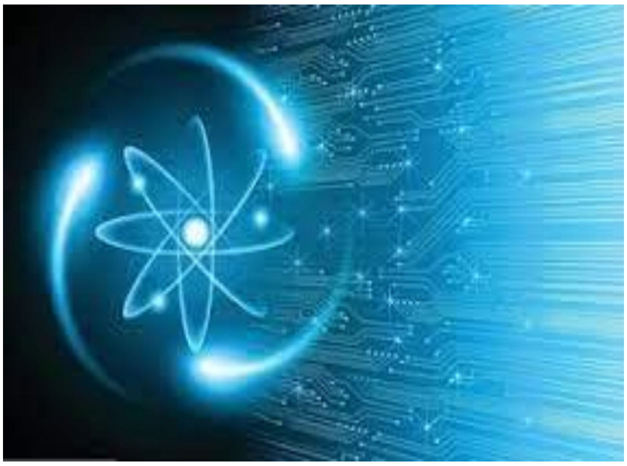


Semi Conductor Physics

B.Tech 1st Year

NOTES

Prepared By: TOPPERWORLD
LEARN & GROW

Unit - 1

CRYSTAL STRUCTURE

A lattice or a crystal lattice is the formation of an asymmetrical 3D structural arrangement of ions, atoms, and molecules to form a unit cell. The significant and distinct geometrical shape of a unit cell defines the type of crystal lattice.

The features of a crystal structure are:

- In a unit cell of a crystal structure, every ion, atom, or molecule represents a point in a 3D space. It means all the constituent particles in that unit can be considered points.
- Every point in a crystal structure is defined as a lattice point or lattice site.
- Two or more lattice points can be joined to form a straight line depending on the shape of a unit cell.
- When two or more straight lines are joined, it will form a 3D unit cell design representing a crystal structure. This arrangement of lattice points in a 3D space is called Bravais Lattices.

Based on their crystal structures, solids can be classified into the following categories:

1. ***Crystalline solids***
2. ***Amorphous solids***

However, crystalline solids can be further classified into molecular, ionic, metallic, and covalent solids. A brief introduction to the classification of solids is provided in this article.

Crystalline Solids

The solids featuring highly ordered arrangements of their particles (atoms, ions, and molecules) in microscopic structures are called crystalline solids.

These ordered microscopic structures make up a crystal lattice that accounts for the structure of the solid at any given point. Examples of crystalline solids include salt (sodium chloride), diamond, and sodium nitrate.

Amorphous Solids

The solids in which the particles are not arranged in any specific order or the solids that lack the overall order of a crystal lattice are called amorphous solids.

The term *'amorphous'*, when broken down into its Greek roots, can be roughly translated to *"without form"*. Many polymers are amorphous solids. Other examples of such solids include glass, gels, and nanostructured materials.

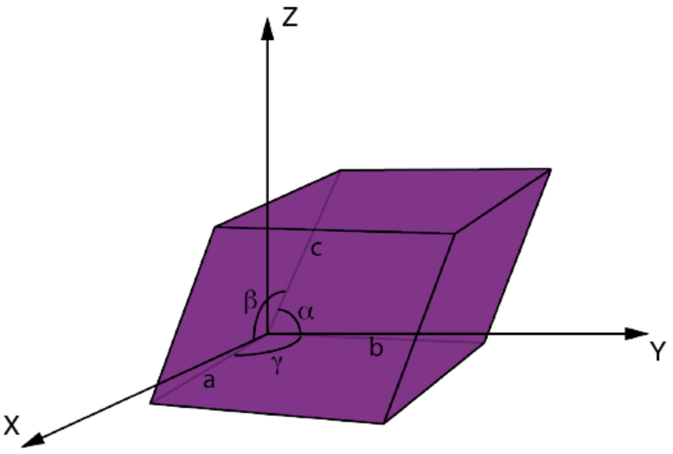
An ideal crystal is defined as an atomic arrangement that has infinite translational symmetry in all the three dimensions, whereas such a definite definition is not possible for an ideal amorphous solid (a-solid).

Unit Cell

The smallest unit of a solid crystal structure is called a unit cell. It can be defined as the structural unit of a solid crystal, and a lattice is generated by its repetition in a particular format.

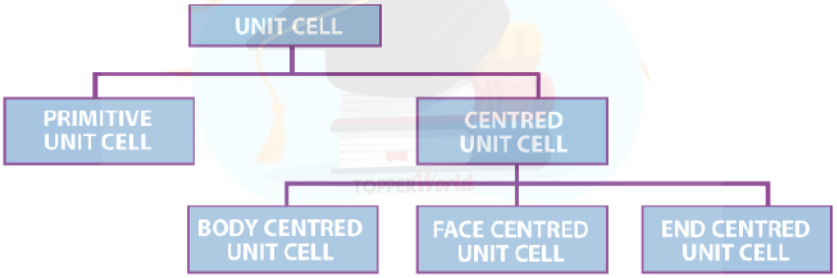
To identify a unit cell of a crystalline structure, the following parameters are considered.

