Annamacharya Institute of Technology and Sciences, Tirupati
(Autonomous)

Applied Physics Material

Subject Code: 20ABS9902
(AK20 Regulation)
(Common to EEE, ECE, CSE, AI&DS, AI&ML and CIC)

Course Outcomes

1. Analyze the wave properties of light and the interaction of energy with the matter.
2. Apply electromagnetic wave propagation in different guided media.
3. Asses the electromagnetic wave propagation and its power in different media
4. Analyze the conductivity of semiconductors.
5. Interpret the difference between normal conductor and superconductor and apply the nanomaterials for engineering applications.

Unit-I
Optics

Introduction:

The English word optics comes from the Greek word ***“Opticks (Οπτικη)”*** which means look (or) appearance.

Optics is the branch of the physics, which deals with nature, properties, generation, and propagation of light and its interaction with matter.

It is classified into three sections.

Geometrical optics or Ray optics:

In this section, it is assumed that light travels in straight lines and we explain the images are formed by mirrors, prism, lenses etc., by geometrical constructions, without considering the nature of light.

Physical optics or wave optics:

In this section, the nature of light considered and we explain various phenomena like reflection, refraction, interference, diffraction and polarization of light etc., by assuming light as a wave form.

Quantum optics:

In this section, the interaction of light with atomic particles is studied and we explain various phenomena like photo electric effect, Compton Effect and atomic excitation, etc.

Electronic optics is applied to study the stream of particles in electronic microscope, cathode ray tubes etc.

Fiber optics deals with the transmission of light through transparent fibers, using the principle of total internal reflection.

UNIT-1 -CHAPTER-1
INTERFERENCE

Introduction:

Is it possible that light added to light produces darkness? Yes, when the two light waves are out of phase.

Light + Light = Brighter light, if two light waves are in phase .This phenomenon is called interference, which is based on the principle of superposition.

Definition of Interference:

The formation of bright and dark fringes on superimposing two coherent light waves is called interference.

Principle of superposition:

When two or more waves meet, they interfere and produce a resultant wave whose properties can be calculated by using the principle of superposition

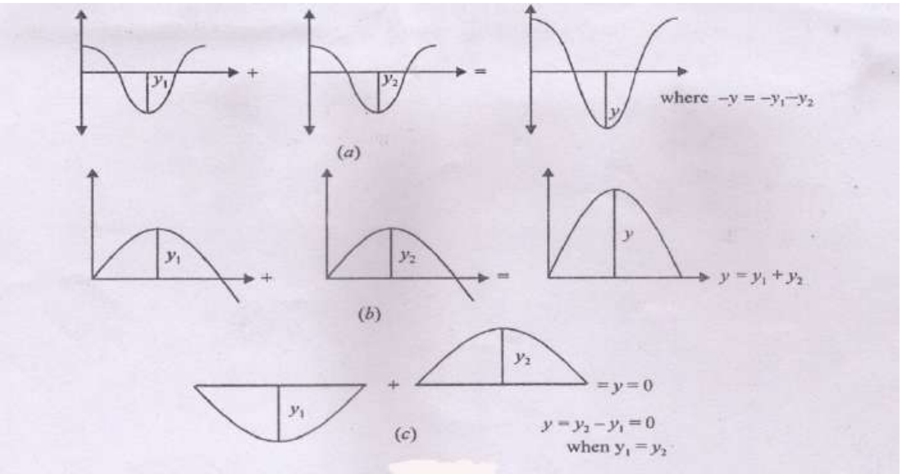


Fig: Superposition of two waves

The principle of superposition states that whenever two or more waves travelling in the same region superpose, the total displacement(y) at any point is equal to the vector sum of their individual displacements(y_1,y_2,y_3…..) at that point .i.e,

$$\mathbf{y = y_1 \pm y_2 \pm y_3 + \ldots..}$$

If there only two waves, then the resultant displacement is given by;

$$\mathbf{y = y_1 \pm y_2}$$

Case1:

- If the two waves are in phase with each other, they undergo constructive interference producing maximum intensity of light called bright band or fringe.

$y = y_1 + y_2$

$A_{max} = a_1 + a_2$ (since $I \alpha A^2$)

$I_{max} =$ ¿)

- The condition for constructive interference to occur is

Path difference $= n\lambda$ (or) phase difference $= 2n\pi$

Case: 2

- If the two waves are out of phase, then they undergo destructive interference producing zero intensity of light called dark band or fringe.

$$y = y_1 - y_2$$

$$A_{min} = a_1 - a_2$$

$$I_{min} = (a_1^2 - a_2^2 ¿$$

- The condition for destructive interference to occur is

$$\text{Path difference} = (2n-1)\frac{\lambda}{2} \quad \text{(or)} \quad \text{phase difference} = (2n-1)\pi$$

Case3:

- If the two waves are in phase with each other, they undergo constructive interference producing maximum intensity of light called bright band or fringe.
- The condition for constructive interference to occur is

Path difference$\Delta = n\lambda$ (or)
Phase difference$\varnothing = 2n\pi$

Case 4:

- If the two waves are out of phase, then they undergo destructive interference producing zero intensity of light called dark band or fringe.
- The condition for destructive interference to occur is

Path difference$\Delta = (2n-1)\frac{\lambda}{2}$ (or)
Phase difference$\varnothing = (2n-1)\pi$

Coherence:

The two light waves are said to be coherent if they have same frequency (wavelength) and maintaining same (or) constant phase relationship.
Coherence is two types. (1)Temporal coherence (2) Spatial coherence.

Conditions for interference:

1. To produce interference, we require two coherent light sources.
2. The two sources must be monochromatic.
3. The two sources must be narrow and parallel.
4. The two sources must emit light in same direction.
5. The distance between the two sources must be small.
6. The distance between the two coherent sources and the screen must be large.
7. Background should be dark to view interference fringes.
8. The amplitudes of the light should be equal or nearly equal.
9. The two sourcs of light must emit contineous light waves of same wavelength and frequency.

Interference in uniform thin films by reflection of light:

Plane parallel thin film:

A transparent thin film of uniform thickness which is bounded by two plane parallel surfaces is known as plane parallel thin film.

Principle:

When light is incident on plane parallel thin film, some portion of light gets reflected from the top surface and the reaming portion is transmitted into the film. Again, some portion of the transmitted light is reflected back into the film by the bottom surface and emerges through the top surface. These reflected light waves superimpose with each other, producing interference and forming interference patterns. This is the principle of interference in thin films by reflection.

Theory:

- Let us consider a transparent film of uniform thickness ‘t’ bounded by two parallel surfaces, as shown in fig.
- Let the refractive index of the material beμ.The film is surrounded by a common medium say, air on both sides.
- Let a monochromatic light ray AB is incident on the top surface at an angle ***‘i’***.
- The ray AB is partly reflected along BC and is partly transmitted into the film along BF makes an angle ***r*** with the normal FG and meets the lower surface.
- At F, it is again partly reflected back into the film along FD, while major portion refracts into the outer medium along FK.
- The reflected ray FD refracts at the outer surface and emerges out along DE which is parallel to the ray BC.
- The waves travelling along directions BC and BFDE are obtained from the incident wave AB.

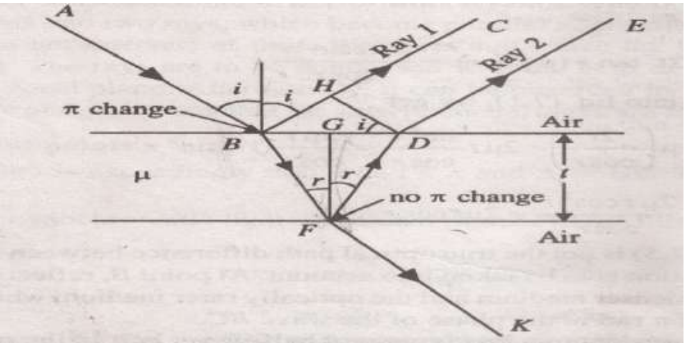


Fig: Interference in thin films by reflection of light

- Therefore, these two light rays superimpose and produce interference patterns. Condition of interference depends on the optical path difference between the rays 1 and 2.
- Let us now calculate the optical path difference between the reflected ray BC (ray1) and the refracted ray BFDE (ray 2).
- A normal DH is drawn on line BC. From points H and D onwards the rays HC and DE equal distances.
- The reflected ray BC travels in air while the refracted ray (BF+FD) travels in the film of refractive indexμ.
- The Optical path difference (OPD**)** between the rays (1) and (2) is

Total path difference(T.P.D) = (μX geometrical path BFD in film) - (μ X geometrical path BH in air)

$$= \mu\,(BF+FD) - \mu_{air}\,(BH) \qquad [\text{Since } \mu_{air} = 1]$$

$$= \mu\,(BF+FD) - BH \text{ ---} \quad (1)$$

Step 1: Calculation of geometrical path BFD in film

Let us calculate path BF+ FD in film,

From fig; ΔBFG, $\text{Cos } r = \frac{GF}{BF}$

$$BF = \frac{t}{\cos r} \text{ -------} \quad (2)$$

Similarly;

From ΔDFG, $\text{Cos } r = \frac{GF}{FD}$

$$FD = \frac{t}{\cos r} \text{-----} \quad (3)$$

$$\therefore BF+FD = \frac{t}{\cos r} + \frac{t}{\cos r}$$

$$\therefore BFD = BF+FD = \frac{2t}{\cos r} \text{ -----} \quad (4)$$

Step 2: Calculation of geometrical path BH in air

To calculate BH air, first BD which is equal to (BG+GD) has to be obtained.

From ΔBGF; $\tan r = \frac{BG}{FG} = \frac{BG}{t}$

$$BG = t.\tan r \text{-----} \quad (5)$$

Similarly;

FromΔDGF; $\tan r = \frac{GD}{FG} = \frac{GD}{t}$

$$GD = t.\tan r \text{-----} \quad (6)$$

$$\therefore BD = BG+GD = t.\tan r + t.\tan r$$

$$= 2t \tan r \text{-----} \quad (7)$$

From ΔBHD; $\text{Sin } i = \frac{BH}{BD}$

$$BH = BD \text{ Sin } i$$

$$= 2t \tan r \text{ Sin} i \text{----} \quad (8)$$

From Snell's law $\mu = \frac{Sini}{Sinr}$

$$\text{Sin } i = \mu \text{ Sin } r$$

From equation (8)

$$BH = 2t \tan r\ \mu \text{ Sin } r$$

$$= 2\mu t \tan r.\ \text{Sin } r$$

$=2\mu t\ \frac{\sin r}{\cos r}$. Sin r

$$\text{T.P.D} = 2\mu t\ \frac{\sin^2 r}{\cos r} \text{ ------ } (9)$$

Substituting equation (4) and equation (9) in equation (1); we have,

$$= \frac{2\mu t}{\cos r} - 2\mu t\ \frac{\sin^2 r}{\cos r} \text{ ----- } (10)$$

$$= \frac{2\mu t}{\cos r}(1 - \sin^2 r)$$

$$= \frac{2\mu t}{\cos r}\cos^2 r$$

$$\text{T.P.D} = 2\mu t \cos r \text{ ---- } (11)$$

This is called Cosine law.

According to stoke principle, if a light wave traveling from rarer medium to denser medium under goes a phase of π or path change $\frac{\lambda}{2}$ (i.e., the wave loses half of wavelength) when it gets reflected at the boundary of a rarer to denser medium.

$$\boxed{\text{T.P.D} = 2\mu t \cos r - \frac{\lambda}{2}} \text{ ---------- } (12)$$

Case 1: Condition for bright band (or) maximum (or) constructive interference:

Bright band occurs when the total path difference (T.P.D) $= n\lambda$ ---- (13)

From equations (12) and (13); we have

$$2\mu t \cos r - \frac{\lambda}{2} = n\lambda$$

$$2\mu t \cos r = n\lambda + \frac{\lambda}{2}$$

$$2\mu t \cos r = (2n+1)\frac{\lambda}{2} \text{ --- } (14)\text{, where } n = 0, 1, 2, 3\ldots$$

This is the condition for maxima. The film appears bright under this condition.

Case2: Condition for dark band (or) minimum (or) Destructive interference:

Dark band occurs when the path difference = $(2n-1)\frac{\lambda}{2}$ ----- (15)

From equations (12) and (15); we have

$$2\mu t \cos r - \frac{\lambda}{2} = (2n-1)\frac{\lambda}{2}$$

$$2\mu t\cos r \;=(2n-1)¿+\frac{\lambda}{2})$$

$$2\mu t\cos r \;=\; n\lambda \quad ------ \quad \text{(7), where n=0, 1, 2, 3…}$$

This is the condition for minima. The film appears dark under this condition.

Newton’s rings:

Newton’s rings are one of the best examples for the interference in a non-uniform thin film.

Definition:

If a Plano-convex lens with its convex surface is placed on a plane glass plate, then air film is formed between the lower surface of the Plano-convex lens and the upper surface of the glass plate. The thickness of the air film is zero at point of contact and gradually increases from the point of contact outer wards.

If a monochromatic light is allowed to fall normally on this air film, a system of alternate bright and dark circular rings is formed in the air film. These rings were discovered by Newton and are called Newton’s rings.

Principle:

Newton’s rings are formed due to interference between the light rays reflected from the top and bottom surfaces of air film between the glass plate and the Plano-convex lens.

Experimental Arrangement:

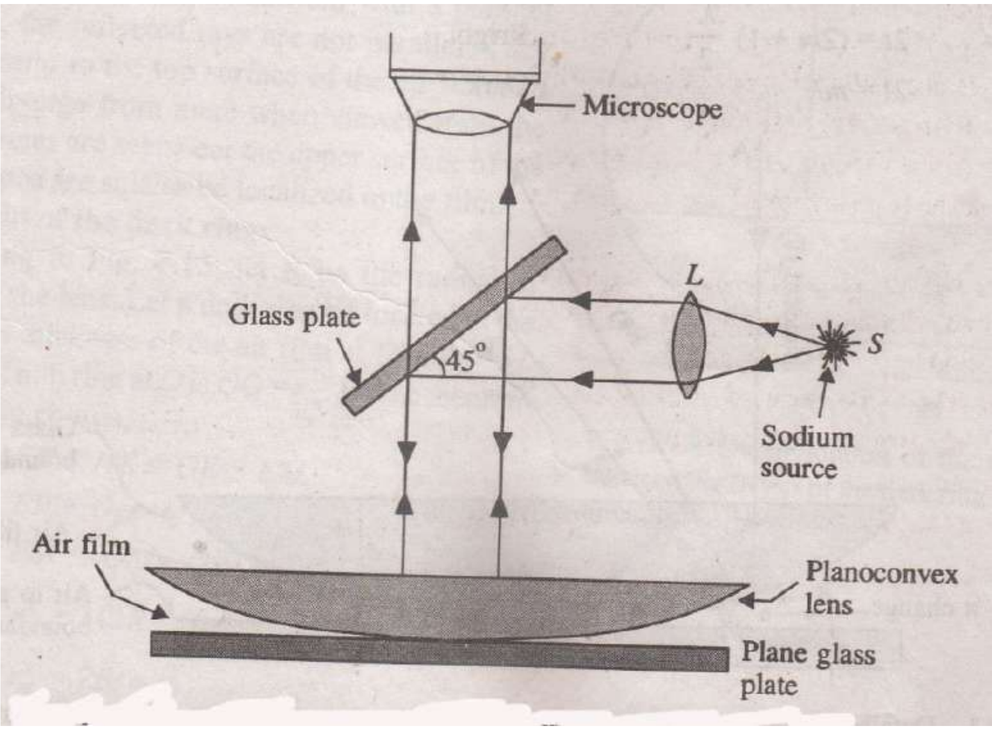


Figure: The experimental arrangement for producing Newton rings.

1. The experimental arrangement for producing Newton rings is as shown in fig (1).
2. Keep the convex surface of the Plano-convex lens over the plane glass plate and arrange glass plate G at an angle of 45^0 over the base set.
3. Switch on the monochromatic light source ‘S’ (Sodium vapor lamp) and it is focus on the Double convex lens (L). This sends parallel beam of light. This beam of light falls on the glass plate B at 45^0.
4. The glass plate ‘G’ reflects a part of light towards the air film enclosed by the Plano-convex lens and the plane glass plate.
5. A part of the light is reflected by the curved surface of the Plano-convex lens and a part is transmitted which is reflected back from the plane surface of the plane glass plate.
6. These reflected light rays superimpose with each other producing interference and forming interference and forming interference patterns in the form of bright dark circular rings.
7. These rings are seen with a microscope (M) focused on the air film.

<u>Explanation of Newton’s rings:</u>

- Newton’s rings are formed due to interference between the light rays reflected from the top and bottom surfaces of air film between the glass plate and the Plano-convex lens.
- The formation of Newton’s rings can be explained with help of fig.
- When a light ray (AB) is incident on the systems, a part of the light is reflected by the curved surface of the lens and a part is transmitted which is reflected back from the plane surface of the glass plate. Ray 1 undergoes no phase change but ray 2 under goes a phase of π or path change $\frac{\lambda}{2}$ (i.e., the wave loses half of wavelength) when it gets reflected at the boundary of a rare to denser medium.
- These reflected light rays superimpose with each other producing interference and forming interference and forming interference patterns in the form of bright dark circular rings as shown in fig.
- We know that; the path difference between two reflected rays in uniform thin film is ,

$$\text{Path difference } \Delta = 2\mu t \cos r \frac{-\lambda}{2} \quad (1)$$

- **<u>For non-uniform thin film :</u>**

For air film $\mu=1$ and

For normal incidence r =0.

Path difference $\boxed{\Delta=2t\frac{-\lambda}{2}}$

Thus the path difference $\Delta=2t\frac{-\lambda}{2}$ ---- (2)

This is the path difference between two reflected rays in non-uniform thin film.

Special Case : Dark central spot

At the point of contact 'O' of the lens and glass plate, the thickness of air film is approximately is zero.

i.e., $t\approx 0$

$$\therefore \text{The path difference } \Delta= 2t\frac{-\lambda}{2}$$

$$= 2(0)\frac{-\lambda}{2}$$

$$\Delta = \frac{-\lambda}{2} \text{----} \quad (3)$$

Thus the two waves are at the centre, out of phase and interference destructively.

Hence the central spot is dark.

Theory of Newton's rings:

To find the diameters (radii) of dark and bight rings, let 'L' be a Plano convex lens placed on a glass plate P.

The convex surface of the lens is the part of spherical surface with centre at C as shown in fig.

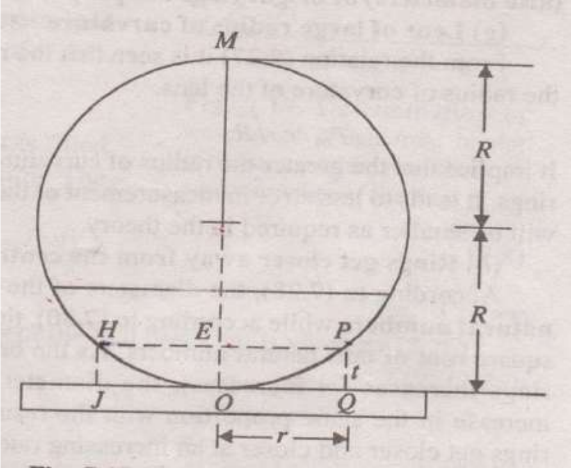


Fig: Theory of Newton's rings

Let 'R' be the radius of curvature of the lens.

Let a dark be located at the point Q is PQ= t.

The radius of the ring at Q is OQ= r.

When the additional path difference due to reflection at Q is taken into account, path difference between the two reflected beams becomes $2t+\frac{\lambda}{2}$.

By the theorem of intersecting chords,

EP x HE=OE x EM

But EP=OQ=HE= r; OE=PQ= t

And EM= OM-OE =2R-t

From equation (1),

$$r \times r = t \times (2R-t)$$

(Or) $r^2 = 2Rt-t^2$

As $2Rt \gg t^2$; t^2 can be neglected.

$$r^2 = 2Rt \quad \text{------−−−−} \quad (4)$$

Case1: Condition for bright Ring :

$$2t + \frac{\lambda}{2} = n\lambda \qquad \text{where } n=1,2,3,\text{-------}$$

$$2t= (2n-1))\frac{\lambda}{2} \quad −−−− \quad (5\ )$$

Case2: Condition for Dark Ring :

$$2t + \frac{\lambda}{2} = (2n+1)\frac{\lambda}{2}$$

$$2t = n\,\lambda \quad \text{-------−−−−} \quad (\ 6)$$

(1) **Calculation of Diameters (Radii) of the bright rings:**

Let us now suppose that a bright ring be located at the point Q.

Therefore, the radius of the nth bright ring will be given by

$$r_n^2 = 2Rt$$

We know that, the condition of the bright band is $2t= (2n-1)\ \frac{\lambda}{2}$

From above equations we have,

$$r_n^2 = (2n-1)\ \frac{\lambda}{2}\ R \ = (n- \frac{1}{2})\ \lambda R$$

$$r_n¿\sqrt{(n-\frac{1}{2})\lambda R} \quad (7)$$

This is the condition for radii of the bright rings.

The diameter of the bright ring is therefore given by

$$\frac{D_n^2}{4} = (n - \frac{1}{2})\ \lambda R \qquad [\text{Since, } r_n = \frac{D_n}{2}]$$

$$\frac{D_n^2}{4} = (\frac{2n-1}{2})\ \lambda R$$

$$D_n¿\sqrt{(2n-1)2\lambda R} \quad (8)$$

∴ The radii (diameters) of the bright rings is directly proportional to (i) $\sqrt{2n-1}$

(ii) $\sqrt{\lambda}$

(iii) $\sqrt{R}$

(2) Calculation of diameter (Radii) of the dark rings:

Thus, the radius of the nth dark ring will be given by,

$$r_n^2 = 2Rt$$

We know that, the condition of the dark ring is $2t = n\lambda$

From the above equations we have,

$$r_n^2 = n\lambda R$$

r_n¿$\sqrt{n\lambda R}$

This is the condition for radii of the dark rings.

The diameter of the dark ring is therefore given by

$$\frac{D_n^2}{4} = n\lambda R \qquad \left[\text{Since, } r_n = \frac{D_n}{2}\right]$$

D_n¿$\sqrt{4n\lambda R}$

D_n¿$2\sqrt{n\lambda R}$ (9)

∴ The radii (diameters) of the dark rings is directly proportional to (i) $\sqrt{n(\textit{natural nuber})}$

(ii) $\sqrt{\lambda}$

(iii) $\sqrt{R}$

Applications:

The theory of Newton's rings can be used

(i) To determine the wavelength of monochromatic light and

(ii) To determine the refractive index of a given liquid.

Application (i): Determination of wavelength of monochromatic light source.

Let 'R' be the radius of curvature of a Plano-convex lens, λ be the wavelength of monochromatic light source.

Let D_m and D_n are the diameters of m^th^ and n^th^ dark rings respectively.

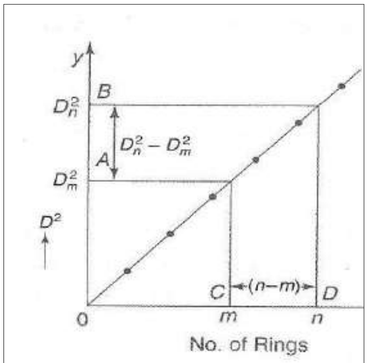


Fig: A graph between (D^2) and number of rings.

Then, D_m^2¿$4m\lambda R$,

D_n^2¿$4n\lambda R$,

and $D_n^2 - D_m^2 = 4n\lambda R - 4m\lambda R$

$= 4(n-m)\lambda R$

Graph:

$$\lambda = \frac{D_n^2 - D_m^2}{4(n-m)R}$$

A graph is drawn with the number of rings on the x-axis and the square of the diameter of the ring (D^2) on the y-axis. The graph is straight line passing through the origin. Form the graph the values of D_m^2 and D_n^2 corresponding to nth and mth rings are found. From the graph, the slope is calculated.

From graph,

$$\frac{D_n^2 - D_m^2}{(n-m)} = \frac{AB}{CD}$$

The radius of curvature 'R' of the Plano-convex lens is found by Boy's method or spherometer. Substituting the values of R and the values of $D_n^2 - D_m^2$ in the above formula (9),λ can be calculated.

Application 2: Determination of refractive index of a given liquid:

To find the refractive index of a given liquid, the Plano convex lens and glass plate set up is placed in a small container 'C'. The transparent liquid of refractive index 'μ'is introduced between the lens L and the glass Plate 'G' as shown as fig. Then a film of liquid formed between the lens and glass plate. The diameters of mth and nth dark rings are determined with help of travelling microscope.
We know that;

The diameters of mth and nth dark rings in the air film are given by

$$D^2{}_m ¿ 4m\lambda R,$$

$$D^2{}_n ¿ 4n\lambda R,$$

$$\text{and } D^2{}_n - D^2{}_m = 4n\lambda R - 4m\lambda R$$

$$= 4(n-m)\lambda R \quad ------ \quad (9)$$

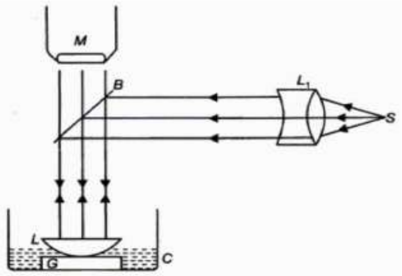


Fig: Determination of refractive index of a given liquid.

With liquid film, the diameters of mth and nth dark rings are determined;

$$D^2{}_m{}^I = \frac{4m\lambda R}{\mu}$$

$$D^2{}_n{}^I = \frac{4n\lambda R}{\mu}$$

$$\text{And } D^2{}_n{}^I - D^2{}_m{}^I = \frac{4n\lambda R}{\mu} - \frac{4m\lambda R}{\mu}$$

$$= \frac{4\lambda R(n-m)}{\mu} \quad ------ \quad (10)$$

From equations (9) and (10); $D^2{}_n{}^I - D^2{}_m{}^I = \frac{D_n^2 - D_m^2}{\mu}$

$$\mu = \frac{D_n^2 - D_m^2}{D_n^2 I - D_m^2 I}$$. Using the above formula μ can be calculated.

Unit-I- chapter-2
Diffraction

Introduction:

The wave nature of light is further confirmed by the optical phenomenon of diffraction .The word ‘diffraction’ is derived from the Latin word ‘dirrfactus’ which means to break pieces. It is common experience that waves bends around the obstacles placed in their path. When light waves encounter an obstacle, they bend round the edges of the obstacle. This bending is predominant when size of the obstacle is comparable to the wavelength of light. The bending of light waves around the edge or corner of an obstacle is diffraction. It was first observed by Gremaldy.

As shown in figure 2.1, when light falls on an obstacle then the corresponding geometrical shadow on the screen should be completely dark. In practice, the geometrical shadow consists of bright dark fringes. These fringes are due to the superposition of bended light waves around the edges or corners of an obstacle. The amount of bending depends upon the size of an obstacle and wavelength of light.

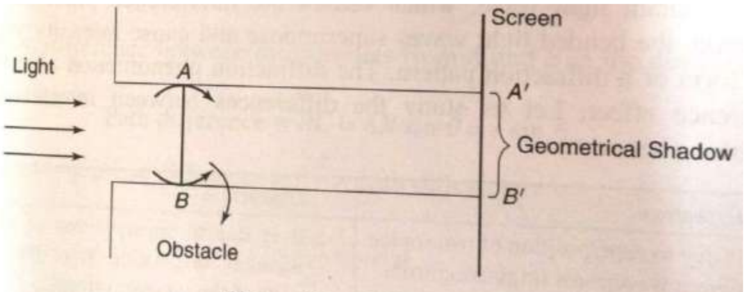


Figure 2.1 Diffraction

Diffraction definition:

When a light falls on an obstacle whose size is comparable with wavelength of light,then light bends around the edges or corners of an obstacle.This bending nature of light tacle is called diffraction.

The diffraction phenomenon is brodly devided into two types.They are

(1) Fresnel’s diffraction
(2) Fraunhoffer’s diffraction.

Difference between Interference and diffraction :

Interference	Diffraction
1.Superposition is due to two separate	1.Superposition is due to secondary wavelets

wavefronts originating from two cohenent sources. 2.The fringenormally have equal widths. 3.All the bright fringes have same intensity. 4.All the dark fringes have zero intensity.	originating from different parts of the same wavefront. 2.The width between frines is never equal. 3.The intensity of bright fringes usually decreases with increase of order. 4.The intensity of dark fringes is not zero.

<u>Difference between Fresnel's diffraction and Fraunhoffer diffraction:</u>

Fresnel's Diffraction	Fraunhoffer diffraction
1.For diffraction to occur,the light source and screen are at finite distance from the obstacle. 2.No lenses are necessary to study the diffractiondiffraction. 3.Study of diffraction is complicated. 4.Either a point source or an illuminated narrow slit is used. 5.Diffraction can be studied only in the direction of propagation of light. 6.In this case the incident wavefronts are either spherical (or) spherical.	1.For diffraction to occur,the light source and screen are at infinite distance from the obstacle. 2. Lenses are necessary to study the diffractiondiffraction. 3.Study of diffraction is easy. 4.Extended source at infinite distance is used.. 5. Diffraction can be studied in any direction of propagation of light. 6.In this case, incident wavefronts are plane.

<u>Fraunhoffer diffraction at Single Slit:</u>

- Consider a slit AB of width 'e' perpendicular to the plane of the paper. Let a plane wave front of monochromatic light of wavelength 'λ' incident at this slit.
- At slit, each and every point is able to generate secondary wavelets. These wavelets are focused through the lens 'L'.
- The wavelets travelling normal to the lens are brought to focus at at P_0 and the wavelets travel with an angle θ are focused at P_1 .Now the intensity at P_O is central maximum with bright fringe ,and intensity at P_1 is either maxima or minima depends upon the angle 'θ'.

- To find the intensity at P_1, first we should calculate the phase difference at point P_1. For this, AC is perpendicular to BC.

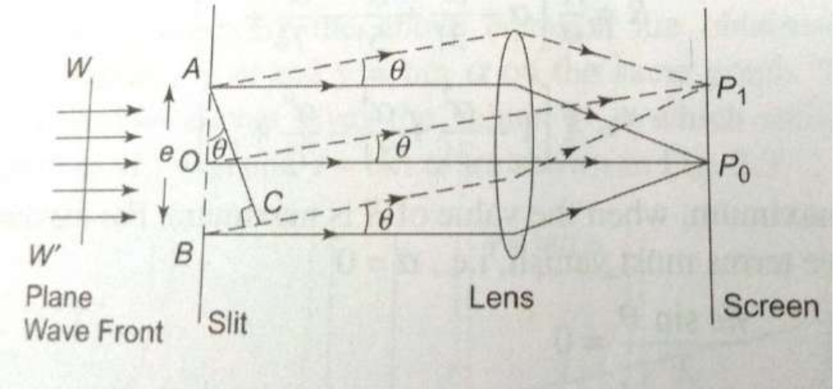


Figure 2.2 Fraunhofer diffraction—single slit

Path difference between the wavelets along P_o & P_1 is given by,

$$\delta = BC$$

From ΔACB, $\mathrm{Sin}\theta = \frac{BC}{AB}$

$$BC = AB\sin\theta$$

$$BC = e\sin\theta$$

$$\therefore\ \delta = BC = e\sin\theta$$

But the phase difference $\phi = \frac{2}{\square}$ (Path difference δ)

$$\therefore\ \phi = \frac{2}{\square}\,(e\sin\theta)\ \text{---------------}> (1)$$

Let the slit ‘s’ is decided into ‘n’ number of small slits of each with amplitude ‘a’, then,

Phase difference at each slit (d) = $\frac{\text{Total phase difference}}{\text{number of slits}}$

$$d = \frac{\square}{n} = \frac{2}{n}\,(e\sin\theta)\ \text{-------------}> (2)$$

After diffraction, the resulting amplitude R can be obtained by using vector addition method with each of amplitude ‘a’.

$$R = \frac{a\sin\left(\frac{nd}{2}\right)}{\sin d/2}\qquad \text{----------------}> (3)$$

Substitute eq(2) in eq(3) we get,

$$R = \frac{a\sin\left(n\left\{\frac{2\,eSin}{n}\right\}\right)/2}{\sin\left\{\frac{2\,eSin}{n}/2\right\}}$$

$$R = \frac{a\sin\left(\frac{eSin}{\square}\right)}{\sin\left(\frac{eSin}{\square}\right)/n}$$

Put $\frac{\square}{\square}$ esinθ =α , then

$$R = \frac{aSin\alpha}{\sin\left(\frac{\alpha}{n}\right)}$$

Since $\frac{\alpha}{n}$ is small, Sin(α/n) ≈ α/n

$$R = a\frac{Sin\alpha}{(\alpha/n)}$$

$$R = n\,a\frac{Sin\alpha}{\alpha}$$

$$R = A\frac{Sin\alpha}{\alpha} \text{ ------------------}\rightarrow (4)$$

Intensity at P_1 is given by,

$$I = R^2 = \left(A\frac{Sin\alpha}{\alpha}\right)^2 = A^2\frac{\sin^2\alpha}{\alpha^2} \text{ -------------------}\rightarrow (5)$$

(1) Principal maximum :

From eq(4) if write the Sinα expansion we have,

$$R = \frac{A}{\alpha}[\alpha - \frac{\alpha 3}{3!} + \frac{\alpha 5}{5!} - \frac{\alpha 7}{7!} \text{----------}]$$

$$= \frac{A\alpha}{\alpha}[\alpha - \frac{\alpha 2}{3!} + \frac{\alpha 4}{5!} - \frac{\alpha 6}{7!} \text{----------}]$$

In above equation for ‘R’ maximum negative terms and higher order terms vanishes i.e, α=0.

But $\frac{\square}{\square}$ esinθ =α =0

$\therefore$ Sinθ =0 $\Rightarrow$θ=0°

Then R = A $\Rightarrow$ $I^2 = A^2 = R^2$

At θ =0°, the maximum intensity is formed at P_o known as principal maximum.

2.Minimum intensity position :

From eq(5) for ‘I_{minimum} , Sinα=0

$\therefore \alpha = \pm m\pi$, m= 0,1,2,3,---------

For m=0, α =0 which gives principal maxima and for α = ±π,±2π,±3π ,------------minimum intensity is observed at P_1,

$\therefore$ α = ±mπ

$\frac{\square}{\square}$ esinθ = ±mπ

esinθ = ±mλ

The above condition is for minimum intensity position.

3. Secondary Maxima:

Between two secondary minima, the secondary maxima can be obtained by differentiating eq(4) w.r.to α and equating to zero , we get

$$\frac{dI}{d\alpha} = 0 \Rightarrow \frac{dI}{d\alpha} = \frac{d}{d\alpha}[A^2\left(\frac{Sin\alpha}{\alpha}\right)^2] = 0$$

$$A^2\left(\frac{2Sin\alpha}{\alpha}\right)[\frac{\alpha Cos\alpha - Sin\alpha}{\alpha^2} ¿ = 0$$

$$A^2\left(\frac{2Sin\alpha}{\alpha}\right) \neq 0, \text{ then } \frac{\alpha Cos\alpha - Sin\alpha}{\alpha^2} = 0$$

$$\alpha Cos\alpha - Sin\alpha = 0$$

α Cosα = Sinα
α = Tanα
which is interms y=α and y=Tanα
Now draw graphs y=α and y=Tanα

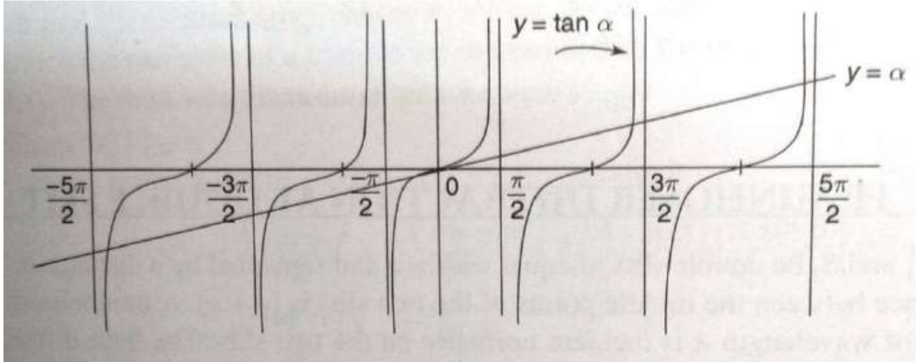


Figure 2.3 Plots of $y = \alpha$ and $y = \tan\alpha$

From graph, it is clear that y=α line touches the y=Tanα curves at

$\alpha = 0, \pm\frac{3}{2}, \pm\frac{5}{2}, \pm\frac{7}{2}$-------.

So, at these α values, secondary maxima can be obtained.We already know that α = 0 is principal maxima. So,

1. $\alpha = \pm\frac{3}{2}$, $I = A^2\left[\frac{\sin(\frac{3}{2})}{(\frac{3}{2})}\right]^2 = \frac{A^2}{\frac{9\square^2}{4}} = \frac{4A^2}{9\square^2}$ ---→ first secondary maxima

2 .$\alpha = \pm\frac{5}{2}$, $I = A^2\left[\frac{\sin(\frac{5}{2})}{(\frac{5}{2})}\right]^2 = \frac{A^2}{\frac{25\square^2}{4}} = \frac{4A^2}{25}$ ---→ second secondary maxima

4. Intensity distribution :

The intensity distribution of diffraction intensity depends on α values, i.e, α = 0, principal maxima.

$\alpha = 0, \pm\frac{3}{2}, \pm\frac{5}{2}, \pm\frac{7}{2}$------- secondary maxima.

For $\alpha = \pm\pi, \pm 2\pi, \pm 3\pi$,------------minimum distribution of intensity.

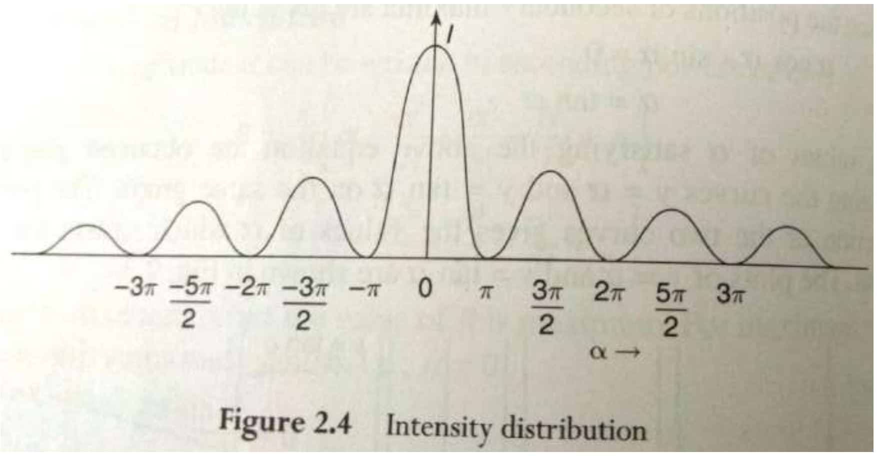


Figure 2.4 Intensity distribution

Fraunhoffer diffraction at double slit:

- Let S1 and S2 be double slits of equal widths 'e' and separated by a distance'd'. The distance between the middle points of the two slits is (e+d).
- A monochromatic light of wavelength of λ is incident normally on the two slits .The light diffracted from these slits is focused by a lens on the screen placed in the focal plane of the lens.

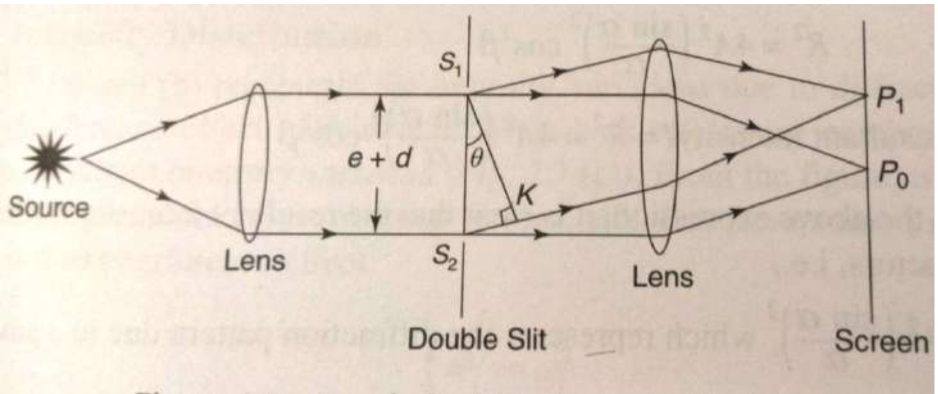


Figure 2.5 Fraunhofer diffraction—double slit

- The diffraction at two slits is the combination of diffraction as well as interference.i.e, the pattern on the screen is the diffraction pattern due to a single slit on which a system of interference fringes is superimposed.
- When a plane wave front is incident normally on two slits, the secondary wavelets from the slits travels uniformly in all directions. The wavelets travelling in the direction of incident light comes to a focus at P_o while the wavelets travelling in a direction making an angle θ, come to focus at P_1 (see figure 2.5).

From the study of diffraction due to single slit, the resultant amplitude = $\frac{ASin\alpha}{\alpha}$ where $\alpha = \frac{\pi eSin\theta}{\lambda}$.

Since we use double slit,from each slit we get a wavelet of amplitude $\frac{ASin\alpha}{\alpha}$ in a direction θ.These two wavelets interfere and meet at a point P_1 on the screen .To calculate the path difference between the wavelets,let us draw a normal S_1K to the wavelet through S_2 .

$$\text{Path difference} = S_2K$$
$$= (e+d)Sin\theta$$

$$\text{Phase difference } \delta = \frac{2\pi}{\lambda}(e+d)\sin\theta$$

To find the resultant amplitude at P_1 we use vector addition method (see fig.2.6) in which the two sides of a triangle are represented by the amplitudes through S_1 and S_2.The third side gives the resultant amplitude.

From figure,

$$(OH)^2 = (OG)^2+(GH)^2+2(OG)(GH)\ Cos\delta$$

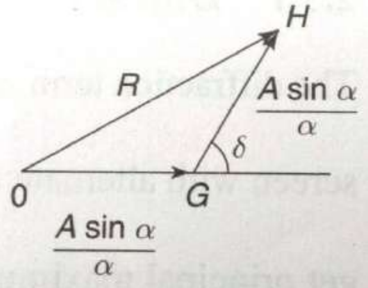


Figure 2.6 Vector addition method

$$R^2 = \left(\frac{ASin\alpha}{\alpha}\right)^2 + \left(\frac{ASin\alpha}{\alpha}\right)^2 + 2\left(\frac{ASin\alpha}{\alpha}\right)\left(\frac{ASin\alpha}{\alpha}\right) Cos\delta$$

$$= A^2\frac{\sin^2\alpha}{\alpha^2}[2+2\cos\delta]$$

$$= 2\left(\frac{ASin\alpha}{\alpha}\right)^2(1+ Cos\delta)$$

$$= 2\left(\frac{ASin\alpha}{\alpha}\right)^2(1+2\ Cos^2(\delta/2)-1)$$

$$R^2 = 4A^2\left(\frac{Sin\alpha}{\alpha}\right)^2 Cos^2\left[\frac{\pi(e+d)Sin\theta}{\lambda}\right]$$

$$\text{Let } \beta = \frac{\pi(e+d)Sin\theta}{\lambda}$$

$$R^2 = 4A^2\left(\frac{Sin\alpha}{\alpha}\right)^2 Cos^2\beta$$

The resultant intensity $I = R^2 = 4A^2\left(\frac{Sin\alpha}{\alpha}\right)^2 Cos^2\beta$

From the above expression it is clear that the resultant intensity is the product of two factors i.e,

1. $4A^2\left(\frac{Sin\alpha}{\alpha}\right)^2$ which represents the diffraction pattern due to a single slit.
2. $Cos^2\beta$ which gives the interference pattern due to wavelets from double slits.

The resultant intensity is due to both diffraction and interference effects.

1. Difraction effect :

The diffraction term $4A^2\left(\frac{Sin\alpha}{\alpha}\right)^2$ gives the principal maximum at the centre of the screen with

alternate minima and secondary maxima of decreasing intensity.
We get principal maximum for $\theta=0$.We get minima for $Sin\alpha=0$.

$$\alpha = \pm m\pi \quad , \quad \text{where } m=1,2,3,4---$$

$$\frac{\pi e Sin\theta}{\lambda} = \pm m\pi$$

$$e Sin\theta = \pm m\lambda$$

The positions of secondary maxima occurs for $\alpha = \frac{\pm 3\pi}{2}, \frac{5\pi}{2}, \frac{\pm 7\pi}{2}, ------¿$

2. Interference effect :

The interference term $Cos^2\beta$ gives the equidistant bright and dark fringes.
The maxima will occur for $Cos^2\beta=1$.

$$\beta = \pm n\lambda \quad \text{where } n=0, 1,2,3-------$$

$$\beta = 0, \pm\pi, \pm 2\pi, \pm 3\pi \text{ ---------}$$

$$\frac{¿\pi(e+d)Sin\theta}{\lambda} = \pm n\pi$$

$$(e+d)Sin\theta = \pm n\lambda$$

The minima will occur for $Cos^2\beta=0$.

$$\beta = \pm(2n+1)\frac{\pi}{2} \qquad \text{where } n=0,1,2,3,-----$$

$$(e+d)Sin\theta = \pm(2n+1)\frac{\lambda}{2}$$

Intensity Distribution:

The figures shown here represents the intensity variations due to diffractions as well as interference effect individually (fig.(2.7 a) and fig.(2.7 b)) . When both effects are combined, then we get the resultant intensity variations(fig.2.7(c). From the figure, it is clear that the resultant minima are not equal to zero still they have some minimum intensity due to interference effect.

