MODULE-1

QUANTUM MECHANICS

Introduction:

Quantum mechanics is a branch of physics which deals with the behaviour of matter and energy on the scale of atoms and subatomic particles / waves. Quantum mechanics is the foundation of several related disciplines including nanotechnology, condensed matter physics, quantum chemistry, structural biology, particle physics, and electronics. The term "quantum mechanics" was first coined by Max Born in 1924.

Classical/Newtonian Mechanics:

- Describes the dynamics of moving objects at macroscopic level.
- Applicable to all objects which have considerable mass i.e., sand stone to asteroids and spherical shell to massive planets.
- Fails to explain phenomena at microscopic level.

Quantum Mechanics:

- Describes the dynamics of moving objects at microscopic level.
- Applicable to particles with negligible mass such as subatomic particles (electron, proton, neutron) and photons, phonons.
- Explains all phenomena where CM fails.

Failures of Classical/Newtonian Mechanics:

Classical mechanics fails to explain following phenomena based on its assumptions that energy emitted or absorbed by a body is continuous and light is a form of wave.

1. Blackbody spectrum
2. Atomic spectra of hydrogen atom
3. Photoelectric Effect and Compton Effect etc.

Black body radiation:

"A body which absorbs radiation of all wavelengths incident on it is called a perfect blackbody". Such a body cannot reflect or transmit a radiation and therefore it appears black. Further, when such a body is placed at constant high temperature, it emits radiation of all wavelengths.

Blackbody radiation does not depend on the type of object emitting it.
Entire spectrum of blackbody radiation depends on only one parameter, the temperature, T.
Black body: A hollow spherical shell whose inner surface coated with lampblack acts as a black body. In practice, a perfect blackbody does not exist.
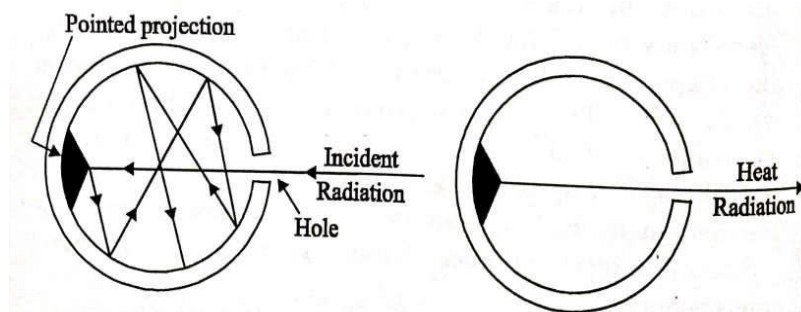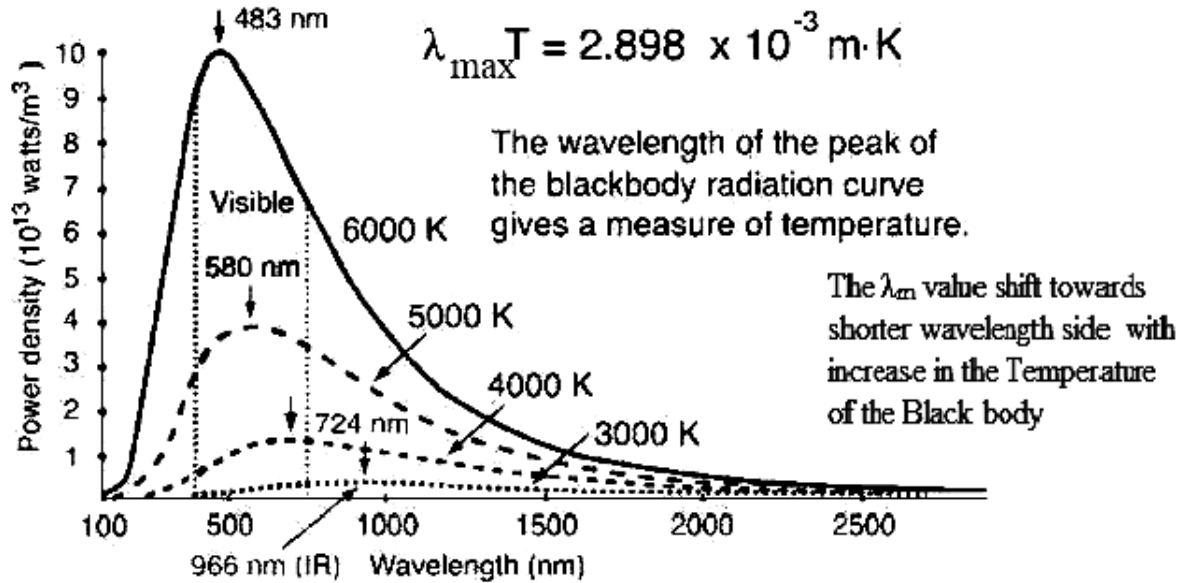
Blackbody radiation: The heat radiation emitted from a blackbody is known as blackbody radiation. The wavelength at which the maximum energy of radiation emitted depends only on temperature of the blackbody and it does not depend on the nature of the material.

Distribution of energy in Blackbody radiation:
The distribution of energy among the different wavelength of blackbody radiation was studied by Lummer and Pringsheim in 1899. The curve in the figure shows the variation of intensity of radiation with wavelength for different temperature of the blackbody.
An examination of these curves lead to following conclusions.The Salient features of black body radiation spectrum are as below,

$$\lambda_{max}T = 2.898 \times 10^{-3} \, m \cdot K$$

The wavelength of the peak of the blackbody radiation curve gives a measure of temperature.

The λ_m value shift towards shorter wavelength side with increase in the Temperature of the Black body

- At a given temperature, the energy distribution of the Black body radiation is not uniform.

- At a given temperature, the blackbody emits continuous range of wavelengths of radiation. As the wavelength increases, the intensity of radiation increases reaches a maximum value at a particular wavelength λ_m and then decreases with further increase in wavelength.

- As the temperature increases, the wavelength 'λ_m' corresponding to maximum energy density will shift towards shorter wavelengths region. The points on the dotted line represent 'λ_m' at various temperatures. It is found that $\lambda_m \propto 1/T$ which is called Wien's displacement law.

- For all wavelengths, an increase in temperature causes an increase in the energy emission or intensity.

- For a given temperature, the area under the curve represents the total energy emitted for the complete spectrum. It is found that the total radiation emitted from a black body is directly proportional to the fourth power of absolute temperature of the blackbody i.e., $E \propto T^4$.This represents Stefan-Boltzmann law.

Planck's radiation law:

Planck's quantum theory is a natural phenomenon of quantum mechanics. The theory was put forward by German physicist Max Planck in 1900. It explains the quantum nature of the energy of electromagnetic waves and the energy distribution among different wavelengths in the blackbody radiation. The Planck's theory is based on the following postulates (or) assumptions.

1. A blackbody consists of a large number of oscillating particles (atoms). These particles can vibrate in all possible energies.

2. An oscillator can have a discrete set of energies which are integral multiples of a finite quantum of energies. ie., E_n = nhv where 'n' is an integer (n = 1, 2, 3, 4…,), 'h' is a Planck's constant and 'v' is the frequency of the oscillating particle.

3. The atomic oscillator can emit or absorb energy in discrete units only by making a transition from one quantum state to another. Thus, the energy of a quantum is directly proportional to the frequency, the quantum traverse in space with the velocity of light.

According to Planck's, the energy density emitted from a blackbody at a temperature T for all wavelength range λ and λ+dλ is given by,

$$E_\lambda d\lambda = \frac{8\pi hC}{\lambda^5}\left[\frac{1}{e^{\left(\frac{hC}{\lambda kT}\right)}-1}\right]d\lambda \quad \text{------------(1)}$$

Where k is the Boltzmann constant, h is Planck's constant, C is velocity of light and T is temperature of the Black body.

Limitation: According to this equation, the distribution of energy in blackbody radiation agrees well with the experimental observations in all wavelength range.
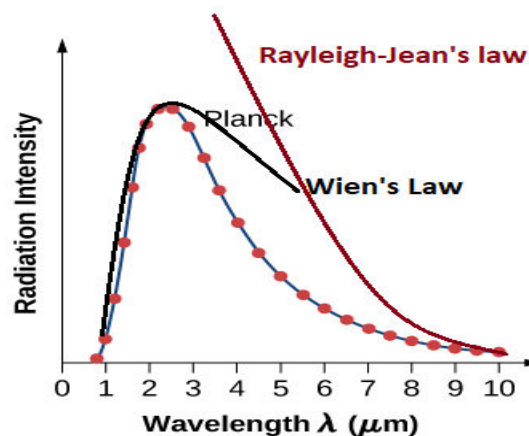
Laws of Blackbody radiation:

Stefan's law or (Stefan-Boltzmann law):
Stefan-Boltzmann law of radiation states "The emissive power (E) of a blackbody is directly proportional to the fourth power of its absolute temperature (T).

$$E \propto T^4 \implies E = \sigma T^4$$

'σ' is a constant known as Stefan is constant. Its value is **5.67x10^8 Wm^{-2} k^{-4}**. This law was suggested experimentally by Stefan and later derived by Boltzmann

Wien's displacement Law:

Wien's law states that "The peak wavelength λ_m corresponding to maximum energy for any temperature is inversely proportional to the absolute temperature 'T' of the blackbody".

$$\lambda_m \propto \frac{1}{T} \quad \text{(OR)} \quad \lambda_m T = b \text{ (Constant)}$$

Where 'b' is the Wien's constant and it is equal to **2.898x10^{-3} m K**. It shows that as temperature increases, λ_m decreases.

Wien's law of Radiation:

Wien deduced the expression for the energy density of Black body radiation for shorter wavelengths in the wavelength range λ and λ+d λ .
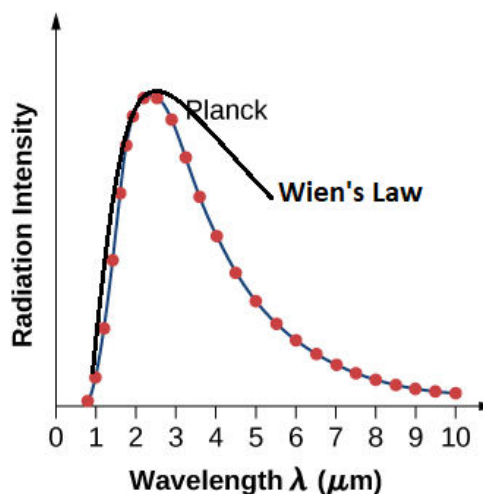
For shorter wavelengths $\dfrac{hC}{\lambda kT} >> 1$

Now equation (1) is reduced to $E_\lambda d\lambda = \dfrac{8\pi hC}{\lambda^5} \dfrac{1}{e^{\frac{hC}{\lambda kT}} - 1} d\lambda$

$$E_\lambda d\lambda = 8\pi hC \lambda^{-5} e^{\frac{-hC}{\lambda kT}} d\lambda \quad (\text{Since } E_\lambda = e^{\frac{hC}{\lambda kT}} >> 1)$$

$$E_\lambda d\lambda = C_1 \lambda^{-5} e^{\frac{-C_2}{\lambda T}} d\lambda \text{ -------- (2)} \quad \text{Where } C_1 \text{ and } C_2 \text{ are constants.}$$

Limitation: This law holds well only for shorter wavelengths and not for longer wavelengths.

Rayleigh-Jeans Law:

Rayleigh-jeans deduced the expression for the energy density of Black body radiation in the for longer wavelengths in the wavelength range λ and λ+d λ .

For longer wavelengths $\dfrac{hC}{\lambda kT} << 1$

$$e^{\frac{hC}{\lambda kT}} = 1 + \frac{hC}{\lambda kT} + \frac{1}{2}\left(\frac{hC}{\lambda kT}\right)^2 + \text{........}$$ (Neglecting higher order terms)
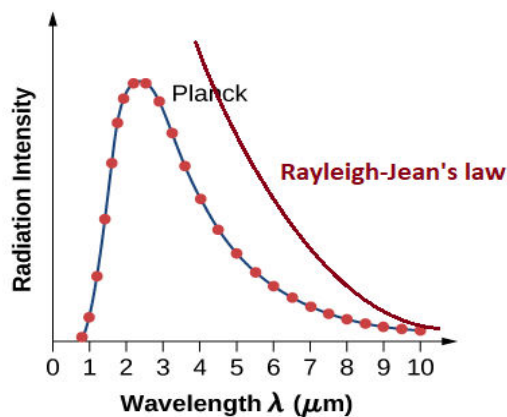
$$e^{\frac{hC}{\lambda kT}} = 1 + \frac{hC}{\lambda kT}$$

Now equation (1) reduces to

$$E_\lambda d\lambda = \frac{8\pi hC}{\lambda^5}\left[\frac{1}{1 + \dfrac{hC}{\lambda kT} - 1}\right] d\lambda$$

$$E_\lambda d\lambda = \frac{8\pi hC}{\lambda^5}\left[\frac{1}{\dfrac{hC}{\lambda kT}}\right] d\lambda$$

$$E_\lambda d\lambda = \frac{8\pi kT}{\lambda^4} d\lambda \text{ ----------(3)}$$

Limitation: This law holds well only for longer wavelengths and not for shorter wavelengths.

de-Broglie's hypothesis of Matter waves:

In 1924 Louis de Broglie suggested that every moving material fundamental particle has a wave associated with it. As electromagnetic waves behave like particles, particles like electrons will behave like waves called matter waves. He derived an expression for wavelength of matter waves.

Based on Planck's Quantum theory of radiation, the energy of a photon,

$$E = h\nu = \frac{hc}{\lambda} \quad \text{---------(1)}$$

Where 'c' is velocity of light in vacuum and λ is wavelength.
According to Einstein's mass energy relation

$$E = mc^2 \quad \text{------ (2)}$$

From (1) and (2) we have , $\quad mc^2 = \frac{hc}{\lambda}$

$$\lambda = \frac{h}{mc} = \frac{h}{p} \quad \text{where p = mc is the momentum associated with photon.}$$

In case of material particles de Broglie wavelength is given by,

$$\lambda = \frac{h}{mv}$$

If 'E' is the kinetic energy of the particle, then $E = \frac{1}{2}mv^2 = \frac{1}{2}\frac{m^2v^2}{m} = \frac{p^2}{2m}$

$$\text{Therefore } p = \sqrt{2mE}$$

$$\text{Hence } \lambda = \frac{h}{\sqrt{2mE}}$$

Consider an electron of rest mass m_0 is moving with velocity 'v' w2hich is accelerated by a potential 'V' volts.

Therefore $\frac{1}{2}m_0v^2 = eV$

$$\Rightarrow \quad v = \sqrt{\frac{2eV}{m_0}}$$

Therefore $\lambda = \frac{h}{m_0v} = \frac{h}{\sqrt{2m_0eV}}$.

By simplifying, we get $\lambda = \frac{12.26}{\sqrt{V}} A^0$.

Properties of Matter waves:
1. Lighter the particle, greater is the wavelength associated with it.
2. Smaller the velocity of the particle greater is the wavelength associated with it.
3. For v= 0, λ= infinity, it means matter waves are associated only with the moving particles.
4. The matter waves are produced whether the particles are charged or uncharged.
5. It is proved that the velocity of matter waves is greater than the velocity of light. We know that

    $E = h\nu$ and $E = mc^2$

    Therefore $h\nu = mc^2$ (or) $\nu = \frac{mc^2}{h}$

   Therefore phase velocity of matter wave is given by $\omega = \nu\lambda = \frac{c^2}{v}$

    As particle velocity 'v' cannot exceed the velocity of light, matter waves can travel with a velocity greater than velocity of light.
6. The wave nature of matter introduces an uncertainty in locating the position of the particle, which leads to the Heisenberg's uncertainty principle.

Davisson-Germer Experiment:

Waves exhibit diffraction. If the de Broglie hypothesis is valid, then the matter waves should exhibit diffraction effects.

➢ In 1927, Davisson and Germer observed the diffraction of an electron beam incident on a nickel crystal. The experiment provided a convincing proof of the wave nature of matter.

➢ The experimental arrangement of Davisson and Germer consists of an electron beam which is generated from a hot tungsten filament 'F' and an anode 'A' connected to a variable voltage source to accelerate the electrons.

➢ The electrons emerging from the filament are made to pass through an opening in the anode and fall normally on the surface of a nickel crystal which can be rotated on the axis and these electrons are scattered by the nickel crystal located at C.

➢ The detector 'D' can be moved on a graduated semi-circular scale and is used to measure the number of electrons scattered by the crystal in different directions. Thus, the intensity of the scattered electron beam was determined as a function of the scattering angle, ϕ.
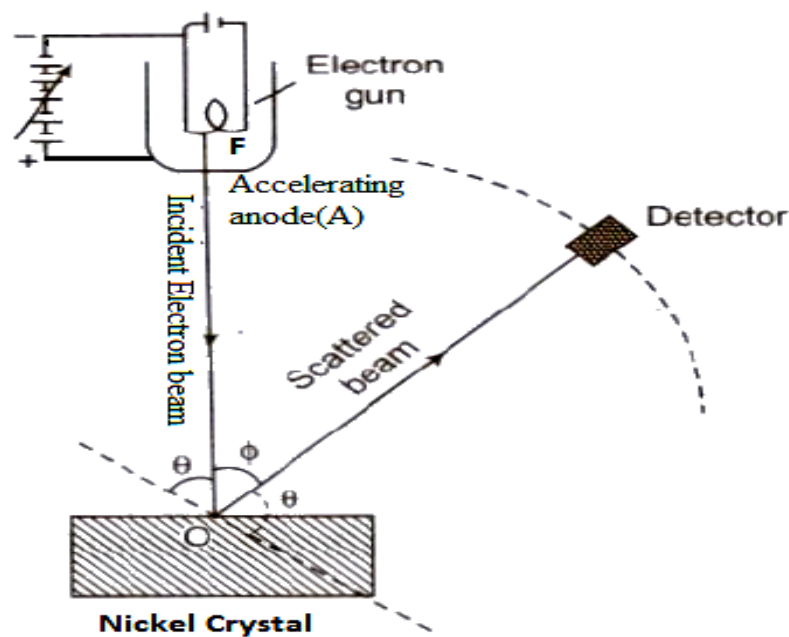
A polar graph was plotted between the detector current and the scattering angle of electron beam at different accelerating voltages. From the curve It was found that for the accelerating voltage of 54 V, the electrons are scattered more pronouncedly at an angle of 50^0 with the direction of incident beam.
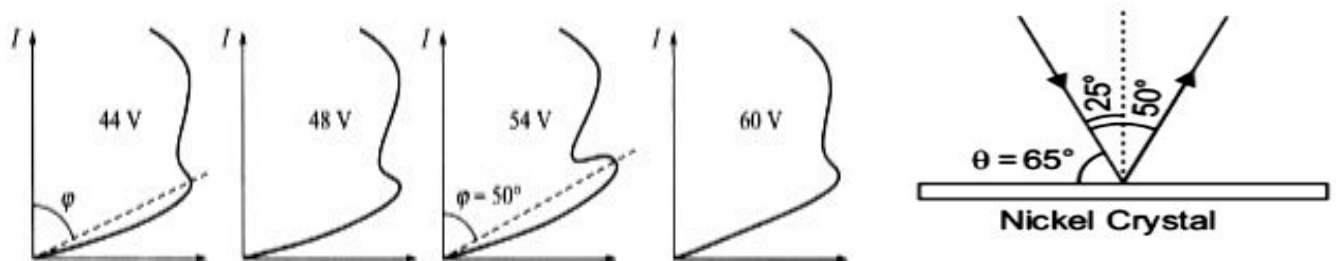
Plots of Intensity of the electron beam against scattering angle at various potentials

The inter planar spacing in nickel crystal is obtained from X-ray analysis as d=0.91 A⁰.
It is seen from the above figure that the glancing angle is $\theta = 65^0$. Applying Bragg equation, the wavelength of electron wave is computed as,

$$\lambda = 2d\sin\theta = 2*0.91A^0*\sin 65^0 = 1.65A^0$$

The wavelength of electron wave can also be computed from,

$$\lambda = \frac{h}{(2meV)^{1/2}} = \frac{6.625*10^{-34}}{(2*9.1*10^{-31}*1.6*10^{-19}*54V)^{1/2}} = 1.66A^0$$

It is seen that the values obtained experimentally using Bragg's equation and de Broglie equation are in good agreement. Hence, Davison-Germer obtained the conclusion that electrons exhibit wave nature.

G.P.Thomson's Experiment:

The de Broglie hypothesis was further confirmed in 1927 by the experiments conducted independently by G.P.Thomson in England and by Kikuchi in Japan. Thomson's experimental arrangement is shown in below.
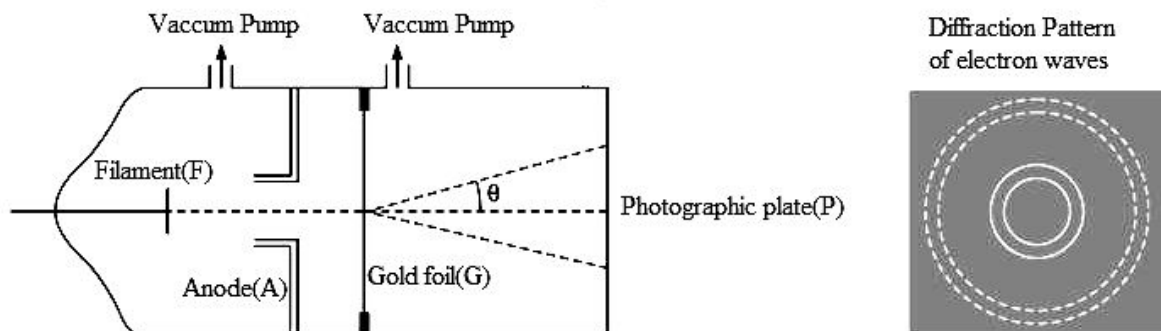
➢ Using a suitable battery the filament F can be heated, so that electrons will be emitted and pass through a high positive potential given to the anode 'A' then the electron beam passes through a fine slit and incidents on the gold foil of thickness 0.1 μm.
➢ The electrons passing through the gold foil are recorded on a photographic plate. Since the gold foil consists a large number of microscopic crystals oriented at random, the electrons striking the gold foil diffracts according to Bragg's law 2d sinθ = nλ.
➢ After developing the photographic plate a symmetrical pattern consisting of concentric rings about a central spot is obtained. The observed rings can only be considered as the diffraction pattern of the incoming beam is due to the diffraction of electrons by the gold foil.
This pattern is similar as produced by X-ray diffraction pattern using powdered samples.
➢ As the diffraction pattern can only be produced by waves and not by the particles. So Thomson concluded that electrons behave like waves. Thus, the experiments of G.P. Thomson and Kikuchi provided irrefutable proof to the existence of de Broglie waves.

Calculation of wavelength of electrons:
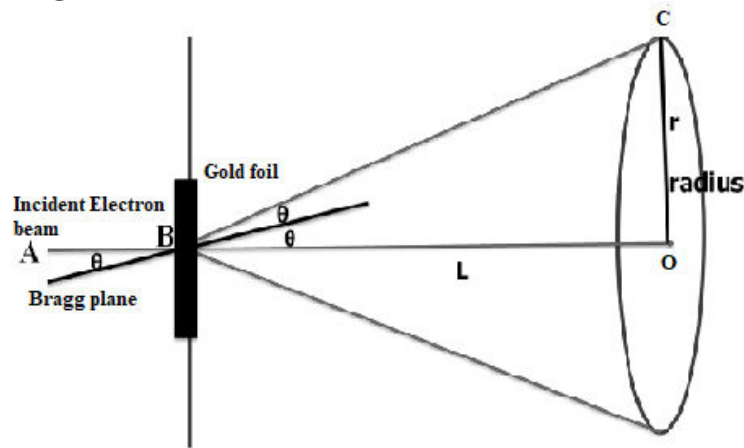
From the figure, OC is the radius 'r' of the ring with O as the center of the ring, B is the point where the electrons strikes a particular point on the gold foil. BO is the distance between the gold foil and the photographic plate and is represented by L.

$$\tan 2\theta = \frac{r}{L}$$

If θ is very small, tan (θ) ≈ θ

Hence $2\theta = r / L$ .... (1)

From Bragg's law, 2d sinθ = nλ

If θ is very small, sin (θ) ≈ θ

$2d\theta = n\lambda$

For first order diffraction n=1

$2\theta = \lambda / d$ .... (2)

From equations (1) and (2), we have r / L = λ / d

$d = \lambda L / r$ .... (3)

From de Broglie wavelength for electrons, we have $\lambda = \dfrac{h}{\sqrt{2m_0 eV}}$ .... (4)

Here, m_0 is the rest mass of electron.

from equations (4) & (3) we get,

$$d = \frac{L}{r}\frac{h}{\sqrt{2meV}} = 4.08 \text{ A}^0$$

From X-ray diffraction method for first order diffraction, 2d sinθ = nλ

$$d = \frac{\lambda}{2\sin\theta} = 4.06 \text{ A}^0$$

Thus the value of 'd' obtained from G.P. Thomson experiment and X-ray diffraction method are the same, demonstrating the de Broglie's concept of matter waves. Thomson also concluded that the wavelength of electron only depends on the accelerating voltage and is independent of the material of the target.

Heisenberg's Uncertainty Principle:

In 1927 Heisenberg proposed the theory of uncertainty. This principle is a result of dual nature of matter. In quantum mechanics a particle is described by a wave packet, which represents and symbolizes all about particle and moves with group velocity. According to Max Born's probability interpretation the particle may be found anywhere within the wave packet. This implies that the position of the particle is uncertain within the limits of wave packet. Moreover the wave packet has velocity spread and hence there is uncertainty about the velocity or momentum of the particle.

According to Heisenberg's Uncertainty Principle, **"It is impossible to determine simultaneously both the position and momentum (or velocity) of a particle with accuracy".**
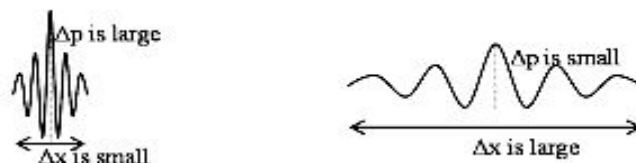
If 'Δx' denotes the error in the measurement of position of the particle and 'Δp' denotes the error in the measurement of momentum, then

$$\Delta x. \Delta p \geq \frac{\hbar}{2}$$

In terms of energy and time the relation is given by, $\Delta E. \Delta t \geq \frac{\hbar}{2}$

Here Position- Momentum, Energy- Time are called canonical conjugate pairs.

Electrons cannot exist inside the nucleus (Application of Uncertainty Principle):
The radiation emitted by radioactive nucleus consists of α,β,γ out of which β-rays are identified to be electrons. We apply uncertainty principle whether electrons are coming out of the nucleus. The radius of the nucleus is of the order of 10^{-15} m. Therefore, if electrons were to be in the nucleus, the maximum uncertainty 'Δx' in the position of the electron is equal to the diameter of the nucleus.
Thus, $\Delta x = 2*10^{-15}$ m

According Heisenberg uncertainty principle, $\Delta x. \Delta p \geq \frac{\hbar}{2}$

Therefore the minimum uncertainty in its momentum is then given by,

$$\Delta p = \frac{h}{4\pi \times 2 \times 10^{-15}} = 2.6 \times 10^{-20} kg - m/s$$

The minimum uncertainty in momentum can be taken as the momentum of the electron. Thus, $\Delta p = 2.6 \times 10^{-20}$ kg-m/s.

The minimum energy of the electron in the nucleus is then given by,

$$E_{min} = p \times C = (2.6 \times 10^{-20} \text{ kg-m/s})(3 \times 10^{8} \text{ m/s}) = 7.8 \times 10^{-12} \text{ J} = 48 \times 10^{6} \text{ e V} = 48 \text{ MeV}$$

It implies that if an electron exists within the nucleus, it must have a minimum energy of about 48 MeV. But the experimental measurements showed that the maximum kinetic energies of β-particles were of the order of 4 MeV only. Hence, electrons are not present in the nucleus. It is subsequently established that emission of β-particles occurs due to the transformation of a neutron into a proton in the nucleus.

Schrodinger Time independent wave equation:

In 1925, Schrodinger gave a mathematical theory known as wave mechanics to describe the dual nature of matter. He derived a wave equation for a moving particle by incorporating de-Broglie wavelength expression in the classical wave equation. If a particle of mass 'm' moving with velocity 'v' is associated with a group of waves.

Let ψ be the wave function of the particle. Also let us consider a simple form ofprogressing wave like the one represented by the following equation,

$$\psi = \psi_0 \sin(\omega t - kx) \text{---------} (1)$$

Where $\psi = \psi(x, t)$ and ψ_0 is the amplitude.

Differentiating Eq.(1) partially w.r.to x,

$$\frac{\partial \psi}{\partial x} = \psi_0(-k)\cos(\omega t - kx) \text{----------}(2)$$

Differentiating Eq.(2) w.r.to x,

$$\frac{\partial^2 \psi}{\partial x^2} = (-k)\psi_0\left[-\sin(\omega t - kx)\right](-k)$$

$$\frac{\partial^2 \psi}{\partial x^2} = -k^2\psi_0 \sin(\omega t - kx)$$

$$\frac{\partial^2 \psi}{\partial x^2} = -k^2\psi \quad \left[\text{From Eq.(1)}\right]$$

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{4\pi^2}{\lambda^2}\psi \quad \text{--------}(3) \quad \left(\text{Since } k = \frac{2\pi}{\lambda}\right)$$

Incorporating de-Broglie wavelength expression in Eq.(3) we get,

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{4\pi^2}{\left(\dfrac{h}{mv}\right)^2}\psi$$

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{4\pi^2}{\left(\dfrac{h^2}{m^2v^2}\right)}\psi$$

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{4\pi^2 m^2 v^2}{h^2}\psi \quad \text{-------}(4)$$

The total energy E of the particle is the sum of its kinetic energy K and potential energy V

i.e., $E = K + V$ -------------- (5)

and $K = \dfrac{1}{2}mv^2$ -------------- (6)

Therefore, $m^2v^2 = 2m(E - V)$ --------- (7)

From Eq.(4) and Eq.(7) we get,

$$\frac{\partial^2 \psi}{\partial x^2} = -\frac{4\pi^2 2m(E-V)}{h^2}\psi$$

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{4\pi^2 2m(E-V)}{h^2}\psi = 0$$

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{2m}{\hbar^2}(E-V)\psi = 0 \text{ ----------(8)} \qquad \left(\text{Since } \hbar = \frac{h}{2\pi}\right)$$

This is known as one-dimensional Schrodinger's time independent wave equation which explains the motion of a charged particle in a potential V. Extending Eq.(8) for a three dimensions, we have,

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{2m}{\hbar^2}(E-V)\psi = 0$$

$$\nabla^2 \psi + \frac{2m}{\hbar^2}(E-V)\psi = 0 \text{ -------- (9)} \quad (\text{Here } \nabla^2 \text{ is called Laplacian operator})$$

This is the Schrödinger Time Independent Wave Equation in three dimensions.

Physical Significance of Wave function (Born's Interpretation):

➤ The wave function is used to identify the state of a particle in a quantum mechanical system.
➤ The wave function 'ψ' measures the variations of the matter wave. Thus it connects the particle and its associated wave statistically.
➤ The probability that a particle will be found at a given place in space at a given instant of time is characterized by the function $\psi(x, y, z, t)$ it is called wave function
➤ The wave function can be either real or complex. According to Max-born, the probability density at any point gives the probability of finding the particle at a given point in space given by Probability density $= |\psi^2| = \psi\psi^*$, Where ψ^* is the complex conjugate of ψ
➤ The probability of finding a particle in a volume dx.dy.dz is ,

$$\therefore \iiint |\psi^2| dxdydz = 1$$

The above equation is called normalized wave equation.

Properties of Wave function 'ψ':

1) It provides all measurable information about the particle in a quantum mechanical system.
2) The wave function ψ must be finite, Single valued and continuous.
3) The value of wave function should vanish at boundaries.

Particle in One Dimensional Potential well :

Consider the case of a particle of mass 'm' which is bound to move in a one dimensional potential well of length 'L'. The particle is free to move inside the well and hence the potential energy of the particle is assumed as zero.
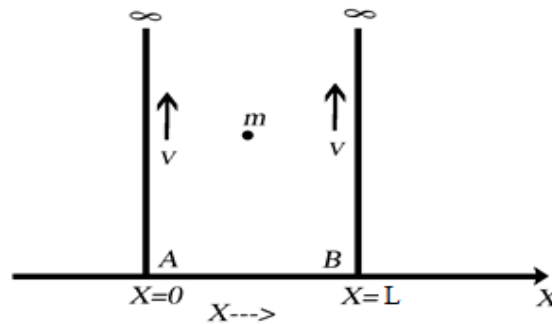
Thus we have

$V(x) = \mathbf{0}$ for $0 < x < L$

And $V(x) = \infty$ for $x \leq 0$ and $\mathbf{x \geq L}$

The One dimensional Schrodinger time independent wave equation for the particle is,

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0$$

Since for a free particle V=0,

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0$$

$$\frac{d^2\psi}{dx^2} + K^2\psi = 0 \text{ ---------- (1)}, \text{ where } K^2 = \frac{2mE}{\hbar^2}$$

The solution of the above equation can be written as,

$$\psi(x) = A\sin Kx + B\cos Kx \text{ ----------(2)}$$

Where A and B are constants.

The values of these constants can be obtained by applying boundary conditions given by

(i) $\psi = 0$ at x=0 and (ii) $\psi = 0$ at x=L.