- Three edges a, b, and c
- The angles between these edges ∝, β, and γ



The angles between these edges can vary. When high-energy electromagnetic waves are passed through a crystal, the unit cells absorb and send signals. These signals can be interpreted to understand the shape of a unit cell. Hence, the mutual arrangements of multiple unit cells can be identified, leading to the expression of a crystal structure.

Types of Unit Cell



1. Primitive Unit Cells

In this arrangement, the constituent lattice points hold the corner positions only.

The atoms in the primitive cubic unit cell are present only at the corners

Every atom at the corner is shared among eight adjacent unit cells

Four unit cells are present in the same layer

Four unit cell in the upper/lower layer

Therefore, a particular unit cell has the only 1/8th of an atom

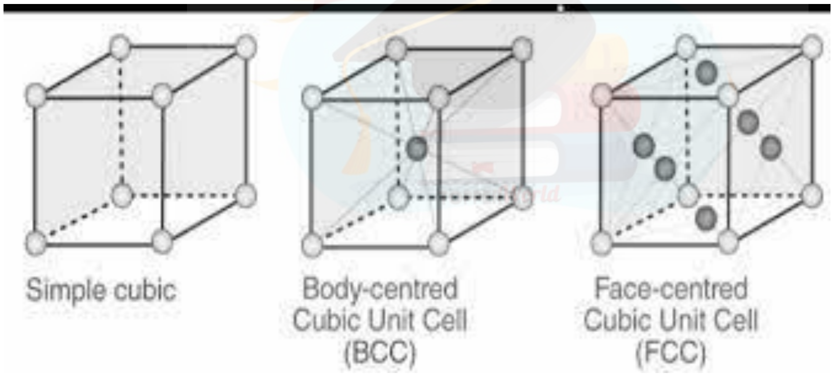
Each small sphere in the following figure represents the centre of a particle which occupies that particular position and not its size

- In each cubic unit cell, there are 8 atoms at the corners. Therefore, the total number of atoms in one unit cell is

8 × 1/8 = **1 atom**.

2. Centered Unit Cells

In this arrangement, the particles occupy the centre along with the corner positions. They are of the following types.



- **Body-Centred Cubic (BCC) Lattice structure**

This is a typical lattice structure where the atomic planes or lattice planes lie within the gaps of the lower planes of the respective atoms. In this type, the unit cell is shaped like a cube. It means that one atom is directly in contact with eight other atoms or lattice points. This type of arrangement can be found in vanadium, tungsten, and chromium.

In BCC unit cell every corner has atoms.

There is one atom present at the centre of the structure

Below diagram is an open structure

According to this structure atom at the body centres wholly belongs to the unit cell in which it is present.

Body centred Cubic (BCC) Unit Cell

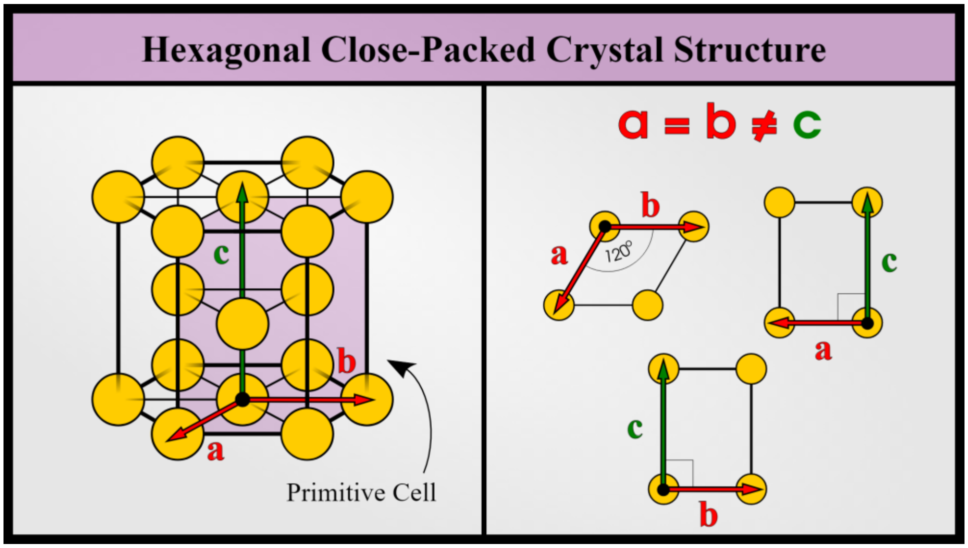
Number of Atoms in BCC Cell:
Thus, in a BCC cell, we have:

- 8 corners × 1/8 per corner atom = 8 × 1/8 = 1 atom
- 1 body centre atom = 1 × 1 = 1 atom

Therefore, the total number of atoms present per unit cell = **2 atoms.**

Hexagonal Closest Packed (HCP) Lattice Structure

The superposed atoms or lattice points create gaps where the underlying plane can fit into the upper one. The atomic levels are closely packed, creating a hexagonal closest packed (HCP) lattice structure. Metals like cobalt, zinc, titanium, and magnesium show such types of lattice arrangement. A point is connected with 12 neighbouring points making its coordination number 12.



