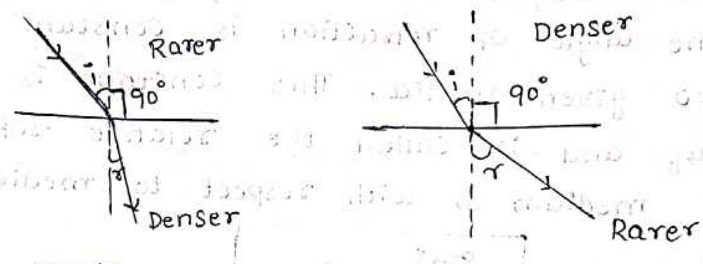
It is also reffered as absolute refractive index of the substance.

Refractive Index of Some Substance Media.

Substance medium	Refractive Index
Ethyl alcohol	1.362
water, H_2O	1.333
Air	1.000293
oxygen, O_2	1.000271

Relative refractive index is a measure of how much light bends, when, it travels from one medium to another medium.

If light travels from optical rarer medium to optical denser medium, then it bends towards the normal, i.e. $i > r$. On the other hand, if light travels from optical denser medium to optical rarer medium, then light bends away from the normal, i.e. $i < r$.

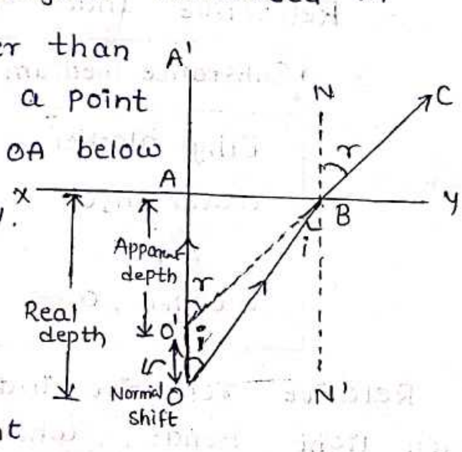


The medium in which the speed of light is higher with respect to other medium, it is said to be optically denser medium. Optical density is the ratio of the speed of light in two media.

Optical density should not be confused with mass density, which is mass per unit volume. It is possible that, mass density of an optically denser medium may be less than that of an optically rarer medium. e.g. Turpentine and water. Mass density of turpentine is less than that of water but its optical density is higher.

Apparent Depth and Normal Shift :-

The depth of an object immersed in water appears to be lesser than its actual depth. Let O be a point object at an actual depth OA below the free surface of water xy.



A ray of light incident normally on xy, along OA passes straight along OAA'. Another ray of light from O incident at $\angle i$ on surface xy along OB deviates away from normal. It is refracted at $\angle r$ along BC. on producing backwards BC meets OA at O'. Therefore O' is virtual image of O.

Apparent depth = AO'

Real depth = OA   Clearly, AO' < OA

Now, $\angle BOA = \angle OBN' = i$ (alternate angles)

$\angle AOB' = \angle CBN = r$ (corresponding angles)

In ΔOAB, $\sin i = \frac{AB}{OB}$

In $\Delta O'AB$, $\sin r = \frac{AB}{O'B}$

As, light ray is travelling from denser medium to rarer medium.

$$^{a}\mu_{\omega} = \frac{\sin r}{\sin i}$$

or. $$^{a}\mu_{\omega} = \frac{AB}{O'B} \times \frac{OB}{AB} = \frac{OB}{O'B}$$

B is close to A (as angles are very small).

So, $OA \approx OB$ and $O'A \simeq O'B$

$$\therefore \boxed{^{a}\mu_{\omega} = \frac{OA}{O'A} = \frac{\text{Real depth}}{\text{Apparent depth}}}$$

If x is the real depth of water surface and $^{a}\mu_{\omega}$ is the refractive index of water with respect to air, then the normal shift (d) in position of point object is given by

$$d = \text{Real depth} - \text{Apparent depth}$$

$$\therefore d = x - \frac{x}{^{a}\mu_{\omega}} \quad \left[\because \text{apparent depth} = \frac{\text{real depth}}{^{a}\mu_{\omega}} = \frac{x}{^{a}\mu_{\omega}}\right]$$

or $$\boxed{d = x\left(1 - \frac{1}{^{a}\mu_{\omega}}\right)}$$

* Ex:- Velocity of light in glass is 2×10^{8} m/s and that in air is 3×10^{8} m/s. By how much would an ink dot appear to be raised, when covered by a glass plate 6 cm thick?

Sol

Given, Velocity of light in glass, $v = 2\times10^8\ m/s$

Velocity of light in air, $c = 3\times10^8\ m/s$

$\therefore$ Refractive index of glass w.r.t air

$${}^{a}\mu_g = \frac{c}{v} = \frac{3\times10^8}{2\times10^8} = 1.5$$

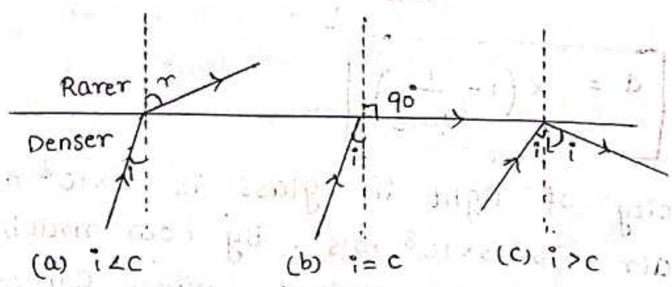
$\therefore$ Normal Shift in the position of ink dot

$$d = t\left(1 - \frac{1}{{}^{a}\mu_g}\right) \quad \because (t = 6\,cm)$$

$$= 6\left(1 - \frac{1}{1.5}\right) = \frac{6\times0.5}{1.5} = 2\,cm$$

Total Internal reflection :-

When a ray of light travels from a denser to rarer medium, it bends away from the normal. The angle of refraction is greater than the angle of incidence. As the angle of incidence increases, the angle of refraction also increases. For a particular angle of incidence, the angle of refraction becomes 90° and the refracted ray grazes the surface separating the two media. This angle of incidence is called the critical angle and is defined for a pair of media. If the angle of incidence is increased beyond the critical angle, the ray is not refracted but gets reflected as shown in Fig. The entire incident light is reflected back into the rarer medium. This phenomenon is called total internal reflection.



Total Internal reflection

Critical angle for a pair of media is that angle of incidence in the denser medium for which the angle of refraction is 90°, i.e. the refracted ray just grazes the surface separating the two media.

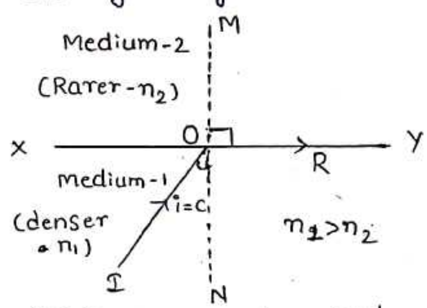
Conditions for total Internal Reflection :-

(1) Light should tend to travel from an optically denser medium to an optically rarer medium.

(2) The angle of incidence in the optically denser medium must be greater than the critical angle for the given pair of media.

Relation between refractive index and critical angle :-

In the fig. xy is a surface separating two media of refractive indices n_1 and n_2 $(n_1 > n_2)$. Let C be the critical angle for the given pair of media. Consider a ray IO incident at an angle C in medium-1. It is refracted as OR grazing the surface of separation xy.



(Refractive Index and Critical angle).

Now $i = c$ and $r = 90°$

From Snell's law,

$$n_1 \sin i = n_2 \sin r$$

$$n_1 \sin C = n_2 \times 1 \quad (\because \sin 90° = 1)$$

$$\sin C = \frac{n_2}{n_1}$$

$$\sin C = \frac{\text{Refractive index of rarer medium } (n_r)}{\text{Refractive index of denser medium } (n_d)}$$

$$\sin C = \frac{n_r}{n_d} \quad . \quad \text{Thus } n_d = \frac{1}{\sin C}$$

If the rarer medium is air or vacuum,

${}_r n_d = n =$ absolute refractive index of the medium

$$n = \frac{1}{\sin C}$$

where C is the critical angle for the pair of media.

* **Applications of Total Internal reflection :-**

(1) Sparkling of diamond :- The refractive index of diamond is about 2.42 and its critical angle is about 24°. The faces of a diamond are Suitably cut so that a ray of light enters the diamond undergoes a series of total internal reflections before it emerges out. In addition, the ray also undergoes dispersion. This accounts for its brilliance.

(2) An empty test tube immersed in water appears to be Silvery when viewed from the Side at a Suitable angle.

(3) Mirages :- A mirage is an optical illusion caused due to refraction and internal reflection. There are two types of mirages namely mirage of inferior type and mirage of Superior type.

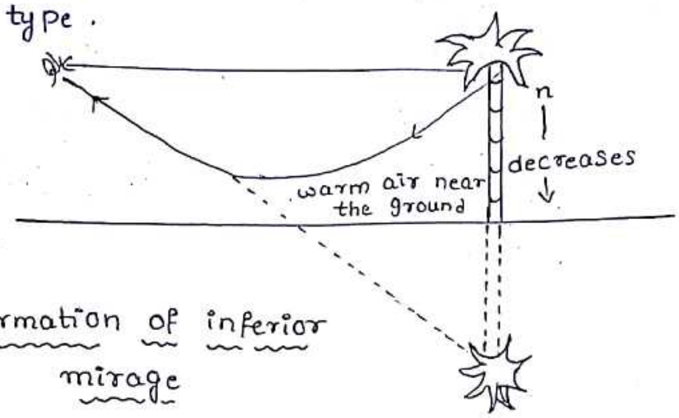
(a) Inferior Mirage :-

An inferior mirage is usually seen in deserts and on tar roads on hot afternoons. The formation of mirages in deserts can be explained as follows.

In deserts, the layers of air closer to the Sand are hotter and hence rarer than those farther away. Rays of light from a distant objects like a tree, travelling downwards undergoes a Series of refractions. At each refraction, the rays travel from a denser to a rarer medium. Hence they bend away from the normal. when they reach the lower layers, they may be incident at angles greater than the critical angle. Hence, they undergo total Internal reflection. Thus, one Can see an inverted image of the tree as if it is reflected by a pool of water.

Scanned by CamScanner

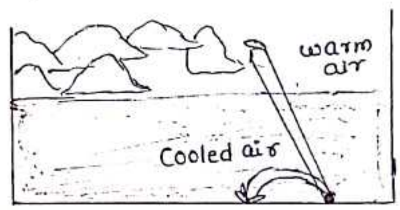
Further due to flumctuations in the density of air, the image appears to quiver, giving an impression of reflection by a disturbed water Surface. This illusion is called a mirage. This is called a mirage of the inferior type.



Formation of inferior mirage

(b) Superior mirage :-

A Superior mirage occurs under reverse atmospheric conditions from that of inferior mirage. For it to be seen, the air close to the Surface must be much Colder than the air above it. This condition is common over Snow, ice and cold water surfaces.



Superior mirage.

when very cold air lies below warm air, light rays are bent downward towards the Surface, thus tricking our eyes into thinking that an object is located higher or taller in appearance than it actually is.

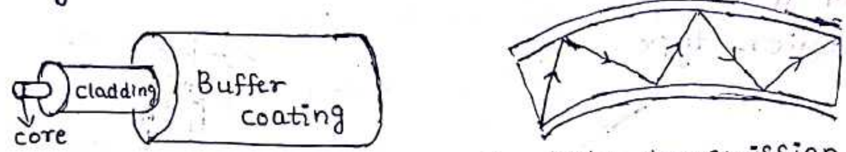
* Optical Fibres :-

John Tyndall (1870), a British physicist, demon-Strated that light could be transmitted along a curved path using a water jet. This remarkable idea was recognized and applied by Kapany and Co-workers (1960) in the form of optical fibres.

An optical fibre is a thin transparent fibre with refractive index greater than that of the Surrounding material. It works on the principle of total internal reflection.

Structure of an optical fibre :-

Fig shows the structure of an optical fibre and light transmission through it.



(a) Structure of optical fibre.

(b) Light transmission through an optical fibre.

Fig. Optical Fibres.

(1) Core :- core is made of glass or silica and its diameter is in the range 10 µm to 100 µm

(2) Cladding :- The core is surrounded by glass or plastic material known as cladding whose refractive index is lower than that of the core.

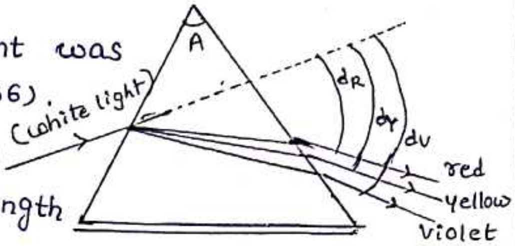
(3) Buffer :- For providing safety and strength, the core-cladding system is covered with a plastic coating known as buffer. This also provides optical insulation when hundreds of fibres are packed into a cable. The refractive index of the buffer material is less than that of the cladding.

A bundle of optical fibres is called a light pipe or an optical cable. For proper working of an optical cable, the optical fibres (a) must be thin (b) insulated from each other and (c) parallel to each other.

When a ray of light entering the fibre from one end is incident on the core-clad interface at an angle greater than the critical angle for the fibre, it undergoes total internal reflection several times before emerging out at the other end of the fibre with negligable loss of light energy.

* ## Dispersion of light :-

Dispersion of sunlight was first observed by Newton (1666). Refractive index of the material of a prism depends on wavelength of light. Thus, when a beam



Dispersion of white light

of composite light is incident on a prism, light of different wavelengths emerge in different directions. The emergent beam consists of an ordered sequence of wavelengths. An ordered sequence of wavelength is called a spectrum.

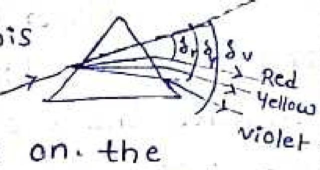
The phenomenon of splitting up of composite light into its constituent wavelengths is called dispersion.

Fig shows dispersion of white light. White light is composed of wavelengths ranging from 400nm to 700nm. This range of wavelengths is called visible range. In this range, a group of certain wavelengths produce the sensation of a particular colour. For instance, wavelength close to 700 nm produce sensation of red and those close to 400nm produce the sensation of violet. Thus, if a screen is inserted in the path of the emergent beam, an ordered bands of colours can be seen on the screen. This array of coloured bands is the spectrum.

Dispersive power :-

Consider a prism of a transparent material. when a beam of white light is passed through the prism, light of different wavelengths are deviated by different amounts. The overall deviation of the light beam is measured by the deviation of the yellow light as this deviation is roughly the average of all deviations. In figure below, this deviation is shown by the symbol δ_y. It is

Clear that if δ_r and δ_v are the deviations for red and violet components, the angular divergence of the transmitted beam is $\delta_v - \delta_r$. This divergence is called angular dispersion.



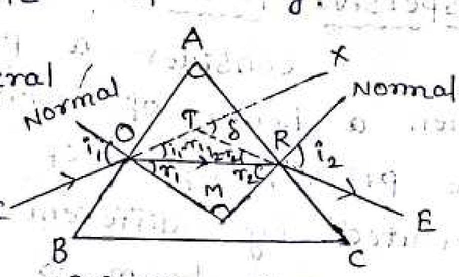
The mean deviation depends on the average refractive index n, where as the angular dispersion depends on the difference $n_v - n_r$.

Def :- The dispersive power of a material is defined as the ratio of angular dispersion to the average deviation when a light beam is transmitted through a thin prism placed in a position so that the mean ray (ray having the mean wavelength) passes symmetrically through it.

When a light ray passes symmetrically through a prism of refracting angle A, it suffers minimum deviations δ.

Refraction of monochromatic light through prism :-

ABC is the principal section of a prism of refractive index n. A ray of monochromatic light IO incident on the refracting face AB, refracts along OR and emerges along RE, as in the fig. let i_1 and r_1 be the angles of incidence and refraction respectively at AB. let r_2 and i_2 be the angle of incidence and emergence respectively at AC. OM and RM are the normals at O and R to AB and AC respectively.



(Refraction of monochromatic light through a prism)

In the cyclic quadrilateral AOMR, $\angle A + \angle M = 180^\circ \longrightarrow (1)$

In ΔORM, $r_1 + r_2 + \angle M = 180^\circ \rightarrow (2)$

from eqs (1) and (2)

$A = r_1 + r_2 \longrightarrow (3)$

The angle d between incident direction IOX and the direction of emergence TRE is the angle of deviation. From fig we have

deviation produced at AB, $\delta_1 = i_1 - r_1$

deviation produced at AC, $\delta_2 = i_2 - r_2$

Total deviation produced, $\delta = \delta_1 + \delta_2 = i_1 - r_1 + i_2 - r_2$

$$\delta = i_1 + i_2 - (r_1 + r_2)$$

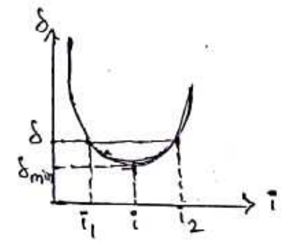
$$\delta = i_1 + i_2 - A \quad (\because A = r_1 + r_2)$$

Therefore, $A + \delta = i_1 + i_2 \longrightarrow (4)$

Angle of minimum deviation :-

A graph of δ vs i as shown in fig. It is found that as the angle of incidence i is increased from a small value, the angle of deviation δ first decreases to a minimum value and then increases.

The smallest angle of deviation for a ray passing through a prism is called the angle of minimum deviation.



From fig, it is clear that for any particular angle of deviation, two angles of incidence are possible.

Hence at minimum deviation, a ray of light passes through the prism symetrically and the refracted ray is parallel to the base. At minimum deviation, that is when $\delta = \delta_m$

$$i_1 = i_2 = i \quad \text{and} \quad r_1 = r_2 = r \quad \text{(from graph)}$$

From Eqs (3) and (4) we get

$$A = 2r \Rightarrow r = \frac{A}{2} \quad \text{and}$$

$$A + \delta = 2i \Rightarrow i = \frac{A+\delta}{2}$$

From Snell's law, $n = \dfrac{\sin i}{\sin r}$

$$\boxed{n = \frac{\sin\left(\frac{A+\delta}{2}\right)}{\sin A/2}}$$

Derivation of a ray due to thin prism :-

A prism whose refracting angle is less than about 10° is considered as a thin prism.

For a thin prism, the angle of minimum deviation δ_m is very small

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

$$n = 1 + \frac{\delta_m}{A}$$

$$\delta_m = (n-1)A$$

Thus, for small angle of incidence i, the deviation δ is almost equal to δ_m. Therefore, $\delta = (n-1)A$

deviation of violet ray $\delta_V = (n_V - 1)A$ and

deviation of red ray, $\delta_R = (n_R - 1)A$

Since $n_V > n_R$ we get $\delta_V > \delta_R$.

Thus, violet is most deviated and red is least deviated.

The angular dispersion is $\delta_v - \delta_r = (n_v - n_r)A$.

The average deviation is $\delta_y = (n_y - 1)A$.

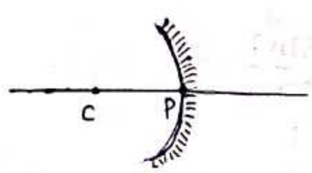
Thus, the dispersive power of the medium is

$$\boxed{\omega = \frac{n_V - n_R}{n_y - 1}}$$

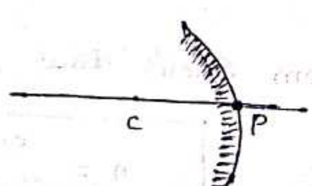
This equation itself may be taken as the definition of dispersive power.

* Spherical Mirrors :-

A Spherical mirror is a part of a hollow sphere which can reflect light. There are two types of spherical mirrors - concave mirrors.



(a) Concave mirror.



(b) Convex mirror.

Scanned by CamScanner

The Surface area of a Spherical mirror available for reflection is called aperture. The geometric centre of the Spherical reflecting Surface (P) is called pole.

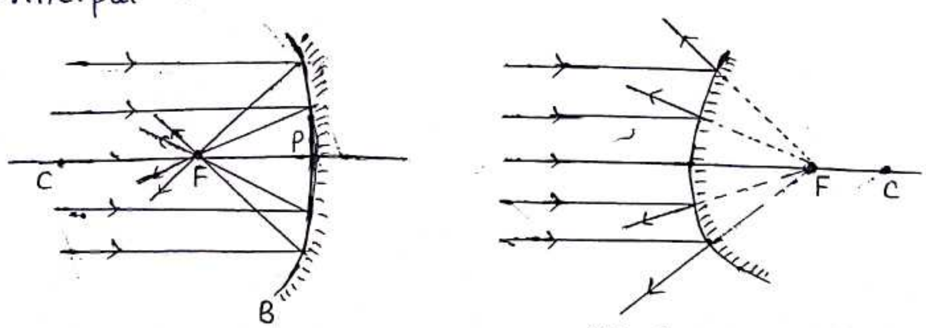
⇒ The centre of the Sphere of which the Spherical mirror is a part (C) is called its center of curvature.

⇒ The radius of the Sphere of which the Spherical mirror is a part (R) is called its radius of curvature.

⇒ A Straight line passing through the pole and Centre of curvature of a mirror is called its principal axis. [Fig (a) & (b)].

When a narrow parallel beam of light is
incident along the principal axis, the reflected rays
either converge to a point on the principal axis of
a concave mirror. (This point is called) or appear to
diverge from a point on the principal axis of
a convex mirror. This point is called the principal
focus (F) of the mirror. The distance between pole
of a mirror and its principal focus is called the
focal length (f) of the mirror.

A plane through the principal focus and
perpendicular to the principal axis is called focal plane.
Rays which are close to the principal axis and make
small angles with it. i.e., almost parallel to the
principal axis, are called paraxial rays.

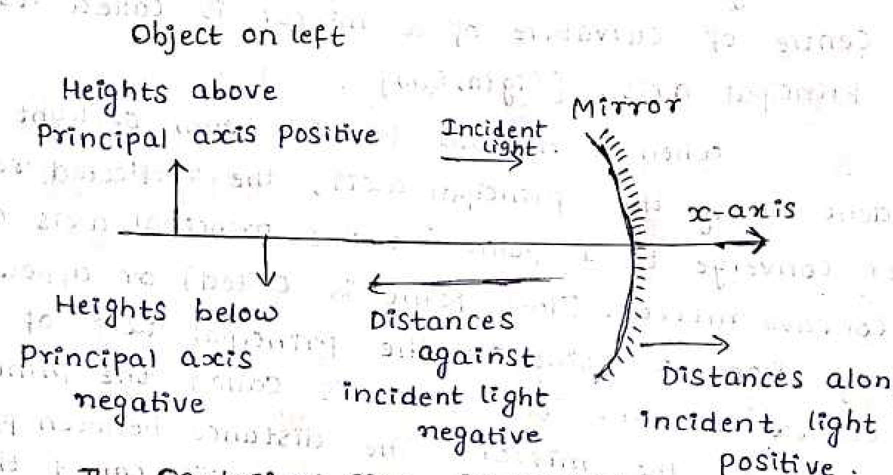


(a) Concave mirror ,   (b) concave mirror.

Fig. principal focus of a spherical mirror.

Sign Convention :-

Use of mirrors for specific application involves a thorough understanding of nature and positions of images and their distances from the mirrors for different positions of object. This is possible only with a well defined convention to deal with signs assigned to various quantities such as object distance, image distance, focal length, etc..



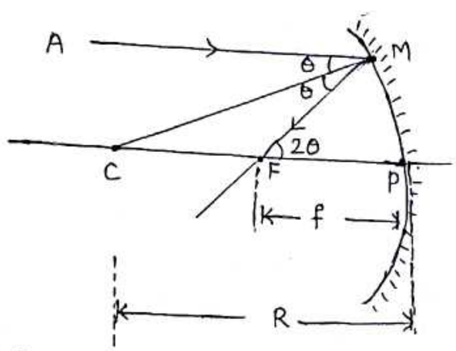
The Cartesian Sign Convention.