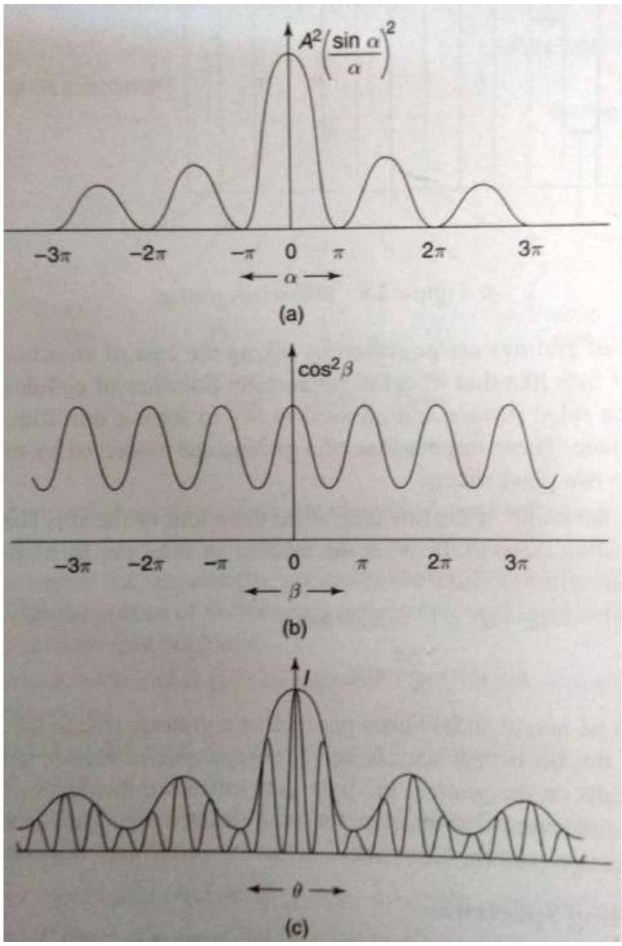


Figure 2.7: (a) Diffraction effect (b) Interference effect (c) Resultant Intensity

Diffraction Grating:

Definition of diffraction grating:

An arrangement which consists of a large number of parallel slits of the same width and separated by equal opaque spaces is known as Diffraction grating.

Explanation :

- ✓ Fraunhoffer used the first grating consisting of a large parallel wires placed side by side very closely at regular intervals.
- ✓ Now gratings are constructed by ruling equidistant parallel lines on a transparent material such as glass with a fine diamond point.
- ✓ The ruled lines are opaque to light while the space between any two lines is transparent to light and acts as a slit (Fig 2.8).

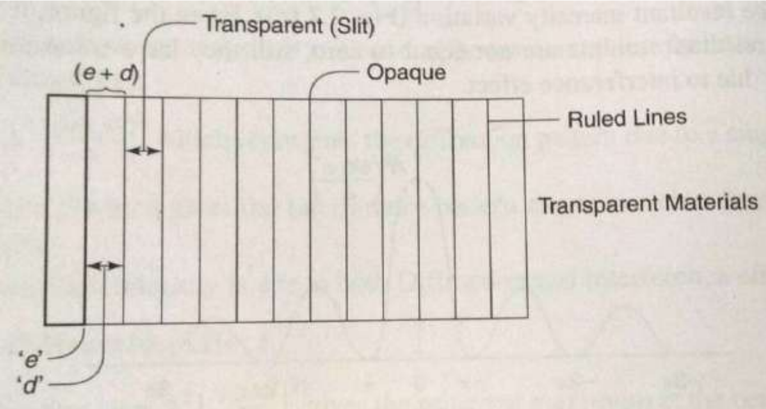


Figure 2.8 Diffraction grating

- ✓ Commercial gratings are produced by taking the cast of an actual grating on a transparent film like that of cellulose acetate.
- ✓ Solution of cellulose acetate is poured on the ruled surface and allowed to form a thin film, detachable from the surface. These impressions of a grating are preserved by mounting the film between two glass sheets.

Let 'e' be the width of the line'd' be the width of the slit. Then (e+d) is called grating element. If 'N' be the number of lines per inch on the grating, then

$$N\,(e+d) = 1 \text{ inch} = 2.54 \text{ cm}$$

$$(e+d) = \frac{2.54}{N} \text{ cm}$$

- ✓ There will be nearly 30000 lines per inch of a grating. Due to above fact, the width of the slit is very narrow and is comparable with the wavelength of light. When light falls on the grating, the light gets diffracted through each slit.
- ✓ As a result, both diffraction and interference of diffracted light gets enhanced and forms a diffraction pattern. This pattern is known as diffraction spectrum.

Grating spectrum:

The condition to form the principal maxima in a grating is given by,

$$(e+d)\sin\theta = n\lambda$$

Where (e+d) is the grating element and the above equation is known as grating equation.
From the grating equation, the following points are clear.

- ✓ For a particular wavelength ' λ' the angle of diffraction θ is different for principal maxima of different for principal maxima of different orders.
- ✓ As the number of lines in the grating are large , maxima appears as sharp , bright parallel lines and are termed as spectral lines.
- ✓ For white light and for a particular order of 'n 'the light of different wavelengths will be diffracted in different directions.
- ✓ At the centre, θ = 0 which gives the maxima of all wavelengths which coincides to form the central image of the same colour as that of the light source. This forms zero order (Fig 2..9).

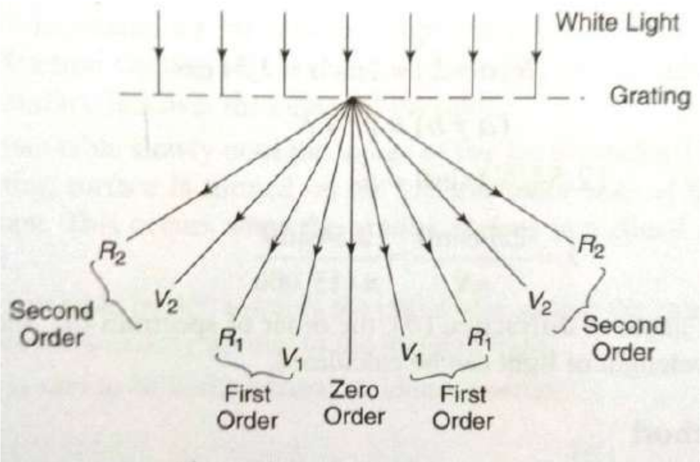


Figure 2.9 Grating spectrum

- ✓ The principal maxima of all wavelengths form the first, second-------order spectra for n=1,2,------
- ✓ The longer the wavelength, greater is the angle of diffraction. Thus, the spectrum consists of violet being in the innermost position and red being in the outermost positions.
- ✓ Most of the intensity goes to zero order and the rest is distributed among other orders. Spectra of different orders are situated symmetrically on both sides of zero order.
- ✓ The spectral lines are more and more dispersed as we go to higher orders ($\sin\theta = \sin 90^0 = 1$)
- ✓ The maximum number of orders available with the grating is $n_{max} = \frac{(e+d)}{\lambda}$.

Unit-I- chapter-3
Electromagnetic Theory

Introduction

Electromagnetic Theory is a branch of Physics in which electric and magnetic phenomena are studied. The devices which come under this are waveguides, antennas, optical fibers, radars, etc.

Scalar and Vector Field

The region of space in which it specifies a physical field is called as a field. It is divided into two types:

(i) Scalar field: The region of space where each point is associated with a scalar point is called as scalar field.

(ii) Vector field:A vector is specified by a continuous vector point having both magnitude and direction.

gradient

To study the rate of change of scalar and vector a common operator del '∇' is used in vector calculus.

$$\nabla = i\frac{\partial}{\partial x}+j\frac{\partial}{\partial y}+k\frac{\partial}{\partial z} \text{ where i, j, k are unit vectors}$$

Divergence of a Field

Divergence is scalar or dot product of operator '∇' with a vector A. It is a scalar.The Divergence of a vector field at any point is defined as the amount of flux per unit volume diverging from that point.

$$\nabla\,.\,\mathrm{A} \;=\left[i\frac{\partial}{\partial x}+j\frac{\partial}{\partial y}+k\frac{\partial}{\partial z}\right].\left(iAx+jAy+kAz\right)$$

$$= i\frac{\partial Ax}{\partial x}+j\frac{\partial Ay}{\partial y}+k\frac{\partial Az}{\partial z}$$

$$\therefore \text{ div A} = \nabla.\,\mathrm{A} = i\frac{\partial Ax}{\partial x}+j\frac{\partial Ay}{\partial y}+k\frac{\partial Az}{\partial z}$$

Curl of a field

The curl of a vector field is defined as the maximum line integral of the vector per unit are. It is a vector quantity and direction is normal to the area.

$$\text{Curl A} = \nabla\,.\,\mathrm{A} = \begin{vmatrix} i & j & k \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ Ax & Ay & Az \end{vmatrix}$$

Stoke's Theorem for Curl

Statement

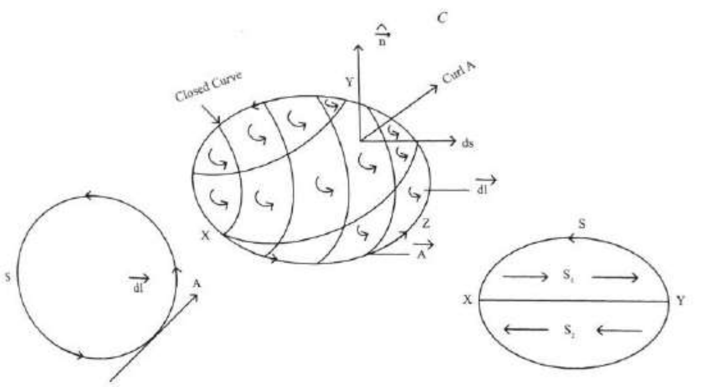
Stoke's theorem states that the line integral of a vector field 'A' around a closed curve is equal o the surface integral of the curl of 'A' taken over the surface 'S' surrounded by the closed curve. It is expressed as

$$\overset{\square}{\oint_C} A\,dl=\overset{\square}{\iint_S} Curl\,A\,ds=\overset{\square}{\iint_S}\left(\nabla\;\; X\,A\right)ds$$

Theorem

Let us consider surface 'S' enclosed in a vector field 'A' as shown in figure in which a closed curve XYZ represents the boundary of surface 'S'.

The line integral of 'A' around the curve XYZ is $\overset{\square}{\oint_C} A\,dl$



Figure

Let the entire surface ‘S’ is divided into large number of square loops, ‘ds’ be the are enclosed by each small loop, and n be the unit positive outward normal upon small loop ‘ds’. Then vector area of the element is

n ds = ds---- (1)

The curl of a vector field at any point is the maximum line integral of the vector computed per unit area, along the boundary of an infinitesimal area at that point. Then, the line integral of A around the boundary of ‘ds’ is

Curl A ds---- (2)

It is applied to all surface elements and hence, the sum of the line integrals of ‘A; around the boundaries of all the area elements can be written as

$$\iint_{S}^{\square} Curl\,A\,ds \text{---- (3)}$$

which is surface integral of ‘A’.

From the figure, we can understand that the line integrals along the common sides of the continuous elements mutually cancel because they are transferred in the opposite direction.

The sides of the elements which lie in the closed curve of the surface contribute to the line integral. The equation (3) gives the sum of the line integrals on the boundary line of the curve.

Therefore,

$$\oint_{C}^{\square} A\,dl=\iint_{S}^{\square} Curl\,A\,ds=\iint_{S}^{\square}(\nabla\ \ X\,A)\,ds \text{---- (4)}$$

Equation (4) is known as Stoke’s Theorem.

Stokes theorem, given by equation (3) is a method to convert a surface integral into a line integral and vice versa.

Gauss’s theorem for divergence

statement

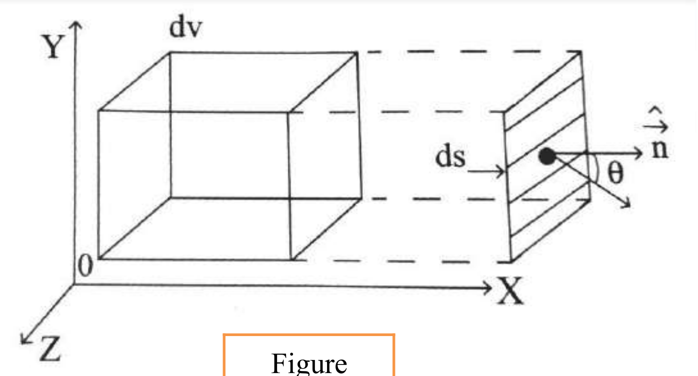
The Gauss Theorem states that the surface integral of the normal component vector ‘A; taken over a closed surface ‘S’ is equal to the volume integral of the divergence of vector ‘A’ over the volume ‘V enclosed by the surface ‘S’, i.e.,

$$\iint_{s}^{\square} A\,ds=\iiint_{V}^{\square} \text{¿} A\,dv=\iiint_{V}^{\square}(\nabla\ \ .A)\,dv$$

The relation between the volume integrals and surface integrals were proved using this theorem.

Theorem

Let us consider a surface ‘S’ in a vector field ‘A’ as shown in figure. Let ‘V’ be the volume enclosed in the surface and the volume is divided into a large number of cubical volume elements ‘dv’



Figure

We have the ‘div A’ represents the amount of flux diverging per unit volume and hence, the flux diverging from the element of volume ‘dv’ will be ‘div A’.

Hence, the total flux coming from the entire volume can be written as

$$\underset{V}{\overset{\square}{\iiint}} ¿\, A\,dv \quad \text{----(1)}$$

Consider a small element of area ‘ds’ with n unit vector drawn normal to ‘ds’. Here, outward drawn normal on a surface is positive.

For a field vector ‘A’ and outward normal n at an angle ‘θ’ the ‘A’ component along n can be written as

$$A \cos\theta = A \,.\, n \text{----(2)}$$

The flux of ‘A’ through the surface element ‘ds’ is given by

$$(A.n)ds = A.ds \quad \text{----(3)}$$

(Since, flux is defined as the product of a normal component of vector and surface area)

The total flux through the entire surface ‘S’ is given by

$$(A.n)\ ds = A\ ds \text{ ----(4)}$$

This must be equal to the total flux diverging from the whole volume ‘V’, enclosed by the surface S.

Therefore, from equation (1) and (4) we have

$$\underset{s}{\overset{\square}{\iint}} (A.n)ds \quad \text{----(5)}$$

Equation (5) (Gauss Theorem) can also be written as

$$\underset{s}{\overset{\square}{\iint}} (A.n)ds = \underset{V}{\overset{\square}{\iiint}} (\nabla\ .A)dv$$

Maxwell’s equations:

Ampere’s law of the form $\nabla\ XB=\mu_0 j$ that is valid only for time independent fields which makes the law inconsistent. Maxwell modified the ampere’s equation by introducing the concept of displacement current

Differential form of Ampere’s law is given by

$$\nabla \ XB=\mu_o j$$

From the relation between Magnetic induction (B) and magnetic field intensity (H),
B= $\mu_o H$
From the above two equations, we get

$$\nabla \ XB=\mu_o j$$

$$\frac{1}{\mu_o}(\nabla \ XB)=j$$

$$\nabla \ XH=j$$

Taking divergence on both sides of the equation,

$$\nabla \ .(\nabla \ XH)=\nabla \ .j$$

As the divergence of the curl of the vector is zero, we have

$$\nabla \ .j=0$$

Above relation states that the net flux of current out of any closed surface is zero, which is known as the equation of continuity for steady fields and it doesnot apply for the time varying fields.
By Gauss law, we have,

$$\nabla \ .E=\frac{\rho}{\in_0}$$

By differentiating above equation, we have

$$\nabla \ .\frac{\partial E}{\partial t}=\frac{1}{\in_0}\frac{\partial l}{\partial t}$$

$$\in_0 \nabla \ .\frac{\partial E}{\partial t}=\frac{\partial l}{\partial t}$$

Adding $\nabla \ .j$ to both sides of the above equation

$$\nabla \ .j+\in_0 \nabla \ .\frac{\partial E}{\partial t}=0$$

$$\nabla \ .(j+\in \text{¿¿}0\nabla \ .\frac{\partial E}{\partial t})=0\text{¿}$$

Displacement vector in free space is given by
D= $\in_0 E$
On substitution, we get

$$\nabla \ .(j+\frac{\partial D}{\partial t})=0$$

On replacing j by $j+\frac{\partial D}{\partial t}$ in Ampere's law, we have

$$\nabla \ XH=j+\frac{\partial D}{\partial t}$$

The term $\frac{\partial D}{\partial t}$ on addition to j is called Displacement current that has same effect as the true current but is effective only when E is a time varying field.

Maxwell's equations in differential form:

1. $\nabla \ .E=\frac{\rho}{\in_0}$

2. $\nabla .B=0$
3. $\nabla XE=\frac{-\partial B}{\partial t}$
4. $\nabla XH=j+\frac{\partial D}{\partial t}$

Maxwell's equations in integral form:

1. $\oint E.ds=\frac{q}{\in_0}$
2. $\oint B.ds=0$
3. $\oint E.dl=\frac{-\partial \O_B}{\partial t}$
4. $\oint B.dl=(I+\in_0\frac{-\partial \O_E}{\partial t})\mu_0$

Electromagnetic wave propogation (nonconducting medium):

Consider a nonconducting, homogeneous and isotropic medium having permeability and permittivity μ and ε respectively.
For this nonconducting medium,

$$B=\mu H H$$

$$\& D=\varepsilon E$$

For a plane electromagnetic wave propogating along z-axis, the field vectors are given by

$$E(z,t)=E(Z)e^{iwt}$$

$$B(z,t)=B(Z)e^{iwt}$$

The constant-phase surface coincides with the surface of constant amplitudes and hence they are called as homogeneous electromagnetic waves.

The electrical conductivity for a dielectric, σ = 0 (J= σμE¿
Now, Maxwell's equations can be written as :

$$\nabla X B=\mu\epsilon\frac{\partial E}{\partial t}$$

$$\nabla X E=\frac{-\partial B}{\partial t}$$

Differentiating above equations w.r.t "t" we get,

$$\frac{\partial(\nabla X B)}{\partial t}=\frac{\partial}{\partial t}[(\in\mu)\frac{\partial E}{\partial t}]$$

$$\nabla X\frac{\partial B}{\partial t}=\in\mu\frac{\partial^2 E}{\partial t^2}$$

Substituting for $\frac{\partial E}{\partial t}$, we get

$$\nabla\ X(-\nabla\ X\,E)=\in\mu\frac{\partial^2 E}{\partial t^2}$$

$$-\nabla\ X(\nabla\ X\,E)=\in^2\mu\frac{\partial^2 E}{\partial t^2}$$

By the property of the operator ∇ ,

$$\nabla\ X(\nabla\ XE)=\nabla\ (\nabla\ .E)-\nabla^2 E$$

On substitution , we have

$$\nabla\ .(\nabla\ .E)-\nabla^2 E=\in^2\mu\frac{\partial^2 E}{\partial t^2}$$

As there are no free charges, because the medium is dielectric, $\nabla\ .E=0$

$$-\nabla^2 E=\in^2\mu\frac{\partial^2 E}{\partial t^2}$$

For the vector field B,

$$-\nabla^2 B=\in^2\mu\frac{\partial^2 B}{\partial t^2}$$

On comparing these equations of electric and magnetic field with the simple harmonic motion of type

$$\frac{\partial^2 y}{\partial t^2}-\mu y=0$$

With velocity of wave propogation $\frac{1}{\sqrt{\mu}}$

Hence, the field vectors satisfy the wave equation in which the wave propogation velocity is Cl=c= $\frac{1}{\sqrt{\mu E}}$

Poynting Theorem:

Consider an electromagnetic wave travelling along x-axis with magnetic and electric fields as E and H along z- and x-axis respectively. It propogates along the direction f the propogation vector (EXH)

By taking divergence of the above equation, we get

$$\nabla\ .(EXH)-H.(\nabla\ XE)-E.(\nabla\ XH)$$

We know that

$$\nabla\ X\,E=\frac{-\partial B}{\partial t}$$

and $\nabla\ X\,H=\frac{\partial D}{\partial t}$

$= -H\frac{\partial B}{\partial t}-E\frac{\partial D}{\partial t}$

$$\text{¿}-[E\frac{\partial D}{\partial t}+H\frac{\partial B}{\partial t}]$$

We have D = $\boldsymbol{\varepsilon}_0$E and B= $\boldsymbol{\mu}_0$H

$$= - \left[\in_0 E\frac{\partial D}{\partial t}+\mu_0 H\frac{\partial B}{\partial t}\right]$$

$$=[\frac{1}{2}\in_0 2E\frac{\partial D}{\partial t}+\frac{1}{2}\mu_0 2H\frac{\partial B}{\partial t}]$$

$$=\frac{1}{2}\in_0 \partial \text{¿¿}$$

$$=-\frac{\partial}{\partial t}\left(\frac{1}{2}\in_0 E^2+\frac{1}{2}\mu_0 H^2\right)$$

As the surface S bounds a volume V, when the above equation is integrated to V

$$\overset{\square}{\int_V}\nabla\ .(EXH).dv=\frac{-\partial}{\partial t}\overset{\square}{\int_V}\left(\frac{1}{2}\in_0 E^2+\frac{1}{2}\mu_0 H^2\right)dv$$

By applying Divergence theorem,

$$\overset{\square}{\int_S}\nabla\ .(EXH).ds=\frac{-\partial}{\partial t}\overset{\square}{\int_V}(\frac{1}{2}\in_0 E^2+\frac{1}{2}\mu_0 H^2)dv$$

Thevector***EXH* = *P***represents the amount of field energy passing through a unit area of surface in unit time, normal to the direction of the flow of energy. This statement is called Poynting's theorem and the vector **P** is known as Poynting vector.

Unit-II
Chapter-1
LASERS

Introduction:

Laser is an acronym for 'Light Amplification by Stimulated Emission of Radiation'. Laser radiation is due to stimulated emission of radiation process which improves (amplifies) the intensity of radiation. In general, when an electron moves from a higher energy orbit to a lower orbit, it emits radiation. This emission of radiation can be explained in terms of energy levels as when the electron transits from a higher energy level to a lower energy level, it emits radiation. According to Planck's Quantum theory, the emission of radiation will be in the form of photon of energy hυ. The frequency (υ) term in the energy of the photon indicates the wave characteristics of the photon. The energy of the emitted photon is equal to the energy difference between the higher and lower energy levels. This loss of energy is attributed to the entire atom. As a result, it can be brought that the atom is moving from a higher energy state to a lower energy state.

When a large number of atoms transit to lower energy level, then each transition emit photons of same energy, same frequency and same wavelength.

$$\upsilon = \frac{E_2 - E_1}{h}$$

All the photons are in phase, and after reinforcement emit a high intensity monochromatic coherent radiation i.e, LASER radiation.

In laser, the intensity of light is amplified by a process called stimulated emission. Lasers find major applications in various fields such as medical, engineering, fiber communication, industries etc.

Lasers are more powerful than ordinary light radiation.

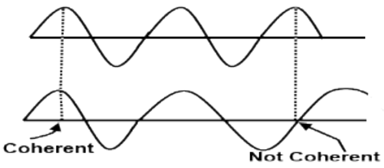
Characteristics of laser:

When compared with any conventional light (sun light, tube light, etc.,), Laser possesses few outstanding characteristics. They are

1] Coherence 2] High directionality 3] High intensity 4] High Monochromaticity

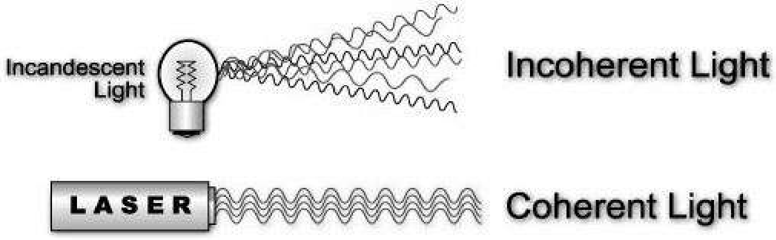
1] Coherence:

The property of existing either zero or constant phase angle difference between two or more waves is called as Coherence. The wave trains which are identical in phase and direction are called 'coherent waves'.



In case of laser, the property of coherence exits between any two or more light waves of same type. On the other hand, the waves emitted by laser source will be in phase and are of the same frequency.

Therefore, light generated by laser is highly coherence. All the constituent photons of laser beam possess the same energy, momentum and propagate in same direction, the laser beam is said to be highly coherent.



Coherence is of two types: Spatial coherence and temporal coherence

Spatial Coherence: If there exists either zero or constant phase angle difference between two points on a wave front then wave is said to have spatial coherence.

Temporal coherence: If there exists either zero or constant phase angle difference between two light fields measure at two instants at the same point then wave is said to have temporal coherence.

2] High directionality:

Lasers emit light only in one direction. Laser travels as a parallel beam; that can travel over a long distance without spreading. During the propagation of laser its angular spreading will be less and occupies less area where it incident. It possesses high degree of directionality.

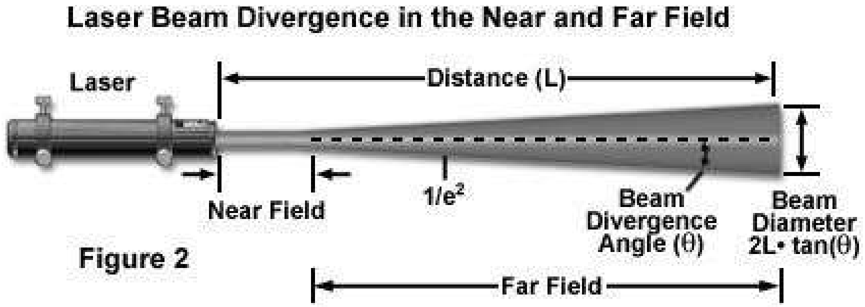

Laser Beam Divergence in the Near and Far Field

Figure 2

The laser light of wavelength ‘λ’ emerges through a laser source aperture diameter ‘2L tanθ’, then it propagates as a parallel beam upto d^2/λ (small value) and gets diverged. The angle of divergence of a laser beam is expressed as,

$$\phi = \frac{arc}{radius} = 10^{-3} \text{ radians.}$$

EX: Laser beam of 10cm diameter is operated on to the moon (which is at a distance of 3, 84,000) when it doesn’t spread over not more than 5km.

3] High intensity or brightness:

A laser emits light in the form of narrow beam which propagates in the form of plane waves. The energy is concentrated in very narrow region; its intensity would be tremendously high. Due to its directionality many beams of light incident in a small area. Therefore the intensity of light is high.

Let there be ‘n’ number of coherent photons of amplitude ‘a’ in the emitted laser radiation These photons reinforce with other and the amplitude of the resulting wave becomes ‘na’ and hence the intensity (I) is proportional to n^2a^2. Thus due to coherent additions of amplitude and negligible divergence, the intensity or brightness increases enormously.

$$I \,\alpha\, A^2$$
$$I \,\alpha\, (na)^2$$
$$I \,\alpha\, n^2a^2$$

Eg: Typically estimated that the light from a 1-mW laser is 10,000 times brighter than the light from the sun at the earth.

4] High Monochromaticity:

The property of exhibiting a single wave length by a light is called monochromaticity.

The light from normal monochromatic source spread over a wave length of the order of 100 Å to 100 Å. The laser light is highly monochromatic. The spread in wavelength is in the order of a few angstroms (<10 Å or 0.001 nm)). The laser emits continuous wave of very long duration.

On the degree of monochromaticity of light from some source we make use of line width (spectral width) of the source $\Delta\upsilon$ which is the frequency spread of a spectral line. Frequency spread $\Delta\upsilon$ is related to the wavelength spread $\Delta\lambda$ as,

$$\Delta\lambda = - [\frac{C}{\upsilon^2}] \Delta\upsilon$$

Interaction of radiation with matter:

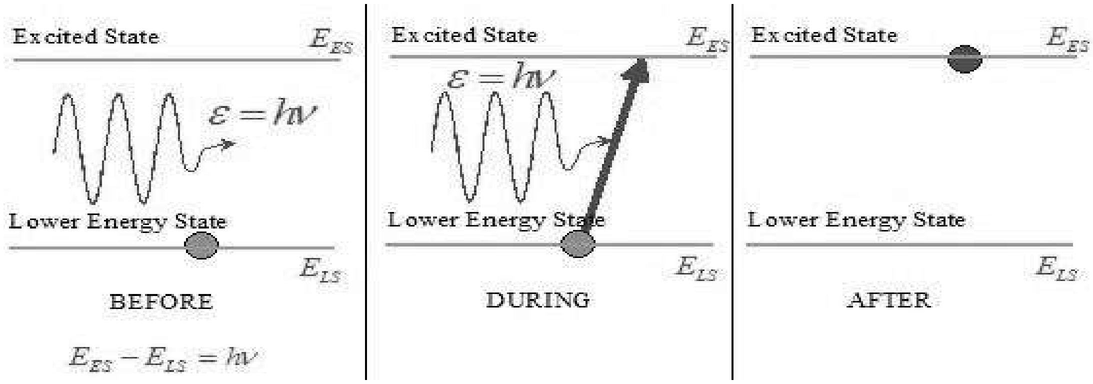
Introduction:

When the incident radiation (Photon) interacts with atoms in the energy levels then three distinct processes can takes place.

- Stimulated absorption of radiation
- Spontaneous emission of radiation
- Stimulated emission of radiation

1] Stimulated absorption of Radiation:

Let us consider a system in which two active energy levels are present whose energies are E_1 and E_2. Usually atoms are in the ground state as long as external forces are not applied. If a photon of energy $(E_1 - E_2 = h\upsilon)$ is incident on the atom in lower state. The atom absorbs incident photon and gets excited to the higher energy state E_2. This is called Stimulated or induced absorption of radiation.



$A + h\upsilon \rightarrow A^*$ A… Atom in the ground state A^*… Atom in higher state

The number of absorptions at any instant will be proportional to the number of density of atoms in lower state E_1 and the photon density (incident radiation density) in the incident beam. The more number of atoms in the lower state can undergo more number of absorption transitions. Similarly, the more number of photons in incident radiation, the more number of atoms can be excited to higher state.

The rate of absorptions R_{12} is proportional to the population of the lower energy level N_1 and to the density of photons in the incident radiation $\rho(\upsilon)$.

R_{12} α N_1

R_{12} α $\rho(\upsilon)$

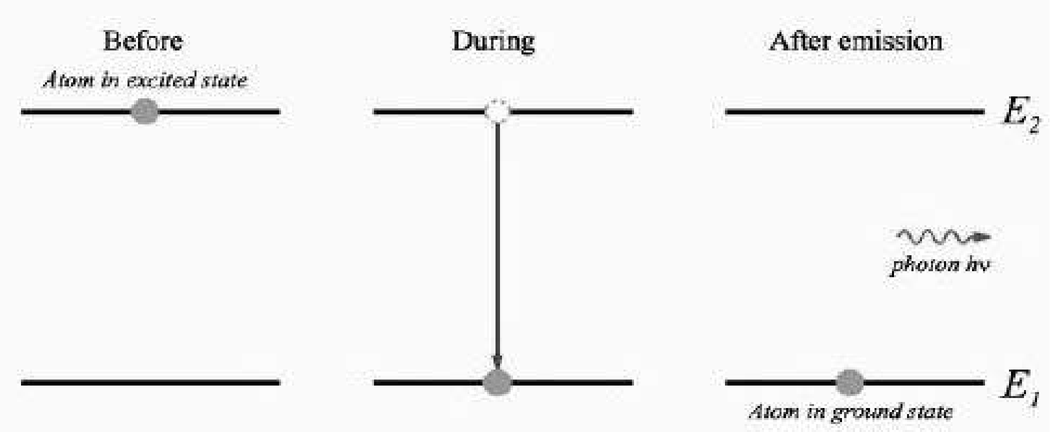
R_{12} α N_1 $\rho(\upsilon)$

$$R_{12} = N_1 \rho(\upsilon) B_{12}$$

B_{12}… Probability of absorption of radiation per unit time (or) Einstein's coefficient of stimulated absorption of radiation.

2] Spontaneous Emission of radiation:

It is the process, in which there is emission of radiation whenever an atom transits from a higher energy state to a lower energy state without external agency". For this process take place an atom initially present in the excited state, since the higher energy level is unstable, the excited atoms in the higher energy level E_2 spontaneously return to lower energy state E_1 with emission of photon of energy $(E_1 - E_2 = h\upsilon)$ without any external agency.



$A^* \rightarrow A + h\upsilon$ A… Atom in the ground state and
A^* … atom in higher state

The instant of transition, the direction of emission of photon, the phase of photon, polarization state of photon are all random quantities. That is different atoms of the medium emit photons at different times and different directions. Hence there is no phase relationship among the emitted photons, so they are incoherent.

EX: Glowing tube light, Electric bulb, Candle flame, etc.,

The light from the conventional sources originates in spontaneous process and is incoherent. It contains a superposition of many waves of random phase.

The rate of spontaneous emission of radiation is proportional to the population at the higher energy level

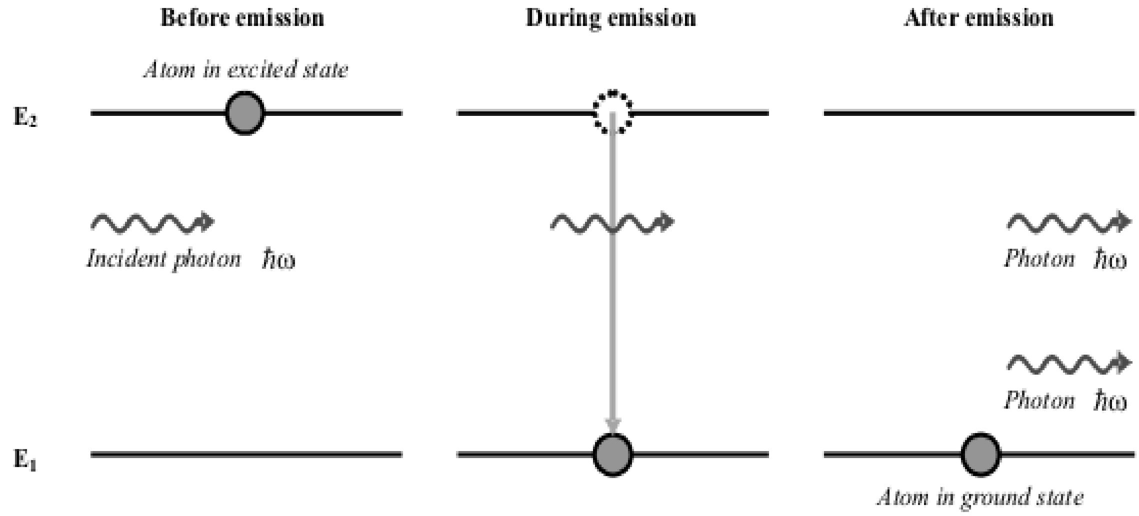
R_{21} α N_2

$$R_{21} = N_2 A_{21}$$

A_{21}…. Probability of spontaneous emission per unit time (or) Einstein's coefficient of spontaneous emission of radiation.

3] Stimulated emission of Radiation:

"It is a process in which there is an emission of photon whenever an atom transmits from higher energy state to lower energy state under the influence of an external agency". For this process also, atom should be already in a excited state. Let a photon having energy $(E_1 - E_2 = h\upsilon)$ interact with an atom in exciting state under such interaction an incident photon stimulates excited state atom in the level E_2 to transmit lower energy state E_1 resulting in the emission of photon of energy $(E_1 - E_2 = h\upsilon)$.



$$A^* + h\upsilon \rightarrow A + 2h\upsilon$$

A… Atom in the ground state and A^* … atom in higher state

Inducing photon and emitted photon have same phase, energy and direction of movement. This type of emission is responsible for laser action. The stimulation emission of radiation is the principle used in the laser action.

The rate (number) of stimulated emission of radiation is proportional to the population at the higher energy state and to the density of inducing photon in incident radiation.

$R_{21} \alpha N_2$ and $R_{21} \alpha \rho(\upsilon)$

$R_{21} \alpha N_2 \rho(\upsilon)$

$$R_{21} = N_2 \rho(\upsilon) B_{21}$$

B_{21} … Stimulated emission of radiation per unit time. (or) stimulated emission of radiation.

Stimulated emissionof Radiation is characterized by some very interesting features:

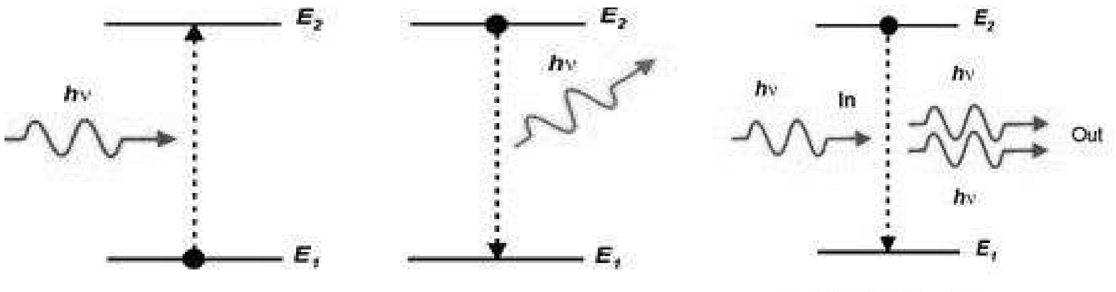
1] The emitted photon is identical to the incident photon. It has same frequency as that of incident photon. It will be in phase with the incident photon. Both the photons travel in same direction. They will be in the same state of outside.
2] The stimulated emission controlled from outside.
3] One photon induces an atom to emit a second photon, these two travelling along the same direction de excite two atoms in their path producing a total four photons which stimulates four atoms generating eight photons and so on. The photons build up in avalanche manner.
4] Electromagnetic waves of extremely high amplitude could be generated by the combined stimulated emission from large samples of atoms.
5] The constructive interference of many waves travelling in same direction with a common frequency and common phase produce an intense coherent light beam.
6] The process of stimulated emission is the key to the operation of a laser.

The external radiation incident on the medium, the three process occur simultaneously. Atoms in the lower level E_1 will occasionally absorb radiation and make a transition to upper level E_2. Atoms in the upper level will occasionally radiate and make a transition to the lower level. In order to maintain N_1 and N_2 constant, the number of upward transitions must be equal to the number of downward transitions. Thus,

$$N_{Absorption} = N_{spontaneous} + N_{Stimulated}$$

➢

➢



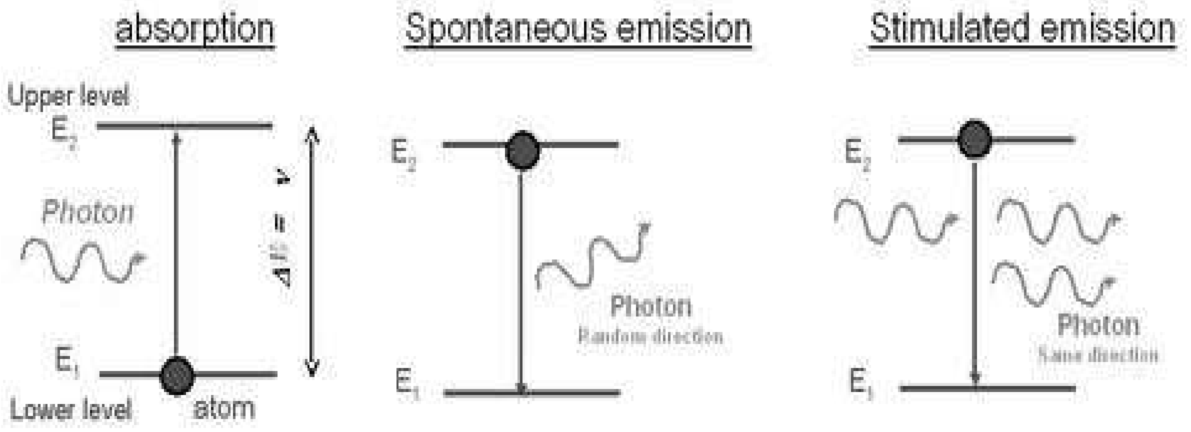
(a) Absorption (b) Spontaneous emission (c) Simulated emission

Difference between spontaneous emission of radiation and stimulated emission of radiation

Spontaneous Emission of radiation	*Stimulated Emission radiation*
1] Spontaneous emission takes place without any stimulus energy.	1] Stimulated emission takes place with the help of stimulus energy
2] It is independent of incident radiation density.	2] It is dependent on density of incident radiation.
3] Spontaneous emission takes place after 10^{-8} Second.	3] Stimulated emission takes place within 10^{-8} second.
4] Emitted radiations are incoherent radiation	4] Emitted radiations are coherent radiation
5] The radiations have low intensity and less directionality.	5] The radiations have high intense and more directionality.
6] Polychromatic radiation.	6] Monochromatic radiation.
7] It is uncontrolled process.	7] It is controlled process.

Einstein Coefficients:

Consider two energy levels E_1 and E_2 of an atomic system such that $E_2 > E_1$. Let N_1 and N_2 be number of atoms per unit volume (population) present at the levels E_1 and E_2 of any radiation of frequency (υ) falls on the atomic system, it can interact with matter in 3-distinct ways.



Absorption:

When the atom makes transition from E_1 to E_2 in the presence of external photon whose energy equal to $(E_1 - E_2 = h\upsilon)$ then the stimulated absorption takes place. The number of stimulated absorption per unit volume per second from level E_1 and E_2

$R_{12} = N_1\rho(\upsilon)B_{12}$ …… [1]

B_{12} …Represent probability or absorption per unit time.

Spontaneous Emission: An atom in level E_2 can make spontaneous transition by jumping into lower energy level E_1. The number of spontaneous emissions per unit volume per second from level levels E_2 to E_1.

$$R_{21} = N_2A_{21} \text{ …… [2]}$$

Stimulated Emission:

When an atom makes transitions from E_2 to E_1 in the presence external photon whose energy equal to $(E_1 - E_2 = h\upsilon)$ stimulated emission takes place. The number of stimulated emission per unit volume per second from levels E_2 to E_1 is given as,

$R_{21} = N_2\rho(\upsilon)B_{21}$ ….. [3]

B_{21} … Represents probability of stimulated emissions per unit volume.

A_{21} … Represents probability of spontaneous emissions per unit volume.

In the equilibrium conditions the number of transitions from E_2 to E_1 must be equal to number of transitions from E_1 to E_2. Thus,

Total number of upward transitions = Total number of downward transitions

$$N_1\rho(\upsilon)B_{12} = N_2\rho(\upsilon)B_{21} + N_2A_{21}$$

$$N_2A_{21} = N_1\rho(\upsilon)B_{12} - N_2\rho(\upsilon)B_{21}$$

$$N_2A_{21} = (N_1B_{12} - N_2B_{21})\rho(\upsilon)$$

$$\rho(\upsilon) = \frac{N_2A_{21}}{(N_1B_{12} - N_2B_{21})}$$

$$\rho(\upsilon) = \frac{N_2A_{21}}{N_2B_{21}\left(\frac{N_1B_{12}}{N_2B_{21}} - 1\right)}$$

$$\rho(\upsilon) = \frac{\left(\frac{A_{21}}{B_{21}}\right)}{\left(\left(\frac{N_1}{N_2}\right)\left(\frac{B_{12}}{B_{21}}\right) - 1\right)} \text{ ………….. [4]}$$

From Boltzmann distribution law $N_1 = N_2e^{\left(\frac{E_2-E_1}{k_BT}\right)}$

$$\frac{N_1}{N_2} = e^{\left(\frac{E_2-E_1}{k_BT}\right)}$$

$$\frac{N_1}{N_2} = e^{\left(\frac{h\upsilon}{k_BT}\right)} \text{ ……………. [5]}$$

From [4] and [5] ,we have

$$\rho(\upsilon) = \frac{\left(\frac{A_{21}}{B_{21}}\right)}{\left(\left(\frac{B_{12}}{B_{21}}\right)e^{\left(\frac{h\upsilon}{k_BT}\right)} - 1\right)} \text{ …………. ……………………….[6]}$$

According to Planck's radiation formula $\rho(\upsilon) = \left(\frac{8\pi h\upsilon^3}{C^3}\right)\left(\frac{1}{e^{\left(\frac{h\upsilon}{k_BT}\right)} - 1}\right)$ ……….. [7]

Compare [6] and [7] ,we get

$$\frac{A_{21}}{B_{21}} = \frac{8\pi h\upsilon^3}{C^3} \quad \textbf{and} \quad \frac{B_{12}}{B_{21}} = 1 \quad \frac{B_{12}}{B_{21}} = 1 \Rightarrow B_{12} = B_{21}$$

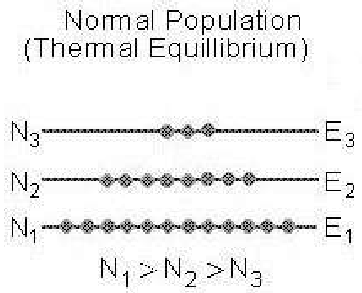
The probability of stimulated emission is same as induced absorption.