- **Face-Centered Cubic (FCC) Lattice Structure**

This lattice structure also packs the maximum number of atomic planes in a stacking sequence. The second layer of atoms is primarily stacked following the HCP structure in the underlying layer gaps. The third layer, on the contrary, lies in the free gaps. The metals like copper, lead, and nickel shows such types of solid crystal lattice arrangements. The coordination number of this crystal structure is 12.

In FCC unit cell atoms are present in all the corners of the crystal lattice

Also, there is an atom present at the centre of every face of the cube

This face-centre atom is shared between two adjacent unit cells

Only 12 of each atom belongs to a unit cell

Number of Atoms in BCC Cell
a) 8 corners × 1/8 per corner atom = 8 × 1/8 = 1 atom
b) 6 face-centered atoms × 1/2 atom per unit cell = 3 atoms

Hence, the total number of atoms in a unit cell = 4 atoms

Thus, in a face-centred cubic unit cell, we have:

- 8 corners × 1/8 per corner atom = 8 × 1/8 = 1 atom

- 6 face-centred atoms × 1/2 atom per unit cell = 3 atoms

Therefore, the total number of atoms in a unit cell = **4 atoms**.

Types of Crystal Systems

1. Triclinic System

In this system, all three axes are inclined to each other and share the same length. Based on their inclination angles, the shape of the crystals varies. Example – kyanite, amazonite, feldspar, etc.

2. Monoclinic System

In this crystalline structure, there are three axes. Two of them are perpendicular to each other, whereas the 3rd one is inclined to form a different angle. The best examples of monoclinic systems are gypsum, petalite, diopside, etc.

3. Orthorhombic System

The crystal lattice structure of this type comprises three axes that are perpendicular to each other but the lengths of these axes vary. Examples of this lattice structure are topaz, iolite, zoisite, etc.

4.Trigonal Structure

The unit cell of this system is basically a three-sided pyramid. It can also attend to shapes like rhombohedral and scalenohedral. Examples are calcite, ruby, agate, etc.

5. Hexagonal System

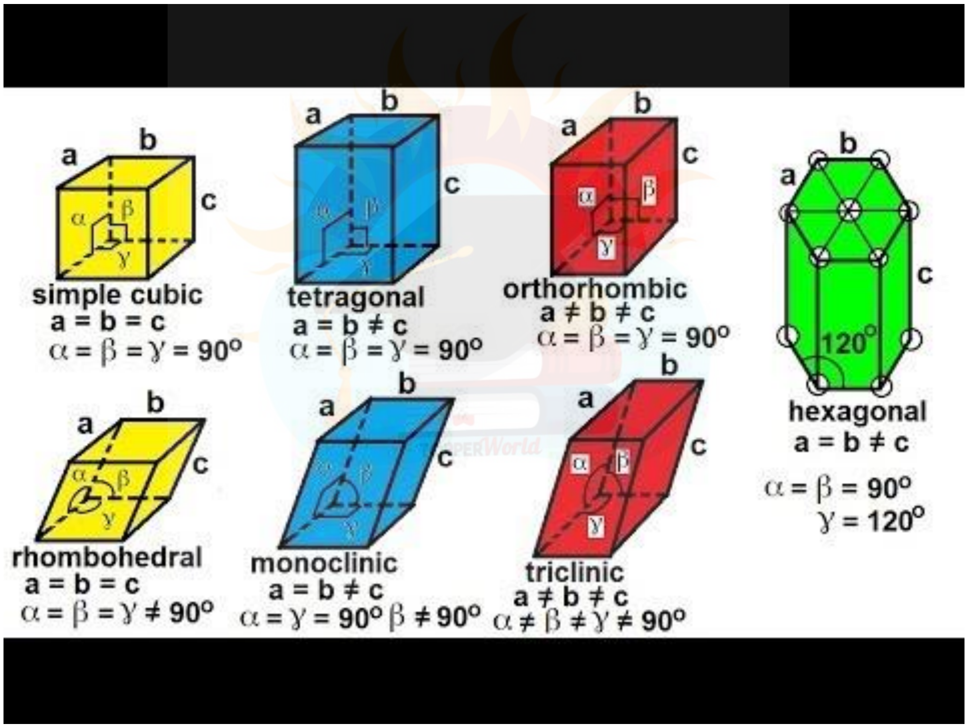
Now that we know what crystal structure is, this system seems a little different from the rest. It has four axes where three of them share the same length. The 4th one intersects these three at right angles. The examples are beryl, apatite, etc.