The following cartesian sign convention is followed in the study of reflection by spherical mirror as well as refraction at spherical surfaces.

* Distances measured along the principal axis are measured from the pole..
* Distances measured in the direction of incident light are taken positive. (Incident light direction is taken from left to right so that distances in positive x-direction are positive).
* Distances measured opposite to the direction of incident light are taken as negative (Distances in negative x-direction are negative)
* Heights of objects and images above the principal axis are taken positive while heights of objects and images below the principal axis are taken (-ve).

Relation between focal length and radius of curvature :-

Consider a concave mirror of a focal length f and radius of curvature R. let P represent its pole, F represent its principal focus and C its centre of curvature. A ray AM parallel to the principal axis of a concave mirror (of small aperture) is incident at M. The ray is reflected along MF.



Reflection from a concave mirror.

AM - incident ray
MF - reflected ray
C - Centre of curvature
F - principal focus
P - pole
CM - Normal to mirror at M
$\angle AMC = \theta$ = angle of incidence.

By law of reflection, angle of reflection and angle of incidence are equal.

$\therefore \quad \angle CMF = \angle AMC = \theta$

From fig, $\angle MCF = \angle AMC = \theta$ (alternate angles)

Also, $\angle MFP = \angle CMF + \angle MCF = 2\theta$ (Exterior angles = Sum of the interior opposite angles)

As the aperture is small and angle θ is small, are PM may be considered as a straight segment.

From ΔPMC, $\tan\theta = \frac{PM}{PC} \Rightarrow \theta = \frac{PM}{-R} \longrightarrow (1)$

$\left[\because \lim_{\theta\to 0}\left(\frac{\tan\theta}{\theta}\right) = 1 \Rightarrow \tan\theta = \theta, \theta \text{ is rad}\right]$

From ΔMFP, $\tan 2\theta = \frac{PM}{PF} \Rightarrow 2\theta = \frac{PM}{-f} \longrightarrow (2)$

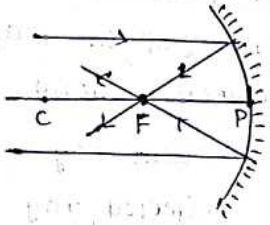
Hence $\frac{(1)}{(2)} \Rightarrow \frac{\theta}{2\theta} = \frac{(PM/-R)}{(PM/-f)} \Rightarrow \frac{1}{2} = \frac{1/R}{1/f}$

$\Rightarrow \frac{1}{2} = \frac{f}{R} \qquad \boxed{\therefore \quad f = \frac{R}{2}}$

* Image formation in a mirror :-

(1) Position of object :- At infinity .

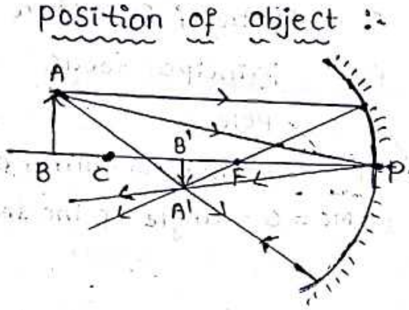
Ray diagram :-



Position of Image :-

(1) At F
(2) Real,
(3) inverted
(4) highly diminished .

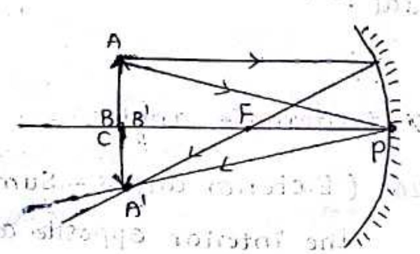
(2) Position of object :- Beyond Centre of Curvature C .



Position of Image :-

(1) Between C and F
(2) Real
(3) Inverted
(4) diminished .

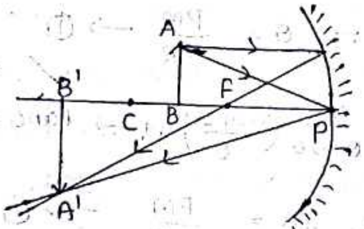
(3) Position of object :- At C .



Position of Image :-

(1) At C
(2) Real
(3) Inverted
(4) Same size .

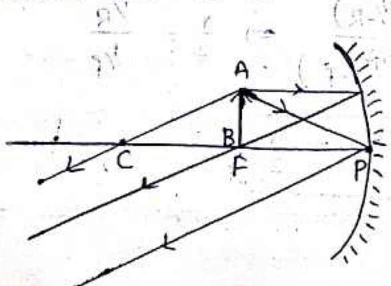
(4) Position of object :- Between F and C .



Position of Image :-

(1) Beyond C
(2) Real
(3) Inverted
(4) enlarged

(5) Position of object :- At F .

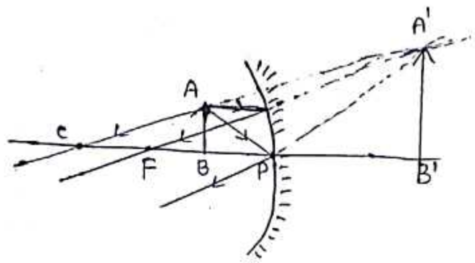


Position of Image :-

(1) At ∞ (infinity)
(2) Real
(3) Inverted
(4) enlarged .

Scanned by CamScanner

(6) Position of object :- within F.

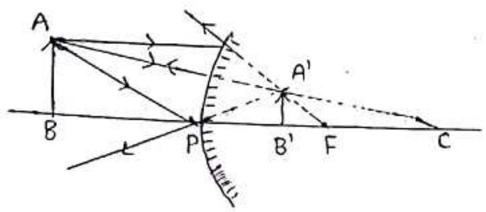


Position of Image :-

(1) Behind the mirror
(2) virtual
(3) errect
(4) enlarged.

* Convex mirror :-

Object any where on the axis :-



Position of Image :-

(1) Behind the mirror
(2) virtual
(3) errect
(4) diminished.

* Mirror equation (mirror formula) :-

The mirror equation gives the relation between object distance, image distance and focal length in the case of a Spherical mirror.

The relationship is $\frac{1}{v}+\frac{1}{u} = \frac{1}{f}$.

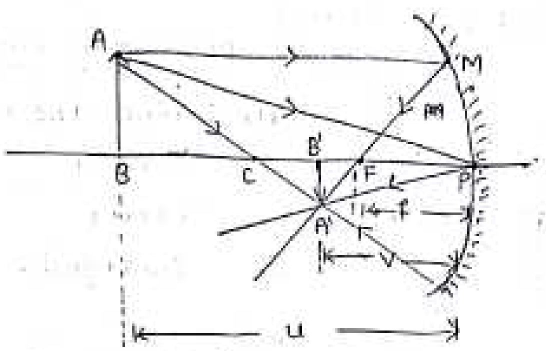
where u is the object distance, v is the image distance and f is the length of Spherical mirror.

(a) Mirror formula in the Case of a Concave mirror producing a real image :-

AB is an object kept a distance u from the pole P beyond the centre of curvature C of a Concave mirror. A ray of light AM parallel to the principal axis after reflection passes through the principal focus F.

A ray AP undergoes reflection such that i=r. A ray through the Centre of Curvature C strikes the mirror normal to it and hence retraces the path. The reflected rays meet at A'. A'B' is the image of AB. The distance of image A'B' is v from the pole P.

Scanned by CamScanner



u	-ve
v	-ve
R	-ve
f	-ve

Concave mirror producing a real image :

ΔABC and $\Delta A'B'C$ are similar

$$\frac{AB}{A'B'} = \frac{BC}{B'C} \longrightarrow (1)$$

ΔABP and $\Delta A'B'P$ are similar

$$\frac{AB}{A'B'} = \frac{PB}{PB'} \longrightarrow (2)$$

From Eqs (1) and (2) we get

$$\frac{BC}{B'C} = \frac{PB}{PB'}$$

$$\Rightarrow \frac{PB - PC}{PC - PB'} = \frac{PB}{PB'} \longrightarrow (3)$$

According to sign convention,

Object distance $u = -PB$, Image distance, $v = -PB'$,

and Radius, $R = -PC$

Hence $PB = -u$, $PB' = -v$ and $PC = -R$

Therefore Eq (3) becomes $\frac{-u+R}{-R+v} = \frac{-u}{-v}$

$$uv - vR = uR - uv$$

$$(u+v) R = 2uv$$

Dividing through out by uvR we get

$$\frac{1}{v} + \frac{1}{u} = \frac{2}{R} \quad \text{But } R = 2f$$

Hence $\frac{1}{u} + \frac{1}{v} = \frac{1}{f}$

Scanned by CamScanner

* **Linear (lateral) magnification (m) :-**

The linear magnification produced by a spherical mirror is defined as the ratio of height of image (h_i) to the height of the object (h_o)

$$m = \frac{h_i}{h_o}$$

h_i and h_o will be taken as positive or negative according to the cartesian sign conventions.

In case of concave mirror producing real image as shown in fig. we find that triangles ABP and A'B'P are similar.

$$\frac{AB}{A'B'} = \frac{PB}{PB'}$$

$$\frac{h_o}{h_i} = \frac{u}{v}$$

Applying sign convention u is -ve, v is -ve, h_o is +ve and h_i is -ve.

$$\therefore \frac{h_o}{-h_i} = -\frac{v}{u} \text{ (Magnification in terms of u and v)}$$

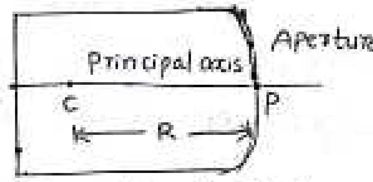
Thus $\boxed{m = -\frac{v}{u}}$

The above equation is also valid for cases of a concave mirror producing a virtual image and a convex mirror producing virtual image.

Relation b/w m, u and f	Relation b/w m, v and f.
1) From the mirror equation $\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$ 2) Multiplying throughout by u $\frac{u}{v} + 1 = \frac{u}{f}$ 3) Magnification, $m = \frac{-v}{u}$ Therefore, $\frac{-1}{m} + 1 = \frac{u}{f}$ $-\frac{1}{m} = u/f - 1$ $= (u-f)/f$ Thus $m = \frac{f}{f-u}$	(1) From the mirror equation, $\frac{1}{v} + \frac{1}{u} = \frac{1}{f}$ (2) Multiplying throughout by v, we get $1 + \frac{v}{u} = v/f$ (3) Magnification, $m = -\frac{v}{u}$ $1 - m = v/f$ $-m = v/f - 1 = \frac{v-f}{f}$ Thus $m = \frac{(f-v)}{f}$

* **Refraction at a Spherical Surface :-**

Spherical Surface :-



Spherical refracting Surface.

Laws of refraction stated for refraction at a plane surface are applicable for refraction at a spherical surface as well. This is because, a small portion of a spherical surface around the point of incidence can be treated as a plane surface. A spherical refracting surface separates two optical media. A spherical surface can be imagined to be a part of a sphere. The definitions given below help in understanding further discussions.

The area of a spherical surface available for refraction is called its aperture. The geometric centre of a spherical refracting surface is called its pole (P). The centre of the sphere of which the refracting surface is a part is called the centre of curvature (C). The radius of the sphere of which the refracting surface is a part is called its radius of curvature (R). The straight line passing through the pole of a spherical refracting surface and its centre of curvature is called its principal axis.

In the discussion that follows, we consider

* a point source on the principal axis.

* Paraxial rays (rays close to principal axis) and

* Small aperture for the refracting surface.

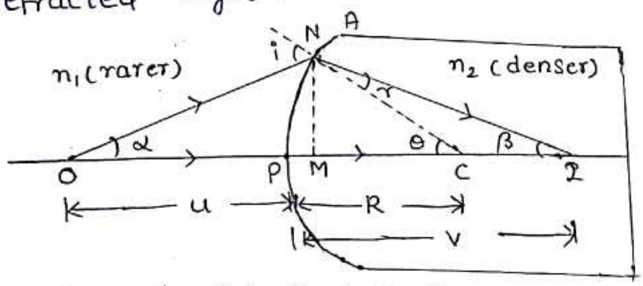
These assumptions enable the angles of incidence and refraction to be small, necessary for simplicity of mathematical calculations.

* **Formula for refraction at a Spherical Surface :-**

Refraction formula for a Spherical refracting Surface is the relation connecting object distance, image distance, radius curvature, refractive indices of the media on the two sides of the refracting surface with the specific sign convention stated earlier the refraction formula retains the same from namely $-\frac{n_1}{u} + \frac{n_2}{v} = \frac{n_2 - n_1}{R}$ in all cases we discuss.
(n_1 = refractive index of the medium containing the incident rays and n_2 = refractive index of the medium containing the refracted rays).

Case - 1 :- Object is in rarer medium and the surface is concave towards denser medium - Real image :-

Consider a spherical refracting surface AB separating two media of refractive indices n_1 and n_2. Let $n_2 > n_1$. Hence here n_1 = refractive index of the medium containing the incident rays and n_2 is the refractive index of the medium containing the refracted rays).



$n_2 > n_1$	
u	-ve
v	+ve
R	+ve

Refraction at Spherical surface - Formation of real image - object in rarer medium and surface is concave towards denser medium.

O is a point object kept in a rarer medium on the principal axis. P is the pole and C is the centre of curvature. The spherical surface is concave towards denser medium. n_1 is the refractive index of rarer medium where the incident rays are present and n_2 is the refractive index of denser medium where refracted rays are present. The aperture AB is assumed to be small. A paraxial ray ON incident on the curved surface undergoes refraction, deviating towards the normal CN as it travels into the denser medium. NI is the refracted ray. A ray OP incident along the principal axis proceeds undeviated along PI into the denser medium as it is incident normal to the surface. I is the real image of O formed in the PI into the denser medium as it is incident normal to the surface. I is the real image of O formed in the denser medium of refractive index n_2.

Fig shows the ray diagram and the angles formed.

i = angle of incidence, r = angle of refraction,
$PO = u$ = object distance, $PI = v$ = image distance,
$PC = R$ = radius of curvature.

ΔONM, $\tan\alpha = \frac{NM}{MO} \approx \frac{NM}{PO}$ [for small aperture $MO \approx PO$]

But for small values of α, we can write

$$\tan\alpha = \alpha \qquad \therefore \alpha = \frac{NM}{PO} \longrightarrow (1)$$

ΔINM, $\tan\beta = \frac{NM}{MI} \approx \frac{NM}{PI}$ $(MI \approx PI)$

But for small values of β, $\beta = \frac{NM}{PI} \longrightarrow (2)$

and ΔCNM, $\tan\theta = \frac{NM}{MC} \approx \frac{NM}{PC}$ $(MC \approx PC)$

But for small values of θ, $\theta = \frac{NM}{PC} \longrightarrow (3)$

The exterior angle in a triangle is equal to the sum of the interior opposite angles.

From ΔNOC, we have $i = \alpha + \theta \longrightarrow (4)$

from ΔNCI, we have $\theta = r + \beta$

$$r = \theta - \beta \longrightarrow (5)$$

From Snell's law, $\frac{n_2}{n_1} = \frac{\sin i}{\sin r} \approx \frac{i}{r}$

$$\frac{n_2}{n_1} = \frac{\alpha+\theta}{\theta-\beta}$$

That is $\quad n_1(\alpha+\theta) = n_2(\theta-\beta)$

$$(n_2-n_1)\theta = n_1\alpha + n_2\beta$$

Substituting for α, β and θ from eqs (1), (2) and (3)

we get $\quad (n_2-n_1)\frac{NM}{PC} = n_1\frac{NM}{PO} + n_2\frac{NM}{PI}$

$$\Rightarrow \frac{n_1}{PO} + \frac{n_2}{PI} = \frac{n_2-n_1}{PC}$$

According to sign convention, $u = -PO$, $v = +PI$ & $R = +PC$

Hence, the refraction formula takes the form

$$-\frac{n_1}{u} + \frac{n_2}{v} = \frac{n_2-n_1}{R} \Rightarrow \boxed{\frac{n_2}{v} - \frac{n_1}{u} = \frac{n_2-n_1}{R}}$$

* **Refraction through a thin lens :-**

A lens is an optical medium bound by two surfaces of which at least one is either spherical or cylindrical. A parallel beam of light can be made to converge or diverge by passing through a lens. Accordingly a lens may be a converging lens or diverging lens. The most commonly employed types of lens are shown in the Fig. below.

Biconvex Biconcave Plano Convex Plano Concave Miniscus lens

Fig :- Commonly used lens.

Let us familiarize a few terminology required to understand the functioning of lens.

⇒ The surface area of a lens surface available for refraction is called its aperture.

⇒ A straight line passing through centres of curvature of two surfaces of a lens is called its principal axis.

when a pencil of rays, close and parallel to the principal axis is incident on a lens, after refraction, the rays converge to a fixed point in case of a convex lens and appear to diverge from a fixed point in case of a concave lens. This fixed point on the principal axis is called principal focus of the lens. The distance between its optic centre of a lens and its principal focus is called its focal length.

⇒ A lens has a pair of principal foci, one on either side of the lens. The two principal foci are equidistant from the lens.

⇒ A lens is considered to be thin when its thickness is negligable compared to the focal length of the lens and the radii of curvature of its two surfaces.

Optic Centre :-

when a ray of light passes through a thin lens, such that the emergent ray is parallel to the incident ray, the refracted ray intersects the principal axis at a unique point called optic centre. A ray passing through the optic centre of a thin lens is undeviated. Further, in case of a thin lens, the lateral shift of the ray is negligable.

* **Ray tracing in a lens :-**

The image formed by a lens can be found by choosing any of the two rays mentioned below. The laws of refraction need to be applied in ray tracing.

(1) A ray from an object parallel to the principal axis after refraction passes through the second principal focus (in a convex lens) or appears to

diverge from the first principal focus (in a concave lens).

(2) A ray of light through the optic centre emerges without any deviation after refraction.

(3) A ray of light passing through the first principal focus (in a convex lens) or appearing to meet at it (in a concave lens) emerges parallel to the principal axis after refraction.

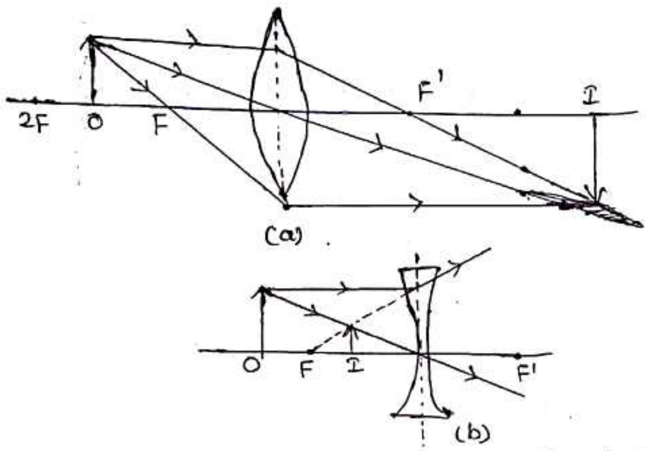


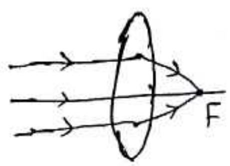
Fig: Illustration of ray tracing for locating an image formed by a lens.

* Different cases of images formed by a lens :-

(1) Position of the object :- At infinity.

Position & Nature of Image

Ray diagram :-

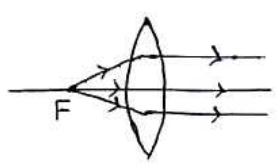


(1) At F
(2) Real
(3) Inverted
(4) diminished

(2) Position of the object :- At F.

Position & Nature of Image:-

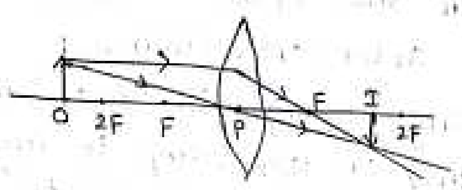
Ray diagram :-



(1) At Infinity.
(2) Real
(3) Inverted
(4) highly enlarged.

(3) Position of the object :- Beyond 2F

Ray diagram :-

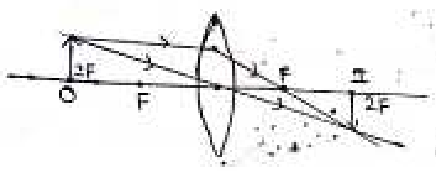


Position and nature of Image :-

(1) Between F and 2F
(2) Real
(3) Inverted
(4) diminished .

(4) Position of the object :- At 2F

Ray diagram :-

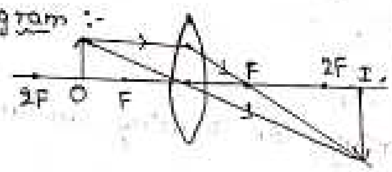


Position and nature of Image :-

(1) At 2F
(2) Real
(3) Inverted
(4) Same size .

(5) Position of the object :- Between F and 2F .

Ray diagram :-

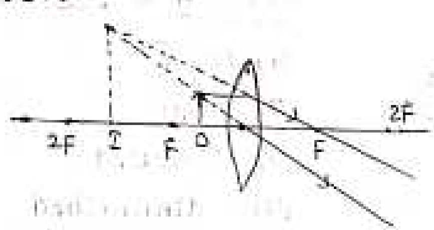


Position & nature of Image :-

(1) Beyond 2F
(2) Real
(3) Inverted
(4) enlarged .

(6) Position of the object :- within F .

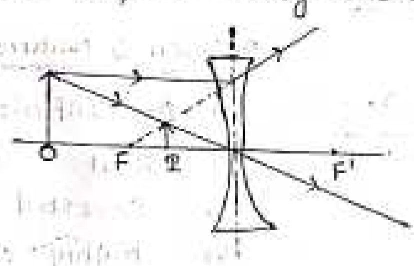
Ray diagram :-



Position & nature of Image :-

(1) Beyond object
(2) virtual
(3) errect
(4) enlarged .

(7) Concave lens : Any where



Position and nature of Image :-

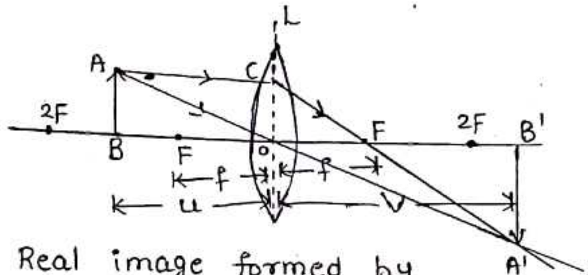
(1) within F
(2) Virtual
(3) Errect
(4) diminished .

* **Lens equation :-**

Lens equation or lens formula is a relation between object distance, image distance and focal length of a lens. The lens equation given by

$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$$

(a) **Convex lens producing real image :-**



Real image formed by convex lens

u	-ve
v	+ve
f	+ve

AB is an object held perpendicular to the Principal axis and placed between F and 2F. Its real inverted, enlarged image A'B' is formed beyond 2F.

Δ ABO and ΔA'B'O are Similar.

$$\frac{A'B'}{AB} = \frac{OB'}{OB} \longrightarrow (1)$$

Δ COF and Δ A'B'F are Similar

$$\frac{A'B'}{CO} = \frac{B'F}{OF} \longrightarrow (2)$$

From Eqs (1) and (2) we get.

$$\frac{OB'}{OB} = \frac{B'F}{OF} \longrightarrow (3)$$

Since $CO = AB$ and $B'F = OB' - OF$

$$\frac{OB'}{OB} = \frac{OB' - OF}{OF}$$

$$(OB')(OF) = (OB)(OB') - (OB)(OF)$$

$$(OB')(OF) + (OB)(OF) = (OB)(OB')$$

$$\frac{1}{OB} + \frac{1}{OB'} = \frac{1}{OF} \longrightarrow (4)$$

According to sign convention,

Object distance $u = -OB$; Image distance, $v = +OB'$; Focal length, $f = +OF$. Therefore, $OB = -u$; $OB' = v$, $OF = f$

Substituting these in Eq (4) we get.