$\frac{A_{21}}{B_{21}} \alpha\ \upsilon^3$. The ratio of spontaneous emission and stimulated emission is proportional to υ^3. This shows that the probability of spontaneous emission increases rapidly with energy difference between two states.

Population

"The number of atoms per unit volume in an energy levels are called as population of that energy level. (OR) The number of active atoms occupying an energy state is called Population of that state."

N_1 And N_2 are the populations of the lower energy state E_1 and upper energy state E_2 respectively. In a state of thermal equilibrium, the populations of energy levels E_1 and E_2 are fixed by Boltzmann factor.



The population ratio is given by

$$\frac{N_2}{N_1} = e^{-\left(\frac{E_2-E_1}{kT}\right)} = e^{-\left(\frac{h\upsilon}{kT}\right)}$$

(OR)

$$\frac{N_1}{N_2} = e^{\left(\frac{E_2-E_1}{kT}\right)} = e^{\left(\frac{h\upsilon}{kT}\right)}$$

The negative exponential indicates $N_1 << N_2$ at equilibrium. It means more atoms are in lower energy state E_1. This state called Normal state.

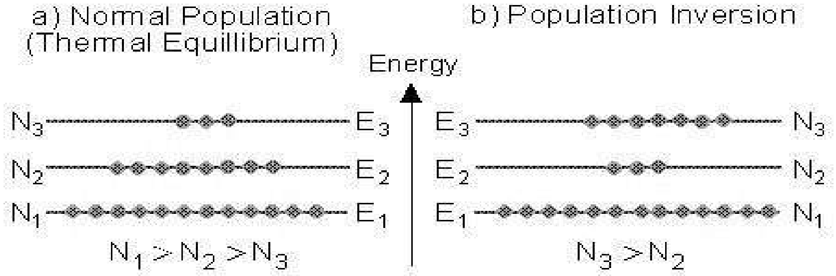
Population inversion

In thermal equilibrium state more atoms in lower energy level than in the upper level. In order to achieve stimulated emission there must be more atoms in the upper level than in lower level. Therefore, a non equilibrium state is to be produced in which the population of the upper energy level exceeds the population in the lower energy level. When this situation occurs, the population distribution between levels E_1 and E_2 is said to be inverted, and the medium is said to have gone into the state of population inversion.

The population inversion is sometimes called as a negative temperature state that means the population inversion is a non-equilibrium state.
In practice the state of population inversion is obtained at ordinary temperature.

Consider 3-energy states in which 3-active energy levels E_1, E_2 and E_3 population in these energy states are, N_1, N_2 and N_3 respectively.

In normal condition $E_1 < E_2 < E_3$ and $N_1 > N_2 > N_3$. E_1 is the ground state, its life time is unlimited. E_3 Is the highest energy level its life time is very less and have is most unstable state. Whereas E_2 is excited state and have more life time. Hence E_2 is metastable state.



When suitable form of energy is supplied to the system then the atoms excited from ground state E_1 to excited states E_2 and E_3. Due to instability, excited atoms will come back to ground state after the life time of respective energy states E_2 and E_3. If this process continued atoms will excite continuously to E_2 and E_3. E_3 is unstable state atoms will fall into E_2 immediately, at the stage the population in E_2 will becomes more than the population in ground state. This situation is called as ‘population inversion’.

<u>Definition of Population inversion :</u>

The stage of making population of the higher energy level to be greater than the population of lower energy level is known as ‘population inversion’.

<u>Definition of Pumping:</u>

The process of sending atoms from lower energy level (E_1) to higher energy level (E_2) to get population is called ‘Pumping’.

<u>Conditions for population inversion:</u>

1] The system should possess at least a pair of energy levels, separated by an energy equal to the energy of photon.

2] There should be a continuous supply of energy to the system such that the atoms must be raised continuously to the excited state.

<u>3-Level pumping Scheme:</u>

The state E_1 is the ground state and E_2 and E_3 are the excited states. When the medium is exposed to radiation of frequency υ a large no. of atoms will be excited to the higher energy level E_3. The E_3 level is not a stable state. Some of the atoms make spontaneous transition to the lower energy state E_1 but many of them make spontaneous transition to the metastable level E_2 through non-radiative transition.

The spontaneous transition from E_2 to E_1 occur rarely, the atoms trapped in the state E_2. The process continuous because of pumping and after same time there will be a large accumulation of atoms at E_2.

More than half of the ground state atoms accumulate at E_2 the population inversion id achieved between the states E_1 and E_2. Now, a photon of energy [$E_2 - E_1 = h\upsilon$] can tiger stimulated emission of atoms at E_2 very high pump power is require in this type of pumping scheme because to achieve population inversion more than half of the ground state atoms pumped to the upper state.

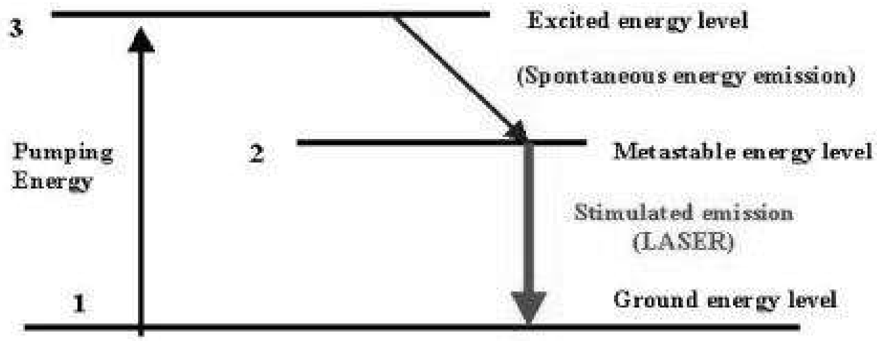


Fig. 3a: *Energy States of Three – level Active Medium*

4- Level pumping scheme:

Pump frequency lifts the active centers from the ground level E_1 to the upper most level E_4. From the pump level E_4 the atoms rapidly fall to the metastable state E_3. The population at this state goes rapidly while the level E_2 is virtually empty. Therefore the population is inversion is achieved between the states E_2 and E_3. The photon of energy $E_3 - E_1 = h\upsilon$ on start a chain of stimulated emission. Hence the atoms into the state E_2 from there atoms undergo non radiative transition subsequently to the ground state E_1 and will be available one to participate in this process.

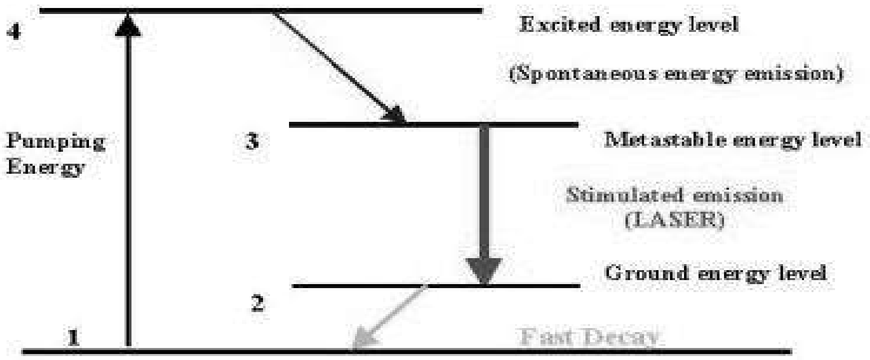
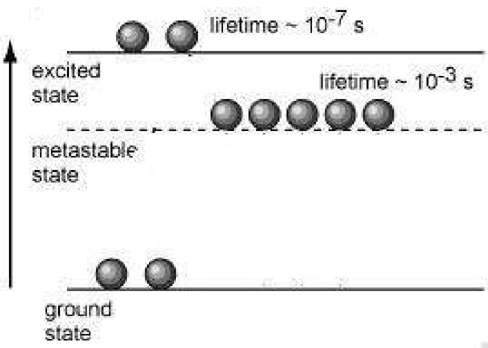


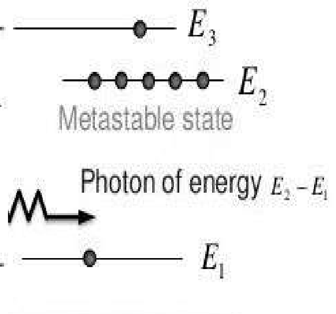
Fig. 3b: *Energy States of Four – level Active Medium*

Meta Stable State

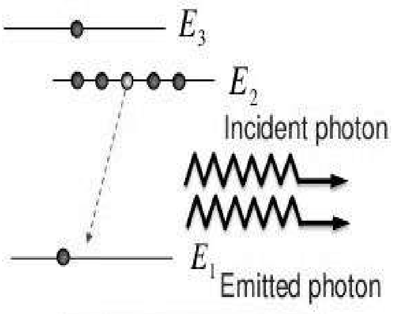
An atom can be excited to a higher level by supplying pumping energy; the excited atoms have short life time and release their energy in a matter of 10^{-9} sec though continuous emission. The atoms do not stay long enough at the excited state. Even though the pumping agent continuously raises the atoms to the excited level, they undergo spontaneous excitations and rapidly return to the lower energy level. Population inversion cannot be established in such circumstances. In order to establish population inversion the excited atoms are required to wait at the upper energy level till a large number of atoms accumulated at that level. It is necessary that the excited state has a large life time.

1] A Meta stable state is such state, atoms excited to meta stable state remains excited for appreciable time, which is of the order of $10^{-6} s$ (or) $10^{-3} s$. This is $10^{3} s$ or $10^{6} s$ times the life time of the ordinary energy level.





Metastable system



Stimulated emission

Excitation Mechanism (Methods to achieve Population inversion)

For laser action to takes place, we need the population inversion between the available energy levels of a system .To obtain population inversion ,the atoms in the lower energy level should be excited to higher energy levels by supply of additional energy in the form of excitation mechanism. The generally used various excitation mechanisms are detailed below.

(1) Optical Pumping :

In optical pumping, a light source is used to supply luminous energy to excite the atoms to higher energy levels to create population inversion for further laser emission of radiation. This type of pumping is widely used in solid state laser i.e., Ruby Laser and YAG: Nd Laser etc.

(2) Electric discharge:

In this method ,atoms are excited by collision with the fast moving accelerated electrons in an electric discharge. This mechanism is well suited for gaseous ion lasers. The electric field (kV/m) between the cathode and anode of discharge tube causes the emitted electrons by the cathode to be accelerated towards the anode. These electrons colloid with the gaseous atoms, ionize the gas and raise them to higher energy levels causing population inversion. The best example is Argon ion Laser.

3. Inelastic atom-atom collisions:

When the gaseous medium consists of two different atoms, then one atom is excited to the corresponding higher energy level by electric discharge. This excited atom colloid with the different atoms in the medium elastically transferring its energy to it and pumping it to the equivalent its higher energy levels. Population inversion is caused by the inelastic collisions between two various atoms of the gaseous medium. The best example is He-Ne Laser.

$$A + e \rightarrow A^* + e1$$

$$A^* + B \rightarrow A + B$$

Here A and B are two different gases.

4. Chemical Reaction:

In this method, the chemical energy released during the chemical process will excite the atoms to higher energy levels causing population inversion in the lasing system. For example, Hydrogen can react with Fluorine liberating heat energy. This heat energy will try to excite the atoms to the higher energy level.

5. Direct Conversion:

When a p-n junction is forward biased, then recombination of electron- hole pair across the junction emits radiations. In this process, the applied electric energy promotes the emission of radiation. The best example is Semiconductor Lasers.

Block diagram of a laser system:

The block diagram of laser system contains three parts, they are

(i) Source of energy

(ii) Active medium and

(iii) Optical resonator

(i) Source of energy:

To achieve population inversion suitable form of energy must be supplied. It supplies suitable form of energy by using any one of the pumping methods.
For example in ruby laser, the pumping source is optical pumping that is helical xenon flash tube. In helium-neon laser, electrical pumping that is electric discharge is used.

(ii) Active medium:

To achieve population inversion medium is necessary The material medium in which population inversion takes place is called as active medium. In which metastable state is present. In metastable state, only the population inversion takes place. It can be a solid, liquid, gas or semiconductor diode junction. The material medium in which the atoms are raised to exited state to achieve population inversion is called as active centers.

For example in ruby laser, the active medium is Aluminum oxide (Al_2O_3) doped with Chromium oxide (Cr_2O_3). In which chromium ions (Cr^{3+}) act as active centers. In helium -neon laser it is the combination of helium and neon in the ratio of 10:1 in which Ne atoms act as active centers.

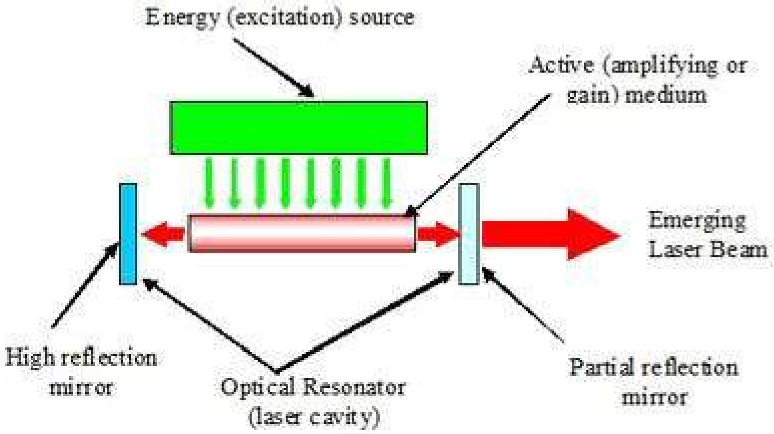


Fig: Components of LASER system.

Optical resonator or optical cavity:
An optical resonator which consists of two mirrors. One mirror is fully reflective and other is partially reflective. An active medium is kept between in them. The light emitted due to the Stimulated emission of radiation bounces back and forth between the two mirrors and hence the intensity of the light is increased enormously. Finally the intense, amplified beam called laser is allowed to come out through the partial mirror as shown in fig.

Different types of laser systems:

On the basis of active medium used in the laser system, lasers are classified into several types, and most popular methods are:
1] Solid state lasers (YAG-Nd Laser, Ruby laser)
2] Liquid lasers (Europium laser)
3] Gaseous lasers (He-Ne laser)
4] Dye lasers (Coumarium laser)
5] Semiconductor lasers (GaAs laser)
Most lasers emit light in the RED or IR region. Lasers can be operated in a continuous wave mode or in a pulsed mode with a higher output power.

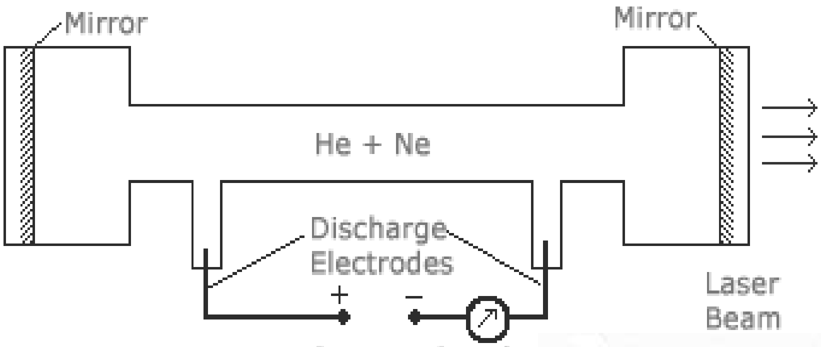
He-Ne Laser

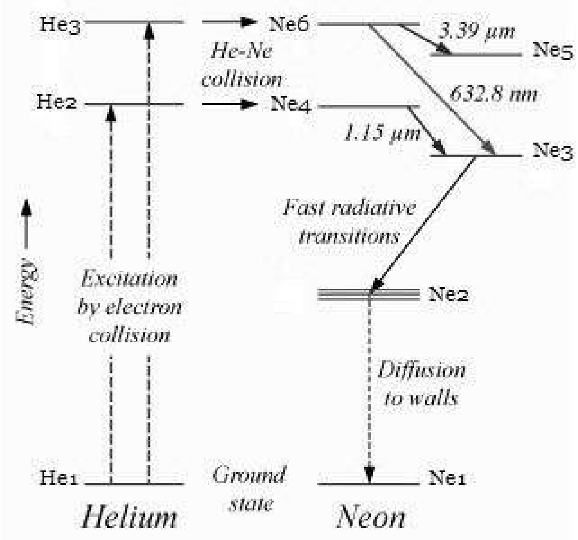
- He-Ne Laser is a gaseous laser system and is used to produce a continuous laser.
- This laser is highly directional, monochromatic, coherent and stable.
- The output of gas lasers is moderate and is generally few mill watts.
- The gas lasers are operating continuously with need of cooling.

- The first laser successfully operated is He-Ne gas laser.

Construction:

The He-Ne gas laser consists of fused quartz tube filled with mixture of neon under the pressure of 0.1mm of mercury and Helium under pressure of 1mm of mercury. The ratio of He-Ne is 10:1 hence the number of Helium atoms is greater than neon atoms. The output power from laser depends upon the length of the discharge tube and pressure of gas mixture.





At one end of the tube, there is a perfect reflector while on the other end is partially reflector. The active medium in this laser is excited by a high frequency generator.

Working:

In He atom three energy levels are present, they are named as He_1, He_2 and He_3. In Neon atom six active energy levels, they are named as Ne_1, Ne_2, Ne_3, Ne_4, Ne_5 and Ne_6.

Here, it should be noted that He_2 and Ne_4 have same energy and life time similarly He_3 and Ne_6.

When a discharge is passed through the gas, the electrons accelerated towards the positive electrode. During their passage they collide with He atoms and excite them into upper states labeled He_2 and He_3. These are Meta stable states in Helium atoms.

Thus the He atoms remain in these energy levels for sufficiently long time. Now these atoms interact with neon atoms, which are in ground state.

The interaction excites the Neon atoms to their metastable states labeled as Ne_4 and Ne_6 while helium atoms return to their ground state. As the energy exchange is continuously between He and Ne atoms, the population of neon atoms in the excited states Ne_4 and Ne_6 increases more and more. At the stage population inversion will be achieved in the Meta stable states Ne_4 and Ne_6.

After achieving population inversion:

1] Few Neon atoms de excite from Ne_6 to Ne_5. During this transition electromagnetic radiation of wavelength of 3390 Å will be emitted.

2] Many other Neon atoms de excite from Ne_6 to Ne_3. During this transition electromagnetic radiation of wavelength 6328 Å will be emitted. This is important and major wavelength in this laser system.

3] In continuation atoms in the excited state Ne_4 de excite to Ne_3, where an electromagnetic radiation of wavelength 1150 Å is emitted.

4] After reaching all the Neon atoms to Ne_3, spontaneously those will de excite to *Ne*. During this transition, an electromagnetic radiation of wavelength 6000 Å will be emitted.

5] Finally Neon atoms take non radiative transition by making collisions with walls of the tube from Ne_2 to Ne_1 .

He-Ne Laser - Characterstics

1] Type:	Gas laser
2] Active medium:	He-Ne gas mixture
3] Active centres:	Ne - Atoms
4] Pumping mechanism:	Electric discharge
5] Output:	100mW
6] Nature of output:	continuous wave
7] Wavelength emitted:	6328 Å

Advantages of He-Ne Laser

1.It is a continuous Laser which emits high monochromatic and high directional laser light when compared to solid state Lasers.

2. Due to end window set at Brewester's angle, we get linearly polarized Laser.

GaAs semiconductor laser:-

The GaAs laser was constructed by Hall.

Characteristics:-

Type	: Semiconductor laser
Active medium	: P-N junction diode
Active centre	: Re combination of electrons and holes
Pumping method	: Direct pumping
Optical resonator	: Junction of Diode ends Polished
Nature output	: pulsed (or) continuous waveform
Power output	: 1mw
Wave length	: 8200Å -9000 Å
Band gap	: 1.44 ev

Active medium:The active medium in GaAs laser is GaAs. But it is also commonly said that depletion region is the active medium in semiconductor laser. The thickness of the depletion layer is usually very small (0.1 μm).

Pumping Source: Forward biasing is used as pumping source. The p-n junction is made forward biased that is p side is connected to positive terminal of the battery and n side to negative terminal. Under the influence of forward biased electric field, conduction electrons will be injected from n side into junction area, while holes will enter will enter the junction from the p side. Thus, there will again be recombination of holes and electrons in depletion region and thus depletion region becomes thinner.

Optical resonator system: The two faces of semiconductor which are perpendicular to junction plane make a resonant cavity. The top and bottom faces of diode, which are parallel to junction plane, are metalized so as to make external connections. The front and back faces are roughened to suppress the oscillations in unwanted direction.

Principle: - When a p-n junction is formed across a p and n-type semiconductor, then it results in the formation of depletion region across the junction. When the junction is forward biased, the width of depletion region decreases allowing more number of electrons from n-type to across the junction and recombine with holes in p-type. This, recombination of hole pairs across the junction emits the radiation (as shown in figure 7.25).

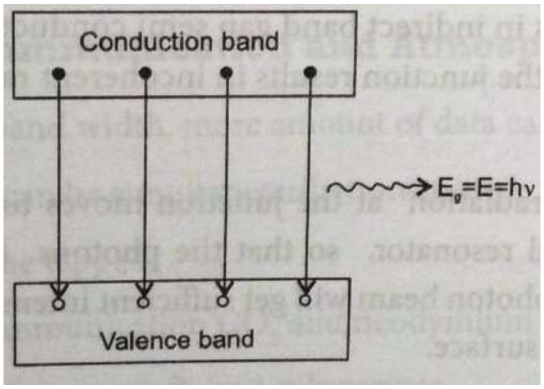
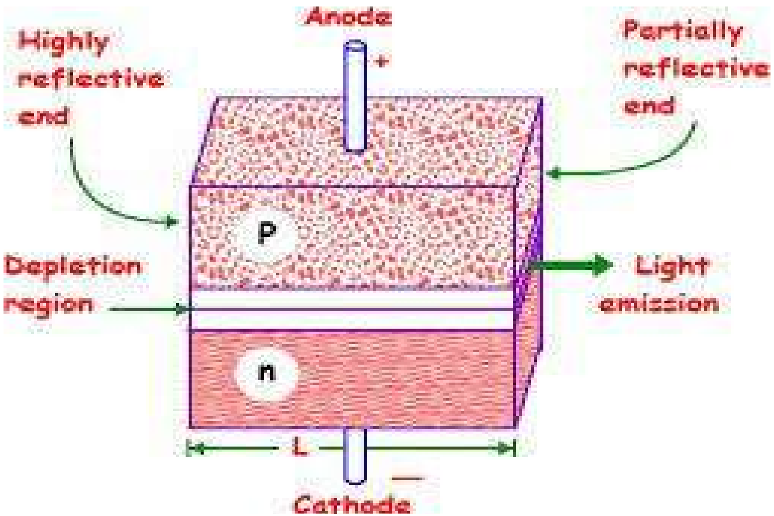


Figure: Principle of semiconductor laser

Construction:-

- A typical GaAs laser is shown in fig.
- A rectangular block of Ga-As semiconductor is converted into p and n-type by proper doping of impurities into the block. The upper region acts as p-type and the lower portion as n-type .Between these two regions, we have a p-n junction. To achieve population inversion p and n- regions are heavily doped with the impurities. The p-n junctions act as active



medium. The two faces of the block, one fully polished and the other partially polished act as an optical resonator (or) cavity.

Figure: Construction of semiconductor diode Laser.

Working:

When the p-type is connected to the positive terminal of the battery and the n-type is connected to the negative terminal then the p-n junction will be in forward biased condition, then there will be injection of electrons into the conduction band along n-side and production of more holes in valence band along p-side of the junction. Thus, there will be more number of electrons in conduction band comparable to valence band, so population inversion is achieved.
Therefore, when the electrons and holes are injected into the junction region from opposite sides with forward biasing, then population inversion is achieved between levels near the bottom of the conduction band and empty levels near the top of the valence band.

When electrons recombine with the holes in junction region, then there will be release of energy in the form of photons. This release of energy in the form of photons happens only in special types of semiconductors like Gallium Arsenide (GaAs). Otherwise in semiconductors like silicon and germanium, whenever holes and electrons recombine, energy is released in the form of heat, thus Si and Ge cannot be used for the production of laser.

The spontaneously emitted photon during recombination in the junction region of GaAs will trigger laser action near the junction diode. The photons emitted have a wavelength from 8200 Å to 9000 Å in the infrared region.

Calculation of wavelength:

Band gap of GaAs =1.44 ev (1ev=1.6×10^{-19} J)

$$E_g = h\upsilon = h\frac{c}{\lambda}$$

$$\lambda = \frac{hc}{E_g} = \frac{6.625\times10^{-34}\times3\times10^{8}}{1.44\times1.6\times10^{-19}}$$

$$\lambda = 8626\,\AA$$

The wavelength is near IR region.

Advantages:-

- It is easy to manufacture
- The cost is low.
- The efficiency of GaAs laser is high.

Disadvantages:

- It produces low power output.
- The beam has large divergence.

Nd:YAG Laser:

- Nd-YAG laser is one of the most popular types of solid state laser. It is a four level laser.
- Yttrium aluminum Garnet ($Y_3Al_5O_{12}$), commonly called YAG is an optically isotropic crystal.
- Some of the Y^{3+} ions in the crystal are replaced by neodymium ions, Nd^{3+}.
- Doping concentrations are typically in the order of 0.725% by weight.
- The crystal atoms do not participate in the lasing action but serve as a host lattice in which the active centers, namely Nd^{3+} ions reside.

Construction:

- Figure illustrates a typical design of Nd:YAG laser. The system consists of elliptically cylindrical reflector housing the laser rod along one of its focus line and a flash lamp along the other focus line.
- The light leaving one focus of the ellipse will pass through the other focus after reflection from the silvered surface of the reflector. Thus the entire lash lamp radiation gets focused on the laser rod.
- The YAG crystal rods are typically of 10cm in length and 12mm in diameter. The two ends of the laser rods are polished and silvered and constitute the optical resonator.

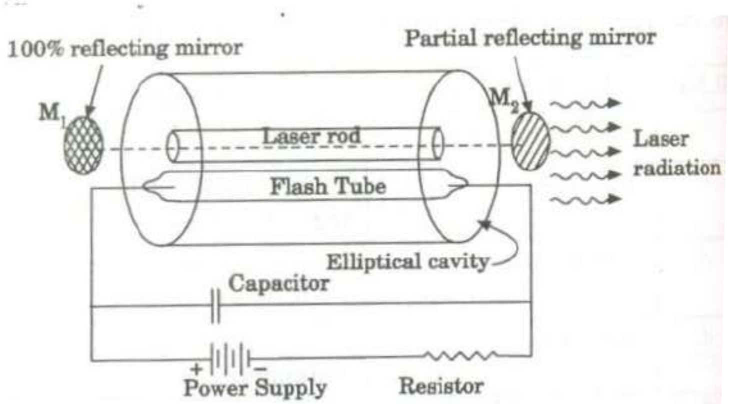


Figure: Construction of Nd:YAG Laser

Working:

A simplified energy level diagram for the neodymium ion in YAG crystal is shown in figure. The energy level structure of the free neodymium atom is preserved to a certain extent because of its relatively low concentration. However the energy levels are split and the structure is complex.

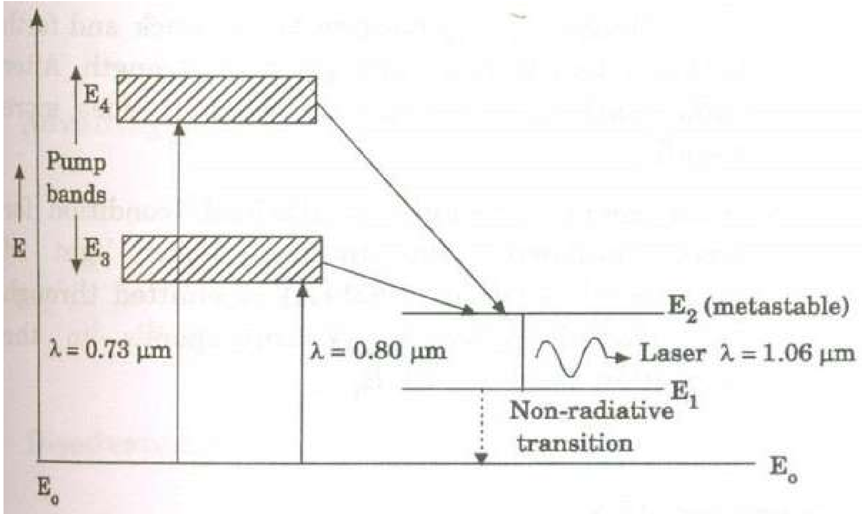


Figure: Energy level diagram for the neodymium ion in YAG crystal

The Pumping mechanism:

- When the krypton flash lamp is switched on, the Nd^{3+} ions are excited to the upper energy band E_4 and E_5.
- The Nd^{3+} ions make a transition from these energy levels to level E_3 by non- radiative transition.E_3 is metastable state.
- The metastable level E_3 is the upper laser level.

Population Inversion:

- The upper laser level E_3 will be rapidly populated, as the excited Nd^{3+} ions quickly make downward transitions from the upper energy bands.
- The lower laser level E_2 is far above the ground level and hence it cannot be populated by Nd^{3+} ions through thermal transitions from the ground level.
- Therefore the population inversion is readily achieved between the E_3 level and E_2 level.

Lasing:

- A chance photon is produced when an Nd^{3+} ion makes a spontaneous transition from E_3 level to E_2 level. This spontaneous photon stimulates another excited atom to make a downward transition.
- This stimulated photon and initial photon trigger many excited atoms to emit photons.
- Photons thus generated travel back and forth between the two end mirrors and gain in strength very rapidly.
- On attaining sufficient energy, the laser beam emerges out through the partially reflecting mirror.
- The laser emission occurs in infrared (IR) region at a wavelength of about 10,600 A^0 (1.6 μm).
- The Nd^{3+} ions return to the ground state E_1 from the lower lasing level E_2 on their own through non-radiative transitions.

Salient Features:

- Uses four level pumping schemes.
- The active centers are Nd^{3+} ions.
- Light from a Xenon or Krypton flash lamp is the pumping agent.
- Low efficiency and moderate power output.
- Operates in CW/pulse mode.

Application of Laser :

Industry :

(1) With increased power output, lasers can be used as a welding tool. Dissimilar metals can be welded using lasers with minimum distortions. Lasers are used to cut glass and quartz.
(2) Lasers are used to drill holes in ceramics.
(3) Lasers are used to drill aerosol nozzles and control orifices within the required precision.
(4) Lasers are used for heat treatment in the tooling and automotive industry.
(5) Lasers are used in electronic industry in trimming the components of ICs
(6) In plastic industry, polymers are obtained by irradiating monomers by lasers.

Medicine:

(1) Ophthalmologists use laser for attaching the retina in retinal-detachment cases.
(2) Lasers are used for cataract removal.
(3) Lasers are used for eye lens curvature corrections.
(4) Lasers are used for bloodless surgery.
(5) Lasers are used in angioplasty for removal of artery block.
(6) Lasers are used in cancer diagnosis and therapy.
(7) Lasers are used in destroying kidney stones and gallstones.
(8) Lasers are used in plastic surgery, skin injuries and to remove moles and tumours developed in skin tissue.
(9) Lasers are used in the treatment of mouth diseases.
(10) Lasers are used in the treatment of liver and lung diseases.
(11) Laser Doppler velocimetry is used to measure blood velocity in blood vessels.

Scientific Field :

(1) Lasers are used for isotope separation.
(2) Lasers are used in recording and reconstruction of a hologram.
(3) Lasers are used to create plasma.
(4)Lasers are used to produce chemical reactions.
(5) Lasers are used to study the internal structure of microorganisms and cells.
(6) Lasers are used in air pollution, to estimate the size of the dust particles.
(7) Lasers are used to develop hidden fingerprints and to clean delicate pieces of art.

Unit-II

chapter-2

Fiber optics

Introduction:

The developments in the fields of communication and information technology demand very easy and transmission of data over longer distances. Fiber optic technology is increasingly replacing wire transmission lines in communication systems and is expected to be as common as electrical wiring even in our vehicles and house very shortly.

Optical fiber lines offer several important advantages over wire lines. Optical fibers are light equitant of microwave guides with the advantage of very high bandwidth and hence very high information carrying capacity .Though at the beginning fiber optic communication systems were more expensive than equivalent wire or radio system, now the situation has changed very much. Fiber optic systems have become competitive with other systems in price and eventually started replacing them.

Fiber optics is a branch of optics which deals with the study of propagation of information in the form of light (rays or modes) through transparent dielectric optical fibers.
The term optical fiber was first coined by N.S.Kapany.

Optical fiber:

The word fiber comes from Latin *fibra* which means a thin thread like piece of material.
Therefore; Optical fiber means a thin thread like piece of visible material.
Optical fiber is a thin and transparent guiding dielectric medium or material which guides or transmits the information as light waves, using principle of total internal refection.

Definition of an optical fiber

Optical fiber is a thin and transparent guiding medium or material which guides the information carrying light waves.

Principle of optical fiber:

An optical fiber works on the principle of total internal reflection.
John Tyndall observed that the propagation of light through the optical fiber will be in the form of multiple total internal reflections.

Definition of Total internal reflection:

when a light ray travels from denser medium to rarer medium and angle of incidence is greater than the critical angle, then the light ray reflects totally, this phenomenon is known as total internal reflection.

Explanation:

Consider a light ray passing from a denser medium of refractive index (n_1) into a rarer medium of refractive index (n_2) as shown in fig. Assuming that $n_1 > n_2$ and that the angle of incidence and refraction with respect to normal to the interface are **i** and **r** respectively.

Then according to Snell's law, $n_1 \sin i = n_2 \sin r$………………… (1).

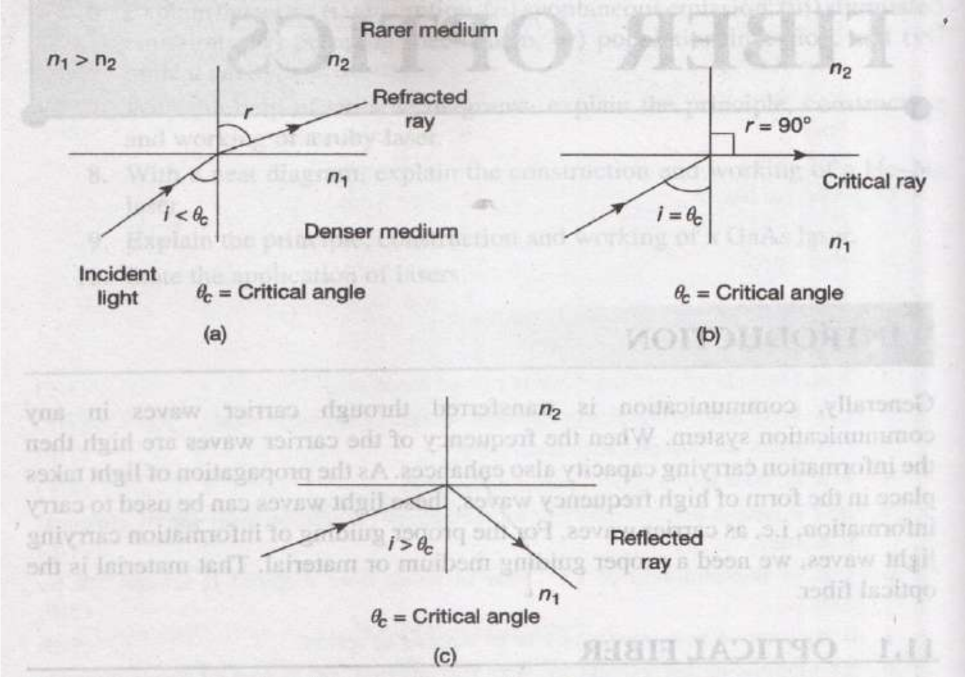


Fig: (a) Normal refraction (b) Critical angle (c) Total internal reflection.

The refracted ray bends away from the normal as it travels from denser medium rarer medium with increase of angle incidence.
As $n_1 > n_2$, if we increase the angle of incidence **i,** the angle of refraction r will go on increasing until a critical situation is reached, when for a certain value of **i=** $\boldsymbol{\theta_c}$, **r** becomes $\frac{\pi}{2}$ and refracted ray passes along interface. This angle **i =** $\boldsymbol{\theta_c}$ is called critical angle [see fig (b)].
If angle of incidence I is further increased beyond $\boldsymbol{\theta_c}$, the ray is no longer refracted but is reflected back into the same medium [see fig (c)].this is called total internal reflection.

From the equation (1); $n_1 \sin i = n_2 \sin r$.

When $i = \theta_c$ then $r = 90^0$.

Therefore; $n_1 \sin \theta_c = n_2 \sin 90^0$.

$n_1 \sin \theta_c = n_2$.

$\sin \theta_c = n_2/n_1$

$$\theta_c = \sin^{-1}(\frac{n_2}{n_1})$$

If the rarer medium is air, then $n_2=1$.

$$\theta_c = \sin^{-1}(\frac{1}{n_1} ¿$$

Conditions for total internal reflection:

1. The light ray should move from denser to rarer medium.
2. When $i < \theta_c$ then the light ray refracts into rarer medium.
3. When $i = \theta_c$ then the refracted light ray passes along interface of the two media.
4. When $i > \theta_c$ then the light ray is reflected back into the denser medium and we get total internal reflection.

Optical fiber structure and construction:

A typical structure of optical fiber as shown in fig.

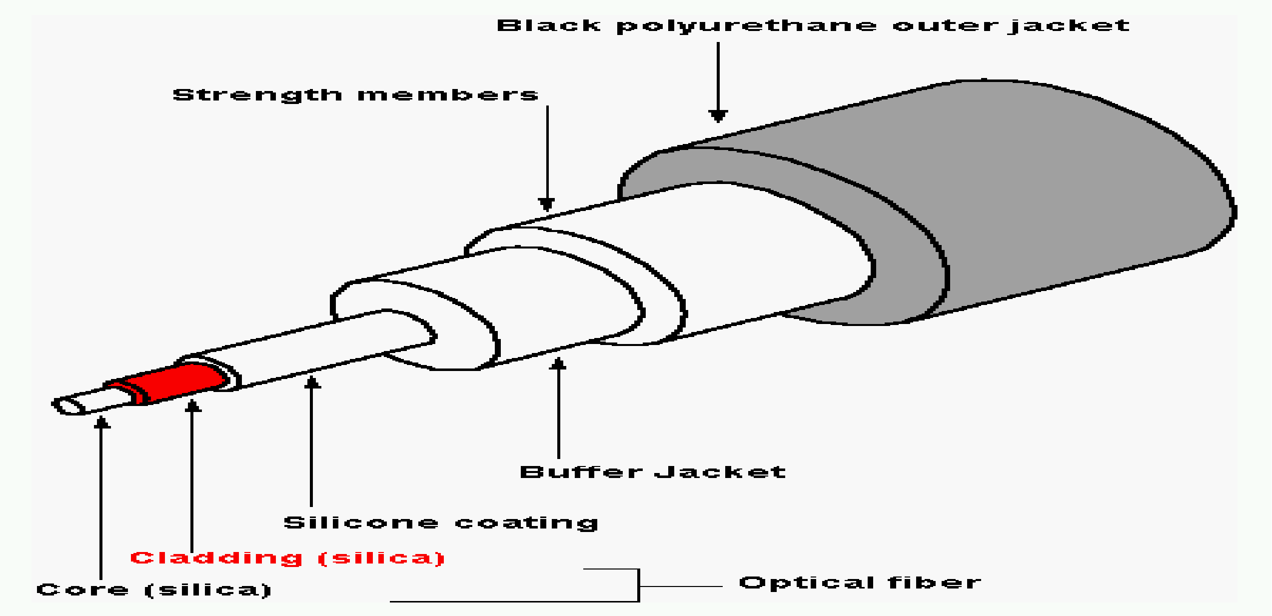


Fig: Structure of optical fiber.

The optical fiber mainly consists of the following six parts as shown in fig.

1. Core and cladding:

An optical fiber consists of a central cylindrical material layer called core with high refractive index n_1, surrounded by a second cylindrical material layer called cladding with a lower of refractive index n_2. ($n_1 > n_2$)

The core is the inner part of the fiber, which guides or transmits the light and cladding keeps the light waves within the core because the refractive index of the cladding is less than that of the core. (Core acts as denser medium and cladding act as rarer medium). The core and cladding are made of either plastic or glass.

2. Silicon coating: it is provided between cladding and buffer- jacket in order to improve the quality of transmission of light.

3. Buffer jacket: it is covered over silicon coating which is made up of plastic material and protects the fiber from moisture and abrasion.

4. Strength members: this layer is arranged over the buffer jacket to provide necessary toughness and tensile strength to the fiber.

5. Black polyurethane outer jacket: Finally a block polyurethane outer layer is provided to avoid damages during hard pulling, bending, stretching or rolling of the fiber in the real field.

Acceptance angle and acceptance cone:

The maximum angle at which the light can suffers total internal refection is called as acceptance angle.
The acceptance cone is derived by rotating the acceptance angle about the fiber axis.

Explanation:

Let us consider a cross-sectional view of an optical fiber. It consists of core of refractive index n_1 and cladding of refractive index n_2 respectively.
Let n_o be the refractive index of the medium (air) in which the optical fiber is placed. The incident ray travels along AO and enters the core at an angle θ_i to the fiber axis. The ray is refracted along OB at an angle θ_r in the core as shown in fig. In core it travels along AB and is incident at point B on core-cladding interface. Let θ be angle of incidence at B.
When θ is greater than the critical angle θ_c then the total internal reflection takes place into core and light takes the path BD. Due to multiple total internal reflections the propagation of light ray takes place through the fiber.

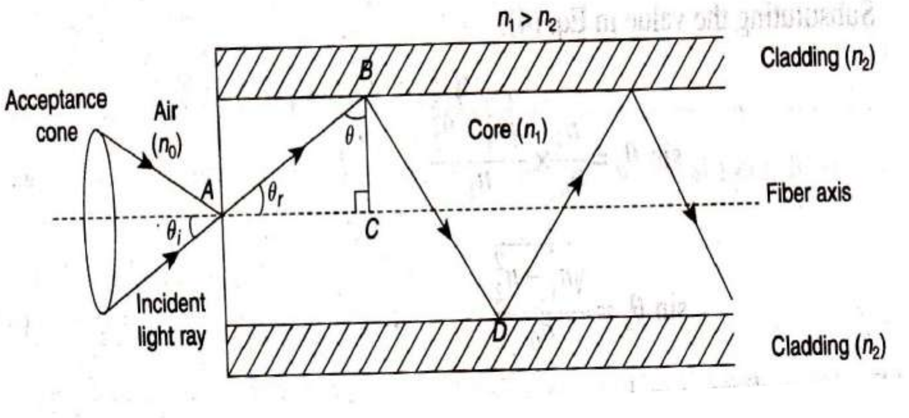


Fig: Cross-sectional view of optical fiber

Applying Snell's law at a (core-air interface)

$$n_0=\frac{1}{\sin \Theta_i}$$

$$n_1=\frac{1}{\sin \Theta r}$$

$$\frac{n1}{n0}=\frac{\sin \Theta i}{\sin \Theta r}$$

$$n_0 \sin \Theta i = n_1 \sin \Theta r \text{ ....(1)}$$

Let a normal BC drawn from the point B to fiber axis. Then the from ΔABC, we get

$$\Theta r+\Theta = 90^0$$

$$\Theta r = 90^0 - \theta \text{ ………………………(2)}$$

Substituting the above value in equation (1)

$$n_0 \sin \Theta i = n_1 \sin ¿90^0-\Theta)$$

$$n_0 \sin \Theta i = n_1 \cos \Theta \text{ …………………………… (3)}$$

To get toatl internal reflection at point b(core-cladding interface) the incident angle Θ should be greater than or equal to Θc(critical angle).

Let the maxium angle of incidence at point A be $\Theta = \Theta a$ for which $\Theta \geq \theta c$.

From equation (3),we get,

$$n_0 \sin \Theta a = n_1 \cos \theta_c$$

$$\sin \Theta a = \frac{n1}{n0} \cos \theta_c \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

$$\text{But} \qquad \sin \theta_c = \frac{n2}{n1}$$

$$\cos \theta_c = \sqrt{1 - \sin^2 \Theta c} = \sqrt{1 - \frac{n_2^2}{n_1^2}}$$

$$= \sqrt{\frac{n_1^2 - n_2^2}{n_1^2}}$$

$$= \frac{\sqrt{n\square_1^2 - n_2^2}}{n_1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Substituting the values in equation (4)

$$\sin \Theta_a = \frac{n1}{n0} \text{ x } \frac{\sqrt{n n_1^2 - n n_2^2}}{n1}$$

$$\sin \Theta_a = \frac{\sqrt{n n_1^2 - n n_2^2}}{n_0} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

For air medium $n_0=1$

$$\sin \Theta a = \sqrt{n n_1^2 - n n_2^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

$$\Theta a = \sin^{-1} \sqrt{n n_1^2 - n n_2^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(8)$$

In the above expression, Θa is the maximum angle of incidence of light at the core and light can suffers total internal reflection.