6. Tetragonal System

This system also consists of three axes. Only the main axis varies in dimension. The other two are of the same length. An example of this system is perovskite.

7. Cubic System

This is a lattice system where all three angles are at right angles to each other with the same length. This crystal structure of metals can be found in gold and silver. Among nonmetals, diamond shows this characteristic crystal shape.



Lattice

A **lattice** is in general defined as a discrete but infinite regular arrangement of points (**lattice sites**) in a vector space.

Bravais Lattice

In solid state physics one usually encounters lattices which exhibit a discrete translational symmetry. If one considers for instance the vector space R^3 this means that a translation of the whole lattice by any translation vector given by

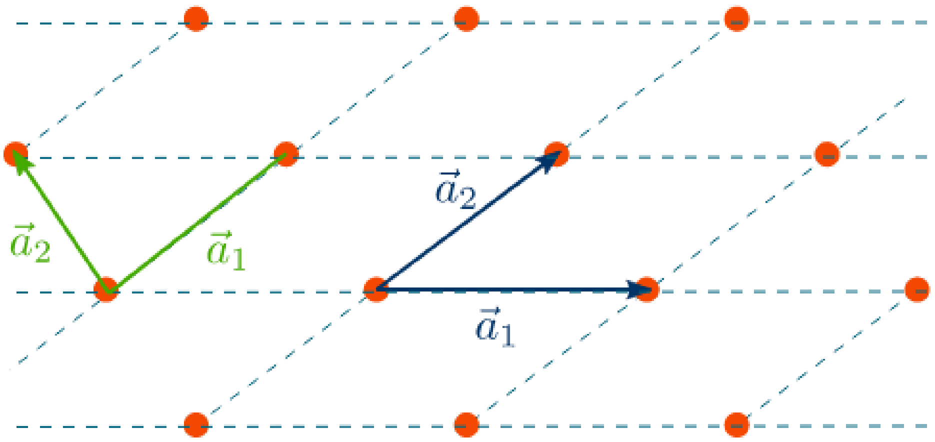
$$Tmno=ma1+na2+oa3 \qquad\qquad m,n,o\in Z$$

A lattice that can be characterised in this way is referred to as a Bravais lattice. All lattice points are equivalent, i.e. all properties remain invariant under translations by any vector →T_{mno}. Also, an observer sitting on one specific lattice point would see the same environment as when sitting on any other. However, if there are lattice points with different environments, they cannot form a Bravais lattice! (This is an equivalent definition of a Bravais lattice.

Primitive Translation Vectors

The linear independent generating vectors →ai of the translations are called **primitive translation vectors** or primitive lattice vectors. However, for a given lattice the choice of the primitive translation vectors is not unique.

Let's illustrate this for the two-dimensional Euclidean space: Any two linear independent vectors →a1 and →a2form a basis and can be used as generators for a two-dimensional lattice:



Basis

The crystal structure is formed by associating every lattice point with an assembly of atoms or molecules or ions, which are identical in composition, arrangement and orientation, is called as the basis. The atomic arrangement in a crystal is called crystal structure. If the basis is substituted for the lattice points, then the resulting structure is called crystal structure.

Thus lattice + basis = crystal structure.

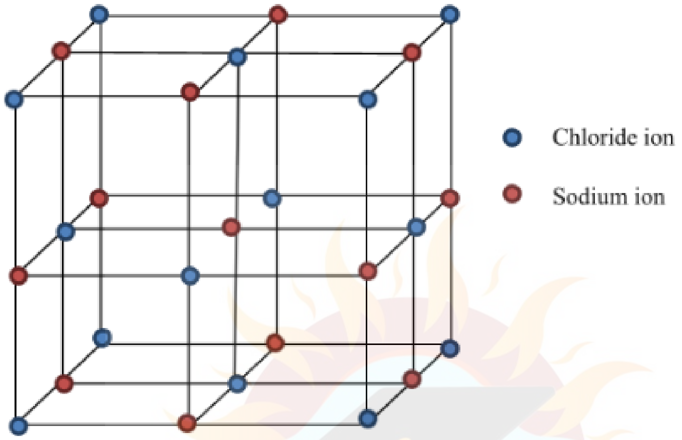
NaCl(Sodium Chloride)

$NaCl$ has a cubic unit cell. It is best thought of as an FCC array of anions with an inter penetrating FCC cation lattice. The cell looks the same whether you start with anions or cation on the corners. Each ion is 6$6$ coordinate and has a local quadrilateral geometry.

The unit cell of NaCl consists of $Na+$ ions and $Cl-$ ions. There are four types of site-unique central position, face site, edge sites and corner site which are used to determine the number of $Na+$ and $Cl-$ ions in the unit cell of NaCl. When counting the number of ions, a corner site would be shared by 7 other unit cells.