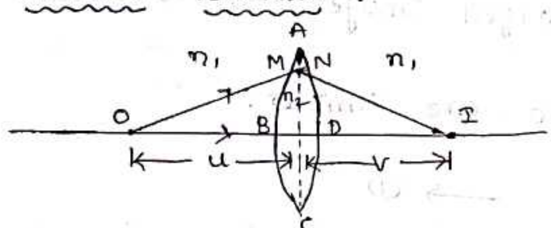
$$\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$$

* Lateral magnification :-

The ratio of the height of the image formed by a lens to the height of the object is called lateral magnification produced by a lens. It is represented by m.

$$\text{Magnification, } m = \frac{\text{height of image}}{\text{height of object}} = \frac{h_i}{h_o} = \frac{v}{u}$$

* Len's maker's formula :-



Formation of real image of a point objected by a thin lens.

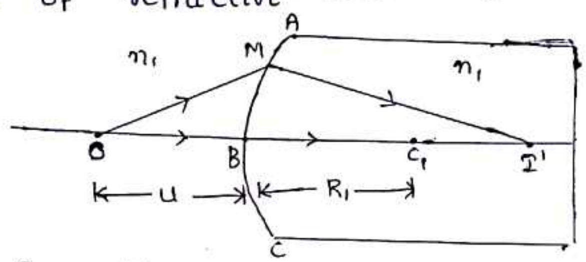
Consider a thin lens of small aperture placed in a medium of refractive index n_1. Let n_2 be the refractive index of the material of the lens. Let f be the focal length of the lens and R_1 and R_2 be the radii of curvature of its two surfaces.

Let O be a point object kept on the principal axis at a distance u from the lens. A ray along the principal axis passes through the lens undeviated. Another ray OM incident on the lens at M undergoes refraction and proceeds along MN. The ray emerges from the lens and intersects the principal axis at I. Therefore, I is the real image of object O. I is formed at a distance v from the lens.

The formation of image I can be considered to occur in two steps.

(1) Refraction at surface ABC only :-

In the absence of the second surface ADC, the refracted ray at M proceeds along MI'. Thus, I' is the real image of point object O formed at a distance v' from the lens in a medium of refractive index n_2.



Formation of Image I' by surface ABC of thin lens.

The general formula for refraction at a spherical refracting surface is $\frac{n_2}{v} - \frac{n_1}{u} = \frac{n_2 - n_1}{R}$ $\longrightarrow$ (1)

Here for refraction at the first surface, let us assume that we have the material of refractive index n_2 to be present to the entire right of ABC. Then an object at O will have its image at I'.

Here, Object distance, $u = -BO$; Image distance, $v = +BI'$ and radius $R_1 = +BC_1$

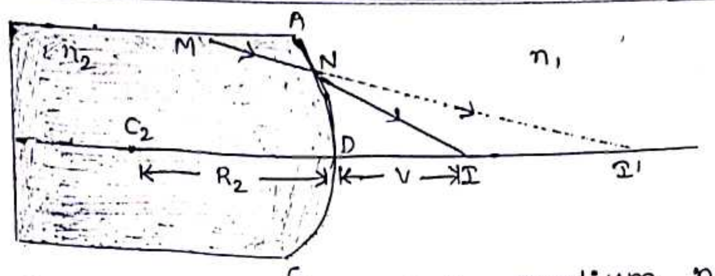
Refractive index of medium in which incident rays are present (in the formula) $n_1 = n_1$

Refractive index of medium in which refracted rays are present (in the formula) $n_2 = n_2$

Substituting these in Eq (1) we get

$$\frac{n_2}{BI'} - \frac{n_1}{(-BO)} = \frac{n_2 - n_1}{BC_1} \longrightarrow (2)$$

(2) Refraction at surface ADC only :- For refraction at second surface ADC, let us assume that to the left of ADC we have only the material of lens of refractive index n_2.



Formation of real image I in medium n_1 by Second surface ADC of a thin lens.

If ADC were to be absent, the refracted ray MN would have to proceed to I'. Thus, I' is the object for refraction at ADC.

Here object distance $u = DI'$; Image distance $v = DI$ and radius $R_2 = -DC_2$

Refractive index of medium in which incident rays are present (in the formula), $n_1 = n_2$

Refractive index of medium in which refracted rays are present (in the formula), $n_2 = n_1$

Substituting these in Eq (1) we get

$$\frac{n_1}{DI} - \frac{n_2}{DI'} = \frac{n_2 - n_1}{-DC_2} \longrightarrow (3)$$

Note that $BI' \approx DI'$ since the lens is thin.

Now adding Eqs (2) and (3) we get

$$\frac{n_1}{BO} + \frac{n_1}{DI} = (n_2 - n_1)\left(\frac{1}{BC_1} + \frac{1}{DC_2}\right) \longrightarrow (4)$$

But $BO = -u$, $DI = v$, $BC_1 = R_1$ and $DC_2 = -R_2$

Substituting these in Eq (4) we get

$$n_1\left(\frac{1}{v} - \frac{1}{u}\right) = (n_2 - n_1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

$$\frac{1}{v} - \frac{1}{u} = \left(\frac{n_2}{n_1} - 1\right)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \longrightarrow (5)$$

But $\frac{1}{v} - \frac{1}{u} = \frac{1}{f}$ and $\frac{n_2}{n_1} = {}_1n_2$

Hence Eq (5) becomes

$$\frac{1}{f} = ({}_1n_2 - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)$$

This equation is known as Lens Maker's formula.

This equation is valid for both converging and diverging lens.

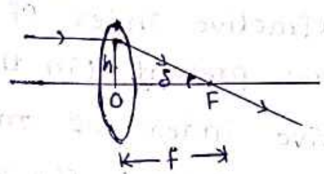
If the lens of a refractive index $n_2 = n$ is present in air, the above equation becomes.

$$\boxed{\frac{1}{f} = (n-1)\left[\frac{1}{R_1} - \frac{1}{R_2}\right]}$$

Note that R_1 is the radius of curvature of the surface on to which the incident rays fall first.

* **Power of a Lens :-**

The ability of a lens to converge or diverge the rays of light incident on it is called the power of the lens.



Power of a lens is defined as the tangent of the angle by which it converges or diverges a beam of light falling at unit distance from the optical centre.

According to the figure,

$\tan\delta = \frac{h}{f}$, if $h = 1$, $\tan\delta = \frac{1}{f}$

For small values of δ, $\tan\delta \approx \delta$ $\Rightarrow$ $\delta = \frac{1}{f}$

$\therefore$ Thus, power of a lens, $P = \frac{1}{f}$

→ SI unit of power of lens is dioptre (D). The power of a lens is measured as the reciprocal of its focal length (in metre) $P = \frac{1}{f(m)}$

If $f = 1m$, then $P = 1m^{-1} = 1$ dioptre (D)

According to the lens Maker's formula for a lens,

$$\frac{1}{f} = (\mu - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right) \qquad \left(\because P = \frac{1}{f}\right)$$

we have $\boxed{P = (\mu-1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right)}$

Here R_1 and R_2 are to be measured in metre.

6. Wave Optics

* Huygen's Principle :-

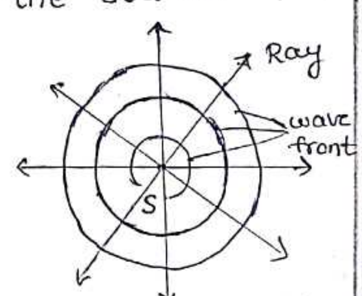
The speed of light in a medium depends upon the nature of the medium. Huygens supposed the existence of a hypothetical medium called "luminiferous ether" which filled the entire space. This medium was supposed to be massless with extremely high elasticity and very low density.

Wavefront :-

It is the locus of points (wavelets) having the same phase (a surface of constant phase) of oscillations. A wavelet is the point of dis-turbance due to propagation of light. A line perpendicular to a wave front is called a ray.

Depending on the shape of source of light, wavefronts can be of three types, which are given below.

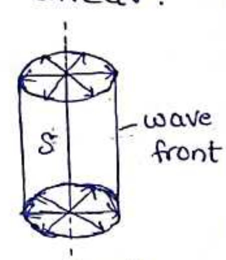
(i) Spherical wavefront :- when the source of light is a point source, the wavefront is a sphere with centre as the source.



Spherical wavefront

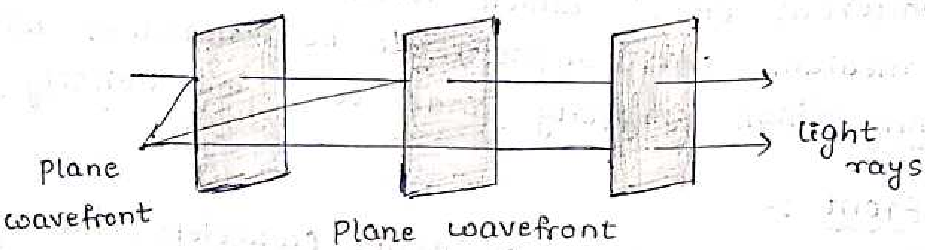
(ii) Cylindrical wavefront :-

when the source of light is linear. e.g. a straight line source, slit, etc.... as shown in the figure. All the points equidistant from the source lie on a cylinder. Therefore, the wavefront is cylindrical in shape.



cylindrical wavefront

(iii) Plane wavefront :- when the point source or linear source of light is at very large distance, a small portion of spherical or cylindrical wave-front appears to be plane. Such a wavefront is called a plane wavefront.



Plane wavefront

Hence, the wavefront is a surface of constant phase. The speed with which the wave front moves outwards from the source is called the speed of the wave. The energy of the wave travels in a direction perpendicular to the wave front.

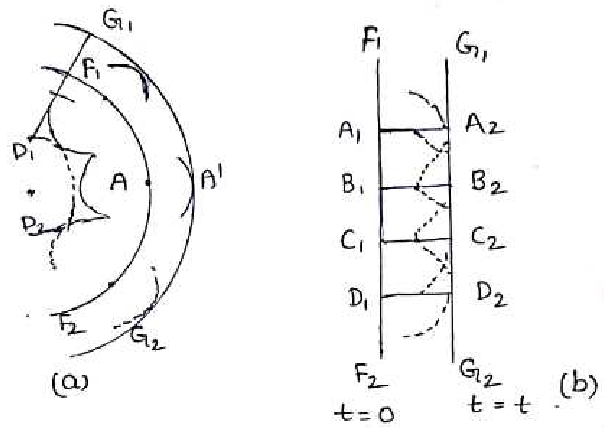
Plane wavefront

* Huygen's Principle :-

It is essentially a geometrical construction, which gives the shape of the wavefront at any time and allows us to determine the shape of the wavefront at a later time. According to Huygen's principle,

(i) Each point on the given wavefront (called primary wavefront) is the source of a secondary disturbance (called secondary wavelets) and the wavelets emanating from these points spread out with the speed of the wave in all directions.

(ii) A surface touching these secondary wavelets, tangentially in the forward direction at any instant gives the new wavefront at that instant. This is called secondary wavefront.



(a) (b)

In fig (a) :- F_1F_2 is the section of the given spherical wavefront and G_1G_2 is the new wavefront in the forward direction.

In fig (b) :- F_1F_2 is the section of the given plane wavefront and G_1G_2 is the new wavefront in the forward direction.

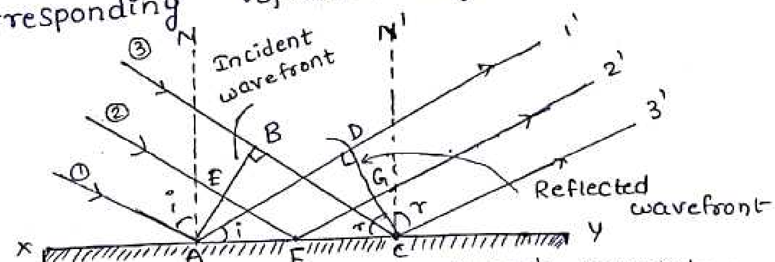
Note :- Huygen's argued that the amplitude of the secondary wavelets is maximum in the forward direction and zero in the backward direction. Hence the backward secondary wavefront is absent.

* Refraction and Reflection of plane waves using Huygen's principle :-

Huygen's principle can be used to explain the phenomena of reflection and refraction of light on the basis of wave theory of light.

Laws of Reflection at plane Surface :-

Let 1, 2, 3 be the incident rays and 1', 2', 3' be the corresponding reflected rays.



Laws of reflection by Hugen's principle.

If 'c' is the speed of the light, t is the time taken by light to go from B to C or A to D or E to G through F, then $t = \frac{EF}{c} + \frac{FG}{c}$ ---- (i)

In $\triangle AEF$, $\sin i = \frac{EF}{AF}$

In $\triangle FGC$, $\sin r = \frac{FG}{FC}$

or $t = \frac{AF \sin i}{c} + \frac{FC \sin r}{c}$

$\Rightarrow t = \frac{AC \sin r + AF(\sin i - \sin r)}{c}$ [$\because FC = AC - AF$]

For rays of light from different parts on the incident wavefront, the values of AF are different. But light from different points of the incident wavefront should take the same time to reach the corresponding points on the reflected wavefront.

So, t should not depend upon AF. This is possible only, if $\sin i - \sin r = 0 \Rightarrow \sin i = \sin r$

$\angle i = \angle r \longrightarrow$ (ii)

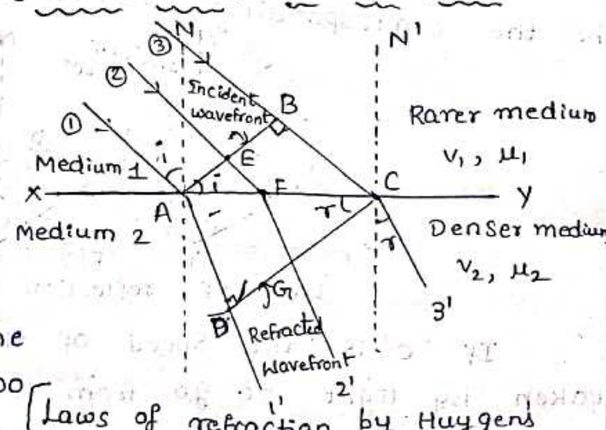
which is the first law of reflection.

Further, the incident wavefront AB, the reflecting surface xy and the reflected wavefront CD are all perpendicular to the plane of the paper. Therefore, incident ray, normal to the mirror xy and reflected ray all lie in the plane of the paper. This proves the second law of reflection.

* Laws of Refraction (Snell's Law) at a Plane Surface :-

Let 1, 2, 3 be the incident rays and $1', 2', 3'$ be the corresponding refracted rays.

If v_1, v_2 are the speed of light in the two media.



[Laws of refraction by Huygens Principle] and t is the time taken by light to go from B to C or A to D or E to G through F, then

$$t = \frac{EF}{v_1} + \frac{FG}{v_2}$$

In Δ AFE, $\sin i = \frac{EF}{AF}$

In Δ FGC, $\sin r = \frac{FG}{FC}$

$$\Rightarrow t = \frac{AF \sin i}{v_1} + \frac{FC \sin r}{v_2} \longrightarrow \text{(iii)}$$

$$\Rightarrow t = \frac{AC \sin r}{v_2} + AF\left(\frac{\sin i}{v_1} - \frac{\sin r}{v_2}\right) \quad [FC = AC - AF)$$

For rays of light from different parts on the incident wavefront, the values of AF are different. But light from different points of the incident wavefront should take the same time to reach the corresponding points on the refracted wavefront. So t should not depend upon AF. This is possible only, if

$$\frac{\sin i}{v_1} - \frac{\sin r}{v_2} = 0$$

$$\frac{\sin i}{\sin r} = \frac{v_1}{v_2} \dashrightarrow \text{(iv)}$$

Now, if c represents the speed of light in vaccum, then $\mu_1 = \frac{c}{v_1}$ and $\mu_2 = \frac{c}{v_2}$ are known as the refractive indices of medium 1 and medium 2 respectively.

In terms of refractive indices, Eq (iv) can be written as $\mu_1 \sin i = \mu_2 \sin r$

$$\Rightarrow \boxed{\mu = \frac{\sin i}{\sin r}}$$

This is known as Snell's law of refraction.

Further, if λ_1 and λ_2 denote the wavelengths of light in medium 1 and medium 2, respectively and if the distance BC is equal to λ_1, then the distance AD will be equal to λ_2.

thus $\frac{\lambda_1}{\lambda_2} = \frac{BC}{AD} = \frac{v_1}{v_2}$ $[v \rightarrow \text{frequency}]$

$$\frac{v_1}{\lambda_1} = \frac{v_2}{\lambda_2} \Rightarrow \nu_1 = \nu_2 \quad \left[\frac{v}{\lambda} = \nu\right]$$

Hence, the frequency does not change on refraction. Thus, frequency (ν) being a characteristic of the source, remains the same as light travels from one medium to another.

Also, wavelength is directly proportional to the (phase) speed and inversly proportional to refractive index.

$$\therefore \lambda' = \frac{\lambda}{\mu}, \quad \mu = \frac{\lambda}{\lambda'} = \frac{c\lambda}{v\lambda} = \frac{c}{v}$$

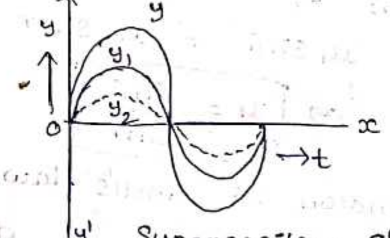
* Interference of Light :-

Superposition principle :-

According to the principle, at a particular point in the medium, the resultant displacement (y) produced by a number of waves is the vector sum of the displacements produced by each of the waves $(y_1 + y_2 + \cdots)$.

i.e $y = y_1 + y_2 + y_3 + y_4 + \cdots\cdots$

Clearly, each wave contributes as if the other wave is not present. The superposition principle which was stated first for mechanical waves is equally applicable to the electromagnetic (light) waves.



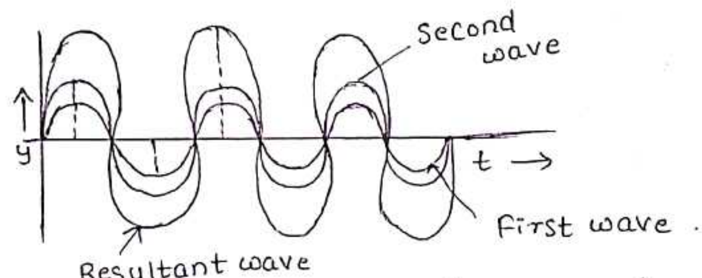
Superposition of waves.

Interference of Light waves :-

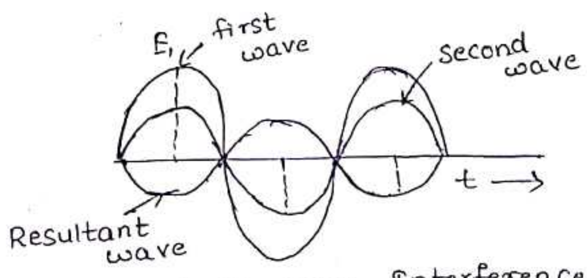
When two light waves of exactly equal frequency having constant phase difference w.r.t. time travelled on same direction and superimpose (overlap) with each other, then intensity of resultant wave does not remain uniform in space.

This phenomenon of formation of maximum intensity at some points and minimum intensity at some other points by two identical light waves travelling in same direction is called the interference of light.

At the points, where the resultant intensity of light is maximum, interference is said to be constructive. At the points, where the resultant intensity of light is minimum, the interference is said to be destructive.



(a) Constructive Interference.



(b) Destructive Interference.

* **Theory of Interference of Waves :-**

Let the waves from two sources of light be represented as $y_1 = a \sin \omega t$ and $y_2 = b \sin(\omega t + \phi)$

where a and b are the respective amplitudes of the two waves and ϕ is the constant phase angle by which second wave leads the first wave.

Applying Superposition principle, the magnitude of the resultant displacement of the waves is .

$$y = y_1 + y_2$$

$$\Rightarrow y = a\sin\omega t + b\sin(\omega t + \phi)$$

$$\Rightarrow y = a\sin\omega t + b\sin\omega t \cdot \cos\phi + b\cos\omega t \cdot \sin\phi$$

$$\Rightarrow y = (a + b\cos\phi)\sin\omega t + b\sin\phi \cos\omega t$$

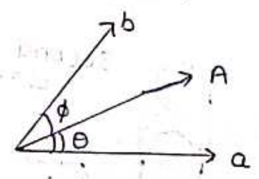
putting $a + b\cos\phi = A\cos\theta$ and $b\sin\phi = A\sin\theta$

we get $y = A\cos\theta \cdot \sin\omega t + A\sin\theta \cdot \cos\omega t$

or $y = A\sin(\omega t + \theta)$

where, A is the resultant amplitude and θ is the resultant phase difference.

$$\therefore A = \sqrt{a^2 + b^2 + 2ab\cos\phi} \quad \text{and} \quad \tan\theta = \frac{b\sin\phi}{a + b\cos\phi}$$



Resultant of amplitudes a and b .

As, intensity is directly proportional to the square of the amplitude of the wave, i.e $I \propto a^2$

So, for two different cases,

$$I_1 = Ka^2, \quad I_2 = Kb^2$$

$$I_R = KA^2 = K(a^2 + b^2 + 2ab\cos\phi)$$

$$\boxed{I_R = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\phi}$$

For Constructive Interference :-

I Should be maximum, for which

$$\cos\phi = \text{maximum} = +1$$

$\therefore$ Phase difference, $\phi = 0, 2\pi, 4\pi, \ldots\ldots$

i.e $\phi = 2n\pi$, where $n = 1, 2, 3, \ldots\ldots$

If Δx be the path difference between the interfering waves, then $\Delta x = \frac{\lambda}{2\pi}(\phi) = \frac{\lambda}{2\pi}(2n\pi)$

$$\Rightarrow \Delta x = n\lambda$$

and $\boxed{I_{max} \propto (a+b)^2}$

Scanned by CamScanner

Applying Superposition principle, the magnitude of the resultant displacement of the waves is

$$y = y_1 + y_2$$

$$\Rightarrow y = a\sin\omega t + b\sin(\omega t + \phi)$$

$$\Rightarrow y = a\sin\omega t + b\sin\omega t \cdot \cos\phi + b\cos\omega t \cdot \sin\phi$$

$$\Rightarrow y = (a + b\cos\phi)\sin\omega t + b\sin\phi\cos\omega t$$

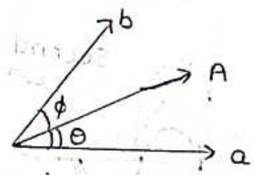
putting $a + b\cos\phi = A\cos\theta$ and $b\sin\phi = A\sin\theta$

we get $y = A\cos\theta \cdot \sin\omega t + A\sin\theta \cdot \cos\omega t$

or $y = A\sin(\omega t + \theta)$

where, A is the resultant amplitude and θ is the resultant phase difference.

$$\therefore A = \sqrt{a^2 + b^2 + 2ab\cos\phi} \quad \text{and} \quad \tan\theta = \frac{b\sin\phi}{a + b\cos\phi}$$



Resultant of amplitudes a and b.

As, intensity is directly proportional to the square of the amplitude of the wave, i.e $I \propto a^2$

So, for two different cases,

$$I_1 = Ka^2, \quad I_2 = Kb^2$$

$$I_R = KA^2 = K(a^2 + b^2 + 2ab\cos\phi)$$

$$\boxed{I_R = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\phi}$$

For Constructive Interference :-

I should be maximum, for which

$$\cos\phi = \text{maximum} = +1$$

$\therefore$ Phase difference, $\phi = 0, 2\pi, 4\pi, \ldots$

$\phi = 2n\pi$, where $n = 1, 2, 3, \ldots$

i.e

If Δx be the path difference between the interfering waves, then $\Delta x = \frac{\lambda}{2\pi}(\phi) = \frac{\lambda}{2\pi}(2n\pi)$

$$\Rightarrow \Delta x = n\lambda$$

and $\boxed{I_{max} \propto (a+b)^2}$

Hence, condition for constructive interference at a point is that, the phase difference between the two waves reaching the point should be zero or an even integral multiple of 2π. Equivalent path difference between the two waves reaching the point should be zero or an integral multiple of full wavelength.