Applying first boundary condition for (2) we have

0=Asin0+Bcos0   ie., B=0

$$\therefore \psi(x) = A\sin Kx \text{ ----------(3)}$$

Applying second boundary condition for (3) we have

$$0 = A \sin KL$$

Here either A=0 or Sin KL=0 but $A \neq 0$

Because if A=0, entire function is zero as B=0

$$\therefore \sin KL = 0 \text{ or } KL = n\pi(n = 0,1,2,3...........)$$

$$K = \frac{n\pi}{L}$$

Now equation (3) becomes

$$\psi(x) = A\sin\frac{n\pi x}{L} \quad\text{-----------(4)}$$

Since $K = \dfrac{n\pi}{L} \Rightarrow K^2 = \dfrac{n^2\pi^2}{L^2}$ and we have

$$\therefore \frac{2mE}{\hbar^2} = \frac{n^2\pi^2}{L^2} \quad \text{(Using E}_n \text{ for E in general)}$$

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2} = \frac{n^2h^2}{8mL^2}$$

$$\therefore E_n \propto n^2$$

This shows that energy of particle is quantized.

The discrete energy values are given by

$$E_1 = \frac{\pi^2\hbar^2}{2mL^2} \quad \text{for n=1}$$

$$E_2 = \frac{4\pi^2\hbar^2}{2mL^2} = 4E_1 \quad \text{for n=2}$$

$$E_3 = \frac{9\pi^2\hbar^2}{2ma^2} = 9E_1 \quad \text{for n=3}$$

$$E_4 = \frac{16\pi^2\hbar^2}{2mL^2} = 16E_1 \quad \text{for n=4 and so on.}$$

The constant A can be obtained by applying normalization condition, i.e,

$$\int_0^L |\psi(x)^2| dx = 1$$

$$\int_0^L A^2 \sin^2\frac{n\pi x}{L} dx = 1$$

$$A^2 \int_0^L \sin^2\frac{n\pi x}{L} dx = 1$$

$$A^2 \int_0^L \frac{1}{2}\left[1 - \cos\frac{2n\pi x}{L}\right] dx = 1$$

$$\frac{A^2}{2}\left[x - \frac{L}{2\pi n}\sin\frac{2n\pi x}{L}\right]_0^L = 1$$

$$\frac{A^2 L}{2} = 1$$

$$A^2 = \frac{2}{L} \quad \text{(or)} \quad A = \sqrt{\frac{2}{L}}$$

$$\therefore \psi_n(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi}{L}\right)x$$

wavefunctions   Probability densities   Energies
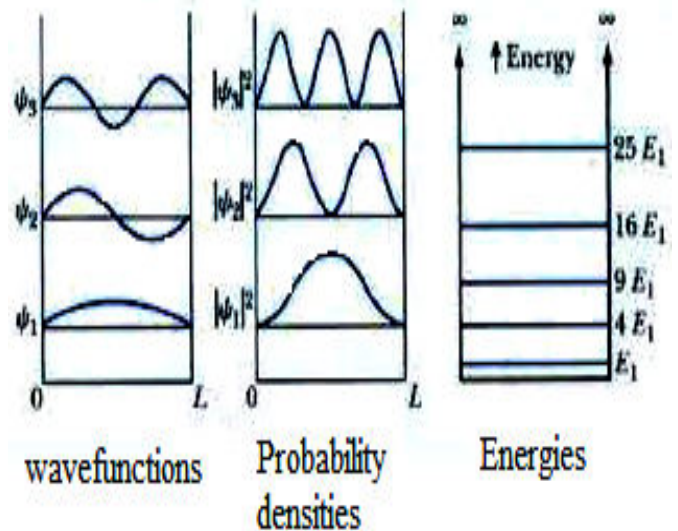
Module-II::BAND THEORY OF SOLIDS

Electron Theory of metals:

The electron theory of metals explain the following concepts

- Structural, electrical and thermal properties of materials.
- Elasticity, cohesive force and binding in solids.
- Behaviour of conductors, semi conductors, insulators etc.

So far three electron theories have been proposed.

1. **Classical Free electron theory**:
   - ✓ It is a macroscopic theory.
   - ✓ Proposed by Drude and Loretz in 1900.
   - ✓ It explains the free electrons in lattice
   - ✓ It obeys the laws of classical mechanics.

2. **Quantum Free electron theory**:
   - ✓ It is a microscopic theory.
   - ✓ Proposed by Sommerfield in 1928.
   - ✓ It explains that the electrons move in a constant potential.
   - ✓ It obeys the Quantum laws.

3. **Brillouin Zone theory or Band theory**:
   - ✓ Proposed by Bloch in 1928.
   - ✓ It explains that the electrons move in a periodic potential.
   - ✓ It also explains the mechanism of semiconductivity, based on bands and hence called band theory.

CLASSICAL FREE ELECTRON THEORY (Drude-Lorentz Theory) [qualitatively]:

This theory was developed by Drude and Lorentz and hence is also known as Drude-Lorentz theory. According to this theory, a metal consists of electrons which are free to move about in the crystal like molecules of a gas in a container. Mutual repulsion between electrons is ignored and hence potential energy is taken as zero. Therefore the total energy of the electron is equal to its kinetic energy.

a) Postulates of Classical free electron theory:

1. All the solids are composed of atoms. Each atom have central nucleus around which there are revolving electrons.
2. The electrons are free to move in all possible directions about the whole volume of metals.
3. In the absence of an electric field the electrons move in random directions making collisions from time to time with positive ions which are fixed in the lattice or other free electrons. All the collisions are elastic i.e.; no loss of energy.

4. When an external field is applied the free electrons are slowly drifting towards the positive potential.
5. Since the electrons are assumed to be a perfect gas they obey classical kinetic theory of gasses.
6. Classical free electrons in the metal obey Maxwell-Boltzmann statistics.

b) Success of classical free electron theory:
1. It verifies ohm's law

2. It explains the electrical and the thermal conductivities of metals.

3. It derives Wiedemann-Franz law. (i.e. the relation between electrical conductivity and thermal conductivity)

4. It explains the optical properties of metals.

c) Drawbacks of Classical free electron theory

1. According to this theory, ρ is proportional to $\sqrt{T}$. But experimentally it was found that ρ is proportional to T.

2. According to this theory, K/ σT = L, a constant (Wiedmann-Franz law) for all temperatures. But this is not true at low temperatures.

3. The theoretically predicted value (4.5R) of specific heat of a metal does not agree with the experimentally obtained value (3R). [R-Universal gas constant]

4. This theory fails to explain ferromagnetism, superconductivity, photoelectric effect, Compton Effect and blackbody radiation.

QUANTUM FREE ELECTRON THEORY [qualitatively]

Classical free electron theory could not explain many physical properties. In 1928, Sommerfeld developed a new theory applying quantum mechanical concepts and Fermi-Dirac statistics to the free electrons in the metal. This theory is called quantum free electron theory. Classical free electron theory permits all electrons to gain energy. But quantum free electron theory permits only a fraction of electrons to gain energy.

a) Postulates:
1. Sommerfeld proposed this theory in 1928 retaining the concept of free electrons moving in a uniform potential within the metal as in the classical theory, but treated the electrons as obeying the laws of quantum mechanics.

2. Based on the deBroglie wave concept, he assumed that a moving electron behaves as if it were a system of waves. (called matter waves-waves associated with a moving particle).

3. According to quantum mechanics, the energy of an electron in a metal is quantized. The electrons are filled in a given energy level according to Pauli's exclusion principle. (i.e. No two electrons will have the same set of four quantum numbers.)

4. Each Energy level can provide only two states namely, one with spin up and other with spin down and hence only two electrons can be occupied in a given energy level.

5. It is assumed that the valance electrons travel in constant potential inside the metal but they are prevented from escaping the crystal by very high potential barriers at the ends of the crystal.

6. In this theory, though the energy levels of the electrons are discrete, the spacing between consecutive energy levels is very less and thus the distribution of energy levels seems to be continuous.

b) Success of Quantum Free Electron Theory:

1. According to classical theory, which follows Maxwell-Boltzmann statistics, all the free electrons gain energy. So it leads to much larger predicted quantities than that is actually observed.

2. But according to quantum mechanics only one percent of the free electrons can absorb energy. So the resulting specific heat and paramagnetic susceptibility values are in much better agreement with experimental values.

3. According to quantum free electron theory, both experimental and theoretical values of Lorentz number $(L = 2.44 \times 10^{-8} W\Omega K^{-1})$ are in good agreement with each other.

c) Drawbacks of Quantum free electron Theory:

1. According to quantum mechanics, a conduction electron in a metal experiences constant(zero) potential and free to move inside the crystal but will not come out because an infinite potential exists at surface.

2. This theory explains electrical conductivity and thermal conductivity but failed to explain the difference between conductors, insulators, and semiconductors.

3. Its fails positive hall coefficient of metals and

4. low conductivities of divalent metals than monovalent metals

| S.No.	Classical Free electron Theory	Quantum Free electron Theory
1	Obeys classical laws	Obeys Quantum laws
2	Energy is continuous	Energy is discrete
3	Pauli's exclusion principle is not considered	Electrons are filled according to Pauli's exclusion principle
4	Electrons are particles	Dual nature of electron (wave and particle)

Brillouin Zone Theory [Band Theory]:

Bloch developed this theory in which the electrons move in a periodic potential due to positive ions in metals. A electron move through these ions, it experience varying potentials. It explains the mechanisms of solids on the basis of energy bands and hence band theory.

BLOCH THEOREM:

A crystalline solid consists of a lattice which is composed of large number of positive ion cores at regular intervals and the conduction electrons that can be move freely throughout the lattice. Let us consider the lattice in one dimensional view, i.e., only an array of ion cores along x-axis. Let a conduction electron has a potential energy V(x) due to its position in the lattice, the variation of potential is shown in fig.
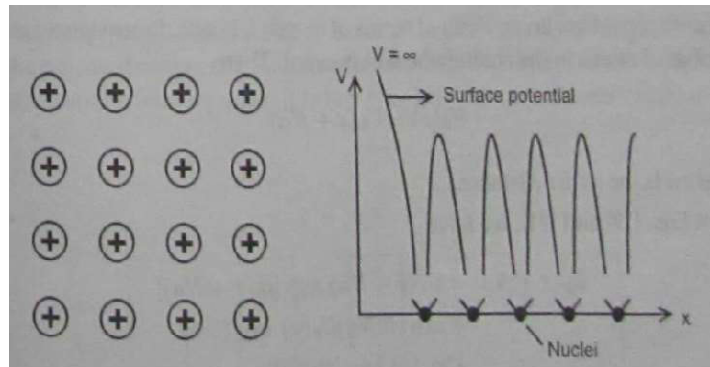
The one dimensional Schrodinger equation corresponding to this can be writer as

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2m}{\hbar^2}(E - V(x))\psi = 0$$

The periodic potential *V(x)* may be defined by means of the lattice constant *a* as

V(x) = V(x+a)

The Bloch theorem states that for a particle moving in a periodic potential, the eigen functions for a conduction electron of the form.

$$\psi_k(x) = U_k(x)\exp(\pm ikx)$$

Where $U_k(x) = U_k(x+a)$

The function $U_k(x)$ has the same periodicity as the potential energy of the electron and is

called the **modulating function.** Let us now consider a linear chain of atoms of length L in one dimensional case with N number of atoms in the chain (N is even), then,

$$\psi_k(x + Na) = U_k(x + Na)\exp(ik[x + Na])$$
$$= U_k(x)exp(ika)\exp(ikNa)$$
$$= \psi_k(x) \quad \text{(Since, } U_k(x+Na)=U_k(x) \text{ and } exp(ikNa) =1)$$
$$\psi_k(x + Na) = \psi_k(x)$$

KRONIG-PENNY MODEL:

It is assumed in quantum free electron theory of metals that the free electrons in a metal express a constant potential and is free to move in the metal. This theory explains successfully using the band theory of solids.

According to this theory, the electrons move in a periodic potential produced by the positive ion cores. The potential of electron varies periodically with periodicity of ion core and potential energy of the electrons is zero near nucleus of the positive ion core. It is maximum when it is lying between the adjacent nuclei which are separated by inter-atomic spacing. The variation of potential of electrons while it is moving through ion core is shown fig.
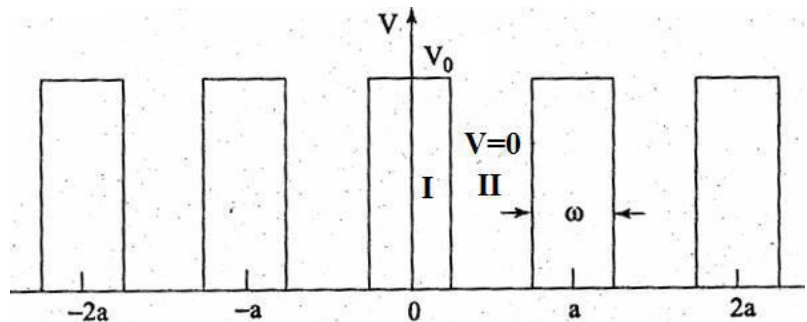
Fig. One-Dimensional Periodic potential

Applying the time independent Schrodinger's wave equation for the above two regions,

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2m}{\hbar^2}(E - V_o)\psi = 0 \text{ in region } \text{--------- I}$$

$$\frac{\partial^2 \Psi}{\partial x^2} + \frac{2m}{\hbar^2}(E)\psi = 0 \text{ In region } \text{--------- II (since V=0)}$$

Substitute the $\frac{2m(E-V_o)}{\hbar^2} = \beta^2$ *and* $\frac{2mE}{\hbar^2} = \alpha^2$

The solution for the eqns. (I) and (II) can be written from the Bloch theorem as,

$$\psi_k(x) = U_k(x) \exp(\pm ikx)$$

Where, $U_k(x) = U_k(x+a)$

Applying the boundary conditions and solve the above equations and we get,

$$p\frac{\sin(\alpha a)}{\alpha a} + \cos(\alpha a) = \cos(ka) \text{ ----------- (1)}$$

Where $p = \dfrac{ma}{\hbar^2} V_o \omega$ is the **scattering power**,

$V_o\omega$ is the **Barrier strength**.

The left hand side of the eq. (1) is plotted as a function of α for the value of P=3π/2, which is shown in fig, the right hand side one takes values between -1 to +1 as indicated by the horizontal lines in fig. Therefore the equation (1) is satisfied only for those values of ka for which left hand side between ± 1.

From fig., the following conclusions are drawn.

Fig Plot of $P\dfrac{\sin \alpha a}{\alpha a} + \cos \alpha a$ versus αa.

1. The energy spectrum of the electron consists of a number of allowed and forbidden energy bands.

2. The width of the allowed energy band increases with increase of energy values i.e., increasing the values of αa. This is because the first term of equation (1)
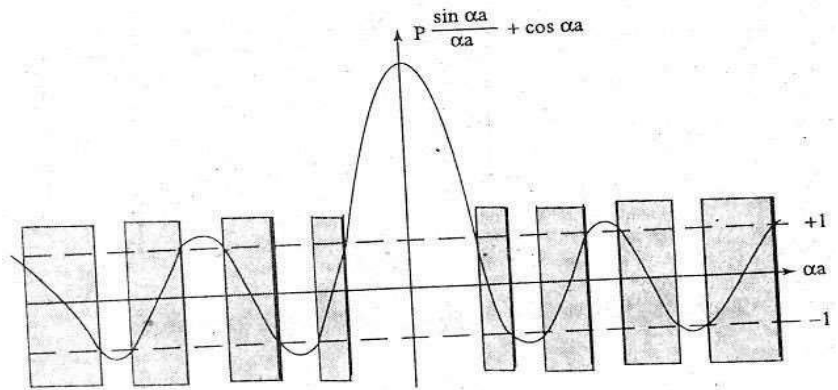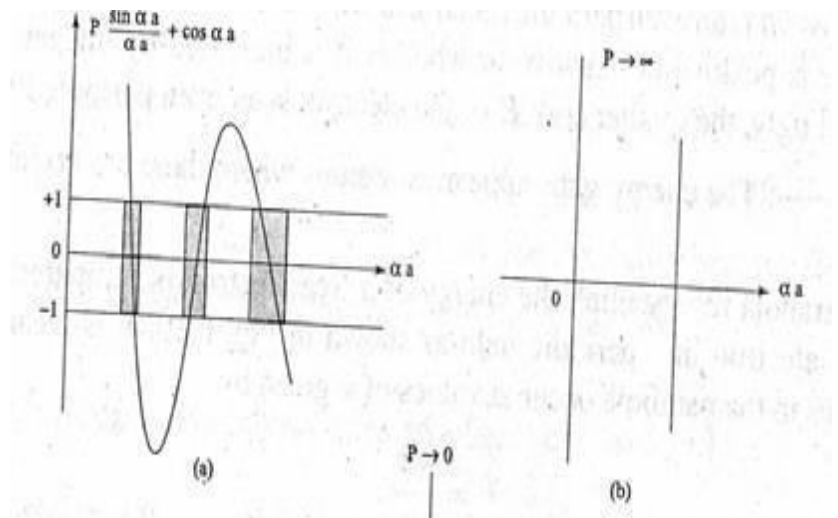
decreases with increase of αa.

3. With increasing P, i.e. with increasing potential barrier, the width of an allowed band decreases. As P→∞, the the allowed energy becomes infinitely narrow and the energy spectrum is a line spectrum as shown in fig (b).

(c)

$p \rightarrow 0$
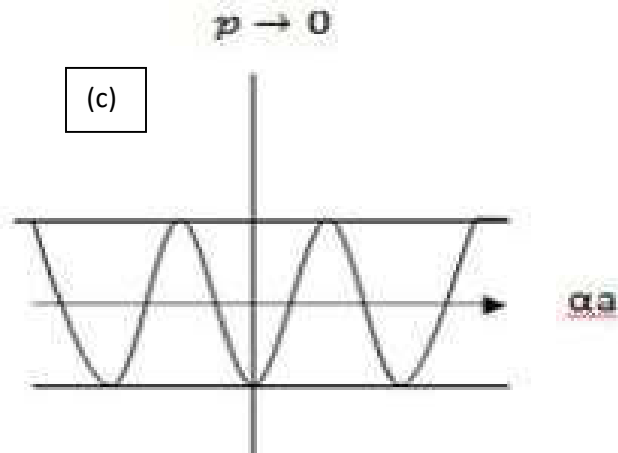
If P→∞, then the equation (1) has solution i.e., This expression shows that the energy spectrum of the electron contains discrete energy levels separated by forbidden regions.

4. When P→0 then,

   cos αa =cos ka,

Therefore, $\alpha^2 = k^2$

But, $\alpha^2 = \dfrac{2mE}{\hbar^2} = k^2 \Rightarrow E = \dfrac{\hbar^2 k^2}{2m} = \dfrac{p^2}{2m} = \dfrac{1}{2}mv^2 \ (since, p = \hbar k)$

The equation (1) shows all the electrons are completely free to move in the crystal without any constraints. Hence, no energy level exists i.e. all the energies are allowed to the electrons and shown in fig (c). This case supports the classical free electrons theory.

BRILLOUIN ZONES OR E-K DIAGRAMS OR BAND DIAGRAMS:

The Brillouin zone is a representation of permissive values of k (wave vector) of electrons in I, II, III dimensions. Thus the energy spectrum of an electron moving in the presence of a periodic potential is divided into allowed and forbidden zones.

When a parabola
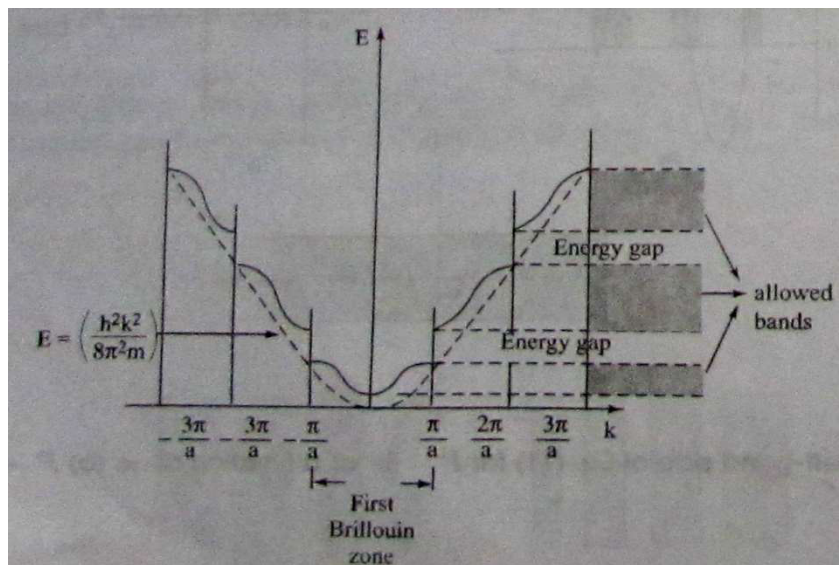
(dashed curve in fig.) representing the energy of a free electron is compared with the energy of an electron in a periodic potential as shown in the above fig. It is clear that the discontinuitites in the parabola occurs at values of **k** given by

$$k = \frac{n\pi}{a}$$ where $n = \pm1, \pm2, \pm3, ...$

From the graph:

- The dashed curve shows the free electron parabola based on Quantum laws.
- The electron has allowed energy values in the region or zone extending from $k = -\pi/a$ to $+ \pi/a$. This zone is called 1st (***First) Brillouin zone.***
- After a discontinuity in energy, called forbidden gap, we get another allowed zone of energy values in the region extended from $k = -\pi/a$ to $-2\pi/a$ and π/a to $2\pi/a$. This zone is called *2nd (second) Brillouin zone*.
- Similarly other higher order Brillouin zones can be defined.

FERMI LEVEL:

- The highest completely filled energy level at absolute zero temperature is known as the Fermi Level.
- The Fermi level lies between the valance band and conduction band because at absolute zero temperature the electrons are all in the lowest energy state.
- Due to the lack of sufficient energy at 0 K, the Fermi level can be considered as the sea of fermions (or electrons) above which no electrons exist.
- The reason for the existence of this energy level is due to Pauli's exclusion principle which states two fermions cannot occupy that same quantum state. So, if a system has more than one fermions, each fermion has a different set of magnetic quantum numbers associated with it.
- This is called Fermi-Dirac distribution function:
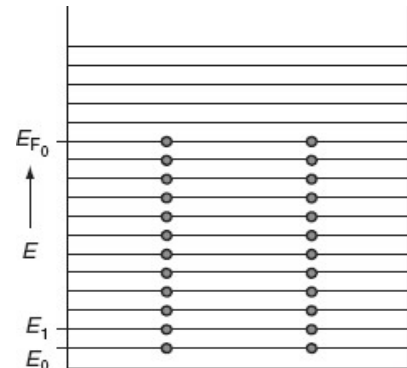
$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}$$

E_F = Fermi energy or Fermi level

k = Boltzmann constant = 1.38×10^{-23} J/K
   = 8.6×10^{-5} eV/K

T = absolute temperature in K

DENSITY OF (ENERGY) STATES:

The number of energy states per unit volume of the crystal which are contained in the energy interval between E and E+dE defines a density of states *Z(E)*.

$$Z(E)dE = \frac{Number\ of\ energy\ states\ between\ E\ and\ E + dE}{Volume\ of\ the\ material(V)} = \frac{D(E)dE}{V} ---(1)$$

Let us consider a sphere of radius "n" in space with Quantum numbers n_x, n_y, and n_z.

$$n^2 = n_x^2 + n_y^2 + n_z^2$$

Number of energy states within a sphere of radius   n

$$= \frac{4}{3}\pi n^3$$

Consider the sphere of radius 'n' due to one octant, n

$$= \frac{1}{8}\left(\frac{4}{3}\pi n^3\right)$$

LIlly sphere of radius 'n+dn'   n+dn =

$$\frac{1}{8}\left(\frac{4}{3}\pi(n + dn)^3\right)$$

Therefore, no. of energy states available in n and n+dn is

$$D(E)dE = \frac{1}{8}\left[\frac{4}{3}\pi(n + dn)^3 - \frac{4}{3}\pi n^3\right]$$

$$D(E)dE = \frac{1}{8}\left[\frac{4}{3}\pi(n^3 + dn^3 + 3n^2 dn + 3ndn^2 - n^3)\right]$$

dn^2 and dn^3 are very very small, it can be neglected.

$$\therefore D(E)dE = \frac{1}{8}\left[\frac{4}{3}\pi(3n^2 dn)\right]$$
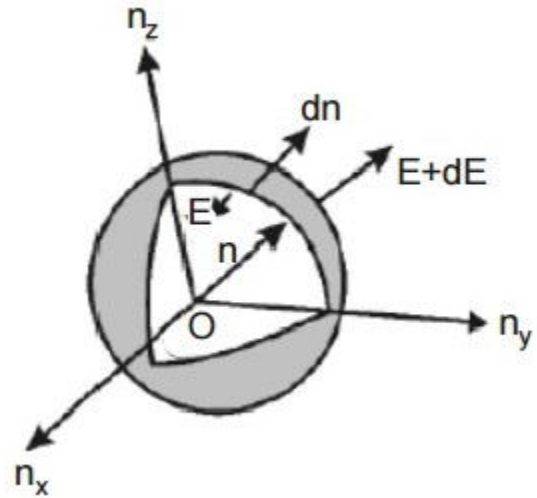
$$D(E)dE = \frac{\pi}{2}n^2 dn -----(2)$$

From Quantum mechanics, the energy of an electron is given by

$$E = \frac{n^2 h^2}{8mL^2}$$

$$i.e.\ n^2 = \frac{8mL^2 E}{h^2} \text{------------------} (3)$$

$$n = \left[\frac{8mL^2 E}{h^2}\right]^{1/2} \text{-------------} (4)$$

Differentiating n^2 in eq. (3), we get,
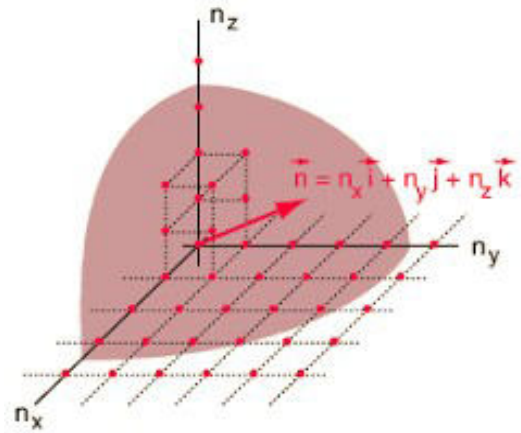
$$2ndn = \frac{8mL^2}{h^2} dE$$

$$ndn = \frac{8mL^2}{2h^2} dE \quad \text{---------------- (5)}$$

Substituting the values of n and ndn from eqs (4) and (5) in eq. (2),

$$D(E)dE = \frac{\pi}{2} \times \left[\frac{8mL^2E}{h^2}\right]^{1/2} \times \frac{8mL^2}{2h^2} dE$$

$$\boldsymbol{D(E)dE = \frac{\pi}{4h^3}(8m)^{3/2}L^3 E^{1/2} \, dE}$$

According to Pauli's exclusion principle, two electrons of opposite spin can accommodate in each energy state and hence the number of energy states available is

$$D(E)dE = 2 \times \frac{\pi}{4h^3}(8m)^{3/2}L^3 E^{1/2} \, dE$$

Therefore,

The density of states (energy) *Z(E)dE* from eq.(1) is,

$$Z(E)dE = \frac{2 \times \frac{\pi}{4h^3}(8m)^{3/2}L^3 E^{1/2} \, dE}{V = L^3}$$

$$\boldsymbol{Z(E)dE = \frac{4\pi}{h^3}(2m)^{3/2} E^{1/2} \, dE}$$

EFFECTIVE MASS OF AN ELECTRON:
      When an electron in a periodic potential is accelerated by an electric field or magnetic field then the mass of the electron varies with velocity. ***This means the mass of the electron is function of velocity and is termed as effective mass of the electron (m*).***

According to quantum theory, an electron moving with a velocity 'v' can be treated as a wave pocket moving with group velocity 'v_g'.

$$v_g = \frac{dw}{dk} \quad \text{------------- (1)}$$

where ω is the angular frequency of the de Broglie's wave and k = 2π/λ is the wave vector and for an electron can be expressed as

$$E = hv = \frac{h}{2\pi} \times 2\pi v$$

$$E = \hbar\omega \qquad \left(since \ \frac{h}{2\pi} = \hbar, 2\pi v = \omega\right) \text{--------- (2)}$$

Differentiating the eqn. (2) w.r. to k,

$$\frac{dE}{dk} = \hbar\frac{d\omega}{dk} \implies \frac{d\omega}{dk} = \frac{1}{\hbar}\frac{dE}{dk} \text{----------- (3)}$$

From eqn. (1) and (3), we get,

$$V_g = \frac{1}{\hbar}\frac{dE}{dk} \text{------------- (4)}$$

Acceleration of the electron be taken as the rate of change of velocity (group)

i.e., $\quad a = \dfrac{dv_g}{dt}$ ----------- (5)

$$a = \frac{d}{dt}\left[\frac{1}{\hbar}\frac{dE}{dk}\right] = \frac{d}{dt}\left[\frac{1}{\hbar}\frac{dE}{dk} \times \frac{dk}{dk}\right]$$

$$a = \frac{1}{\hbar}\frac{d^2E}{dk^2} \times \frac{dk}{dt} \text{------------- (6)}$$

But, the momentum of electron is, $p = \hbar k$

$$\frac{dp}{dt} = \hbar\frac{dk}{dt} \implies F = \hbar\frac{dk}{dt} \quad \left(since, \frac{dp}{dt} = F\right)$$

$$\frac{dk}{dt} = \frac{F}{\hbar} \text{----------- (7)}$$

Substitute eqn. (7) in eqn. (6), we get,

$$a = \left[\frac{1}{\hbar}\frac{d^2E}{dk^2} \times \frac{F}{\hbar}\right]$$

$$F = \frac{\hbar^2}{\left[\frac{d^2E}{dk^2}\right]}. a \text{----------- (8)}$$

But, Force (F) = effective mass of the electron $(m^*) \times$ acceleration (a)

Hence, $\qquad \boldsymbol{m^* = \dfrac{\hbar^2}{\left[\dfrac{d^2E}{dk^2}\right]}}$

(a)

(b)

(c)

According to band theory of solids, the variation of E with k is shown in above fig.
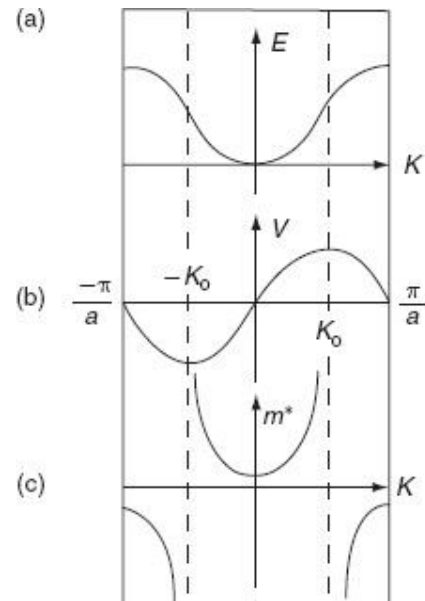
Using, E-k diagram and velocity equation, the velocity of electron can be calculated. From the above fig., it is clear that, the mass of the electron (positive) increases with k and maximum value is observed at $k=k_O$. Further, the m* is negative due to declaration of the electron velocity.

ORIGIN OF ENERGY BANDS IN SOLIDS:

Solids are usually moderately strong, slightly elastic structures. The individual atoms are held together in solids by inter-atomic forces or bonds. In addition to these attractive forces, repulsive forces also act and hence solids are not easily compressed. The attractive forces between the atoms are basically electrostatic in origin. The bonding is strongly dependent on the electronic structure of the atoms. The attraction between the atoms brings them closer until the individual electron clouds begin to overlap. A strong repulsive force arises to comply with Pauli's exclusion principle. When the attractive force and the repulsive force between any two atoms occupy a stable position with a minimum potential energy. The spacing between the atoms under this condition is called equilibrium spacing.

In an isolated atom, the electrons are tightly bound and have discrete, sharp energy levels. When two identical atoms are brought closer, the outermost orbits of these atoms overlap and interact. When the wave functions of the electrons on different atoms begin to overlap considerably, the energy levels corresponding to those wave functions split. if more atoms are brought together more levels are formed and for a solid of N atoms, each of the energy levels of an atom splits into N levels of energy. The levels are so close together that they form an almost continuous band. The width of this band depends on the degree of overlap of electrons of adjacent atoms and is largest for the outermost atomic electrons. In a solid many atoms are brought together so that the split energy levels form a set of bands of very closely spaced levels with forbidden energy gaps between them.
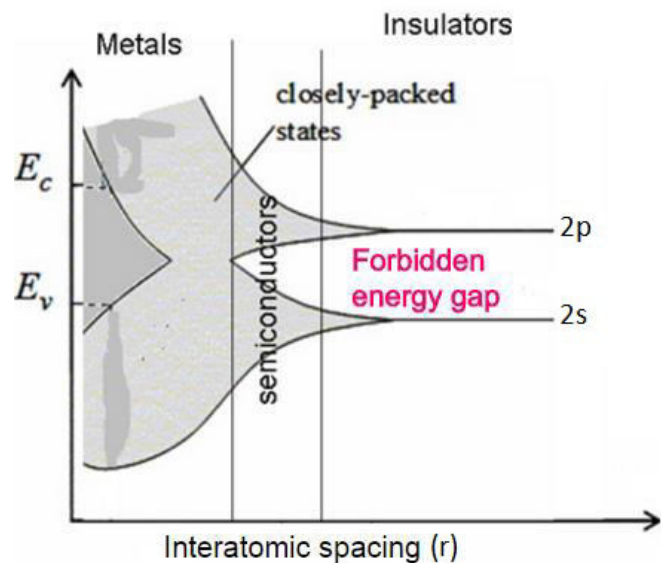
CLASSIFICATION OF MATERIALS:

The electrons first occupy the lower energy bands and are of no importance in determining many of the electrical properties of solids. Instead, the electrons in the higher energy bands of solids are important in determining many of the physical properties of solids. Hence the two allowed energy bands called valence and conduction bands are required. The gap between these two allowed bands is called forbidden energy gap or band gap since electrons can't have any energy values in the forbidden energy gap. The valence band is occupied by valence electrons since they are responsible for electrical, thermal and optical properties of solids. Above the valence band we have the conduction band which is vacant at 0 K. According to the gap between the bands and band occupation by electrons, all solids can be classified broadly into two groups.