Therefore, 1 corner would be 18 of an ion. For a face site, it is shared by other unit cells and for an edge site, the ion is shared by 3 other unit cells. NaCl is a FCC unit cell which has four cations and four anions. This can be shown by counting the number of ions and multiplying them in relation to their position.

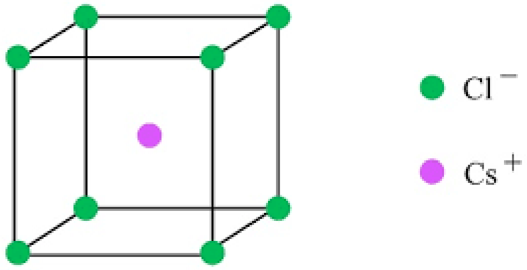
Na+: 1center + 12edge×1/4=4
Cl− : 4face×12+8corner×1/8=4



In addition to this, NaCl also known as salt or halite. A halite is an ionic compound with the chemical formula NaCl , representing a *1:1* ratio of sodium and chloride ions. Sodium chloride salt occurs in oceans and sea waters. It is also formed as Rock salt.

CsCl(Cesium Chloride)

CsCl crystallize in a primitive cubic lattice which means the cubic unit cell has nodes only at its corners. The structure of CsCl can be seen as two interpenetrating cubes, one of Cs^+ and one of Cl^-. The ions are not touching one another. Touching would cause repulsion between the anion and cation. Some may mistake the structure type of CsCl with NaCl, but really the two are different. CsCl is more stable than NaCl, for it produces a more stable crystal and more energy is released.



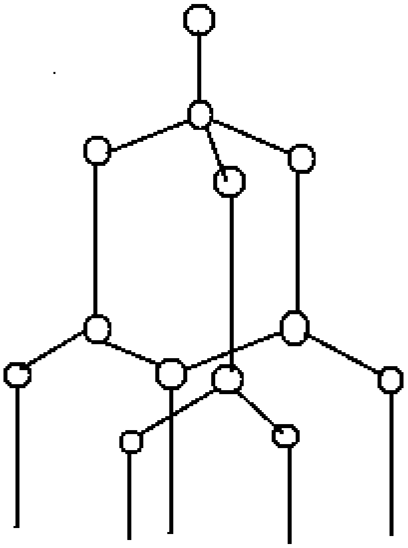
Unit cell - Cesium Chloride

Anions and cations have similar sizes. Each Cs^+ is surrounded by 8 Cl^- at the corners of its cube and each Cl^- is also surrounded by 8 Cs^+ at the corners of its cube. The cations are located at the center of the anions cube and the anions are located at the center of the cations cube. There is one atom in CsCl. To determine this, the following equation is given:

8 Corners of a given atom x 1/8 of the given atom's unit cell = 1 atom

Diamond

Diamond is a crystal structure with a face centred cubic space lattice and two identical atoms in the basis. One atom sits on the lattice point whereas the other one is shifted by 14 along each axis. This forms a tetrahedral structure where each atom is surrounded by four equal-distanced neighbours. Each atom forms bonds with four nearest neighbours. The enclosed angles between two angles are 109.47°. The distance between the two atoms equals one-quarter of the body diagonal of the cube. The diamond lattice contains also 4 Bravais points per unit cell but contains 8 atoms per unit cell tetrahedrally bonded atoms in each primitive cell.



Face-centred cubic lattice (fcc or cubic-F), like all lattices, has lattice points at the eight corners of the unit cell plus additional points at the centres of each face of the unit cell. It has unit cell vectors a=b=c

and interaxial angles α=β=γ=90. Whereas Body-centred cubic lattice (bcc or cubic-I), like all lattices, has lattice points at the eight corners of the unit cell plus an additional point at the centre of the cell. It has unit cell vectors a=b=c and interaxial angles α=β=γ=90o.

Point Defects

Point defects explain about the **imperfections of solids** along with the types of point defects. Crystalline solids are formed by joining many small crystals. Different types of defects are found in crystals after the process of crystallization.

Point defects are accounted for when the crystallization process occurs at a very fast rate. These defects mainly happen due to deviation in the arrangement of constituting particles. In a crystalline solid, when the ideal arrangement of solids is distorted around a point/ atom it is called a point defect.

Defects or Imperfections in crystalline solid can be divided into four groups namely ***line defects, point defects, volume defects and surface defects.*** Historically, crystal point defects were first regarded in ionic crystals, not in metal crystals that were much simpler.

There are 3 types of point defects:

3. **Stoichiometric defect**
4. **Frenkel defect**
5. **Schottky defect**

1. Stoichiometric Defect:

In this kind of point defect, the ratio of positive and negative ions (Stoichiometric) and electrical neutrality of a solid is not disturbed. Sometimes it is also known as intrinsic or thermodynamic defects.