Thus the maximum angle at which the light can suffers total internal refection is called as acceptance angle.

Acceptance cone:

Rotating the acceptance about the fiber axis, we get an incident light cone with semi- vertical angle asΘa . This incident light cone at core of an optical fiber will be accepted by the fiber for guidance through it, and is known as acceptance cone.

Numerical aperture (N.A):

Numerical aperture represents the light gathering power of an optical fiber. It is a measure of the amount of light that can be accepted by a fiber. The value of NA ranges from 0.13 to 0.50.

A large NA implies that a fiber will accept large amount of light from the source.

Numerical aperture is proportional to acceptance angle. So, numerical aperture is equal to the sine of acceptance angle**.**

$$\text{N.A} = \sin \Theta a.$$

$$= \sqrt{n_1^2 - n n_2^2}$$

$$= \sqrt{(n_1 + n2)(n1 - n2)} \quad \ldots\ldots (9)$$

Numerical aperture is also defined as fractional index change **Δ**. It is the ratio of refractive index difference in core and cladding to the refractive index of core**.**

$$\text{i.e.,} \qquad \Delta = \frac{n1 - n2}{n1} \qquad \ldots\ldots\ldots (10)$$

This parameter is always positive because n1 must be larger than n2 for the total internal reflection condition. In order to guide light rays effectively through a fiber, Δ<<1.Typically Δ is of the order of 0.01.

From the equation (9)

n_1-n_2= Δ n_1……….. (11)

Substituting equation (10) in equation (9)

$$N.A= \sqrt{(n1+n2)\Delta n1}$$

For all optical fibers, $n_1 \approx n_2$, so

$$N.A=\sqrt{2nn_1^2\Delta}$$

$$N.A= n_1\sqrt{2\Delta} \ldots\ldots\ldots\ldots\ldots (12)$$

This is the relation between Numerical aperture and fractional index change.

Types of optical fibers:

Based on type of material used in core and cladding, optical fibers can be classified into two types. They are

(1) Glass fiber
(2) Plastic fiber

(1) Glass fiber :

If optical fibers are made by fusing mixtures of metal oxides and silica glasses, then it is known as 'glass fiber. The resulting material is randomly connected by molecular network rather than well defined ordered structures as found in crystalline materials.

The most common material used in glass fiber is silica (oxide glasses).It has a refractive index of 1.458 at 850 nm. To produce two similar materials having slightly different indices of refraction for the core and cladding, either fluorine or various oxides such as B_2O_3 , GeO_2 or P_2O_5. are added to silica.

Examples of fiber compositions are:

1. $GeO_2 – SiO_2$ core; SiO_2 cladding
2. P_2O_5-SiO_2 core ; SiO_2 cladding
3. SiO_2 core : P2O5-SiO_2 cladding

Another type of silica glasses are the low melting silicates .Such optical fibers are made of soda Lime silicates, germanosilicates and various borosilicates.

(2) Plastic fibers :

The plastic fibers are typically made of plastics are of low cost and can be easily handled without special care due to their toughness and durability.

Examples of plastic fibers are as follows:

1. Polystyrene core (n_1 = 1.6) and methylmethacrylate cladding (n_2=1.49)
2. Polymethylmethacrylate core (n_1= 1.49) and cladding made of its co-polymer (n_2=1.40)

Again based on variation of refractive index of core of an optical fiber, optical fibers are classified into two types. They are

(1) Step index optical fiber
(2) Graded index optical fiber

Again based on number of paths (Modes) , available for the light rays inside the core, these optical fibers are further divided into two types , they are

(1)Single mode optical fiber
(2) Multimode optical fiber

In single mode optical fiber, the width or diameter of the core is smaller when compared to the width of the cladding. As a result, only a single path (mode) is available for the light ray through the optical fiber.

When the width of the core is greater than the cladding, then the large number of paths (modes) is available for the light ray through the fiber and fiber and it is known as multimode optical fiber.

(1) Step index optical fiber – Refractive index profile:

- In a step index optical fiber, the refractive index of the core remains constant throughout the core and decreases from step to n_1 to n_2 at the core cladding interface. Thus it is known as Step-index optical fiber.
- The transmission of information will be in the form of signals or pulses.
- For single mode step index optical fiber, a single light ray from the signal enters into the fiber and traverses a single path and forms the output signal. In this case two signals match with each other as shown in figure 11.5
- In a multimode step index optical fiber, due to large width of core, greater number of light rays from the input signals enters into the core and takes multipath as shown in figure 11.6.
- The light ray (1) which greater angle with the fiber axis suffers more reflections through the fiber and takes more time to traverse the optical fiber, where as the light ray (2) makes less angle with the axis suffers less number of reflections and within a short time, it traverses the optical fibers.

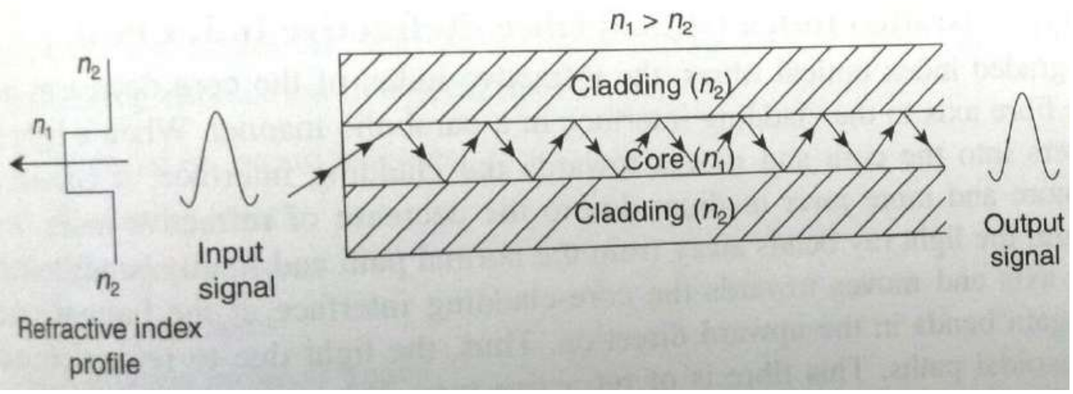


Figure: Single mode step index optical fiber.

- At the output end we receive ray (2) first and later we get ray (1). Due to the path difference between the light rays when they superimpose to form the output signals, the signals are overlapped. In this we get signal distortion known as ‘ intermodal dispersion’.

- It is difficult to retrieve the information carried by the distorted output signal. In a step index fiber, the propagation of light ray is due to multiple reflections, so it is reflective type.

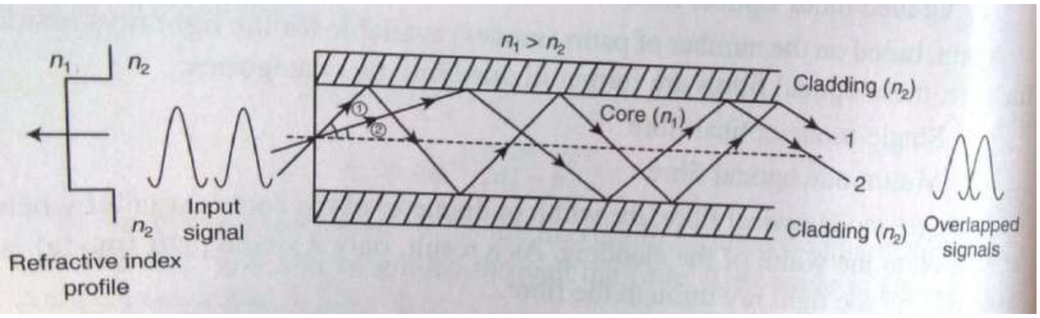


Figure: Multimode step index optical fiber.

The number of possible propagation modes in the core is given by the V-number as,

$$V = \frac{2\pi}{\lambda} a\ (NA)$$

Where λ = Wavelength of light
a = radius of the core
NA = Numerical aperture

Number of modes through step index fiber = $\frac{v^2}{2}$

Note :

Inter-modal dispersion: When more than one mode is propagating through a fiber, then the inter-modal dispersion will occur. Since, many modes are propagating; they will have different wavelengths and will take different time to propagate through the fiber, this results in elongation or stretching of data in the pulse. This is known as inter-modal dispersion.

(2) Graded index optical fiber – Refractive index profile:

- In graded index optical fiber, the refractive index of the core decreases from the fiber axis to the cladding surface in a parabolic manner.
- When a light ray enters into the core and moves towards the cladding interface, it encounters a more and more rarer medium due to decrease of refractive index.
- As a result, the light ray bends more away from the normal and finally bends towards the axis and moves the core-cladding interface at the bottom. Again it bends it bends in the upward direction.
- Thus the light due to refraction takes sinusoidal paths. This fiber is of refractive type. When two light rays (1) and light ray (2) making different angles with the axis enters into the fiber , they adjust their velocities (due to variations of refractive index) and come to focus at the same point.
- As a result, all the light rays will be received at the output end at the same time. There is no intermodal dispersion and the output signals match with input signal. It is easy to retrieve the information from the signals. In this fiber we get a refocusing effect of light rays.

- The number of possible modes through graded index fiber = $\frac{V^2}{2}$ where V is the V-number.

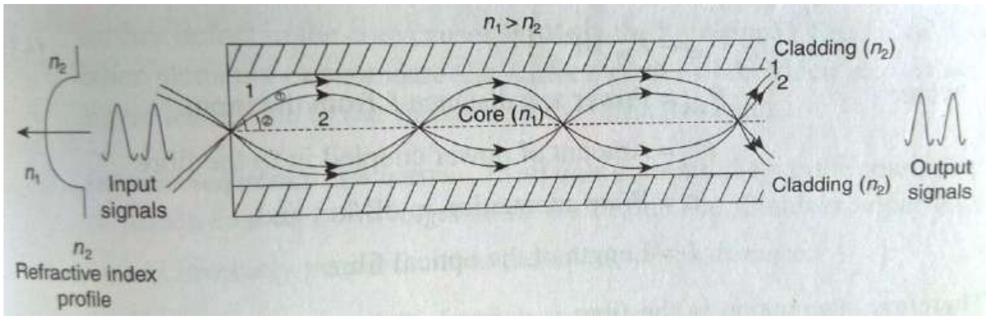


Figure: Graded index optical fiber.

Attenuation and losses in fibers:

When the light signal propagates in the optical fibers, losses arise due to different factors and these losses are referred to attenuation in optical fiber.

Losses are expressed in decibles per kilometer (dB/km). The attenuation loss (α) is given by,

$$\alpha = \frac{10}{l}\log ¿¿) \text{ dB/km}$$

Where P_{out} = Power at a distance L from the input

P_{in} = Amount of power coupled in the fiber

A = Fiber attenuation in dB/km

L = length of the optical fiber

Definition of Attenuation in optical fiber :

The ratio of the optical power input (P_{in}) fed to the optical fiber to the optical power output (P_{out}) obtained from a fiber of length ‘L’ is called Attenuation of an optical fiber.
The various factors causing attenuation in optical fiber are:

1) Material (or) impurity losses
2) Scattering losses
3) Absorption losses
4) Bending losses
5) Radiation induced losses
6) Inherent defect losses
7) Inverse sequence law losses
8) Transmission losses

9) Core and cladding losses.

1. Material (or) impurity losses :

The doped impurities present in the fabrication of an optical fiber in order to vary the refractive index causes losses in the light signal propagation through the fiber.

2. Scattering losses :

In glass fiber, the fiber glass contains many microscopic inhomogeneities and material content. Due to this, a portion of light signal passing through the glass fiber gets scattered. This scattering loss varies inversely with the fourth power of the wavelength.

$$\text{Scattering loss } \alpha \ \frac{1}{\lambda^4}$$

3. Absorption losses :

Absorption loss is caused by the nature of the core material and varies inversely to the transparency of the material. For glass fibers, ion-resonance absorption ultra violet absorption and infrared absorption are three separate mechanisms which contribute to total absorption losses.

4. Bending losses:

Whenever a fiber deviates from a straight line path , radiative losses occur.These losses are prominent for improperly installed single mode optical cable

5. Radiation induced losses:

When the glass molecular matrix interacts with the electrons, neutrons, X-rays and gamma rays, the structure of the molecules is altered and the fiber darkens. This introduces additional losses which increase with amount, type, dose and exposure time of radiation..

6. Inherent defect losses:

The inherent defect present in the core and cladding causes losses of the propagating light signal through it. The surface defect in the core causes losses in the light signal. Grease, oil and other contaminates on the surface of the fiber also causes signal losses due to variation of refractive index.

7. Inverse sequence law losses:

In all light systems, there is the possibility of losses caused by divergence of the beam. The illuminance per unit area is inversely proportional to the square of the distance.

8. Transmission losses:

The losses are caused by light which is caught in the cladding material of the optical fibers. This light is either lost to the outside or is trapped in the cladding layer and is thus not available for propagation in the core of the fiber.

9. Core and cladding losses:

In a fiber, core and cladding have different refractive indices as they have different compostions.so the core and the cladding have different attenuation coefficients, causing the power losses in the fiber.

Block diagram of optical fiber communication system:

An optical fiber communication system mainly consists of three parts viz., (1) transmitter section (2) optical fiber (3) receiver section as shown in fig.

Transmitter:

The information signal source may be audio, video, and data etc, which is in analog form to be transmitted, is converted from analog signal to electrical signal.

The transmitter consists of a drive circuit and a light source. The drive circuit transfers the electric input signal into digital pulses and the light source (LED/LASER) converts that optical pulses and are focused into the optical fiber.

Fiber–optic cable:

It acts as a waveguide and transmits the optical pulses towards receiver, by the principle of total internal reflection.\

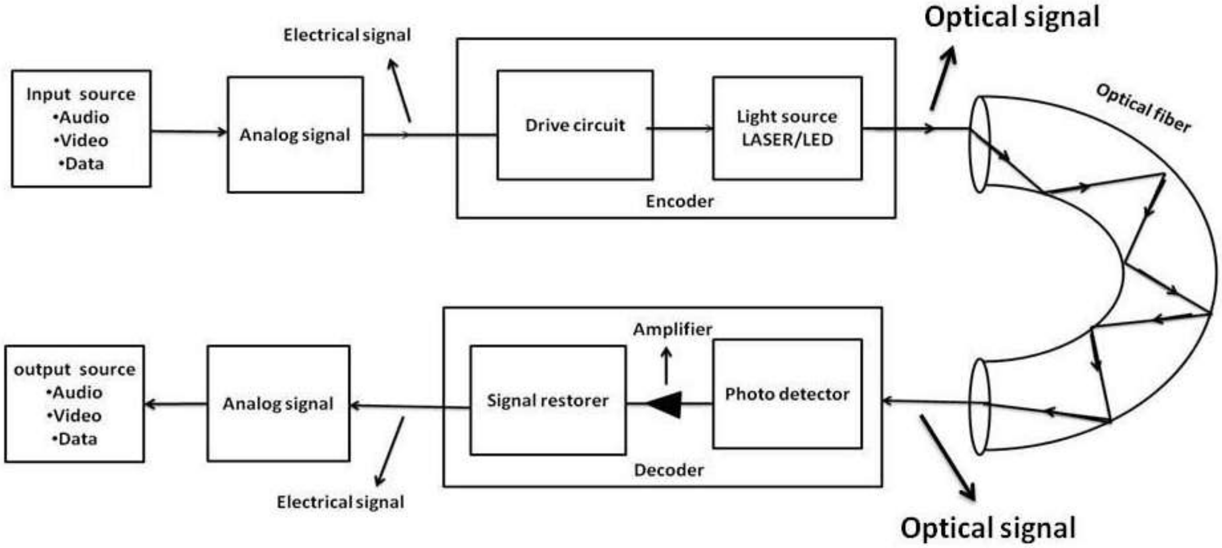


Fig: Fiber optical communication system

Receiver:

The receiver consists of a photo detector, amplifier and signal restorer. The photo detector receives the optical pulses and converts into electrical pulses. Further the signals are *amplified (distortion & noise are filtered out)* by an amplifier. These electrical signals are decoded i.e., converted from digital to analog signal. Thus original signal is obtained, in analog form, with the same information. In this way information is transmitted from one end to other end.

Advantages of Optical fiber in communication:

Let us see the advantages of optical fiber communication over conventional communication system.

(1) Enormous band width:

The information carrying capacity of a transmission system is directly proportional to the carrier frequency of the transmitted signals. The optical carrier frequency is in the range of 10^{14} Hz while the radio frequency is about 10^6 Hz and microwave frequency is about 10^{10} Hz. Thus the optical fibers have enormous transmission bandwidths and high data rate. Using wavelength division multiplexing operation, the data rate or information carrying capacity of optical fibers is enhanced to many orders of magnitude.

(2) Low transmission loss :

Due to the usage of ultra low loss fibers and the erbium doped silica fibers as optical amplifiers, one can achieve almost loss less transmission .Hence for long distance communication fibers of 0.002 dB/km are used. Thus the repeater spacing is more than 100 km.

(3)Immunity to interference and less cross talk:

Since optical fibers are dielectric wave guides, they are free from any electromagnetic interference (EMI) and radio frequency interference (RFI) . Since optical interference among different fiber is not possible , cross talk is negligible even many fibers are cabled together.

(3) Electrical isolation :

Optical fibers are made from silica which is an electrical insulator. Therefore they do not pick up any electromagnetic wave or any high current lightening.It is also suitable in explosive environment.

(4) Small size and weight :

The size of the fiber ranges from 10μm to 50 μm which is very small. The space occupied by the fiber cable is negligibly compared to conventional electric cables. Optical fibers are light in weight. These advantages make them to use in aircrafts and satellites more effectively.

(5) Signal security:

The transmitted signal through the fiber does not radiate. Unlike in copper cables, a transmitted signal cannot be drawn from fiber without tampering it. Thus the optical fiber communication provides 100 % signal security.

(7) Ruggedness and flexibility :

The Fiber cable can be easily bend or twisted without damaging it. Further the fiber cables are superior than the copper cables in terms of handling, installation, storage, transportation, maintenance, strength and durability.

(8) Low cost and availability :

Since the fibers are made of silica which is available in abundance. Hence there is no shortage of material and optical fibers offer the potential for low cost communication.

(9) Reliability :

The optical are made from silicon glass which does not undergo any chemical reaction or corrosion .Its quality is not affected by external radiation. Further due to its negligible attenuation and dispersion, optical fiber communication has high reliability. All the above factors also tend to reduce the expenditure on its maintenance.

Industrial applications of optical fibers:

(i) Applications of Optical fibers in Communication :

Optical fibres are used in exchange of information between different networks of computers.

- For example, a local area network (LAN) is a computer network that interconnects computers in a limited area such as a home, school, computer laboratory, or office building using network media to exchange the information. They are used for short distances about 1to 2 km.
- **Long haul communication**: They are used for long distances, 10 km or more. Tele phone cables in which Optical fibres are used in to exchange of information between various places.
- They are used for exchange of information in cable television, space vehicles, submarines, etc**.**
- Nearly 10000 information carrying signals can be transmitted simultaneously through the optical fiber.
- They are used for guiding weapons and submarine communication systems.
- As the optical fiber is highly immune to temperature, moisture etc, without any environmental effects the information can be delivered.
- During the war time they are used for secret communication.

(ii) Applications of Optical fibers in Medicine :

1. Fiberscope in endoscopy is one of the widely used optical techniques to view the internal parts of the disease affected body. In this optical fibre plays a major role in visualization of internal portions of human body but also in the selective cauterisation of tissues using laser beam.

2. This technique is widely used for the diagnosis of interior of lungs, stomach and other human body parts.
3. This method is used for the examination of gastrointestinal tract for diagnosis of ulcers, cancers etc.
4. Optical fibres are used in photodynamic thrapy for cancer.
5. They are used in treatment of lung disorders.
6. They are used in treatment of bleeding ulcers.
7. They are used in arthroscopic surgery for damaged cartilage, ligaments, and tendons in major joints such as knees and shoulders.
8. They are used in the investigation of heart , respiratory system and pancreas.

Optical fiber in sensors:

Fiber optic sensor is a transducer which converts any form signal into optical signal in measurable form. They can used for measuring physical parameters such as Temperature, pressure, Flow, liquid level, displacement, velocity, acceleration, Force, Rotation, Vibration, Radiation, pH, humidity, strain, Acoustic fields, Electric fields and Magnetic fields etc. A fiber optic sensor consists of a light source which generates light signals. These signals passes through the suitable optical fiber placed in the sensing fields and then passes through the light detector. The variation in the light signal is caused by sensing field and is detected by the detector as shown in fig.

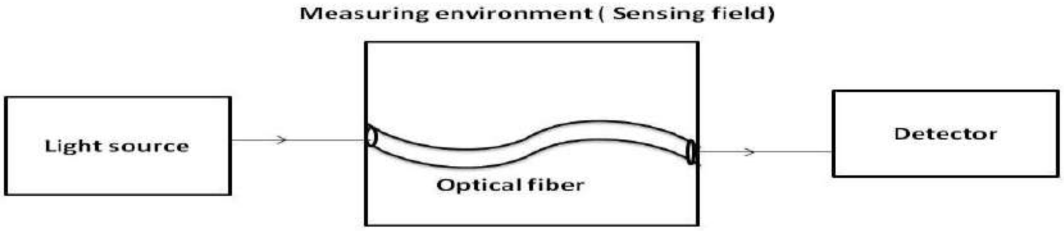


Fig .Block diagram of fiber optic sensor

(1) Temperature and pressure sensor:

It is an example of active sensors.

Principle:

It is based on the principle of interference between the beams emerging out from the reference fiber and the test fiber kept in the measuring environment.

Description:

- The block diagram of Temperature and pressure sensor as shown in fig.
- It consists of a laser source to emit light.
- A beam splitter arranged at 45^0.It splits the incident beam into two beams: Main beam and split beam.
- In which two fibers are used.

- Reference fiber which is isolated from the measuring environment.
- Test fiber kept in the environment to be sensed.
- Double convex lens are provided to convergent the beam.

Working:

- A monochromatic source of light is emitted from the laser source.
- The beam splitter kept at 45^0 which divide the beam emerging from the laser source into beams (i) main beam (ii) spitted beam, exactly at right angles to each other.
- The main beam passes through the lens L_1 and is focused onto the reference fiber which is isolated from the environment to be sensed.
- The beam after passing through the reference fiber then falls on the lens L_2.
- The split beam passes through the lens L_3 and is focused onto the test fiber kept in the environment to be sensed.
- The beam after passing through the test fiber then falls on the lens L_2.
- The two beams after passing through the fibers, produces a path difference due to the change in parameters such as pressure, temperature etc., in the environment.
- Therefore a path difference is produced between the two beams, causing the interference pattern as shown in fig.
- Thus the change in pressure or temperature can be accurately measured with the help of the interference pattern obtained.

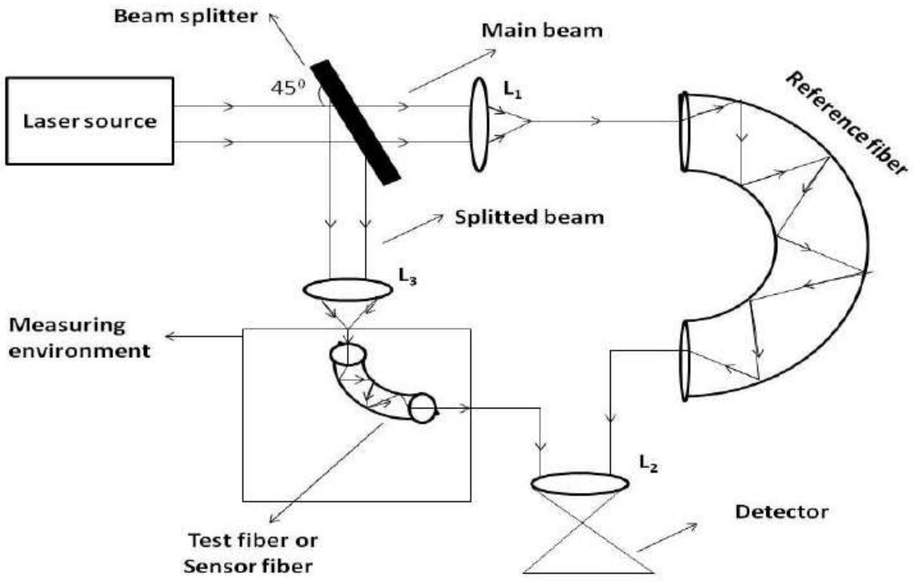


Fig: Temperature and pressure sensor

(3) Liquid flow level detector:

It is an example of active sensor.

Principle:

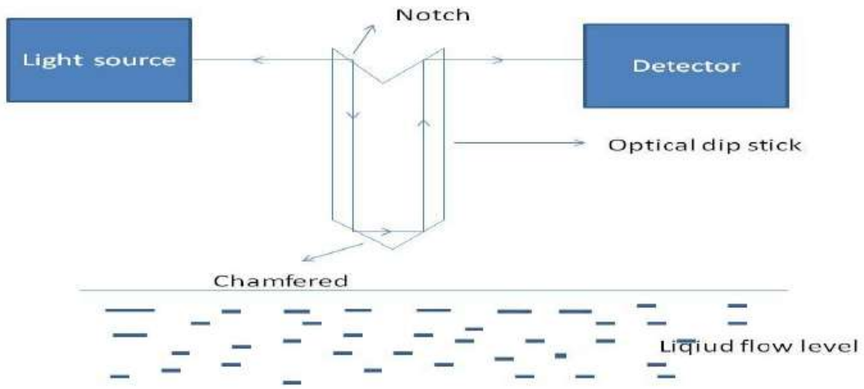
The liquid level detector described here is based on the principle of total internal reflection.

Description:

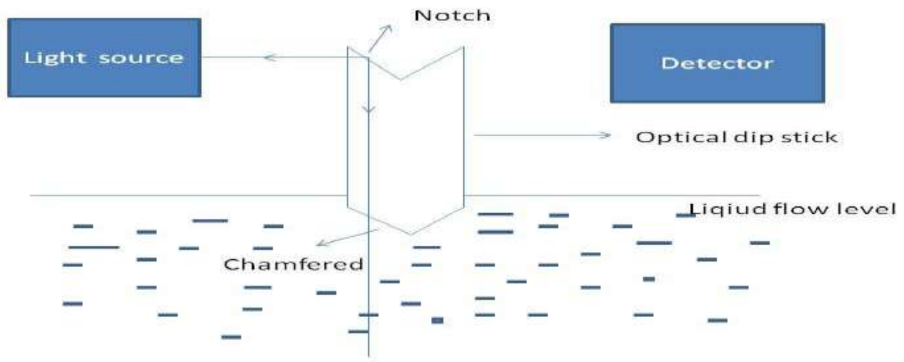
- A simple liquid detector is shown in fig.
- A notch is made of one end of multimode optical fiber and its other end is chamfered as shown in fig.
- A light source sends light onto the fiber and a photo detector on the other side measures light emerging out from the fiber.

Working:

- The optical fiber is arranged at the desired height in a vessel.
- The refractive index of the fiber is chosen to be less than that of the liquid whose level is to be detected.
- Light from the light source is made to be incident on one of the inclined faces of the notch.
- The light turns through 90^0 and travels through the fiber.



- On reaching the chamfered end of the fiber, it gets internally reflected, if the liquid is below the desired level. Then, it is again turned through 90^0 at the opposite face, travels back through the fiber to be turned once again through 90^0 and is detected at the detector.
- When the liquid rises and touches the fiber end, total internal reflection ceases and light is transmitted into the liquid. Hence photo detector does not receive any light.
- Thus, an indication of the liquid level is obtained at the detector.



(4) Displacement sensor:

It is an example of a passive sensor.

Principle:

In this method, two separate sensors are used for transmitting the light on the moving object and other, to receive the reflected light from the object. Light is sent through a transmitting fiber and is made to fall on a moving object. The reflected light from the target is sensed by a detector. Based on the intensity of light reflected from it the displacement of the object is measured.

Description:

- The displacement sensor consists of a light source, a transmitting fiber, a receiving fiber, an object and the detector as shown in fig.
- The transmitting optical fiber consists of bundle of fibers coupled to the laser source and receiving optical fiber is also consists of bundle of fibers coupled to the detector.

Working:

- The light energy from the He-Ne laser source is transmitted through the transmitting fiber and is made to incident on the moving object.
- The light reflected from the moving object is made to pass through the receiving fiber and the same is detected by the detector.
- Based on the intensity of light received ,the displacement of object can be measured,
  For example, if the object is moving towards the sensor, the intensity of light increases and if the object is moving away from the sensor, the intensity of light decreases.
- Thus the change in intensity of the light is used to measure the displacement of the object.

UNIT – III

Chapter – I

Dielectric Materials

Introduction

Dielectrics are insulating or non-conducting ceramic materials and are used in many applications such as capacitors, memories, sensors and actuators. Dielectrics are insulating materials that exhibit the property of electrical polarization, thereby they modify the dielectric function of the vacuum. A dielectric material is any material that supports charge without conducting it to a significant degree. In principle all insulators are dielectric, although the capacity to support charge varies greatly between different insulators. Although these materials do not conduct electrical current when an electric field is applied, they are not inert to the electric field. The field may cause a slight shift in the balance of charge within the material to form an electric dipole.

Thus the materials is called dielectric material.

Dielectric materials are used in many applications, from simple electrical insulation to sensors and circuit components.

Faraday was carried out the first numerical measurements on the properties of insulating materials when placed between the two parallel plates (capacitor), those materials, he called as dielectrics. He has found that the capacity of a condenser was dependent on the nature of the material separating the conducting surface. This discovery encouraged further empirical studies of insulating materials aiming at maximizing the amount of charge that can be stored by a capacitor. In search of suitable dielectric materials for specific applications, these materials have become increasingly concerned with the detailed physical mechanism governing the behavior of these materials.

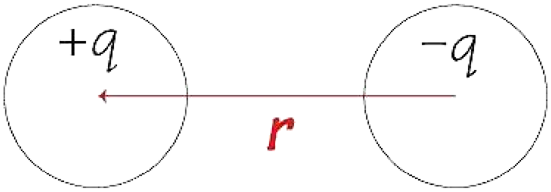
The difference between dielectric material and insulator depends on its application. Insulating materials are used to resist flow of current through it, on the other hand dielectric materials are used to store electrical energy. In contrast to the insulation aspect, the dielectric phenomena have become more general and fundamental, as it has the origin with the dielectric polarization.

Electric dipoles:

Upon application of a dc or static electric field, there is a long range migration of charges. However, there is a limited movement of charges leading to the formation of charge dipoles and the material, in this state, is considered as polarized. These dipoles are aligned in the direction of the applied field. The net effect is called Polarization of the material.

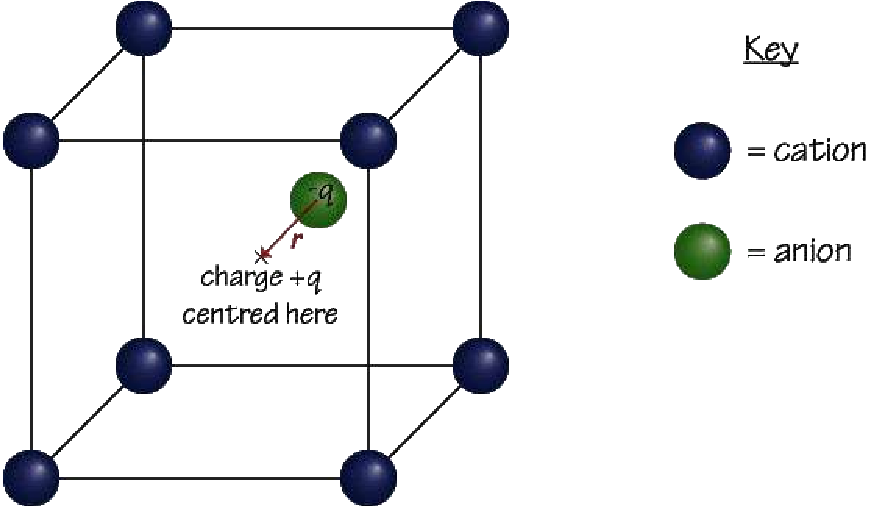
A dielectric supports charge by acquiring a polarisation in an electric field, whereby one surface develops a net positive charge while the opposite surface develops a net negative charge. This is made possible by the presence of electric dipoles – two opposite charges separated by a certain distance – on a microscopic scale.

1. If two discrete charged particles of opposite charges are separated by a certain distance, a dipole moment ***μ*** arises.



$$\mu = qr$$

2. If the centre of positive charge within a given region and the centre of negative charge within the same region are not in the same position, a dipole moment ***μ*** arises. For example, in the diagram below the centre of positive charge from the 8 cations shown is at X, while the centre of negative charge is located some distance away on the anion.



$$\mu = qr$$

The second view of dipole moment is more useful, since it can be applied over a large area containing many charges in order to find the net dipole moment of the material.

The dipoles can be aligned as well as be induced by the applied field.

Note that in the equation for dipole moment, ***r*** is a vector (the sign convention is that ***r*** points from negative to positive charge) therefore the dipole moment ***μ*** is also a vector

Dielectric loss:

Dielectric loss is the dissipation of energy through the movement of charges in an alternating electromagnetic field as polarisation switches direction.

An efficient dielectric supports a varying charge with minimal dissipation of energy in the form of heat is called dielectric loss. There are two main forms of loss that may dissipate energy within a dielectric. In conduction loss, a flow of charge through the material causes energy dissipation.

Dielectric loss is especially high around the relaxation or resonance frequencies of the polarisation mechanisms as the polarisation lags behind the applied field, causing an interaction between the field and the dielectric's polarisation that results in heating. This is illustrated by the diagram below (recall that the dielectric constant drops as each polarisation mechanism becomes unable to keep up with the switching electric field.)

Dielectric loss quantifies a dielectric material's inherent dissipation of electromagnetic energy into, e.g., heat.

It can be represented in terms loss tangent tan δ and is defined:

$$\tan \delta_e = \frac{\epsilon''}{\epsilon'}$$

Dielectric Breakdown : The dielectric breakdown is the sudden change in state of a dielectric material subjected to a very high electric field , under the influence of which , the electrons are lifted into the conduction band causing a surge of current , and the ability of the material to resist the current flow suffers a breakdown .

Or

When a dielectric material loses its resistivity and permits very large current to flow through it, then the phenomenon is called dielectric breakdown

Or

At high electric fields, a material that is normally an electrical insulator may begin to conduct electricity – i.e. it ceases to act as a dielectric. This phenomenon is known as dielectric breakdown.

Frequency dependence of polarizability:

On application of an electric field, polarization process occurs as a function of time. The polarization P(t) as a function of time. The polarization P(t) as a function of time t is given by

$$P(t) = P[\ 1- \exp\ (\ - t\ /\ t_r\ )]$$

Where P – max. Polarization attained on prolonged application of static field. t_r - relaxation time

for particular polarization process

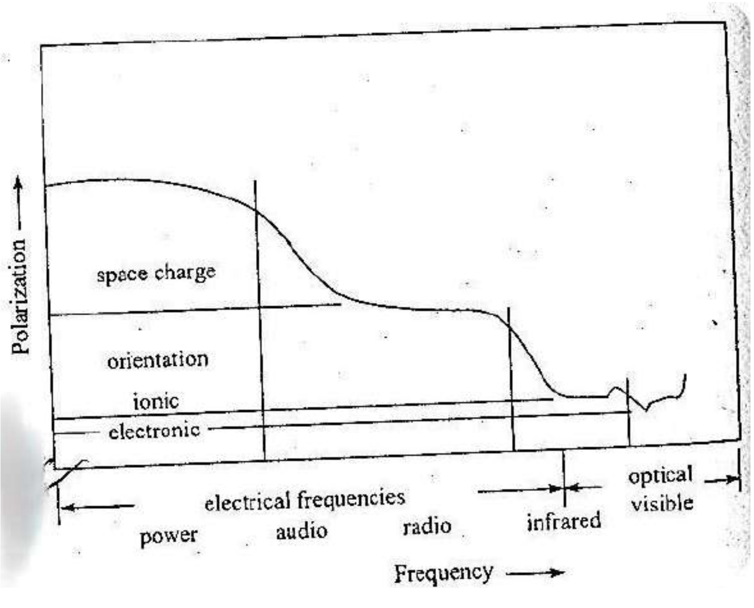
The relaxation time t_r is a measure of the time scale of polarization process. It is the time taken for a polarization process to reach 0.63 of the max. value.

Electronic polarization is extremely rapid. Even when the frequency of the applied voltage is very high in the optical range (≈10^{15} Hz), electronic polarization occurs during every cycle of the applied voltage.

Ionic polarization is due to displacement of ions over a small distance due to the applied field. Since ions are heavier than electron cloud, the time taken for displacement is larger. The frequency with which ions are displaced is of the same order as the lattice vibration frequency (≈1013Hz). Hence, at optical frequencies, there is no ionic polarization. If the frequency of the applied voltage is less than 1013 Hz, the ions respond.

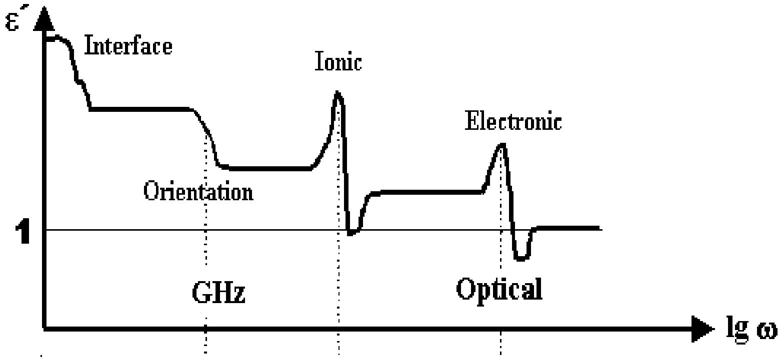
Orientation polarization is even slower than ionic polarization. The relaxation time for orientation polarization in a liquid is less than that in a solid. Orientation polarization occurs, when the frequency of applied voltage is in audio range (1010 Hz).

Space charge polarization is the slowest process, as it involves the diffusion of ions over several interatomic distances. The relaxation time for this process is related to frequency of ions under the influence of applied field. Space charge polarization occurs at power frequencies (50-60 Hz).



Frequency Dependence of dielectric constant

When a dielectric material is subjected to an alternating field, the polarization component required to follow the field in order to contribute to the total polarization of the dielectrics. The relative permittivity which is a measure of the polarization also depends on the frequency. The dependence of ε_r on frequency of the electric field is shown in the figure.



At very low frequency, the dipoles will get sufficient time to orient themselves completely with the field and all types of polarization exist. Since the dielectric is characterized by polarisability α ($\alpha = \alpha_e + \alpha_i + \alpha_o$) at low frequency i.e at radiofrequency region the dielectric constant will be due to all polarisability.

The orientation polarization, which is effective at low frequencies, is damped out for higher frequencies. In the microwave region the dipoles fail to follow the field and the polarisability reduces to ($\alpha = \alpha_e + \alpha_i$), as a result ε_r decreases to some amount.

In the IR region the ionic polarization fails to follow the field so the contribution of ionic polarization dies away. In this region only electronic polarization contributes to the total polarization. Therefore ($\alpha = \alpha_e$) the ε_r still decreases and only electronic polarization exist.

We know that $$P = N\alpha_e E = \varepsilon_o(\varepsilon_r - 1)E$$

Then the relative permittivity is $$\varepsilon_r = 1 + \frac{N\alpha_e}{\varepsilon_o} \quad \ldots(1)$$

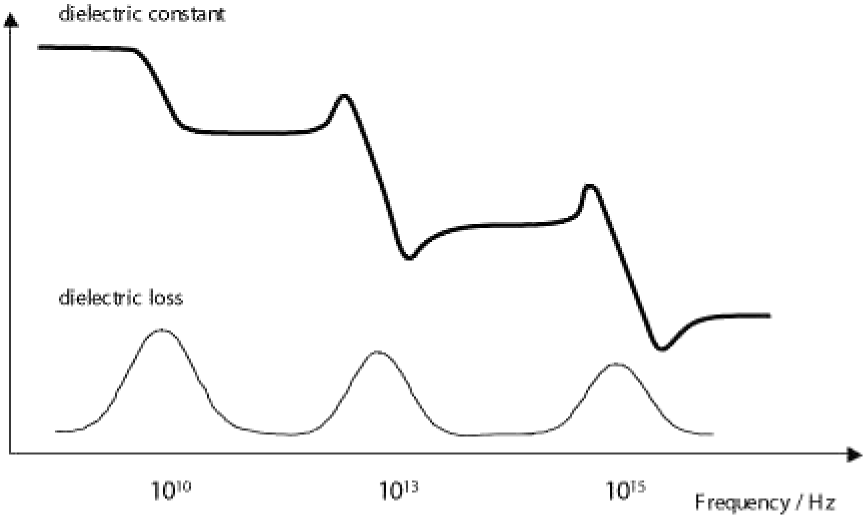
In the ultraviolet region even the electron cloud could not follow the field and electronic polarizability becomes almost zero and the permittivity becomes one.

$$[\varepsilon_r]_{X-ray} = 1$$

For example at low frequency the dielecric constant of water at room temperature is about 80, but it fall to about 1.8 in the optical region.

Frequency Dependence of dielectric loss:

Dielectric loss tends to be higher in materials with higher dielectric constants. This is the downside of using these materials in practical applications. Dielectric loss is utilised to heat food in a microwave oven: the frequency of the microwaves used is close to the relaxation frequency of the orientational polarisation mechanism in water, meaning that any water present absorbs a lot of energy that is then dissipated as heat. The exact frequency used is slightly away from the frequency at which maximum dielectric loss occurs in water to ensure that the microwaves are not all absorbed by the first layer of water they encounter, therefore allowing more even heating of the food.



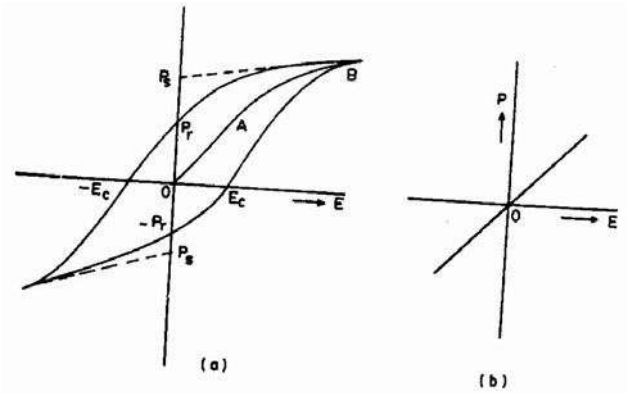
Ferroelectrics

Below certain temperature it is found that some materials spontaneously acquire an electric dipolemoment. These materials are called as ferroelectric materials or ferroelectrics.The temperature at which ferroelectric property of the material disappears is called as ferroelectric Curie temperature.

Ferroelectric materials are anisotropic crystals which exhibit a hysteresis curve P versus E which can be explained by domain hypothesis.

Ferro electricity: Ferro electric materials are an important group not only because of intrinsic Ferro electric property, but because many possess useful piezo electric, birefringent and electro optical properties.

The intrinsic Ferro electric property is the possibility of reversal or change of orientation of the polarization direction by an electric field. This leads to hysteresis in the polarization P, electric field E relation , similar to magnetic hysteresis. Above a critical temperature, the Curie point T_c, the spontaneous polarization is destroyed by thermal disorder. The permittivity shows a characteristic peak at T_c.



Piezo – Electric Materials and Their Applications: Single crystal of quartz is used for filter, resonator and delay line applications. Natural quartz is now being replaced by synthetic material.

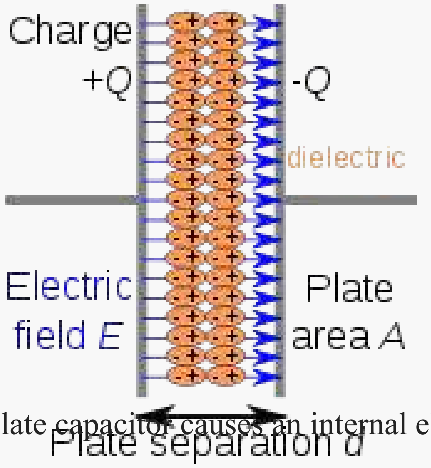
Rochelle salt is used as transducer in gramophone pickups, ear phones, hearing aids, microphones etc. the commercial ceramic materials are based on barium titanate, lead zirconate and lead titanate. They are used for high voltage generation (gas lighters), accelerometers, transducers etc. Piezo electric semiconductors such as GaS, ZnO & CdS are used as amplifiers of ultrasonic waves.

Applications of Dielectric Materials:

Almost any type of electrical equipment employs dielectric materials in some form or another. Wires and cables that carry electrical current, for example, are always coated or wrapped with some type of insulating (dielectric) material. Sophisticated electronic equipment such as rectifiers, semiconductors, transducers, and amplifiers contain or are fabricated from dielectric materials. The insulating material sandwiched between two conducting plates in a capacitor is also made of some dielectric substance.