In the first group of solids called *metals* there is a partially filled band immediately above the uppermost filled band .this is possible when the valence band is partially filled or a completely filled valence band overlaps with the partially filled conduction band.

In the second group of solids, there is a gap called band gap between the completely filled valence band and completely empty conduction band. Depending on the magnitude of the gap we can classify **insulators** and **semiconductors**.

Insulators have relatively wide forbidden band gaps. For typical insulators the band gap $E_g > 7$ eV. On the other hand, semiconductors have relatively narrow forbidden bands. For typical semiconductors $E_g \leq 1$ eV.
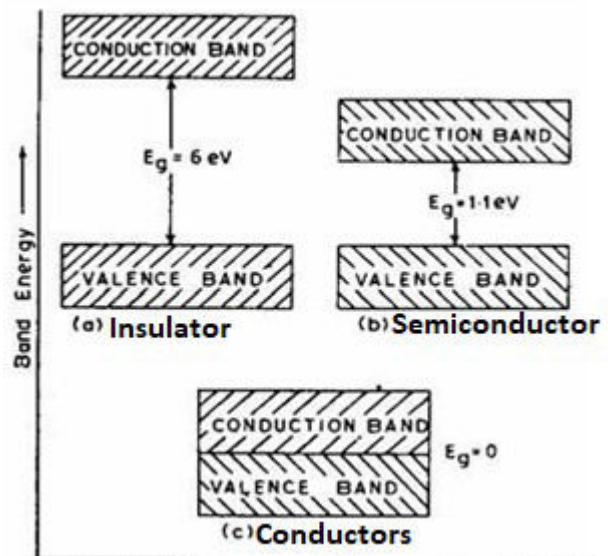
SEMICONDUCTORS

The materials in which the electrical conductivity lies between conductors and insulators are called semiconductors. They have resistivity value between 10–4 to 0.5 m . The electrical conductivity of semiconductor increases, when we added to impurities and by increasing the temperature and it is contrary to the metals. They have negative temperature Coefficient of resistance and which are formed by covalent bonds.

Elemental and Compound Semiconductors:

Semiconductors are classified as **elements or compounds**. Elemental semiconductors such as silicon (Si) and germanium (Ge) found in group IV of the Periodic Table serves as the chief component for commercially produced semiconductors. Besides pure element semiconductors, many alloys and compounds are semiconductors. Compound semiconductor materials are formed from special combinations of group III and group V elements. Compound semiconductors include a range of other materials, such as InSb, InAs, GaP, GaSb, GaAs, SiC, and GaN. Compound semiconductors have the advantage of wide range of energy gaps and nobilities, so sensing materials are available with properties that meet individual requirements.

S.No	Elemental semiconductors	Compound semiconductors
1.	These are made from single element.	These are made from compound (mixed) element.
2.	These are made from IV group and VI group elements	These are made from III and V [or] II and VI elements.
3.	These are called as indirect band gap semiconductor (electron-hole recombination takes place through traps)	These are called as direct band gap semiconductor (electron-hole recombination takes place directly)
4.	Heat is produced in the recombination recombination	Photons are emitted during
5.	Life time of charge carriers is more due to indirect recombination	Life time of charge carriers is less due to direct recombination.
6.	Current amplification is more	Current amplification is less.
7.	These are used for making diodes, transistor, etc.	These are used for making LED, laser diodes, etc.
8.	Example : Ge, Si	Example : GaAs, GaP, CdS, MgO

SEMICONDUCTORS

Introduction:- It has been observed that certain substances like Germanium (Ge), silicon (Si) etc., have resistivity (10^{-4} to $0.5\,\Omega\text{-m}$) between good conductors like copper having resistivity $1.7\times10^{-8}\,\Omega\text{-m}$ and Insulators like glass having resistivity $9\times10^{11}\,\Omega\text{-m}$. These substances are known as Semiconductors. "Thus a substance which has resistivity in between Conductors & Insulators is known as Semiconductors"

Semiconductors are broadly classified as under

Semiconductors

- Intrinsic (or) Pure Semiconductors
- Extrinsic (or) Impure Semiconductor
  - P-type
  - N-type
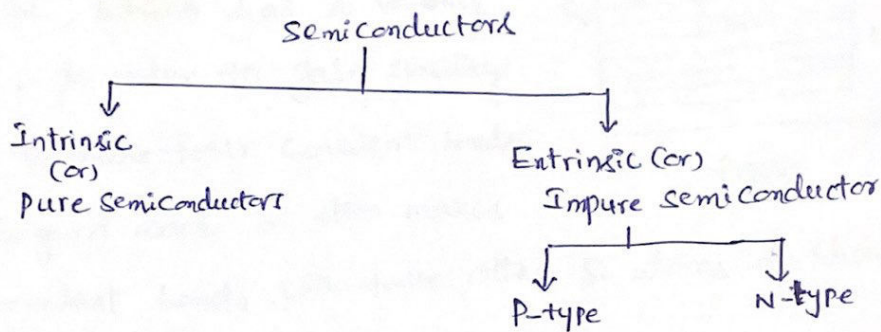
* Semiconductors have the following properties

i) They have resistivity lessthan Insulators & morethan Conductors.

ii) The Resistivity of Semiconductor decreases with the increase in temperature and vice versa. i.e. They have negative Temperature Coefficient of Resistance (-ve TCR).

iii) When suitable mettalic impurity like arsenic, gellium etc is added to a semiconductor, its current conducting properties change appreciably. This is the most imp. property.

i) **Intrinsic Semiconductor :-**

A semiconductor which does not have any kind of impurities, behaves as an insulator at $T = 0K$ and behaves as a conductor at high temperature is known as Intrinsic Semiconductor.

Ex: Si, & Ge, etc.

In order to get insight view of an intrinsic Semiconductor, let us Consider Si, which has 4 Valency electrons. In order to gain stability it has to make four Covalent bonds. In this regard each Si atom makes four covalent bonds with four other Si atoms as shown in fig ①.

Covalent bond

fig(1): Si crystal

fig ②

* The $\bar{e}$'s which are participating in the covalent bonds are known as valence $\bar{e}$'s.

* If some energy is supplied then Covalent bond break, $\bar{e}$'s will come out and move freely, Resulting in the formation of vacant sites in the covalent bonds. These are known as Positive charge carriers named as 'holes' as shown in ②

* The no. of free or Conduction $\bar{e}$'s will be equal to the no. of vacant sites or holes in the valence band $(n = P)$
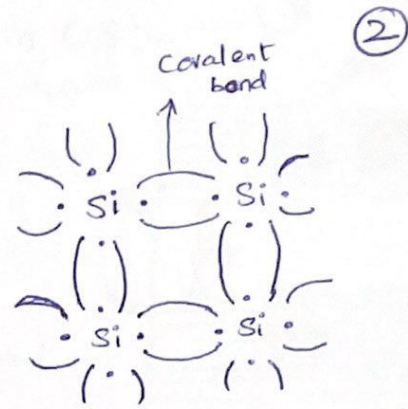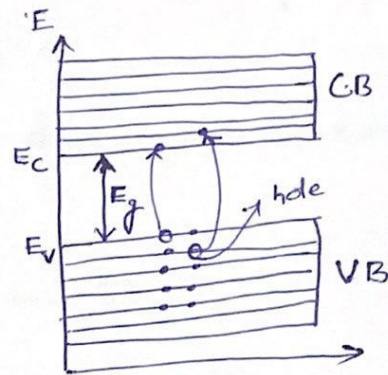
Carrier Concentration (density) of $\bar{e}$'s in C.B :-

Let 'n' be the no.of $\bar{e}$'s in the

Conduction band per unit volume of a

homogeniously doped semiconductor crystal

in equilibrium. Let us consider a

small element in the energy range of

E and E+dE. Let 'dn' is the no. of electrons in the Range

E and E+dE in Conduction band.

$$\therefore \quad dn = Z(E) \cdot F(E) \, dE \quad —— ①$$

Where , $Z(E)$ is the no. of energy states per unit volume and $F(E)$

is the Probability of electron in Conduction band in the

Range E and E+dE.

$$Z(E) \, dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} E^{1/2} \, dE$$

Energy level in the CB is above the E_c $(E > E_c)$

$$\therefore \quad Z(E) dE = \frac{4\pi}{h^3} (2m_e^*)^{3/2} (E - E_c)^{1/2} \, dE \quad —— ②$$

Fermi dirac distribution function, $F(E) = \dfrac{1}{1 + exp\left(\frac{E - E_F}{KT}\right)}$

But, $E - E_F \gg KT \Rightarrow F(E) = \left[1 + exp\left(\frac{E - E_F}{KT}\right)\right]^{-1}$

$$\therefore F(E) = exp\left[\frac{-(E - E_F)}{KT}\right]$$

$$\therefore F(E) = exp\left(\frac{E_F - E}{KT}\right) \quad —— ③$$
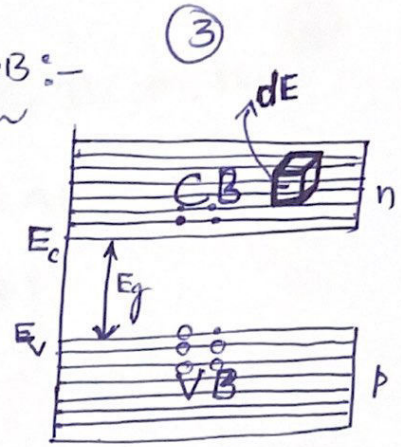
If the conduction band extends from E_c to ∞, then the electron concentration can be written as

From eq① $\Rightarrow n = \int_{E_c}^{\infty} dn = \int_{E_c}^{\infty} Z(E)\, dE\, F(E)$

From eq② & ③ $\Rightarrow$

$$n = \int_{E_c}^{\infty} \frac{4\pi}{h^3}(2m_e^*)^{3/2}(E-E_c)^{1/2} \exp\left(\frac{E_F - E}{kT}\right) dE$$

$$n = \frac{4\pi}{h^3}(2m_e^*)^{3/2} \exp\left(\frac{E_F}{kT}\right) \cdot \int_{E_c}^{\infty}(E-E_c)^{1/2} \exp\left(\frac{-E}{kT}\right) dE$$

Put $E - E_c = x \Rightarrow dE = dx$, Limits $x = 0$, & $x = \infty$

$\Rightarrow E = E_c + x$

$$\therefore n = \frac{4\pi}{h^3}(2m_e^*)^{3/2} \exp\left(\frac{E_F}{kT}\right) \int_0^{\infty} x^{1/2} \exp\left(\frac{-(E_c + x)}{kT}\right) dx$$

$$n = \frac{4\pi}{h^3}(2m_e^*)^{3/2} \exp\left(\frac{E_F}{kT}\right) \cdot \exp\left(\frac{-E_c}{kT}\right) \cdot \int_0^{\infty} x^{1/2} \exp\left(\frac{-x}{kT}\right) dx$$

But from Standard Integral, $\int_0^{\infty} x^{1/2} \exp\left(\frac{-x}{kT}\right) dx = (kT)^{3/2} \frac{\sqrt{\pi}}{2}$

$$\therefore n = \frac{4\pi}{h^3}(2m_e^*)^{3/2} \exp\left(\frac{E_F - E_c}{kT}\right) \cdot (kT)^{3/2} \frac{\sqrt{\pi}}{2}$$

$$\boxed{n = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{kT}\right)}$$

$$\boxed{n = N_c \exp\left(\frac{E_F - E_c}{kT}\right)}$$
Where $N_c = 2\left(\frac{2\pi m_e^* kT}{h^2}\right)^{3/2}$

This expression is called the carrier concentration of e^-s in CB

Carrier Concentration of holes in VB :-

Let dp be the no. of holes (vacancies) in the energy Interval E and E+dE in the Valence band,

$$dp = Z(E) [1 - F(E)] dE \qquad ①$$

Where $Z(E) dE$ is the density of states per unit volume in the energy interval E and E+dE and $[1-F(E)]$ is the probability of existance of a hole. (Since $F(E)$ is the probability of e^- occupancy, $1-F(E)$ gives the probability of an e^- not occupying the state. i·e, presence of a hole).

$$1 - F(E) = 1 - \left[1 + \exp\left(\frac{E-E_F}{KT}\right)\right]^{-1} \qquad \left(\because x = e^{\frac{E-E_F}{KT}} \right)$$

Since $E-E_F \gg KT$, $1-F(E) = 1-[1+x]^{-1}$

$$= 1 - \left(1 - \frac{x}{1!} + x^2 - x^3 + x^4 + \cdots\right)$$

$$= 1 - (1-x) = x$$

$$\therefore 1 - F(E) = \exp\left(\frac{E-E_F}{KT}\right) \qquad ②$$

density of states in the interval E & E+dE in the Valence band is

$$Z(E) dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} E^{1/2} dE$$

Since E_V is the Energy of the Top of the Valence band,

$$\therefore Z(E) dE = \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_V-E)^{1/2} dE \qquad ③$$

To calculate the no. of holes in the V.B., the eqn ① has to be integrated from $-\infty$ to E_V

⑥

The no. of holes in the V.B. per unit volume is given by.

$$P = \int_{-\infty}^{E_v} dp = \int_{-\infty}^{E_v} \frac{4\pi}{h^3} (2m_h^*)^{3/2} (E_v - E)^{1/2} \exp\left(\frac{E - E_F}{KT}\right) dE$$

$$P = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp\left(\frac{-E_F}{KT}\right) \int_{-\infty}^{E_v} (E_v - E)^{1/2} \cdot \exp\left(\frac{E}{KT}\right) dE$$

put $E_v - E = x \Rightarrow E = (E_v - x) \Rightarrow dE = -dx$

Limits $E = -\infty \Rightarrow x = \infty$ and $= E_v \Rightarrow x = 0$

$$\therefore P = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp\left(\frac{-E_F}{KT}\right)(-)\int_{\infty}^{0} x^{1/2} \exp\left(\frac{E_v - x}{KT}\right) dx$$

$$= \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp\left(\frac{-E_F}{KT}\right) \cdot \exp\left(\frac{E_v}{KT}\right) \cdot \int_{0}^{\infty} x^{1/2} \exp\left(\frac{-x}{KT}\right) dx$$

$$P = \frac{4\pi}{h^3} (2m_h^*)^{3/2} \exp\left(\frac{E_v - E_F}{KT}\right) \cdot (KT)^{3/2} \frac{\sqrt{\pi}}{2} \quad \text{(using standard Integral)}$$

Hence

$$\boxed{P = 2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2} \exp\left(\frac{E_v - E_F}{KT}\right)}$$

i.e., The no. of holes in the V.B. per unit volume is given by,

$$\boxed{P = N_v \exp\left(\frac{E_v - E_F}{KT}\right)} \qquad \therefore N_v = 2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2}$$

Intrinsic Carrier Concentration :-

In intrinsic Semiconductors, $n = p$

Hence $n = p = n_i$ is called intrinsic carrier concentration.

Therefore $\boxed{n_i^2 = n \cdot p}$

$$n_i^2 = n \cdot p = 4\left(\frac{2\pi KT}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{E_v - E_c}{KT}\right)$$

$$= 4\left(\frac{2\pi KT}{h^2}\right)^3 (m_e^* m_h^*)^{3/2} \exp\left(\frac{-E_g}{KT}\right)$$

Where $E_c - E_v = E_g$ is the forbidden energy gap,

Hence,

$$\boxed{n_i = 2\left(\frac{2\pi KT}{h^2}\right)^{3/2} (m_e^* m_h^*)^{3/4} \exp\left(\frac{-E_g}{2KT}\right)}$$

This is the Expression for intrinsic Carrier Concentration.

Fermi Level :—

Since $n = p$ in intrinsic Semiconductors,

$$2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2} \exp\left(\frac{E_F - E_c}{KT}\right) = 2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2} \exp\left(\frac{E_v - E_F}{KT}\right)$$

or $(m_e^*)^{3/2} \exp\left(\frac{E_F - E_c}{KT}\right) = (m_h^*)^{3/2} \exp\left(\frac{E_v - E_F}{KT}\right)$

$$\exp\left(\frac{2E_F}{KT}\right) = \left(\frac{m_h^*}{m_e^*}\right)^{3/2} \cdot \exp\left(\frac{E_c + E_v}{KT}\right)$$

Taking logarithms on both sides,

$$\frac{2E_F}{KT} = \frac{3}{2} \log\left(\frac{m_h^*}{m_e^*}\right) + \frac{E_c + E_v}{KT}$$

$$\therefore E_F = \frac{3KT}{4} \log\left(\frac{m_h^*}{m_e^*}\right) + \left(\frac{E_c + E_v}{2}\right)$$

If we assume that $m_e^* = m_h^* \Rightarrow \boxed{E_F = \frac{E_c + E_v}{2}}$

Thus fermi level is located half
way B/w the VB & CB and its position is independent of Temperature.

ii) Extrinsic Semiconductors:-

An extrinsic semiconductor is an intrinsic semiconductor doped by a specific impurity (B, In, P, etc.) which is able to modify deeply its electrical properties, making it suitable for electronic or opto electronic applications.

Extrinsic semiconductors are basically of two types:

    b). p-type semiconductor

    a) n-type semiconductor

a). n-type semiconductor:- when pentavalent impurity is added to an intrinsic semiconductor (Si or Ge), then it is said to be an n-type semiconductor. Pentavalent (+5) impurities such as P, As, Sb, etc, are called **donor impurity**

Let us consider, pentavalent impurity P is added to Si as shown in fig①. P atom has 5 valence $\bar{e}s$ and Si has 4 valence $\bar{e}s$. P atom has one excess valence e^- than Si. The 4 valence $\bar{e}s$ of each P atom form 4 covalent bonds with 4 neighbouring Si atoms.
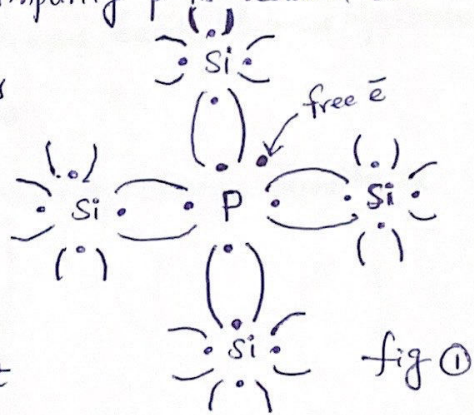
The fifth valence e^- of P atom does not involve in the formation of covalent bonds. Hence it is free

fig ①
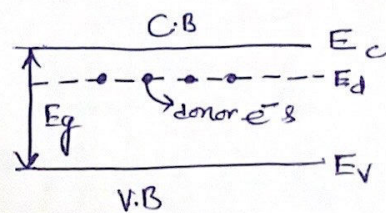
to move and not attached to the parent atom. This shows that each p atom donates one free $\bar{e}$. Therefore, all the pentavalent impurities are called donors. The no. of donor electrons are depends on the amout of impurity (P) added to the Si.

In case of n-type semiconductor, the following points should be remembered:

i) In n-type semiconductor, the $\bar{e}$'s are majority carriers while holes are minority carriers

ii) n-type semiconductor has excess $\bar{e}$'s but it is electrically neutral. This is due to the fact that $\bar{e}$'s are created by the addition of neutral V-valent impurity atoms to the S.C.

iii) In n-type S.C, In addition to the free $\bar{e}$'s of impurity atoms, the density of free $\bar{e}$'s is given by

$$n = n_i + N_d$$

Where n_i is the density of intrinsic S.C. $\bar{e}$'s and N_d, the density of ionised donor atoms formed due to thermal aggitation.

$$N_d > n_i \qquad \therefore n \simeq N_d$$

In a purely n-type crystal, the minority charge density is given by

$$P_n = \frac{n_i^2}{n} = \frac{n_i^2}{N_d}$$

Fermi level in n-type S.C. and variation with temperature :-

The energy level diagram of n-type S.C is shown in fig① N_d is the donor concentration

fig①
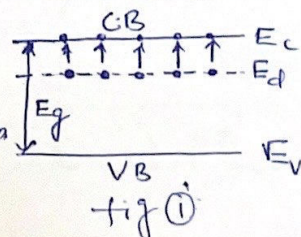
i.e., the no. of donor atoms per unit volume of the material and E_d is the donor energy level. At very low temperatures all the donor levels are filled with electrons. With increase of temperature more & more donor atoms gets ionized & the density of $\bar{e}$'s in the CB. increases. Density of electrons in the Conduction Band according to pure Si atoms (intrinsic).

$$n = 2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2} exp\left(\frac{E_F - E_c}{KT}\right) \quad —\text{①}$$

donor atom ($\bar{e}$) Concentration in the energy level E_d is n_d

$$n_d = N_d[1 - F(E_d)] \approx N_d exp\left(\frac{E_d - E_F}{KT}\right)$$

At very low temperatures, $n = n_d$

$$2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2} exp\left(\frac{E_F - E_c}{KT}\right) = N_d exp\left(\frac{E_d - E_F}{KT}\right)$$

Taking logarithm on both sides, we get

$$\left(\frac{E_F - E_c}{KT}\right) - \left(\frac{E_d - E_F}{KT}\right) = lg N_d - log 2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2}$$

$$2E_F - (E_d + E_c) = KT log\left[\frac{N_d}{2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2}}\right]$$

$$E_F = \frac{E_d + E_c}{2} + \frac{KT}{2} log\left[\frac{N_d}{2\left(\frac{2\pi m_e^* KT}{h^2}\right)^{3/2}}\right]$$

At $T = 0K$,

$$\boxed{E_{F_n} = \frac{E_d + E_c}{2}}$$

With increase of temperature, more and more donor atoms are ionised and also starts the generation $\bar{e}$-hole pairs. The fermilevel moves gradually towards the intrinsic fermi level E_{Fi} as shown in fig ②

fig ②
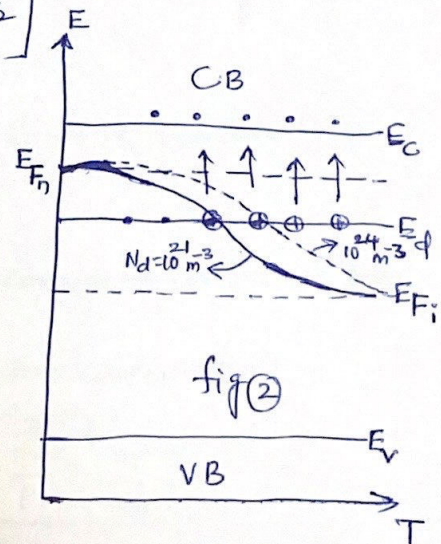
(b) **P-TYPE Semiconductor :-**

When a small amount of trivalent impurity is added to a pure (intrinsic) Semiconductor. Then it is said to be p-type Semiconductor.
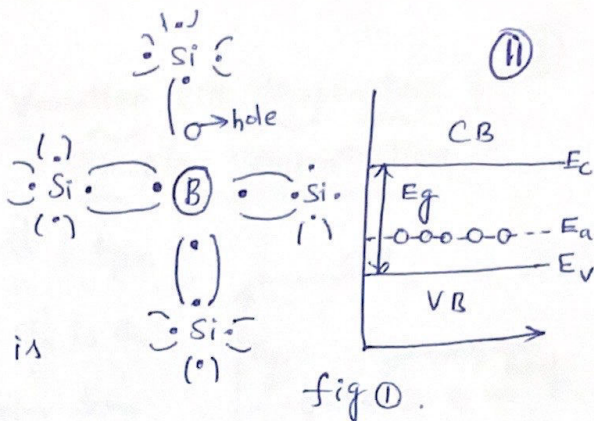
Trivalent impurities (+3) Such as B, Ga, In, etc.,. Let us consider the case when trivalent Boron (B) is added to pure Si crystal as shown in fig ①. As shown in fig ① each atom of B fits into Si Crystal with only three Covalent bonds. In the forth covalent bond, only Si atom Contributes one valence e^- and there is a deficiency of one electron which is called a 'Hole'. The addition trivalent impurity atoms, produces large no. of holes are termed as "acceptors" because they accept e^-s from Si atoms. Since Current carriers are +vely charged (holes), This type of Semi-conductor is called P-type s.c.

i) In P-type S.C. materials, the majority carriers are 'holes' and minority carriers are e^-s

ii) In p-type S.C., there are holes due to acceptor impurity in @ addition to free e^-s & holes of intrinsic S.C.

iii) The density of holes in P-type-s.c. at Room temperature is

$$P = P_i + N_a$$

where N_a is the acceptor concentration, P_i – concentration of intrinsic holes.

In P-type – S.C., $N_a \gg P_i$

$\therefore$ $\boxed{P \approx N_a}$

free electron Concentration
(Minority)

$$\boxed{n_p = \frac{n_i^2}{P} = \frac{n_i^2}{N_a}}$$

fig ①.

Fermi level in p-type S.c & Its Variation with temperature & acceptor concentration :-

The energy level diagram of p-type Semiconductor is shown in fig ①. N_a is the acceptor concentration (The no. of acceptor atoms Per unit volume of the material) and E_a is the acceptor energy level slightly above the V.B. edge E_v as shown in fig ①.

fig ①
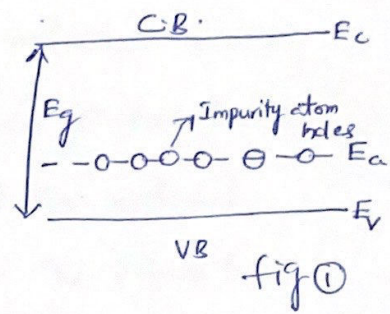
Concentration of holes (intrinsic) in Valence band is given by,

$$p = 2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2} exp\left(\frac{E_v - E_F}{KT}\right)$$

If some of the doped atoms get ionized by accepting the electrons, let N_a be the acceptor Concentration,

$$= N_a F(E_a) = N_a\left[1 + exp\left(\frac{E_a - E_F}{KT}\right)\right]^{-1}$$

$$= N_a exp\left(\frac{E_F - E_a}{KT}\right)$$

Since the Concentration of holes in the V.B. is equal to the Concentration of ionized acceptors,

$$2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2} exp\left(\frac{E_v - E_F}{KT}\right) = N_a exp\left(\frac{E_F - E_a}{KT}\right)$$

$$exp\left(\frac{E_v + E_a - 2E_F}{KT}\right) = \frac{N_a}{2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2}}$$

Taking logarithm,

$$E_{F_p} = \frac{E_v + E_a}{2} - \frac{KT}{2} log \frac{N_a}{2\left(\frac{2\pi m_h^* KT}{h^2}\right)^{3/2}}$$

fig ②
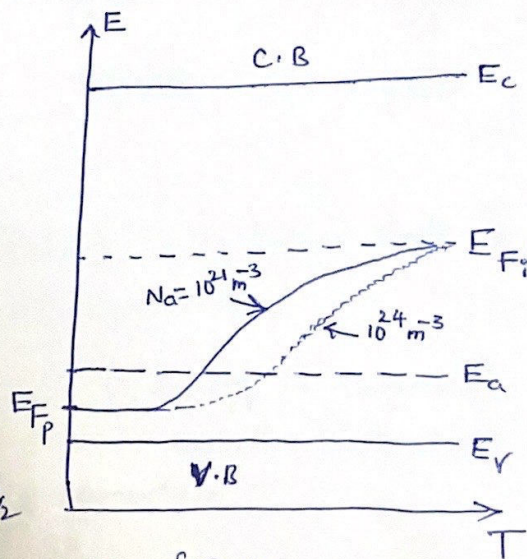
At $T = 0K$,

$$E_{F_p} = \frac{E_v + E_a}{2}$$

As the temperature is increased, more and more acceptor atoms are ionized and results in generation of e^--hole pairs due to breaking of covalent bonds and the material tends to behave ~~intaim~~ in intrinsic manner. The Fermi level of Extrinsic p-type S.C gradually moves towards the intrinsic fermi level E_{F_i} as shown in fig ②.

——— * ———

Direct and Indirect Band gap Semiconductors:-

1) Direct Band gap:- The minimum energy state in the CB; and the maximum energy state in the ② V.B. are characterized by same k-vector, it forms direct band gap semiconductor as shown in fig ①ⓐ.
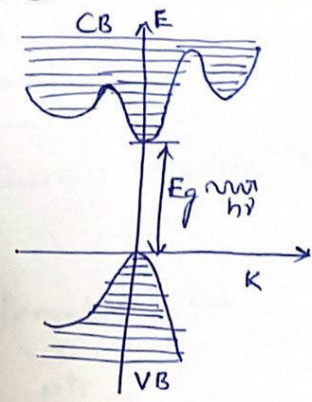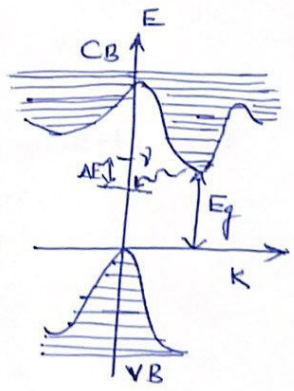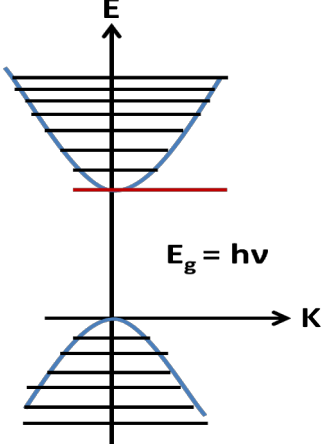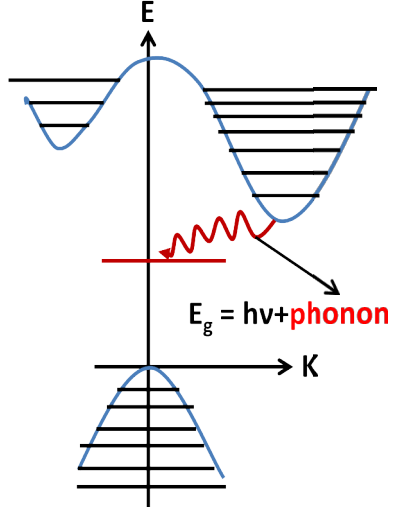
fig ①ⓐ

fig ①ⓑ

* An e^- can shift from the lowest state of CB to higher state of V.B without any change in Crystal momentum.

* When an e^- and hole combine the excess energy is released in the form of photon & this process is called radiative recombination.

* It obeys conservation of Energy & Momentum

Ex:- GaAs, InP etc,

Module-III-B: Semiconductors

DIRECT AND INDIRECT BAND GAP SEMICONDUCTORS

Direct band gap semiconductor	Indirect band gap semiconductor
1. These are impure or Extrinsic or compound semiconductors 2. Examples: InP, GaAs, GaAsP etc 3. $E_g = h\nu$	1. These are pure or intrinsic or elemental Semiconductors. 2. Examples: Ge, Si. 3. $E_g = h\nu + \text{phonon}$
4. The minimum energy of Conduction band (CB) and maximum energy of valence band (VB) have the same value of wave vector , i.e. $k_1 = k_2$ 5. Here an electron from CB to can recombine with a hole in VB directly by emitting light of photon of energy 'hϑ'. 6. They are used to fabricate LEDs, Laser Diodes etc. 7. Life time (recombination rate) of charge carriers is less. 8. Emission of light has energy gap is $E_g = hC/\lambda$ eV	4. The minimum energy of Conduction band (CB) and maximum energy of valence band (VB) have the different values of wave vector i.e. $k_1 \neq k_2$ 5. Here an electron from CB cannot recombine directly with holes in VB. But can recombine through traps by emitting light without emission of photon or light. 6. They are used to amplify the signals in electronic devices like rectifiers, transistors, amplifiers etc. 7. Life time of charge carriers is more. 8. No emission of light. It conducts only Electricity.

Carrier Generation and Recombination:

Generation:

Carrier generation due to light absorption occurs if the photon energy is large enough to raise an electron from the valence band into an empty conduction band state, thereby generating one electron-hole pair. The photon energy needs to be larger than the bandgap energy to satisfy this condition. The photon is absorbed in this process and the excess energy, E_{ph} - E_g, is added to the electron and the hole in the form of kinetic energy.
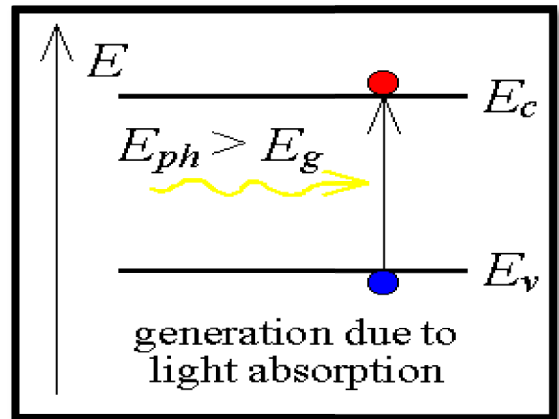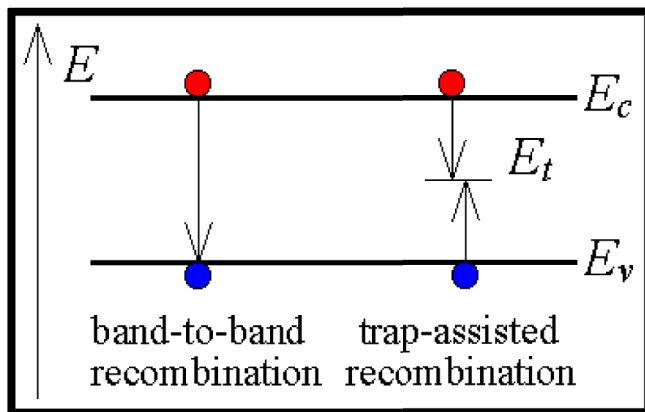

Recombination:

Recombination and generation are always happening in semiconductors, both optically and thermally, and their rates are in balance at equilibrium. Carrier recombination mechanisms in semiconductors mainly by two ways: i) band-to-band (direct) recombination and ii) trap-assisted (Indirect) recombination.

i) Band-to-band(direct) recombination:

Band-to-band recombination occurs when an electron moves from its conduction band state into the empty valence band state associated with the hole. This band-to-band transition is typically also a radiative transition in direct bandgap semiconductors.



ii) Trap-assisted (Indirect) recombination:

Trap-assisted recombination occurs when an electron falls into a "trap", an energy level within the bandgap caused by the presence of a foreign atom or a structural defect. Once the trap is filled it cannot accept another electron. The electron occupying the trap, in a second step, moves into an empty valence band state, thereby completing the recombination process. One can envision this process as a two-step transition of an electron from the conduction band to the valence band or as the annihilation of the electron and hole, which meet each other in the trap. We will refer to this process as Shockley-Read-Hall (SRH) recombination.