Fundamentally, they are of two types:

- *Vacancy defect:* When an atom is not present at their lattice sites, then that lattice site is vacant and it creates a vacancy defect. Due to this, the density of a substance decreases.
- *Interstitial defect:* It is a defect in which an atom or molecule occupies the intermolecular spaces in crystals. In this defect, the density of the substance increases.

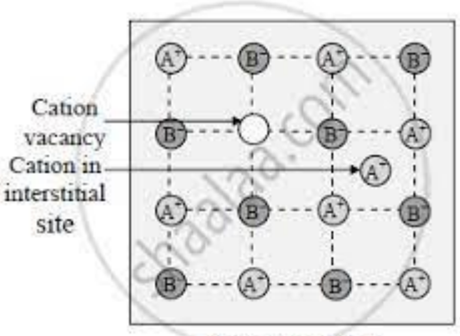
A non-ionic compound mainly shows vacancy and interstitial defects. An ionic compound shows the same in Frenkel and Schottky defect.

2. Frenkel Defect:

In ionic solids generally, the smaller ion (cation) moves out of its place and occupies an intermolecular space. In this case, a vacancy defect is created on its original position and the interstitial defect is experienced at its new position.

- It is also known as dislocation defect.
- The density of a substance remains unchanged.

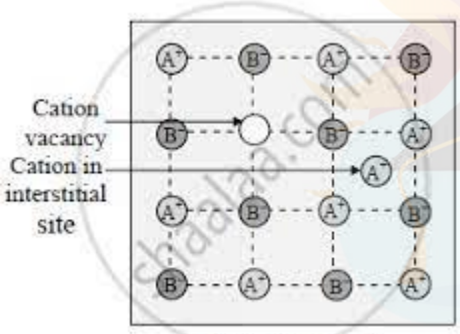
- It happens when there is a huge difference in the size of anions and cations.
- ***Example:*** ZnS and AgCl.



Frenkel defect

3. **Schottky Defect**

- This kind of vacancy defects is found in Ionic Solids. But in ionic compounds, we need to balance the electrical neutrality of the compound so an equal number of anions and cations will be missing from the compound.
- It reduces the density of the substance.
- In this, the size of cations and anions are of almost the same.



Frenkel defect

1 Impurity Defect: Let's understand the impurity defect by an example. If molten NaCl is crystallized with $SrCl_2$ compound then the Sr^{2+} ions replace two Na^+ ions and occupy the place of one Na^+ In this way the lattice site of one Na^+ is vacant and it creates an impurity defect.

2 Non-Stoichiometric Defect: In this defect, the cations and anions ratio is disturbed either because of adding or removing of ions.

Types of Non-Stoichiometric Defect:

Metal deficiency defect: In this, the solids have less number of metals relative to the described Stoichiometric proportion.

Metal excess defect: There are two types of metal excess defect:

- *Metal excess defect due to anionic vacancies:* This occurs due to the absence of anions from its original lattice site in crystals. Therefore, instead of anions, electrons occupy their position
- *Metal excess defect due to the presence of extra cations at interstitial sites:* Here, on heating the compound, it releases extra cations. These cations occupy the interstitial sites in crystals and the same number of electrons goes to neighbouring interstitial sites.

There is so much to learn about the crystalline solids, its structure and its defects. This is just the overview of point defects in solids.

UNIT - 2

QUANTUM THEORY

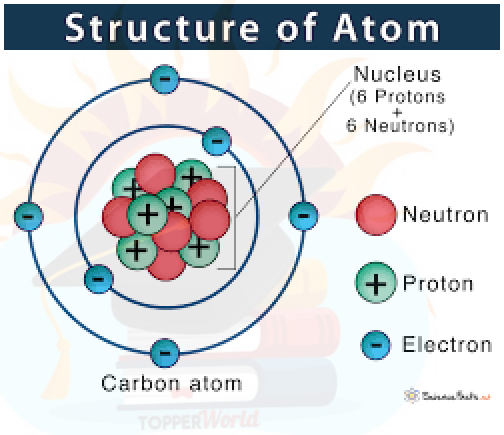
INTRODUCTION

Quantum mechanics is a physical science dealing with the behaviour of matter and energy on the scale of atoms and subatomic particles or waves.

The term "quantum mechanics" was first coined by Max Born in 1924. The acceptance by the general physics community of quantum mechanics is due to its accurate prediction of the physical behaviour of systems, including systems where Newtonian mechanics fails.

Structure of atom

Atoms consist of an extremely small, positively charged nucleus surrounded by a cloud of negatively charged electrons. Although typically the nucleus is less than one ten-thousandth the size of the atom, the nucleus contains more than 99.9% of the mass of the atom.