Liquid dielectrics are also employed as electrical insulators. For example, transformer oil is a natural or synthetic substance (mineral oil, silicone oil, or organic esters, for example) that has the ability to insulate the coils of a transformer both electrically and thermally.

1. **Capacitors**



Charge separation in a parallel-plate capacitor causes an internal electric field. A dielectric (orange) reduces th

Commercially manufactured capacitors typically use a solid dielectric material with high permittivity as the intervening medium between the stored positive and negative charges. This material is often referred to in technical contexts as the *capacitor dielectric*.

The most obvious advantage to using such a dielectric material is that it prevents the conducting plates, on which the charges are stored, from coming into direct electrical contact. More significantly, however, a high permittivity allows a greater stored charge at a given voltage. This can be seen by treating the case of a linear dielectric with permittivity ε and thickness d between two conducting plates with uniform charge density σ_ε. In this case the charge density is

$$\sigma_\varepsilon = \varepsilon \frac{V}{d}$$

given by

and the capacitance per unit area by

$$c = \frac{\sigma_\varepsilon}{V} = \frac{\varepsilon}{d}$$

From this, it can easily be seen that a larger ε leads to greater charge stored and thus greater capacitance.

Dielectric materials used for capacitors are also chosen such that they are resistant to ionization. This allows the capacitor to operate at higher voltages before the insulating dielectric ionizes and begins to allow undesirable current.

2. **Dielectric resonator**

A *dielectric resonator oscillator* (DRO) is an electronic component that exhibits resonance of the polarization response for a narrow range of frequencies, generally in the microwave band. It consists of a "puck" of ceramic that has a large dielectric constant and a low dissipation factor. Such resonators are often used to provide a frequency reference in an oscillator circuit. An unshielded dielectric resonator can be used as a Dielectric_Resonator Antenna (DRA).

3. **Insulators-**

Required Qualities of Good Insulating Materials: The required qualities can be classified as under electrical, mechanical, thermal and chemical applications.

i) Electrical: 1. electrically the insulating material should have high electrical resistivity and high dielectric strength to withstand high voltage.

2 .The dielectric losses must be minimum.

3. Liquid and gaseous insulators are used as coolants. For example transformer oil, hydrogen and

ii) Mechanical: 1. insulating materials should have certain mechanical properties depending on the use to which they are put.

2. When used for electric machine insulation, the insulator should have sufficient mechanical strength to withstand vibration.

iii) Thermal: Good heat conducting property is also desirable in such cases. The insulators should have small thermal expansion and it should be non-ignitable.

iv) Chemical: 1. chemically, the insulators should be resistant to oils, liquids, gas fumes, acids and alkali's.

2. The insulators should be water proof since water lowers the insulation resistance and the dielectric strength.

Other Applications:

Solid dielectrics are perhaps the most commonly used dielectrics in electrical engineering, as very good insulators. Some examples include porcelain, glass, and most plastics.
Air, nitrogen and sulfur hexafluoride are the three most commonly used gaseous dielectrics.

Industrial coatings such as parylene provide a dielectric barrier between the substrate and its environment.

Mineral oil is used extensively inside electrical transformers as a fluid dielectric and to assist in cooling. Dielectric fluids with higher dielectric constants, such as electrical grade castor oil, are often used in high voltage capacitors to help prevent corona discharge and increase capacitance.

Because dielectrics resist the flow of electricity, the surface of a dielectric may retain stranded excess electrical charges. This may occur accidentally when the dielectric is rubbed (the triboelectric effect). This can be useful, as in a Van de Graaff generator or electrophorus, or it can be potentially destructive as in the case of electrostatic discharge.

Piezoelectric materials are another class of very useful dielectrics which are used for transducers and sensors.

Ferroelectric materials often have very high dielectric constants, making them quite useful for capacitors.

Annamacharya Institute of Technology and Sciences, Tirupati
(Autonomous)
Applied Physics Material (20ABS9902)
(AK20 Regulation)

Unit-3
Chapter-2
Magnetic Materials

Magnetism

A substance that attracts pieces of iron (or) steel is called "Magnet". This property of a substance is called "magnetism.

Basics of Magnetism

Magnet:

The magnets are materials which produces a magnetic field. The magnets are material which attracts the ferromagnetic materials like iron, co, Ni etc.

Properties of Magnets:

1. Magnets attract ferromagnetic objects.
2. The magnetic field at the poles of a magnet is greater than at the middle of magnet.
3. The like poles of magnet repel each other while the opposite poles attract each other.
4. The magnet has a property that when it is suspended freely it's South Pole and North Pole moves to the earth's North and South Pole, respectively.

Magnetic Poles:

When a bar magnet is dipped in a heap of iron filings and taken out, it is seen that maximum amount of iron filings are seen to cling to the magnet at the two ends of the magnet. Practically no iron filings cling in the middle. The regions of the magnet at which maximum amount of iron filings cling are called poles of the magnet.*Poles of magnet are regions near the two ends of a magnet with maximum power of attraction.* The strength of the pole is called pole strength denoted by*m*. The S.I. unit of pole strength is *Ampere Meter.*

The distance between two magnetic poles is called "magnetic length" (2*l*).

Magnetic Dipole:

- Magnetic dipoles are found to exist in magnetic materials, analogous to electric dipoles.
- Two equal and opposite charges separated by a small distance is called an electric dipole. Similarly a north pole and south pole separated by a small distance 2*l* **(magnetic length***)* constitute a magnetic dipole.

- *For example:* A bar magnet, a compass needle etc. are the magnetic dipoles.
  And also a current loop behaves as a magnetic dipole.

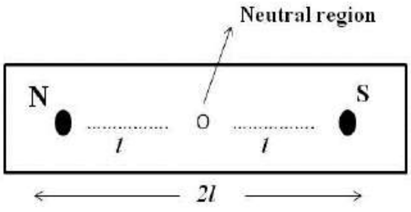


Fig : Magnetic dipole

Magnetic Dipole Moment:

The behavior of magnetic dipole is described by the magnetic dipole moment.

(a)In the case of bar magnet:

It is defined as the product of pole strength (m) and magnetic length (2*l*).

$$\boxed{\vec{\mu}_m = m\,(2l)}$$

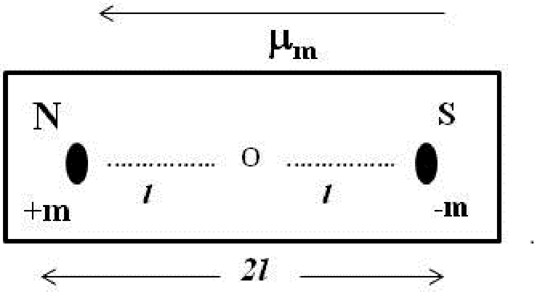


Fig : Magnetic dipole moment

It is a vector quantity. It is directed from South Pole to North Pole.
The S.I.Unit of magnetic dipole moment: Ampere – meter ^{2} (A-m^2).

(b) In the case of current loop:

A current carrying loop behaves as a magnetic dipole.
Consider a current carrying conductor loop of wire as shown fig. The current (I) establishes a magnetic field around the loop. By right hand palm rule, the upper face of the loop acts a N-pole and the lower face act as S- pole.

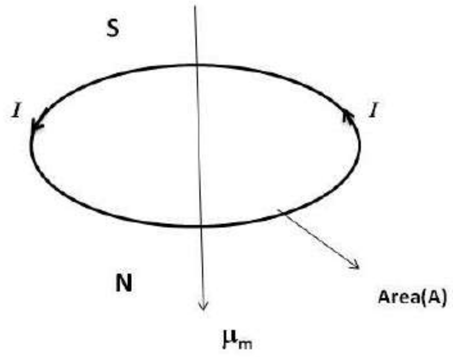


Fig : Current carrying conductor loop

The magnitude of dipole moment of current loop (μ_m) is
(i) Directly proportional to current (I) through the loop.
(ii) Directly proportional to the area of cross –section (A).

$\mu_m \alpha\ I.A$

$\mu_m = K\ IA$

$$\boxed{\mu_m = IA}$$

Where K is a proportionality whose value is one

Magnetic Field:

.The space surrounding a magnet where magnetic force is experienced is called a magnetic field.

A magnetic field can be represented by drawing lines called "magnetic lines of force".The lines go from North to South on the magnet.

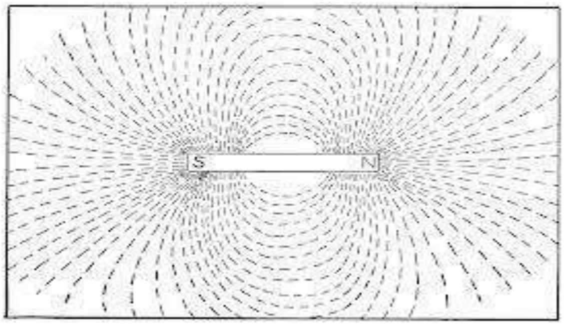


Fig: Magnetic field.

Magnetic Flux:

A group of magnetic lines of force is called "magnetic flux".

The symbol for magnetic flux is Φ (phi).

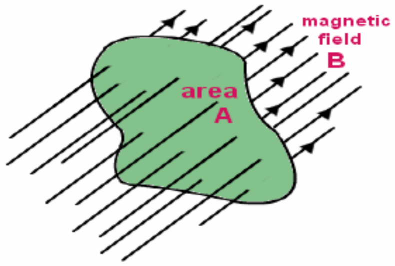
The SI unit of magnetic flux is the Weber (Wb).

One Weber is equal to 1 x 10^8 magnetic fieldlines.

Magnetic Flux Density:

Magnetic flux density is the amount of magnetic flux per unit area of a section, perpendicular tothe direction of flux.

$$\text{Magnetic flux density (B)} = \frac{\text{Magneticflux (Weber)}}{\text{Area (m}^2\text{)}} \qquad \boxed{B = \frac{\emptyset}{A}\ \text{Tesla}}$$



Magnetization:

Magnetization in magnetic field is analogues to polarization of dielectric material in electrostatic field.

The process of converting a non-magnetic material into a magnetic material is known as "magnetization".

Intensity of Magnetization (I or M)

When a material medium is placed in a magnetic field, it gets magnetized. To magnetize the material medium is to create the magnetic dipole moments.

The magnetic dipole moment per unit volume of the material is called the intensity of magnetization I (or simply magnetization).

$$I = \frac{\text{Magnetic dipole moment}(\mu_m)}{\text{Volume (V)}} = \frac{\text{Length of magnet } (2l) \times \text{Pole Strength (m)}}{\text{Length of Magnet } (2l) \times \text{Area of cross –section (A)}}$$

$$= \frac{\text{Pole Strength (m)}}{\text{Area of Cross Section (A)}}$$

The S.I. Unit of magnetization is ampere / meter

Magnetic Field Strength (H):

The ability of magnetic field to magnetize a material medium is called its magnetic intensity or field strength. It is denoted by **H**.

The S.I. Unit of magnetic field strength is ampere / meter.

Magnetic Susceptibility (χ_m):

The word Susceptibility comes from the Latin word “susceptible” means the easily affected

The magnetic susceptibility of a material medium indicates how easily a material medium can be magnetized in the presence of magnetic field..

The intensity of Magnetization is directly related to the applied field strength H.

$$M \alpha H$$

$$M = \chi_m H$$

$$\boxed{\text{Magnetic Susceptibility } (\chi_m) = \frac{M}{H}}$$

Therefore; the magnetic susceptibility of a material is defined as the ratio of intensity of magnetization (I) developed in the material to the applied magnetic field (H).

Magnetic Permeability (μ):

- It is the natural property of material.
- In Latin, *per* means *through* and *meare* means *to pass.*
- It is defined as the ability of the material to permit the passage of magnetic lines of force through it.
- The Magnetic induction B is proportional to the applied Magnetic field intensity H.

$$B \alpha H$$

$$B = \mu H$$

$$\boxed{\text{Magnetic Permeability } (\mu) = \frac{B}{H}}$$

Where “μ” is the permeability of a medium.

For vacuum,

$$\boxed{B = \mu_0 H}$$

Where is the proportionality constant and is also called permeability of the free space and its value is $4\pi \times 10^{-7}$ H m^{-1}.

Relative permeability (μ_r):

The ratio of permeability of medium to the permeability of free space is called relative permeability μ_r of the medium

$$\boxed{\mu_r = \frac{\mu}{\mu_o}}$$

$\mu_r = 1$ for vacuum. It has no units.

The Relation between Relative Permeability and Magnetic Susceptibility:

When a magnetic material is magnetized by placing it in a magnetic field, the resultant field inside the material is the sum of the field due to the magnetization of the material and the original magnetizing field. The resultant field is called magnetic induction or magnetic flux density **B.**

$$B = \mu_0 H + \mu_0 M$$

$$B = \mu_0 (H+M)$$

$$\mu H = \mu_0 (H+M) \qquad (B=\mu H \text{ and } \mu = \mu_0 \mu_r)$$

$$\mu_0 \mu_r H = \mu_0 (1+\frac{M}{H})$$

$$\boxed{\mu_r = (1+\chi_m)}$$

$$\text{Where } \chi_m = \frac{M}{H}$$

This is the relation between Relative Permeability and Magnetic Susceptibility.

Origin of Magnetism:

- All substances are made of atoms or molecules. An atom which consists of '+' vely charged nucleus at the centre and negatively charged electrons revolving around the nucleus in different orbits. This motion of electrons is called orbital motion as shown in fig. The orbiting elections constitute tiny current loops. These loops behave as the magnetic dipoles.
- The orbital motion of electrons around the nucleus gives rise to the orbital magnetic dipole moment (μ_{orbit}).
- The electrons also rotate around their own axes. This motion of electrons is called spin motion as shown fig.
  The spinning motion of electrons around their axes gives rise to the spin magnetic dipole m oment (μ_{spin}).

- The motion of the protons and neutrons within the nucleus also contributes to the total magnetic moment (($\mu_{nucleus}$). But the magnitude of the nuclear magnetic moment is (about 10 $^{-3}$ times) very small compared with the magnetic moment of electron and is usually neglected.

Orbital magnetic dipole moment of electron (μ_{orbit}):

According to electron theory of matter;

All substances are made of atoms or molecules. An atom which consists of '+' vely charged nucleus at the centre and negatively charged electrons revolving around the nucleus in different orbits. This motion of electrons is called orbital motion as shown in fig. The orbiting elections constitute tiny current loops. These loops behave as the magnetic dipoles. The orbital motion of electrons around the nucleus gives rise to the orbital magnetic dipole moment (μ_{orbit}).

Let us consider an electron of mass **'m'** and charge e revolving around the nucleus in a circular orbit o radius **'r'** with linear velocity **'v'** as shown in fig.

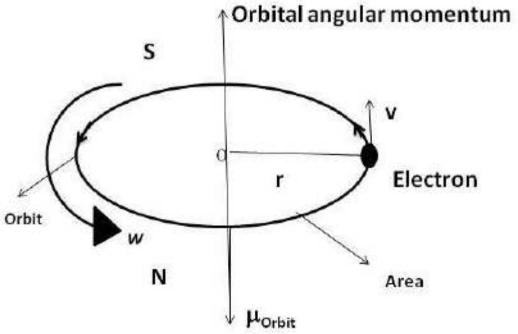


Fig : Orbital motion of electron

The revolving electron in circular orbit establishes a current is given by

$$I = \frac{\text{Charge of electron}}{\text{time period}} = \frac{-e}{T} \longrightarrow (1)$$

Where 'T' is the time taken by the electron to make one revolution around the nucleus

$$\text{i.e., } T = \frac{2\pi}{w} \longrightarrow (2)$$

Where *'w'* is the angular frequency of the electron

But relation between linear velocity 'v' and angular velocity can be written as,

$$v = r\,w$$

$$\text{and } \omega = \frac{v}{r} \longrightarrow (3)$$

Substituting the equation (3) in (2),

$$T = \frac{2\pi r}{v} \longrightarrow (4)$$

Further, substituting the equation (4) in (1),

$$I = \frac{-ve}{2\pi r} \longrightarrow (5)$$

The current 'I' establishes a magnetic field around the circular orbit, so that the upper surface acts as South Pole and the lower surface acts as North Pole.

The Area of the orbit is **A** = πr^2 ⟶ (6)

Then the corresponding magnetic dipole moment is given by

$$\mu_{\text{Orbit}} = IA$$

$$= \frac{-ve}{2\pi r} \text{X} \pi r^2$$

$$= \frac{-evr}{2} \longrightarrow (7)$$

Dividing and multiplying the equation (7) by the mass "m' of electron.

$$\mu_{\text{Orbit}} = \frac{-evr}{2} \text{X} \frac{m}{m}$$

$$= \frac{-e(mvr)}{2m}$$

$$= \frac{-e\,(L)}{2m} \qquad \text{(But L= mvr)}$$

$$\boxed{\mu_{\text{Orbit}} = \frac{-e}{2m} L \text{ (Orbital angular momentum)}}$$

The – ve sign indicates that the orbital angular momentum and orbital magnetic dipole moment are in opposite directions.

An electron has an intrinsic orbital angular momentum ($\vec{L}$) itself cannot be measured. However, its component along any axis can be measured. And that component is quantized, with possible values of the same magnitude but different sings.

Let us assume that the component of orbital angular momentum ($\vec{L}$) is measured along the z-axis of a coordinate system. Then the measured component L_Z can have only the values is given by

$$L = l\frac{h}{2\pi}$$

Where ***l*** is called orbital magnetic quantum number = $0, \pm 1, \pm 2, \pm 3$ ....

Let us assume that the component of orbital magnetic dipole moment ($\mu_{Orbit)}$ of electron is measured along the z- axis of a coordinate system. Then the measured component μ_{orbit}, zcan have only the two values is given by

$$\mu_{Orbit,z} = \frac{-e}{2m} L$$

$$\mu_{Orbit,\,z} = -\left(\frac{e}{2m}\right) l \frac{h}{2\pi}$$

$$\mu_{Orbit,\,z} = -\left(\frac{eh}{4\pi m}\right) l$$

$$\boxed{\mu_{Orbit\,,z} = -\,\mu_B\,.\,\boldsymbol{l}}$$

Where $\mu_B = \frac{eh}{4m\pi}$ is known as Bohr magneton and its value is 9.27x 10-^{24} A-m^2.

Spin magnetic dipole moment of electron (μ_{spin}):

The electrons also rotate around their own axes. This motion of electrons is called spin motion as shown fig.

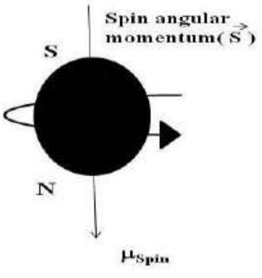


Fig: Spin motion of electron

The spinning electron constitutes a tiny current loop. This loop behaves as a magnetic dipole.

The magnetic dipole moment arises due to its spin motion is called spin magnetic moment (μ_{spin}) and is given by

$$\mu_{Spin} = \gamma \left(\frac{e}{2m}\right) \vec{S} \text{ (Spin angular momentum)}$$

$$S = m_s \frac{h}{2\pi}$$

Where m_s is called magnetic spin quantum number = $\pm\frac{1}{2}$

Let us assume that the component of spin magnetic dipole moment ((μ_{Spin}) of electron is measured along the z- axis of a coordinate system. Then the measured component μ_{spin}, zcan have only the two values is given by

$$\mu_{Spin} = \gamma\left(\frac{e}{2m}\right) S_z$$

$$\mu_{Spin} = \gamma\left(\frac{e}{2m}\right) m_s\frac{h}{2\pi}$$

$$\mu_{Spin} = \gamma\left(\frac{eh}{4\pi m}\right) m_s$$

Where $\mu_B = \frac{eh}{4m\pi}$ is known as Bohr magneton and its value is 9.27x 10^{-24} A-m^2.

Nuclear spin magnetic dipole moment ($\mu_{Nuclear}$):

The atomic nucleus contains protons and neutrons. They have intrinsic spin.
The spin motion of the protons and neutrons within the nucleus also contributes to the total spin magnetic dipole moment and is given by

$$\mu_{nuclear,\ spin} = \frac{eh}{4\pi M_p} = 5.525X1O^{-27}\ A\text{-}m^2$$

Where M_p is the Mass of the proton

But the magnitude of the nuclear magnetic dipole moment is(about 10^{-3} times) very small compared with the magnetic dipole moment of electron and is usually neglected.

Therefore, the magnetism mainly arises due to the orbital and spin magnetic dipole moments of electron.

Classification magnetic materials:

Magnetic materials are classified based on presence or absences of the permanent magnetic dipoles in a material. They are

1. Dia magnetic material
2. Para magnetic material
3. Ferro magnetic material
4. Anti Ferro magnetic material and
5. Ferri magnetic material

1. Diamagnetic materials:

Those materials which when placed in a magnetic field are weakly or feebly magnetized in a direction opposite that of the applied magnetic field are called diamagnetic materials

Examples:

Bismuth, Copper, Zinc, Gold, Water, etc

Cause of diamagnetism:

In the Diamagnetic materials, there exist paired electrons, so the spins in two opposite directions are equal and hence magnetic dipole moments cancel with each other. i.e., the resultant magnetic dipole moment is equal to zero. Therefore, most of these materials do not have magnetism in the absence of magnetic field.

Effect of external magnetic field:

a) In the absence of external magnetic field, the atoms/molecule/ions of the diamagnetic substance have no net magnetic dipole moment. Hence, the material does not exhibit diamagnetism

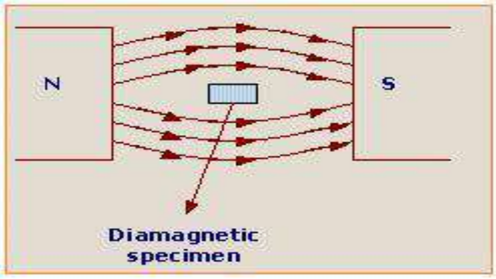
.

b) When a diamagnetic material is placed in an external magnetic field, currents are induced in the current loops of atom/molecule/ion according to Faraday's law of electromagnetic induction. According to Lenz's law, these currents give rise to a magnetic field which opposes the applied magnetic field. Hence, the induced magnetic moments of atoms/molecule/ions are opposite to the applied magnetic field.

Properties:

- They don't possess permanent magnetic dipole moment.

- When a diamagnetic material is placed in a magnetic field, it is feebly magnetized in a direction opposite to that of the applied magnetic field.
- When a diamagnetic material is placed in a magnetic field, the magnetic lines force prefers to pass through the surroundings air rather than through the diamagnetic magnetic material.



- The magnetic flux density inside is small than that in the free space. Hence the relative permeability $\mu_r < 1$.
- The magnetic susceptibility (χ_m) is negative and small.
- The magnetic susceptibility (χ_m) is independent of temperature.
- When a rod of diamagnetic material is suspended freely in a uniform magnetic field, the rod comes to rest with its axis perpendicular to the direction of the applied field.

2. Para magnetic materials:

Those materials which when placed in a magnetic field are weakly or feebly magnetized in the direction of the applied magnetic field are called Para magnetic materials.

Examples:

Aluminum, platinum, copper sulphate($CuSo_4$),manganese, chromium etc.

Cause of paramagnetism:

In the case of paramagnetic materials, the spins in two opposite directions will not be equal. There exist some unpaired electrons which gives rise to spin magnetic dipole moment.

Hence the resultant magnetic dipole moment will not be equal to zero. i.e., they possess permanent magnetic dipole moment.

Effect of external magnetic field:

c) In the absence of external magnetic field ,the dipoles of the paramagnetic material are randomly oriented and ,therefore, the net magnetic dipole moment of the material is zero. Hence, the material does not exhibit paramagnetism.

d) When a paramagnetic material is placed in an external magnetic field,the magnetic dipoles are partially aligned in the direction of the applied magnetic field. Therefore, the material is weakly or feebly magnetized in the direction of the applied magnetic field.

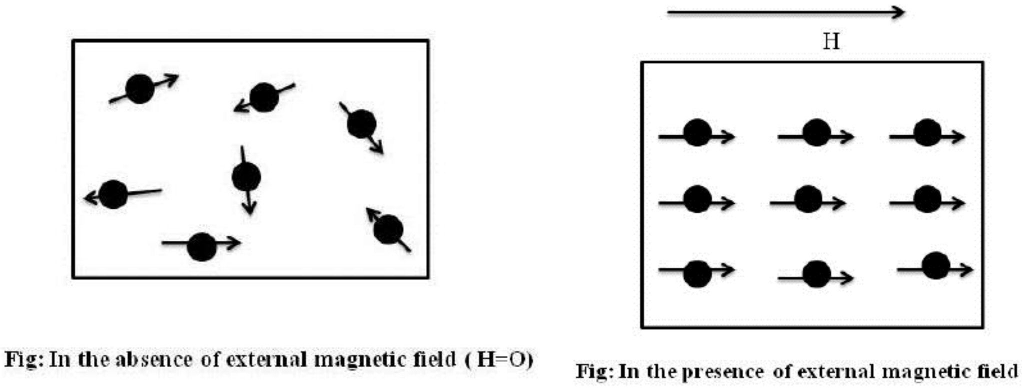
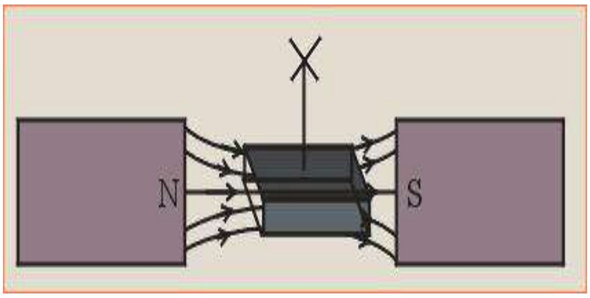


Fig: In the absence of external magnetic field (H=O)

Fig: In the presence of external magnetic field

Properties:

- They possess permanent magnetic dipole moment.
- When a paramagnetic material is placed in a magnetic field, it is feebly or weekly magnetized in the direction of applied magnetic field.
- When a paramagnetic material is placed in a magnetic field, the magnetic lines force prefers to pass through the paramagnetic magnetic material rather than air.

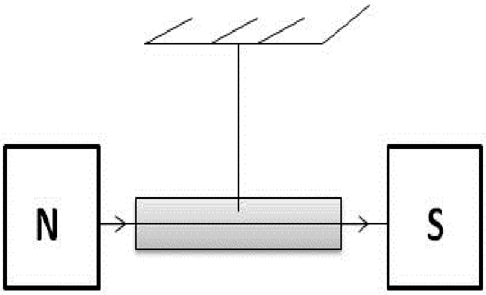


- The magnetic flux density inside is greater than that in the free space. Hence the relative permeability $\mu_r > 1$.
- The magnetic susceptibility (χ_m) is positive and small.
- The magnetic susceptibility (χ_m) is inversely proportional to the temperature.

$\chi_m = \frac{C}{T-\theta_C}$ (Curie-Weiss law)

Where C → Curie constant
T → Absolute temperature and
θ_C → Curie temperature

- When the temperature is less than the Curie temperature, paramagnetic materials becomes diamagnetic material.
- When a rod of paramagnetic material is suspended freely in a uniform magnetic field, the rod comes to rest with its axis parallel to the applied field.



3. Ferro magnetic materials:

Those materials which when placed in a magnetic field are strongly magnetized in the direction of the applied magnetic field are called Ferro magnetic materials.

Examples:
Iron, Steel, Nickel, Cobalt, etc

Cause of Ferro magnetism:

- In a Ferro magnetic material, the number of unpaired electrons is more and most of the magnetic dipole moments align parallel to each other even in the absence of magnetic field..Hence they possess permanent magnetic dipole moment even in the absence of magnetic field.
- In Ferro magnetic materials, atoms grouped into regions called *domains,* instead of acting independently like paramagnetic materials.
  *The region of space over which the magnetic dipole moments are aligned is called domain.*A typical domain contains 10^{17} to 10^{21} atoms and occupies a volume of 10^{-12} to 10^{-8} m^3.

Effect of external magnetic field:

a) In the absence of external magnetic field, the domains of a ferromagnetic material are randomly oriented. In other words, within the domain, all magnetic dipole moments are aligned, but the direction of alignment varies from domain to domain. The result is that there is no net magnetic dipole moment. Therefore, a Ferro magnetic material does not exhibit magnetism in the normal state.

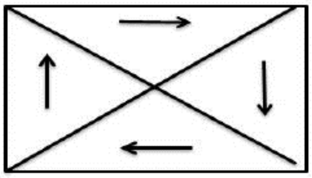


Fig: Without field

b) When a Ferro magnetic material is placed in an external magnetic field, a net magnetic dipole moment develops. This can occur in two ways:

i) By the movement of domain walls

ii) By the rotation of domain walls.

i) By the movement of domain walls:

- The movement of domain walls takes place in weak magnetic fields.
- Due to weak magnetic field applied to the material the magnetic dipole moments increases and hence the boundary of domains displaced, so that the volume of the domains changes as shown in fig.

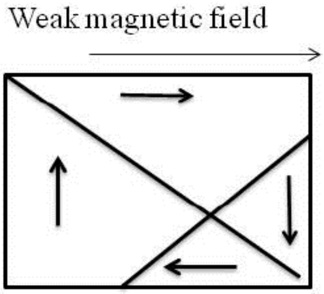


Fig: Displacement of domain walls With weak magnetic field

ii) By the rotation of domain walls

- The rotation of domain wall takes place in strong magnetic fields.
- Due to strong magnetic field applied to the material the magnetic dipole moments increases enormously and hence the domains rotate, so that the magnetic dipole moments are aligned in the direction of applied magnetic field as shown in fig..

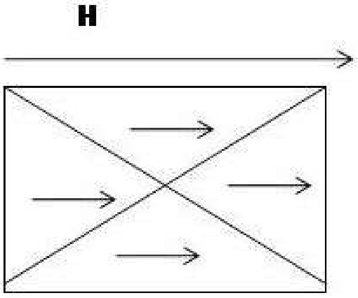
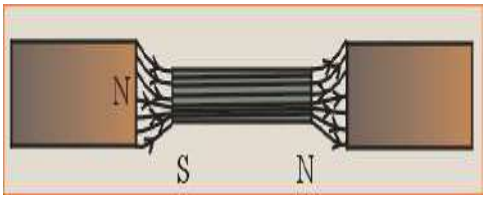


Fig: Rotation of domain walls in strong magnetic field

Properties:

- They possess permanent magnetic dipole moment.
- When a Ferro magnetic material is placed in a magnetic field, it is strongly magnetized in the direction of applied magnetic field.
- When a Ferro magnetic material is placed in a magnetic field, the magnetic lines force tend to crowed into the Ferro magnetic material.

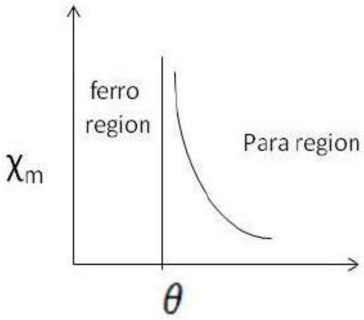


- The magnetic flux density inside is very greater than that in the free space. Hence the relative permeability $\mu_r >> 1$.
- The magnetic susceptibility (χ_m) is positive and very high.
- The magnetic susceptibility (χ_m) is inversely proportional to the temperature.

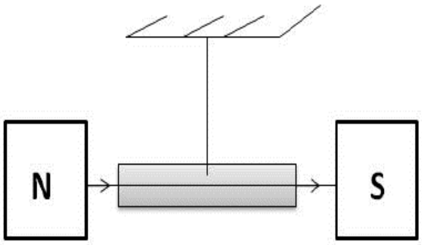
$\chi_m = \frac{C}{T-\theta_C}$ (Curie-Weiss law)

Where C → Curie constant
T → Absolute temperature and
θ_C →Curie temperature

- When the temperature is greater than the Curie temperature, ferromagnetic materials becomes Para magnetic material.



- When a rod of Ferro magnetic material is suspended freely in a uniform magnetic field, it quickly aligns itself in the direction of the applied magnetic field.



4. Antiferromagnetic materials:

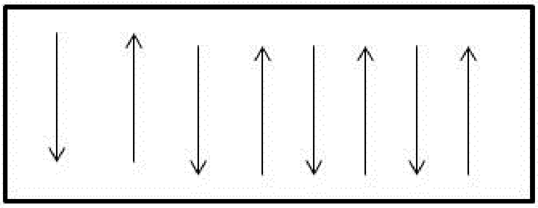
The materials which consist of anti parallel spin magnetic dipole moment with same magnitudes are known as anti ferromagnetic materials.

Examples:

Ferrous oxide (FeO),
Manganese oxide (MnO_4),
Manganese sulphideMnS),
Chromium Oxide (Cr_2O_3),
Ferrous Chloride ($FeCl_2$) etc

Properties:

- In this materials spin magnetic dipole moments are aligned in anti parallel manner.

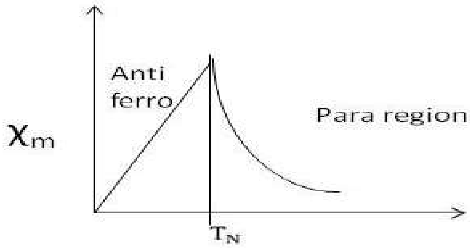


- The magnetic susceptibility is very small and positive
- The magnetic susceptibility is inversely proportional to temperature. The variation of susceptibility with temperature is shown in fig.

$\chi_m = \frac{C}{T \pm \theta_C}$

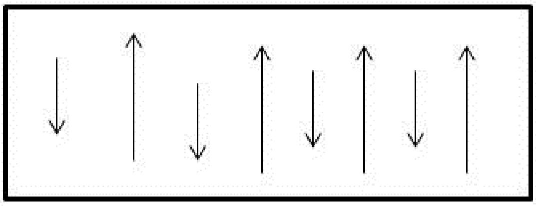
Where C → Curie constant
T → Absolute temperature and
θ_C → Curie temperature

$\chi_m$is increases gradually with temperature and attains a maximum value at Neel temperature(T_N) and then decreases with increase in temperature.



5. Ferrimagnetic materials:

The materials which consist of anti parallel magnetic dipole moments of different magnitudes are known as ferrimagnetic materials.



Examples:

Ferrites-general formula: $Me^{+2}\ Fe^{2}O_4$

Where Me^{+2} =divalent metal ions(Zn,Cu,Ni).

Properties:

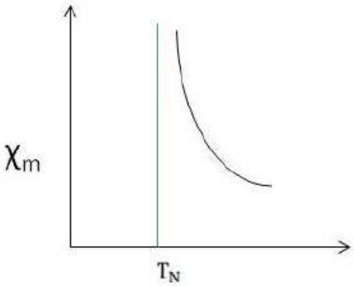
- In this materials spin magnetic dipole moments of different magnitudes are aligned in anti parallel manner.
- The magnetic susceptibility is very high and positive
- The magnetic susceptibility is inversely proportional to temperature. The variation of susceptibility with temperature is shown in fig.

$$\chi_m = \frac{C}{T \pm \theta_C}$$

Where C → Curie constant

T → Absolute temperature and

θ_C → Curie temperature



HYSTERISIS (BH-Curve):-

When a Ferro magnetic substance (e.g. iron) is subjected to a cycle of magnetization (i.e. it is magnetized first in one direction and then in the other), it is found that flux density B in the material lags behind the applied magnetizing force H. This phenomenon is known as hysteresis.

The phenomenon of lagging of flux density (B) behind the magnetizing force (H) in a ferromagnetic material subjected to cycles of magnetization is known as hysteresis.

The term hysteresis is derived from the Greek word hysterein meaning to lag behind..If a piece of ferromagnetic material is subjected to one cycle of magnetization ,the resultant B-H curve is a closed loop a b c d e f a Called hysteresis loop.

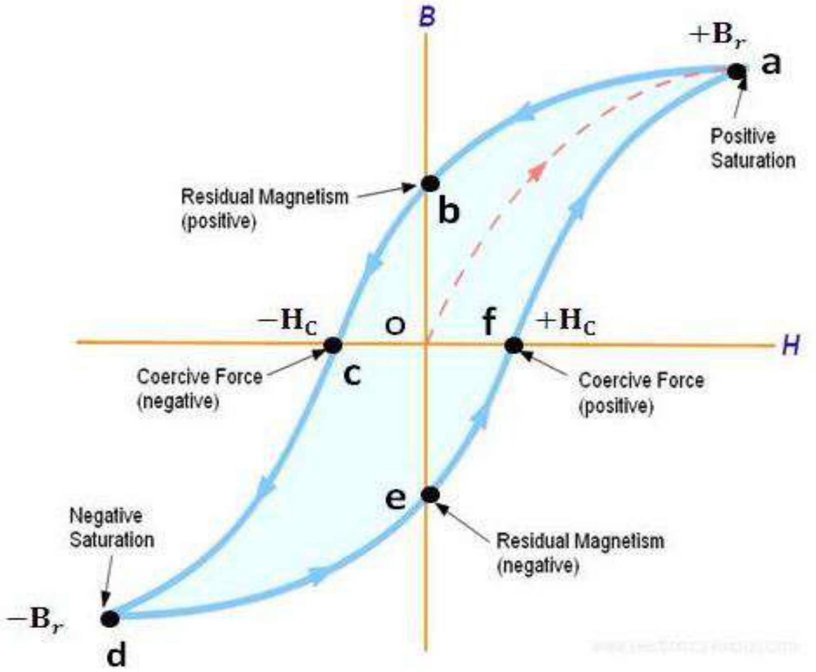


Fig: B-H Curve

Hysterisis loop:

Consider an iron-cored toroid carrying current I. If N is the total number of turns and ℓ the length of toroid, then magnetizing force is H= NI/ℓ .The value of H can be changed by varying current in the coil. We shall see that when the iron-cored toroid is subjected to a cycle of magnetization, the resultant B-H curve traces a loop a b c d e f a called hysteresis loop.

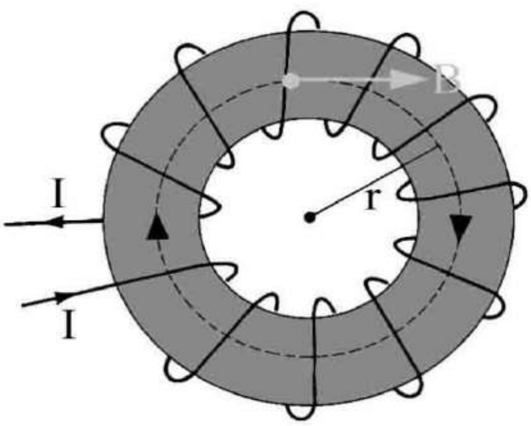


Fig : Toroid

.

(i) To start with, the toroid is unmagnetised and its situation is represented by point O in fig. As H is increased (by increasing current I), B increases along oa and reaches its saturation value B_{max} at a .At this stage, all the domains are aligned.

(ii) If now H is gradually reduced by decreasing the current in the toroid, it is found that curve follows the path ab instead of ao .At point b, H=0 but flux density in the material has a finite value $+B_r$ (=ob) called *residual flux density*. It is also called remanence or retentivity. Note that B lags behind H. This effect is called *hysteresis.*

(iii) In order to reduce flux density in the material to zero, it is necessary to apply H in the reverse direction. This can be done by reversing the current I the toroid .When H is gradually increased in the reverse direction, he curve follows the path bc. At point c, B=0 and $H= -H_C$ .The value of H needed to wipe out residual magnetism is called coercive force (H_C).

(iv) Now H is further increased in the reverse direction until point d is reached where the sample is saturated in the reverse direction ($-B_{max}$).If H is now reduced to zero point e is reached and the sample again retains magnetic flux density ($-B_r$).The remaining part of the loop is obtained by increasing current to produce H in the original direction. The curve "*a b c d e f a*" is called hysteresis loop. Thus hysteresis loop results because the domains do not become completely unaligned when H is made zero. The area enclosed by the hysteresis loop represents loss in energy. This energy appears in the material as heat.

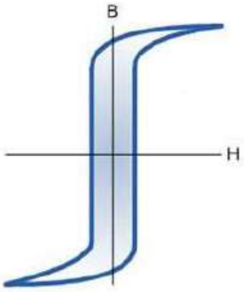
Based on the area of the hysteresis loop, the magnetic materials are classified into soft and hard magnetic materials.

Soft magnetic materials:-

The materials which can be easily magnetized and demagnetized are called Soft magnetic materials.

Properties:

- They can be easily magnetized and demagnetized and hence they show high values of susceptibility and permeability.
- Movement of domain wall is easy and hence even for small applied field large magnetization occurs.
- The nature of hysteresis loop is very narrow
- The hysteresis loop area is very small hence the hysteresis loss is also small as shown in fig.



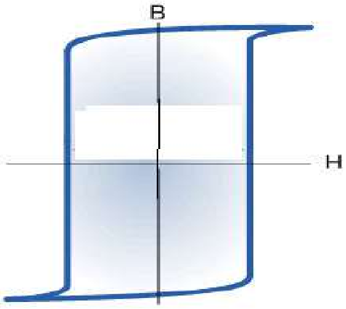
- The coercivity and retentivity values are small
- These materials are free from irregularities or impurities or imperfections
- Examples:
  Fe- Si alloys, Ni-Fe alloys, Fe-Co alloys, Ferrities and Garnets etc
- **Applications:**
  - They are used in switching devices, electromagnets,
  - They are used in matrix storage of computers.
  - They are used in motors, relays and sensors
  - They are used to make the temporary magnets.

Hard magnetic materials:-

The materials which can't be easily magnetized and demagnetized are called hard magnetic materials.

Properties:

- They can't be easily magnetized and demagnetized and hence they show low values of susceptibility and permeability.
- Movement of domain wall is not easy due to presence of impurities and hence large magnetic field is required for magnetization
- The nature of hysteresis loop is very broad.
- The hysteresis loop area is large hence the hysteresis loss is also large as shown in fig.



- The coercivity and retentivity values are high
- These materials are have irregularities or impurities or imperfections
- **Examples:**
  Carbon steel, tungsten steel, chromium steel,
  Cu-Ni-Fe alloys

Cu-Ni-Co alloys
Al-Ni-Co alloys

- **Applications:**
  - They are used in magnetic detectors ,
  - They are used in microphones.
  - They are used in magnetic separators.
  - They are used to make the permanent magnets.

Unit-4
Semiconductors

Origin of Energy bands

- We know that in a real solid, there is very large number of atoms closely packed together.
- An isolated atom consists of positively charged nucleus and negatively charged electrons revolve around the nucleus in various orbits as shown in figure.
- In the innermost orbits, some electrons are tightly bounded with nucleus are called as bound electrons or inner electrons.
- In the outermost orbits, some electrons are loosely bounded with nucleus are called as valance electrons and the force of attraction between the nucleus and these valance electrons will be very less. So that the electrons can be easily detached from their outermost orbits and become free and hence these electrons are known as free electrons.
- Let us consider the case of sodium atom with atom with atomic number 11 and mass number 23. There are 11 electrons. Two electrons revolve in the first orbit, eight electrons in the second orbit and one electron in the third orbit. The first, second and third orbits are also known as K, L and M orbits respectively.
- Each orbit has amount of energy associated with it i.e., each electron in an orbit will posses discrete energy called energy level.
- A convenient way of representing the different energy orbits of electron in an isolated atom is known as energy level diagram as shown in figure.
- The first orbit represents the first energy level; the second orbit represents the second energy level and so on.
- When large numbers of atoms are closely packed in a solid, then their corresponding energy levels are closely packed together and formed as regions called energy bands. Hence, energy band can be defined as the range of energies possessed by an electron in a solid or the large number of discrete but closely spaced energy levels.
- A convenient way of representing the different energy bands of electrons in a solid is called known as energy band diagram.

- The range of energies possessed by bound electrons or inner electrons in a solid is called inner bands.
- The range of energies possessed by valance electrons or outermost electrons in a solid is called valance band. This band completely or partially filled.
- The range of energies possessed by free electrons or conduction electrons in a solid is called conduction band.
- Energy bands are separated by regions called forbidden energy bands.
- The separation between conduction band and valance band in the energy level diagram is known as forbidden energy gap.
- No electron can allow or stay in this region, because there are no energy levels in this region.