DRIFT & DIFFUSION:

The net flow of current through a SC material in two ways: i) Drift, & ii) Diffusion.

i) Drift Current: The movement of charge carriers in a particular direction by the application external filed, E, constitute the electric current called **drift current** and this phenomenon is called **'drift'**.

Let us cossider an n-type SC with large no. free electrons in it and is placed in an electric filed, E. So the electrons and holes in SC experiences a force.
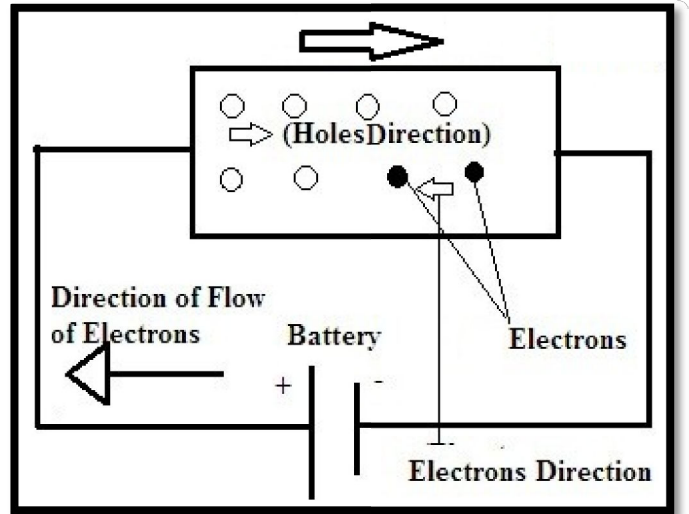
The charge carriers (electrons) gains the valocity called **drift velocity(v_{dn})**, which is proportional to applied field, E.,

$$v_{dn}\, \alpha - E \Longrightarrow v_{dn} = -\mu_n E ----- (1) \quad \mu_n\text{-mobility of electrons (m}^2\text{/V-s)}$$

The current density due to electron Drift is given by, $J_n = nev_{dn}$ ------ (2) n-concentration of e⁻s

Therefore, $J_n = -ne(-\mu_n E) \Longrightarrow \boldsymbol{J_n = ne\mu_n E}$ ---------- (3)

Conductivity, $\sigma_n = \dfrac{J_n}{E} = \dfrac{ne\mu_n E}{E} \Longrightarrow \boldsymbol{\sigma_n = ne\mu_n}$-------- (4)

The drift current density and conductivity in p-type SC due to holes is,

$$\boldsymbol{J_p = pe\mu_p E} \text{ ------- (5) and } \boldsymbol{\sigma_p = pe\mu_p}\text{---------- (6)}$$

The total drift current density is, $\boldsymbol{J_{drift} = J_n(drift) + J_p(drift)}$

Therefore, from eq(3&5) and eq(4&6), we get, $\boldsymbol{J_{drift} = ne\mu_n E + pe\mu_p E}$

$$\Longrightarrow \boldsymbol{J_{drift} = eE(n\mu_n + p\mu_p)} \text{ and } \boldsymbol{\sigma_{drift} = e(n\mu_n + p\mu_p)}$$

For an intrinsic SC, $\boldsymbol{n = p = n_i}$

Drift **currrent density** and **conductivity** in intrinsic SC $\boxed{\boldsymbol{J_{drift} = n_i eE(\mu_n + \mu_p)}}$ &

$$\boxed{\boldsymbol{\sigma_i(drift) = n_i e(\mu_n + \mu_p)}}$$

ii) Diffusion Current:

Let us consider a non-uniformly doped n-semiconductor which has more number of electrons at one end and less at another end. The process by which, charge carriers (electrons

or holes) in a semiconductor moves from a region of higher concentration to a region of lower concentration is called **diffusion**.

The diffusion of electrons and holes across the junction depend on the concentration gradient of electrons and holes across the function. If we consider the direction of current flow as x – direction and $\frac{dn}{dx}$ is the concentration gradient of electrons.
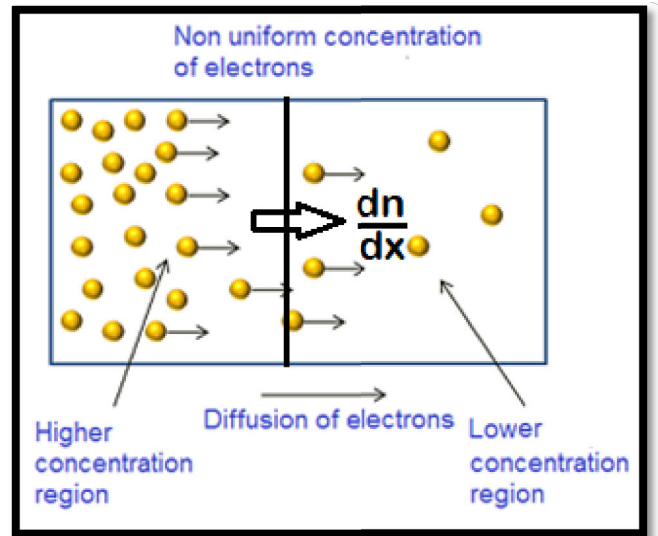
According to Fick's Law of diffusion, the rate of diffusion of electrons is proportional to $-\frac{dn}{dx}$.

Therefore, diffusion rate of e⁻s $= -D_n\frac{dn}{dx}$ ------- (1), where D_n is the diffusion coefficient of electrons in SC.

Therefore, the diffusion current density of e⁻s is given by, $J_n = -e\left(-D_n\frac{dn}{dx}\right) = eD_n\frac{dn}{dx}$ ----- (2)

Similarly, the diffusion current density of **holes** is given by, $J_p = +e\left(-D_p\frac{dp}{dx}\right) = -eD_p\frac{dp}{dx}$ -- (3)
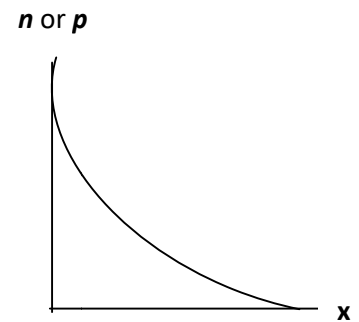
Therefore,

Total current density due to electrons in a SC is given by

$$J_n = J_n(drift) + J_n(diffusion)$$

$$\boxed{J_n = ne\mu_n E + eD_n\frac{dn}{dx}}$$

Similarly, **for holes,** $\boxed{J_p = pe\mu_p E - eD_p\frac{dp}{dx}}$

The variation of excess carrier concentration with respect to x is shown in fig.

CONTINUITY EQUATION:

Equation of continuity describes the behavior of charge carriers in Semiconductor. Consider a block of semiconductor material of cross-sectional area, A and an infinitesimal region within the semiconductor of
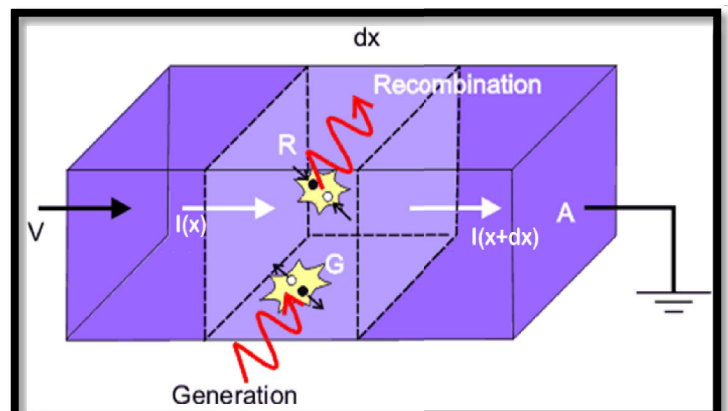
thickness, dx at x. The mathematical relations formed from the excess carriers in unit volume of semiconductor due to **generation**, **recombination**, **drift** and **diffusion** are called <u>continuity equations in semiconductors</u>.

Let us consider the p-type semiconductor. Let **'p'** be the **no. of holes** in the given piece of SC per unit volume and **'τ_p'** is the mean life time of holes.

No. of holes lost due to recombination per unit time per unit volume $= \dfrac{p}{\tau_p}$

Therefore, No. of holes lost due to **recombination** in entire volume

$$= \textbf{charge*volume*recombination rate} = e.Adx.\dfrac{p}{\tau_p} \text{ ------ (1)}$$

No. of holes increases due to **generation = charge*volume*generation rate**

$$= e.Adx.\textbf{g} \text{ ---------- (2)} \text{ where, '}g\text{' is the generation rate}$$

If **'I'** is the current entering at x and **'I+dI'** is the current leaving at $x+dx$. Therefore, the decrease of charge carriers per sec. from entire volume = **dI** ------ (3)

But, the rate of change of hole concentration in the total volume $= e.Adx.\dfrac{dp}{dt}$ ---------------- (4)
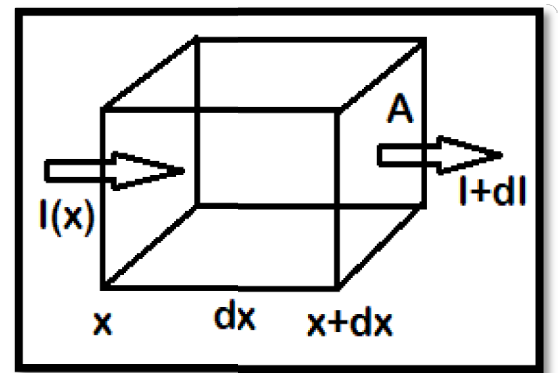
According to **Law of conservation of charge,**

Rate of change of charge carriers [eq(4)] = Generation [eq(2)]– Loss [eq(1+3)]

$$\Rightarrow e.Adx.\dfrac{dp}{dt} = e.Adx.\textbf{g} - \left[e.Adx.\dfrac{p}{\tau_p} + dI\right]$$

$$\Rightarrow \dfrac{dp}{dt} = \textbf{g} - \dfrac{p}{\tau_p} - \dfrac{dI}{e.Adx} \text{ --------- (5)}$$

When No external filed is applied and at thermal equilibrium, the hole concentration attains a constant value, i.e. $p = p_o$.

$$\Rightarrow \textbf{dI = 0 and } \dfrac{dp}{dt} = \dfrac{dp_o}{dt} = \textbf{0,}$$

$$\therefore \text{ From } eq(5), \text{we get}, \textbf{g} = \dfrac{p_o}{\tau_p} \text{ ----- (6)}$$

Substitute **eq(6) in eq(5), we get,**

$$\Rightarrow \dfrac{dp}{dt} = \dfrac{p_o - p}{\tau_p} - \dfrac{1}{e}\dfrac{dJ_p}{dx} \text{ ------ (7)} \text{ where } J = \dfrac{I}{A} \Rightarrow dI = A.dJ$$

The resultant current density of holes due to **drift** and **diffusion** is given by,

$$J_p = J_p(drift) + J_p(diffusion) \implies J_p = pe\mu_p E - eD_p \frac{dp}{dx} \text{ ------- (8)}$$

Therefore, from **eq(7&8),** we get,

$$\implies \frac{dp}{dt} = \frac{p_o - p}{\tau_p} - \frac{1}{e}\frac{d}{dx}\left(pe\mu_p E - eD_p \frac{dp}{dx}\right)$$

$$\boxed{\frac{dp}{dt} = \frac{p_o - p}{\tau_p} - \mu_p E \frac{dp}{dx} + D_p \frac{d^2p}{dx^2}}$$ **Continuity eq.** for **p-type** semiconductor

Similarly for electrons in n-type SC,

$$\boxed{\frac{dn}{dt} = \frac{n_o - n}{\tau_n} - \mu_n E \frac{dn}{dx} - D_n \frac{d^2n}{dx^2}}$$ **Continuity eq.** for **n-type** semiconductor

P-N JUNCTION DIODE:

A PN Junction diode is two-terminal semiconductor device, which allows the electric current in only one direction while blocks the electric current in opposite or reverse direction. If one side of a single crystal of pure semiconductor (Ge or Si) is doped with acceptor impurity atoms and the other side is doped with donor impurity atoms, a PN junction is formed as shown. The symbol of the p-n junction diode is
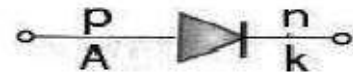
In practice a PN junction diode may be obtained in any of the following three methods.

1. Grown junction type

2. Fused (or alloyed) junction type and

3. Diffused junction type

1) Grown junction type:

When an extrinsic semiconductor is grown from melt, during the middle of the growth process impurities of opposite kind are added to the melt so that the opposite type of crystal grows further for example: when an n type Si crystal containing phosphorous atoms as pentavalent impurity is grown, the growth process is stopped temporarily and efficient boron atoms are added to the melt as trivalent impurity. This results in growth of p-type thereafter.

2) Fused (or Alloyed) junction type:

In this type p-type and n-type materials are kept in contact and fused together by proper heat treatment to form the junction. For example, a small dot of indium (p-type material) is pressed on thin water of (n-type) germanium. During few minutes of heat treatment, indium
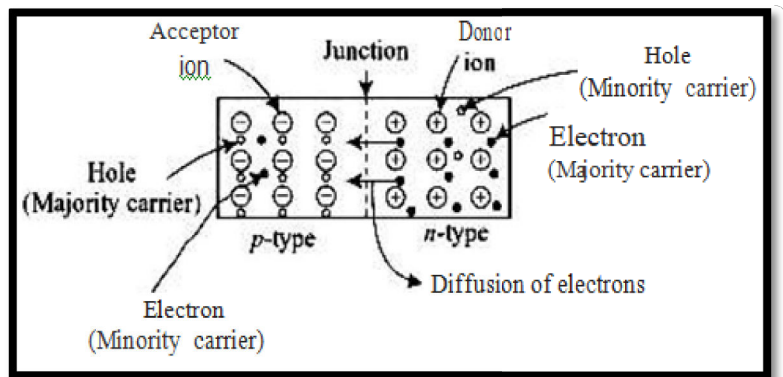
atoms fuse in to the surface of germanium and produce p-region inside germanium. This p-type region with n-type germanium water forms a p-n junction.

3) Diffused Junction type:

Diffused junctions are formed by impurity diffusion technique. The diffusion process employs either gas diffusion method or solid diffusion method. For example, in gas diffusion method, a wafer of n-type silicon is heated at about 1000°C in a gaseous atmosphere of high concentration gradient the boron atoms. At the temperature due to concentration gradient the boron atom diffuse in to silicon forming p-n junction. In solid diffusion process a p-type impurity (say indium) is painted on a n-type substrate and both are heated. Now impurity atoms diffuse into n-type substrate for a short distance and form p-n junction.

Formation of p-n Junction diode:

The formation PN junction is represented in a figure. Let us consider the formation of a sharp junction when two separate semiconductors of P and N-type are brought together. As soon as the junction is formed, free electrons and holes cross through the

junction by the process of diffusion. During this process, the electrons crossing the junction from N-region into the P-region recombine with holes in the P-region very close to the junction. Similarly holes crossing the junction from the P-region into the N-region, recombine with electrons in the N-region very close to the junction. Thus a region is formed, which does **not have any mobile charges** very close to the junction. This region is called **depletion region (potential barrier)**. The thickness of this region is in the older of 6×10^{-6} cm.
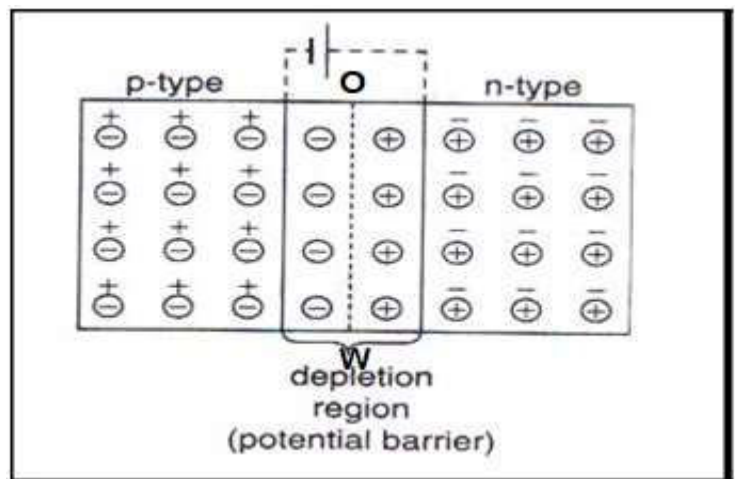
This leaves n-region near the boundary positively and p-region negatively charged as a result, electric field E_B appears in a small region W on either side of the junction O. Due to electric field E_B, potential difference appears across the depletion region and this potential V_B is called "Contact Potential or Barrier Potential". The barrier potential is approximately **0.7V & 0.3V** for Si & Ge PN junctions respectively. The width of the potential barrier depends upon the nature of the material.

I-V Characteristics of p-n junction diode:

There are two operating regions and three possible "biasing" conditions for the standard **p-n Junction Diode** and these are:


i. **Zero Bias:**– No external voltage is applied to the PN junction diode. Then the current is zero.
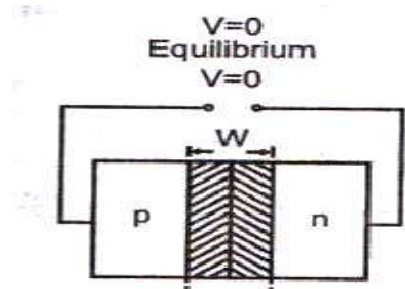
ii. **Forward Bias:**
When a diode is connected in a Forward Bias condition, a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material. If this external voltage becomes greater than the value of the potential barrier, approx. 0.7 volts for silicon and 0.3 volts for germanium, the potential barriers opposition will be overcome and current will start to flow.

The application of a forward biasing voltage on the junction diode results in the depletion layer becoming very thin and narrow which represents a low resistance path through the junction thereby allowing high currents to flow. The point at which this sudden increase in current takes place is represented on the static I-V characteristics curve above as the "**knee" point**.
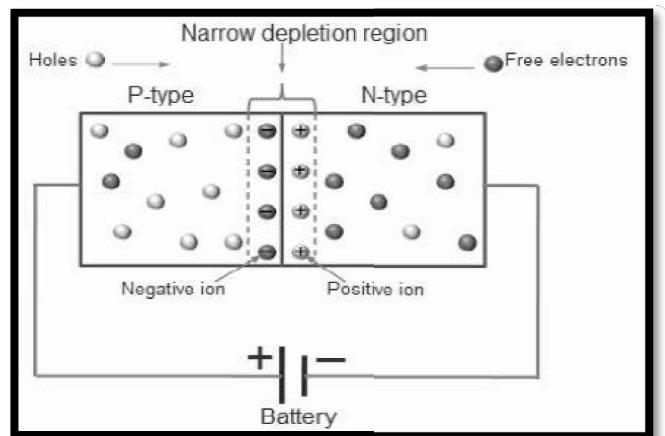
iii. **Reverse Bias:**
When a diode is connected in a **Reverse Bias** condition, a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material. The positive voltage applied to the N-type material attracts electrons towards the positive electrode and away from the junction, while the holes in the P-type end are also attracted away from the junction towards the negative electrode.

This condition represents a high resistance value to the PN junction and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small **reverse leakage current** does flow through the junction which can normally be measured in micro-amperes, (μA).
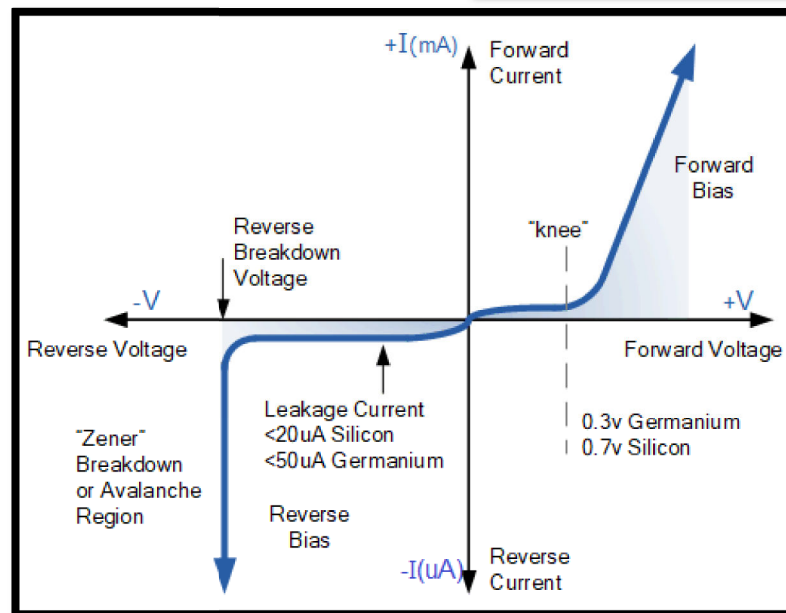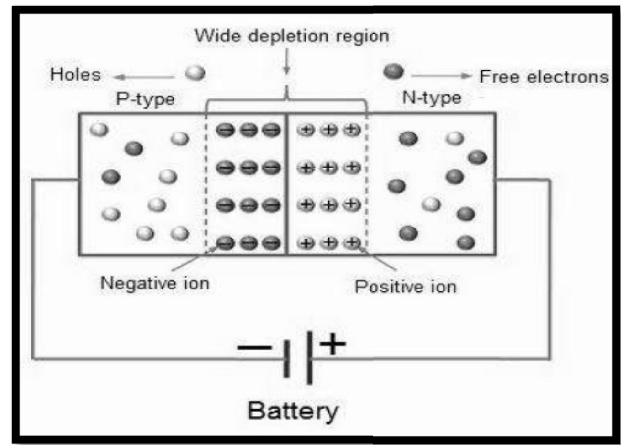
Applications:

The applications of a PN junction diode are as follows:

- We can use the p-n junction diode as a **photodiode**, as this diode is sensitive to the light when the configuration of the diode is changed to reverse-biasing.
- As a **solar cell**.
- When the diode is forward-biased, we can use it in **LED** lighting applications.
- We can use the PN junction diode as **rectifiers** in many electric circuits and as a voltage-controlled oscillator in the device like **varactors**.
- We can use it in **clipping circuits** as wave shaping circuits in **computers, radios, radars**, etc.
- In **detectors** and **demodulator** circuits.
- As a **switch** in **digital logic designs**.

ENERGY BAND DIAGRAM OF P-N JUNCTION DIODE:

The energy levels of valance band, conduction band and Fermi level of both p-type and n- type semiconductors are shown in fig(a&b). When P-n junction is formed the Fermi levels become common for both the types formation of potential barrier is represented in fig(c). The potential, VB across the junction is due to the potentials of depletion region [fig(d)] on n- side denoted by V_n and p – side denoted by V_p,

i.e., $V_B = V_n - V_p$

The contact potential separates the energy bands in p-type and n- type crystals. Since there is no net current flow at equilibrium i.e under unbiased condition, it should have common Fermi level. Hence in PN junction diode the valance and conduction band energy levels Evp and Ecp of p- type at higher level compared to the valance and conduction band energy levels Evn and Ecn of n-type . The electric field E_B across the junction is given by

$$E_B = E_{vp} - E_{vn} = E_{cp} - E_{cn} = eV_B$$

LIGHT EMITTING DIODE (LED):

LED is just a normal SC p-n junction diode that works under forward bias condition which are fabricated with direct band gap SC(compound SC) such as GaAs, InP, GaAsP etc.,. But these devices can emit EM radiation by 'electroluminescence'(Emission of radiation from a solid when it is supplied with some form of energy is known as LUMINESCENCE.) in the UV, visible or Near IR regions. The Quanta
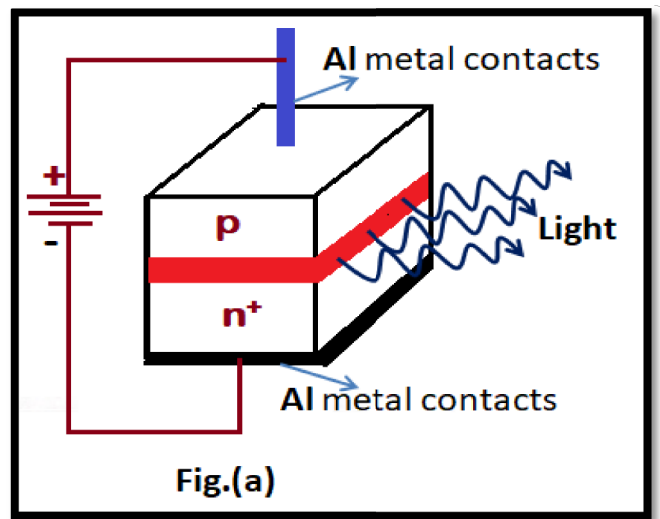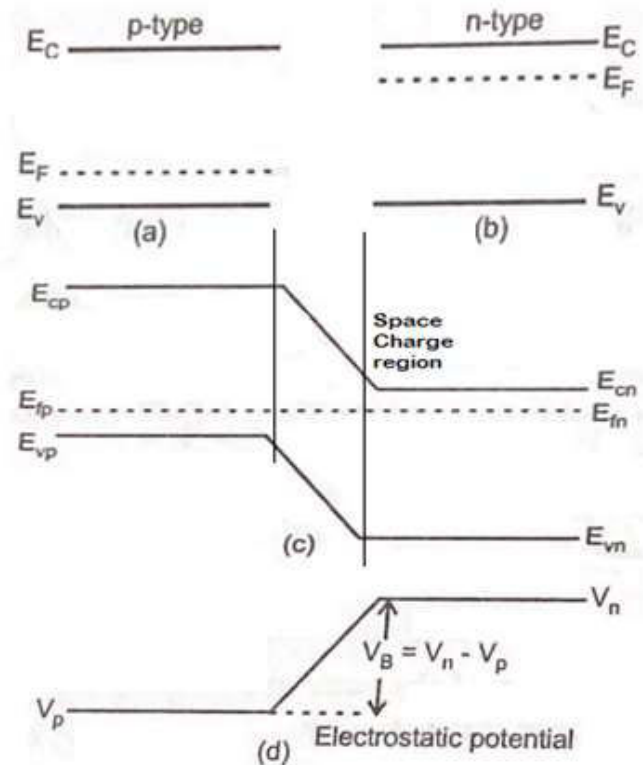
(Photon) of light energy is released and it is proportional to the band gap of that SC. The symbol of LED is shown below.

Fig(b)

As shown in fig.(a), the LED is developed on a suitable substrate. A heavily doped n-type SC on which a p-type layer is formed by different fabrication methods to develop the junction between them. An aluminum layer is coated on top and bottom to provide the metal contacts.
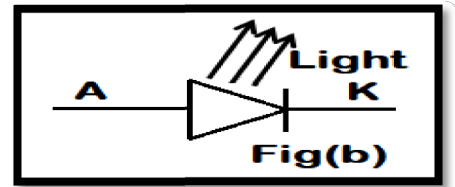
Working of LED:

PN junction diode operated under forward bias gives rise to emits the EM radiation by electroluminescence. Under **forward bias**, **majority carriers** from both sides of the junction cross the potential barrier and enter the other side of the junction where they are **minority carriers**. This process is called **minority carrier injection.** These excess minority carriers undergo radiative recombination with majority carriers as illustrated in fig (c), emitting **photons**. In **reverse bias**, no carrier injection takes place and consequently **no photons** are emitted.

Fig.(c)
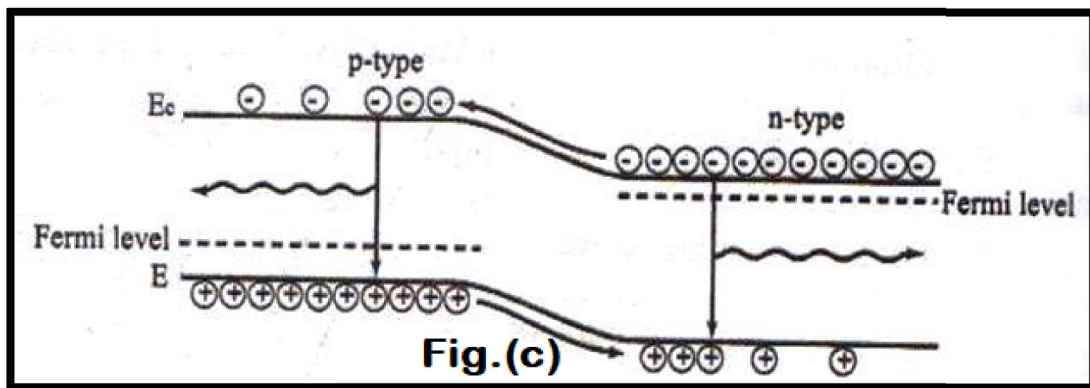
- To get the desired colored LED, we must choose suitable material with right Eg, such that

$$E_g = \frac{hc}{\lambda}$$

LED materials: GaAs - Red

GaAsP - Red/Yellow

GaP - Red/Green

Applications:

1. For Instrument Display,
2. Calculators,
3. Digital clocks
4. In traffic signal lights
5. For optical switching circuits,
6. Light detectors,
7. In optical communications,
8. In semiconductor Lasers, etc.,

PHOTO DIODE:

Photo diode is a semiconductor device which **converts light energy into electric current**. When it is exposed to light, it absorbs the light and forms the electron hole pairs, thus the current is generated. It is also called as **photo detector** or **light detector** and **photo sensor** or **light sensor**. The function of the **photo diode** is opposite to an LED function. In an LED, photons are released in response to the current flow through the junction. **In a photo diode, the photons are absorbed resulting in the generation of the carriers that manifest as current through the junction**. The symbol of photo diode is similar to the PN junction diode. The difference is the arrow pointing towards the diode.
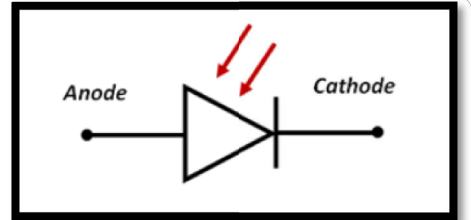

Construction:

The most common semiconductor material used for photodiodes is **silicon**. It has an energy gap of 1.14eV. It provides excellent photodiodes with quantum efficiencies up to 80% at wavelength between 0.8-0.9µm. It shows a typical silicon photodiode structure for photo-conductive operation.

A junction is formed between heavily doped p- type material (p^+) and fairly lightly doped n-type material so that the depletion region extends well into the n- material. The p^+ layer is made fairly thin. Metallic contracts can be made directly to the p+ material but to obtain an ohmic contact to the n – material an intermediate n^+ layer is formed.
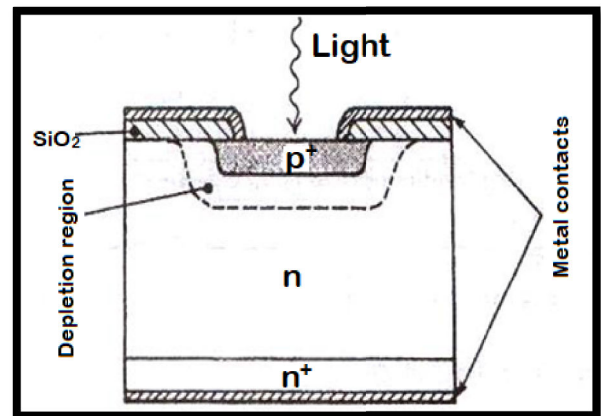

Working of photo diode:

When the device is **exposed to radiation** or illuminated the photons are absorbed by the diode. The electrons in the valence band get energized and they jump to the conduction band leaving hole in the valence band. Thus, electron hole pair is produced at the junction.
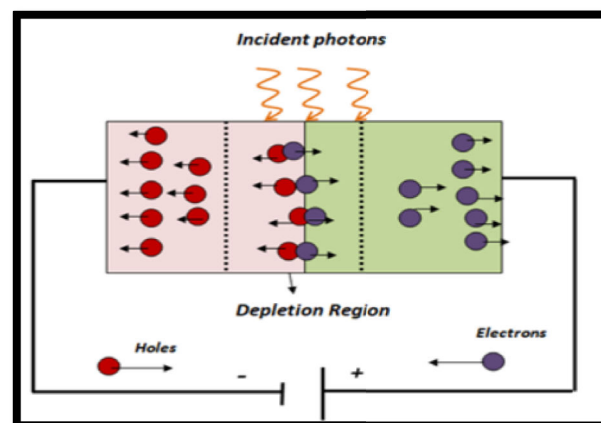
When the photo diode is **reverse biased**, the electrons are attracted towards the positive terminal and the holes are attracted towards negative terminal. If the reverse biased voltage is increased the electron hole pair breaks and holes move toward the negative terminal and the electrons move towards positive terminal which results in high reverse current. This is called **photo current**.
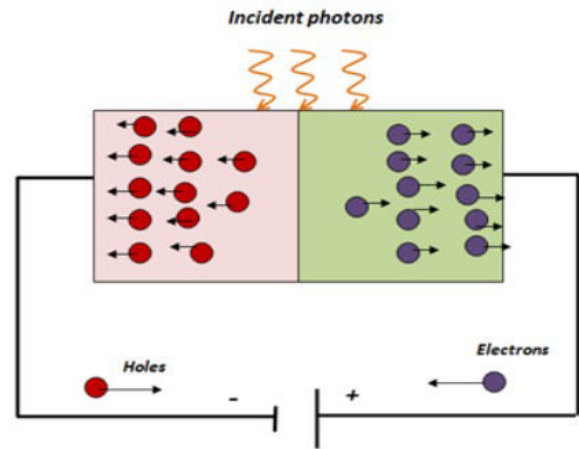
Advantages:

1. Low noise
2. Highly sensitive to noise
3. It operates at high frequencies
4. Long lifetime
5. High voltage is not required
6. It operates very fast

Disadvantages:

1. The active area is very small
2. Temperature stability is less, when the temperature increases the dark current also increases
3. Change in current is very small
4. Should not exceed the working temperature limit

Application of Photodiode:

1. Used in counters and switching circuits.
2. Used widely in optical communication system.
3. Used in detection of both visible and invisible light rays.
4. Used in safety electronics like fire and smoke detectors.
5. Used as photo sensors in camera.
6. Used in medical applications like pulsed oximeters, in instruments which analyze sample.