For destructive interference :-

I should be minimum, for which

$$\cos\phi = \text{minimum} = -1.$$

$\therefore$ phase difference, $\phi = \pi, 3\pi, 5\pi, \ldots$

i.e $\phi = (2n-1)\pi$

where $n = 1, 2, \ldots$

The corresponding path difference between the two waves is $\Delta x = \left(\frac{\lambda}{2\pi}\right)\phi = \left(\frac{\lambda}{2\pi}\right)(2n-1)\pi$

$$\Rightarrow \Delta x = (2n-1)\lambda/2$$

and $\boxed{I_{min} \propto (a-b)^2}$

Hence, condition for destructive interference at a point is that, the phase difference between the two waves reaching the point should be an odd integral multiple of π or path difference between the two waves reaching the point should be an odd integral multiple of half wavelength.

Comparison of Intensities of Maxima and Minima :-

As, $I_{max} \propto (a+b)^2$ and $I_{min} \propto (a-b)^2$

$$\therefore \frac{I_{max}}{I_{min}} = \frac{(a+b)^2}{(a-b)^2} = \frac{\left(\frac{a}{b}+1\right)^2}{\left(\frac{a}{b}-1\right)^2}$$

$$\Rightarrow \boxed{\frac{I_{max}}{I_{min}} = \frac{(r+1)^2}{(r-1)^2}}$$

where $r = \frac{a}{b}$ (ratio of amplitudes).

* Coherent and Incoherent Sources :-

Light sources are of two types, i.e. coherent and non-coherent light sources. The sources of light which emit light waves of same wavelength, same frequency and are in same phase or having constant phase difference are known as coherent sources.

Two such sources of light, which do not emit light waves with constant phase difference are called incoherent sources.

Need of Coherent Sources for the production of Interference pattern :-

As discussed earlier, when two monochromatic waves of intensity I_1, I_2 and phase difference ϕ meet at a point, then the resultant intensity is given by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos\phi$$

Here, the term $2\sqrt{I_1 I_2}\cos\phi$ is called interference term. There are two possibilities.

(i) If $\cos\phi$ remains constant with time, then the total intensity at any point will be constant. The intensity will be maximum $(\sqrt{I_1}+\sqrt{I_2})^2$ at points, where $\cos\phi$ is 1 and minimum $(\sqrt{I_1}-\sqrt{I_2})^2$ at points, where $\cos\phi$ is -1. The sources in this case are coherent.

(ii) If $\cos\phi$ varies continuously with time assuming both positive and negative value, then the average value of $\cos\phi$ will be zero over time interval of measurement. The interference term averages to zero. There will be same intensity $I = I_1 + I_2$ at every point. The two sources in this case are incoherent.

Note :- In practice, Coherent Sources are produced either by dividing the wavefront or by dividing the amplitude of the incoming waves.

* **Requirements for obtaining Two Coherent Sources of Light :-**

Following are the requirements (conditions) for obtaining two Coherent Sources of light.

(i) Coherent Sources of light Should be obtained from a Single Source by Some device. Two Coherent Sources can be obtained either by

(a) the Source and its virtual image (Lloyd's mirror).

(b) the two virtual images of the Same Source (Fresnel's biprism)

(c) two real images of the Same Source (Young's double Slit).

(ii) The two Sources Should give monochromatic light.

(iii) The path difference between light waves from two Sources Should be Small.

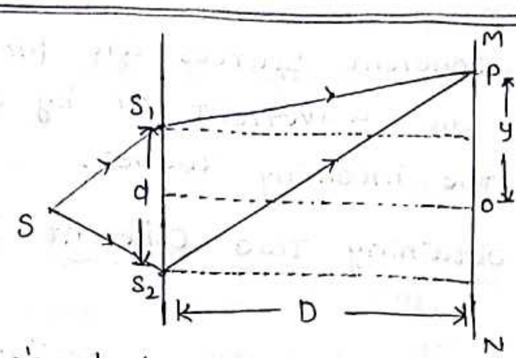
(iv) Coherent Sources can be produced by two methods :

(a) By division of wavefront (Young's double Slit experiment).

(b) By division of amplitude (partial reflection or refraction in thin films).

* **Young's Double Slit Experiment :-**

Suppose S_1 and S_2 are two fine slits, a Small distance d apart. They are illuminated by a Strong Source S of monochromatic light of wavelength λ. MN is a Screen at a distance D from the Slits.



Young's double slit arrangement to produce interference pattern.

Consider a point P at a distance y from O; the Centre of Screen.

The Path difference between two waves arriving at point P is equal to $S_2P - S_1P$.

Now, $(S_2P)^2 - (S_1P)^2 = \left[D^2 + \left(y+\frac{d}{2}\right)^2\right] - \left[D^2 + \left(y-\frac{d}{2}\right)^2\right]$

$= 4y\left(d/2\right) = 2yd.$

Thus, $S_2P - S_1P = \dfrac{2yd}{S_2P + S_1P}$

But $S_2P + S_1P = 2D$

$\therefore S_2P - S_1P \approx \dfrac{dy}{D}$

For Constructive interference (Bright fringes) :-

Path difference $= \dfrac{yd}{D} = n\lambda$ where $n = 0, 1, 2, \dots$

$\therefore y = \dfrac{nD\lambda}{d}$

Hence for $n=0$, $y_0 = 0$ at O Central bright fringe.

for $n=1$, $y_1 = \dfrac{D\lambda}{d}$ for 1st bright fringe.

for $n=2$, $y_2 = \dfrac{2D\lambda}{d}$ for 2nd bright fringe

for $n=n$, $y_n = \dfrac{nD\lambda}{d}$ for nth bright fringe.

The Separation between two Consecutive bright fringes is $\beta = \dfrac{nD\lambda}{d} - \dfrac{(n-1)D\lambda}{d} = \dfrac{D\lambda}{d}$.

$\therefore \boxed{\beta = \dfrac{\lambda D}{d}}$

For destructive interference (Dark fringes) :-

Path difference $= \frac{yd}{D} = (2n-1)\lambda/2$

or $y = (2n-1)\frac{D\lambda}{2d}$, where $n = 1, 2, 3, \ldots$

Hence, for $n=1$, $y_1' = \frac{D\lambda}{2d}$ for 1st dark fringe.

for $n=2$, $y_2' = \frac{3D\lambda}{2d}$ for 2nd dark fringe.

for $n=n$, $y_n' = (2n-1)\frac{D\lambda}{2d}$ for nth dark fringe.

The separation between two consecutive dark fringes is $\beta' = (2n-1)\frac{D\lambda}{2d} - [2(n-1)-1]\frac{D\lambda}{2d}$

$$\therefore \boxed{\beta' = \frac{\lambda D}{d}}$$

Fringe width :-

The distance between two consecutive bright or dark fringes is called fringe width ω.

$\therefore$ Fringe width, $\boxed{\omega = \frac{\lambda D}{d}}$

The above formula is free from n that means the width of all fringes is same.

Fringe width is directly proportional to λ. Hence, the fringes of red light (longer wavelength) are broader than the fringes of blue light (shorter wavelength).

Intensity of the fringes :-

$\Rightarrow$ For a bright fringe, $\phi = 2n\pi$

and $\cos\phi = \cos 2n\pi = 1$

So, $I_R = I_{max} = I_1 + I_2 + 2\sqrt{I_1 I_2} = 4I$

[as $I_1 = I_2 = I$ in YDSE]

$\therefore$ Intensity of a bright fringe $= 4I$

$\Rightarrow$ For a dark fringe, $\phi = (2n-1)\pi$

$\cos\phi = -1$

$I_R = I_{min} = I_1 + I_2 - 2\sqrt{I_1 I_2} = 0$

$\therefore$ Intensity of a dark fringe $= 0$.

* **Conditions for Sustained Interference :-**

In order to obtain a well-defined observable interference pattern, the intensity at points of constructive and destructive interference must be maintained maximum and almost zero, respectively.

For this, following conditions must be satisfied

(i) The two sources producing interference must be coherent.

(ii) The two interfering waves must have the same plane of polarisation.

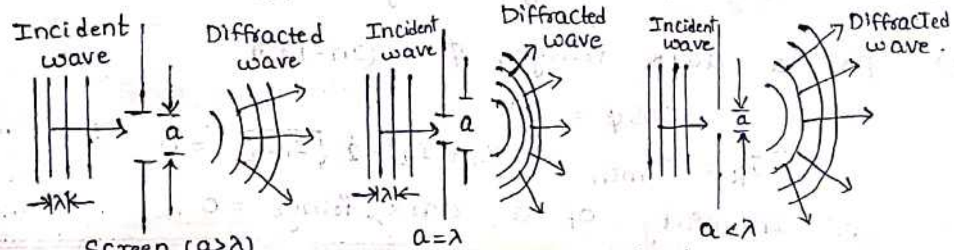
(iii) The two sources must be very close to each other and the pattern must be observed at a larger distance to have sufficient width of the fringe $\left(\frac{D\lambda}{d}\right)$.

(iv) The sources must be monochromatic, otherwise the fringes of different colours will overlap.

(v) The two waves must be having same amplitude for better contrast between bright and dark fringes.

* **Diffraction of Light :-**

The phenomenon of bending of light around the sharp corners and the spreading of light within the geometrical shadow of the opaque obstacles is called diffraction of light. The light thus deviates from its linear path. The deviation becomes much more pronounced, when the dimensions of the aperture or the obstacle are comparable to the wavelength of light.



Diffraction of waves for slits of different width.

Note :- Diffraction is a general characteristic exhibited by all types of waves. For visible light, λ is very small ($\approx 10^{-6}$m). Therefore, diffraction of visible light is not so common as obstacles / apertures of this size are hardly available.

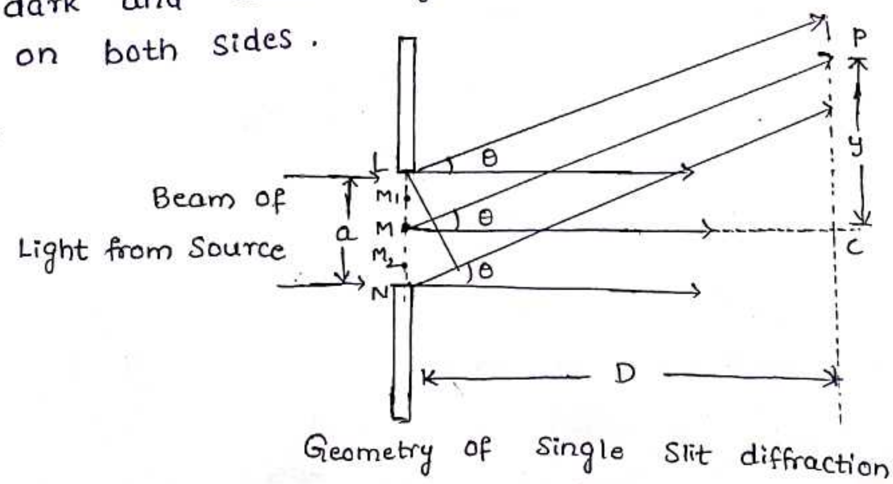
According to Fresnel, diffraction occurs on the account of mutual interference of secondary wavelets starting from portions of the wavefront which are not blocked by the obstacle or from portions of the wavefront which are allowed to pass through the aperture.

* **Diffraction of Light at a Single Slit :-**

A parallel beam of light with a plane wavefront is made to fall on a single slit LN. As width of the slit LN = a is of the order of wavelength of light, therefore diffraction occurs when beam of light passes through the slit.

The wavelets from the single wavefront reach the centre C on the screen in same phase. Hence, interfere constructively to give central maximum (bright fringe).

The diffraction pattern obtained on the screen consists of a central bright band, having alternate dark and weak bright bands of decreasing intensity on both sides.



Geometry of single slit diffraction.

Consider a point P on the screen at which wavelets travelling in a direction, make an angle θ with MC. The wavelets from points L and N will have a path difference equal to NQ.

From the right angled Δ LNQ, we have

$$NQ = LN \sin\theta$$

$$NQ = a \sin\theta$$

To establish the condition for secondary minima, the slit is divided into 2, 4, 6, ... equal parts such that corresponding wavelets from successive regions interfere with path difference of $\lambda/2$.

or for nth secondary minima, the slit can be divided into 2n equal parts.

Hence, for nth secondary minima,

$$\text{Path difference} = \frac{a}{2} \sin\theta = \lambda/2$$

$$\text{or} \quad \boxed{\sin\theta_n = \frac{n\lambda}{a}, \quad (n = 1, 2, 3, \dots)}$$

To establish the condition for secondary maxima, the slit is divided into 3, 5, 7, ... equal parts such that corresponding wavelets from alternate regions interfere with path difference of $\lambda/2$.

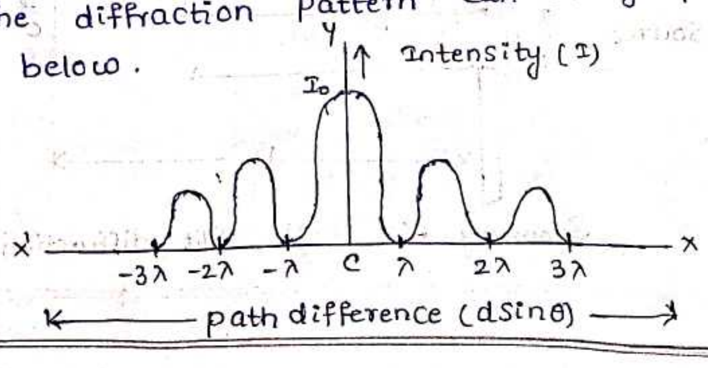
or for nth secondary maxima, the slit can be divided into (2n+1) equal parts.

Hence, for nth secondary maxima,

$$a \sin\theta_n = (2n+1)\frac{\lambda}{2} \quad [n = 1, 2, 3, \dots]$$

$$\text{or} \quad \boxed{\sin\theta_n = (2n+1)\frac{\lambda}{2a}}$$

Hence, the diffraction pattern can be graphically shown as below.



The point C corresponds to the positions of central maxima. And the position $-3\lambda, -2\lambda, -\lambda, \lambda, 2\lambda, 3\lambda$- are secondary minima. The above conditions for diffraction maxima and minima are exactly reverse of mathematical conditions for interference maxima and minima.

Width of Central Maximum :-

It is the distance between first secondary minimum on either side of the central bright fringe C.

For first secondary minimum,

$$a \sin\theta = \lambda \quad \text{or} \quad \sin\theta = \frac{\lambda}{a} \longrightarrow \text{(i)}$$

If θ is small, $\sin\theta \approx \theta = \frac{y}{D} \longrightarrow \text{(ii)}$

From Eqs. (i) and (ii), we get

$$\frac{y}{D} = \frac{\lambda}{a} \quad \text{or} \quad y = \frac{D\lambda}{a}$$

width of central maximum $\boxed{2y = \frac{2D\lambda}{a}}$

As, the slit width a increases, width of central maximum decreases.

$\therefore$ Angular width of central maxima, $\boxed{2\theta = \frac{2\lambda}{a}}$

* Polarisation :-

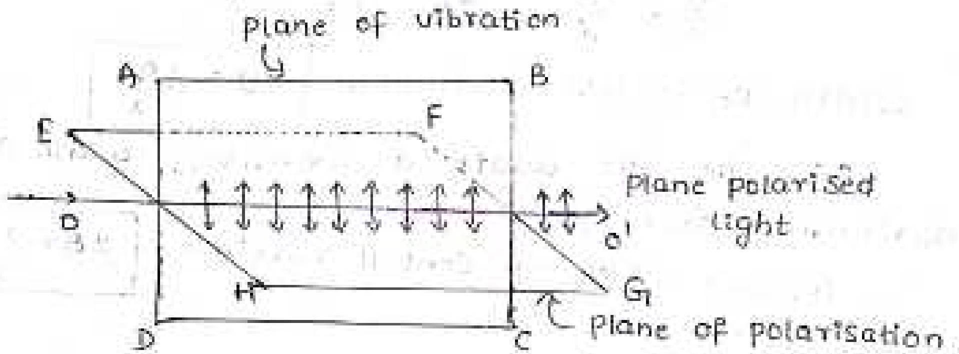
Light is an electromagnetic wave in which
electric and magnetic field vectors vary sinusoidally
perpendicular to each other as well as perpendicular
to the direction of propagation of wave of light.
In an ordinary or unpolarised light, the vibrations
of electric vector occur symmetrically in all
possible directions in a plane perpendicular to the
direction of propagation of light. This phenomenon
of restricting the vibrations of light (electric vector)
in a particular direction, perpendicular to the
direction of wave motion is called polarisation
of light.

Plane of vibration :-

The plane containing the direction of vibration of the electric vector and the direction of propagation of light is called the plane of vibration. In fig below, OO' is the direction of propagation of light and ABCD is the plane of vibration.

Plane of Polarisation :-

The plane containing the direction of propagation of light and perpendicular to the plane of vibration is called the plane of polarisation. Clearly, there are no vibrations of light in the plane of polarisation. In the figure, EFGH is the plane of polarisation.



Plane of vibration, polarisation and Plane polarised light.

* Polaroids :-

It is a material which can polarises light. Tourmaline is a natural polarising material. These are now artificially made. They are also used for the identification of a given light, i.e. whether the light is polarised or unpolarised.

A polaroid on which unpolarised light is incident is called polariser and on which polarised light is incident is called analyser.

when we pass the light through a analyser and rotate it about assuming the incident light as axis and examine the emergent light :-

(i) Incident light is unpolarised, if there is no change in intensity of emergent light.

(ii) If the change in intensity of emergent light is minimum but not equal to zero, then the incident light is said to be partially polarised.

(iii) If change in intensity of emergent light is minimum and equal to zero, then the incident light is said to be plane polarised or linearly polarised.

* Uses of polaroids :-

Polaroids are used

(i) in sunglasses.
(ii) to prepare filters.
(iii) for laboratory purpose.
(iv) in head lights of automobiles.
(v) in three dimensional motion pictures.
(vi) polaroids are fitted on the wind shield of the cars.
(vii) to improve colour contrast in old paintings.

* Note :- when angles between the principal sections of two nicols are 0° and 180°, they are referred to as parallel nicols. when this angle is 90° they are said to form crossed nicols.

7. Dual Nature of Matter and Nuclei

* PhotoElectric Effect :-

Max planck held the view that energy is emitted or absorbed as quanta. Einstein extended this view and proposed that light energy is composed of quanta which are called photons.

The concept of photon was so revolutionary that several leading scientists of that time includ-ing planck hesitated to accept Einstein's proposition, Ultimately the experimental verification of Einstein's Photo electric equation followed by the experimental evidence provided by Compton's effect lead to the universal acceptance of Einstein's ideas. Einstein was awarded the Nobel prize for his explanation of photo electric effect using his photon concept.

* Hertz's observation :-

Heinrich Hertz discovered the phenomenon of Photo electric effect in 1887 while conducting experi-ments on electromagnetic waves. He found that the electric sparks between two electrodes passed more readily when the electrodes are illuminated by UV rays.

Experiments carried out later, showed that several substances such as the alkali elements namely Li, Na, K, Rb emit electrons, when light falls on them. Zinc and magnesium responded only to ultraviolet radiation. This phenomenon is called photoelectric effect and the liberated electrons are called photoelectrons.

→ Photoelectric effect is the phenomenon of emiss-ion of electrons from mainly metal surfaces exposed to light energy (X-rays, γ-rays, UV-rays, visible light and even infra Red (IR) rays) of suitable frequency falls on it.

Conclusion from Hertz's observation :-

Light falling on a metal surface knocks off Some electrons from it. These electrons absorb energy from the incident radiation and Overcome the attraction of the positive ions in the Surface of the material.

* **Hallwach's and Lenard's observations :-**

Lenard and Hallwachs conducted experiments on the effect of radiation on a photo cathode. Their observations can be summarized as follows:

→ Until the frequency of the incident radiation is higher than a certain minimum value no photo electrons are emitted. This frequency is called treshold frequency, which depends on the material of the photocathode.

→ For most metals the treshold frequency corresponds to the UV region (wavelength range around 200nm to 300 nm) while for alkali metals (Na, K, Li, CS, etc) photoelectric effect occurs when even with visible light.

* **Experimental Study of Photo electric effect :-**

Fig Shows an experimental setup for studying the photoelectric effect. L and M are two plates fused into an evacuated tube (T) with a quartz window (ω). The plates are connected to a Variable dc Source and a micro ammeter in the Series.

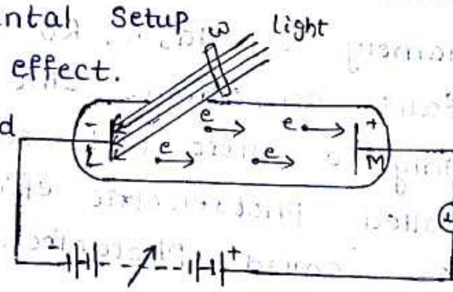


Fig. Experimental Study of Photo electric effect.

The Plate L is the Photosensitive surface connected to the negative of the dc Source and is called the photocathode.

The Plate M connected to the positive pole of the dc Source is the Collector plate. when radiation of Suitable frequency is incident on the photo cathode (through the window w) electrons are ejected from this plate. They are received by the Collector plate and a current flows through the circuit, which is registered by the microammeter.

* **Experimental Results :-**

The following experimental observations are often called laws of photoelectric effect.

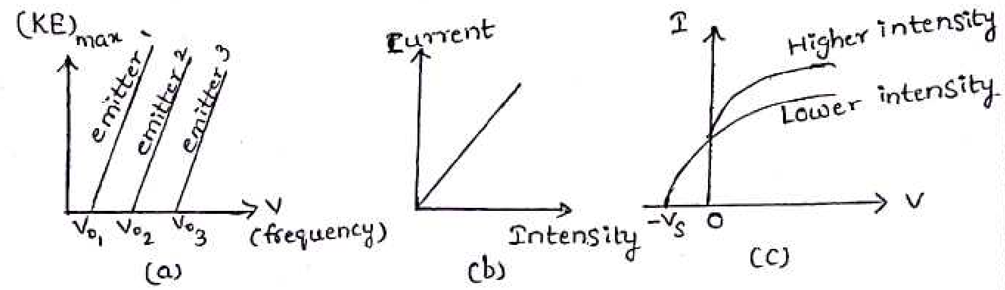


Fig. Graphs of (a) maximum KE of Photo electrons vs frequncy (b) photo current vs Intensity and (c) photocurrent vs plate potential.

(1) For a given photosensitive cathode, there is a certain frequency of radiation, below which no Photoelectric emission takes place, whatever may be the intensity. The minimum frequency is called treshold frequency (ν_0). The Corresponding wavelength is called limiting wavelength (λ_0). This is the maximum wavelength beyond which Photoemission is not possible. [Fig a] The treshold frequency is found to be different for different materials.

(2) The photoelectric effect is almost instantaneous (time lag is of the order of 10^{-9} s).

(3) when the frequency of radiation incident on a photoemissive Surface is increased above the treshold frequency, the kinetic energy of the photoelectrons increases linearly.

(4) For a given frequency of radiation higher than the treshold frequency and a given photo cathode the photoelectric current is proportional to the intensity of incident radiation.

The variation of photoelectric current with the intensity of incident radiation is as shown in Fig (b).