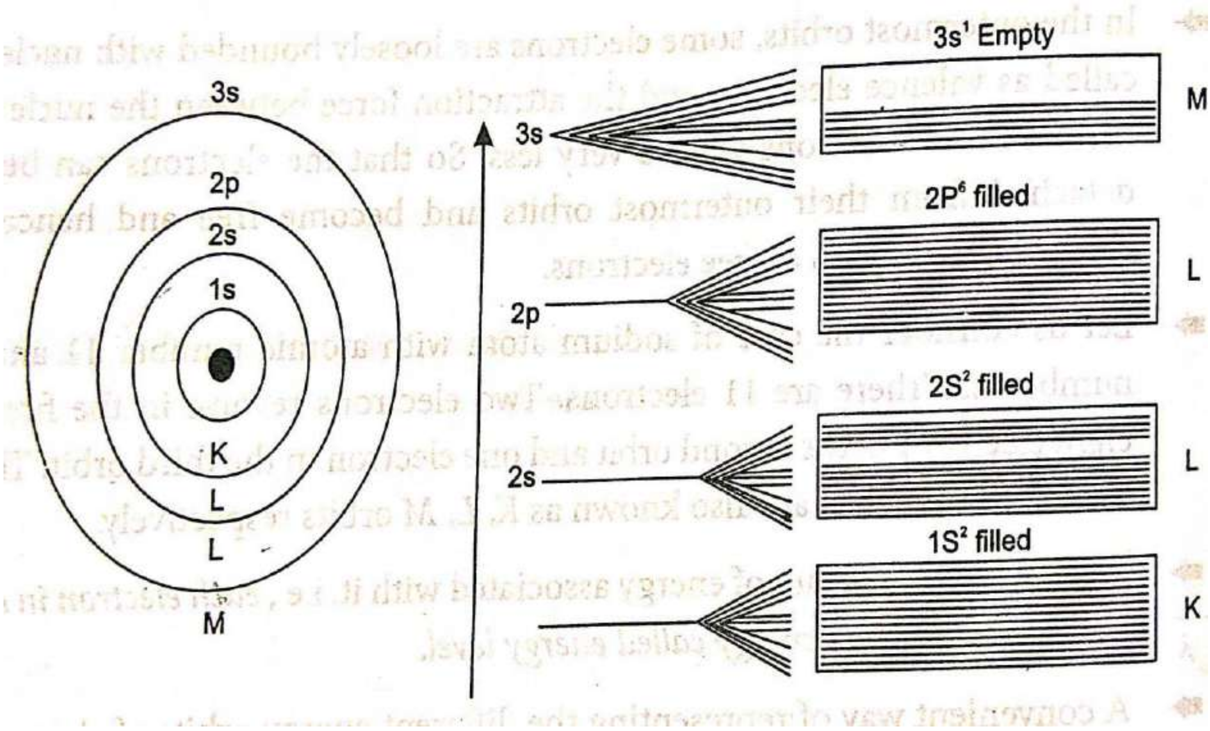
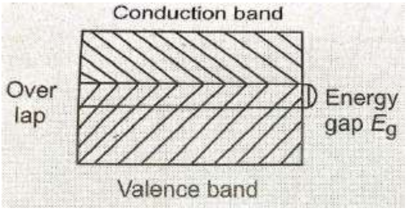
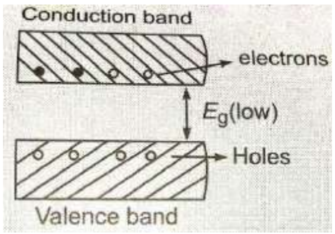
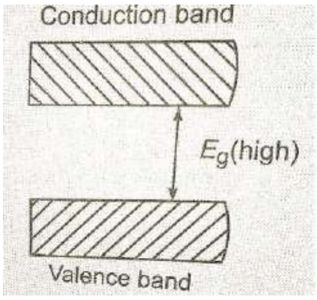


Figure: Energy levels and energy bands in sodium material.

Difference between conductor, semiconductor and insulator

Conductor	Semiconductor	Insulator
		
1. For conductor, the energy band structure will be overlapped valence and conduction band.	1. For semiconductors the energy band structure consists of separated valance band and conduction band.	1. For insulator, the energy band structure consists of well separated valance band and conduction bands with band gaps.
2. Due to overlapping of bands, there is no band gap i.e, E_g=0.	2. The band gap E_g is low.	2. The energy band gap E_g is high.
3. The electrons from the valance band freely enter into the conduction band due to overlapping of bands.	3. In the presence of external electric field, the electrons enter into the conduction band by acquiring sufficient energy to overcome the band gap.	3. Even in the presence of external electric field, the electrons does not enter into the conduction band since they are tightly bound to their nucleus of respective atoms.
4. Due to absence of bandgap, free electrons are available in the conduction band at room temperature.	4. Due to low band gap, only few electrons are available in the conduction band at room temperature.	4. Due to high band gap, no electrons are available in the conduction band at room temperature.
5. The electrical conductivity is mainly due to electrons only.	5. The electrical conductivity may be either electrons or holes depending upon the type of semiconductors.	5. No charge carriers are available for the electrical conductivity in insulator.
6. The large free electrons of conductor contribute to high electrical conductivity.	6. The large number of electrons or holes contributes to the electrical conductivity depending upon the type of semiconductors.	6. Electrical conductivity is almost zero for insulator.

Semiconductors:

A Semiconductor is a substance whose conductivity or resistivity (10^{-4} to 0.5 ohm –m) lies in between conductors and insulators.

Examples: Si, Ge, GaAs etc.

Properties:

- They have crystalline structure.
- They are formed by covalent bonds.
- They have negative temperature of coefficient of resistance.
- At T=0K, they behaves as insulators.
- When the impurities (III group or V group elements) added to the pure semiconductor, then electrical conductivity increases.

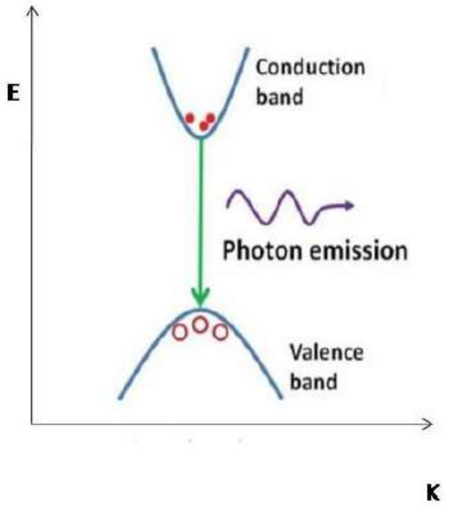
Direct Band Gap And Indirect Band Semi Conductors:

Based on the type of energy emission the semiconductors can be of two groups.

1. Direct band gap semiconductors
2. Indirect band gap semiconductors

Direct Band Gap Semiconductors:

The semiconductor, in which the minimum of the conduction band lies directly above the maximum of the valance band in the E-K space diagram(as shown in fig), is known as direct band gap semiconductor.



In a direct band gap semiconductor such as GaAS and InP, when an excited electron falls back into the valence band, the electrons and holes recombine to produce light energy.

$$\bar{e} + hole \longrightarrow h\nu \ (photon)$$

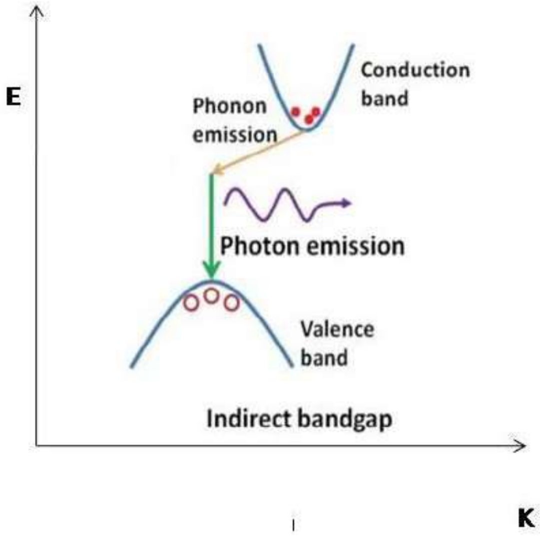
This process is known as radiative recombination and also called spontaneous emission. The life time of charge carriers is very less. These are mostly from compound semiconductors.

These direct band gap semiconductors are used to make LED's and lasers of different colours.

Ex: GaAs and InP.

Indirect Band Gap Semiconductors:

The semiconductor, in which the minimum of the conduction band does not lies directly above the maximum of the valance band in the E-K space diagram(as shown in fig), is known as indirect band gap semiconductor.



In an indirect gap semiconductor such as Si, Ge and GaP, when an excited electron falls back into the valance band, the electrons and holes recombine to generate heat and are dissipated within the material.

$$\bar{e} + hole \longrightarrow phonon$$

This process is known as non-radiative recombination. The life time of charge carriers is more. These are mostly from elemental semiconductors.

Ex: Si, Ge etc

Intrinsic semiconductors and extrinsic semiconductors

Based on purity semiconductors are classified into types,

1. Intrinsic semiconductors
2. Extrinsic semiconductors

Intrinsic semiconductor:

A semiconductor in an extremely pure form is known as intrinsic semiconductor.

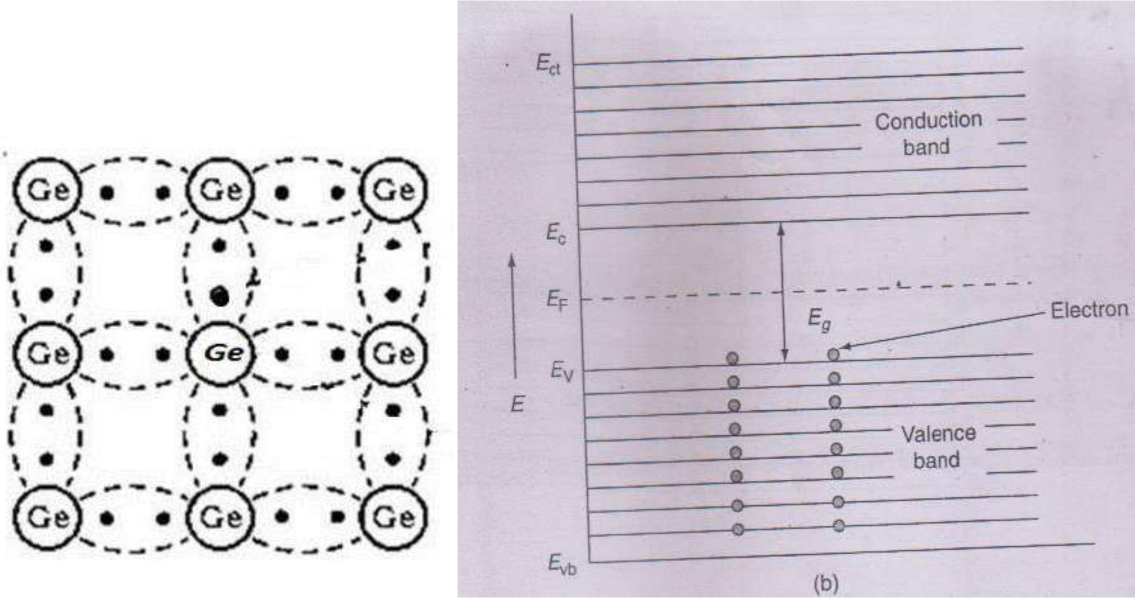
Examples: Si, Ge etc.

Explanation:

Consider a pure (semi conductor) Si or Ge Crystal. We know that Ge is a forth group element, i.e., it has four valence electrons; to get stability each Ge atom will make four covalent bonds with neighboring Ge atoms as shown in fig.

Effect of temperature on intrinsic semiconductor:

(i) At absolute zero temperature (0K):



a) Structure b) Energy band diagram

Fig: At absolute zero temperature (0K)

At absolute zero temperature, all the electrons are tightly bound by the semi conductor atoms. The inner orbit electrons are bound to the nucleus where as the valence electrons are engaged in covalent bonding. At this temperature, the covalent bonds are very strong and there are no free electrons. Therefore, the semiconductor crystal behaves as a perfect insulator at absolute zero temperature.

(ii) Above absolute zero (> 0 K)

When the temperature is raised, some of the covalent bonds in the semi conductor break due to the thermal energy supplied. The breaking of bonds set those electrons free, at the same time vacant site produced in the covalent bond called “hole”. The result is that a few free electrons exist in the semi conductor. These free electrons can constitute a tiny electric current if potential difference is applied across the semi conductor as shown in fig(c).

Fig (d) shows the energy band diagram. As the temperature is raised, some of the valence electrons acquire sufficient energy to enter into the conduction band and thus become free electrons. Under the influence of electric field, these free electrons will constitute electric current. It may be noted that each time a valence electron enters into the conduction band; a hole is created in the valence band, which acts a positive charge. For one electron set free one hole is created. Therefore, thermal energy creates hole- electron pairs;

In an intrinsic semi conductor, even at room temperature electron-hole pairs are created. Every free electron creates one hole in the valence band. Therefore the number of electrons in the conduction band and the number of holes in the valence band become equal.

When electric field is applied across an intrinsic semi conductor, the current conduction takes place by two processes namely: by free electrons and holes. Therefore the total current inside the semiconductor is the sum of currents due to free electrons and holes.

$$\boxed{I = I_{Electron} + I_{Hole}}$$

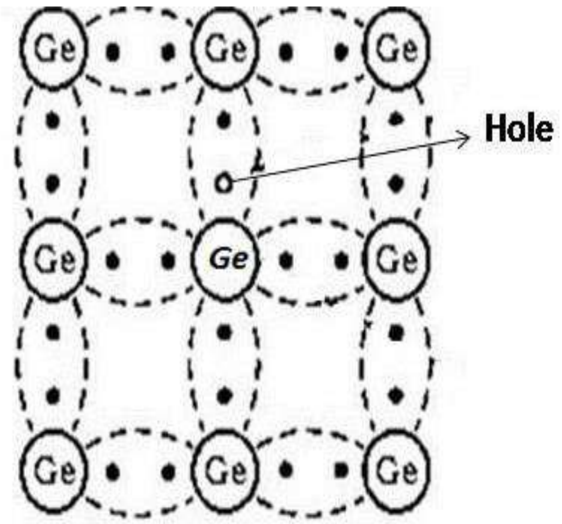
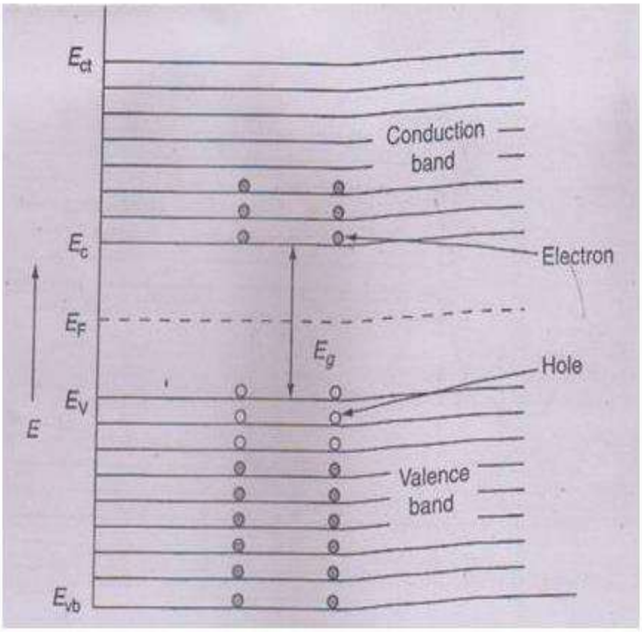


Fig: At T > 0K



Extrinsic Semiconductor:

The electrical conductivity of a pure semiconductor is very small. To increase the conductivity, impurities are added. The impurity added semi conductor is called “Extrinsic Semiconductor”.

(OR)

The conductivity of intrinsic semiconductor can be increased enormously by adding small amounts of impurity atoms [such as III or V group atoms]. This impure semiconductor is called “Extrinsic Semiconductor”.

Doping:

The process of adding impurities to a semiconductor is called doping and the impurity added is called “Dopant”.

Depending upon the type of impurity added, Extrinsic Semiconductors are classified into two types.

1. n - type Semi Conductor.
2. p - type Semi Conductor.

n – type Semiconductor:

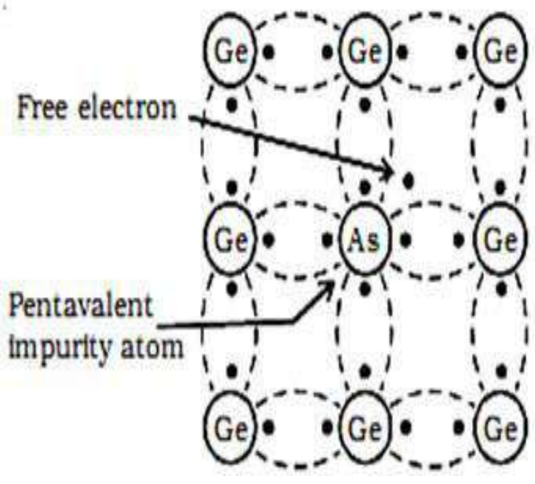
When a small quantity of pentavalent impurity is added to a pure semi conductor, it is known as ‘n – type semiconductor’.

The logical relation is

Pure Semiconductor + V Group Element = n – type Semi Conductor

Explanation:

- To explain the formation of n-type semi conductor, consider a pure (semi conductor) Germanium Crystal.
- We know that germanium atom has four valence electrons, when a small amount of pentavalent impurities like arsenic (As) is added to a germanium crystal, a large number of free electrons available in the crystal.
- The reason is simple. Arsenic is pentavalent i.e., its atom has five valance electrons. The four of the five valence electrons of Arsenic will make covalent bonds with four germanium atoms. The fifth valence electron of arsenic atom finds no place in covalent bonds and is thus free as shown in fig.
- Therefore for each arsenic atom is added, one free electron will be available in the germanium crystal.
- Thus, a small amount of arsenic impurities (As) are added to a germanium crystal, a large number of free electrons available in the crystal.
- Pentavalent impurities are also known as donor impurities because they donate free electron to semi conductor crystal.



Energy Band Description:

Fig (2) shows the energy band description of n-type semi conductor.

The addition of donor impurity to an intrinsic semi conductor creates extra energy level (called donor energy level) just below the bottom of conduction band.

1. At Absolute Zero:

At absolute zero temperature the fifth electrons of arsenic atoms occupy the donor level (represented by dashed lines) as shown in fig.

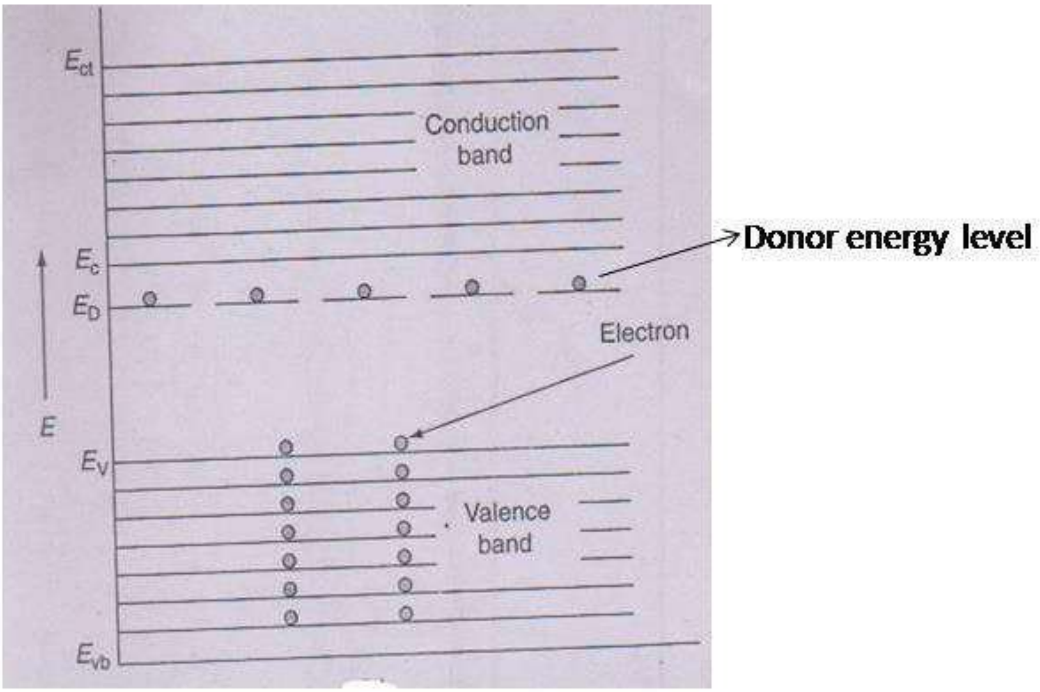
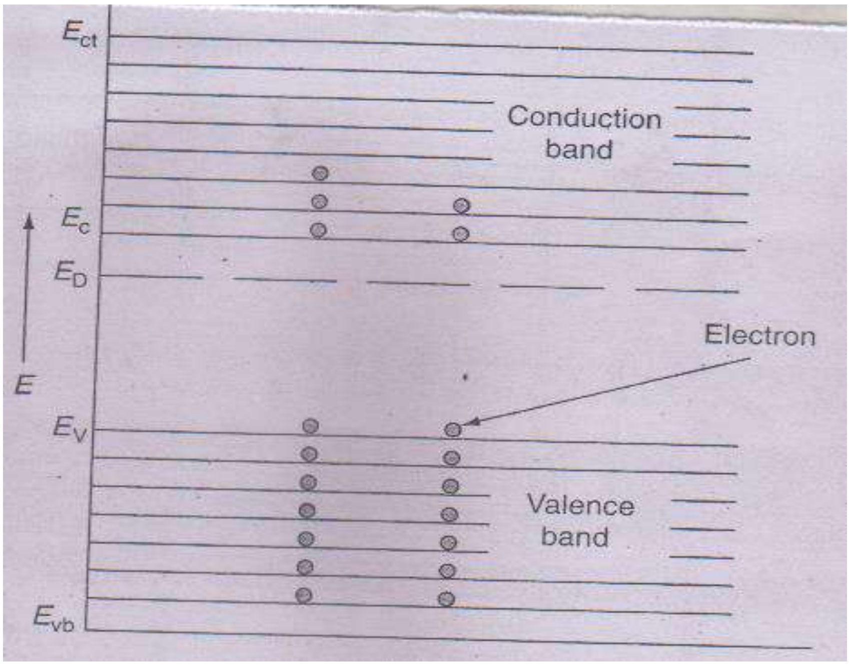


Fig: Energy band structure

2. At above Absolute Zero:

Even room temperature provides enough thermal energy to push the electrons sitting at donor energy level to conduction band. (The result is that at room temperature, there are a large number of free electrons in the conduction band as shown in fig) and also few of the covalent bonds in the semi conductor break due to the thermal energy applied. Due to the breaking of covalent bond, there will be a few holes formed in the valance band at this temperature. The result is that at room temperature, there are a large number of free electrons in the conduction band compared to the holes in valence band, i.e., Electrons are majority carriers and holes are minority carriers. So this material is called n-type semi conductor.



P-Type Semiconductor:

When a small quantity of trivalent impurity (such as indium (In) Boron (B), aluminum (Al) etc. is added to a pure semi conductor, it is called P-type semi conductor.

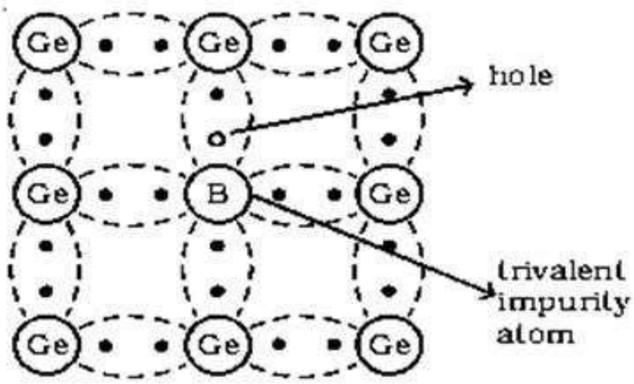
The logical relation is

Pure Semi Conductor + III Group Element = P – type Semi Conductor

Explanation:

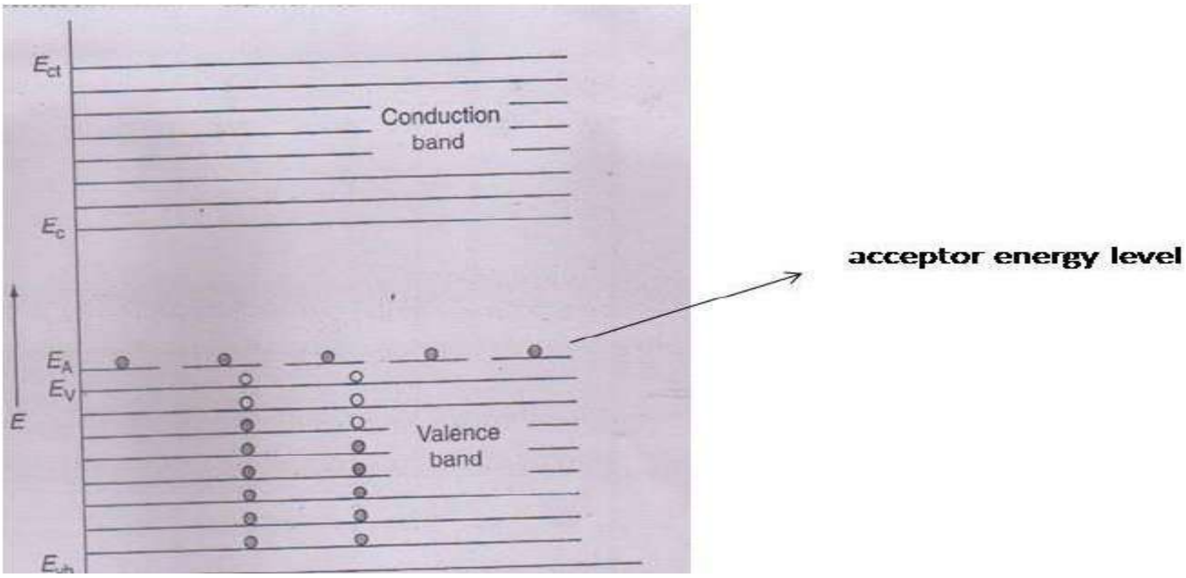
- To explain the formation of p-type semi conductor, consider a pure (semi conductor) Germanium Crystal.

- We know that germanium atom has four valence electrons, when a small amount of trivalent impurities like boron atoms (B) are added to a germanium crystal, a large number of holes available in the crystal.
- The reason is simple. Boron is trivalent i.e., its atom has thee valence electrons. The three valence electrons of Boron atom will make three covalent bonds with surrounding Ge atoms.
- The fourth bond is incomplete because short of one electron. This missing electron is called a “hole”.
- Therefore for each arsenic atom is added, one free electron will be available in the germanium crystal.
- Thus, a small amount of boron impurities (B) are added to a germanium crystal, a large number of free holes available in the crystal.
- Trivalent dopant impurities also called acceptors. Because they have need of extra electrons to complete their fourth covalent bonds formation with Ge atoms.



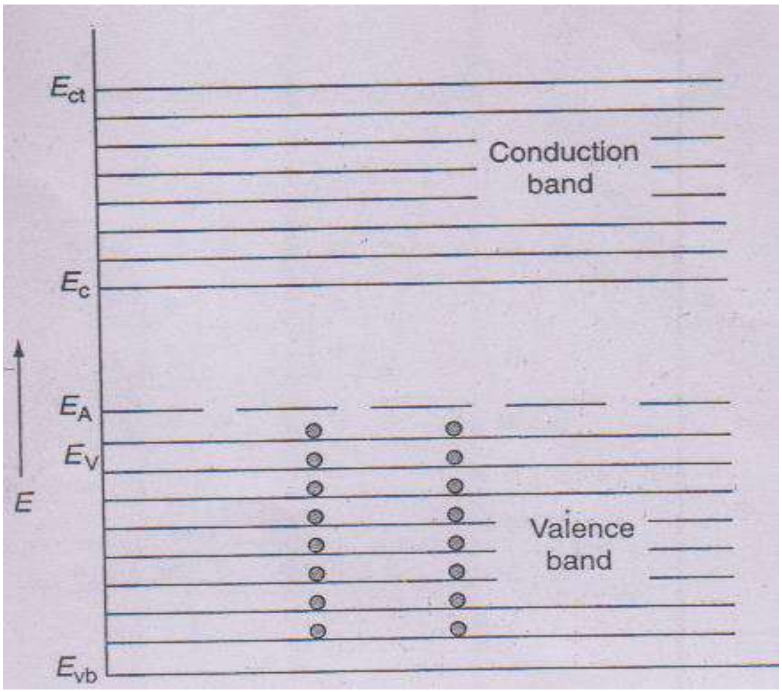
Energy Band Description:

Figure shows the energy band description of p-type semi conductor. The addition of acceptor impurity to an intrinsic semi conductor creates extra energy level (celled acceptor energy level) just above the top of the valance bond [see fig].



Since the energy difference between acceptor level and valence band energy level is very small so even at low temperature the valence band electrons gain sufficient thermal energy to occupy the acceptor energy level states as shown in fig.

After receiving an electron the, Boron atom becomes a negative ion. But at room temperature provides enough thermal energy t push the electrons in the valence band to the acceptor energy level and also due to thermal agitation a few electrons jumps from B to C.B. The result is that at room temperature there are a large number of holes in the V.B compared to the electrons in the C.B. i.e., holes are majority carriers and electrons are minority carries. So, this material is called P-type semi conductors.



Carrier concentration:

In general, the number of charge carriers (free electron or holes) per unit volume of a semi conductor material is called "carrier concentration or density."

We know that, in the semiconductors, free electron and holes are the charge carriers. So, the carrier concentration can be divided into two types.

1. Electrons carrier concentration and
2. Holes carrier concentration.

Electrons carrier concentration :

The number of free electrons present per unit volume of conduction band of the semi conductor is called "electros carrier concentration or density."

Let 'n' be the electron carrier concentration in the conduction band and is given by

$$n = 2\left(\frac{2m_e^*\pi K_B T}{h^2}\right)^{3/2} \exp^{-(E_C - E_F)}/K_B T$$

$$\text{(Or)} \quad n = N_c \exp^{-(E_C - E_F)}/K_B T \xrightarrow{\qquad} (1)$$

Where $N_c = 2\left(\frac{2m_e^*\pi k_B T}{h^2}\right)^{3/2}$ is called Pseudo constant and m_e^* is called effective mass of electron.

Holes carrier concentration:

The number of holes present per unit volume of valence band of the semi conductor is called "holes carrier concentration or density."

Let P be the holes – carrier concentration in the valence band and is given by

$$p = 2\left(\frac{2m_h^*\pi K_B T}{h^2}\right)^{3/2} \exp^{-(E_F - E_v)}/K_B T$$

$$\text{(Or)} \ P = N_V \exp^{-(E_F - E_V)}/K_B T \xrightarrow{\qquad} (2)$$

Where $N_V = 2\left(\frac{2m_h^*\pi K_B T}{h^2}\right)^{3/2}$ is a Pseudo constant.

And m_h^* is called effective mass of hole.

Carrier concentration in an intrinsic semi conductor:

We know that, at OK in an intrinsic semi conduct, the valence bond is completely filled and conduction band is completely empty. The Fermi level is exactly mid way between the valance band and conduction band as shown in fig.

As the temperature of intrinsic semi conductor is increased, electrons from the valence band get thermally excited to the conduction band. These electrons in the conduction band behave like a free particle with an effective mass m_e^*.

Similarly, the holes created by these electrons in the valance bond also behave like a free particle with an effective mass m_h^*.

Therefore, at any temperature, the number of free electron (electron carrier concentration) in the conduction band is equal to the number of holes (holes carrier concentration) in the valence band.

i.e., $n = p = n_i$

where n_i is called intrinsic carrier concentration.

Also the product $n.p = n_i^2 \xrightarrow{\qquad} (3)$

This is known as law of mass action.

Let ‘n’ be the electron carrier concentration in the conduction band and is given by

$$n = 2\left(\frac{2m_e^*\pi K_B T}{h^2}\right)^{3/2} \exp^{-(E_C - E_F)/K_B T}$$

$$\text{(Or)} \quad n = N_c \exp^{-(E_C - E_F)/K_B T} \longrightarrow (1)$$

$$\text{Where} \quad N_c = 2\left(\frac{2m_e^*\pi kBT}{h^2}\right)^{3/2} \text{ is called Pseudo constant}$$

and m_e^* is called effective mass of electron.

Let P be the holes – carrier concentration in the valence band and is given by

$$p = 2\left(\frac{2m_h^*\pi K_B T}{h^2}\right)^{3/2} \exp^{-(E_F - E_v)/K_B T}$$

$$\text{(Or)} \; P = N_V \exp^{-(E_F - E_V)/K_B T} \longrightarrow (2)$$

$$\text{Where } N_V = 2\left(\frac{2m_h^*\pi K_B T}{h^2}\right)^{3/2} \text{ is a Pseudo constant.}$$

andm$_h$* is called effective mass of hole.

Substituting equations (1) and (2) in (3) gives

$$n_i^2 = \left[N_c \exp^{-(E_C - E_F)/K_B T}\right]\left[N_V \exp^{-(E_F - E_V)/K_B T}\right]$$

$$= N_C\, N_V \exp\left[\frac{-E_C + E_F - E_F + E_V}{K_B T}\right]$$

$$= N_C\, N_V \exp^{(-E_C + E_V)/K_B T}$$

$$= N_C\, N_V \exp^{-(E_C - E_V)/K_B T}$$

$$\text{But } E_C - E_V = E_g$$

$$n_i^2 = N_C\, N_V \exp^{-(E_g)/(K_B T)}$$

$$n_i = (N_C\; N_V)^{1/2} \exp^{-(E_g)/(K_B T)} \longrightarrow (4)$$

This equation represents the intrinsic carrier concentration.

Fermi level in an intrinsic semiconductor :- .

At temperature T k, the electrons concentration (n) is equal to the holes concentration ‘P’ in the intrinsic semiconductor i.e. $n = P$

Therefore [from equations (1) and (2)]

$$N_c \exp^{-(E_C - E_F)/K_B T} = N_V \exp^{-(E_F - E_V)/K_B T}$$

$$\frac{\exp^{-(E_C - E_F)/K_B T}}{\exp^{-(E_F - E_V)/K_B T}} = \frac{N_V}{N_c} \exp^{-(E_C - E_F)/K_B T} \times \exp^{(E_F - E_V)/K_B T} = \frac{N_V}{N_c}$$

$$\exp^{[-E_c + E_F + E_F - E_V]/K_B T} = \frac{N_V}{N_c}$$

$$\exp\left[\frac{2E_F - (E_C + E_V)}{K_B T}\right] = \frac{N_V}{N_c}$$

Taking log on both sides, we have

$$\frac{2E_F - (E_C + E_V)}{K_B T} = \log\left(\frac{N_V}{N_c}\right)$$

$$E_F - \frac{(E_C + E_V)}{2} = \frac{1}{2} \log\left(\frac{N_V}{N_c}\right). K_B T$$

$$E_F - \frac{(E_C + E_V)}{2} = \frac{1}{2} K_B T \log\left(\frac{N_V}{N_c}\right)$$

$$E_F = \frac{E_C + E_V}{2} + \frac{1}{2} K_B T \log\left(\frac{N_V}{N_C}\right) \longrightarrow (5)$$

If $N_V = N_C$ then $\log\left(\frac{N_V}{N_C}\right) = 0$

Eq(5) becomes

$$E_F = \frac{E_C + E_V}{2}$$

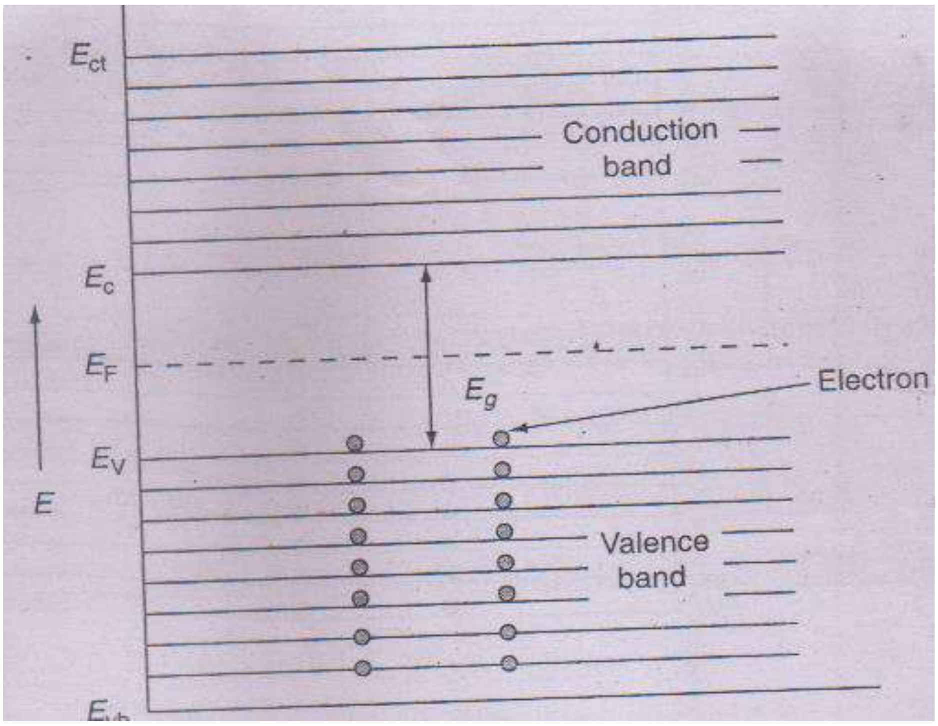
Thus, the Fermi level in an intrinsic semi conductor lies in midway between E_C and E_V.

Dependence of E_F on temperature:

The variation of Fermi level with temperature for intrinsic semiconductor is shown in fig.

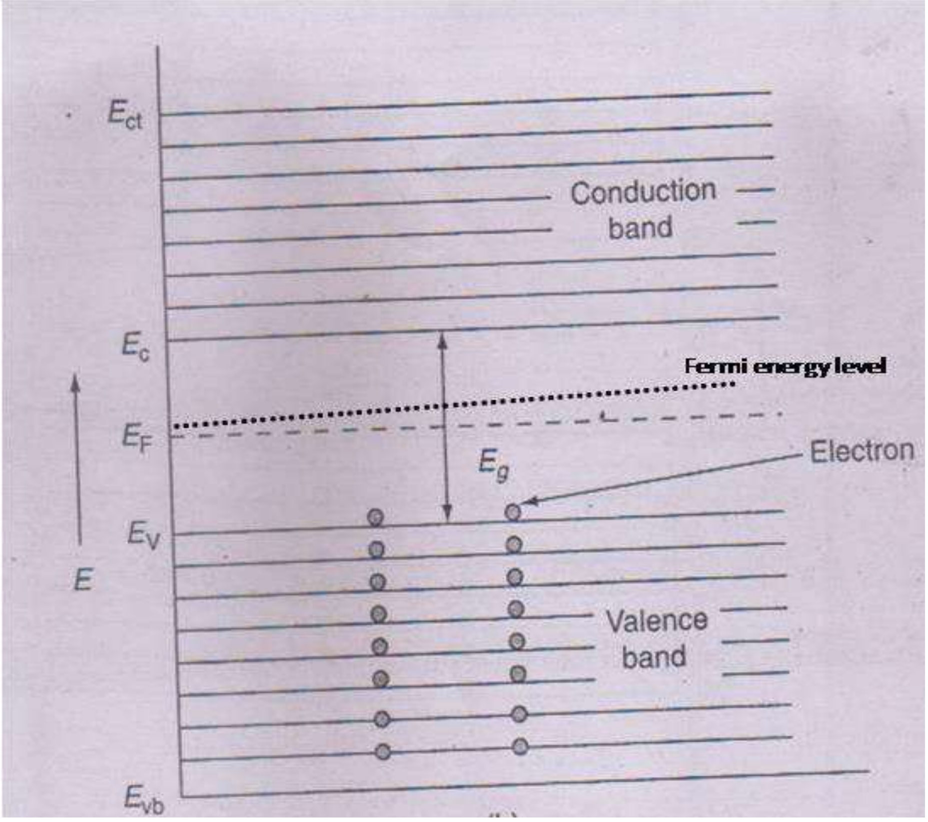
Case (a):

At 0 k, the Fermi level lies exactly in the middle of forbidden gap as shown in fig.



Case(b) :

The Fermi level slightly increases with the increase in temperature as shown in figure.



Intrinsic conductivity

Consider an intrinsic semiconductor to which a potential difference V is applied. It establishes an electric field E and the charge carriers are forced to drift in the respective directions to constitute an electric current I. The drift velocity acquired by the charge carriers is given by,

$V_d = \mu E$

Where 'μ' is the mobility of charge carriers.

Let n be the concentration of electrons in the semiconductor. Then the current density due to the electron is given by,

$J_n = neV_d$

$= ne\mu_n E \longrightarrow (1)$

Where μ_n is mobility of electron.

Similarly, current density due to holes is given by,

$J_p = pe\mu_p E \longrightarrow (2)$

Where 'p' is hole carrier concentration and μ_p is mobility of hole.

Total current density $J = J_n + J_p$

$J = ne\mu_n E + pe\mu_p E$

$= eE(n\mu_n + p\mu_p) \longrightarrow (3)$

But total current density $J = \sigma E \longrightarrow (4)$

Where σ is total conductivity.

From eq(3) and (4) we have,

$\sigma = e(n\mu_n + p\mu_p)$

For an intrinsic semiconductor, $n = p = n_i$

$\sigma = en_i(\mu_n + \mu_p)$

But, $n_i = (N_c N_V)^{1/2} e^{\frac{-E_g}{2K_B T}}$

$\sigma = (N_c N_V)^{1/2} e(\mu_n + \mu_p) e^{\frac{-E_g}{2K_B T}}$

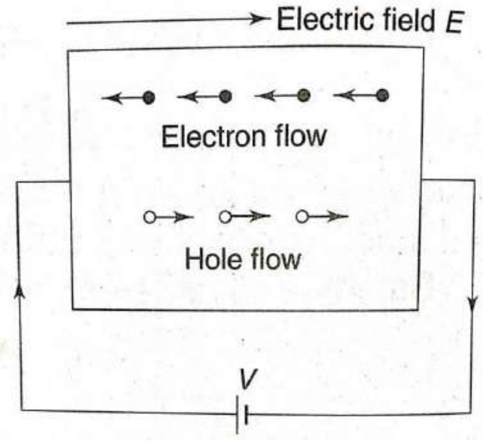


Figure: conductivity in intrinsic semiconductor

$\sigma = A e^{\frac{-E_g}{2K_B T}} \longrightarrow (5)$

Where A = $(N_c N_V)^{1/2}$ e $(\mu_n + \mu_p)$ = constant

Determination of energy band gap (E_g)

The energy gap between valance band and conduction band is called energy band gap (E_g).

From the equation (5), $\sigma = A\, e^{\frac{-E_g}{2K_BT}}$ we get,

$\rho = \frac{1}{\sigma} = \frac{1}{A}\, e^{\frac{-E_g}{2K_BT}}$

$= B\, e^{\frac{-E_g}{2K_BT}}$

Where $B = \frac{1}{A}$ = constant.

Taking log on both sides, we get

$\log\rho = \log B + \frac{E_g}{2K_BT}$

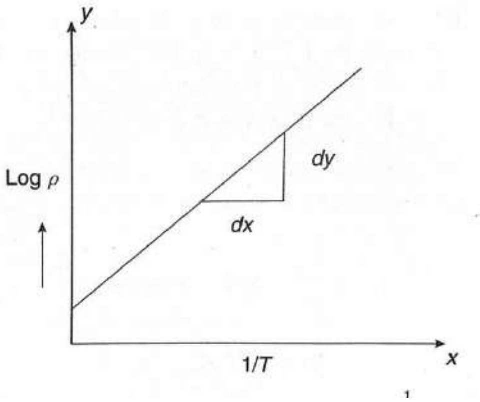


Figure: logρ vs (1/T) graph

Log of the resistivity when plotted with (1/T), produces a straight line as shown in figure. The slope of the line gives the energy band gap of the semiconductor.

Electrical conduction in semiconductors

In the semiconductors electrical conduction arises due to two processes namely

1. Drift
2. diffusion

Drift:

Definition:-

Under an applied electric field both the charge carriers i.e., electrons and holes acquire an average velocity called the drift velocity (V_d) and this process is called drift.

The drift velocity V_d is small. It is directly proportional to the electrical field E.

i.e., $V_d \propto E$

$V_d = \mu E \longrightarrow (1)$

Where μ is a constant and is called the mobility of charge carriers.

If μ_n and μ_p are the electron mobility and hole mobility respectively, then

$V_n = \mu_n E \longrightarrow (2)$ and

$V_p = \mu_p E \longrightarrow (3)$

We know that, Current $I = neAV_d$.
But in a semi conductor, the total current is the sum of electrons current and holes current.
Let n be the electron concentration and p be the hole concentration.

Therefore $I = I_n + I_p$

$= neAV_n + pAV_p$

$= eA(nV_n + pV_p)$ [From equations (2) &(3)]

$= eA(n\mu_n E + p\mu_p E)$

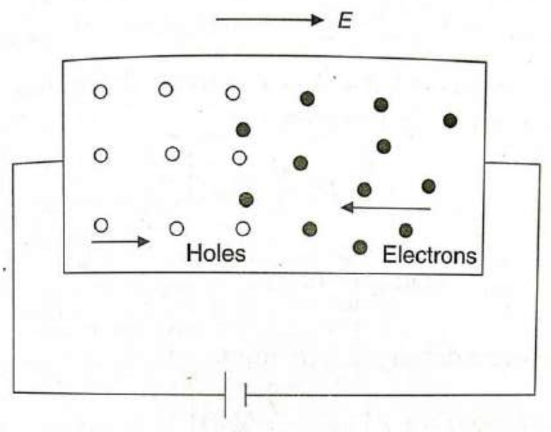
$= eA(n\mu_n + p\mu_p)E$

Current density $J = \frac{I}{A}$

$J_n = \frac{neA\mu_n E}{A}$ and

$J_p = \frac{peA\mu_p E}{A}$

Therefore, total current density $J = J_n + J_p$



$= \frac{eA(n\mu_n + P\mu_p]E}{A} = e(n\mu_n + P\mu_p)E$

Figure: Drifting of charge carriers.