SOLAR CELL:

A **solar cell** (also known as a photovoltaic cell or PV cell) is defined as an electrical device that converts light energy into electrical energy through the **photovoltaic effect**. A solar cell is basically a p-n junction diode. In general, the photovoltaic effect means 'the generation of a potential difference at the junction of two different materials in response to visible or other radiation'.

Individual solar cells can be combined to form modules commonly known as solar panels. The common single junction silicon solar cell can produce a maximum open-circuit

voltage of approximately 0.5 to 0.6 volts. By itself this isn't much – but remembers these solar cells are tiny. When combined into a large solar panel, considerable amounts of renewable energy can be generated.

Construction:

A solar cell is basically a junction diode, although its construction it is little bit different from conventional p-n junction diodes. A very thin layer of p-type semiconductor is grown on a relatively thicker n-type semiconductor. Then we apply metal finger electrodes on the top of the p-type semiconductor layer.
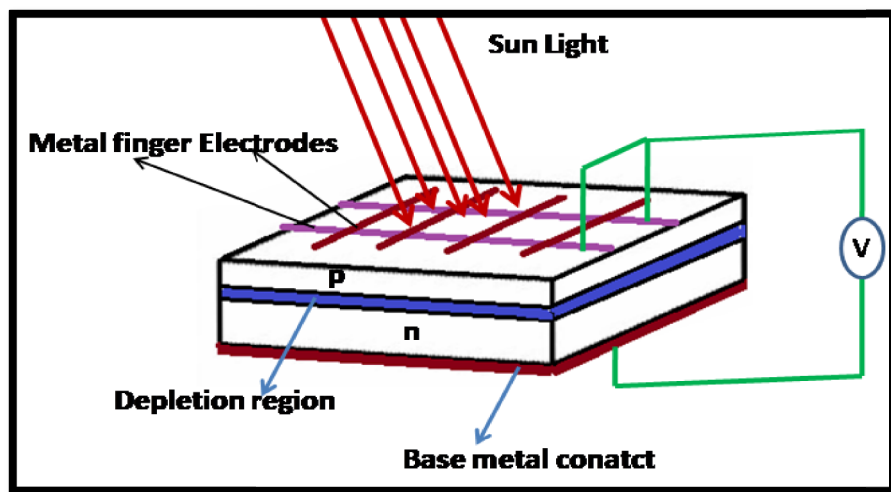

These electrodes do not obstruct light to reach the thin p-type layer. Just below the p-type layer there is a p-n junction. We also provide a current collecting electrode at the bottom of the n-type layer. We encapsulate the entire assembly by thin glass to protect the **solar cell** from any mechanical shock.

Working:

When a p – n junction diode is exposed to light, the photons are absorbed and electron pairs are generated in both in both the p – side and n – side of the junction, as shown in the figure below. The electrons and holes that are produced over a small distance from the junction reach the space charge region X by diffusion. The electron – hole pairs are then separated by the strong barrier field that exists across the region X. The electrons in the p – side slide down the barrier potential to move to the n – side while the holes in the n – side while the holes in the n – side move towards the p – side.

When the p – n junction diode is open circuited, the accumulation of electrons and holes on the two sides of the junction gives rise to an open – circuit voltage V0. If a load resistance is

connected across the diode, a current will flow in the circuit. The maximum current, called the short – circuit current is obtained when an electric short is connected across the diode terminals. Note that the current flows as long as the diode is exposed to sunlight and the magnitude of the current is proportional to the light intensity. Solar cells are used extensively in satellites and space vehicles as most important long duration power supply. Solar cells are constructed with Silicon, Germanium, Gallium arsenide, Cadmium sulphide and with many other semiconductors, and in various device configurations. The following diagram shows the I-V characteristics of the solar cell.
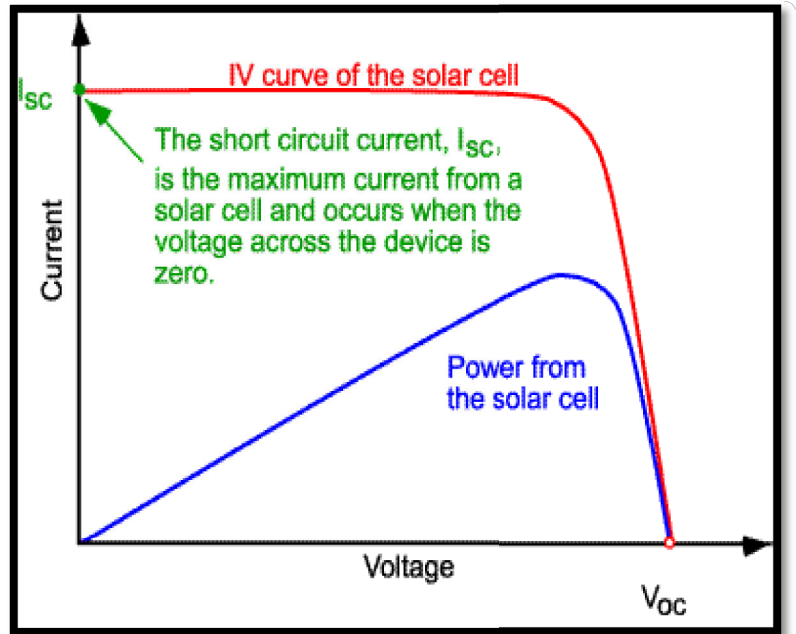
Advantages of Solar Cell	Disadvantages of Solar Cell
1. No pollution associated with it.	1. It has high cost of installation.
2. It must last for a long time.	2. It has low efficiency.
3. No maintenance cost.	3. During cloudy day, the energy cannot be produced and also at night we will not get solar energy

Uses of Solar Generation Systems

1. It may be used to charge batteries.
2. Used in light meters.
3. It is used to power calculators and wrist watches.
4. It can be used in spacecraft to provide electrical energy.

MODULE-IV: LASER AND FIBER OPTICS

A. LASERS

Introduction:

LASER is a device that emits light through a process of optical amplification based on the Quantum effect called stimulated emission of electromagnetic radiation. LASER is an acronym for Light Amplification by Stimulated Emission of Radiation.

Distinguish between conventional light and Laser light:

Conventional light	Laser light
1. It emits photons in all directions with wide range of wavelengths.	1. It emits photons in a narrow, well-defined directional beam.
2. These are incoherent (No fixed phase among the photons emitted by the source).	2. These are highly coherent (constant phase relationship among the photons)
3. May be multi-wavelength	3. Single wavelength or color.
4. Intensity is low	4. Intensity is very high.
5. Ex.:- Electric Bulb, candle, etc.,	5. Ex.:- Ruby Laser, He-Ne Laser, etc.,

Characteristics of Laser light

Like ordinary light laser light is electromagnetic in nature. However, there are few characteristics not processed by the normal light. Some of the main characteristics of laser beam are mentioned below:

1. **High Monochromaticity,**
2. **Extremely Directional,**
3. **Extremely Coherence,**
4. **High Intensity and brightness**.

1. **Monochromaticity**: The light emitted from a laser is highly *monochromatic*, that is, it is of one unique wavelength (color). The light from a laser typically comes from one atomic transition with a **single precise wavelength**. So the laser light has small broadening width ($\Delta\lambda$) and is almost the purest monochromatic light.
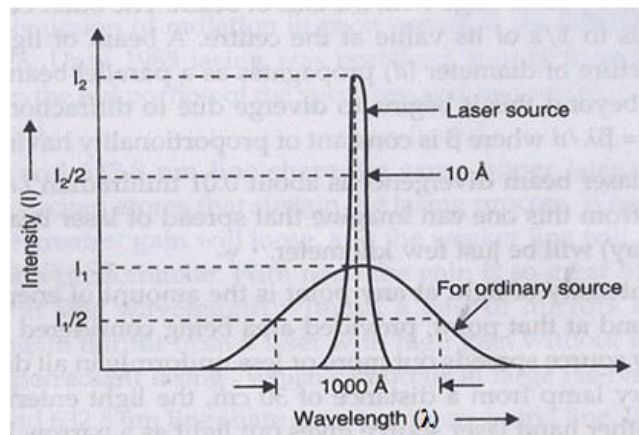

For Ordinary white light, $\Delta\lambda$= 300 nm, For monochromatic light, $\Delta\lambda$= 5-10 nm
For Laser light, $\Delta\lambda$= 0.2 nm

2. **Directionality:** We know that when light travels then it tends to spread out, this spreading of light is called divergence and the angle at which the light spreads is called angle divergence (ϕ). Generally divergence is more in ordinary light and lesser in laser light.

Usually a laser generates less than one mille radian (0.001) divergent beam. This means that a beam from the laser will spread to less than 1cm diameter at a distance of 10 m from the laser. The laser beam divergence is illustrated in the adjacent figure.

divergence, $\phi = \frac{r_2 - r_1}{d_2 - d_1}$
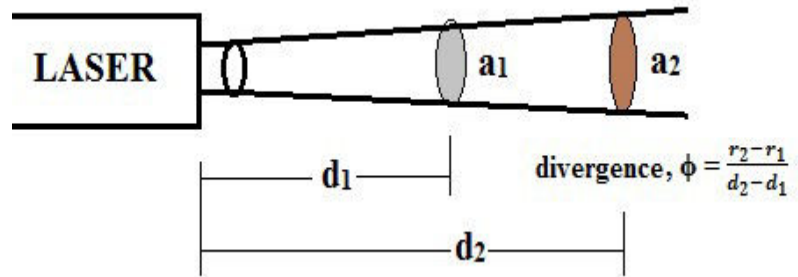
3. **Coherence:** The light coming from the Laser is said to be coherent, which means the group of photons (Laser beam) are in phase in space (Spatial) and time (Temporal). Laser light is much more coherent than ordinary light.
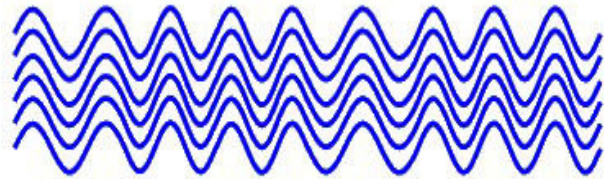

Figure: Coherent Light Waves

4. **Intensity and Brightness:** Intensity of a wave is defines as energy per unit time flowing through a unit normal area. In Laser light, energy is concentrated in small region of space with small wavelengths with greater intensity. A laser beam has brightness many times in magnitude greater than that of conventional sources due to high directional property of laser beam.

> **Therefore, Laser light is an extra ordinary light emitted under stimulated and amplified conditions, so that the beam is characterized by high intensity, specific directionality, high monochromaticity and high degree of coherence.**

ABSORPTION, SPONTANEOUS AND STIMULATED EMISSION:

To describe the phenomenon of Absorption, spontaneous or stimulated emission, let us consider two energy levels, 1 and 2, of some atom or molecule of a given material, their energies being E_1 and E_2 ($E_1 < E_2$).

i. **Absorption:** Let us now assume that the atom is initially lying in level 1 (Adj. Fig.). If this is the ground level, the atom will remain in this level unless some external stimulus (photons) is applied to it. We shall assume that, a photon of frequency v is incident on the material. In this case there is a finite probability that the atom will be raised to level 2. The energy difference $E_2 - E_1$ required by the atom to
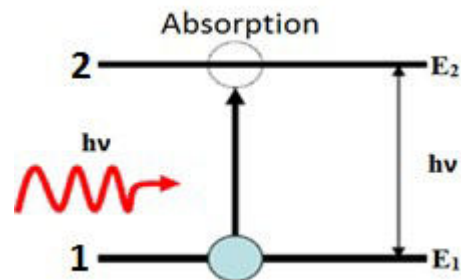
undergo the transition is obtained from the energy of the incident photon. This is the *absorption* process.

ii. **Spontaneous emission:** Let us now assume that the atom is initially in level 2. Since $E_2 > E_1$, the atom will tend to decay to level 1. The corresponding energy difference, E_2–E_1, must therefore be released by the atom. When this energy is delivered in the form of an electromagnetic wave called photon, the process will be called *spontaneous* (or *radiative*) *emission*.


Spontaneous emission
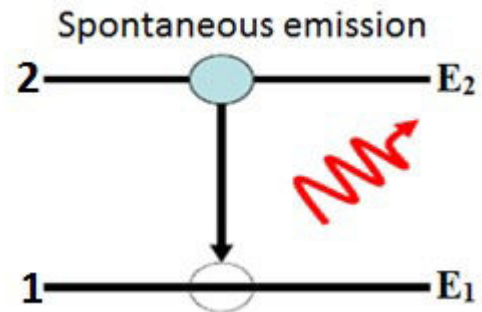
iii. **Stimulated emission:** Let us now suppose that the atom is found initially in level 2 and that a photon of frequency ν_o is incident on the material (Adjacent Fig.). Since this photon has the same frequency as the atomic frequency, there is a finite probability that this photon will force the atom to undergo the transition 2→1. In this case the energy difference E_2–E_1 is delivered in the form of a photon that adds to the incident one. This is the phenomenon of *stimulated emission*.

Stimulated emission
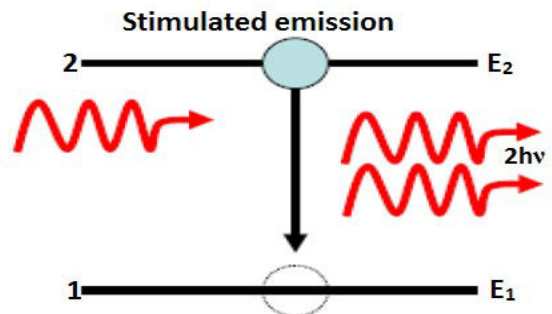
There is a fundamental difference between the spontaneous and stimulated emission processes. In the case of spontaneous emission, the atoms emits a photon that has no definite phase relation with that emitted by another atom. Furthermore, the photon can be emitted in any direction. In the case of stimulated emission, since the process is forced by the incident photon, the emission of any atom adds in phase to that of the incoming photon and along the same direction.

EINSTEIN'S COEFFICIENTS:

In 1916, **Albert Einstein** proposed that there are three processes occurring in the formation of an atomic spectral line. The three processes are referred to as **spontaneous emission, stimulated emission, and absorption**. With each is associated an Einstein coefficient which is a measure of the probability of that particular process occurring. Einstein considered the case of isotropic radiation of frequency v, and **energy density $\rho(v)$**.
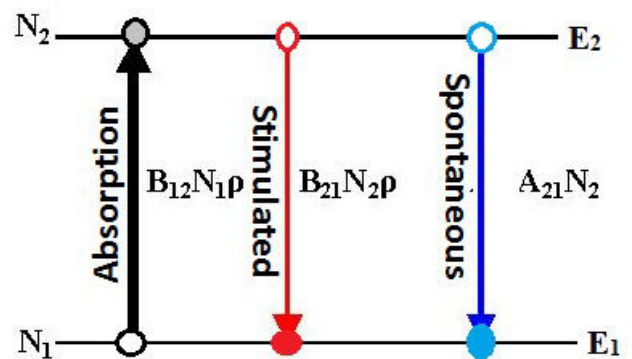
Let us consider N_1 and N_2 be the populations in the energy levels of energies E_1 and E_2 respectively in a system of atoms at a thermal equilibrium of temperature T.

Upward transition:

Absorption is the process by which a photon is absorbed by the atom, causing an electron to jump from a lower energy level E_1 to a higher one E_2. The process is described by the Einstein coefficient B_{12},

The absorption rate is directly proportional to N_1 and $\rho(v)$

$$\text{Therefore, } \textit{Rate of absorption} = B_{12}N_1\rho(v) \qquad \text{------------- (1)}$$

Here B_{12} is **Einstein's coefficient of absorption**.

Downward transition:

Spontaneous emission is the process by which an electron "spontaneously" (i.e. without any outside influence) decays from E_2 to E_1. The process is described by the Einstein coefficient A_{21}.

Spontaneous emission rate is directly proportional to N_2 only.

$$\text{Therefore, } \textit{Rate of spontaneous emission} = A_{21}N_2 \quad \text{------------- (2)}$$

Here, A_{21} is the **Einstein's coefficient of spontaneous emission**.

Stimulated emission is the process by which an atomic electron in the excited E_2 is interacting with a photon of certain frequency may drop to a lower energy level E_1, transferring its energy to that photon. A new photon created in this manner has the same phase, frequency and direction of travel as same as the incident photon. The process is described by the Einstein coefficient B_{21}.

Stimulated emission rate is directly proportional to N_2 and $\rho(v)$.

$$\text{Therefore, } \textit{Rate of stimulated emission} = B_{21}N_2\,\rho(v) \quad \text{---------- (3)}$$

Here, B_{21} is the **Einstein's coefficient of stimulated emission**.

Consider an ideal material with only two non-degenerate energy levels, at thermal equilibrium,

$$\textbf{Absorption = spontaneous emission + Stimulated emission} \qquad \text{------ (4)}$$

$$\text{i.e. } B_{12}N_1\rho(v) = A_{21}N_2 + B_{12}N_2\rho(v) \qquad \text{------- (5)}$$

$$B_{12}N_1\rho(v) - B_{12}N_2\rho(v) = A_{21}N_2$$

$$[B_{12}N_1 - B_{21}N_2]\,\rho(v) = A_{21}N_2$$

$$\rho(v) = \frac{A_{21}N_2}{[B_{12}N_1 - B_{21}N_2]} \quad \text{or} \quad \rho(v) = \frac{A_{21}N_2}{B_{21}N_2\left\{\left[\frac{N_1}{N_2}\left(\frac{B_{12}}{B_{21}}\right)\right]-1\right\}}$$

$$\text{or} \quad \rho(v) = \frac{A_{21}}{B_{21}}\frac{1}{\left\{\left[\frac{N_1}{N_2}\left(\frac{B_{12}}{B_{21}}\right)\right]-1\right\}} \quad\quad\quad \text{-------------- (6)}$$

But, in thermal equilibrium, the Boltzmann's distribution law applies, so,

$$N_1 = N_0\exp\left(\frac{-E_1}{k_BT}\right) \text{ and } N_2 = N_0\exp\left(\frac{-E_2}{k_BT}\right)$$

where, N_o is population in ground state and k_B is the Boltzmann's constant.

Therefore, $\frac{N_1}{N_2} = \exp\left[\frac{E_2-E_1}{k_BT}\right] = \exp\left[\frac{hv}{k_BT}\right]$ (Since, $E_2-E_1 = hv$) -------- (7)

Substituting eq. (7) in eq. (6), we get

$$\rho(v) = \frac{A_{21}}{B_{21}}\frac{1}{\left\{\left[\exp\left(\frac{hv}{k_BT}\right)\left(\frac{B_{12}}{B_{21}}\right)\right]-1\right\}} \quad\quad\quad \text{------------- (8)}$$

According to Plank's law of black body radiation at temperature T we have the energy density $\rho(v)$ at frequency v is

$$\rho(v) = \frac{8\pi hv^3}{c^3}\frac{1}{\left\{\exp\left(\frac{hv}{k_BT}\right)-1\right\}} \quad\quad\quad \text{------------- (9)}$$

Comparing eq. (8) and eq. (9), we get

$$\boxed{\frac{A_{21}}{B_{21}} = \frac{8\pi hv^3}{c^3}} \quad \text{----- (10)} \quad\quad\quad \boxed{\frac{B_{12}}{B_{21}} = 1} \text{ -------- (11)}$$

Equations (10) and (11) show the relations between Einstein's coefficients B_{12}, B_{21} and A_{21}.

From the above relation eq. (10) shows that the ration of rate of spontaneous (A_{21}) to rate of stimulated emission (B_{21}) is **proportional to cube of the frequency** of the incident radiation. The eq. (11) shows that the rate of absorption is equal to the rate of stimulated emission, when the system in equilibrium.

POPULATION INVERSION:
At thermal equilibrium, the no of atoms in level 2 is less than the no of atoms in level 1 ($N_2 < N_1$). **The process of making of higher population level 2 than the population in level 1 is**

known as population inversion ($N_2 > N_1$) This is highly non equilibrium situation. When the population inversion exists, rate of stimulated emission is greater than rate absorption occurs and the light passing through the material is amplified.
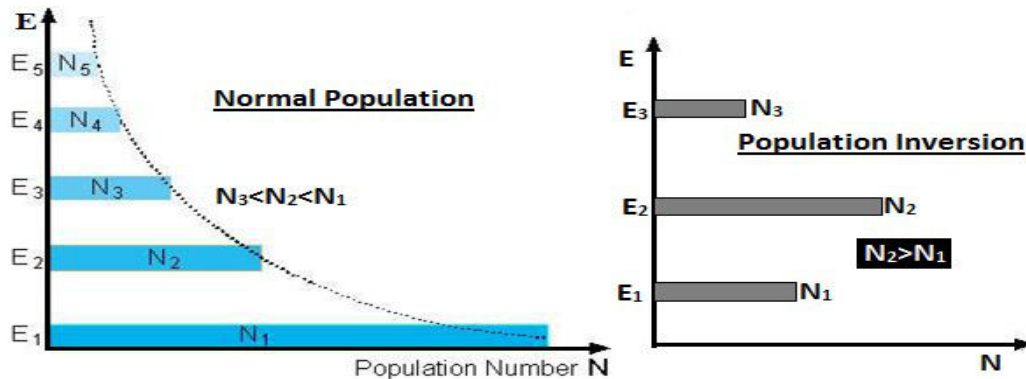
LASER PUMPING:

The process of achieving the population inversion by sending the atoms from lower level to higher levels through supply of energy to the lower level atoms is called **laser pumping**. The supply of energy is usually provided in the form of **light (Optical) or electric current (Electrical)**. But, more exotic sources such as chemical or nuclear reactions can also be used.

i. **Optical Pumping:** Optical pumping is a process in which light is used to raise or pump electrons from lower energy level in an atom or molecule to higher one. It is commonly used in solid state lasers (Ex.: Ruby Laser, etc.).
ii. **Electrical Pumping:** Electric discharge is common in gas lasers. Gas lasers have very narrow absorption band, pumping then in any flash lamp is not possible. In most of the cases population inversion is created by means of electric discharge. (Ex.: He-Ne Laser etc.)
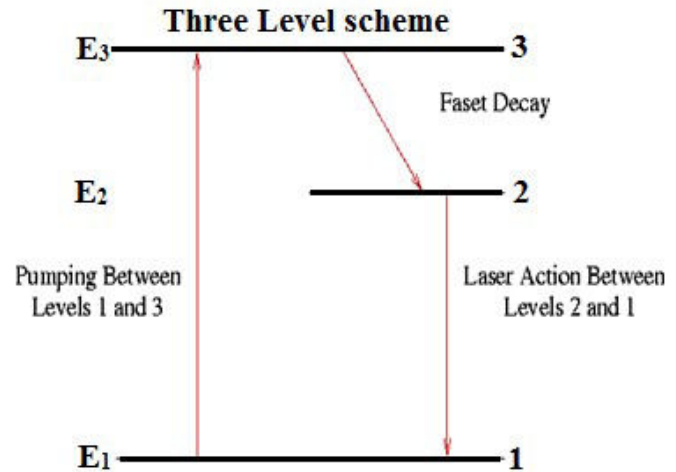
META-STABLE STATE:

Meta-stable state is an excited state of an atom or other system with a longer lifetime than the other excited states. However, it has a shorter lifetime than the stable ground state. Atoms in the meta-stable state remain excited for a considerable time in the order of 10^{-6} to 10^{-3}s. During meta-stable state, all the parameters associated with state hold stationary values. A large number of excited atoms are accumulated in the meta-stable state.

The population of meta-stable state can exceed the population at a lower level thereby establishing population inversion in a lasing medium. Population inversion could not be created without a meta-stable state.

THREE LEVEL SCHEME:

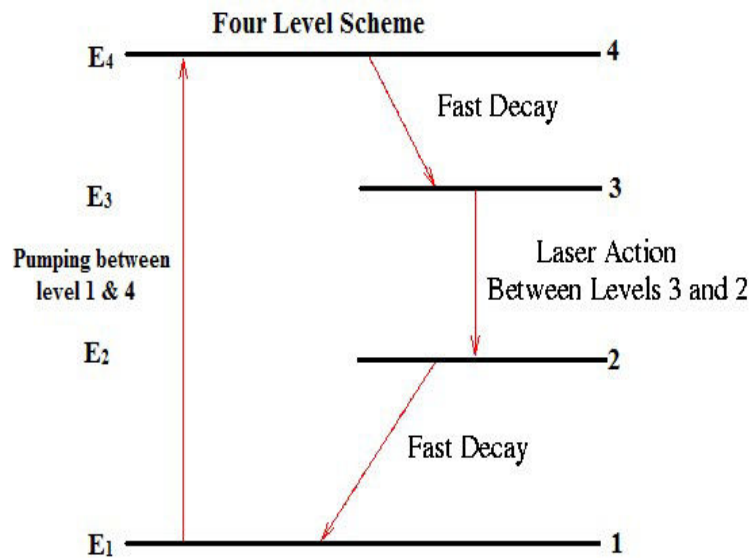We have seen that the key to laser action is to obtain a population inversion between two levels of energy E_1 and E_2 with $E_2 > E_1$, so that more atoms are in the level 2 than in the level 1. In the three-level lasers figure, we look for three levels in an atom such that $E_3 > E_2 > E_1$, with a fast decay between levels 3 and 2 and a slow decay between 2 and 1. Incident radiation of angular

frequency $v_{31} = (E_3 - E_1)/\hbar$ is used to raise as many atoms as possible from the level 1 to level 3. Note that a population inversion cannot be obtained between levels 3 and 1, because when the number of atoms N_3 in level 3 equals the number N_1 in level 1, absorption will be balanced by stimulated emission and the material will become transparent at the frequency v_{31}. If level 3 decays rapidly to level 2, a population inversion can be obtained between levels 2 and 1. An example is the ruby laser.

Three Level scheme

E3 ——————————— 3

Faset Decay

E2 ———————— 2

Pumping Between Levels 1 and 3

Laser Action Between Levels 2 and 1

E1 ——————————— 1

FOUR LEVEL SCHEME:

Except in special cases, such as the ruby, it is difficult to produce a population inversion between a ground state and an excited state, because initially all the atoms are likely to be in the ground state, and we have to get more than half the atoms into level 2 before a population inversion can be achieved. An easier approach is to use a four-level system in the adjacent figure and attempt to create a population

Four Level Scheme

E4 ———————————— 4

Fast Decay

E3 ———————— 3

Pumping between level 1 & 4

Laser Action Between Levels 3 and 2

E2 ———————— 2

Fast Decay

E1 ———————————— 1

inversion between two excited levels. We start with all the atoms in the ground state1, and none in the excited states 2, 3 and 4 ($E_2 < E_3 < E_4$). Level 4 is chosen so that it has a fast decay to level 3, and pumping between levels 1 and 4 immediately produces a population inversion between levels 3 and 2. As level 2 begins to fill up by stimulated emission at the frequency ($E_3 - E_2$)/h, the population inversion will decrease. To minimize this, level 2 is chosen so that it has a fast decay to the ground state.

REQUIREMENTS FOR THE LASER SYSTEM:

The basic requirements of the Laser system are shown in below figure. It consists of three basic parts as given below:

    i. Pumping System
    ii. Active medium
    iii. Optical Resonator

i. **Pumping System:** Pumping source is the basic energy source for a laser. It gives energy to various atoms of laser medium & excites them, so that population inversion can take place and it is maintained with time. The excitation of atom occurs directly or through atom or atom collision.

There are various types of pumping systems depending on the nature of the active medium.

ii. **Active Medium:** Active medium is the heart of the laser system and is responcible for producing population inversion (gain) and subsequent generation of laser. It can be a solid, liquid, semiconductor or gas medium and can be pumped to higher energy state.
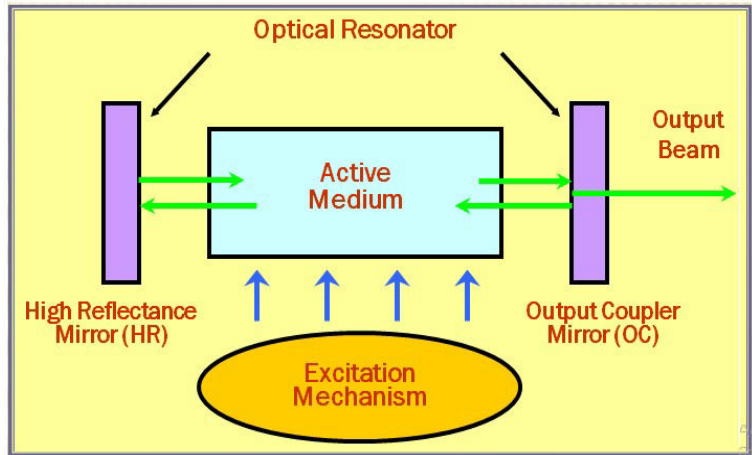
iii. **Optical resonator:** It plays a very important role in the production of laser beam from the laser system. It is a setup used to obtain amplification of slimulated photons, by oscillating them back and forth between two extreme limits. It consists of two plane or concave mirrors placed co-axially. One mirror is fully reflecting and other is partially reflecting.
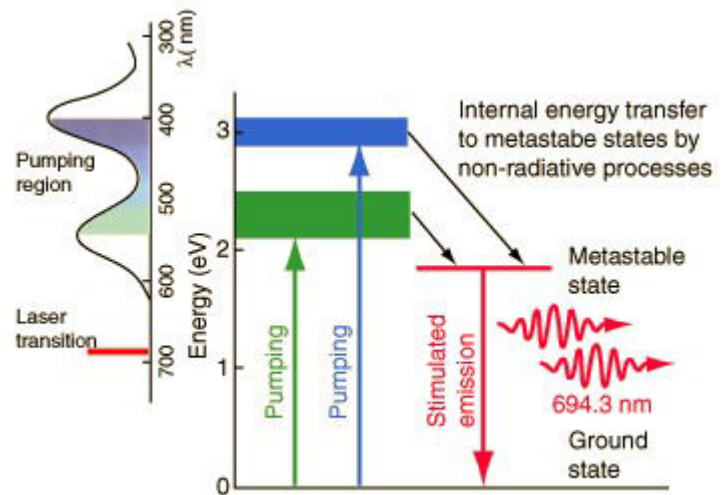
RUBY LASER:

The first working laser was built in 1960 by T.H. Maiman using a Ruby crystal and so called the Ruby Laser. This is also called **solid state laser or three level Laser**. Ruby belongs to the family of gems consisting of Al_2O_3 with various types of impurities. For example in Ruby laser used the pink ruby contains 0.05% Cr atoms (Al_2O_3+0.05%Cr_2O_3). The schematic diagram of the Ruby laser is as follows:
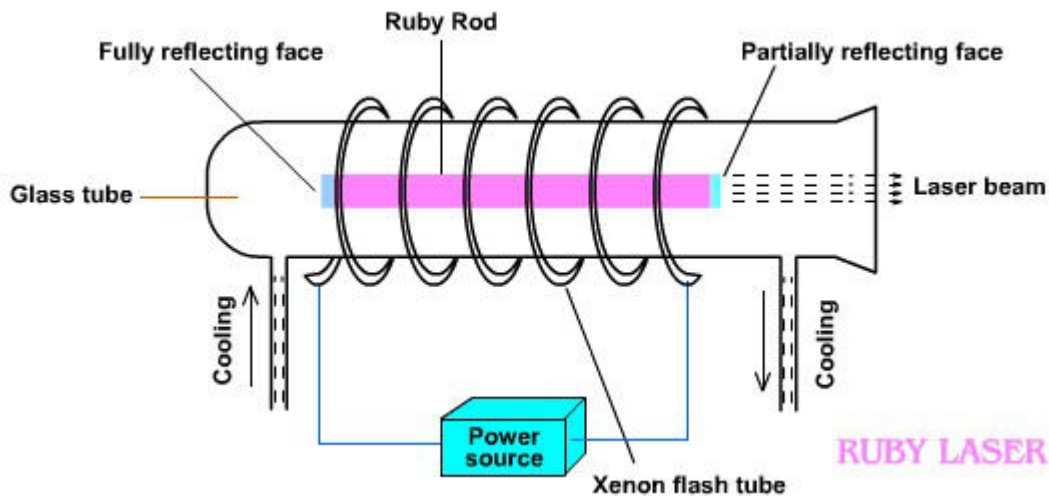
Construction: The Ruby laser consists of a ruby rod whose length is few centimeter and diameter is 0.5 cm. This is made up of with chromium (Cr^{3+}) doped Al_2O_3 material. Both the ends of the ruby rod are silvered such that one end is fully reflecting and the other end is partially reflecting. The ruby rod is surrounded by helical xenon flash lamp tube which provides the optical pumping to raise Cr^{+3} ions to upper energy level. The light from the xenon flash tube was focused by the cylindrical cavity onto the ruby rod, thereby exciting the chromium atoms which were responsible for the laser action.