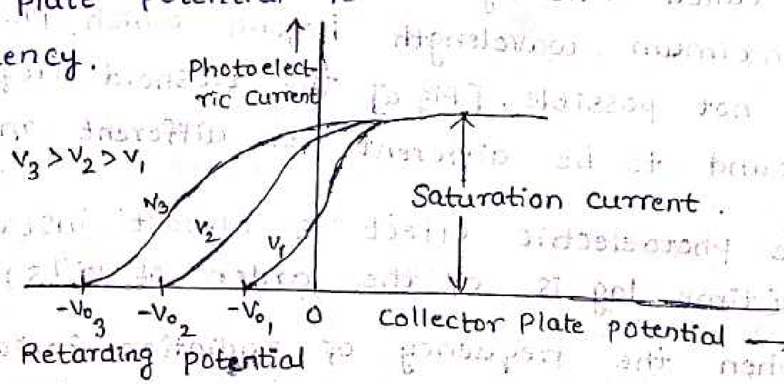
(5) As the plate potential (V) is increased, the photoelectric current (I) increases gradually upto a certain constant value called the saturation current (I_S).

The saturation current and hence the number of photoelectrons increases with the intensity of the incident radiation as shown in Fig (c).

(6) (i) If the plate potential is made negative, current still flows, but decreases as the negative potential is increased. At a particular negative potential (V_S), the photoelectric current becomes almost zero. This potential is called stopping potential as shown in Fig-(c).

→ Stopping potential is the negative voltage on the collector (anode) place at which the photo electric current just becomes zero.

(ii) Fig (2) shows the variation of photoelectric current with plate potential for radiation of different frequency.



Fig(2): Variation of photoelectric current with collector plate potential for different frequencies of incident radiation.

→ It is Seen that the Stopping potentials are different, but the Saturation current is same.

→ The energy of the emitted electrons depends on the frequency of the incident radiations.

→ The stopping potential is more negative for higher frequencies of incident radiation. This implies that greater the frequency of incident light, greater is the maximum Kinetic energy of photoelectrons. As a result, greater retarding potential is required to stop them.

* **Failure of wave theory of light to explain photoelectric effect :-**

According to wave theory, when a beam of radiation falls on the metal Surface, the free electrons at the surface absorb the radiant energy Continuously. If the intensity of the radiation is increased, the amplitude of electric and magnetic field vectors is also increased. As a result, the energy absorbed by the free electron Should also increase. This Should increase the max KE of the Photoelectron emitted, no matter what the frequency of incident radiation is. Therefore a Sufficiently intense beam of radiation over a Sufficient duration of time Should be able to supply enough energy to the free electrons in the metal surface. So that they can Come out of the metal Surface. Hence, there Should not be any treshold frequency. But, the experimental results of photo electric effect, directly Contradict these expectations of the wave theory. The fact, that the max KE of the photoelectrons (Stopping potential) is independent of intensity of radiation could not be explained on the basis of wave theory.

Further, according to wave theory, the electron Should absorb the incident energy continuously

Over the entire wavefront of radiation. As the number of electrons that absorb energy is very large, the energy absorbed per electron per unit time will be very small. Calculations show that it can take several hours or more for a single electron to pick up sufficient energy to overcome the work function and come out of the metal. This conclusion contradicts the experimental observation that the phenome photoemission is instantaneous.

Hence, the wave theory of light fails to explain the most basic observations of photoelectric effect.

* **Einstein's explanation of photoelectric effect (photon hypothesis) :-**

To account for the experimental facts of photoelectric emission, Albert Einstein, in 1905, made a simple but remarkable assumption, based on Planck's quantum theory of radiation. Einstein assumed that a beam of light is composed of discrete packets of energy called quanta or photons.

He proposed that when a photon interacts with matter, it imparts its energy (E) completely to an individual electron. Part of this energy is utilized by the electron to come out of the metal surface. The remaining energy appears as the kinetic energy of the photoelectron with which it is emitted. Hence, the energy equation can be written as

$$\begin{bmatrix}\text{Energy of the}\\ \text{incident photon}\end{bmatrix} = \begin{bmatrix}\text{The energy required}\\ \text{by electron to come}\\ \text{out of the metal}\end{bmatrix} + \begin{bmatrix}\text{max KE of}\\ \text{the photoelectron}\end{bmatrix}$$

Thus $E = W + KE_{max}$

$$h\nu = \omega + \frac{1}{2}mv^2_{max} \longrightarrow (1)$$

This is called Einstein's photoelectric equation.

→ The minimum energy required to remove an electron from the metal surface is called photoelectric work function.

It is the characteristic of the emitting material. If the frequency of the incident light is decreased, the kinetic energy of the photoelectron also decreases and becomes zero for a particular frequency called treshold frequency (v_0). Then from Eq (1), we have $h\vartheta_0 = \phi$

Substituting for ω in the Eq-(1) we get.

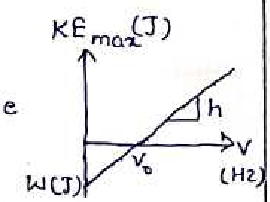
$$h\vartheta = h\vartheta_0 + \frac{1}{2}mv^2 \quad \longrightarrow (2).$$

Using this equation, the experimental observations of Photoelectric effect can be explained as follows :-

(1) Photo electric effect is the instantaneous because the collision between a photon and an electron results in instantaneous transfer of energy.

(2) If the incident photon has a frequency equal to the treshold frequency (ϑ_0) of the material, the energy is just sufficient to free the electron. If $\vartheta < \vartheta_0$ then kinetic energy is negative which is physically not possible. Hence, electrons are not liberated. This explains the existence of treshold frequency.

(3) when the intensity of incident light is increased, the number of photons increase and hence, more photo electrons are liberated. Hence, the saturation photoelectric current increases, with the increase in intensity of light.



(4) A graph of kinetic energy against the frequency is a straight line, whose slope is a constant equal to h, the Planck's constant.

(5) we have $h\vartheta = h\vartheta_0 + \frac{1}{2}mv^2_{max}$

when v is increased, ϑ_0 being constant for a given metal, kinetic energy of photoelectrons increases with frequency of the incident radiation. Hence, corresponding magnitude of the stopping potential also increases.

kinetic energy of the photoelectrons emitted is independent of the intensity of incident radiation.

* **Particle nature of light : The Photon**

Photoelectric effect gave evidence to the fact that in the interaction of light with matter, light behaves as if it is made of quanta or packets of energy, each of energy $h\nu$. Einstein associated a momentum $\left(\frac{h\nu}{c}\right)$ for the light quantum.

A definite value of energy and momentum strongly suggests that the light quantum behaves as a particle. This particle was later named Photon. The particle-like behaviour of light was further confirmed in Compton scattering of x-rays.

→ The photon picture of electromagnetic radiation can be summarized as follows :-

According to the quantum theory of radiation, emission or absorption of radiant energy takes place in discrete amount called quanta. Einstein called these quanta of light as photons. Thus, light radiation consists of photons. According to this theory, photon possess the following properties.

(1) Photons travel at a speed of 3×10^8 m/s in vacuum and a photon can never be at rest.

(2) The energy of photon is given by $E = h\nu$, where ν is frequency of light.

(3) Photons possess momentum and thus behave like particles. The linear momentum of a photon is given by $p = \frac{h}{\lambda}$ where $h = 6.6 \times 10^{-34}$ J s is Planck's constant and λ is wavelength of light.

(4) The rest mass of a photon is zero.

(5) During interaction of radiation with matter, photons behave like particles.

(6) In a photon - particle interaction, total energy and total momentum are conserved. But the number of photons is not conserved.

(7) Photons are electrically neutral. Therefore, photons cannot be deflected by electric and magnetic fields.

* Wave Nature of Matter :-

In 1924, the French physicist, Louis de Broglie, proposed that material particles can also exhibit wave like properties while in motion. A wavelength, which is a characteristic property of a wave, can be assigned to a material particle. Louis de Broglie was lead to this conclusion by the following considerations.

(1) Nature loves Symmetry :-

Entire universe manifests in two forms : Energy and Matter. Nature loves Symmetry. Nature cannot be partial to existence of one form. If light can have dual nature so should matter have. That is, matter should behave both as a particle and as a wave.

(2) Matter and Energy are equivalent :-

Einstein, in his special theory of relativity developed in 1905, had established the equivalence of mass and energy, through the relation $E = mc^2$; where c is the speed of light. This suggests the possibility that mass or matter can exhibit wave like properties.

(3) Bohr's theory of atomic Structure :-

In Bohr's theory of atomic structure the orbital radii and angular momenta of electrons are associated with integers. This suggests some periodicity associated with electrons. But periodicity is normally associated with wave phenomena like interference; stationary waves, etc...

Thinking on these lines, de Broglie concluded that a material object is associated with a wave. Under suitable conditions, material particles should exhibit a behaviour which can be explained only on the basis of waves associated with them.

The wave associated with a particle in motion to describe its wave like character is called a matter wave.

Scanned by CamScanner

* **Expression for de Broglie wavelength :-**

An expression for the wavelength of matter waves may be arrived at as follows.

A moving body behaves in certain ways as though it has a wave nature. A photon of light of frequency ν has the momentum.

$$p = \frac{h\nu}{c}$$

This can be expressed in terms of wavelength λ as $\quad p = \frac{h}{\lambda} \quad$ (since $c = \lambda\nu$)

The wavelength of a photon is therefore specified by its momentum according to the relation. $\quad \lambda = \frac{h}{p}$

de Broglie suggested that the above equation is completely general and that it applies equally to any material particle as well as to a photon. The momentum of a particle of mass 'm' and velocity 'v' is $p = mv$ and its de Broglie wavelength is accordingly.

$$\lambda = \frac{h}{mv}$$

The greater the particle's momentum, the shorter is its wavelength.

In the case of bodies that we come across in everyday life, the momentum which is the product of mass and velocity is quite large and the resulting de Broglie wavelengths are too small to measure. In the case of sub atomic particles like electrons, the momenta are quite small and the de Broglie wavelength ranges from a fractions of a nanometer to a few hundred nanometre and is measurable.

* **de Broglie wavelength of electron :-**

Consider an electron of mass m and charge e accelerated from rest through a potential difference V. The kinetic energy gained by the electron is

k = Charge × potential difference = eV.

The kinetic energy and linear momentum are related by $k = \frac{p^2}{2m} \Rightarrow p = \sqrt{2mK}$

The de Broglie wavelength of electrons is, thus,

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mK}} = \frac{h}{\sqrt{2meV}}$$

$(h = 6.625 \times 10^{-34}\ Js,\ m = 9.1 \times 10^{-31}\ kg,\ e = 1.6 \times 10^{-19}\ C)$

we get $\lambda = \frac{1.227}{\sqrt{V}}$ nm.

* **Heisenberg's uncertainty principle :-**

In classical physics both the position and momentum of a particle can be measured simultaneously to any specified accuracy. However, with de Broglie's hypothesis of wave - particle duality this possibility is ruled out for subatomic particles.

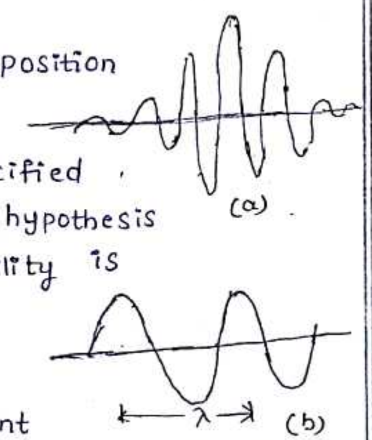


Fig: uncertainity principle.

Imagine a subatomic particle like an electron. Assume that we want to measure its momentum accurately. Since the de Broglie wavelength $\lambda = \frac{h}{p}$; for p being obtained with a desired accuracy, λ must be accurate.

→ λ can be accurate only if we have a large number of wavelengths involved in the measurement of λ.

→ If the wave is so spread, momentum is accurate but the position of electron is not defined.

It can be anywhere from $-\infty$ to $+\infty$. Thus, to overcome this situation we can superpose a number of waves to form a compact wave packet to locate the position with reasonable accuracy.

If Δx is the uncertainity in position, obviously there is also a finite uncertainty in the measurement of momentum. If $\Delta x \to 0$, then $\Delta p \to \infty$. Thus both position and momentum of a particle cannot be simultaneously measured accurately. This principle is called Heisenberg's uncertainty principle stated as follows :

Statement :- It is not possible to measure Simultaneou-Sly both the position and momentum of a particle (say electron) to any arbitary accuracy.

Mathematically it can be Shown that,

$$\Delta x \cdot \Delta p \approx \frac{h}{2\pi}$$

where h= Planck's constant and Δx and Δp are the uncertainities in the Simultaneous measurement of position and momentum of a particle.

* **Nucleus and Its Composition :-**

In every atom, the positive charge and mass are densely concentrated at the centre of the atom forming its Nucleus. The overall dimen-sions of a nucleus are much Smaller than those of an atom. The radius of the nucleus is Smaller than the radius of an atom by a factor of 10^4. This means the volume of a nucleus is about 10^{-12} times the volume of the atom.

* **Composition of Nucleus :-**

The nucleus was first discovered in 1911 by Lord Rutherford and his associates by experi-ments on Scattering of α- particles by atoms. He found that the Scattering results Could be explai-ned, if atoms Consist of a Small, central, massive and positive core surrounded by orbiting electrons. The experimental results indicated that the size of the nucleus is of the order 10^{-14}m. The Study of radioactivity revealed that nucleus is not a composite body, but it is made of nucleons. The positive charge in the nucleus is that of the protons. A proton carries one unit of fundamen-tal charge. A free proton is Stable.

* Atomic Mass unit :-

The mass of an atom is very Small. Kilogram Cannot be used to measure Such Small quantity of mass. It is measured by a unit called atomic mass unit (amu i.e. u).

It is defined as

$$1u = \frac{\text{mass of one } ^{12}C \text{ atom}}{12} = \frac{1.992647 \times 10^{26} \text{ kg}}{12}$$

$$= 1.660539 \times 10^{27} \text{ kg}$$

Atomic masses are measured by an instrument Called mass Spectrometer.

* Atomic Number (Z) :-

Atomic number of an element is the number of protons present inside the nucleus of an atom of the element. It is also equal to the number of electron revolving in various orbits around the nucleus of the neutral atom.

Atomic number, Z = Number of Protons
= No. of electrons (in a neutral atom).

* Mass Number (A) :-

Mass number of an element is the total number of protons and neutrons inside the atomic nucleus of the element.

Mass number, A = Number of protons + Number of Neutrons.

= Number of electrons (in a neutral atom) + Number of neutrons

= Atomic number + Number of neutrons = Z+N

The term nucleon is also used for Neutron and proton. Thus, the number of nucleons in an atom is its mass number A.

Nuclear Species or nuclides are shown by the notation $^A_Z X$, where X = Chemical Symbol of the Species.

* Ex-1 :- In a nucleus of $_{92}U^{238}$, find the number of protons and the number of neutrons.

Sol) Number of Protons, Z = 92

∴ Number of Neutrons, N = A-Z = 238 - 92 = 146.

* **Classification of nuclides :-**

Nuclides (or nuclear species) are denoted by $^{A}_{Z}X$, where X is the chemical symbol of the species. For Example, one specie of uranium is denoted by $^{235}_{92}U$. It contains 92 protons, 143 neutrons and 235 nucleons.

Isotopes :-

Isotopes of an element are the atoms of that element which have the same atomic number but different mass number.

[Iso means Same, tope means place] ⇒ occupy Same place in the periodic table; istopes of an element will have Same proton number].

Ex:-

Sl. no.	Name of the element	Atomic number	Isotope (s)
1.	Hydrogen	1	^{1}H, ^{2}H, ^{3}H
2.	Helium	2	^{3}He, ^{4}He, ^{6}He
3.	Beryllium	4	^{7}Be, ^{8}Be, ^{9}Be, ^{10}Be
4.	Silver	47	^{107}Ag, ^{108}Ag, ^{109}Ag
5.	Gold	79	^{197}Au

Isotopes of an element have identical chemical properties as number of electrons in the isotopes are the same. As the masses of isotopes differ, they possess different physical properties. Isotopes may have same name or different name. For example : ^{1}H is called hydrogen, ^{2}H is called deuterium and ^{3}H is called tritium. The isotope ^{14}N, ^{15}N are called by the same name - nitrogen.

Isobars :-

ISobars are the atoms of different elements which have the Same atomic weight and Same mass number but different proton number.

[ISo means Same, bar means weight].

ISobars Contain different number of protons, different number of electrons and also different number of neutrons. The total number of nucleons in them is the same.

Examples : (i) ^{14}C, ^{14}N (ii) ^{235}U, ^{235}Th

(iii) ^{58}Fe, ^{58}Ni, ^{58}Co

The chemical properties of isobars are widely different.

Isotones :-

ISotones are the atoms whose nuclei contain Same number of neutrons.

[This name was arrived at by replacing p (for photon) in the word isotopes by n (for neutron)].

$N = A - Z$ is the Same for atoms which are isotopes.

Ex :- (i) $^{3}_{1}H$, $^{4}_{2}He$ are isotones. Each contains 2 neutrons

(ii) $^{15}_{7}N$, $^{16}_{8}O$ are isotones. Each Contains 8 neutrons.

* Characteristics of a nucleus :-

(i) Radius of a nucleus :-

Instead of α- particles, Rutherford used fast electrons as projectiles which bombarded tangents made of various elements. It was found that the volume (v) of a nucleus is directly proportional to the number of nucleons (A - the mass number of the nucleus).

If R is the radius of the nucleus, then its volume $\propto A$ $\Rightarrow \frac{4}{3}\pi R^3 \propto A$, we have $R^3 \propto A$.

or $R \propto A^{1/3}$ (or) $\boxed{R = R_0 A^{1/3}}$ (where $R_0 = 1.2\,fm$, $R_0 = 1.2 \times 10^{-15}\,m$)

→ Ex :- Radius of $^{8}_{4}Be = 1.2 \times (8)^{1/3} = 2.4\,fm$

→ AS A is different for different elements, therefore atomic nuclei of different elements have different sizes.

(ii) Nuclear mass :-

Nuclear mass is found to be less than the Sum of protons and neutrons. Neglecting the mass of electrons, the nuclear mass can be taken to be the mass of atom itself. Mass Spectrographs are used to determine atomic masses.

Atomic masses are measured in Comparison with the atomic mass of Carbon -12 isotope, taken as 12 atomic mass units (12u).

Nuclear mass $< [Zm_p + (A-Z)m_n]$

where Z = atomic number, A = mass number, m_p and m_n are the masses of proton and neutron respectively.

(iii) Nuclear density :-

The nucleus occupies a very Small Space inside the atom and almost the entire mass of the atom is confined to the nucleus. Hence, nucleus has a very high density.

We know $\left(\text{density} = \frac{\text{mass}}{\text{Volume}}\right)$ of the nucleus is of the order of $10^{17} kgm^{-1}$.

Assuming the nucleus to be Spherical in Shape, its volume is proportional to R^3.

$[\because$ Volume of a sphere, $V = \frac{4}{3}\pi R^3$

Nuclear mass $= Zm_p + (A-Z)m_n \simeq Am_p \quad [\because m_n \simeq m_p]$

Nuclear Volume $= \frac{4}{3}\pi R^3 = \frac{4}{3}\pi R_0^3 A$

Nuclear density, $\rho = \frac{\text{nuclear mass}}{\text{nuclear volume}} = \frac{Am_p}{(\frac{4}{3})\pi R_0^3 A}$

$$\rho = \frac{3m_p}{4\pi R_0^3} = \frac{3\times 1.67\times 10^{-27}}{4\pi \times (1.2\times 10^{-15})^3} = 2.3\times 10^{17} kg\, m^{-3}$$

Thus, the nuclear density is of the order $10^{17} Kg\, m^{-3}$. This expression shows that the nuclear density is approximately a constant for all nuclei and is independent of mass number.

(iv) Nuclear charge :-

The positive charge of the nucleus is due to the protons contained in it. Each proton has a charge e, equal in magnitude to that of electron i.e 1.6×10^{-19}C. If the atomic number of the nucleus is Z. then,

Nuclear charge = Ze

* Einstein's Mass - Energy Relation :-

In classical physics, mass and energy were considered to be two separate entities. In 1905, Einstein established from his theory of relativity, that mass and energy are equivalent. The energy equivalent of mass m is given by.

$$E = mc^2$$

where c is the velocity of light in vacuum.

This equation is called Einstein's mass - energy relation. Mass can be converted into energy and energy can be converted into mass. The factor c^2 appears as the conversion factor between mass and energy. According to this relation, mass and energy can be considered as two forms of the same physical entity.

Nuclear fission, nuclear fusion, pair production and annihilation in cosmic rays provide experimental evidence for Einstein's mass energy relation.

Examples :-

(1) During nuclear reactions such as fission and fusion, the total mass of the products is less than the total mass of the reactants. It is this difference in mass which is liberated as energy during the reaction.

(2) When an electron and a positron come very close to each other, they annihilate. Their mass, thus appears as energy and is radiated in the form of γ- rays.

(3) Under Suitable conditions, a γ-ray (of energy ≥ 1.02 Mev) Can result in electron - positron pair production.

* **Atomic mass unit :-**

Atomic mass unit (u) is defined as $(1/12)^{th}$ the mass of a Carbon - 12 atom.

i.e. $1u = \frac{1}{12}$ (mass of one atom of C - 12 isotope)

The number of atoms in 12kg of C - 12 isotope is equal to Avagadro's number, [$N_A = 6.0221367 \times 10^{26}$ per kilo mole]

$$\text{Mass of one atom} = \frac{12}{6.0221367 \times 10^{26}} \text{ kg}$$

$$\text{Hence } 1u = \frac{1}{12}\left[\frac{12}{6.0221367 \times 10^{26}}\right] \text{kg} = 1.6605402 \times 10^{-27} \text{ kg}$$

we take $1 \text{ amu} = 1.66 \times 10^{-27}$ kg or $1u = 1.66 \times 10^{-27}$ kg.

* **Electron volt :-**

Since mass and energy are equivalent, the masses of particles such as electron, proton, neutron, are expressed in terms of an energy unit called electron volt (eV).

→ One electron volt is defined as the energy gained or lost by an electron in travelling through a potential difference of one volt.

It can be shown that $1 \text{ eV} = 1.6 \times 10^{-19}$ joule.

* **Relation between u and ev :-**

According to Einstein's mass energy relation $E = mc^2$. It can be shown that the energy equivalent of 1 amu = 931.48 Mev

$1 \text{ amu} = 1.6605402 \times 10^{-27}$ kg and $c = 2.99 \times 10^8$ m/s

$$\therefore E = \frac{1.6605402 \times 10^{-27} \times (2.99 \times 10^8)^2}{1.6 \times 10^{-19}} \text{ ev} \qquad (E = mc^2)$$

$$E = 931.5 \times 10^6 \text{ ev}$$

$$E = 931.5 \text{ Mev}$$

* **Some masses in Various units :-**

Particle	Mass (Kg)	Mass(u)	Mass (MeV/c^2)
Proton	1.6726×10^{-27}	1.007276	938.28
Neutron	1.6750×10^{-27}	1.008665	939.57
Electron	9.1095×10^{-31}	5.486×10^{-4}	0.511
^1_1H atom	1.6736×10^{-27}	1.007825	938.79

* **Mass Defect :-**

It is found that the sum of the masses of the constituent particles of an atom is always slightly greater than the mass of the atom. The difference between the sum of the masses of the constituent particles and the mass of an atom is called mass defect.

Thus, mass defect $\Delta m = \left[Zm_p + (A-Z)m_n + Zm_e\right] - M$

where Z is the atomic number, A is the mass number, M is the mass of the atom and m_p, m_n, and m_e are the free state rest masses of proton, neutron and electron respectively.

* **Nuclear Binding energy :-**

An atomic nucleus is a stable structure. A stable nucleus has a smaller mass than the combined masses of its constituent particles.

For example, Mass of deuterium atom = 2.014102 u

Mass of hydrogen atom plus that of neutron

$= m_H + m_n$

$= 1.007825 + 1.008665$

$= 2.016490\,u$

Mass defect $= 0.002388\,u$

Thus, when a deuteron is formed from a free proton and neutron, mass is lost which appears as binding energy of deuteron.