Conductivity $\sigma = \frac{J}{E} = \frac{e(n\mu_n + P\mu_p]E}{E} = e(n\mu_n + P\mu_p)$

Note:-

For an intrinsic semiconductor: $n = p = n_i$

Where n_i is the carrier concentration of the intrinsic semiconductor.

$I_{inc} = eA(n_i\mu_n + n_i\mu_p)E$

$= eAn_i(\mu_n + \mu_p)E$

$\therefore$ Current density $J_{in} = \frac{I_{in}}{A} = \frac{eAn_i(\mu n + \mu p]E}{A} = en_i(\mu n + \mu p)E$

Conductivity $\sigma_{in} = \frac{J_{in}}{E} = \frac{en_i(\mu n + \mu p]E}{E} = en_i(\mu n + \mu p)$

FOR AN EXTRINSIC SEMICONDUCTOR:-

a) **For n-type semiconductor:-** The hole current is neglected, as the holes are minority carriers and the electrons are the majority carriers.

$n \simeq N_d$ is the donor electron concentration.

[∴ Where N_a can be neglected]

$$I = eAV_nN_d$$

$$= eA\mu_nEN_d$$

$$J = \frac{I}{A} = e\mu_nEN_d$$

$$\sigma = \frac{J}{E} = e\mu_nEN_d$$

b) **For P-type semiconductor :-**The electron current is neglected and $P=N_{ai}$, where N_a is the acceptor hole concentration. [∴ When N_d can be neglected].

$$I = eA\,V_P\,N_a$$

$$=eA\mu_p\,E\,N_a$$

$$J = \frac{I}{A} = e\,\mu_p\,E\,N_a$$

$$\sigma = \frac{J}{E} = e\,\mu_p\,N_a$$

Diffusion :-

Diffusion takes place due to existence of non-uniform concentration of charge carriers.

Definition:-

Due to nonuniform carrier concentration in a semiconductor, the charge carriers moves from a region of higher concentration to a region of higher concentration to a region of lower concentration. This process is known as diffusion of charge carriers.

- → According to Fick's law, the diffusion flux F is directly proportional to concentration gradient (i.e. carrier concentration per unit length) of the particles and is in the opposite direction.i.e;

Let Δn and Δp be the excess concentration in a semiconductor material.

Then Fick's law says,

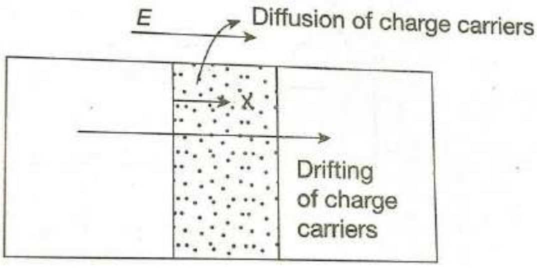


Figure: Diffusion of charge carriers

$$F \propto \frac{-\partial(\Delta n)}{\partial x} \quad \text{(For electrons)}$$

$$F = -D_n\frac{\partial(\Delta n)}{\partial x} \longrightarrow (1)$$

$$\text{And } F \propto \frac{-\partial(\Delta p)}{\partial x} \quad \text{(For holes)}$$

$$F = -D_p\frac{\partial(\Delta p)}{\partial x} \longrightarrow (2)$$

- → But the current density due to diffusion depends on charge of particle and rate of diffusion.

∴ Current density (J) = Charge of particle × Rate of diffusion $\longrightarrow$ (3)

→ The diffusion flux density due to electron is given by

$J_{n\,(Diff)} = -e\left[-D_n\frac{\partial(\Delta n)}{\partial x}\right]$ [∴ from(3)]

$J_{n\,(Diff)} = eD_n\frac{\partial(\Delta n)}{\partial x}$ $\longrightarrow$ (4)

The diffusion flux density due to hole is given by

$J_{p\,(Diff)} = +e\left[D_p\frac{\partial(\Delta p)}{\partial x}\right]$ $\longrightarrow$ (5)

Where +e and –e are charge of the hole and electron respectively. ∴ The total current density due to diffusion is given by

$J_{TOTAL,(DIFF)} = J_{n(diff)} + J_{p(diff)}$

$J_{TOTAL(DIFF)} = eD_n\frac{\partial(\Delta n)}{\partial x} - eD_P\frac{\partial(\Delta p)}{\partial x}$ $\longrightarrow$ (6)

NOTE:- The total current density due to excess electrons is the sum of the current densities due to drift and diffusion of electrons.

$J_n = J_{n(drift)} + J_{n(diff)}$

$=\Delta n\mu_n E + eD_n\frac{\partial(\Delta n)}{\partial x}$ $\longrightarrow$ (7)

Similarly,

The total current density due to excess holes is the sum of the holes densities due to drift and diffusion of holes.

$J_P = \Delta Pe\mu_P E - eD_P\frac{\partial(\Delta p)}{\partial x}$ $\longrightarrow$ (8)

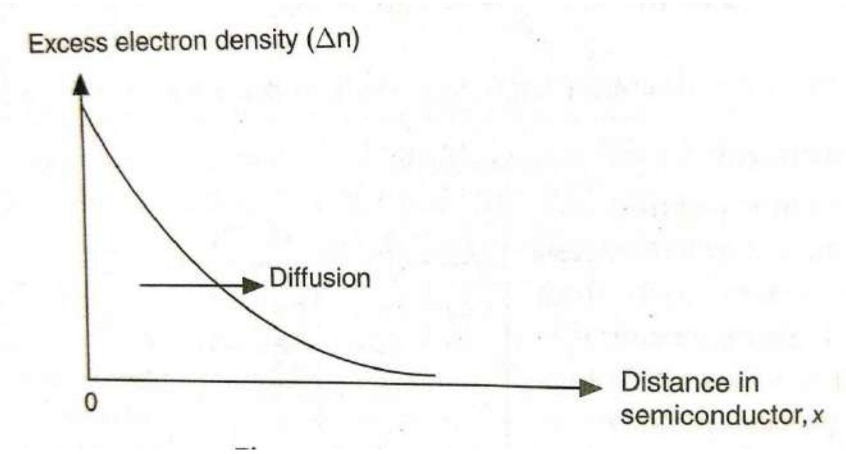


Figure: Carrier diffusion

EINSTEIN'S RELATION:-

Einstein showed the direct relation between the mobility (μ) and diffusion coefficient (D) of a semiconductor. This relation is called Einstein relation.

At any disturbance in equilibrium, the drift and diffusion current densities due to carrier concentration are equal.

$$\text{i.e.;}\ J_{n(drift)} = J_{n(diff)}$$

$$(\Delta n)eE\mu_n = D_n e\frac{\partial(\Delta n)}{\partial X} \longrightarrow (1)$$

The force on excess electrons $(\Delta n)$to restore equilibrium is given by the product of Δn and electric forces eE on each electron.Thus,

$$F = (\Delta n)eE \longrightarrow (2)$$

Equation (1) can be written as

$$F.\mu_n = D_n e\frac{\partial(\Delta n)}{\partial X}$$

$$F = \frac{D_n}{\mu_n} e \frac{\partial(\Delta n)}{\partial X} \longrightarrow (3)$$

From the kinetic theory of gases, the force on gas molecules (charge carriers) is given by

$$F = K_B T \frac{\partial(\Delta n)}{\partial X} \longrightarrow (4)$$

Comparing (3) and (4),we get

$$K_B T \frac{\partial(\Delta n)}{\partial X} = \frac{D_n}{\mu_n} e \frac{\partial(\Delta n)}{\partial X}$$

$$K_B T = \frac{D_n}{\mu_n} e$$

$$D_n = \frac{\mu_n k_B T}{e}$$

$$\frac{D_n}{\mu_n} = \frac{k_B T}{e} \longrightarrow (5)$$

Similarly for holes,

$$\frac{D_p}{\mu_p} = \frac{k_B T}{e} \longrightarrow (6)$$

From equations (5) and (6)

$$\frac{D_n}{\mu_n} = \frac{D_P}{\mu_P}$$

$$\frac{D_n}{D_P} = \frac{\mu_n}{\mu_P}$$

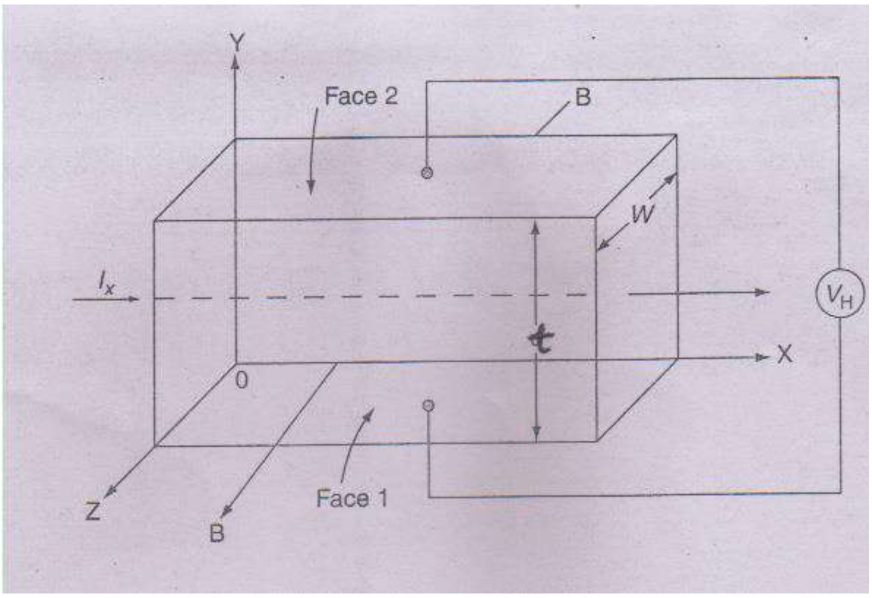
The above relation is known as Einstein’s relation.

Hall Effect:

When a current carrying semiconductor is placed in the magnetic field, then a potential difference (voltage) or electric field is developed in the material perpendicular to both the current and the applied magnetic field .this effect is known as Hall Effect and the developed potential difference (voltage) is known as hall voltage.

Explanation:

Consider a semiconductor slab of thickness‘*t*’ and width ‘*w*’ in which current (I) flowing along X-direction and magnetic field is applied along Z- direction as shown in fig.



The charge carriers inside the semiconductor experience a force due to the applied magnetic field called Lorentz force.

As a result, electrons move (drift) towards downward direction with velocity ‘V_d’ and accumulate at bottom surface (face 1) and the corresponding positive charge on upper surface (face2) of the material as shown in fig.

Thus a potential difference and electric field developed across the surface. This developed electric field along Y-direction exerts a force (F_e= eE_H) on the electrons to prevent further accumulation at the bottom surface.

Therefore, the force due to potential difference $F_e = eE_H$

The force due to magnetic field $F_L = BeV_d$

At steady state,

$F_e = F_L$

$eE_H = Bev_d$

$E_H = Bv_d \longrightarrow (1)$

For n-type:

If ‘n’ is the electrons carrier concentration in the semiconductor, then current density

$$J = -neV_d$$

$V_d = \frac{J}{-ne} \longrightarrow (2)$

Substituting eq (2) in eq (1) we get,

$$E_H = \frac{BJ}{-ne}$$

$$\boxed{E_H = R_H BJ} \longrightarrow (3)$$

Where $R_H = \frac{1}{-ne}$ is known as Hall coefficient.

For p-type:

If ‘p’ is the electrons carrier concentration in the semiconductor, then current density

$$J = peV_d$$

$V_d = \frac{J}{pe} \longrightarrow (4)$

Substituting eq (4) in eq(1) we get

$$E_H = \frac{BJ}{pe}$$

$$\boxed{E_H = R_H BJ} \longrightarrow (5)$$

Where $R_H = \frac{1}{pe}$ is known as Hall coefficient

Hall coefficient in terms of hall voltage:

If the thickness of the sample is‘*t’* and the developed voltage is V_H, then the Hall voltage

$$V_H = E_H.t \longrightarrow (6)$$

Substituting eq (6) in eq(3) we get

$$V_H = R_H B J\, t \xrightarrow{\quad} (7)$$

If '*w*' is the width of the semiconductor,

then area of the sample $A = w.t$

But current density $J = \frac{I}{A} = \frac{I}{w\,t} \xrightarrow{\quad} (8)$

Substituting eq (8) in eq(7); we get

$$V_H = \frac{R_H I B}{W}$$

$R_H = \frac{V_H w}{IB}$

$$\boxed{R_H = \frac{V_H w}{IB}}$$

This is the relation between Hall coefficient and hall voltage.

a) The electrical conductivity in a semiconductor is given by

$\sigma = ne\mu$

$\mu = \frac{\sigma}{ne}$

$$\boxed{\mu = R_H \sigma}$$

Applications of Hall effect:

1. Knowing R_H and σ, the mobility of charge carriers (μ) can be measured.
2. It is used to fine the carrier concentration

$$\boxed{n = -\frac{1}{eR_H} \text{ and } p = \frac{1}{eR_H}}$$

3. It is used to determine whether the material is p-type or n-type semiconductor.
   i.e., if R_H is negative then the material is n-type.
   if R_H is positive then the material is p-type.
4. Magnetic field (B) can be determined by knowing Hall voltage, Hall coefficient and electric current.
5. Electron carrier concentration (n) and hole carrier concentration (p) can be determined using Hall effect.

Applications of semiconductors

1. Semiconductors are used in pocket calculators, televisions and portable radios.

2. They are used in research laboratories of electronic instruments to perform tests, measurements, and numerous numerous other experimental tasks.

3. They are used in industrial control systems and automatic telephone exchanges.

4. They are used to convert large amounts of power into electric railroads.

5. They are used in military equipments, data display systems and data processing units.

6. They are used in consumer applications of space systems, mobile phones and computers.

UNIT-5

Chapter-1

Superconductivity

Introduction:

Discovery of Superconductivity:

Every material has its own resistivity or resistance. It is an inherent property of materials. This is mainly due to the scattering of electrons while interacting with the positive ions (atoms or molecules) present in the materials.
When the temperature of material is decreased, then the scattering of electrons decreases and as result resistance or resistivity decreases (i.e., conductivity increases).

Before the discovery of super conductivity, it was thought the electrical resistance of material becomes zero only at absolute zero temperature. But it is found that, in some materials the electrical resistance becomes zero, when they cooled to very low temperatures.

For example, the electrical resistance of pure mercury suddenly drops to zero, when it is cooled below 4.2 Kelvin and becomes a superconducter.This was first observed by the Dutch physicist, Heike KammerlinghOnnes on April 8, 1911. Further, the theory of super conductivity was developed in 1957 by three American physicists-John Bardeen, Leon Cooper, and John Schrieffer, through their Theories of Superconductivity, known as the BCS Theory. Therefore; Super conductivity is the phenomenon in which the electrical resistance of certain materials becomes zero at very low temperatures.

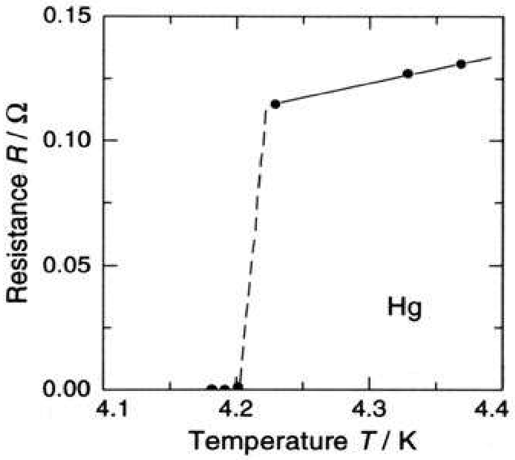


Figure: The resistance of mercury measured by Onnes.

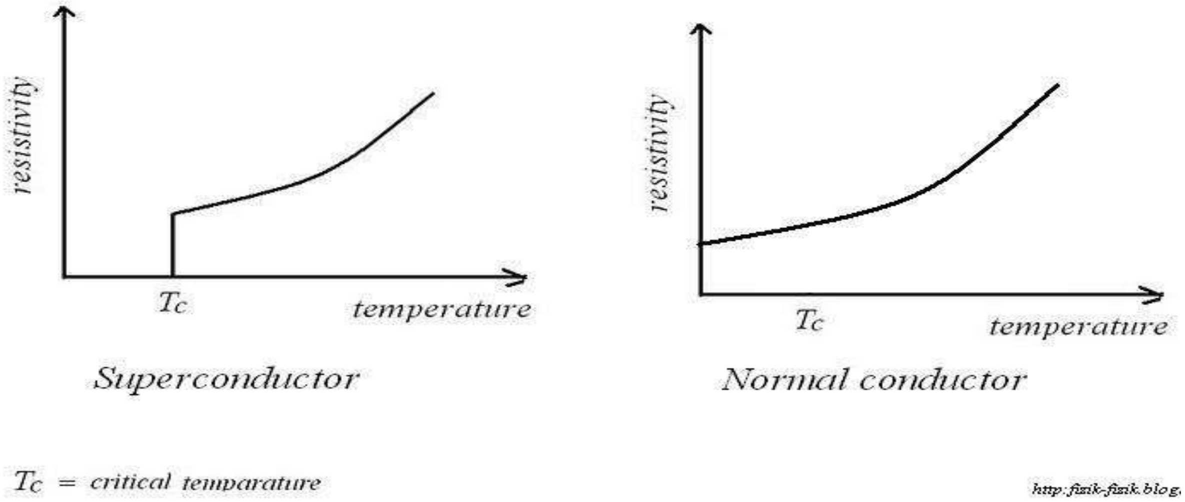


Fig: The variation of electrical resistance with temperature.

<u>Superconductors:</u>

A Superconductor is a material that loses all its resistance to flow of electric current, when it is cooled below a certain low temperature.

Examples:

Material	Type
Tungsten	Metal

Zinc	Metal
Aluminum	Metal
Tin	Metal
Mercury	Metal
Lead	Metal
NbTi	Inter metallic compound
Nb_3Sn	Inter metallic compound
Nb_3Ge	Inter metallic compound
$YBa_2Cu_3O_7$	Ceramic
TlBaCaCuO	Ceramic

Note:

- Good conducters of electric current such as silver,copper and gold are not superconducters.In fact, Super-conducting materials are not good conductors at room temperature.
- Non-transition metals such as Be and Al and transtion metals such as Nb,Mo and Zn exhibit superconductivity.
- Semiconducters like Si,Ge,Se and Te transform to a mettalic phase when subjected to high pressure and then become superconducting at low temperatures.

General properties of superconductor:

- Superconductivity is the physics of the Cooper Pairs.It occurs due to the movement of electron pairs called cooper pairs through the lattice points.These cooper pairs are formed due to the **electron –lattice –electron interaction.**
- Super conductivity is a low-temperature phenomenon.
- The electrical resistivity of a super conducting material is very low and is the order of 10^{-5} Ωcm.
- The temperature at which a normal conductor loses its resistivity and becomes a super conductor is known as transition temperature or critical temperature (T_C).Different materials will have different critical temperatures.

- When impurities are added to superconducting elements, the super conducting property is not lost, but the T_c value is lowered.
- The critical temperature(Tc) decreases with increasing isotopic mass **M**.This effect is known as isotope effect.

$$T_c \propto M^{\frac{-1}{2}}$$

- The magnetic field at which super conductor loses its super conductivity and becomes normal conductor is known as critical magnetic field H_c. Different materials will have different Hc values.
- Super conductors do not allow magnetic flux through them and behave as a diamagnetic. This property of expulsion of magnetic flux is known as meissner effect.
- Ferromagnetic materials like Fe, Co, Ni do not show superconductivity.
- The magnetic flux (Φ) lines passing through a super conducting ring due to persistent current is quantized in terms of integral multiples of $\frac{h}{2e}$

$$\Phi = \frac{nh}{2e}$$

Where n = 1, 2, 3………

Where $\frac{h}{2e}$ is known as fluxion (or) fluxiod = 2.07 X 10 $^{-15}$ Webers.

Effect of temperature-Critical temperature:

In the year 1911, Kammerligh Onnes observed that the electrical resistance of pure mercury suddenly drops to zero, when it is cooled below 4.2 Kelvin and becomes a superconductor.

The temperature at which a normal conductor loses its resistivity and becomes a super conductor is known as transition temperature or critical temperature (T_C) as show in fig.

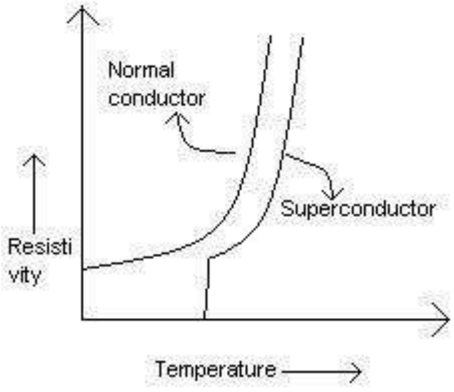


Fig: The variation of electrical resistance with temperature

Figure shows the variation of electrical resistivity with temperature. Below T_C the material is said to be in the superconducting state and above T_C the material is said to be in non-superconducting state (i.e., normal state). The value of this critical temperature varies from material to material.

Material	Type	T_c(K)
Tungsten	Metal	0.01
Zinc	metal	0.88
Aluminum	metal	1.19
Tin	metal	3.72
Mercury	metal	4.15
Lead	metal	7.2
NbTi	Inter metallic compound	9.5
Nb_3Sn	Inter metallic compound	21
Nb_3Ge	Inter metallic compound	23.2
$YBa_2Cu_3O_7$	ceramic	90
TlBaCaCuO	ceramic	125

High temperature superconductors:

Super conductors are divided into two types based on their transition temperatures.

a) Low T_C super conductors

b) High T_C super conductors

If the transition temperature is low (below 30 K), then the superconductors are known as low temperature superconductors.

If the transition temperature is high (above 30 K), then the superconductors are known as high temperature superconductors.

The first high-T_c superconductor was discovered in 1986 by Georg Bednorz and Muller, in ceramics. They found that the mixed metallic oxide of lanthanum-barium-copper (La_1Ba_2 Cu_3 O_7) exhibited superconductivity at about 30 K.

Further it has been developing by many scientists and co-research scholars.

Some examples are:

S.No	Material	T_C K
1	Ba-pbBi-O_3	38
2	$YBa_2Cu_3O_7$	92
3	$Bi_2Sr_2CaCu_2O_8$	85
4	$Bi_2Sr_2Ca_2Cu_3O_6$	110
5	$Tl_2Ba_2CuO_6$	80
6	$Tl_2Ba_2CaCu_2O_8$	108
7	$Tl_2Ba_2Ca_2Cu_3O_{10}$	125
8	$TlBa_2Ca_3Cu_4O_{11}$	122
9	$HgBa_2CuO_4$	94
10	$HgBa_2CaCu_2O_6$	128
11	$HgBa_2Ca_2Cu_3O_8$	134

Properties:

Some of the properties of high T_C superconductors as follows:

1. They have high transition temperature.
2. They are brittle in nature.
3. They are highly anisotropic.
4. They are reactive, brittle and can't be easily formed or joined.
5. They are oxides of copper in combination with other elements.
6. The hall coefficient is positive indicating that the charge carriers are holes.
7. Their behavior can't be explained by BCS theory.

Effect of magnetic field - Critical magnetic field:

Kammerlingh Onnes observed in 1913 that superconductivity vanishes if a suffiently strong magnetic field is increased.

When a magnetic field is applied to a super conductor, then particular value of applied field and below its critical temperature, it loses super conductivity and becomes a normal conductor. This minimum magnetic field required to destroy the super conducting state is called the critical magnetic field H_C.

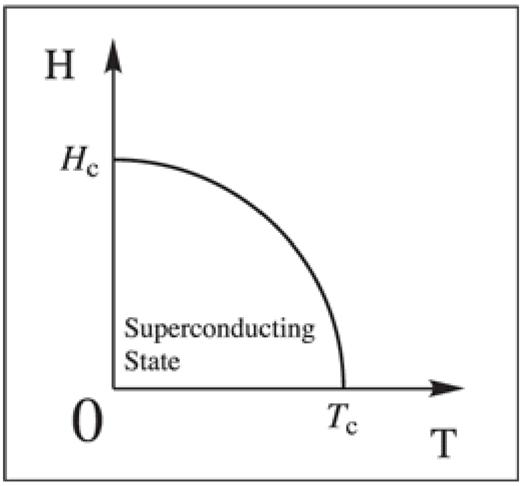


Figure: Effects of temperature and magnetic field on the superconducting state.

The critical magnetic field of a superconductor is a function of temperature. The variation of Hc with temperature is given by

$$H_C = H_0\,[1-(\frac{T}{T_C})^2]$$

Where H_C =critical magnetic field,

H_0= critical magnetic field at T=ok, and

T_C= critical temperature.

Figure shows the variation of critical magnetic field H_C as a function of temperature. The material is said to be in the superconducting state within the curve and is non super conducting (i.e., normal state) in the region the outside the curve.

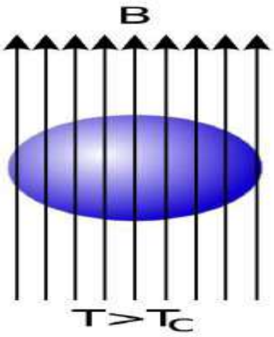
Meissner effect:

When the supper conducting material is placed in magnetic field, under the condition when $T \leq T_C$ and with $H \leq H_C$ the flux lines are expelled from the material. Thus the material exhibits perfect diamagnetism. This phenomenon is called as Meissner effect.

Explanation:

Consider a normal conducting material at room temperature. When a magnetic field H is applied to it, then it allows the magnetic lines of force to pass through it. Thus we have a magnetic tic field B in a conductor as shown in fig (1).

Now, when the material is cooled below its critical temperature ($T \leq T_C$) and with $H \leq H_C$, then the magnetic lines of forces are expelled or ejected out from the material as shown in fig (2).



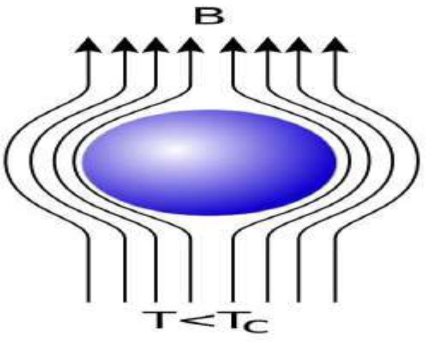


Fig (1): Normal state ($B \neq 0$ at $T > Tc$). Fig (2): Superconducting state. ($B=0$ at $T < Tc$)

Proof:

We know that, for a magnetic material the magnetic flux density is given by

$$B = \mu_0 (M+H) \dots\dots\dots\dots\dots\dots (1)$$

Where μ_0 is the permeability of free space=$4\pi x 10^{-7}$ H/ m
M is the intensity of magnetization
and, H is applied magnetic field strength.

But, for superconductor B=0

Therefore, equation (1) can be written as

$$0 = \mu_0 (M+H)$$

$$[\because \mu_0 \neq 0]$$

$$M+H=0$$

or $M=-H$

or $\frac{M}{H} = -1$

Hence,$\chi_m = \frac{M}{H} = -1$ is called the magnetic susceptibility.

Thus this means that, for a superconductor the susceptibility is negative .i.e., a super conductor exhibits perfect diamagnetism.

Types of superconductors:

Depending upon their behavior in an external magnetic field, superconductors are divided into two types:

1) Type I superconductors and 2) Type II superconductors

Let us discuss them one by one:

1) Type I superconductors:

Those superconductors which loose their superconductivity very easily or abruptly when placed in the external magnetic field are known as Type I superconductors.

Explanation:

When the super conductor is kept in the magnetic field and if the field is increased the super conductor becomes a normal conductor abruptly at critical magnetic field as shown in fig. These types of materials are termed as Type – I super conductors. After Hc, the Type I superconductor will become conductor.

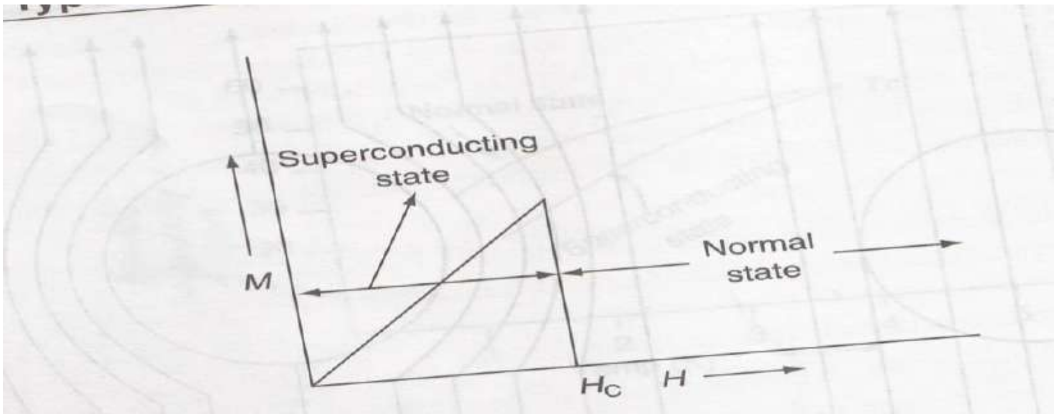


Fig: The variation of magnetization (M) with external magnetic field H in a type I Superconductor.

Properties:

- These superconductors are also known as **soft superconductors** because of this reason that is they loose their superconductivity easily and relatively small magnetic field requirement to bring them back to their normal state. These superconductors exhibit perfect and complete Meissner effect.
- Only one critical field exists for these superconductors.
- The critical magnetic field value is very low and highest value is 0.1Tesla.
- No mixed states exist in these superconductors.
- Materials with pure form are Type I superconductors.
- Examples: :

Material	Hc value in Tesla
Zinc	0.0054
Aluminum	0.0105
Mercury	0.014
Stransiam	0.03

Lead	0.08

2) Type II superconductors:

Type II superconductors was discovered by Schubnikov in 1930’s and was explained by Abrikosov in 1957.

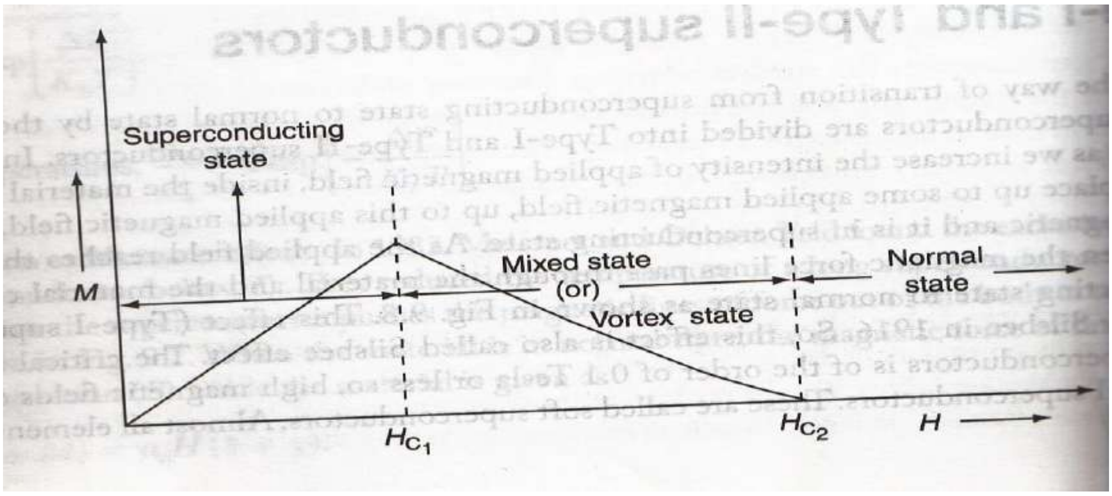
Those superconductors which lose their superconductivity gradually but not easily or abruptly when placed in the external magnetic field are known as Type II superconductors.

Explanation:

When the super conductor is kept in the magnetic field and if the field is increased, below the lower critical field Hc_1, the material exhibits perfect diamagnetism i.e., it behaves as a super conductor and above Hc_1, the magnetization decreases and hence the magnetic flux starts penetrating through the material. The specimen is said to be in a mixed state (or) vortex state between Hc_1 and Hc_2, above Hc_2 (upper critical field) it becomes a normal conductor as shown in fig.

Type – II Super conductors are also called as hard super conductors because of relatively sma

ll magnetic field requirement to bring them back to their normal state.



Properties:

- These superconductors are also known as **hard superconductors** because of this reason that is they lose their superconductivity gradually but not easily and relatively large magnetic field requirement to bring them back to their normal state.
- These superconductors exhibit Meissner effect but not completely.
- Two critical fields Hc_1 (lower critical magnetic field) and Hc_2 (upper critical magnetic field) exist for these superconductors.
- The critical magnetic field value is very high.
- Mixed states exist in these superconductors.
- Materials with impurities or alloys are of Type II superconductors.

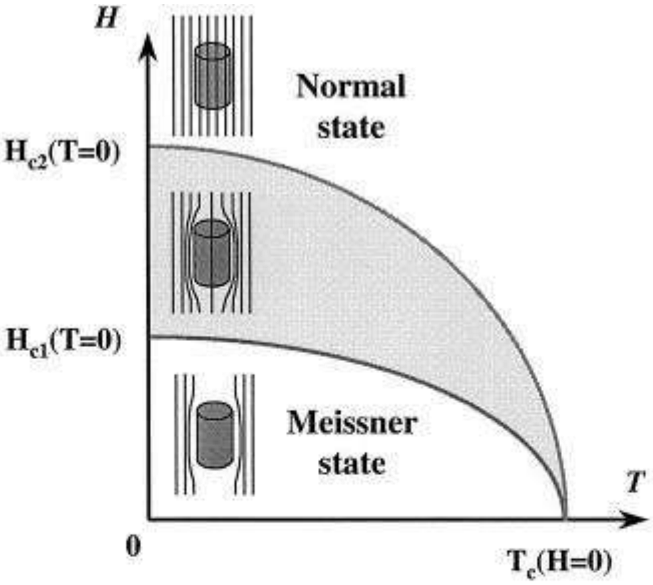


Fig: The variation of magnetization (M) with external magnetic field H in a type II Superconductor.

- Examples:

Material	Hc value in Tesla
NbN	8×10^6
Babi₃	59×10^3
Nb_3Sn	24.5
Nb_3Ge	38
$Y_1Ba_2Cu_3O_7$	300

Penetration depth:

According Meissner effect, when the material is cooled below its critical temperature ($T \leq T_C$) and with $H \leq H_C$, then the magnetic lines of forces are expelled or ejected out from the super conducting material .i.e., magnetic field drop to zero at the surface of the super conducting material.

But, in 1935, F. London and H.London observed that the applied magnetic field does not drop to zero at the surface of the super conductor but decreases exponentially as given by the equation.

$$H = H_0 \exp^{-x/\lambda} \ldots\ldots\ldots\ldots (1)$$

Where H is the intensity of magnetic field at a depth x from the surface,

Ho is the intensity of magnetic field at the surface,
and λ is called London Penetration depth.

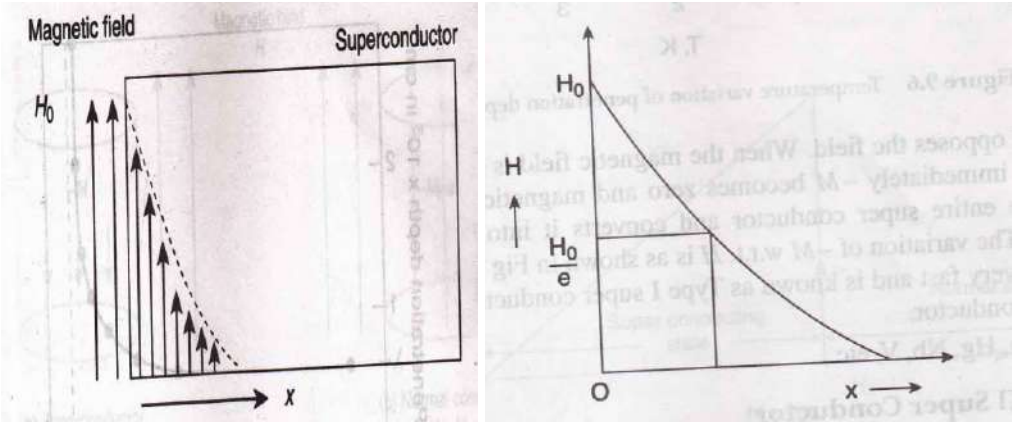


Fig: London penetration depth

To define the penetration depth, in the equation (1), let x = λ. then

$$H = H_0 \exp^{-1}$$

$$H = \frac{H}{e} \text{........................} \quad (2)$$

Hence penetration depth can be defined as the distance from the surface of the superconductor to a point inside the material at which the intensity of magnetic field is (1/e) of the magnetic field at the surface [i.e., Ho/e]. The magnetic field is likely to penetrate to a depth of 10 to 100nm from the surface of a superconductor.

The value of for some materials is given below:

Material	**λ (in nm)**
Mercury	70
Aluminum	50
Tin	50
Lead	39
Indium	64

London penetration depth depends strongly on the temperature and becomes much larger as T approaches critical temperature Tc. The relation is

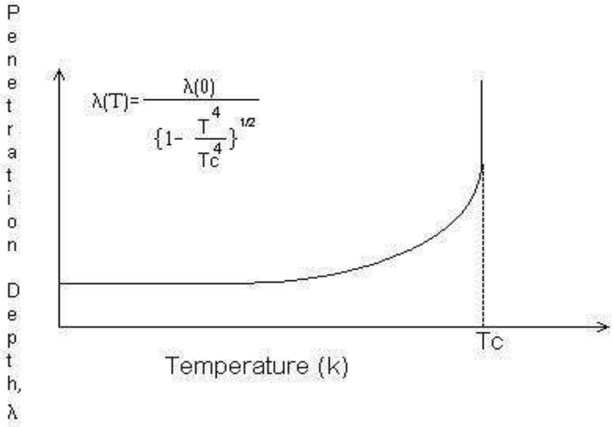
$$\lambda(\mathrm{T}) = \frac{\lambda(0)}{\left[1 - \left(\frac{\mathrm{T}}{\mathrm{T_C}}\right)^4\right]^{\frac{1}{2}}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Where λ (T) is the penetration depth at Tk

λ (0) is the penetration depth at 0k.

The variation of penetration depth with temperature as shown in fig.

According to Eq(3), λ increases with increase of T and at $T = T_C$, it becomes infinite. At $T = T_C$ the substance changes from superconducting state to normal state and hence the field can penetrate to the whole specimen, i.e., the specimen has an infinite depth of penetration.



Flux Quantization:

Generally, the magnetic flux produced in an ordinary transformer or solenoid coil is not quantized.

In 1957, A.A.Abrikosov predicted the existence of magnetic flux is quantized and experimentally confirmed by Deaver and Fairbank.

The quantization of magnetic flux is a special property of superconductors.

We know that electric charge is quantized in terms of integral multiples of charge on an electron.

$$\boxed{Q = \pm ne}$$

Where n = 1, 2, 3 ....

Similarly, the magnetic flux (Φ) lines passing through a super conducting ring due to persistent current is quantized in terms of integral multiples of $\frac{h}{2e}$

$$\boxed{\Phi = \frac{nh}{2e}}$$

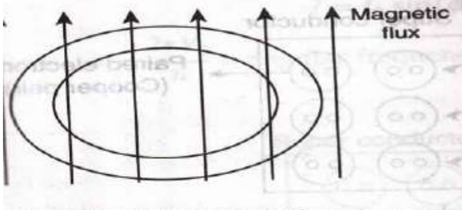
Where n = 1, 2, 3………

Where $\frac{h}{2e}$ is known as fluxion (or) fluxiod. = 2.07 X 10 $^{-15}$ Weber's.

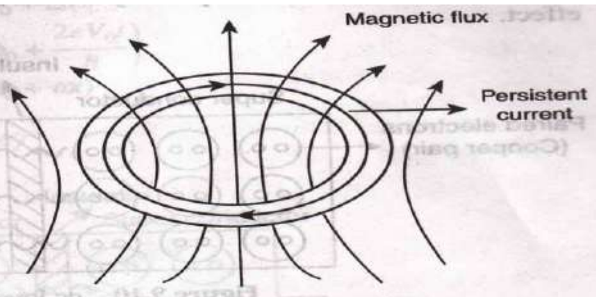
This phenomenon is known as flux quantization of super conductor.

Explanation:

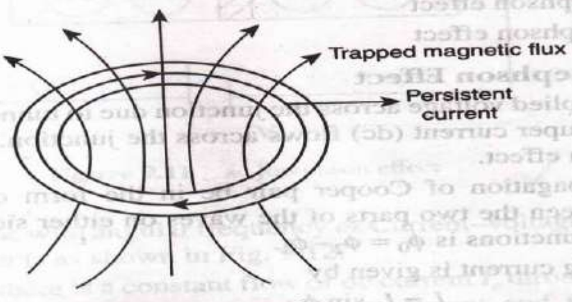
Consider a hollow normal conducting ring in a magnetic field. It allows the magnetic flux due to this field. So, we observe magnetic flux outside and inner hollow space and also on the ring as shown in fig.



(a) Hollow normal conducting ring in a magnetic field (H) ($T>T_c$)



(b) Hollow super conducting ring in a magnetic field (H) ($T<T_c$)



(c) Hollow super conducting ring in a removed magnetic field (H) ($T<T_c$)

When the temperature of the ring is lowered to its critical temperature (T_C); the ring expels the magnetic flux and it becomes a super conductor.

As a result, persistent current is set up in the material; this current will remain in the material even if we remove the applied magnetic field afterwards. This persistent current sets up magnetic force lines (i.e., flux) in the ring. This magnetic flux adjusts itself such that the total flux through the cylinder is quantized in integral multiples of $\frac{h}{2e}$.

$$\Phi = \frac{nh}{2e}$$

Where $\frac{h}{2e}$ is known as fluxion (or) fluxiod. = 2.07 X 10 $^{-15}$ Weber's.

This phenomenon is known as flux quantization of super conductor.

Occurrence of superconductivity:

Bardeen, cooper and Schrieffer (BCS) theory:

This is the first microscopic theory, based on quantum mechanics.

In 1957 John Bardeen, Leon Cooper, and John Robert Schrieffer wrote a paper called 'Theories of Superconductivity'. In this paper they proposed a theory to explain why materials lose all resistance and become superconductors at their critical temperatures. That theory for superconductivity has come to be known as the BCS theory, after the initials of the surnames of the developers. John Bardeen, Leon Cooper, and John Robert Schrieffer shared the Nobel Prize in physics in 1972 for the theory's development.

John Bardeen, Leon Cooper and J. Robert Schrieffer.

Principle:

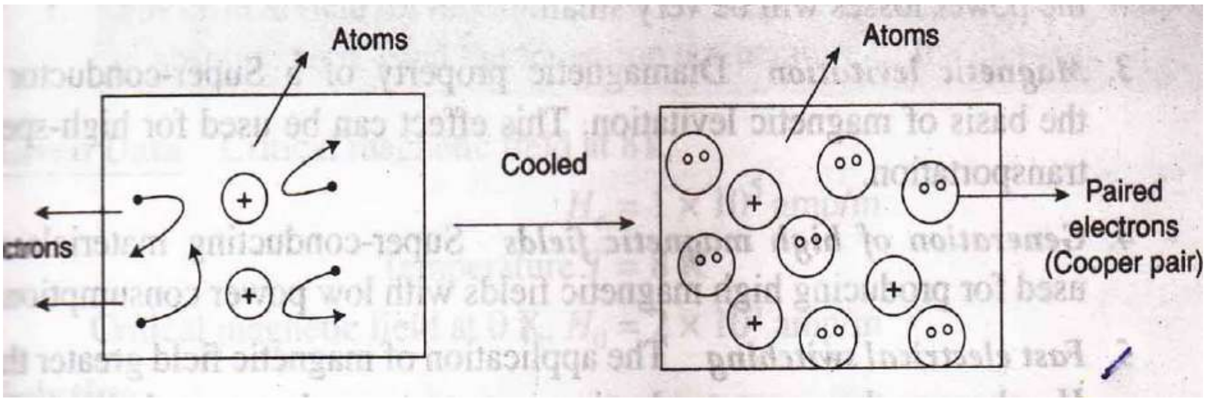
According to this theory, superconductivity occurs due to the movement of electron pairs called cooper pairs through the lattice points. These cooper pairs are formed due to electron-lattice –electron interaction.

In this interaction, electrons experience a special kind of attractive interaction, overcoming the coulomb forces of repulsion between them. As result cooper pairs (electron pairs) are formed. At low temperature, these pairs move without scattering through the lattice points and as result resistance or resistivity decreases (i.e., conductivity increases).