Working: The ruby laser is a three-Level system (Adjacent Fig.) since only three energy levels are involved in the process of stimulated emission. The depopulation of the ground state for population inversion is achieved by exciting the atoms of the ruby crystal with intense light from a xenon flash lamp. Thus the atoms are excited from the ground state (level 1) to an upper state (level 3) by means of absorption. From the energy level 3, the atoms are transferred to energy level 2 without emitting radiation (non-radiative transfer). The energy level 2 is called met stable level since the atoms stay at this level for a longer interval of time. Finally, the atoms return to the ground state from the meta-stable level through the process of stimulated emission giving rise to an intense laser light at **6943A⁰**. The laser beam comes out in the form of a pulse of very short duration (about a millisecond).

Applications: They are still used in a number of applications where short pulses of red light are required.

i. Holography's around the world produce holographic portraits with ruby lasers, in sizes up to a meter squared.
ii. Many non-destructive testing labs use ruby lasers to create holograms of large objects such as aircraft tires to look for weaknesses in the lining.
iii. Ruby lasers were used extensively in **tattoo and hair removal**.

Drawbacks:

i. The laser requires high pumping power because the laser transition terminates at the ground state and more than half of ground state atoms must be pumped to higher state to achieve population inversion.
ii. The efficiency of ruby laser is very low because only green component of the pumping light is used while the rest of components are left unused

iii. The laser output is not continues but occurs in the form of pulses of microseconds duration.

HELIUM-NEON (He-Ne) LASER:

The He-Ne laser active medium consists of two gases which do not interact form a molecule. Therefore He-ne laser is one type of **atomic gas laser** and also called as **Four level laser**.

Construction:
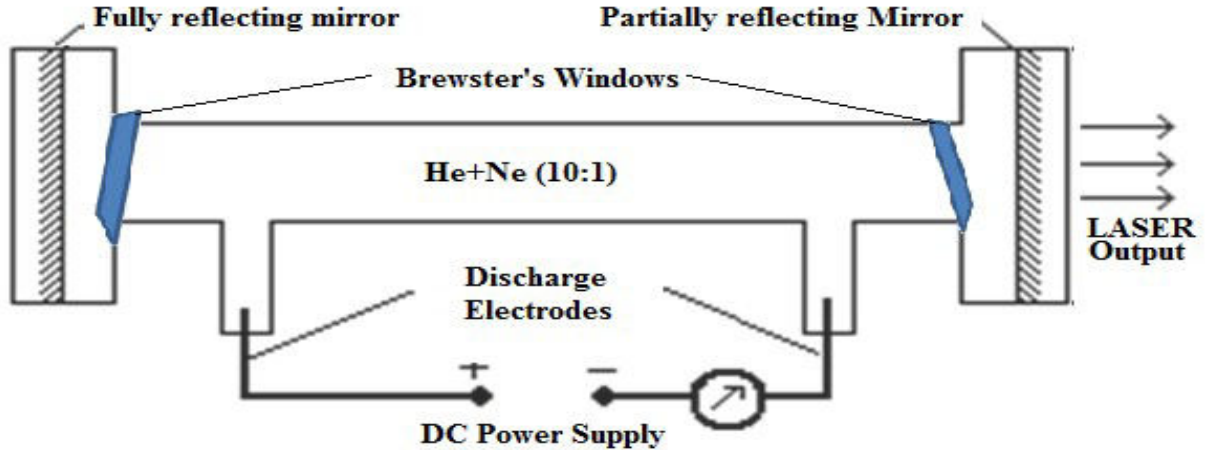
He - Ne gas laser consists of a gas discharge tube of length 30cm and diameter of 1.5cm. The tube is made up of quartz and is filled with a mixture of Neon under a pressure of 0.1mm of Hg. The Helium under the pressure of 1mm of Hg, the ratio of He-Ne mixture of about 10:1, hence the no. of helium atoms are greater than neon atoms. The mixtures is enclosed between a set of parallel mirrors forming a resonating cavity, one of the mirrors is completely reflecting and the other partially reflecting in order to amplify the output laser beam.

Working:

In the He-Ne laser, the light produced by atomic transitions within the Neon atom. The Helium does not directly produce laser light but it acts as a buffer gas, this purpose of which is to assist/help the atoms of the other gas to produce lasing action.

The active energy levels of He and Neon atoms are show in adjacent figure. In helium there are three active energy levels named as F_1, F_2 and F_3 where as in Neon, there are six active energy levels named as E_1, E_2, E_3, E_4, E_5, and E_6. In Helium, the metastable states are F_2 and F_3, where as in Neon, E_4 and E_6. When a discharge is passed through the gaseous mixture electrons are accelerated in the tube these
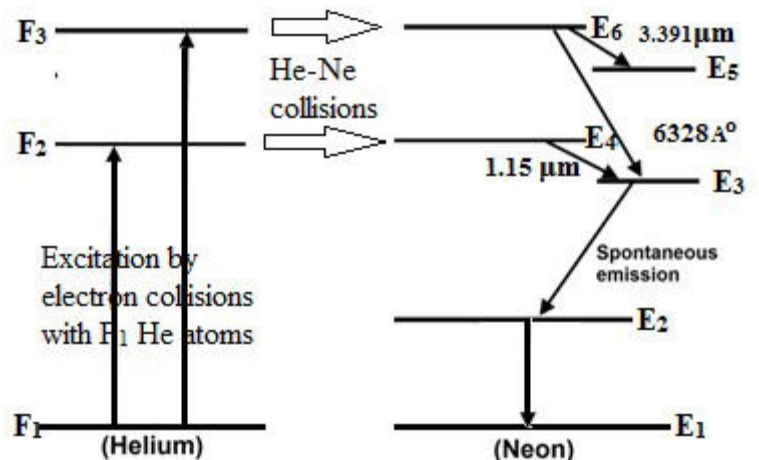
accelerated electrons collide with the helium atoms and excite them to higher energy levels (F_2 and F_3) since the levels are metastable energy levels, He atoms spend sufficiently long time. Inelastic collision of the excited helium atoms (F_2 and F_3) with the ground state (E_1) Neon atoms results in transfer of energy to the neon atoms, exciting them into E_4 and E_6 states. This is due to the coincidence of energy levels between the helium and neon atoms. Since E_4 and E_6 are metastable states, hence the population inversion takes place between the E_6 and E_4 with respect to E_5 and E_3. The stimulated emission takes place between $E_6 \rightarrow E_3$ gives a laser light of wave length **6328A⁰** and the stimulated emission between $E_6 \rightarrow E_5$ gives a laser light wavelength of **3.39μm**. Another stimulated emission between $E_4 \rightarrow E_3$ gives a laser light wavelength of **1.15μm**. The neon atoms undergo spontaneous emission from $E_3 \rightarrow E_2$ and $E_5 \rightarrow E_2$. Finally the neon atoms are returned to the ground state E_1 from E_2 by non-radiative diffusion and collision process.

After arriving to the ground state, once again the neon atoms are raised to $E6$ & $E4$ by excited helium atoms thus we can get **continuous wave** output from He-Ne laser. But some optical elements placed insides the laser systems are used to absorb the infrared laser wavelengths 3.39μm and1.15μm. Hence the output of He-Ne laser contains only a single wavelength of 6328A° and the output power is about few milliwatts.

Applications:

i.   The Narrow red beam of He-Ne laser is used in supermarkets to read bar codes (Bar-code scanners).
ii.   The He-Ne Laser is used in Holography in producing the 3D images of objects.
iii.   He-Ne lasers have many industrial and scientific uses, and are often used in laboratory demonstrations of optics.

SEMICONDUCTOR (Diode) LASER:

Semiconductor lasers also known as *quantum well lasers* are smallest, cheapest, can be produced in mass, and are easily scalable. They are basically p-n junction diode, which produces light of certain wavelength by recombination of charge carrier when forward biased, very similar to the light-emitting diodes (LEDs). LEDs possess spontaneous emission, while laser diodes emit radiation by stimulated emission.

Principle: In the case direct band gap semiconductors there is a large possibility for direct recombination of hole and electron emitting a photon. GaAs is a direct band gap (1.44 eV) semiconductor and hence it is used to make lasers and light emitting diodes. The wave lengths of the emitted light depend on the band hap of the material.

There are two types of semiconductor diode lasers
    i.   Homo junction laser
    ii.   Hetero- Junction laser.

If a p-n junction is formed in a single crystalline material, then it is called as homo-junction laser.

Example: **Gallium Arsenide (Ga-As) Laser**

If p-n junction is formed with different semiconducting materials, then it is known as Hetero- Junction laser. It is also called modern laser diode.

Example: Hetero- Junction Laser can be formed between **Ga-As and Ga-Al-As.**

Construction:

i. Homo-Junction:

The P^+ and N^+ regions of the diode are obtained by heavily doped p-and n-regions of GaAs. The thickness of the p-n junction layer is very narrow at the junction, the side walls are well polished and parallel to each other. Since the refractive index of GaAs is high, the reflectance at the material air interface is sufficiently large so that the external mirrors are not necessary to produce multiple reflections. The p-n junction is forward biased by connecting positive terminal to p-type and negative terminal to n-type.
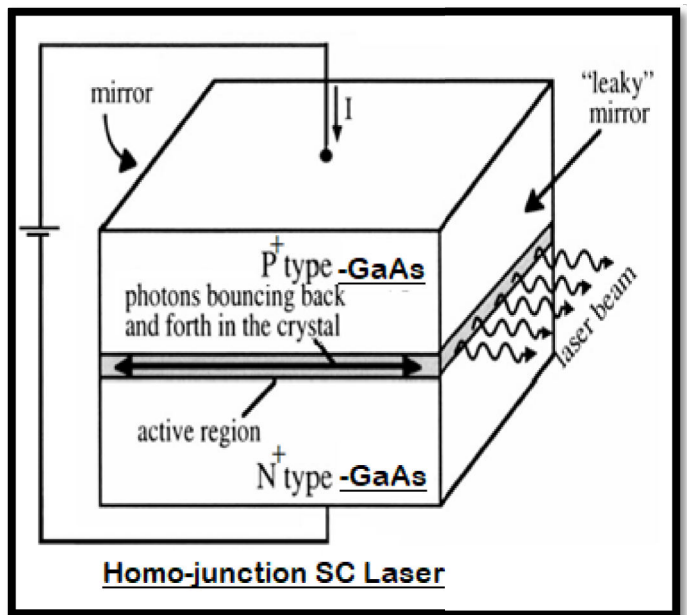

Homo-junction SC Laser

ii. Hetero-junction:

This laser consists of five layers as shown in the above figure. A layer of Ga-As p–type (3rd layer) will act as the active region. This layer is sand-witched between two layers having wider band gap viz. GaAlAs p–type (2nd layer) and **GaAlAs** n-type (4th layer). The end faces of the junctions of 3rd and 4th layer are well polished and parallel to each other. They act as an optical resonator.
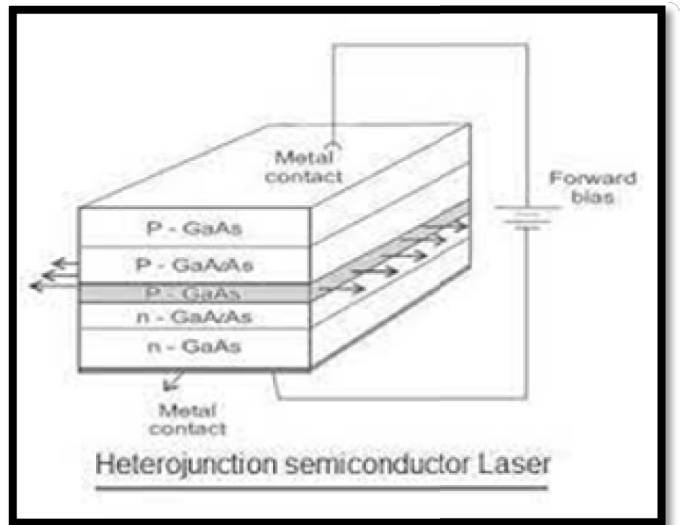
Heterojunction semiconductor Laser

Working:

The population inversion can be obtained by injecting electrons and holes in to the junction from the n-region and p-region by means of forward bias voltage. When the forward bias is not connected, no electrons and holes present in the depletion region. When small forward bias voltage is given to the p-n junction then small number of electrons and holes will injected into the depletion region from respective regions. When relatively a large current of the order of 10^4 A/cm^2 is passed through the junction then large number of electrons and holes will be injected into the depletion region as shown in above figure. Then the direct recombination processes take place between holes and electrons in the depletion region and release the photons. Further the emitted photons increase the rate of recombination. Thus more number of photons produced having same phase and frequency of the induced photons.
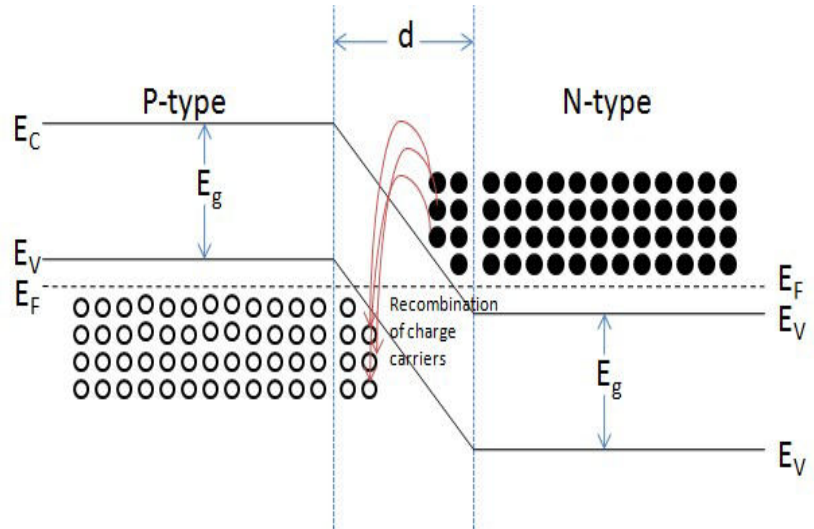
The wavelength of the emitted radiation depends on the energy band gap of the semiconductor material. The energy gap of the GaAs Semiconductor is 1.44 eV then it emits laser light of wavelength ~8600 A°.

$$\lambda = \frac{hc}{E_g} = \frac{6.625 \times 10^{-34} \times 3 \times 10^8}{1.44} = 8626 \text{ A}°$$

The efficiency of the laser emission increases when a cooling arrangement provided to the laser diodes.

APPLICATIONS OF LASER: There are many scientific, military, medical, industrial and commercial laser applications which have been developed since the invention of the laser.

i. **Laser Drilling:** Laser drilling is one of the few techniques for producing holes in hard materials with desired diameter. The diameter of these holes as small as 5 microns.
Laser-drilling is used in many applications, including the oil gallery of some engine blocks, aerospace turbine-engine cooling holes, laser fusion components and printed circuit board micro-vias.

ii. **Laser Welding and Cutting:** The highly collimated beam of a Laser can be further focused to a microscopic dot of extremely high energy density for welding and cutting. The automobile industry makes extensive use of CO_2 laser with powers up to several KW for computer controlled welding on auto assembly lines.

iii. **Lasers in Data Storage:** The reading and writing (Store) of the data on a compact disc (CD or DVD) is done with semiconductor laser.

iv. Lasers in scientific research:
   a) Lasers are used to clean delicate pieces of art, develop hidden finger prints
   b) Laser are used in the fields of 3D photography called holography
   c) Using lasers the internal structure of micro organisms and cells are studied very accurately
   d) Lasers are used to produce certain chemical reactions.

v. Laser in Medicine:
   a) The heating action of a laser bean used to remove diseased body tissue
   b) Lasers are used for elimination of moles and tumours, which are developing in the skin tissue.
   c) Laser beam is used to correct the retinal detachment by eye specialist.

vi. It is used in biomedical science.

vii. It is used in 3D photography.

viii. It can be used for measuring velocity of light, to study spectrum of matters, to study Raman effect.

ix. It is used in holography.

x. It is also used in military, like LIDAR.

xi. It is used to accelerate some chemical reactions.

**xii. It is also used in nuclear fusion reactions.**

B. FIBER OPTICS

STRUCTURE OF AN OPTICAL FIBER:

Optical Fiber: A very thin, long, flexible, transparent, cylindrical dielectric medium which guide the light signal propagated through it.

It consists of three parts:

1. The core
2. The cladding,
3. The outer jacket.

Core   Cladding   Sheath

- The core is the inner part of the fiber, which guides the light signal.
- The cladding surrounds the core completely.
- The refractive index of the core (n_1)is greater than the cladding(n_2) to satisfy the total internal reflection ($n_1 > n_2$)
- The outer jacket provides the mechanical protection to the fiber.
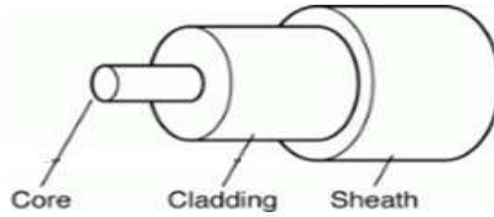
The core and cladding diameters are about few microns. Optical fibers are arranged in bundles called optical cables and used to transmit light signals over long distances.

PRINCIPLE AND WORKING OF OPTICAL FIBER:

Principle: The transmission of light in optical fiber is based on the principle of **Total Internal reflection**.

Let n_1 and n_2 be the refractive indices of core and cladding respectively such that $n_1 > n_2$. Let a light ray travelling from the medium of refractive index n_1 to the refractive index n_2 be incident with an angle of incidence θ and the angle of refraction r. By Snell's law,

$$n_1 \sin\theta = n_2 \sin r \qquad \text{---------- (1)}$$

The refracted ray bends towards the normal as the ray travels from rarer medium to denser medium. On the other hand, the refracted ray bends away from the normal as it travel from denser medium to rarer medium. In Optical fiber, the light ray travelled from core (denser) to cladding (rarer) medium, there is possibility of total internal reflection, if the angle of incidence is greater than the critical angle (θ_c).

Critical Angle: When a light ray moves from high refractive index (core) medium to low refractive index (cladding) medium and for a particular angle of incidence the refraction angle (r) is 90° then the angle of incidence is known as critical angle (θ_c).

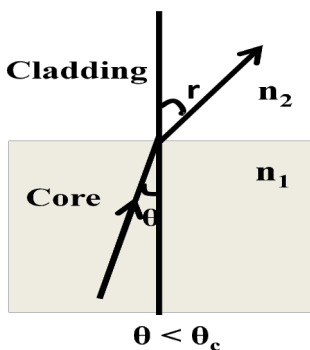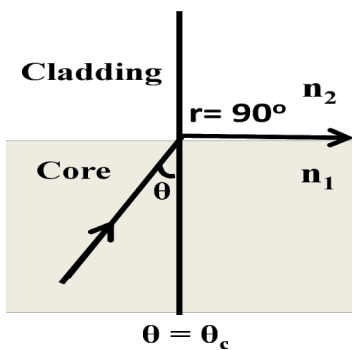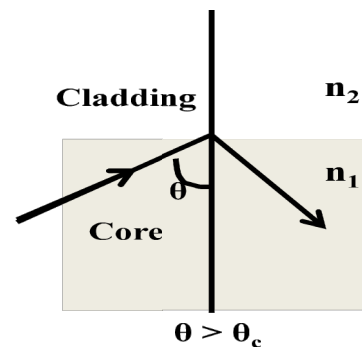|		
$\theta < \theta_c$	$\theta = \theta_c$	$\theta > \theta_c$
Fig.1	Fig.2	Fig.3

i.   When $\theta < \theta_c$, then the ray refracted into the second medium as shown in above fig.1.
ii. When $\theta = \theta_c$, then the ray travels along the interface of two media as shown fig. 2.
iii. When $\theta > \theta_c$, then the ray totally reflects into the same medium (core) as shown in fig.3.
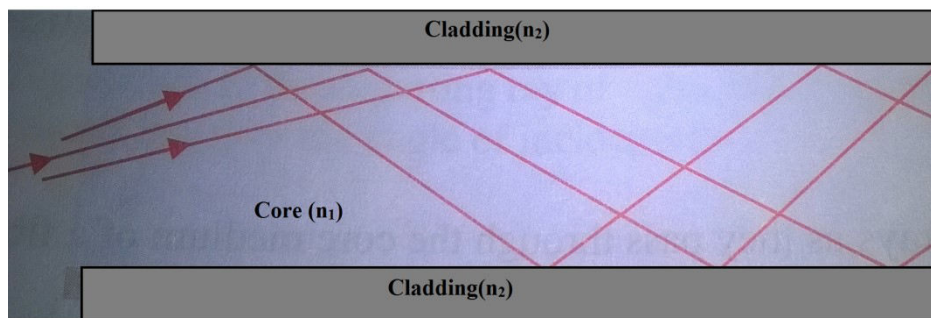The critical angle can calculated from Snell's law, suppose if $\theta = \theta_c$ then $r = 90°$, hence

$$n_1 \sin\theta_c = n_2 \sin 90°$$

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

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

This is known as critical angle of mediums of refractive indices n_1 and n_2 ($n_1 > n_2$).
The light signal once entered the fiber and takes total internal reflection with in the core as shown in fig.4, it will continue to propagate till the other end of the fiber.

ACCEPTANCE ANGLE:

The maximum angle of incidence to launch the beam at its one end to enable the entire light to pass through the core is called **acceptance angle**.

When we launch the light beam in to the fiber at its one end the entire light may not pass through the core and propagate. Only the rays which make the angle of incidence greater than critical angle undergo total internal reflection and propagate through the core and all other rays are lost. Let us consider a ray enters the core of refractive index n_1 from air medium of refractive index n_o with an angle of incidence α_i at the interface of air and core and incident at the interface of core and cladding with an angle of incidence θ as shown in below figure.
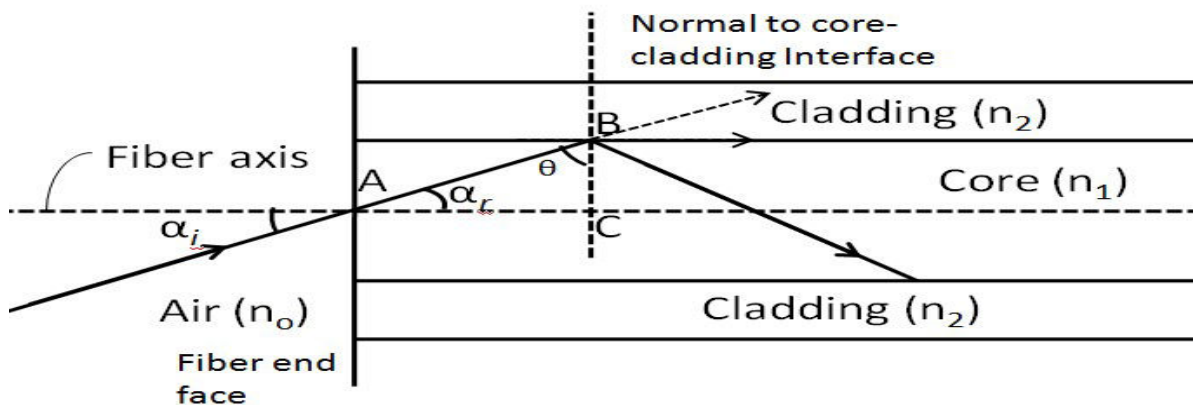
If the angle of incidence at the interface of air and core is maximum ($\alpha_i = \alpha_{max}$), the angle of incidence at the interface of core and cladding is equal to critical angle ($\theta = \theta_c$). If the angle α_i is less than the α_{max}, the angle of incidence at the interface of core and cladding would be greater than the critical angle ($\theta > \theta_c$), and further total internal reflections cascades through the fiber.

$$From\ fig.,\ \Delta ABC,$$
$$\alpha_r = 90 - \theta$$

$$At\ air - core\ interface, the\ Snell's\ law\ is,\quad \frac{\sin \alpha_i}{\sin \alpha_r} = \frac{n_1}{n_o}$$

$$\sin \alpha_i = \frac{n_1}{n_o} \sin \alpha_r$$
$$\sin \alpha_i = \frac{n_1}{n_o} \sin(90 - \theta)$$
$$\sin \alpha_i = \frac{n_1}{n_o} \cos \theta$$
$$\boldsymbol{\sin \alpha_{max}} = \frac{n_1}{n_o} \boldsymbol{\cos\ \theta_c}\ \text{(When } \alpha_i = \alpha_{max}, \theta = \theta_c)\qquad \text{----- (1)}$$

Snell's law at core and cladding interface gives,

At critical angle($\theta = \theta_c$), the angle of refraction is 90°

$$n_1 \sin \theta_c = n_2 \sin 90$$
$$\sin \theta_c = \frac{n_2}{n_1}$$

$$\cos \theta_c = \sqrt{1 - sin^2\theta_c} = \sqrt{1 - \frac{n_2^2}{n_1^2}}\qquad \text{------ (2)}$$

From Eq. (1), we get, $\boldsymbol{\sin \alpha_{max}} = \frac{n_1}{n_o}\sqrt{1 - \frac{n_2^2}{n_1^2}} = \frac{\sqrt{n_1^2 - n_2^2}}{n_o}$

Thus,

$$\boldsymbol{\sin \alpha_{max}} = \frac{\sqrt{n_1^2 - n_2^2}}{n_o}$$

$$\boldsymbol{\alpha_{max} = sin^{-1}\left(\frac{\sqrt{n_1^2 - n_2^2}}{n_o}\right)}\qquad \text{----- (3)}$$

This $\boldsymbol{\alpha_{max}}$ is known as **Acceptance angle.**

ACCEPTANCE CONE: The acceptance cone is derived by rotating the Acceptance Angle about the fiber axis. Light launched at the fiber end within this acceptance cone alone will be accepted and propagated to the other end of the fiber by total internal reflection. Larger acceptance angles make launching easier.
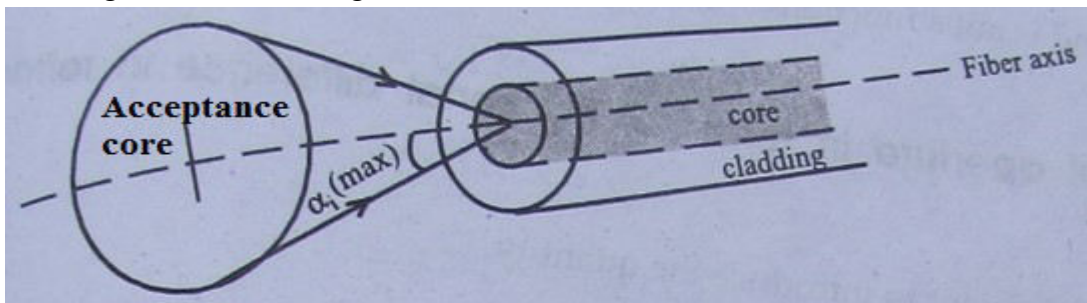
Numerical Aperture: The numerical aperture (*NA*) of an optical fiber is defined as sin of acceptance angle and is dimensionless number that characterizes the range (ability) of angles over which the system can accept light.

Therefore, $NA = \sin(\alpha_{max}) = \dfrac{\sqrt{n_1^2 - n_2^2}}{n_o}$ -------- (4)

If the refractive index of air is unity ($\mu_{air}=1$), thus the Eq. (4) reduces as,

$$NA = \sqrt{n_1^2 - n_2^2} = n_1\sqrt{2\Delta} \qquad \text{where } \Delta = \dfrac{n_1 - n_2}{n_1} \text{ called Fractional Index difference}$$
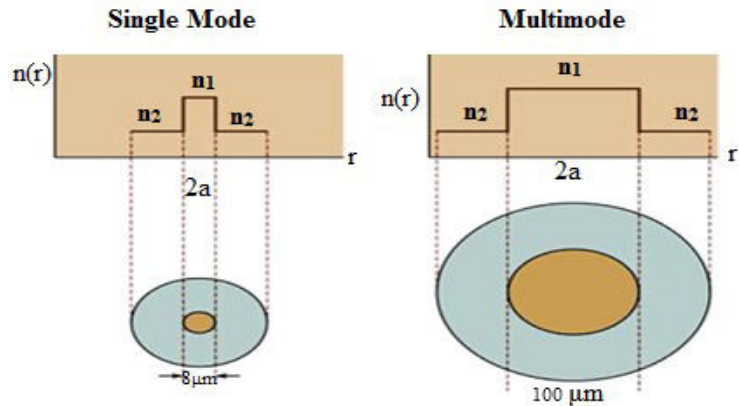
FIBER TYPES AND REFRACTIVE INDEX PROFILES:

Depending upon the refractive index profile of the core, optical fibers are classified into two types. They are:

      i. Step Index (SI) Fiber   ii. Graded Index (GI) fiber

i. Step Index (SI) Fiber: In step index fibers, the refractive index of the core is uniform throughout the medium and undergoes an abrupt (Step) change at the interface of core and cladding.

The light in the fiber propagates by bouncing back and forth from core-cladding interface. The step index fibers propagate both single and multimode signals within the fiber core. The light rays propagating through it are in the form of meridinal rays which will cross the fiber core axis during every reflection at the core – cladding boundary and are propagating in a zig – zag manner.

The variation of refractive index profile with radial distance *r* for the SI fiber is shown in the above figure.

$$n(r) = n_1 (\text{core})$$
$$= n_2 (\text{cladding})$$

The number of possible propagation modes in the core depends on the radius of the core and NA of the fiber and is given by V-number as

$$V = \dfrac{2\pi}{\lambda} a(NA) \quad \text{Where '}a\text{' is the radius of the core and NA-Numerical Aperture.}$$

The number of modes (paths) through the SI fiber is $= \dfrac{V^2}{2}$

The propagation modes in the SI Fiber in both the modes are shown below:

Single mode SI fiber

Multimode SI Fiber
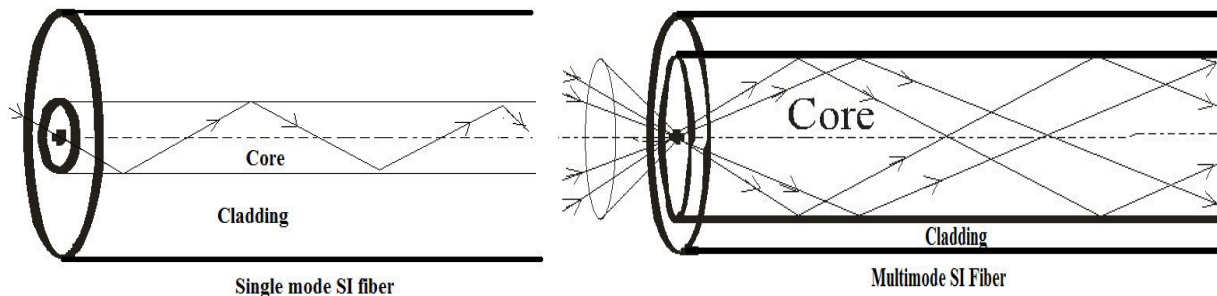
ii. **Graded Index (GI) Fiber:** In graded index fibers, the refractive index of the core varies gradually as a function of radial distance from the fiber center.

The refractive index of the core decreases as we move away from the centre. The refractive index of the core is made to vary in the form of parabolic manner such that the maximum refractive index is present at the centre of the core.

The variation of refractive index profile with radial distance r for the SI fiber is shown in the adjacent figure.

$$n(r) = n_1\sqrt{1 - 2\Delta\left(\frac{r}{a}\right)^{\alpha}}\ (0 \leq r \leq a)\ \text{(core)}$$

$$= n_2\ \text{(cladding)}$$

Here α is the grading parameter which decides the variation of RI in core.

$\alpha = 1$ for linear grading,

$\alpha = 2$ for **parabolic** grading

$\alpha = \infty$ for Step Index grading.

The transmitted light signals travel through the core medium in the form helical (sine waves) rays, which will not cross the fiber axis at any time.

The number of modes propagated through the GI Fiber depends on the radius of the core and NA of the fiber. Therefore,

The possible number of modes propagated through the GI Fiber is $= \frac{V^2}{4}$

The propagation modes in the GI Fiber are shown below:

Graded Index Fiber
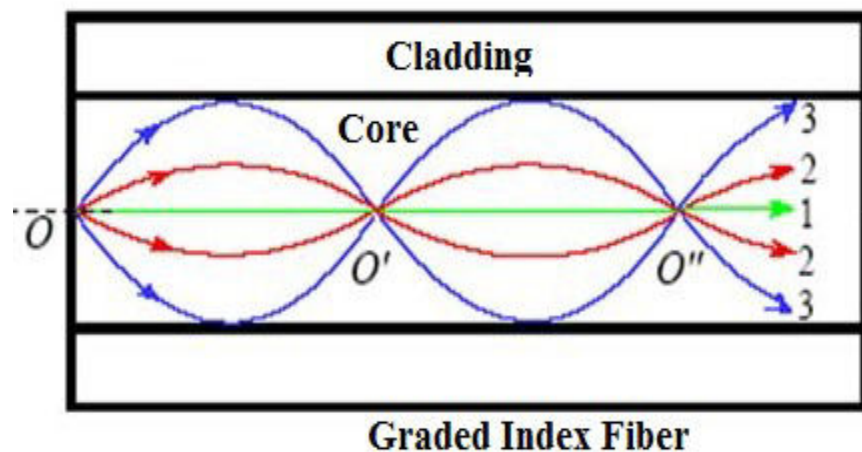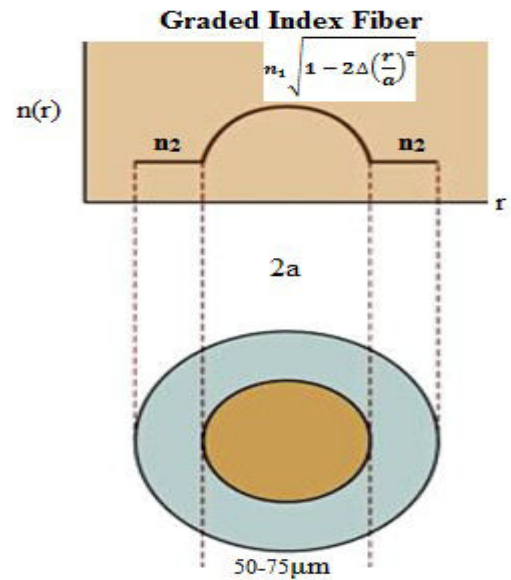
ATTENUATION:

Attenuation means loss of light energy as the light pulse travels from one end of the fiber cable to the other. It is also called as signal loss or fiber loss. It is directly proportional to the length of the cable. It limits the optical power which can reach the receiver at the destination end of the fiber.