The energy released in the formation of deuteron is $0.002388 u \times 931 \ MeV/u = 2.23 \ MeV$. This is the energy with which proton and neutron are bound inside the deuterium nucleus. Conversely, an amount of energy equal to 2.23 MeV has to be supplied from an external source to break a deuteron into a proton and a neutron.

The energy equivalent of the mass defect in a nucleus is called binding energy and is a measure of the stability of the nucleus. Binding energies arise from the action of the forces that hold nucleons together to form nuclei.

Definition :- The binding energy of a nucleus is defined as the energy required to be provided from external agency to separate the nucleons an infinite distance apart from the nucleus, so that they may not interact with one another.

The binding energy is related to mass defect (ΔM) by Einstein's mass-energy relation is

$$B = (\Delta M) c^2$$

If a nucleus ${}^{A}_{Z}X$ has an atomic mass M then the mass defect is given by

$$\Delta M = \left[Z m_p + (A-Z) m_n + Z m_e \right] - M$$

where m_p is the mass of proton.

m_n is the mass of neutron.

m_e is the mass of electron.

Hence, $B = \left\{ \left[Z m_p + (A-Z) m_n + Z m_e \right] - M \right\} c^2$

If M_N represents nuclear mass, we can also write $B = \left[Z m_p + (A-Z) m_n - M_N \right] c^2$

where masses are expressed in atomic mass unit (u).

Specific binding energy (or Average binding energy) :-

The Specific binding energy of a nucleus is the ratio of total binding energy to the total number of nucleons. $B_{av} = \frac{B}{A}$

The Specific binding energy i.e., binding energy per nucleon gives a measure of Stability of the nucleus. The Specific binding energy is more useful than total binding energy.

Definition :- The Specific binding energy of a nucleus is the average energy with which each nucleon is bound to the nucleus and is defined as the energy required to be Supplied from an external agency to Separate a nucleon from the nucleus to infinite distance.

Fig Shows a graph of Specific binding energy against mass number for a large number of nuclei.

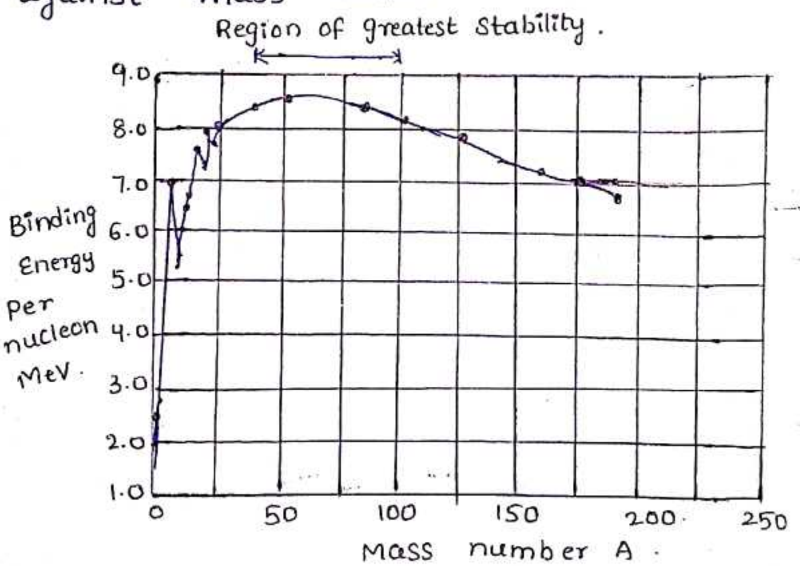


Fig :- Specific binding energy curve.

* Observations of main features of the graph are the following :-

(1) Specific binding energy for light nuclei like ^2H, ^3H is Small.

(2) For $A < 20$, the Specific binding energy curve Shows variations with well defined maxima and minima.

(3) The specific binding energy is approximately a constant for $A = 30$ to $A = 120$. The specific binding energy is maximum for ^{56}Fe and has a value about 8.8 MeV / nucleon.

(4) For $A > 120$, the specific binding energy decreases gradually falling to about 7.6 MeV / nucleon for ^{238}U.

Conclusions from the observations of graph :-

(1) The nuclear force is attractive and sufficiently strong to produce binding energy of few MeV per nucleon.

(2) Larger the specific binding energy, greater is the stability of the nucleus. In the case of light nuclei, nuclei lying at a maxima are relatively more stable than the other nuclei in the neighbourhood.

(3) Generally, light nuclei and heavy nuclei have low average binding energy and hence they are relatively less stable.

(4) The approximately constancy of specific binding energy in the range $30 < A < 120$ shows that the average specific binding energy is large, which means more energy is required to break. Thus, these nuclei are more stable.

(5) The binding energy curve shows that if a nucleus with $A = 240$ splits into two nuclei with $A = 120$, the specific binding energy increases. This implies that energy would be released in the fission process.

(6) If two nuclei with $A < 10$ were to be combined to produce heavier nucleus, then we find that the specific binding energy increases. This indicates that energy would be released in the fusion process.

* Radioactivity :-

A French Physicist H Becquerel discovered radioactivity in 1896, by accident. Radioactivity is a Spontaneous nuclear phenomenon in which an unstable nucleus undergoes a decay with the emission of some particles (α, β) and electromagnetic radiation (γ- rays).

In nature, three types of radioactive decay occurs, which are given below:

(i) α- decay :- In α-decay, a helium nucleus ${}^{4}_{2}H$ is emitted.

(ii) In β- decay, electrons or positrons with the Same mass as electron are emitted.

(iii) In γ- decay, high energy ≈ 100 keV Photons are emitted.

* Properties of α- particles :-

An α- particle carries +2e charge and mass equal to that of proton, i.e. mass of four times that of hydrogen atom equal to the mass of the helium nucleus.

Some important properties of α- particle are as follows:

(i) α- particles are deflected in electric and magnetic fields.

(ii) The velocity of α- particles varies from 0.01 - 0.1 times of c (velocity of light).

(iii) α- particles have low penetrating power which varies 8.6 cm for different radioactive elements.

(iv) α- particles have high ionisation power. Their ionising power is 100 times greater than that of β-rays and 1000 times greater than that of γ- rays.

(v) α- particles produce fluorescence in substances like zinc sulphide and barium platinocyanide. They produce scintillation or fluorescent screen.

(vi) α- particles feebly affect photographic plate. They also produce heating effect, when stopped and cause incurable burns on human body.

* **Properties of β- particles :-**

A β- particle ($_{-1}e^{0}$) has a charge of electron. Actually it is a fast moving electron. (Not the orbital electron of the atom but it is emitted from the nucleus).

Some important properties of β- particles are as follows :

(i) β- particles are also deflected by electric and magnetic fields. Their deflection is much larger than the deflection of α- particles.

(ii) The velocity of β- particles varies from 1% to 99% of the velocity of light.

(iii) As, the velocity of β- particles is of the order of the velocity of light, its mass increases with increase in their velocity. Mass of β- particles is given by Einstein's theory of relativity, $m = \frac{m_0}{\sqrt{1-(v^2/c^2)}}$

where m_0 = rest mass of β- particle with velocity v.

(iv) β- particles have high penetrating power (100 times larger than the penetrating power of α- particles). They can pass through 1mm thick sheet of aluminium.

(v) β- particles have low ionising power, i.e (1/100)th of the ionising power of α- particles.

(vi) β- Particles produce florescence in calcium tungstate, barrium platinocyanide and zinc sulphide. And it affects photographic plate more than α- particles.

* **Properties of γ- Rays :-**

γ- rays $(_0\gamma^0)$ are high energy electromagnetic radiation of nuclear origin and short wavelength ($\approx 0.01 \mathring{A}$). It is about (1/100) th part of the wavelength of x-rays. Radiation of γ-rays is a nuclear property whereas x-rays is due to atomic property.

Some important properties of γ-rays are as follows :

(i) As, the γ-rays do not have any charge, they are not deflected by electric and magnetic fields.

(ii) γ- rays travel with the speed of light (3×10^8 m/s)

(iii) γ- rays have highest penetrating power, i.e. more than that of α and β - particles. They can pass through 30cm thick iron sheet.

(iv) γ- rays have least ionisation power as compared to that of α and β particles.

(v) γ- rays produce fluorescence. They also affect photographic plate more than β- particles.

(vi) As, the γ-rays have high energy, they can give rise to the phenomenon of pair production. When a photon of γ-rays strikes the nucleus of an atom, its energy is converted into an electron and a positron (positively charged electron).

$$\underset{\text{(photon)}}{\gamma} \longrightarrow \underset{\text{(electron)}}{e^-} + \underset{\text{(positron)}}{e^+}$$

Note :- α- particle is a doubly ionised helium atom or a helium nucleus. and β- particle is a fast moving electron from the nucleus.

* Nuclear Fission :-

Nuclear fission is the phenomenon of splitting of a heavy nucleus (usually $A > 230$) into two or more lighter nuclei by the bombardment of proton, neutron, α- particle, etc...

Energies associated with nuclear processes are about a million times larger than chemical process.

In fission, a heavy nucleus like $^{235}_{92}U$ breaks into two smaller fragments by the bombardment of thermal neutron (low energy or slow moving).

e.g :- $^{235}_{92}U + {}^{1}_{0}n \longrightarrow {}^{141}_{56}Ba + {}^{92}_{36}Kr + 3\,{}^{1}_{0}n + Q$ (Energy released)

Q-value here refer to the energy released in the nuclear process, which can be determined using mass-energy relation, $E = mc^2$. (Here 'm' is the difference of mass of products and reactants).

Energy released per fission of $^{235}_{92}U$ is 200.4 MeV. The fragment nuclei produced in fission are highly unstable. They are highly radioactive and emit β- particles in succession until each reaches to a stable end product.

Nuclear Chain Reaction :-

In the nuclear fission reaction, there is a release of extra neutrons. The extra neutrons in turn initiate fission process, producing still more neutrons and so on. Thus, a chain of nuclear fission is set up called nuclear chain reaction. The chain reactions may be of two types :-

(1) Uncontrolled, (2) Controlled chain reactions.

(1) Uncontrolled Chain Reaction :-

During fission reaction, neutrons released are again absorbed by the fissile isotopes, the cycle repeates to give a chain reaction, i.e self - sustaining and gives off energy at a rate that increases rapidly with time leading to large amount of radiation. This is called uncontrolled chain reaction.

(2) Controlled Chain Reaction :-

If by some means, the reaction is controlled in such a way that only one of the neutrons emitted in a fission causes another fission, then the fission rate remains constant and the energy is released steadily. Such a chain reaction is called a controlled chain reaction. It is used in a nuclear reactor.

The sustained fissibility of nuclear chain reaction depends on the multiplication factor or reproduction factor K.

$$K = \frac{\text{Rate of Production of neutrons}}{\text{Rate of loss of neutrons}}$$

→ If $K=1$, the operation of reactor is said to be critical. It is what we wish to be for steady power operation.

→ If $K>1$, the reaction rate and reactor power increases exponentially. In this case, reaction is super - critical and can even explode.

→ If $K<1$, the reaction gradually stops. And the condition is called sub- critical.

* Nuclear Fusion :-

Nuclear fusion is the phenomenon of fusing two or more lighter nuclei forming a single heavy nucleus.

For fusion to take place, the two nuclei must come close enough so that, attractive short range nuclear force is able to affect them. Since both the nuclei are positively charged particles, so they experience coloumb's repulsion. Therefore, they must have enough energy to overcome this Coloumb barrier.

e.g: $_1H^1 + _1H^1 \longrightarrow _1H^2 + _1e^+ + \nu + 0.42\,MeV$

$_1H^2 + _1H^2 \longrightarrow _2He^3 + _0n^1 + 3.27\,MeV.$

$_1H^2 + _1H^2 \longrightarrow _1H^3 + _1H^1 + 4.03\,MeV.$

Fusion of hydrogen nuclei into helium nuclei is the source of energy of most of the stars including the sun.

* **Energy Generation in Stars :-**

Thermonuclear fusion is the source of energy output in the interior of stars. The interior of the sun has a temperature of $1.5 \times 10^7 K$, which is considerably less than the estimated temperature required for fusion of particles of average energy.

Fusion in the sun involves protons whose energies are much above the average energy, i.e. protons which are in the high velocity fall of Maxwell - Boltzmann distribution.

The fusion reaction in the sun is a multi-step process in which the hydrogen is fused into helium. The proton-proton (P, P) cycle by which this occurs is represented by the following sets of reactions :-

$${}^1_1H + {}^1_1H \longrightarrow {}^2_1H + e^+ + \nu + 0.42\,MeV \longrightarrow (i)$$

$$e^+ + e^- \longrightarrow \gamma + \gamma + 1.02\,MeV \dashrightarrow (ii)$$

$${}^2_1H + {}^1_1H \longrightarrow {}^3_2He + \gamma + 5.49\,MeV \longrightarrow (iii)$$

$${}^3_2He + {}^3_2He \longrightarrow {}^4_2He + {}^1_1H + {}^1_1H + 12.86\,MeV \longrightarrow (iv)$$

For the fourth reaction to occur, the first three reactions must occur twice, in case two light helium nuclei unite to form ordinary helium nucleus. If we consider the combination 2(i) + 2(ii) + (iv), the net effect is $4\,{}^{1}_{1}H + 2e^- \rightarrow {}^{4}_{2}He + 2\nu + 6\gamma + 26.7\ MeV$.

(or) $(4\,{}^{1}_{1}H + 4e^-) \rightarrow ({}^{4}_{2}He + 2e^-) + 2\nu + 6\gamma + 26.7\ MeV$.

Thus, four hydrogen atoms combine ${}^{4}_{2}He$ atom with a release of 26.7 MeV of energy.

As the hydrogen in the core gets depleted and becomes helium, the core starts to cool. The star begins to collapse under its own gravity, which increases the temperature of the core.

⇒ The age of the Sun is about 5×10^{9} yr and it is estimated that there is enough hydrogen in the sun to keep it going for another 5 billion years.

* **Distinction between Nuclear Fission and Nuclear Fusion :-**

(i) Fission is the splitting of large nucleus into two or more smaller ones, on the other hand, fusion is the combining of two or more lighter nuclei to form larger one.

(ii) Fission does not normally occur in nature but fusion occurs in stars such as the Sun.

(iii) Fission requires critical mass of the substance and high speed neutrons but in fusion, high density and high temperature environment are required.

(iv) In fission, energy released is million times greater than in chemical reactions, but lower than energy released by nuclear fusion.

* Difference between nuclear fission and fusion :-

Nuclear fission	Nuclear fusion .
(1) The process of splitting of a heavy nucleus into two light nuclei of comparable masses is called fission.	(1) The process of combining two light nuclei to produce a heavy nucleus is called fusion.
(2) Fission is initiated by a slow neutron. Hence easier to start fission.	(2) Fusion is initiated at very high temperatures of about 10^8 K. Hence very difficult to achieve fusion.
(3) Fission can be controlled	(3) Fusion cannot be controlled.
(4) Energy released is small. It is less than 1MeV per nucleon.	(4) Energy released is large. It is greater than 1MeV per nucleon.
(5) Heavy nuclei are required for fission which are not abundantly available. Hence the source of energy is restricted.	(5) Light nuclei are abundantly available in nature. Hence an infinite source of energy exists.
(6) Fission products are generally radioactive and harmful to the environment.	(6) The fusion products are stable and hence they are friendly to the environment.
(7) Used to construct nuclear bombs.	(7) Used to construct hydrogen bombs. (hydrogen bombs use both fission and fusion reaction).

8. Electronic Devices

* **Classifications of Metals, Conductors and Semi Conductors on the basis of conductivity :-**

On the basis of the relative values of electrical conductivity (σ) or resistivity ($\rho = 1/\sigma$) the solids are broadly classified as,

(i) Metals :- They possess very low resistivity (or high conductivity)

$\rho \sim 10^{-2} - 10^{-8}\ \Omega m$, $\sigma \sim 10^{2} - 10^{8}\ Sm^{-1}$

(ii) Semiconductors :- They have resistivity or conductivity intermediate to metals and insulators.

$\rho \sim 10^{-5} - 10^{6}\ \Omega m$, $\sigma \sim 10^{+5} - 10^{-6}\ Sm^{-1}$

(iii) Insulators :- They have high resistivity (or low conductivity). $\rho - 10^{11} - 10^{19}\ \Omega m$, $\sigma = 10^{-11} - 10^{-19}\ Sm^{-1}$

The values of ρ and σ given above are indicative of magnitude and could well go outside the ranges as well.

Our interest in this chapter is in the study of semiconductors, which can be of the following types

(i) Element Semiconductors :- These semiconductors are available in natural form.
e.g. Silicon and Germanium.

(ii) Compound Semiconductors :- These semiconductors are made by compounding the metals.
e.g. (a) Inorganic Semiconductors are CdS, GaAs, CdSe, InP, etc...

(b) Organic Semi Conductors are anthracene, doped Pthalocyanines, etc..

(c) Organic Polymer Semiconductors are Poly-Pyrole, Polyaniline, Polythiophene, etc..

* Energy Bands in Solids (Conductor, Insulator and Semiconductor).

Energy Band :-

According to Bohr's atomic model and Conce-
pt of electronic configuration in an isolated atom,
the electrons have certain definite discrete amounts
of energy corresponding to different shells and
subshells, i.e. there are well-defined energy levels
of electrons in an isolated atom.

But in a crystal due to interatomic
interaction, valence electrons are shared by more
than one atom. Due to this, splitting of energy
level takes place.

The collection of these closely spaced
energy levels is called an energy band. These
bands are formed due to the continuous energy
variation in different energy levels.

These different energy levels in different
electrons are formed because inside the crystal,
each electron has a unique position and no
two electrons is exactly at the same pattern
of surrounding charges.

Valence Band :-

The energy band, which includes the
energy levels of the valence electrons is called
valence band. This band may be partially or
completely filled with electrons but is never
empty.

Conduction Band :-

The energy band above the valence
band is called conduction band. At room
temperature, this band is either empty or parti-
ally filled with electrons.

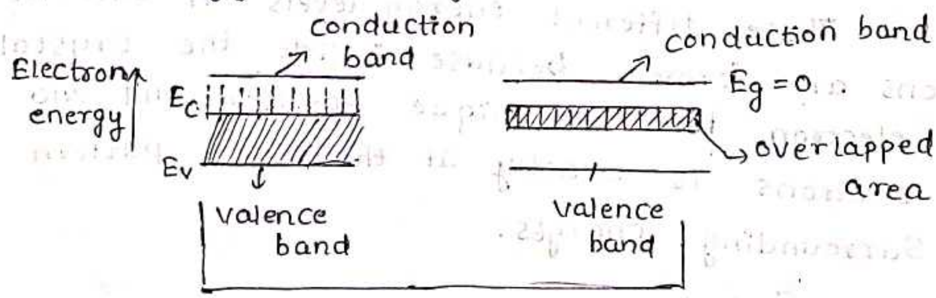
Electrons can gain energy from external electric field, then jump from valence to conduction band and contribute to the electric current.

* **Difference between Conductor, Insulator and Semiconductor on the basis of Energy Bands**

Conductor (Metal) :-

In conductor, either there is no energy gap between the conduction band which is partially filled with electrons and valence band or the conduction band and valence band overlap each other.

Thus, many electrons from below the fermi level can shift to higher energy levels above the fermi level in the conduction band and behave as free electrons by acquiring a little more energy from any other sources.

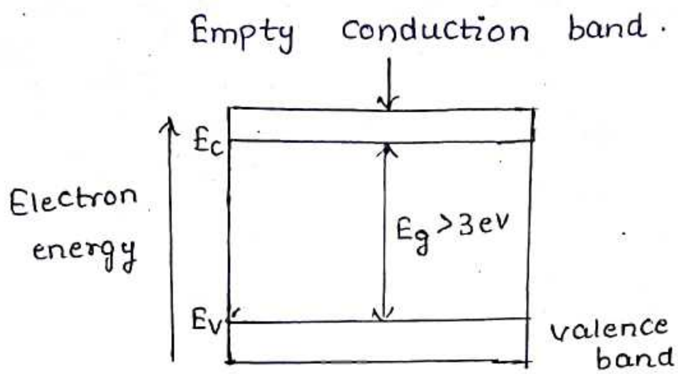


For metals.

Insulator :-

In insulator, the valence band is completely filled, the conduction band is completely empty. In this, energy gap is quite large and even energy from any other source cannot help electrons to overcome it.

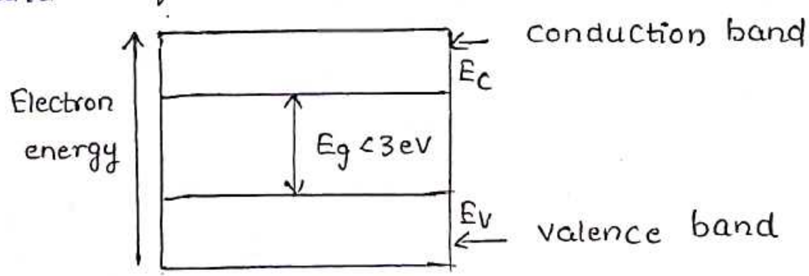
Thus, electrons are bound to valence band and are not free to move. Hence, electric conduction is not possible in this type of material.



* **Semiconductor :-**

In semiconductor, the valence band is totally filled and the conduction band is empty but the energy gap between counduction band and valence band, unlike insulators is very small.

Thus, at room temperature, some electrons in the valence band acquire thermal energy greater than energy band gap and jump over to the conduction band where they are free to move under the influence of even a small electric field and acquire small conductivity.



* **Energy Band Gap :-**

The minimum energy required for shifting electrons from valence band to conduction band is called energy band gap (E_g). It is the gap between the top of the valence band and bottom of the conduction band. It can be zero, small or large depending upon the material.

Note :- If λ is the wavelength of radiation used in shifting the electron from valence band to conduction band, then energy gap is

$$\boxed{E_g = h\nu = hc/\lambda}$$

where h is called Planck's constant and c is the velocity of light.

* **Fermi Energy :-**

It is the maximum possible energy possessed by free electrons of a material at absolute zero temperature (i.e 0K). The value of fermi energy is different for different materials.

* **Semi Conductors :-**

The materials whose conductivity lie between metals and insulators are known as semi conductors. They are characterised by narrow energy gap (less than 3eV) between the valence band and conduction band.

→ At absolute zero temperature, all states in valence band are filled and all states in conduction band are empty.

→ An applied electric field cannot give so much energy to the valence electrons that they could cross the gap and enter the conduction band. Hence, at low temperatures, pure semi conductors are insulators.

* **Electrons and Holes in Semi conductors :-**

At room temperature, however some of the valence electrons acquire thermal energy greater than E_g and move into conduction band. A vacancy is created in the valence band at each place where an electron was present before moving into conduction band. This vacancy is called hole.

Scanned by CamScanner

It is a seat of positive charge of magnitude equal to the charge of an electron. Thus, free electrons in the conduction band and the holes are created in the valence band, which can move even under a small applied field. The solid is therefore conducting.

On the basis of purity, semiconductors are of two types.

Intrinsic Semiconductors :-

This type of semiconductor is also called an undoped semiconductor. or i- type semiconductor. It is a pure semiconductor without any significant presence of dopant species.

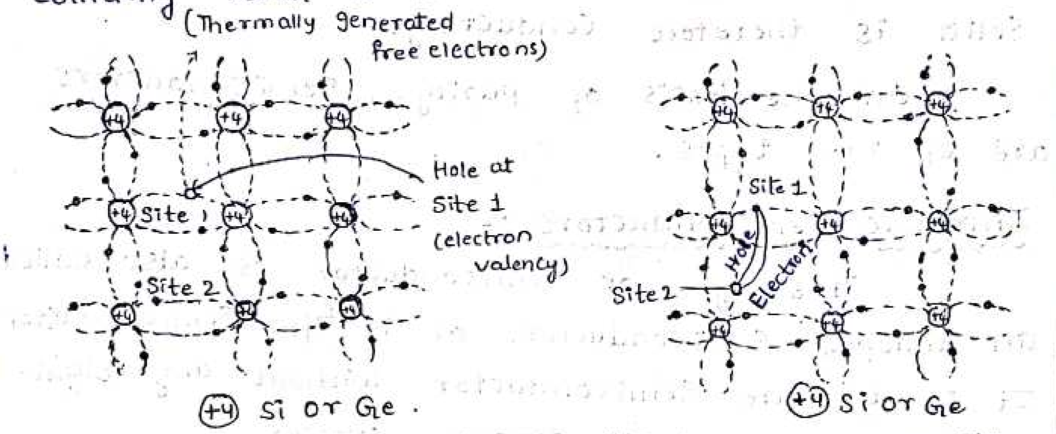
→ Pure germanium, silicon in their natural states are intrinsic semiconductors.