Explanation:

Under normal condition, the ions in the lattice vibrate about their equilibrium positions due to thermal energy. These vibrations are lattice vibrations. When electrons pass through the lattice ions in the normal state, they collide and scatter with the lattice and with each other. As a result, resistance arises in the material.

When the temperature decreases below its critical temperature, due to decrease in energy the scattering of electrons by vibrating lattice ions also decreases. As a result, electron pairs (cooper pairs) are formed. These cooper pairs move without scattering through the lattice ions and as result resistance or resistivity decreases (i.e., conductivity increases) and material becomes superconductor. These cooper pairs are formed due to electron-lattice –electron.



$\rho \neq 0$ when T > Tc $\rho = 0$ when$T < Tc$

Fig: (a) scattering of electrons in normal state (T > Tc), b) movement of cooper pairs without scattering in superconducting sate (T < Tc).

Formation of cooper pairs:

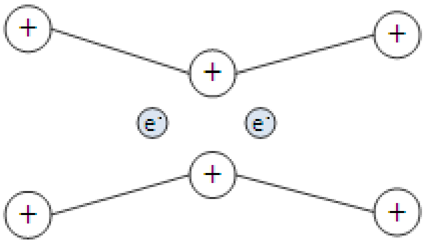
Electron-Lattice (phonon) - Electron interaction:

According to BCS theory,

Suppose an electron (1st electron) moves through the lattice, it will be attracted by the positive ion core. It suffers attractive coulomb interaction. . Due to this attraction, positive ion core is disturbed and it is called as lattice distortion. This is shown in the figure below. The lattice vibrations are quantized in terms of phonons.

At that instant, if another electron (2nd electron) moves through the distorted lattice, it will be attracted by the greater concentration of positive ion core. It also suffers attractive coulomb interaction.

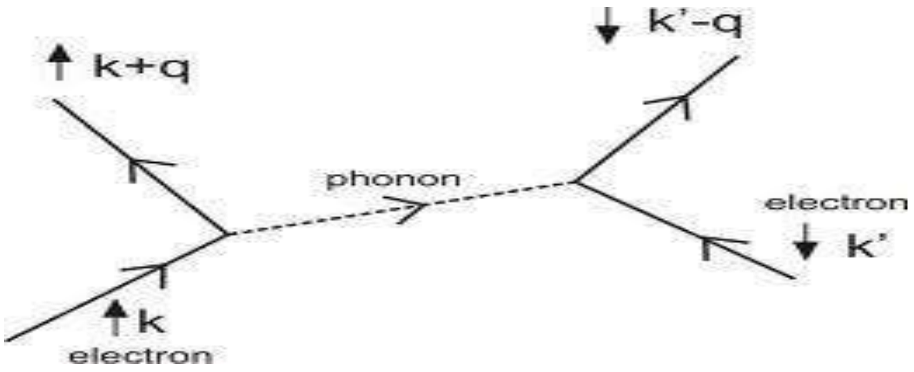
Therefore, the two electrons interact via lattice distortion or the phonon field, lowering the energy of electrons. This lowering of energy implies that the force between the two electrons is attractive. This type of interaction is called electron-lattice-electron interaction. This interaction can also be interpreted as electron –electron interaction through phonons.



Cooper pair of electrons.

Copper Pairs:

Cooper showed that the pair of electrons formed by the interaction between the electrons with opposite spins and momenta are known as cooper pairs. This interaction can be represented in terms of the wave vector of electrons as shown in fig.



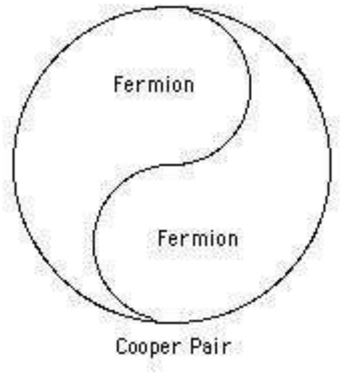
Consider the 1st electron with wave vector **k** distorts the lattice, there by emitting a phonon of wave vector **q**. This results in the wave vector for k-q for the 1st electron.

Now, if the 2nd electron with wave vector k^I, seeks the lattice, it takes up the energy from the lattice and wave vector changes to k^I +q as shown in fig.

Two electrons with wave vectors k-q and k^I +q form a pair of electrons known as cooper pairs.

Therefore, the pair of electrons formed due to electron-lattice-electron (force of attraction) by overcoming the electron-electron (force of repulsion), with equal and opposite momentum and spins with wave vectors k-q and k^I +q, and are called cooper pairs.

Occurrence of superconductivity:



©Dale Pond, 2003

When the temperature decreases below its critical temperature, due to decrease in energy the scattering of electrons by vibrating lattice ions also decreases and the pair of electrons formed due to electron-lattice-electron (force of attraction) by overcoming the electron-electron (force of repulsion), with equal and opposite spins and momenta.These cooper pairs move without scattering through the lattice ions and as result resistance or resistivity decreases (i.e., conductivity increases) and material becomes superconductor. Therefore, superconductivity occurs due to the movement of electron pairs (cooper pairs) through the lattice points.

Josephson Effect:-Introduction:

According to quantum mechanics, every moving electron has always associated with a wave nature and it can be described by wave function ψ that is always associated with a moving electron at any position(x, y, z) and at any time**'t'** and it relates the probability of finding the particle at that point and at that time.

Ivan Giaever, the American physicist demonstrated in 1960`s that an electron can tunnel through a thin insulator barrier. A significant tunnel current can exist only when the de-Broglie wavelength of the electron is comparable to or greater than the barrier thickness ($50A^O$ or less).

In1962 Brain D. Josephson the English physicist, theoretically showed that tunneling of cooper pairs was likely as tunneling of unpaired electrons.

Josephson Effect:

When two super conductors are separated by a very thin insulator (oxide layer of about $20A^o$), forms a Joseph junction and then cooper pairs can tunnel or penetrates through the thin insulator and constitute a small super current .This effect is called Josephson Effect.

Explanation:

Consider two superconducters which are joined together with help of a thin insulating layer and forms a junction called josephson junction.These super conductors consist of paired electrons known as cooper pairs in the superconducting state.These cooper pairs will try to penetrate or tunnel through the thin insulator and constitute a small super current as shown in fig.

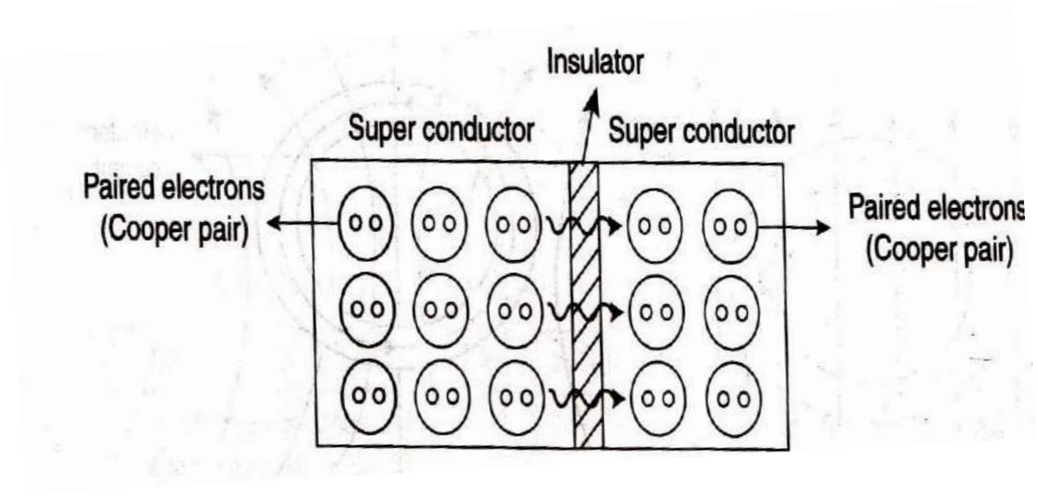


Fig: D.C. Josephson effect

Josephson Effect can be dived into two types .They are:

a) D.C. Josephson Effect
b) A.C. Josephson Effect.

D.C. Josephson Effect:

Josephson observed that the cooper pairs can tunnel from one film of superconductor into another through the thin insulator and a small direct super current flows across the junction without applied voltage across the junction. This effect is known as D.C. Josephson Effect.

Josephson showed that the dc current through the in junction is given by

$$I=I_O \mathrm{Sin}\Phi_O \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Where $\Phi_O = \Phi_2 - \Phi_1$ is the phase difference the wave functions describing cooper pair4s on both sides of the barrier, and I_O is the critical current which the junction can support and is dependent on the thickness and width of the insulating layer.

A.C. Josephson Effect:

If we apply the voltage across the junction, then ac current is produced and is given by

$$I=I_O \mathrm{~Sin}~(wt+\Phi)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Where w = angular frequency $=\frac{4\pi eV}{h}$

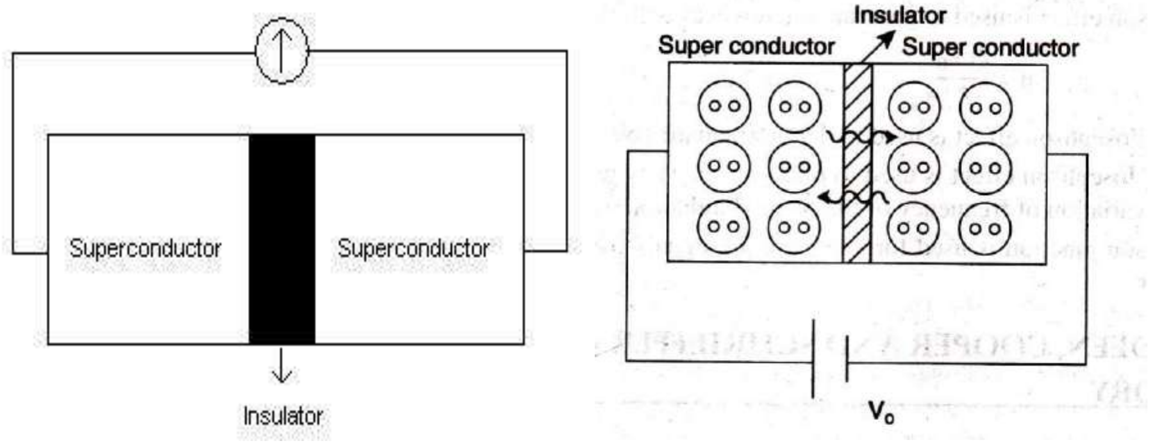


Fig: A.C. Josephson effect

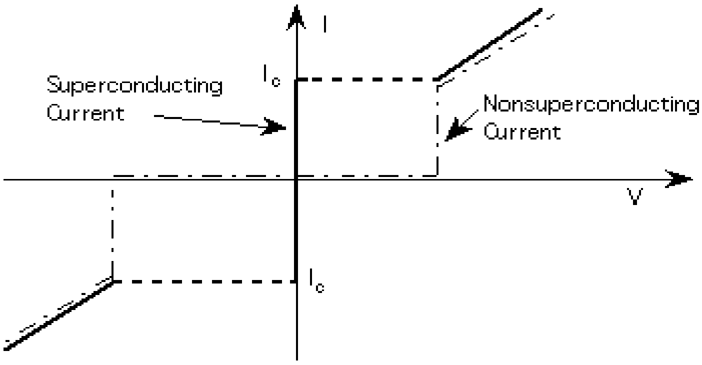


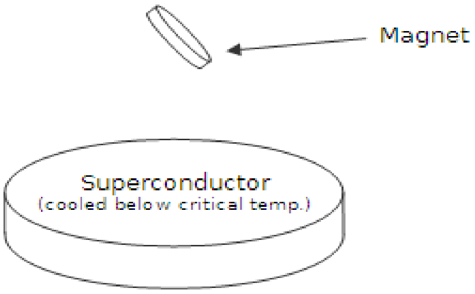
Fig: I-V characteristics of a Josephson junction

Applications of Superconductors:

Superconductors are used in the following applications:

1. MagneticLevitation

Magnetic levitation, maglev, or magnetic suspension is a method by which an object is suspended with no support other than magnetic fields.



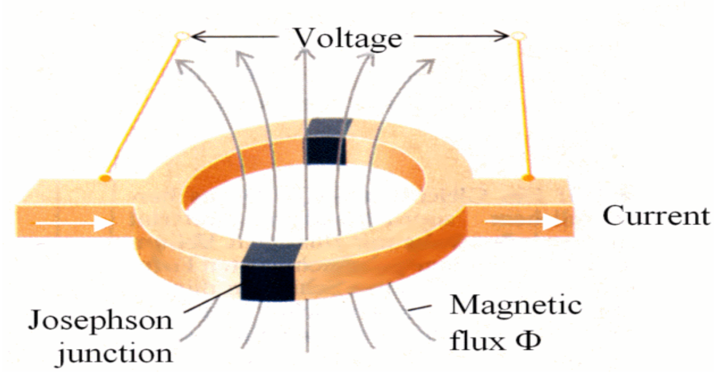
Magnetic levitation is used for high speed transportation.

For example, Maglev (magnetic levitation) trains travel 500 km/h. These work because a superconductor repels a magnetic field so a magnet will float above a superconductor (Messiner effect)– this virtually eliminates the friction between the train and the track.

Fig: Maglev train

2. SQUID :

The superconducting quantum interference device (SQUID) consists of two superconductors separated by thin insulating layers to form two parallel Josephson junctions as shown fig.



Uses:

- This device used as sensitive magnetometer to detect small magnetic fields from brain and heart.

  Threshold for SQUID: 10^{-14}
  Magnetic field of heart: 10^{-10} T
  Magnetic field of brain: 10^{-13} T.

- They are used in mine detection equipment to help in the removal of land mines.

3. **Large hadron collider or particle accelerator:**

This use of superconductors was developed at the Rutherford Appleton Laboratory in Oxford shire, UK in the 1960s. The latest and biggest large hadron collider is currently being built in Switzerland by a coalition of scientific organizations from several countries. Superconductors are used to make extremely powerful electromagnets to accelerate charged particles very fast (to near the speed of light).

4. Magnetic Resonance Imaging (MRI)

MRI is a technique developed in the 1940s that allows doctors to see what is happening inside the body without directly performing surgery.

5. Efficient Electricity Transportation:

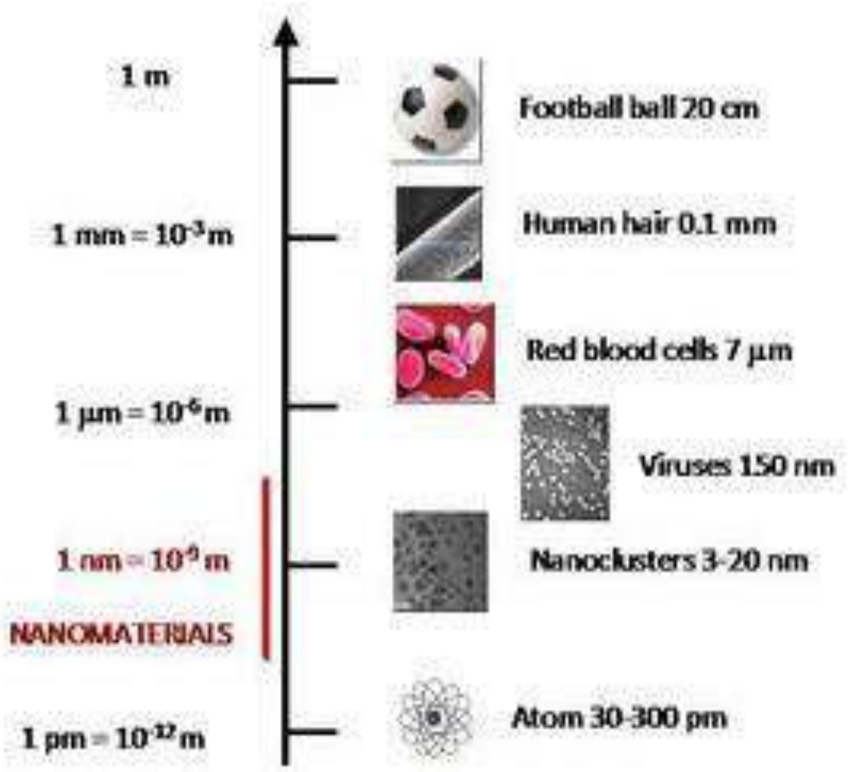
Superconducting magnets are also more efficient in generating electricity than conventional copper wire generators - in fact, a superconducting generator about half the size of a copper wire generator is about 99% efficient; typical generators are around 50% efficient.

Unit-5
Chapter-2
Nanomaterials

Introduction:-

All materials are composed of grains, which in turn comprise many atoms. The visibility of these grains depends on their size. The materials possessing grains of size ranging from 1 to 100 nm, are known as “Nanomaterials” and the developed technology is called nano technology. This was first discovered by Feynman in 1959.

- Nano technology deals with the design, manufacturing and application of devices (or) materials which lie in the sizes of 1 to 100 nm.
- Nano means 10^{-9} (one-billionth)
- A nanometer 1nm = 10^{-9} m(one-billionth of meter)
- Nano materials are found in both biological systems and man-made structures.
- Nature has been using nano materials for millions of years.



Nano materials:-

All materials are composed of grains, which in turn comprise many atoms. The visibility of these grains depends on their size. The materials possessing grains of size ranging from 1 to 100 nm, are known as "Nanomaterials".

Examples: zno, cu-Fe alloys, Ni, Pd, Pt etc.

Classification of Nanomaterials:

According to Siegel, Nanostructured materials are classified as Zero dimensional,

one dimensional, two dimensional, three dimensional nanostructures.

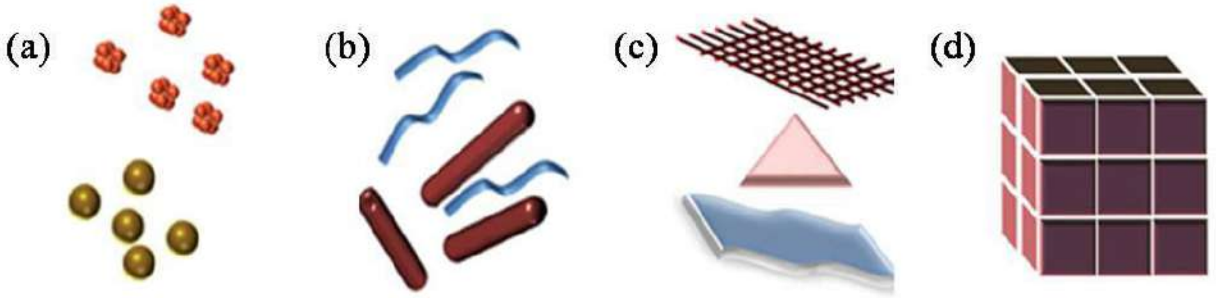


Fig: Classification of Nanomaterials (a) 0D spheres and clusters, (b) 1D nanofibers, wires, and rods, (c) 2D films, plates, and networks, (d) 3D nanomaterials.

- **One-dimensional Nanomaterial: -** In these Nanomaterial, grains will be layered in the form of multi-layers such as thin films or surface coatings.
- **Two dimensional Nanomaterials: -** This consists of ultrafine grains layered over layers or buried layers, which include nano wires and nano tubes.
- **Three dimensional Nanomaterials**: - This consists of nano meter sized grains.

Examples: precipitates, colloids and quantum dots.

Why the properties of nano materials are different?
(Or)
Basic principles of Nanomaterials:-

Two principle factors cause the properties of nano materials to differ significantly from other materials increased in surface area to volume ratio , and quantum confinement . These factors can change or enhance the properties such as reactivity, strength and electrical characteristics.

(1) INCREASE IN SURFACE AREA TO VOLUME RATIO :-

Nanomaterials have a relatively larger surface area when compared to the larger form of the material of same volume. Let us consider a sphere of radius 'r".

Its surface area = $4\pi r^2$

Its volume = $\left(\frac{4}{3}\right)\pi r^3$

Surface area to its volume ratio = $\frac{4\pi r^2}{\frac{4}{3}\pi r^3} = \frac{1}{\frac{r}{3}} = \frac{3}{r}$

Thus, when the radius of the sphere decreases, its surface area to volume ratio increases. Let us consider another example. For a one cubic volume shown in fig .4.The surface area is 6m^2.When it is divided in to eight pieces its surface area becomes 12m^2.When the same volume is divided in to 27 pieces its surface area becomes 18m^2 .Thus we find that when the given volume is divided in two smaller pieces ,the surface area increases. Hence as particle size decreases a greater proportion of atoms are found at the surface compared to those in-sides.

For example, a particle of size 30 nm has 5% of atoms its surface, at 10nm 20% of its atoms, and at 3nm has 50% of its atoms.

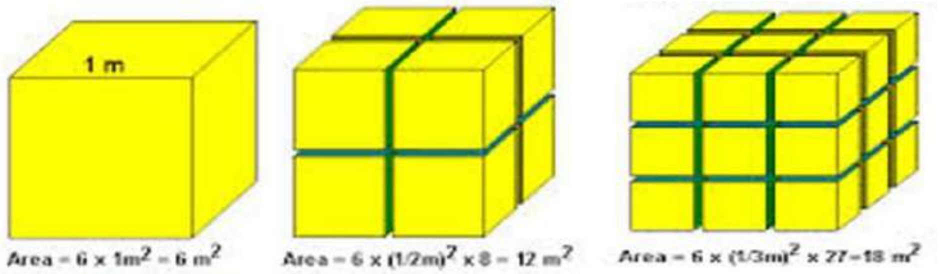
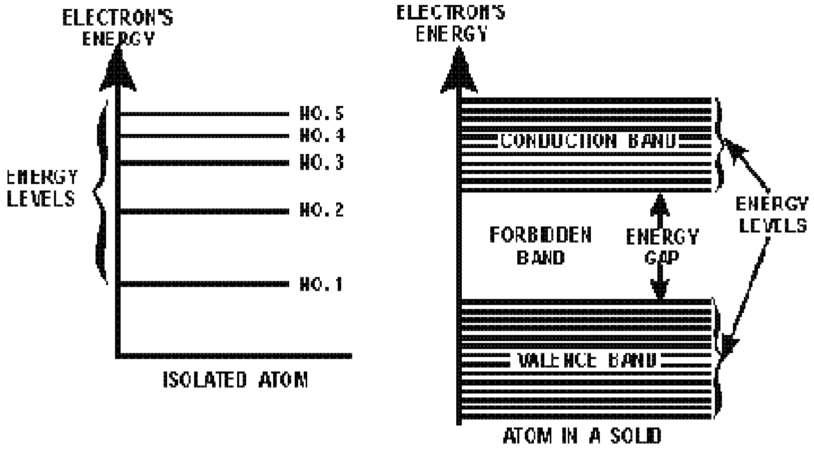


Fig: when the given object divided into smaller particles, the surface area increases.

(2) **QUANTUM CONFINEMENT**: -



The properties of materials can be studied based on their energy levels. When atoms are isolated, their energy levels are discrete. When very large number of atoms closely packed to form a solid, the energy levels split and form bands. .Nanomaterial is intermediate to the above cases. As a result, the energy levels change, when the material size is reduced to nano scale, energy levels of electrons changes. This effect is called “Quantum confinement “. This affects the optical, electrical and magnetic properties of nano materials.

Properties of Nanomaterials: -

As the size of thenano particle is very less, these materials exhibits peculiar physical, electronic, magnetic, mechanical and chemical properties.

1) **Physical properties** :-

- Since the size of the particle is very less, the particles are very close to each other and hence the inter particle spacing is very less in nano materials.
- Because of its very less size, these nano materials cannot be further divided into smaller particles and it does not have any dislocation in it. Thus we can say that they have high strength, hardness and toughness.
- The melting point of nanomaterials will be very less.

(2**) Electronic properties: -**

- The energy bands in these materials will be very narrow.
- The electrical conductivity and energy band width of same materials change when they pass from bulk phase to nano particle phase. For example, bulk silicon is an insulator; it becomes a conductor in nano phase.
- Nano material semiconductors such a Ge and Si do not snow semi conducting property.

(3) Magnetic properties: -

The nano materials shows variation in their magnetic property, when they change from bulk state to nano phase state. Some of the examples are listed in table (1).

S.No	Material	Bulk state	Nano-phase state
1	Iron, Ni,Cobalt	Ferro-magnetic	Super Para-magnetic
2	Sodium, Potassium	Para-magnetic	Ferro magnetic
3	Chromium	Anti-ferromagnetic	Para-magnetic

4. Mechanical properties :-

- Mechanical properties such as hardness, toughness, elastic modulus varies from material to material.
- They exhibit super plastic behavior.

5. Chemical properties:-

- The nano structures in chemistry are collides and these are formed in a condensed phase having size in the range from 1 to 100nm.

- The chemical reactivity of a cluster depends on the cluster size. This is useful in the preparation of catalytic agents. Some chemically inert bulk materials become good chemical catalyst in the nano phase, for example: Platinum and gold.
- In chemical reactions, the rate of increase in mass transport increases as the particle size decreases.

Fabrication of Nanomaterials:-

Nano materials can be generally fabricated in two broad ways, namely

(i) Top-down approach
(ii) Bottom – up approach

i) Top-down approach: - In which, the bulk solid materials are breaked in to nano sizes.

(ii)Bottom – up approach: - In which, the nano materials are build up from the bottom: atom by atom, molecule by molecule or cluster by cluster..

The schematic representation of the synthesis and building of nano structured materials are shown in fig (1) and fig (2) respectively.

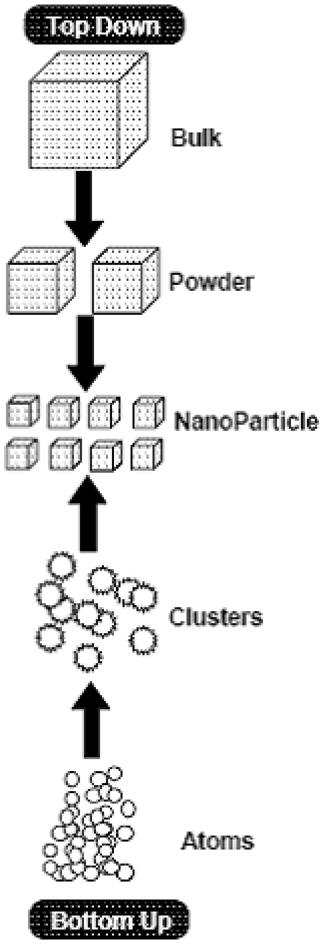


Figure: Schematic representation of the building up of Nanostructure.

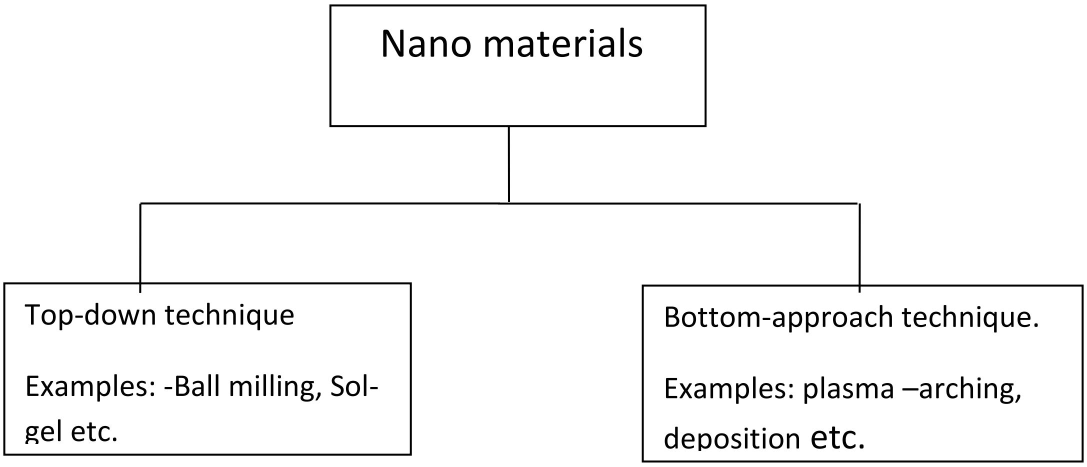


Fig:-Synthesis of nano materials.

Ball milling (or) Mechanical crushing:-

- The ball milling method is a typical example of top-down method of fabrication of nano materials.
- A **ball mill** is a type of grinder used to grind materials into extremely fine powder in order of nm.
- In this method, small balls of the material are allowed to rotate around the inside of a drum and then fall on a solid with gravity force. The balls are broken in to nano crystallites. This is also known as mechanical crushing.
- This method is used to prepare a wide range of elemental powders.For example iron nano particles of sizes 13 to 30nm can be formed.
- This method is used for produce metal oxide nano crystals (cerium oxide CeO_2 and Zinc oxide ZnO_2 .

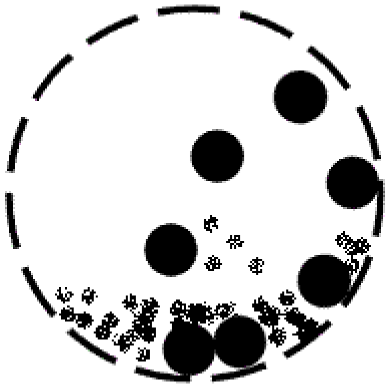


Fig: Ball milling method

Applications of ball milling method:

1. This method is useful in the preparation of elemental and metal oxide nanocrystals such as Co, Cr, AlFe, AgFe and Fe.

2. A variety of intermetallic compounds of Ni and Al can be formed.

3. This method is useful in producing new types of building materials, fireproof materials, glass ceramics, etc.

Chemical Vapour Deposition (CVD) Method:

Chemical vapour deposition (CVD) is one of the bottom up approach which is a chemical reaction followed by the deposition.

Principle:

The reactant particles are mixed with carrier gas and allowed to pass on through the hot solid surface. When the reactant particles and substrate come in contact, the heat energy initiates the chemical reaction and form nano film on the substrate surface (under vacuum condition).

In this method, nanoparticles are deposited from the gas phase.

Thermal CVD is most commonly used method. Some of the typical reaction to grow semiconducting crystals are mentioned here.

SiH_4 (gas) → Si (solid) + H_2 (gas)

For silicondioxide, deposition can be achieved (via) the reaction at 200-500°C using N_2 carrier gas.

$SiH_4 + 2O_2 \rightarrow SiO_2 + 2H_2O$

For silicon Nitrate at 750-850°C,

$3\ SiH_4 + 4NH_3 \rightarrow Si_3N_4 + 12H_2$

For polysilicon, the reaction takes at 500°C,

$SiH_4 \rightarrow Si + 2H_2$

Figure shows the apparatus used for CVD with horizontally mounted substrates.

Nanopowders of oxides and carbides of metals can also be formed if vapours of carbon or oxygen are present with the metal.

In some cases LASER may be used to assist the decomposition of the carrier gas. This is is known as LASER assisted Chemical vapour deposition method (LACVD). This method can be used to produce nanowires.

It is a well known process in which vapour of the material gets deposited on a heated surface through a chemical reaction route.

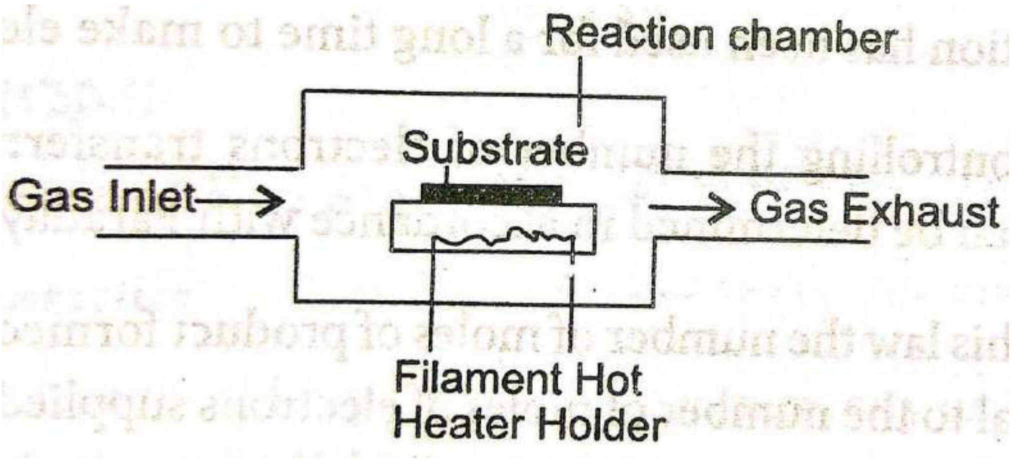


Figure: Chemical vapour deposition with substrates mounted horizontal.

Characterization of nanomaterials:

1. X-ray diffraction (XRD):

Principle:

When the X-ray beam encounters the regular three-dimensional arrangements of atoms in a crystal, most of the X-rays will destructively interfere with each-other and cancel each-other out, but in some specific directions the X-ray beams interfere constructively and reinforce one another. It is these reinforced diffracted X-rays that produce the characteristic X-ray diffraction pattern that is used for crystal structure determination.

These reflections occur only under certain conditions which satisfy the equation:

$$2d\sin\theta = n\lambda$$

The above equation is also known as Bragg's equation. Here *n* is an intiger (1,2,3,...n), lambda is the wavelength, *d* is the distance between the atomic planes, and theta is the angle of incidence of the X-ray beams. An X-ray beam takes a longer (but parallel) path because if 'reflects' off an adjacent atomic plane. This path length difference must equal an integer value of the 1 of the incident X-ray beams for the constructive interference to occur such that a reinforced diffracted beam is produced.

Explanation:

X-ray diffractometers consist of three basic elements: an X-ray tube, a sample holder, and an X-ray detector. X-rays are generated in a cathode ray tube by heating a filament to produce electrons, accelerating the electrons toward a target by applying a voltage, and bombarding the target material with electrons. When electrons have sufficient energy to dislodge inner shell electrons of the target material, characteristic X-ray spectra are produced. These spectra consist of several components, the most common being K_α and K_β. K_α consists, in part, of $K_{\alpha1}$ and $K_{\alpha2}$. $K_{\alpha1}$ has a slightly shorter wavelength and twice the intensity as $K_{\alpha2}$. The specific wavelengths are characteristic of the target material (Cu, Fe, Mo, Cr). Filtering, by foils or crystal monochrometers, is required to produce monochromatic X-rays needed for diffraction.

$K_{\alpha 1}$ and $K_{\alpha 2}$ are sufficiently close in wavelength such that a weighted average of the two is used. Copper is the most common target material for single-crystal diffraction, with CuK_α radiation = 1.5418Å. These X-rays are collimated and directed onto the sample. As the sample and detector are rotated, the intensity of the reflected X-rays is recorded. When the geometry of the incident X-rays impinging the sample satisfies the Bragg Equation, constructive interference occurs and a peak in intensity occurs. A detector records and processes this X-ray signal and converts the signal to a count rate which is then output to a device such as a printer or computer monitor.

The geometry of an X-ray diffractometer is such that the sample rotates in the path of the collimated X-ray beam at an angle θ while the X-ray detector is mounted on an arm to collect the diffracted X-rays and rotates at an angle of 2θ. The instrument used to maintain the angle and rotate the sample is termed a goniometer. For typical powder patterns, data is collected at 2θ from ~5° to 70°, angles that are preset in the X-ray scan.

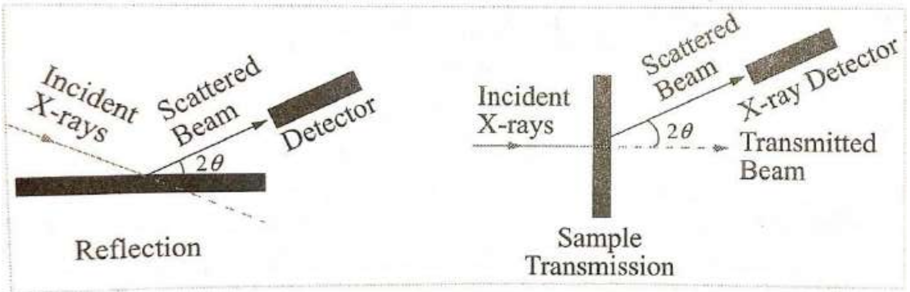


Figure: X-ray diffraction technique

A powder XRD scan from $K_2Ta_2O_6$ sample is shown below as a plot of scattering intensity vs scattering and 2θ or the corresponding d-spacing. The peak positions, intensities, widths and shapes all provide important information about the structure of the material.

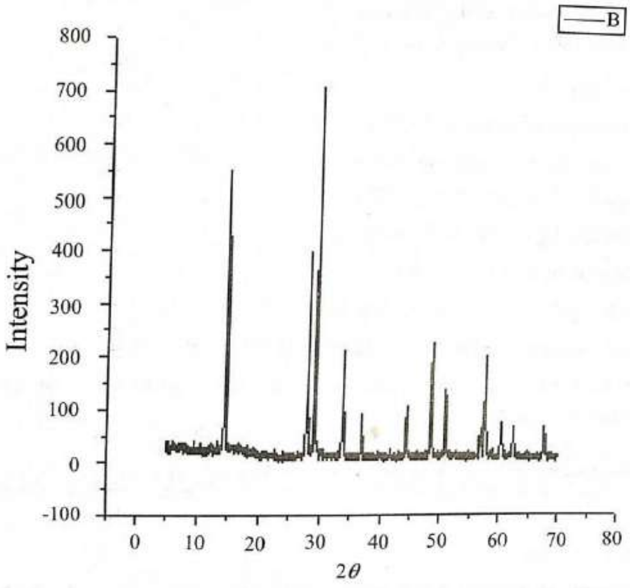


Figure: XRD patterns of $K_2Ta_2O_6$ sample.

Applications of XRD

1. Characterization of crystalline materials.

2. Identification of fine-grained minerals such as clays and mixed layer clays that are difficult to determine optically.

3. Determination of unit cell dimensions.

4. Measurement of sample purity.

5. Determine crystal structures.

2. Scanning Electron Microscope (SEM):

Principle:

When the specimen is irradiated with a fine electron beam, secondary electrons are emitted from the surface of the specimen. Topography of the surface can be observed by two dimensional scanning of the electron probe over the surface and acquisition of an image from the detected secondary electrons.

Instrumentation and Working

The SEM consists of following components include:

- Source of electrons
- Column down which electrons travel with electromagnetic lenses
- Electron detector
- Sample chamber
- Computer and display to view the images

Electrons are produced at the top of the column, accelerated down and passed through a combination of lenses and apertures to produce a focused beam of electrons which hits the surface of the sample. The sample is mounted on a stage in the chamber area and, unless the microscope is designed to operate at low vacuums, both the column and the chamber are evacuated by a combination of pumps. The level of the vacuum will depend on the design of the microscope.

The position of the electron beam on the sample is controlled by scan coils situated above the objective lens. These coils allow the beam to be scanned over the surface of the sample. This beam rastering or scanning, as the name of the microscope suggests, enables information about a defined area on the sample to be collected. As a result of the electron-sample interaction, a number of signals are produced. These signals are then detected by appropriate detectors.

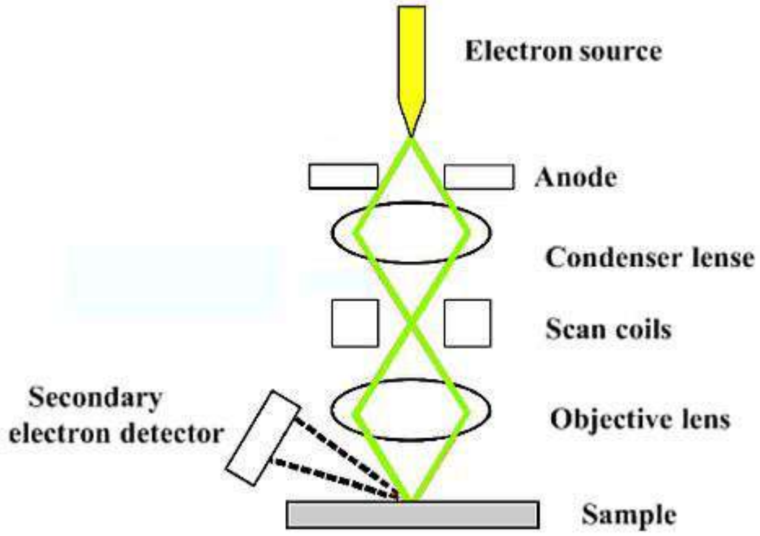


Figure: Schematic representation of scanning electron microscopy.

Sample-Electron Interaction:

The scanning electron microscope (SEM) produces images by scanning the sample with a high-energy beam of electrons. As the electrons interact with the sample, they produce secondary electrons, backscattered electrons, and characteristic X-rays.

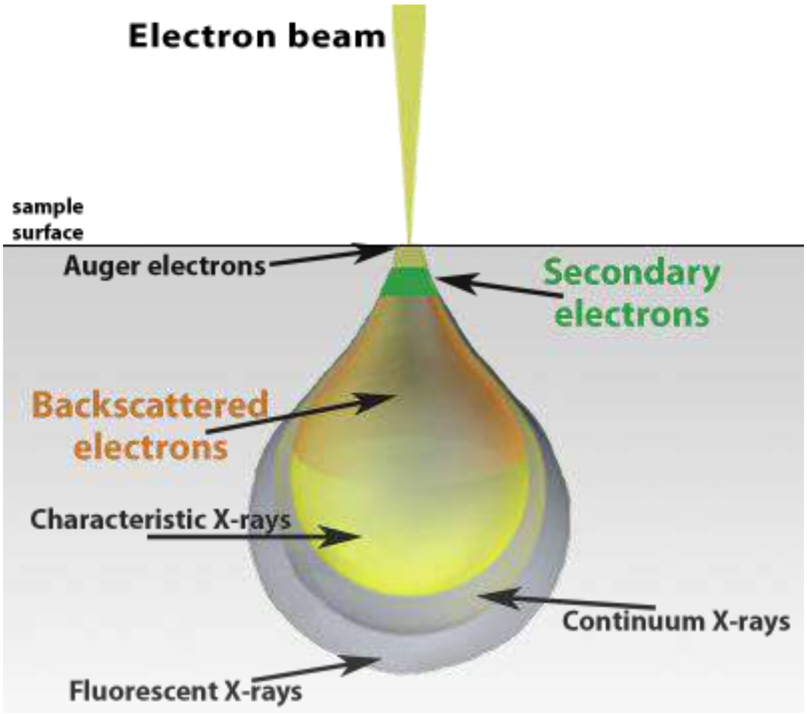


Figure: Emission of various electrons and electromagnetic waves from the specimen.

These signals are collected by one or more detectors to form images which are then displayed on the computer screen. When the electron beam hits the surface of the sample, it penetrates the sample to a depth of a few microns, depending on the accelerating voltage and the density of the sample. Many signals, like secondary electrons and X-rays, are produced as a result of this interaction inside the sample.

The maximum resolution obtained in an SEM depends on multiple factors, like the electron spot size and interaction volume of the electron beam with the sample. While it cannot provide atomic resolution, some SEMs can achieve resolution below 1 nm. Typically, modern full-sized SEMs provide resolution between 1-20 nm whereas desktop systems can provide a resolution of 20 nm or more.

Applications of SEM:

1. In addition to topographical, morphological and compositional information, SEM can detect and analyze surface fractures, provide information in microstructures, examine surface contaminations, reveal spatial variations in chemical compositions, provide qualitative chemical analyses and identify crystalline structures.

2. SEMs can be as essential research tool in fields such as life science, biology, gemmology, medical and forensic science, metallurgy.

3. In addition, SEMs have practical industrial and technological applications such as semiconductor inspection, production line of miniscule products and assembly of microchips for computers.

4. SEM data is used for semiconductor investigation.

Applications of Nano materials:-

(i) Mechanical Engineering :-

(a) Since they are stronger, lighter etc; they are used to make hard metals.
(b) Nano-MEMS (Micro-Electro Mechanical Systems) are used in ICS, optical switches, pressure sensors, mass sensors etc.

(ii) Electrical Electronics and Communication Engineering:-

(a) Orderly assembled nano materials are used as quantum electronic devices and photonic crystals.
(b) Some of the nano materials are used as sensing elements. Especially the molecular nano materials are used to design the robots, assemblers etc.
(c) They are used in energy storage devices such as hydrogen storage devices, magnetic refrigeration and in ionic batteries.
(d) Dispersed nano materials are used in magnetic recording devices, rocket propellant, solar cells, fuel cells etc.
(e) Recently nano robots were designed which are used to remove the damaged cancer cells and also to modify the neuron network in human body.

(iii) Computer Science Engineering and IT:-

(a) Nano materials are used to make CD’s and Semiconductor laser.

(b) These materials are used to store the information in smaller chips.
(c) They are used in mobiles, lap-tops etc.
(d) Further they are used in chemical/ optical computer.
(e) Nano-dimensional photonic crystals and quantum electronic devices plays a vital role in the recently developed computers.

(iv) Bio-Medical and Chemical Engineering:-

(a) Consolidated state nano particles are used as catalyst, electrodes in solar and fuel cells.
(b) Bio-sensitive nano particles are used in the production of DNA-chips, bio-sensors etc.
(c) Nano-structured ceramic materials are used in synthetic bones.
(d) Few nano materials are also used in adsorbents, self-cleaning glass, fuel additives, drugs, Ferro fluids, paints etc.
(e) Nano-metallic colloids are used as film precursor.