Attenuation is mainly caused as a result of

      i.   Absorption loss

     ii.   Scattering loss and

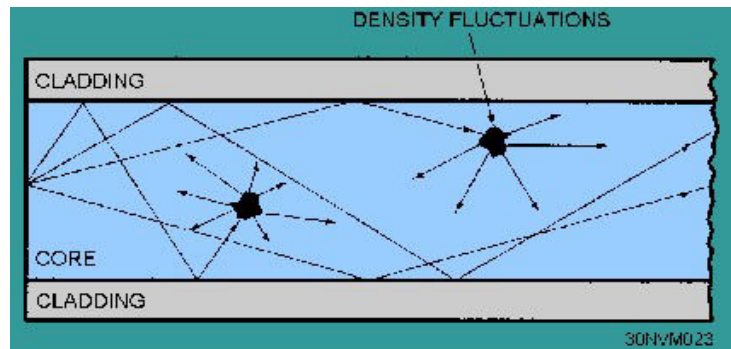   iii. Bending losses.

Attenuation is defined as the ration of input optical power (P_i) to the output optical power (P_o)

The following equation defines the signal attenuation as a unit of length:

$$\alpha(dB/km) = -\frac{10}{L}\log_{10}\left(\frac{P_O}{P_i}\right)$$

i. **Absorption loss:** Every material has a characteristic of absorbing a fraction of the incident light. Optical fibers are also no exception. This property is called intrinsic absorption. Besides the intrinsic absorption, the impurities whatsoever present in optical fiber also absorb light which is called impurity (Extrinsic) absorption. Such types of absorptions result in the reduction of the strength of the light signal propagating through the optical fiber cable.

ii. **Scattering loss:** Light signal scattering can be thought of as the deflection of a ray from a straight path, for example by irregularities in the propagation medium, particles or in the interface between the two media. Irregularities and defects (which are produced when optical fibers are manufactured) are main causes for the scattering of light in unexpected directions.


iii. **Bending losses:** This loss induced by physical stress on the fiber. Bending loss is classified according to the bend radius of curvature:

         a) Macrobend Loss   b) Microbend Loss

a) Macroscopic Bending: Macro-bend Losses are observed when a fiber bend's radius of curvature is large compared to the fiber diameter. These bends are a great source of loss when the radius of curvature is less than several centimeters.

Macroscopic Bending
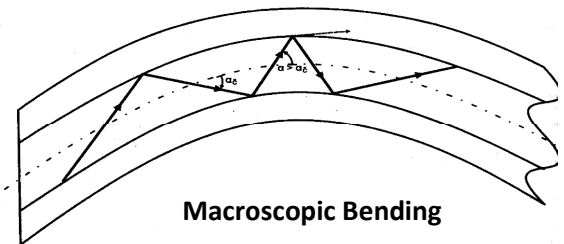
b) Microscopic Bending: Micro-bend Loss are caused by small discontinuities or imperfections in the fiber. Uneven coating applications and improper cabling procedure increases micro bend loss. External forces are also a source of micro bends.

Microscopic Bending
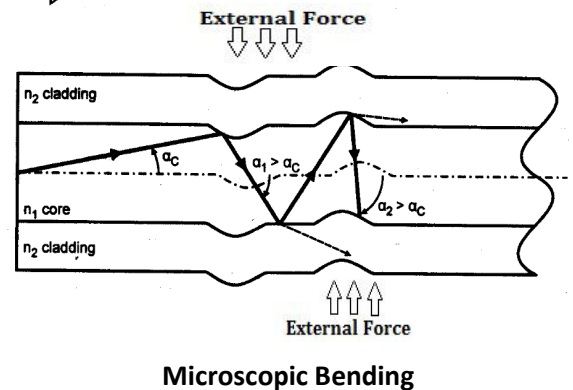
APPLICATIONS OF OPTICAL FIBERS:

Due to its variety of advantages optical fiber has a wide range of application in different fields namely:

    i.   Communication:

    ii. Medicine and

    iii. Sensors etc.,

i. **COMMUNICATION:** Optical fibers are used as wave guides in the communication system. A typical block diagram of optical fiber communication system (OFCS) is shown in the following figure. It mainly consists of the following parts:

    a) Encoder, b) Transmitter, c) Waveguide, d) Receiver and d) Decoder.
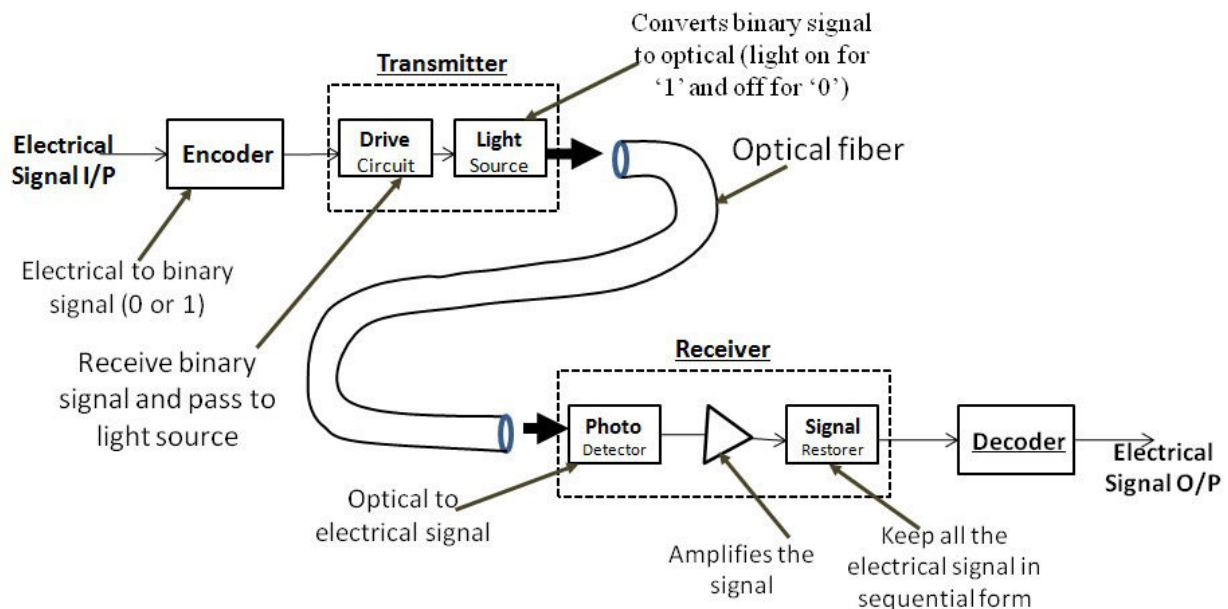

Fig.: Block Diagram of Optical Fiber Communication System

a) Encoder: The audio signal (i.e., the words spoken by us) is converted into electrical signal which is an analog signal. Encoder is an electronic circuit that converts this analog signal into binary or digital signals.

b) Transmitter: The digital signal from the encoder is fed to the transmitter which consists of two parts- Drive circuit and Light source. Drive circuit receives the digital signal from encoder and feeds it to the light source. Light source is usually LED or a Diode LASER. If digital '0' is received then light source will be turned OFF. If digital '1' is received then the light source will be turned ON. Thus light source converts electrical signals into optical signals.

c) Waveguide: Now the Optical signals generated by the transmitter are fed to an optical fiber which acts as waveguide. The signal traverses over longer distances through these waveguides.

d) Receiver: On the other side of the waveguide, he optical signal is received by the receiver which consists of Photo detector, amplifier and a signal restorer. The Photo detector

receives the optical signal and generates the equivalent electrical signals. These electrical signals are amplified by the amplifier. The signal restorer keeps all the electrical signals in a sequential form and supplies to decoder.

e) **Decoder:** It is an electronic system that converts the digital signal to analog signal.

ii. **MEDICAL:**

Optical fibers are generally used in Endoscopy. They are also used in LASER Angioplasty (Laparoscopic Surgery) which is usually used for operations in the stomach area such as appendectomies. A LASER Angioplasty usually makes use of three channels (bundles) of optical fibers. Channel 1(One optical fiber cable) is used to observe where exactly the cholesterol deposits are present. LASER of suitable power is sent through channel 2 to destroy the cholesterol deposits. Channel 3 is used to suck out the debris.

iii. **SENSORS:**

Another important application of optical fibers is in sensors. If a fiber is stretched or squeezed, heated or cooled or subjected to some other change of environment, there is usually a small but measurable change in light transmission.

Level Sensors:

An Optical fiber, containing of core alone is used in a level sensor. The condition here is that the refractive indices of air, core and liquid should such that $n_{air} < n_{core} < n_{liquid}$.

A light signal from the source is fed to the fiber core as shown in fig. (i). The light signal reaches the detector after getting subjected to total internal reflections in the core. It is observed here that the liquid level did not touch the tip of the optical fiber yet and hence the light signal smoothly reaches the detector.

When the liquid raises to sufficient level as shown in figure (ii), at the point of core-liquid interface, total internal reflection cannot take place. The reason is $n_{core} < n_{liquid}$. Therefore the light signal gets leaked into liquid without reaching the detector.

Thus the light signal reaching the detector indicates lower liquid level while the detector does not receive the light signal, indicates sufficient level of the liquid.
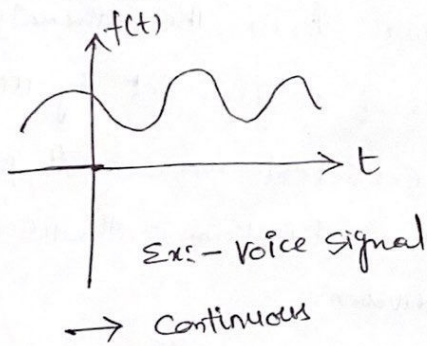

Fig (i)

Fig (ii)
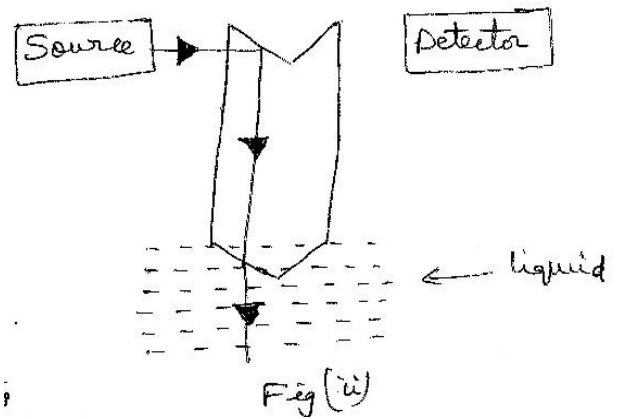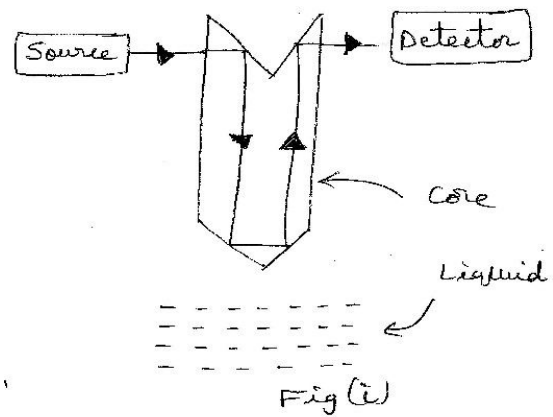
Digital Electronics.
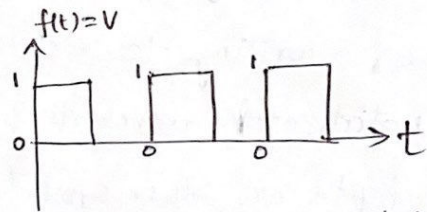
<u>Introduction</u>:— Digital electronics is a field of electronics involving the study of digital signals and the engineering of devices and that use or produce them.

→ Digital electronic Circuits are usually made from large assemblies of logic gates, often packaged in integrated circuits.

→ Digital electronics deals with the electronic manipulation of Numbers, or with the manipulation of varying quantities by means of number.

→ Digital Circuits are collections of devices that perform logical operations on two logical states (0 & 1), by means of AND, OR, NOT, etc.

<u>Analog Signal</u> :—

$f(t)$

Ex:— Voice Signal

→ Continuous

<u>Digital signal</u>

$f(t) = V$

Ex:— Computers, mobile phones, etc.

→ Discreate

<u>Characteristics of Digital Systems</u>:—

i). Digital Systems are the systems that use discreate values such as numbers (from 0, to 10) and letters (Alphabets)

ii). The signal in most digital systems use two values '0 & 1' called a bit'

iii). Discrete elements of information are represented with a group of bits called 'binary codes'.

iv). Therefore, Digital system is a system that manipulates discrete elements of information represented internally in binary form.

Advantages of Digital System over Analog system:—

i). Ease of programmability

ii). Reduction in Cost of hardware

iii) High speed.

iv). High Releability

v). ~~speed~~ Easy design,

vi). Easy Results.

Number Systems:— A Number system is defined as a system of writing to express Numbers. It is the mathematical notation for representing numbers of a given set by using digits or other symbols in a consistent manner. It provides a unique representation of every number and perform arithmetic operations like addition, subtraction, and division.

   The value of any digit in a number can be determined by
       → The digit ⇒ Its position ⇛ The Base of the number system.

There are various types of number systems. The four most common number systems are: 1. Decimal Number System (Base-10) ⟶ 0,—9.

   2. Binary Number system (Base: 2) → 0,1.

   3. Octal Number System (Base÷8) → 0, to 8

   4. Hexadeimical NumberSystem (Base: 16) → 0 to 9 & A,B,C,D,E&F.

① <u>Decimal Number System</u>:- Everyone is familiar with one number system known as decimal Number system. The 'Deci' means 'ten' so this system has 10 distinct digits or symbols as follows:

$$0, 1, 2, 3, 4, 5, 6, 7, 8, 9.$$

→ The decimal numbers fall in the category of positional number system. Since the position of a digit indicates the significance to be attached to that digit.

Ex:① 7639, ⇒ $7 * 1000 + 6 * 100 + 3 * 10 + 9 * 1$

$$⇒ 7 \times 10^3 + 6 \times 10^2 + 3 \times 10^1 + 9 \times 10^0 = 7639$$

② $5367.42 ⇒ 5 \times 10^3 + 3 \times 10^2 + 6 \times 10^1 + 7 \times 10^0 + 4 \times 10^{-1} + 2 \times 10^{-2}$

② <u>Binary Number System</u>:- It has the Base of $\overset{r=}{2}$ (radix) & its elements or digits will be '0' & '1' only. This system is known as Binary number system as its radix is '2' (Binary means 2). The digits '0' & '1' of this system are known as "bits". This number system finds extensive use in digital Electronics.

Decimal	Binary		Decimal	Binary
0	0		8	1000
1	01		9	1001
2	10		10	1010
3	11		11	1011
4	100		12	1100
5	101		13	1101
6	110		14	1110
7	111		15	1111

→ Binary number can be pronounced as individual digits.

Ex:- 10 can be pronounced as 'one' 'zero' not 'ten'.

(a) Conversion of Binary to Decimal number:-

$$Ex(i):- (10110)_2 = 1 \times 2^4 + 0 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 0 \times 2^0$$

$$= 16 + 0 + 4 + 2 + 0 = (22)_{10}$$

$$\Rightarrow (10110)_2 = (22)_{10}$$

$$Ex(ii):- (11011.0101)_2 = 1 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4}$$

$$= 16 + 8 + 0 + 2 + 1 + 0 + 0.25 + 0 + 0.0625$$

$$= (27.3125)_{10}$$

(b) Conversion of Decimal to Binary Number:-

$$Ex:- (i)\ (35)_{10} \longrightarrow (100011)_2 \quad \Big| \quad ii)\ (43.6525)_{10} \rightarrow (101011.101001)_2$$

2	35	
2	17	— 1
2	8	— 1
2	4	— 0
2	2	— 0
2	1	— 0
	0	— 1

2	43	
2	21	— 1
2	10	— 1
2	5	— 0
2	2	— 1
2	1	— 0
	0	— 1

fractional part

$$0.6525 \times 2 = 1.3050 - 1$$
$$0.3050 \times 2 = 0.6100 - 0$$
$$0.6100 \times 2 = 1.220 - 1$$
$$0.220 \times 2 = 0.440 - 0$$
$$0.440 \times 2 = 0.880 - 0$$
$$0.880 \times 2 = 1.760 - 1$$

(3) Octal Number system:- The radix or Base of the octal number system is '8' (octal means '8') and its digits will be '0 to 7'

Decimal	Octal		Decimal	Octal
0	0		7	7
1	1		8	10
2	2		9	11
3	3		10	12
4	4		11	13
5	5		12	14
6	6		13	15
			14	16

Octal Numbers are commonly used in computer Applications. Converting an octal number to decimal Number is the same as decimal Conversion. Explanation is as follows:

$$(215)_8 \longrightarrow (\quad)_{10}$$

$$(215)_8 = 2*8^2 + 1*8^1 + 5\times 8^0 = 2*64 + 1*8 + 5*1$$
$$= 128 + 8 + 5 = (141)_{10}$$

$$\therefore (215)_8 \longrightarrow (141)_{10}$$

© Conversion of octal Number to Decimal Number:-

i) Ex: $(7126 \cdot 45)_8 \longrightarrow (\quad)_{10}$

$$(7126 \cdot 45)_8 = 7*8^3 + 1*8^2 + 2*8^1 + 6*8^0 + 4*8^{-1} + 5*8^{-2}$$
$$= 7*512 + 1*64 + 2*8 + 6*1 + 0.5 + 0.078125$$
$$= 3584 + 64 + 16 + 6 + 0.5 + 0.078125$$
$$= (3670.57812)_{10}$$

ii),

ⓓ Conversion of Decimal Number to octal Number:-

Ex:-i) $(567)_{10} \longrightarrow (\quad)_8$

ii) $(567 \cdot 6525)_{10} \longrightarrow (\quad)_8$

```
8 | 567
8 |  70  — 7
8 |   8  — 6
8 |   1  — 0
      0  — 1
```

$= (1067)_8$

$0.6525 \times 8 = 5.2200 — 5$
$0.2200 \times 8 = 1.7600 — 1$
$0.7600 \times 8 = 6.0800 — 6$

$\Rightarrow (1067 \cdot 516)_8$

④ Hexadecimal Number System :- In Hexadecimal number system, the radix or base is 16 and its digits will be 16 distinct elements.

In this number system, the numbers are first represented just like in decimal system, i.e from 0 to 9. Then, the remaining are represented using the Alphabets from A to F.

Decimal	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Hexa Decimal	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F

ⓔ Conversion of hexadecimal Number to Decimal Number :-

Ex:- $(3BC7 \cdot 46)_{16} \longrightarrow (\quad)_{10}$

$(3BC7 \cdot 46)_{16} = 3*16^3 + 11*16^2 + 12*16^1 + 7*16^0 + 4*16^{-1} + 6*16^{-2}$

$= 12288 + 2816 + 192 + 7 + 0.25 + 0.02343$

$= (15303 \cdot 2734)_{10}$

$\Rightarrow (3BC7 \cdot 46)_{16} = (15303 \cdot 2734)_{10}$

ⓕ Conversion of Decimal Number to Hexadecimal Number :-

Ex:- $(8537)_{10} \rightarrow (\quad)_{16}$ Ex:- $(98765)_{10} \rightarrow (\quad)_{16}$

```
16 | 8537
16 |  533 — 9
16 |   33 — 5
16 |    2 — 1
        0 — 2
```
$= (2159)_{16}$

```
16 | 98765
16 |  6172 — D
16 |   385 — C
16 |    24 — 1
16 |     1 — 8
        0 — 1
```
$= (181CD)_{16}$

Ex:- Do the following Conversions:-

    i) $(965.125)_{10} \rightarrow ($ $)_8$ (ii) $(8765.025)_{10}$ to $($ $)_{16}$

    iii) $(6754.05)_8$ to $($ $)_{10}$

(g) Conversion of octal to Binary :- since the octal system has a base of '8' $(2^3 = 8)$, so each octal Number Cem be represented as a group of '3' bits. Therefore, starting with the least significant bit of the binary Number, the successive three bits are arranged together in the form of groups. These groups of three bits are replaced by their octal equivalents as shown in fig. below.

octal	Binary		OCTAL	BINARY
0	000		4	100
1	001		5	101
2	010		6	110
3	011		7	111

Ex:- $(2467.534)_8 \rightarrow ($ $)_2$

Sd:-

2	4	6	7	.	5	3	4
010	100	110	111	.	101	011	100

$\therefore (2467.534)_8 \rightarrow (10100110111.101011)_2$

(h) Conversion of Binary to octal system:- Binary numbers are Converted to the octal Numbers by making the groups of three bits from right to left in the integer part and from left to right on the fractional part of Binary number. If the need arises for making the groups of three bits, one or two Zeros may be added to the left of most significant bits and/or to the right of the least significant bit of the fractional part of the binary number.

    Ex:- $(11001101111110.1011)_2$ to $\rightarrow ($ $)_8$

Sd:-

    001 100 110 111 110 . 101 100

    $\rightarrow$ 1   4   6   7   6 . 3   4 $= (14676.34)_8$

ⓘ Conversion of Hexadecimal to Binary :- As the Hexadecimal

system has a base 16 ($2^4 = 16$) so every hexadecimal digit can be represented as a group of 4 bits as shown in below table.

Hexa Decimal	Binary	Hexadecimal	Binary
0	0000	8	1000
1	0001	9	1001
2	0010	A	1010
3	0011	B	1011
4	0100	C	1100
5	0101	D	1101
6	0110	E	1110
7	0111	F	1111

Ex:- $(2ABGE7.5D4)_{16} \longrightarrow (\quad)_2$

Sol:-

2	A	B	6	E	7	.	5	D	4
0010	1010	1011	0110	1110	0111	.	0101	1101	0100

$\therefore (2ABGE7.5D4)_{16} \longrightarrow (101010101101101110 0111 . 010111 0101)_2$

ⓙ Conversion of Binary to Hexadecimal :- Its very similar to Binary to octal Conversion.

Ex:- $(1100110101101111110.101)_2 \longrightarrow (\quad)_{16}$

Sol:-

0011	0011	0101	1011	1110	.	1010
3	3	5	B	E	.	A

$\therefore (1100110101101111110.101)_2 = (335BE.A)_{16}$

(K) Conversion of Hexadecimal to octal & vice-versa:-

Ex:- ① $(29AF \cdot 1EA)_{16} \rightarrow (\quad)_8$

Sol:-

2	9	A	F	.	1	E	A	
0010	1001	1010	1111	.	0001	1110	1010	→ Group of four Bits

000 010 100 110 0101 111 . 000 111 101 010 → Group of Three Bits

0 2 4 6 5 7 . 07 5 2 → octal equivalent

$\therefore (29AF \cdot 1EA)_{16} = (24657 \cdot 0752)_8$

② $(6327 \cdot 405)_8 \rightarrow (\quad)_{16}$

Sol:-

6	3	2	7	.	4	0	5	
110	011	010	111	.	100	000	101	→ Group of 3 bits

1100 1101 0111 . 1000 0010 1000 → Group of 4 bits.

C D 7 . 8 2 8

$\therefore (6327 \cdot 405)_8 = (CD7 \cdot 828)_{16}$

Binary Logic :- Binary logic consists of binary variables and logic operations. The variables are designated by the alphabets such as A, B, C, X, Y, Z, etc., with each variable having only two distinct values: 1 (one) & 0 (zero). The logic & Mathematics of this system were developed by mainly by the British Mathematician George Boole (1815-64), whose contributions were so important that this concept is often referred to as "Boolean Algebra"

Binary logic is the basis of electronic systems, such as

Computers, and Cellphones. It works on `0`'s and 1's. It Involves Addition, Subtraction, multiplication, devision ob zeros & ones. It includes Logic gate functions AND, OR, and NOT which translates I/p Signals into specific O/p.

Boolean Algebra operations:- The basic operations ob Boolean algebra are as follows: 1. AND operation (conjunction)
2. OR operation (Disjunction)
3. NOT operation (Negation)

Logic gates are elementary building blocks of a digital circuit. Most logic-gates have two I/p's and only one O/p. The Basic logic gates are AND, OR & NOT. Binary logic deals with true and false. You know that `0` is usually associated with 'FALSE' and '1' is associated with 'TRUE'.

① AND operation:- It Also called as Boolean multiplication. It corresponds to logical Conjuption of an AND gate, as well as to series switch contacts. . The AND operation is represented by a dot (•) or by the Absence of an operator.

Ex:- A·B = Y (or) AB = Y is read as "A AND B is equal to Y".

The logic operation AND is interpreted that Y=1 if and only If A=1 & B=1; otherwise Y=0.

A	B	Y = AB
0	0	0
0	1	0
1	0	0
1	1	1

A B Y=AB

Y A B V
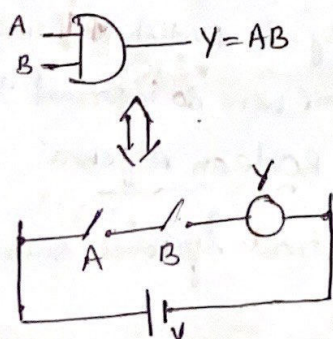
② **OR operation :-** It is also called as Boolean Addition. It corresponds to the logical Junction of an OR gate, as well as to parallel switch Contacts.

The OR operation is represented by a plus (+) sign. for Ex.,

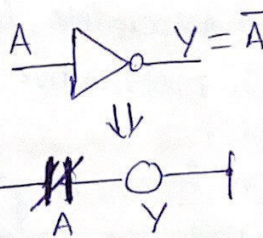A+B = Y is read "A OR B equal to Y", meaning that Y=1 if A =1 or B =1 or if Both A=B=1.

A	B	Y=A+B
0	0	0
0	1	1
1	0	1
1	1	1
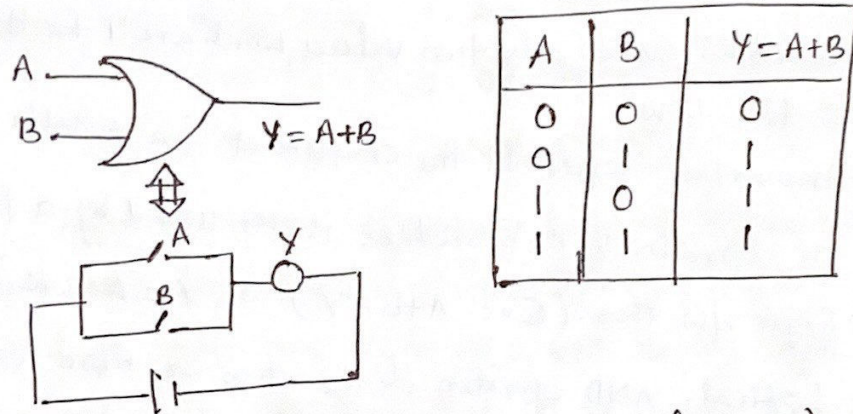
③ **NOT operation (Boolean Complementary function) :-**

NOT operation corresponds to logical NOT gate, as well as closed Switch or relay contact.

The NOT operation is represented by a prime (sometimes by a bar).

Ex:- If Input is A , The output is $Y = \bar{A}$ (or) A'

If A =1 , then $\bar{A}$ =0 .
A =0 Then $\bar{A}$ =1

A	$Y=\bar{A}$
D	1
1	0

Boolean Algebra Truth tables:- The Truth tables represent the mathematical representation of above operators such as AND, OR & NOT

A	B	Y = AB	Y = A+B (OR)
F	F	F	F
F	T	F	T
T	F	F	T
T	T	T	T

A	Y = $\bar{A}$ (NOT)
0(F)	T
T	F

→ truth table is a table that gives all possible values of logical variables and the combinations of it. It is possible to convert the Boolean equation into truth table. The number of rows in the truth table should be equal to "2^n", where n is the number of variables in the eq.

Boolean Algebra Rules:- (Properties):-

i). Variables used only two values, which are '1' for HIGH and '0' for LOW.

ii). An overbar represents the complement of a variable. ($\bar{A}$).

iii). OR operation of the variables is represented by a plus (+) sign b/w them. (Ex: A+B = Y) or Y = A+B+C)

iv). Logical AND operation of the two or more variables are represented by putting dot (•) between them. (Ex: A·B·C) (OR) ABC)

LAWS OF BOOLEAN ALGEBRA:- There are six types of Boolean algebric laws. They are:

1. Commutative law
2. Associative law
3. Distributive law
4. AND law
5. OR law
6. NOT law (Inversion)

① . Commutative law:- Any Binary operation, It states that, when we add or multiply two variables (Binary) then the resultant value remains same, even if we change the position of the two variables.

→ Changing the sequence of the variables does not have any effect on the o/p of a logic circuit.

Law 1 :- A·B = B·A

LAW 2 :- A+B = B+A

② Associative law :— It states that the order in which the logic operations are performed is irrelevant as their effect is the same.

LAW 1: $(A \cdot B) \cdot C = A \cdot (B \cdot C)$

LAW 2: $(A+B)+C = A+(B+C)$

③ Distributive law :— It states that the product of an expression and a sum is equal to the sum of the products of the logical expression and each term in the sum.

LAW 1: $A \cdot (B+C) = (A \cdot B) + (A \cdot C)$

LAW 2: $A+(B \cdot C) = (A+B) \cdot (A+C)$

④ AND law :— These laws use the AND operation. Therefore, they are called AND laws.

→ Annulment law ⓐ   → Identity law ⓑ   → Idempotent law ⓒ

i). $A \cdot 0 = 0$, ii) $A \cdot 1 = A$ iii) $A \cdot A = A$ iv) $A \cdot \overline{A} = 0$
↳ complement law ⓓ

⑤ OR Law :— These laws use the OR operation. Therefore, they are called OR laws.

i). $A+0 = A$ ii) $A+1 = 1$, iii) $A+A = A$ iv) $A+\overline{A} = 1$
     ↳ⓑ   ↳ⓐ   ↳ⓒ   ↳ⓓ

⑥ Inversion law :— This law uses the NOT operation. The Inversion law states that double inversion of Variable results in the original Variable itself.

If $A = 0$, $\overline{A} = 1$, $\overline{\overline{A}} = 0 = \overline{A}$

OTHER Boolean Algebric laws :—

⑦ Absorption (&) Redundancy law :—

i) Absorption law

ⓐ $A+AB = A$

ⓑ $A(A+B) = A$

ii) Redundancy law

ⓐ $A+\overline{A}B = A+B$

ⓑ $A \cdot (\overline{A}+B) = AB$

BOOLEAN ALGEBRE THEOREMS:-

The Two Important theorems, which are extremely used in Boolean Algebra, are deMorgan's laws (theorems). There are Two DeMorgan theorems Used to change the Boolean expression. These theorems help to reduce the given Boolean expression in the simplified form.

i) De Morgan's first theorem:-

$$\overline{A \cdot B} = \overline{A} + \overline{B} \qquad \left((A \cdot B)' = A' + B' \right)$$

The first theorem states that, the complement of the pro-duct of the variables is equal to the sum of their individual complements of a variable.

The truth table shows that the verification of Demorgan's first theorem is given as follows:

A	B	A'	B'	(AB)'	A' + B'
0	0	1	1	1	1
0	1	1	0	1	1
1	0	0	1	1	1
1	1	0	0	0	0

The last two columns show that $(A \cdot B)' = A' + B'$

Hence, Demorgan's first theorem is proved.

Demorgans first theorem interms of logic gates:-

$$\overline{AB} = \overline{A} + \overline{B}$$

NAND Gate = Bubbled OR gate.
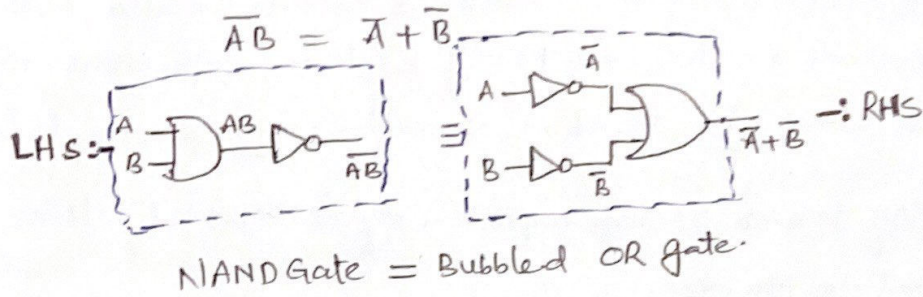
ii) DEMORGAN'S SECOND THEOREM:-

Demorgan's second theorem states that $\overline{(A+B)} = \overline{A} \cdot \overline{B}$

It states that, the complement of the sum of Boolean variables is equal to the product of their individual complements of a variable.