→ The number of charge carries is determined by the properties of the material instead of the amount of impurities.

In intrinsic semiconductors, the number of excited electrons is equal to number of holes, i.e $n_h = n_i$ (where n_i is called intrinsic carrier concentration). At temperature 0K, the valence band is filled. The energy gap is 0.72 eV and the conduction band is totally empty.

Under the action of an ~~of~~ electric field, holes move towards negative potential giving hole current I_h. The total current I is the sum of the electron current I_e and the hole current I_h i.e $I = I_e + I_h$

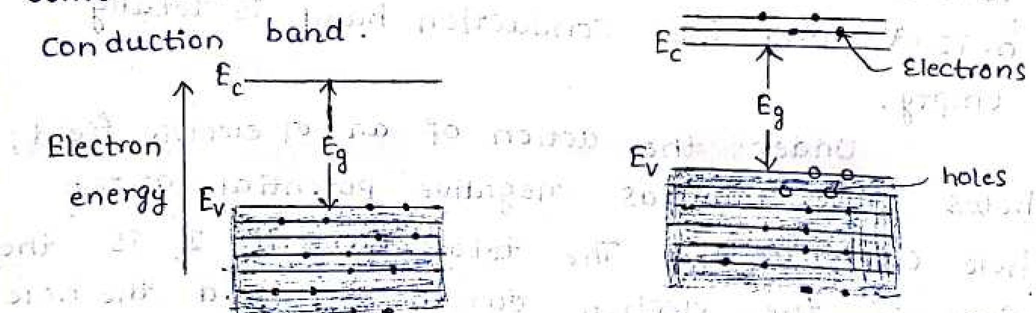
It may be noted that apart from the process of generation of conduction in electrons and holes, a simultaneous process of recombination occurs in which the electron recombine with the holes. At equilibrium, the rate of generation is equal to rate of combination of charge carriers. The recombination occurs due to an electron colliding with a hole.



(a) Representing the generation of hole at site 1 and conduction electron due to thermal energy at modern temperatures.

(b) Representing possible thermal motion of a hole. The electron from the lower left hand covalent bond (site 2) goes to the earlier hole site 1, leaving a hole at its site indicating an apparent movement of the hole from site 1 to site 2.

An intrinsic semiconductor behaves like an insulator at T = 0K. The thermal energy at higher temperature is the only reason which excites some electrons from the valence band to the conduction band.



(a) an intrinsic semiconductor at T=0K behaves like insulator

(b) Representing four thermally generated electron-hole pairs at T>0K

In fig (b) these thermally excited electrons at T>0K partially occupy the conduction band. They have come from the valence band leaving equal number of holes there.

* **Extrinsic Semiconductors :-**

The conductivity of intrinsic semiconductors is very low at room temperature. But it can be significantly increased, if some pentavalent or trivalent impurity is mixed with it. Hence, those semiconductors in which some impurity atoms are embedded are known as extrinsic or impurity semiconductors.

* **Note :-** when some desirable impurity is added to intrinsic semiconductors deliberately then this process is called dobing and the impurity are called dopants. (The process of adding impurity to an intrinsic semiconductor in a controlled manner is called doping).

There are two types of dopants used in doping.

(i) Trivalent (valency 3) atoms :- e.g. Indium (In), Boron (B), Aluminium (Al), etc..

(ii) Pentavalent (valency 5) atoms :- e.g. Arsenic (As), Antimony (Sb), phosphorous (P), etc...

Extrinsic semiconductors are basically of two types (i) n-type semiconductors (ii) P-type semiconductors.

n-Type Semiconductors :-

This type of semiconductor is obtained when pentavalent impurity is added to Si or Ge.

During doping, four electrons of pentavalent element bond with the four silicon neighbours while fifth remains very weakly bound to its parent atom. Also, the ionisation energy required to set this electron free is very small.

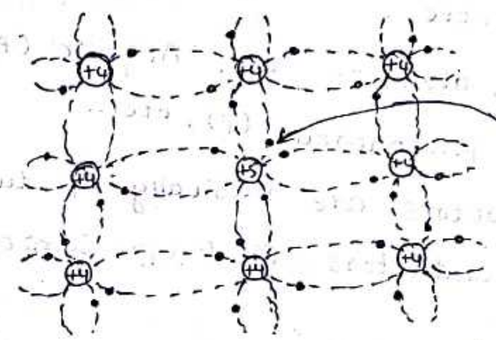
Hence, these electrons are almost free to move. In other words, we can say that these electrons are donated by the impurity atoms. So, these are also known as donor atoms.

and the conduction inside the semiconductor will take place with the help of the negatively charged electrons. Due to this negative charge, these semiconductors are known as n-type semiconductors.

when the semiconductors are placed at room temperature, then the covalent bond breakages takes place. So, more free electrons are generated. As a result, same number of holes generation takes place. But as compared to the free electrons, the number of holes are comparatively less due to the presence of donated electrons, i.e $n_e >> n_h$.

Therefore, major conduction in n-type semiconductors is due to electrons. So, electrons are known as majority carriers and the holes are known as minority carriers.

This means $n_e >> n_h$; $I_e >> I_h$



e^- unbonded free electron donated by Pentavalent (+5 valency) atom.

Fig (a) :- Pentavalent donar atom (AS, Sb, P, etc) doped for tetravalent Si or Ge giving n-type Semiconductor

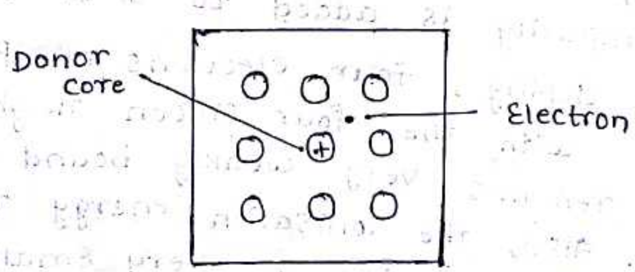


Fig (b) :- Commonly used Schematic representation for n-type material which shows only the fixed cores of the subsit.

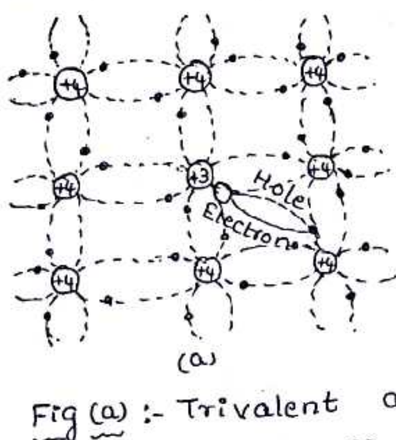
P - Type Semiconductors :-

This type of Semiconductor is obtained when a trivalent impurity is added to Si or Ge.

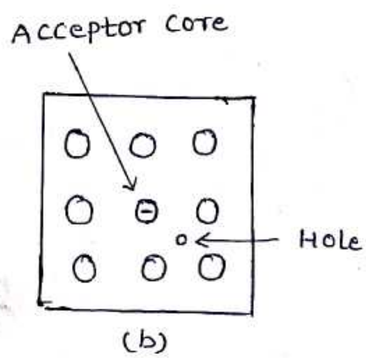
So, the three valence electrons of the doped impure atoms will form the covalent bonds with Silicon atoms but Silicon atoms have four electrons in its Valence Shell. Hence, one Covalent bond will be improper.

This means, one more electron is needed for the proper covalent bonding. This need of one electron is fulfilled from any of the bond between the Silicon and impurity atoms will be Completed. After bond formation, the doped' impurity will get ionised. As we know that, ions are negatively charged. So, the impurity will also get negative charge.

As, hole was created when the electron Come from Silicon - Silicon bond moved to complete the bond between the doped impurity and Silicon. Due to this, an electron will now move from any one of the covalent bond to fill the empty hole. This will further result in a new hole formation.



(a)



(b)

Fig (a) :- Trivalent acceptor atom (In, Al, B, etc.) doped in tetravalent Si or Ge lattice giving P-type Semiconductor. and Fig (b) :- Commonly used Schematic representation of P-type material which Shows only the fixed core of the Substituent acceptor with one effective additional negative Charge and its associated hole.

So, in p-type Semiconductor, the holes movement results in the formation of the current. This means, in this type of Semiconductor majority Charge carriers are electrons, i.e. $n_h >> n_e$; $I_h >> I_e$. Hence, these Conductors are known as p-type Semiconductors or acceptor type Semiconductors.

The electron and hole concentration in a Semiconductor in thermal equilibrium is given by $n_e n_h = n_i^2$

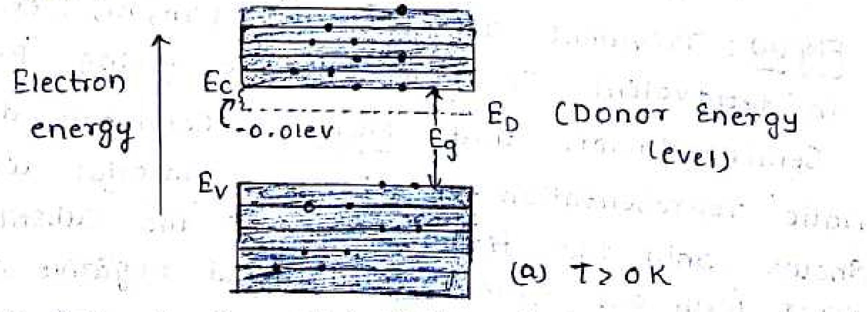
Note :- The energy gaps of C, Si and Ge are 5.4 ev, 1.1 ev and 0.7 ev respectively.

Sn is a group IV element as its energy gaps is zero.

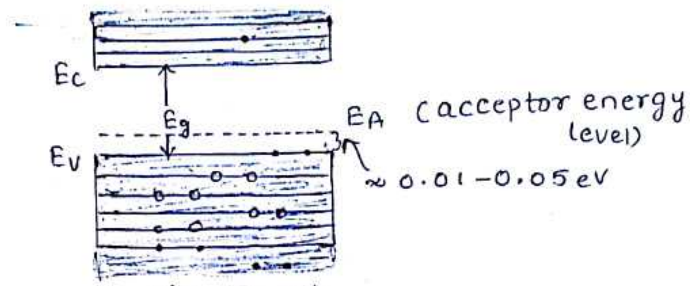
* Energy Band in Extrinsic Semiconductors :-

In extrinsic Semiconductors, additional energy States due to donor impurities (E_D) and acceptor impurities (E_A) also exist. In the energy band diagram of n-type Semiconductor, the donor energy level E_D is slightly below the bottom E_C of Conduction band and the electrons from this level move into conduction band with very small supply of energy.

In p-type Semiconductors, the acceptor energy level E_A is slightly above the top energy level E_V of the valence band. with very small supply of energy an electron from the valence band can jump to the level E_A and ionise the acceptor negatively.



one thermally generated electron-hole pair +9 electrons from donor atoms.



(b) T > 0K

Energy bands of (a) n-type semiconductor at T > 0K, (b) p-type semiconductor at T > 0K.

* **Ex-1:-** The number of silicon atoms per m^3 is 5×10^{28}. This is doped simultaneously with 5×10^{22} atoms per m^3 of arsenic and 5×10^{20} atoms per m^3 of indium. Calculate the number of electrons and holes. Given that $n_i = 1.5\times10^{16}\,m^{-3}$. Is the material n-type or p-type?

Sol For each atom doped with arsenic, one free electron is received. Similarly, for each atom doped of indium, ~~Calculate~~ a vacancy is created.

So, number of free electrons introduced by pentavalent impurity is

$$N_{AS} = 5\times10^{22}\,m^{-3}$$

The number of holes introduced by trivalent impurity added is $N_1 = 5\times10^{20}\,m^{-3}$

So, net number of electrons added is

$$n_e = N_{AS} - N_1 = 5\times10^{22} - 5\times10^{20}$$
$$= 4.95\times10^{22}\,m^{-3}$$

We know that, $n_e n_h = n_i^2$

$$\text{So, } n_h = \frac{n_i^2}{n_e} = \frac{(1.5\times10^{16})^2}{4.95\times10^{22}} = 4.54\times10^{9}\,m^{-3}$$

As, $n_e > n_h$ (number of holes). So, the material is n-type semiconductor.

* **P-n Junction :-**

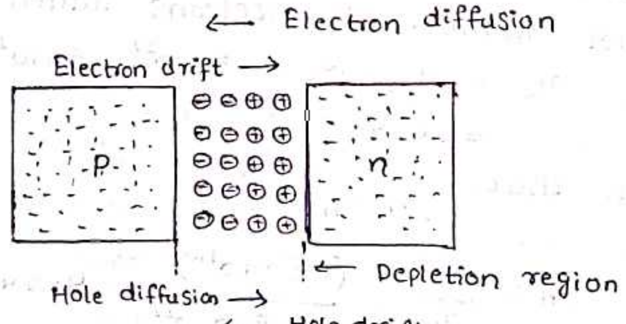
It is an arrangement made by a close contact of n-type semiconductor and p-type semiconductor.

There are various methods of forming p-n junction. In one method, an n-type germanium crystal is cut into thin slices called wafers. An aluminium film is laid on an n-type wafer, which is then heated is an oven at temperature of about 600°C. Aluminium then diffuses into the surface of water. In this way p-n junction is formed.

* Formation of Depletion Region in p-n Junction :-

In an n-type semiconductor, the concentration of electrons is more than that of holes. Similarly, in a p-type semiconductor, the concentration of holes is more than that of electrons. During the formation of p-n junction and due to the concentration gradient across p and n-sides, holes diffuse from p-side to n-side ($p \rightarrow n$) and electrons diffuse from n-side to p-side ($n \rightarrow p$). The diffused charge carriers combine with their counterparts in the immediate vicinity of the junction and neutralise each other.

Thus, near the junction positive charge is built on n-side and negative charge on p-side.



p-n junction formation Process

This sets up potential difference across the junction and an internal electric field E_i directed from n-side to p-side. The equilibrium is established when the field E_i becomes strong enough to stop further diffusion of the majority charge carriers (however, it helps the minority charge carriers to diffuse across the junction).

The region on either side of the junction which becomes depleted (free) from the mobile charge carriers is called depletion region or depletion layer. The width of depletion region is of the order of $10^{-6} m$.

The potential difference developed across the depletion region is called the potential barrier. It depends on dopant concentration in the semiconductor and temperature of the junction.

* **Note :-**

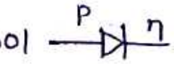
(1) Due to the diffusion of holes from p-side to n-side and electrons from n-side to p-side at the junction, a current rises from p-side to n-side which is called diffusion current.

(2) If an electron - hole pair is created on the depletion region due to thermal collision, the electrons are pushed by the electric field towards the n-side and the holes towards the p-side, which gives rise to a current from n-side to p-side known as drift current.

(3) In steady state, diffusion current = drift current.

* **Semiconductor Diode or p-n Junction diode :-**

It is basically a p-n junction with metallic contacts provided at the ends for the application of an external voltage. It is a two terminal device.

It is represented by the symbol 

The direction of arrow indicates the conventional direction of current.

* **Forward Biasing and Reverse Biasing of Junction Diode :-**

Biasing is the method of connecting external battery or emf source to a p-n junction

Scanned by CamScanner

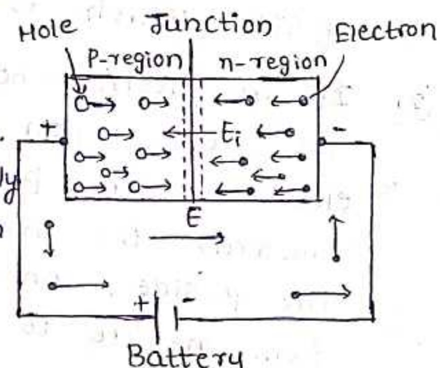
diode. The junction diode can be connected to an external battery in two ways, called forward biasing and reverse biasing of the junction.

Forward Biasing :-

A junction diode is said to be forward biased when the positive terminal of the external battery is connected to the p-side and negative terminal to the n-side of the diode.

Flow of Current in Forward Biasing :-

In this situation, the forward voltage opposes the potential barrier, due to which both the potential barrier and width of the depletion layer decreases. Under the effect of external electric field, holes in the p-region and electrons in the n-region, both move towards the junction. These holes and electrons mutually combine just near the junction and cease to exist. For each electron-hole combination, a covalent bond breaks up in the p-region near the positive terminal of the battery. Out of the hole and electron so produced, the hole moves towards the junction, while the electron enters the positive terminal of the battery through the connecting wire.



[Forward biasing of junction diode]

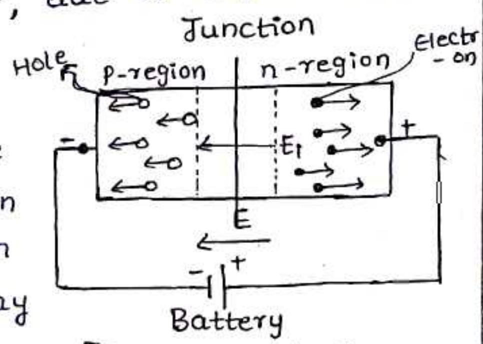
Just at this moment, an electron is released from the negative terminal of the battery which enters the n-region to replace the electron lost by combining with a hole at the junction. Thus, a current called forward current, is constituted by the motion of majority charge carriers across the junction. In forward bias, the junction diode offers low resistance.

* ## Reverse Biasing :-

A junction diode is said to be reverse biased when the positive terminal of the external battery is connected to the n-side and negative terminal to the p-side of the diode.

* ## Flow of current in Reverse Biasing :-

In this situation, the reverse voltage supports the potential barrier, due to which both the potential barrier and width of the depletion layer increases. Under the effect of external electric field, holes in the p-region and electrons in the n-region are pushed away from the junction i.e they cannot be combined at the junction. So, there is almost no flow of current due to majority charge carriers.



[Reverse biasing of junction diode]

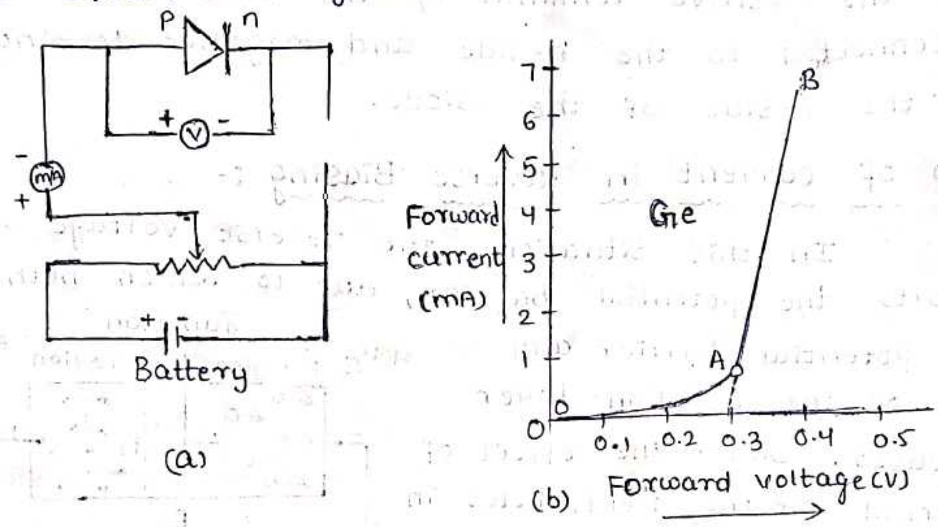
However, a very small current due to minority charge carriers, flow across the junction. This current is called reverse current.

* ## I-V (current - voltage) Characteristics of p-n Junction Diode :-

The graphical relations between voltage applied across p-n junction and current flowing through the junction are called I-V characteristics of junction diode.

Forward Biased Characteristics :-

The circuit diagram for studying forward biased characteristics is shown in the fig (a). Starting from a low value, forward bias voltage is increased step by step (measured by voltmeter) and forward current is noted (by ammeter). A graph is plotted between voltage and current is shown in fig (b).



Forward biased characteristic of a diode

At the start when applied voltage is low, the current through the diode is almost zero. It is because of the potential barrier, which opposes the applied voltage.

Till the applied voltage exceeds the potential barrier, the current increases very slowly with increase in applied voltage (OA portion of the graph).

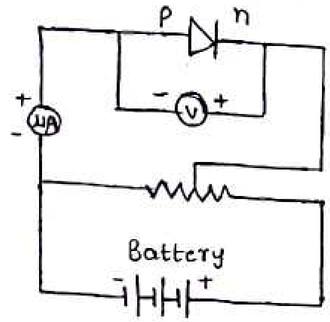
With further increase in applied voltage, the current increases very rapidly (AB portion of the graph), in this situation the diode behaves like a conductor. The forward voltage beyond which the current through the junction starts increasing rapidly with voltage is called knee voltage or treshold voltage.

If line AB is extended back, it cuts the voltage axis at potential barrier voltage.

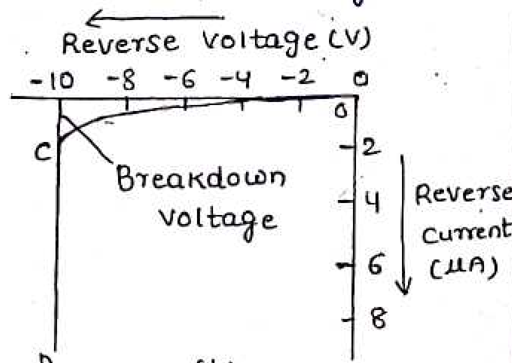
Scanned by CamScanner

* Reverse Biased Characteristics :-

The circuit diagram for studying reverse biased characteristics is shown in the figure (a).



(a)



(b)

Reverse biased characteristic of a diode.

In reverse biased, the applied voltage supports the flow of minority charge carriers across the junction. So, a very small current flows across the junction due to minority charge carriers. Motion of minority charge carriers is also supported by internal potential barrier, so all the minority carriers cross over the junction.

Therefore, the small reverse current remains almost constant over a sufficiently long range of reverse bias, increasing very little with increasing voltage (OC portion of the graph). This reverse current is voltage independent upto certain voltage known as breakdown voltage and this voltage independent current is called reverse saturation current.

Note :- If the reverse bias is equal to the breakdown voltage, then the reverse current through the junction increases very rapidly (CD portion of the graph), this situation is called avalanche breakdown and the junction may get damaged due to excessive heating if this current exceeds the rated value of p-n junction.

In diodes, a resistance is offered by the function which depends on the applied voltage, which is called dynamic resistance. It is the ratio of small change in voltage to the small change in current produced.

$$\text{Dynamic resistance, } r_d = \frac{\Delta V}{\Delta I}$$

* Diode as a Rectifier :-

The process of converting alternating voltage/ current into direct voltage / current is called rectification. Diode is used as a rectifier for converting alternating current / voltage into direct current / voltage.

Principle :-

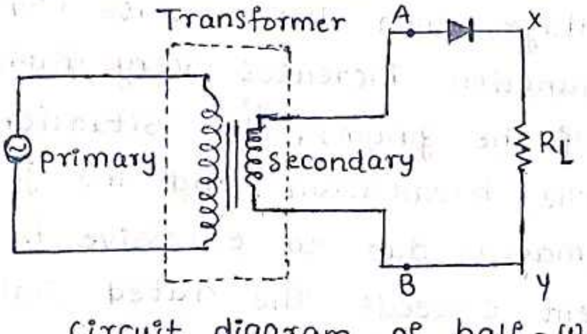
From the V-I characteristic of a junction diode, we see that it allows current to pass only when it is forward biased. So, if an alternating voltage is applied across a diode, the current flows only in that part of the cycle when the diode is forward biased. This property is used to rectify the current/ voltage.

There are two ways of using a diode as a rectifier, i.e.

(i) Diode as a half-wave rectifier.

(ii) Diode as a full wave rectifier.

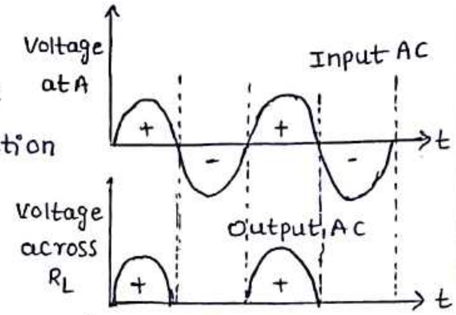
Diode as a Half-Wave Rectifier :-



Circuit diagram of half-wave rectifier

In this, the AC voltage to be rectified is Connected to the primary coil of a Step-down transformer and Secondary Coil is connected to the diode through resistor R_L across which, output is Obtained.

Working :-



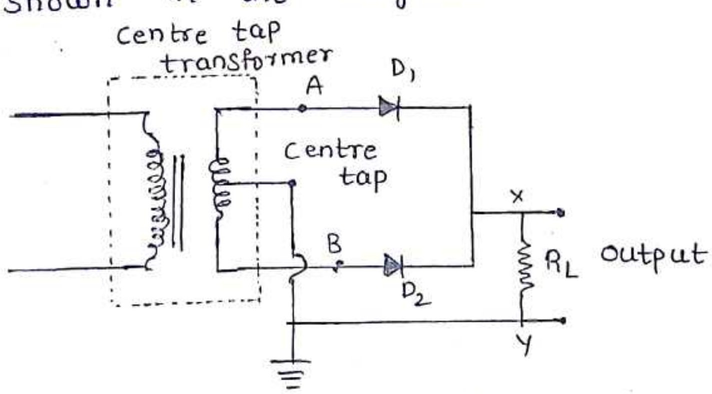
Input and output waveforms.

During positive half cycle of the input AC, the p-n junction is forward biased. Thus, the resistance in p-n junction becomes low and current flows. Hence we get output in the load.

During negative half cycle of the input AC, the p-n junction is reverse biased. Thus, the resistance of p-n junction is high and current does not flow. Hence, no output is in the load.

* **Diode as a full wave Rectifier :-**

In the full wave rectifier, two p-n junction diodes D_1 and D_2 are used. This arrangement is Shown in the diagram below.

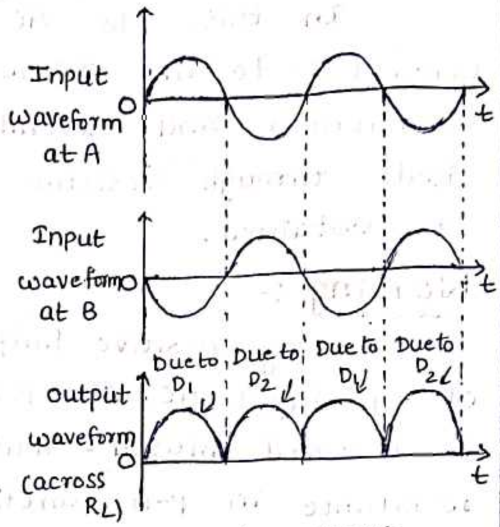


Circuit diagram of full wave rectifier.

Working :-

During the positive half cycle of the input AC, the diode D_1 is forward biased and the diode D_2 is reverse biased. The forward current flows through diode D_1.

During the negative half cycle of the input AC, the diode D_1 is reverse biased and diode D_2 is forward biased. Hence, current flows through diode D_2. Hence, we find that during both the halves, current flows in the same direction.

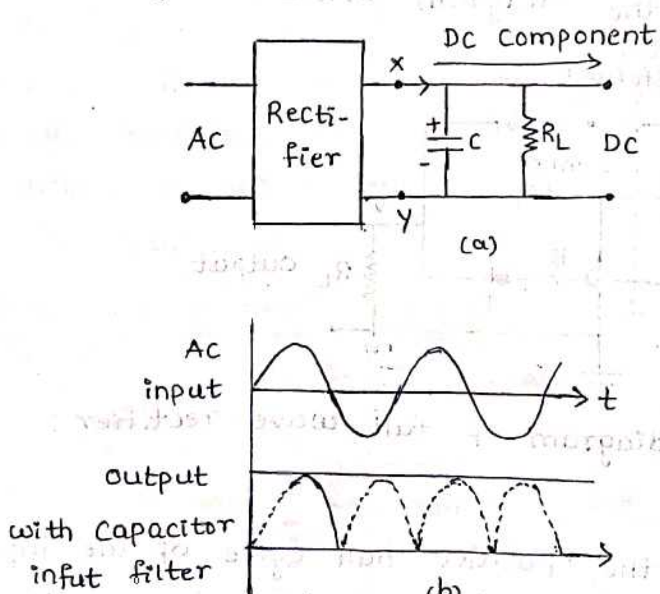


[Input and output waveforms].

Role of Filters :-

In order to get the steady DC output from the pulsating voltage normally, a capacitor is connected across the output terminals (parallel to load R_L). An inductor can also be used in series for the same purpose.

As these additional circuits appear to filter out the AC ripple and provide a pure DC voltage, so they are called filters.



Fig(a) :- A full wave rectifier with capacitor filter and Fig(b) :- input and output voltage of rectifier.

Let us discuss the role of capacitor in filtering. when the voltage across the capacitor is rising, it gets charged. If there is no external load, it remains charged to the peak voltage of the rectified output. When there is a load, it gets discharged through the load and the voltage across it begins to fall. In the next half cycle of the rectified output, it again gets charged to the peak value (see the above figure).

The rate of fall of voltage across the capacitor depends upon the inverse product of capacitor C and the effective resistance R_L used in the circuit and is known as time constant. To make the time constant large value of C should be large. So, capacitor input filters use large capacitor input filter is nearer to the peak voltage of the rectified voltage.

* Special purpose P-n Junction Diodes

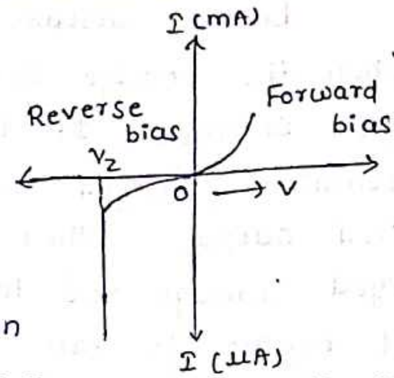
Zener Diode :-

It is a reverse biased heavily doped p-n junction diode. It is designed to operate in the reverse breakdown voltage continuously without being damaged.

This can be achieved by changing the thickness of the depletion layer to which the voltage is applied. Current through this diode is controlled by an external resistance. It is represented by the symbol.

Symbol of zener diode.

* V-I Characteristics :-



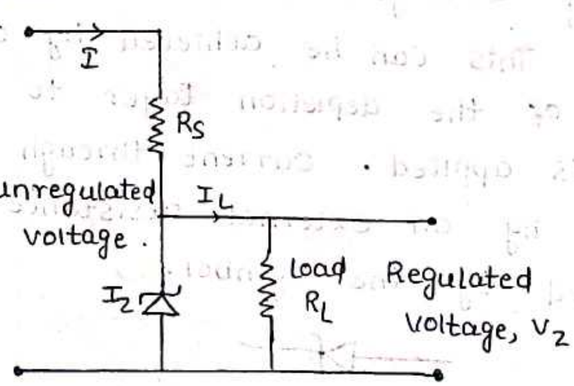
[V-I characteristics of a zener diode]

The V-I characteristics of zener diode is shown in the figure. Here, we observe that when the applied reverse voltage (V) reaches the breakdown voltage (V_z) of the zener diode, there is a large change in the current. But after the breakdown voltage V_z, a large change in the current can be produced by almost insignificant change in the reverse bias voltage.

Zener Diode as a Voltage Regulator :-

This is the most important application of a zener diode.

Principle :- From the above V-I characteristic of zener diode, we can say that, zener voltage remains constant even though the current through the diode varies over a wide range. If a zener diode is joined in reverse bias to the fluctuating DC input voltage through a resistance R then, the constant output voltage is taken across a load resistance connected in parallel with zener diode.



Circuit diagram of zener diode as voltage regulator.

Working :- Here, when input DC voltage increases beyond a certain limit, the current through the circuit sharply, causing a sufficient increase in the voltage drop across the resistor R_S. Thus, the voltage across the zener diode remains constant and also the output voltage remains constant at V_Z.

When the input DC voltage decreases, the current through the circuit goes down sharply causing sufficient decrease in the voltage drop across the resistance. Thus, the voltage across the zener diode remains constant and also the output voltage across R_L remains constant at V_Z.

Hence, the output voltage remains constant in both conditions.

* **Optoelectronic junction devices :-**

Semiconductor diodes which generate charge carries by photons are called optoelectronic devices.

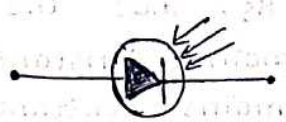
Optoelectronics is the technology that combines optics and electronics. This includes many devices based on the function of a p-n junction diode.

Some of these are.

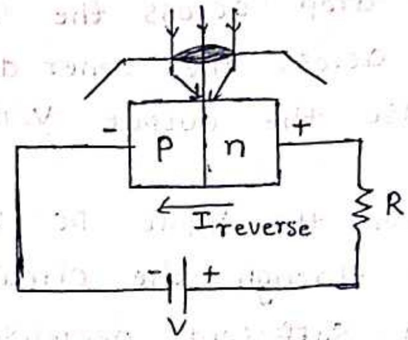
* Light emitting diodes (LED) which converts electric energy into light.
* Photo diodes used for detecting optical signals.
* Photovoltaic devices (Solar cells) which convert light into electrical energy.

* **Photodiode :-**

A Photodiode is a special purpose photosensitive semiconducting diode operated in the reverse biased condition. When light is incident on the junction, the junction resistance decreases rapidly with the increase in the intensity of light.



Symbol of a photo diode



Basic biasing arrangement of Photodiode.

A Photodiode is fabricated differently when compared to the conventional diode. A photodiode is designed to have a large depletion region, very often with a layer of an intrinsic (undoped) semi-conductor material within the depletion region. A diode of this type is often called a pin diode (for p, intrinsic semiconductor, n). The reason for the large depletion region or the intrinsic semi-conductor is to provide a large number of electrons when light is incident.

In the case of a photodiode in reverse bias, there is a small leakage current due to the minority charge caries, as in the case of a normal diode (rectifier diode).

In addition, when light is incident on a photodiode, a current is caused by the release of charge carries in the junction. The large increase in the current from the non-lit to lit condition (non-illuminated to illuminated condition) means that there is a sharp fall in its resistance. Typically, the dark current, namely the current in the non-lit condition is about 5 nA, whereas the light current namely, the current in the lit condition is about 15 μA. Thus increases by a factor of 3000 which means resistance has droped by a factor of 3000.

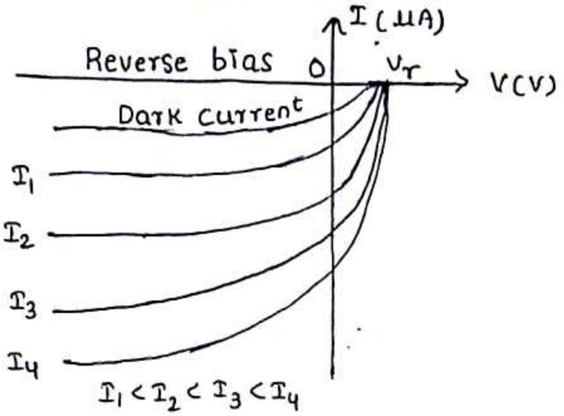


Fig :- Characteristics of a photodiode.

The change in resistance depends on the intensity of incident light, its wavelength and the reverse bias voltage.

Solar cells :-

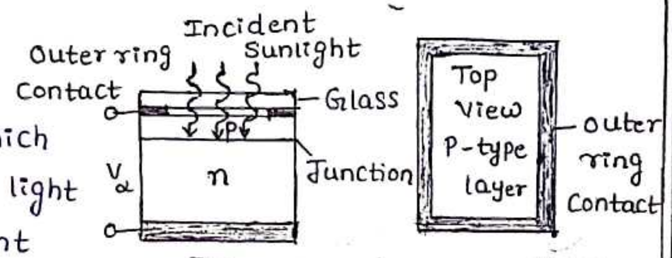


Fig :- p-n junction Solar cell.

A Solar cell is a p-n junction which generates emf when light radiation is incident on it. The basic construction of a Silicon p-n junction Solar cell is Shown in the Fig.

The top p-type material is made very thin So that the photons reach the junction. The contact to the p-type material is made through an outer ring deposited on it while a metallic contact is established to the bottom of the n-type material. A protective glass covering is provided for the top Surface.

When Solid radiation falls on the top Surface and reach the junction, the photon energy will be imparted to the valence electron in the atoms near the junction. Electrons are liberated from the atom and the result is a generation of electron-hole pairs. This phenomenon will occur on each Side of the junction.

In the p-type material, the newly generated electrons are minority carriers and will move rather freely across the junction to n-type material because the polarity of the junction potential favour this. Similarly, the newly generated holes in the n-type material which are minority carriers are swept across the junction to the p-type material by the potential. The result is an increase in the minority carrier flow, which is opposite in direction to the conventional forward current of a p-n junction. This increase in reverse current is shown in Fig for different illuminations.

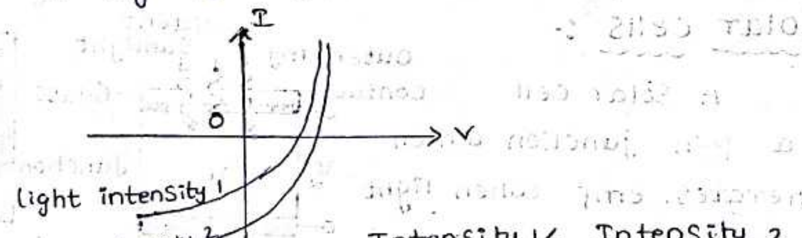


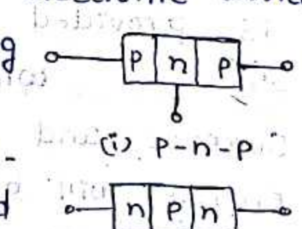
Fig. output characteristics of Si or Ge

* Uses of Solar Cells :-

(i) Used to charge storage batteries in day time.
(ii) Used in calculators for power supply unit.
(iii) Used to supply power in artificial satellites and space vehicles.

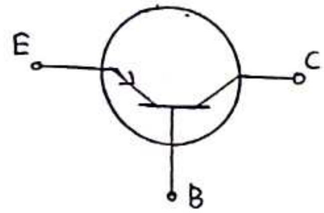
* Transistor :-

Transistor is the modern electronic device which is capable of doing anything in the electronics. Transistor can does, rectification, amplification, oscillation etc. Transistor was discovered by J. Bardeen and W.H Brattain of USA in 1948. A transistor consists of a p-type crystal sandwiching between two n-type crystals or a n-type crystal sandwiching between two

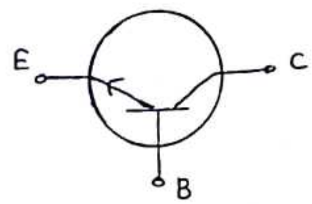


p- type crystals. Accordingly these are two types of transistors : p-n-p and n-p-n. These are :

* **Transistor Symbols :-**



(i) Symbol of p-n-p transistor.



(ii) Symbol of n-p-n transistor.

* **Parts of a transistor :-**

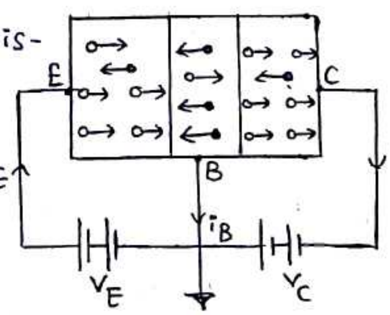
(i) Emitter :- It is the heavily doped crystal which supplies the majority charge carriers. It is made thinner than collector and thicker than base

(ii) Base :- The base is the lightly doped crystal and made very thin. It passes most of the emitter injected charge carriers to the collector.

(iii) Collector :- It is the moderately doped crystal and made thickest one. Its function is to remove charges from its junction with the base. For the sake of convenience, it is customary to show emitter and collector to be of equal size.

* **Working of p-n-p transistor :-**

Consider a p-n-p transistor connected in the circuit in such a way that base is common with emitter and collector. Thus there are two loops in the circuit. The



emitter- base junction is given small forward bias, while collector - base junction is given large reverse bias. Because of the forward bias of emitter - base region, holes move from emitter towards base. Most of them (nearly 98%) cross base and enter into collector region, while very few (nearly 2%) combine with the electrons in base. As soon as a hole combines with the electron, an electron leaves the negative terminal of battery V_E and enters into base. This causes a small base current i_B. The holes entering into the collector region combine with the electrons coming from the negative terminal of V_C. This causes collector current i_C. Both the currents together constitute emitter current i_E.

Thus $i_E = i_B + i_C$

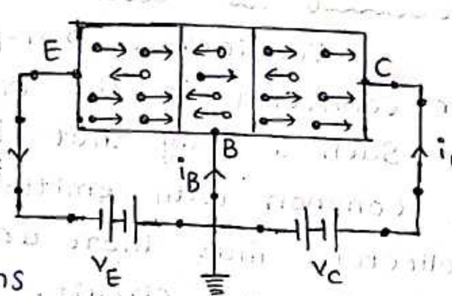
It should be remembered that:

(i) The base current may be nearly 2 to 10% of the emitter current depending on the dopping level. Similarly collector current may be nearly 90 to 98% of the emitter current.

(ii) The holes are the charge carriers within the transistor while electrons are charge carriers in external circuit.

* Working of n-p-n transistor :-

Consider a n-p-n transistor connected in a circuit in such a way that base is common with emitter and collector. The emitter-base junction is given a small forward bias V_E while collection base junction is given large reverse bias.



Because of forward bias in emitter - base region, electrons from emitter region move towards base. Most of them (nearly 98%) cross the base and enter into collector region while very few (nearly 2%) combine with the holes in base.

As soon as an electron combines with the hole, a bond break in base region and produces a pair of hole and electron.

This electron is captured by positive terminal of the battery V_E and send it towards emitter region. This causes a small base current i_B. The electrons entering into collector region are attracted by the positive terminal of V_C. This causes collector current i_C. These two currents combine together constituent emitter current i_E. Thus $i_E = i_B + i_C$

It should be remembered that

(i) The base current may be 2 to 10% of the emitter current depending on the doping level. Similarly collector current may be 90 to 98% of the emitter current.

(ii) The electrons are the charge carriers within the transistor as well as in external circuit.

* **Basic logic gates :-**

The three basic laws of Boolean algebra are (1) AND (2) OR and (3) NOT laws. Correspondingly we have three basic logic gates.

Logic 1 and logic 0 :-

The two possible outputs of a logic circuit are usually represented by the. Symbols 1 and 0. These logic symbols have no numerical values. They have only logical significance. They may represent either YES or NO, TRUE or FALSE etc... In a positive logic system, a high output state of the circuit is represented by logic 1 and the low output state by logic 0.

Truth table :-

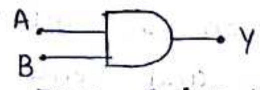
A table giving the output for different combination of the input is called the truth table of the logic circuit.

AND Gate :-

Table 1 :- 2-input AND operation.

A	B	$Y = A \cdot B$
0	0	0
0	1	0
1	0	0
1	1	1

A logic circuit which has a high output only when all the inputs are high is called an AND gate is low when any one input is low. The logic equation for a two input AND gate is $Y = A \cdot B$. It can also be represented as $Y = AB$



Fig(1) :- 2-input AND gate Symbol.

The Symbol of AND gate and its truth table are as shown in fig(1) and Table 1.

Example (1) - Switch circuit

The bulb will glow (logic 1) only when both K_1 and K_2 are closed (logic 1). In the circuit (Fig(2)) the bulb glows (logic 1) only when both the switches are closed (logic 1 state). whether one switch is open (logic 0), the bulb does not glow (logic 0).

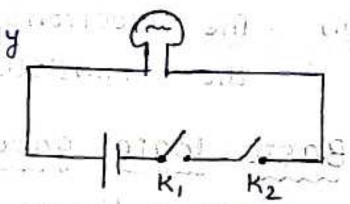


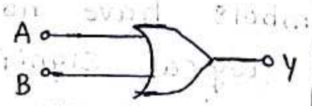
Fig 2 :- Switch circuit for AND gate

OR gate :-

Table 2 :- input OR operation

A	B	$Y = A + B$
0	0	0
1	0	1
0	1	1
1	1	1

A logic gate whose output is high (logic 1) when any one of the inputs is high is called an OR gate. Its output is low when all the inputs are low. The logic equation for a two input OR gate is $Y = A + B$.



Fig(a) :- 2-input OR gate Symbol

The symbol of OR gate and truth table are as in Fig(a) and Table 2.

Example :- Switch Circuit

In the circuit fig (b) when any one of the keys is closed (logic 1) or both the keys are closed, the bulb glows (logic 1). when both the keys are open (logic 0) the bulb does not glow (logic 0).

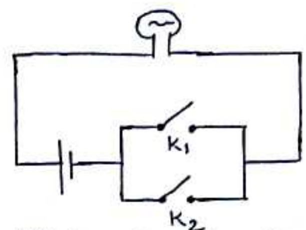
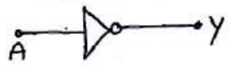


Fig b :- Switch circuit for OR gate

NOT gate :-

A logic gate whose output is the complement of the input is called a NOT gate. If the input is 1 the output is 0 and vice-versa. NOT gate is also called an inverter. The logic equation for NOT gate is $Y = \bar{A}$. The symbol of NOT gate and its truth table are as in fig(1) and Table 3.



Fig(1):- NOT gate symbol

Table 3: NOT operation

A	$Y=\bar{A}$
1	0
0	1

Example :-

In the circuit Fig(2), when the key K is closed (logic 1) the bulb is off (logic 0). when the key is open (logic 0) the bulb is on (logic 1).

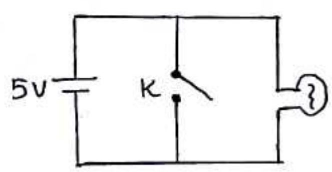


Fig 2 :- Switch circuit for NOT operation

NAND gate :-

A logic gate whose output is low (0) when all of the inputs are high is called a NAND gate. Its output is high when any one of the inputs is low or all the inputs are low. NAND gate may be treated as an AND gate followed by a NOT gate. The logic equation for a two input NAND gate is $Y = \overline{A \cdot B}$. The symbol of NAND gate and its truth table are as in fig and Table 4.

Table-4 :- NAND operation

A	B	$Y=\overline{AB}$
0	0	1
1	0	1
0	1	1
1	1	0

Fig :- NAND gate Symbol

NOR gate :-

A logic gate whose output is high when all the inputs are low is called a NOR gate. Its output is low when one of the inputs is high, or all the inputs are high. NOR gate may be treated as an OR gate followed by a NOT gate. The logic equation for a two input NOR gate is $Y = \overline{A+B}$. The Symbol of NOR gate and its truth table are as shown in fig and Table 5:

Table 5: NOR operation

A	B	$Y = \overline{A+B}$
0	0	1
1	0	0
0	1	0
1	1	0

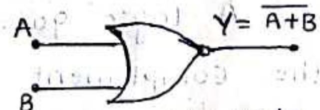


Fig:- NOR gate Symbol

→ NOR and NAND gates are universal gates.