In terms of Logic gates:- $\overline{(A+B)} = \overline{A} \cdot \overline{B}$

NOR Gate = Bubbled AND gate.
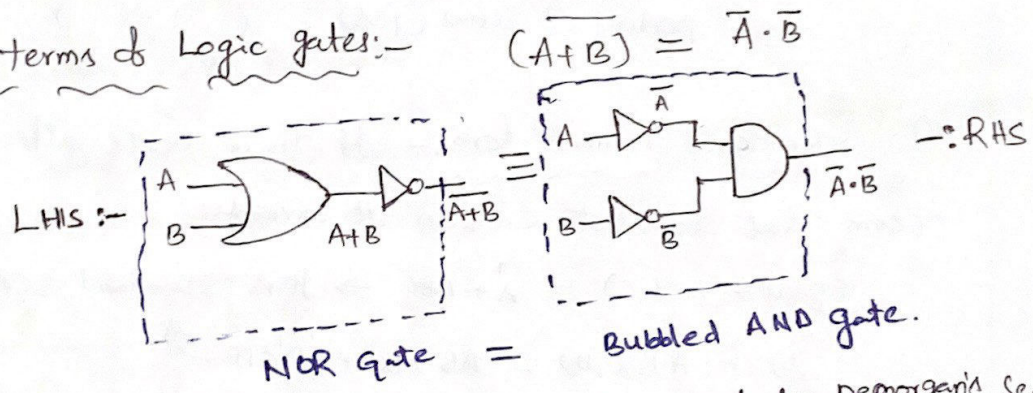
The following truth table shows the proof for Demorgan's second theorem:

The last Two columns shows that

$\overline{A+B} = \overline{A} \cdot \overline{B}$

A	B	$\overline{A}$	$\overline{B}$	A+B	$\overline{A+B}$	$\overline{A} \cdot \overline{B}$
0	0	1	1	0	1	1
0	1	1	0	1	0	0
1	0	0	1	1	0	0
1	1	0	0	1	0	0

BOOLEAN FUNCTIONS:- (1) Boolean functions are represented in various forms). Boolean functions consists of Boolean variables, Boolean constants and Boolean operators. (2). Mainly Boolean expressions represented in two forms: i) standard (Normal) form ii) Canonical form.

Any form of Boolean expression may be expressed interms of either minterms or Maxterms.

@ Minterms (SOP):- They are the products of I/p variables such that value of the product is '1'.

Ex:- Sum of products (SOP): (   ) + (   )
                              1   1

(b) Maxterms (POS):- They are the sum of I/p variables such that the total value of sum is zero(0).

Ex:- product of sums (POS): (   ) · (   )
                              0   0

i). Standard (Normal) form:- If there exists atleast one term that does not contain all variables.

Ex:-a) $F(A, B, C) = A' + ABC$ → It is standard SOP exp.

b). $F(A, B, C, D) = AB + BC + A'B'CD$ ↗

ii) Canonical form:- Each term of Boolean expression contain all I/p variables either in true form or in complement form.

Ex:-a) $F(A, B, C) = A'B'C + ABC$ → It is Canonical SOP exp

b) $F(A, B, CD) = A'BCD + AB'C'D + ABCD$ ↗

Boolean expression in the form of standard or canonical form can be evaluated by using Truth tables or Venn Diagrams.

Ex:- i) Midterm & Maxterms for given Boolean variables.

   ii) Identify the midterm & Maxterm expressions for the following truth tables.

Sol:-(i)

A	B	C	minterm Exp	Maxterm Exp
			1	**'0'**
0	0	0	A'B'C'	A+B+C
0	0	1	A'B'C	A+B+C'
0	1	0	A'BC'	A+B'+C
0	1	1	A'BC	A+B'+C'
1	0	0	AB'C'	A'+B+C
1	0	1	AB'C	A'+B+C'
1	1	0	ABC'	A'+B'+C
1	1	1	ABC	A'+B'+C'

Sol:-ii)

S. NO.	A	B	C	F
0	0	0	0	0
1	0	0	1	1
2	0	1	0	1
3	0	1	1	0
4	1	0	0	1
5	1	0	1	0
6	1	1	0	0
7	1	1	1	1

$F(A,B,C) =$
$\Sigma = m_1 + m_2 + m_4 + m_7$
$= A'B'C + A'BC' + AB'C' + ABC$

$F(A,B,C) = M_0 \cdot M_3 \cdot M_5 \cdot M_6$

$\pi() = (A+B+C) \cdot (A+B'+C') \cdot$
$(A'+B+C') \cdot (A'+B'+C)$

→ Evaluate the following Boolean Expression using truth tables:

$$Z = AB' + A'C + A'B'C$$

Sol:- There are three Boolean Variables; A, B, & C. Hence, there will be
$2^3 = 8$ combinations of boolean values.

   To evaluate Z in the above function, knowing the values for A, B, & C at each row of the truth-table, we would first generate the values for A', B' and then AB', A'c & A'B'c by AND operation. Finally, we would derive the Z value by OR operation among above values.

A	B	C	A'	B'	AB'	A'c	A'B'c	Z
0	0	0	1	1	0	0	0	0
0	0	1	1	1	0	1	1	1
0	1	0	1	0	0	0	0	0
0	1	1	1	0	0	1	0	1
1	0	0	0	1	1	0	0	1
1	0	1	0	1	1	0	0	1
1	1	0	0	0	0	0	0	0
1	1	1	0	0	0	0	0	0

LOGIC GATES:-

A logic gate is a basic building block of a digital circuit that has two or more inputs and one output. The relationship between the I/p and the o/p is based on a certain logic.

I/p's ⟹ [LOGIC Operation] ⟹ O/p.
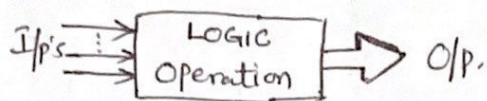
The logic gates are classified into seven types; in which 3 are Basic logic gates, 2-are Universal gates & 2 are Exclusive gates.

i) Basic [AND, OR, NOT] gates

ii) Universal [NAND & NOR] gates

iii) Exclusive [Ex-OR & Ex-NOR] gates.

ⓐ AND gate:- An AND gate is a digital logic gate with n-i/p's and one o/p, which performs logical Multiplication based on the combination of its I/p's.

→ The o/p of this gate is "True"(1) only when all the I/p are True(1)

→ The symbol & Truth table of an AND gate is shown below.

I/p's A, B ⟹ o/p $Y = A \cdot B$
Symbol   B·Exp.
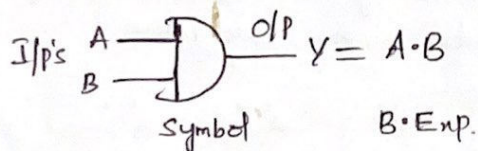
A	B	$Y = A \cdot B$
0	0	0
0	1	0
1	0	0
1	1	1

← Truth Table

ⓑ OR gate:- The OR-gate is a digital logic gate with n-I/p's and one -O/P, that performs logical Addition based on the combinations of its I/p's.

→ The O/p of the OR-gate is high(true) only one or more I/p's are True(1). otherwise false.

I/p's A, B ⟹ o/p $y = A+B$
Symbol.   B·Exp.
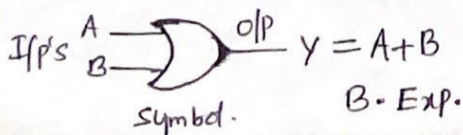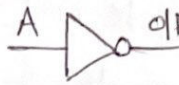
A	B	$Y = A+B$
0	0	0
0	1	1
1	0	1
1	1	1

← Truth table

© NOT-gate:- The NOT-gate is a digital logic gate with one I/p and one o/p that operates an inverter operation of the I/P.

→ The o/p of the NOT-gate is the reverse of the I/P. when the I/p of the NOT gate is '1' (True) - then the o/p will be '0' (false) & vice versa.

I/p A $\longrightarrow$ o/p $Y = \overline{A}$
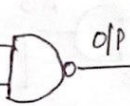
A	$Y = \overline{A}$
0	1
1	0

← Truth table.

Symbol   B. Exp.

④ NAND gate:- The NAND gate is a digital logic gate with n-I/p's and one o/p, that performs the operation of the AND gate followed by operation of the NOT-gate. NAND gate is designed by combining the AND & NOT gates.

→ If the I/p's of the NAND gate is high, then the o/p of the gate will be low, otherwise it will be high.

I/p A B $\longrightarrow$ o/p $Y = \overline{A \cdot B}$
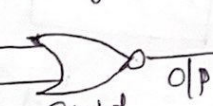
A	B	$Y = \overline{AB}$
0	0	1
0	1	1
1	0	1
1	1	0

← Truth table

Symbol.   B. Exp.

ⓔ NOR-gate:- It is a digital logic gate with n-I/p's and one o/p; that performs the operation of OR-gate followed by NOT-gate oper. NOR gate is designed by combining the OR & NOT gates.

→ when any one of the I/p's of the NOR-gate is true, then the o/p of the NOR-gate is false (0).

I/p's A B $\longrightarrow$ o/p $Y = \overline{A+B}$

A	B	$Y = \overline{A+B}$
0	0	1
0	1	0
1	0	0
1	1	0

← Truth Table.

Symbol   B. Exp.

ⓕ Exclusive-OR (Ex-OR) gate:-

The Exclusive-OR gate is a digital logic gate with two or more I/p's and one o/p. It performs based on the operation of OR-gate. If Any one of the I/p's are high, then the o/p of the Ex-OR gate will be high.

$$Y = A \oplus B$$
$$= A\bar{B} + \bar{A}B$$

Symbol   B. Exp.

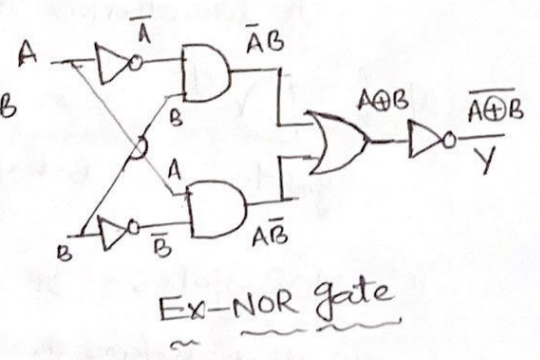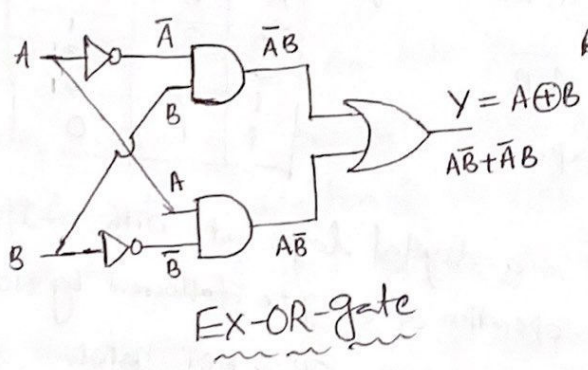| A	B	Y = A⊕B
0	0	0
0	1	1
1	0	1
1	1	0

← Truth Table.

⑨ Ex-NOR gate :— It is a digital logic gate with two or more I/p's and one O/p. It performs based on the operation of NOR gate.

→ when both the I/p's are high, (or) Both are low, then the O/p is high. But, If any one of the I/p's is high, then the O/p will be low.
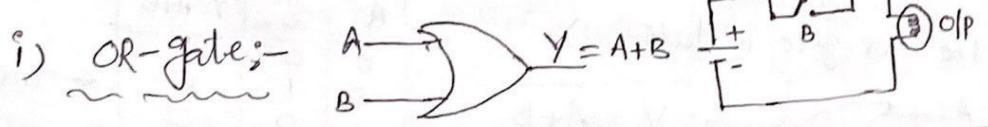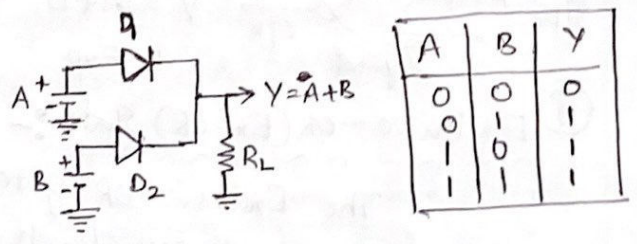

$$Y = \overline{A \oplus B}$$
$$= A \odot B$$

B. Exp.

SYMBOL
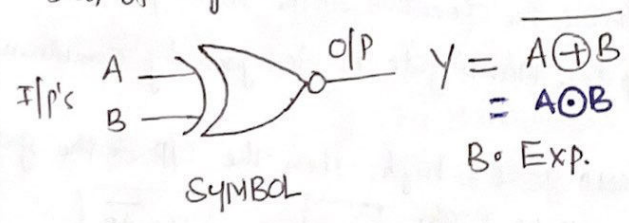
A	B	Y = A⊕B
0	0	1
0	1	0
1	0	0
1	1	1

← Truth Table.

$$Y = A \oplus B$$
$$\bar{A}B + A\bar{B}$$

EX-OR-gate

EX-NOR gate

LOGIC GATES Using discrete Components :—

i) OR-gate :—

$$Y = A + B$$

Diode OR gate :-

Fig. shows the simple circuit using diodes to build a two I/p OR gate. The working of this circuit can be explained as follows :

$$Y = A + B$$

A	B	Y
0	0	0
0	1	1
1	0	1
1	1	1

Case (i)   A = 0 & B = 0 ⇒ when Both A & B are at zero level, (i.e) low, the O/p voltage will be low, because the diodes are non-conducting.

case ii): $A=0$, & $B=1$ ⇒ When A is low, & B is high, Diode D_2 is forward biased so that the o/p is high. (D_1–open, D_2–Shorted)

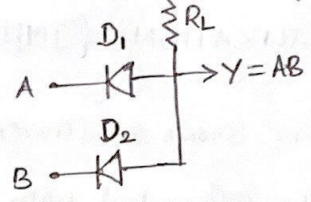Case-iii): $A=1$ & $B=0$ ⇒ When A is high & B is low, Diode D_1 conducts and The o/p is high (D_1–shorted, & D_2–open)

case-iv):– $A=1$ & $B=1$ ⇒ when A & B ^Both^ are high, both diodes D_1 & D_2 are Conducting & the o/p is high. (D_1–shorted & D_2–shorted)

ii) AND –gate:– A ⟶ [gate] ⟶ $Y = A \cdot B$ $+5$v V_{cc}
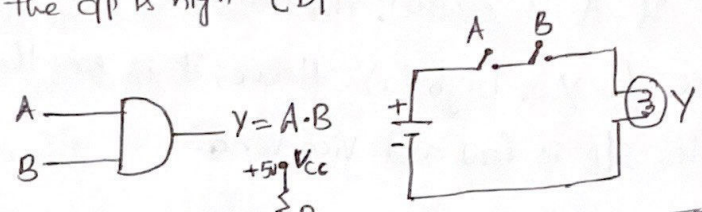
Diode AND gate :–

Fig. Shows a simple Circuit using diodes to Build 2- I/P AND gate. The working of the circuit can be explained as follows:

A	B	Y
0	0	0
0	1	0
1	0	0
1	1	1

CASE–i):– $A=0$ & $B=0$ ⇒ when A & B are zero, Both diodes (D_1 & D_2) are in forward biased condition and they conduct and hence the o/p will be zero, because the supply voltage V_{cc} will be dropped across R_L only. Therefore, $Y=0$.

Case –ii):– $A=0$ & $B=1$ ⇒ when $A=0$ & B is high, Diode D_1 is F. Biased & D_2 is reverse Biased. The D_1 will now conduct and D_2 is open. Therefore o/p is 0

CASE –iii):– $A=1$ & $B=0$ ⇒ When $A=1$ & $B=0$, diode D_2 will be conducting & hence the o/p $Y=0$

CASE –iv):– $A=1$ & $B=1$ ⇒ In this case, both the diodes are not–conducting. Since D_1 & D_2 are in OFF condition, The o/p is equal to the supply voltage. Therefore $Y=1$
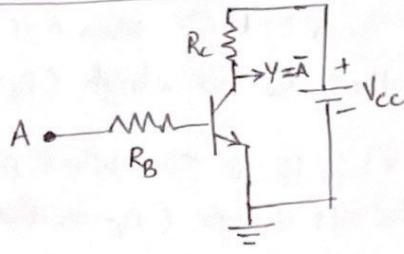
iii) NOT–gate:– (INVERTER) A ⟶ [inverter] ⟶ $Y = \bar{A}$
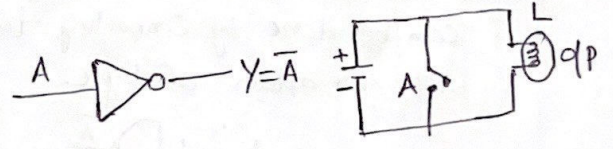
The NOT gate is the gate with only one I/p and one o/p. It is so called because its o/p is complement to the I/p. It is also known as Inverter.

→ The Transistor in CE mode, which is used as NOT gate as shown in fig.

→ when the I/p A is high,
The transistor is driven into
Saturation (ON-state) & hence
the o/p is low. (Y=low).

A	$Y=\bar{A}$
0	1
1	0

If A is low (0), the transistor is in Cutoff (OFF-state) & hence
the o/p Y is high (1). Hence, it is seen that, whenever Input is high,
the o/p is low and Vice-versa.

NAND GATE REALIZATION:- [IMPLEMENTATION]

NAND is known as "Universal gate" because any digital
Circuit (Basic) can be implemented with NAND gates alone. To prove the above
statement, it suffices to show that AND, OR & NOT gates can be implemented
using NAND gates only.

a) NOT gate Design using NAND:- The designing of the NOT gate is very simple
by simply connecting both the inputs as one.

$$A \longrightarrow Y = \overline{A \cdot A} = \bar{A} \Rightarrow A \longrightarrow Y = \bar{A}$$
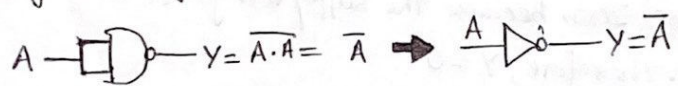
b) AND gate Design using NAND :- The designing of AND gate using NAND gate
can be done at the NAND gate's o/p to reverse it and obtain AND logic.

$$A, B \longrightarrow \overline{AB} \longrightarrow AB \Rightarrow A, B \longrightarrow Y = A \cdot B$$
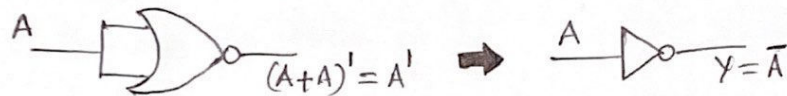
c) OR gate design using NAND:- The designing of OR gate using NAND gates
can be done by connecting two NOT gates using NAND gates at the NAND's
I/p's to obtain OR logic.

$$\overline{\bar{A} \cdot \bar{B}} = A + B \Rightarrow Y = A + B$$
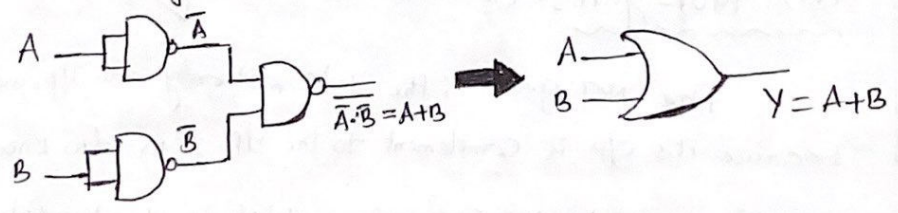
NOR Gate REALIZATION :- [IMPLEMENTATION]

NOR gate is known as a "universal" gate because any digital circuit (Logic gate) can be implemented with NOR gate alone. The following are the implementation of NOT, AND and OR gates using NOR only.

ⓐ NOT gate design using NOR:- The designing of NOT gate with NOR gate is simply by connecting both the I/p's as one.

$$A \quad (A+A)' = A' \quad \Rightarrow \quad A \quad Y = \bar{A}$$

ⓑ OR gate design using NOR :- The designing of OR gate with NOR gate is simply by connecting NOR gate Inverter at the o/p of the NOR gate and obtain the OR logic.

$$A, B \quad \overline{A+B} \quad A+B \quad \Rightarrow \quad A, B \quad y = A+B$$
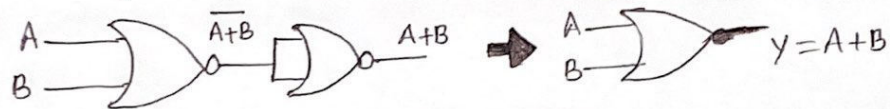
ⓒ AND gate design using NOR:- The designing of AND-gate using NOR gate can be done by connecting two NOR gate Inverters at the NOR inputs to obtain AND logic.

$$A \quad \bar{A} \quad B \quad \bar{B} \quad \overline{\bar{A}+\bar{B}} = AB \quad \Rightarrow \quad A, B \quad Y = A \cdot B$$
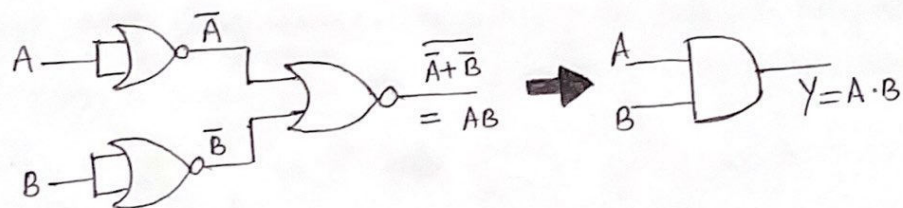
Integrated Circuits:

The basic logic gates discussed above were designed using discrete components like diodes, transistors, and resistances etc. In the recent past, it has been possible to fabricate many hundreds of thousands of active and passive components on a small silicon chip. Such fabricated devices are known as integrated circuits (ICs). The Integrated circuits are broadly classified in two categories namely Linear or analog ICs and digital ICs. The analog ICs mainly contain amplifiers, operational amplifiers, audio and power amplifiers etc. However, the digital ICs contain logic gates etc. The variety of logic gates are fabricated in digital ICs using various technologies.

The digital ICs may further be classified into following categories depending upon their level of integration:

1. **Small Scale Integrated Circuits (SSI):**

   Twelve gates per IC are fabricated in SSI and total number of components per chip is less than 100.

2. **Medium Scale Integrated Circuits (MSI):**

   These ICs contain 12 to 100 gates per IC and total number of components per IC is 100 to 1000.

3. **Large Scale Integrated Circuits (LSI):**

   The large-scale integrated circuits contain 100 to 1000 gates per IC and number of components is 1000 to 10000 per IC.

4. **Very Large Scale Integrated Circuits (LSI):**

   These ICs contain more than 1000 and less than 10000 gates per IC and total number of components per chip is 10000 to 100000.

5. **Ultra Large Scale Integrated Circuits (LSI):**

   More than 10000 gates per IC are fabricated and total components are more than 100000 per chip.

Logic ICs

Typically, standard logic gates are available in 14 pin or 16 pin DIL (dual in line) chips. The number of gates per IC varies depending on the number of inputs per gate. Two–input gates are common, but if only a single input is required, such as in the 7404 NOT(or inverter) gates, a 14 pin IC can accommodate 6 (or Hex) gates.
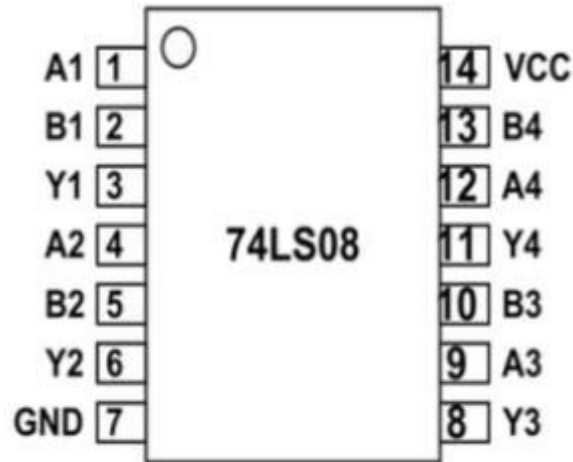
IC logic gate-AND:

IC 7408 is a **QUAD 2-Input AND gate** and contains four independent **gates** each of which performs the logic AND function. It comes in 14-pin DIP package.
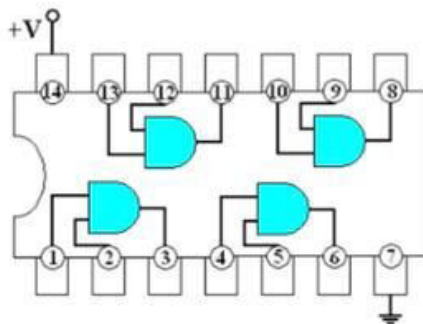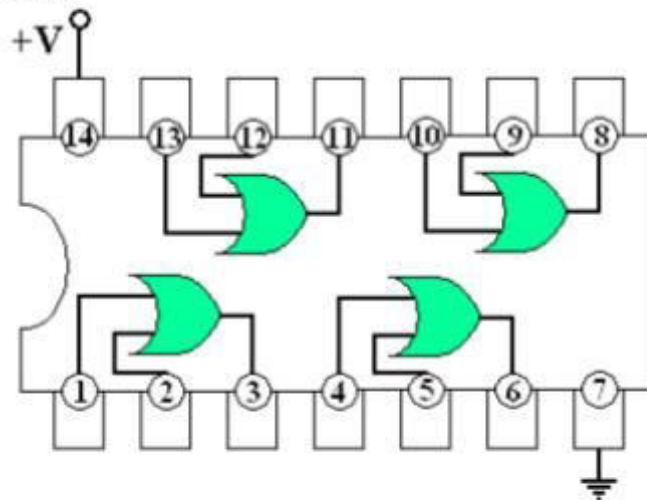

Pin description: Here A and B represents input pins and Y represents output pin. Pins 14 and 7

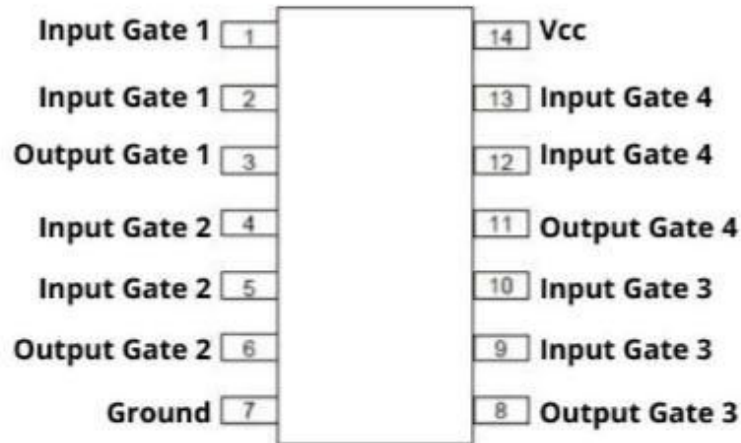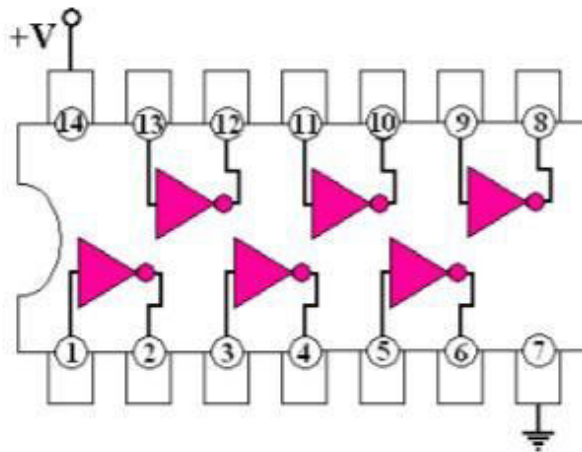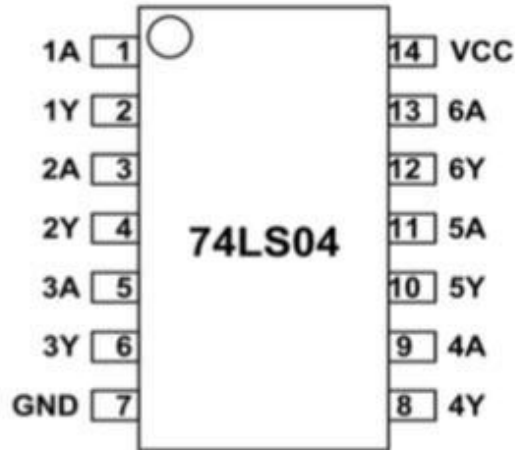provide power for all four logic gates.

IC logic gate-OR:

The **7432** is a Dual **Input OR Gate** with the **Quad** package. It contains four independent gates each of which performs the logic OR function. Each **gate** has **two inputs** that's why it is named **Quad 2-Input OR Gate.**
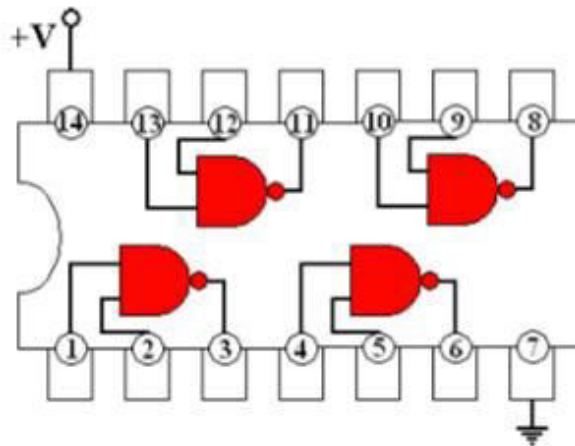

Pin description: Here A and B represents input pins and Y represents output pin. Pins 14 and 7 provide power for all four logic gates.

7432

Input Gate 1 — 1	14 — Vcc
Input Gate 1 — 2	13 — Input Gate 4
Output Gate 1 — 3	12 — Input Gate 4
Input Gate 2 — 4	11 — Output Gate 4
Input Gate 2 — 5	10 — Input Gate 3
Output Gate 2 — 6	9 — Input Gate 3
Ground — 7	8 — Output Gate 3

IC logic gate-NOT:

   7404 is a Hex inverter **IC**. This is a 14 pin DIP version of the popular NOT Gate **IC**. It consists of six inverters (NOT gates), which perform logical invert action. The output of an inverter is the complement of its input logic state, i.e., when input is high its output is low and vice versa.

Pin Description: Here A represents input pin and Y represents output pin. Pins 14 and 7 provide power for all four logic gates.
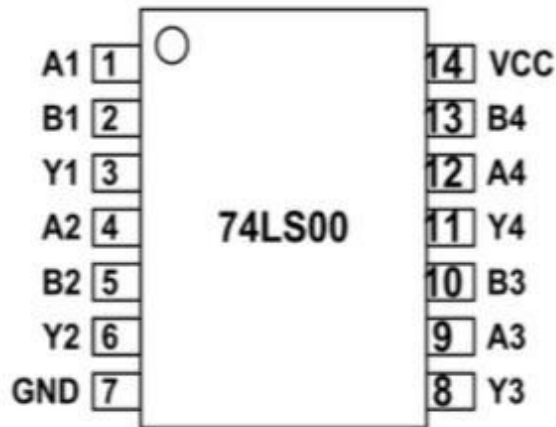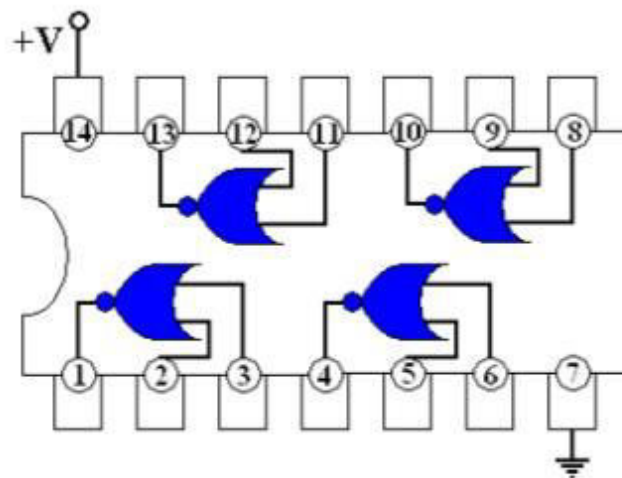
IC logic gate-NAND:

   7400 IC is a **Quad 2-Input NAND Gate** that contains four independent **gates** each of which performs the logic **NAND** function.

Pin description: Here A and B represents input pins and Y represents output pin. Pins 14 and 7 provide power for all four logic gates.
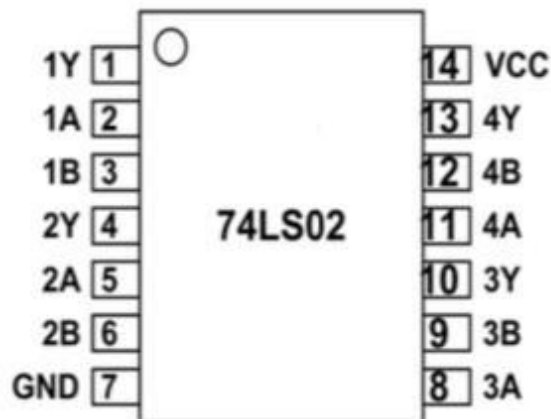
IC logic gate-NOR:

The **7402 Quad-2 input NOR gate IC** and contains four independent positive logic NOR GATES. Pins 14 and 7 provide power for all four logic gates.
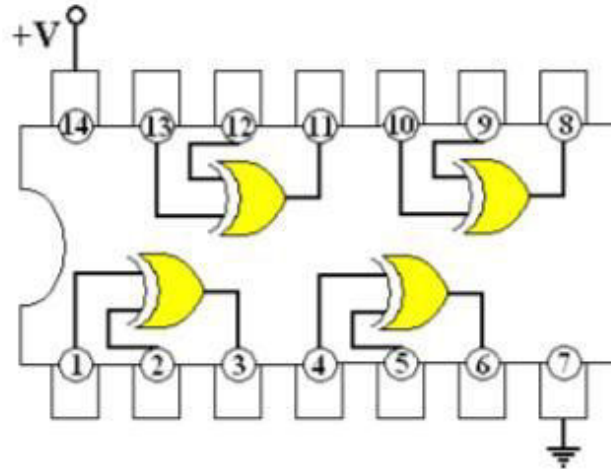

Pin Description: Here A and B represents input pins and Y represents output pin.

IC logic gate-XOR:

The 7486 Quad 2-input XOR gate and contains four independent **gates** each of which performs the logic XOR function.

Pin description: Here A and B represents input pins and Y represents output pin. Pins 14 and 7 provide power for all four logic gates.