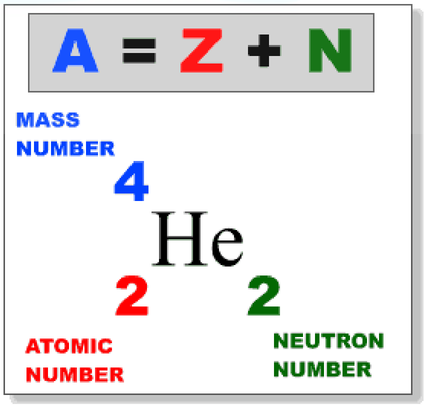
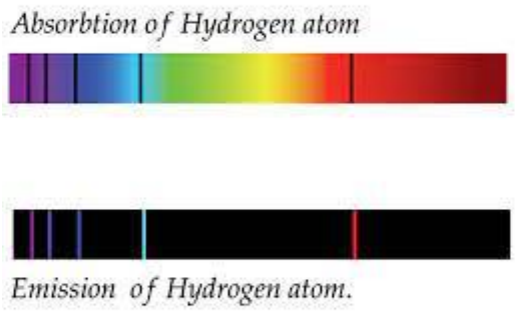


Figure 2

Hydrogen Spectrum

We all know that electrons in an atom or a molecule absorb energy and get excited, they jump from a lower energy level to a higher energy level, and they emit radiation when they come back to their original states. This phenomenon accounts for the emission spectrum through hydrogen too, better known as the hydrogen emission spectrum.



The hydrogen spectrum is an important piece of evidence to show the quantized electronic structure of an atom. The hydrogen atoms of the molecule dissociate as soon as an electric discharge is passed through a gaseous hydrogen molecule. It results in the emission of electromagnetic radiation initiated by the energetically excited hydrogen atoms.

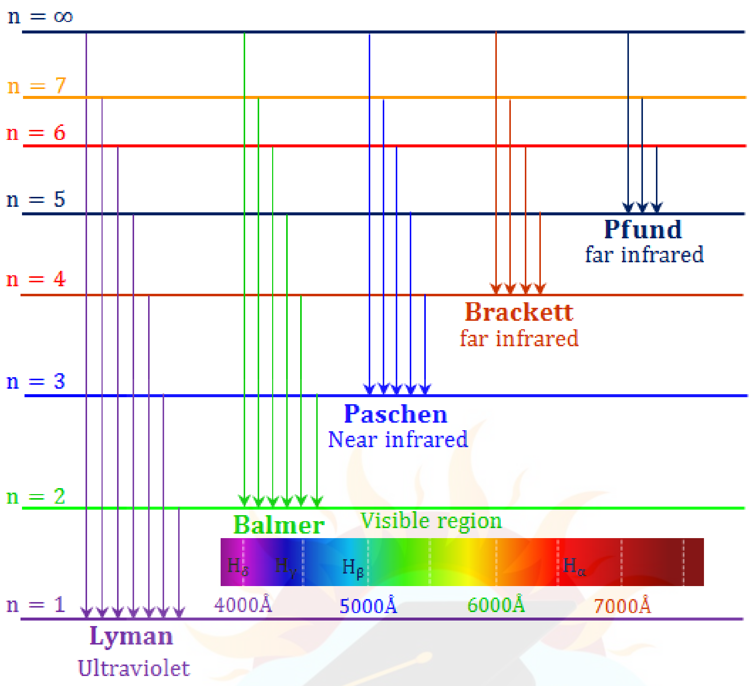
Hydrogen spectrum wavelength

When a hydrogen atom absorbs a photon, it causes the electron to experience a transition to a higher energy level, for example, n = 1, n = 2. When a photon is emitted through a hydrogen atom, the electron undergoes a transition from a higher energy level to a lower, for example, n = 3, n = 2. During this transition from a higher level to a lower level, there is the transmission of light occurs. The quantized energy levels of the atoms cause the spectrum to comprise wavelengths that reflect the differences in these energy levels. For example, the line at 656 nm corresponds to the transition n = 3 n = 2.

Hydrogen emission spectrum:

In the year 1885, on the basis of experimental observations, Balmer proposed the formula for correlating the wave number of the spectral lines emitted and the energy shells involved. This formula is given as:

$$\bar{\nu} = 109677\left(\frac{1}{2^2} - \frac{1}{n^2}\right)$$



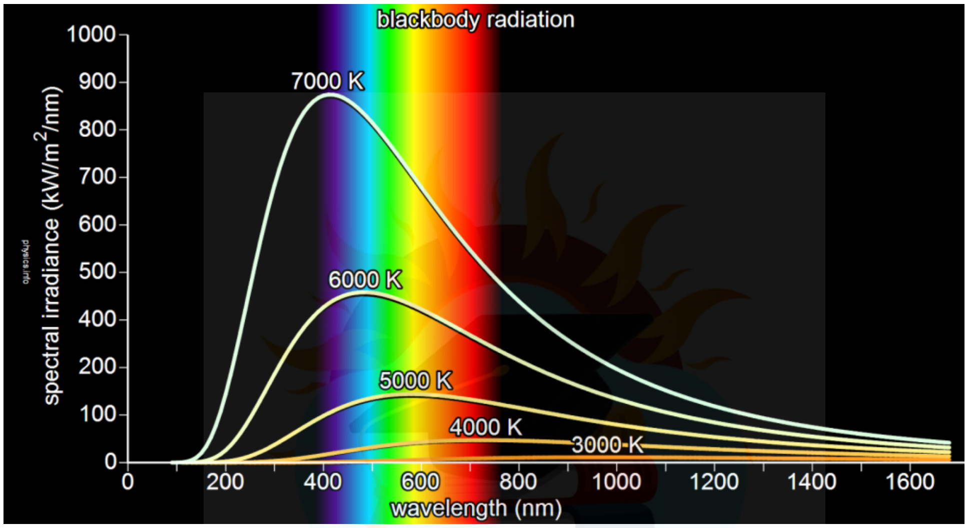
- The transition from the first shell to any other shell – Lyman series
- The transition from the second shell to any other shell – Balmer series
- The Transition from the third shell to any other shell – Paschen series
- The transition from the fourth shell to any other shell – Brackett series
- The transition from the fifth shell to any other shell – Pfund series

Black Body

A black body is an idealization in physics that pictures a body that absorbs all electromagnetic radiation incident on it irrespective of its frequency or angle. In this article, we will be learning about blackbody radiation and some important laws related to it.

Black Body Radiation

To stay in thermal equilibrium, a black body must emit radiation at the same rate as it absorbs, so it must also be a good emitter of radiation, emitting electromagnetic waves of as many frequencies as it can absorb, i.e., all the frequencies. The radiation emitted by the blackbody is known as **blackbody radiation**.



Characteristics of Blackbody Radiation

The characteristics of the blackbody radiation are explained with the help of the following laws:

- Wien’s displacement law
- Planck’s law
- Stefan-Boltzmann law

Wien's Displacement Law

Wien's displacement law states that

The blackbody radiation curve for different temperature peaks at a wavelength is inversely proportional to the temperature.

Wien's Law Formula

Wien's Law Formula	$\lambda_{\text{max}} = \frac{}{T}$	• T is the temperature in kelvins • b is the Wien's displacement constant = $2.8977*10^3$ m.K

Planck's Law

Using Planck's law of blackbody radiation, the spectral density of the emission is determined for each wavelength at a particular temperature.

Read More: Planck's Law

Planck's Law Formula

Planck's law	c = speed of light = $2.997925 * 10^{\circ}\ m / sec$ λ = wavelength (m) k = Boltzmann's constant = $1.381 * 10^{-23}\ J/K$ T = temperature (K)	• E_λ is the wavelength • T is the absolute temperature

Stefan-Boltzmann Law

The Stefan-Boltzmann law explains the relationship between total energy emitted and the absolute temperature.

Stefan-Boltzmann Law Formula

Stefan-Boltzmann Law	$E \propto T^4$	• E is the total energy emitted • T is the absolute temperature

DUAL NATURE OF LIGHT

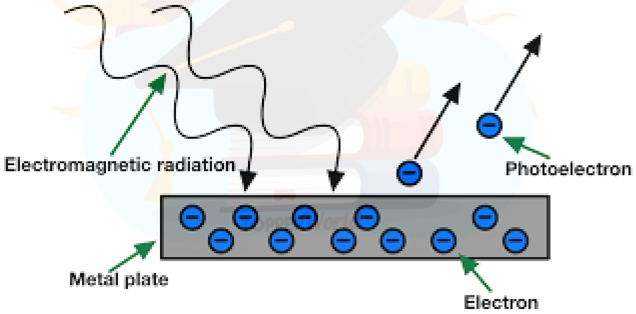
There are some phenomena such as interference, diffraction and polarization which can be explained by considering light as wave only.

On the other hand, phenomenon such as photoelectric effect and Compton Effect can be explained by considering light as a particle only.

When we visualize light as a wave, we need to forget its particle aspect completely and vice versa. This type of behavior of light as a wave as well as particle is known as dual nature of light.

Photo electric effect

The photoelectric effect is a phenomenon in which electrons are ejected from the surface of a metal when light is incident on it. These ejected electrons are called **photoelectrons**. It is important to note that the emission of photoelectrons and the kinetic energy of the ejected photoelectrons is dependent on the frequency of the light that is incident on the metal's surface. The process through which photoelectrons are ejected from the surface of the metal due to the action of light is commonly referred to as **photoemission**.



The photoelectric effect occurs because the electrons at the surface of the metal tend to absorb energy from the incident light and use it to overcome the attractive forces that bind them to the metallic nuclei

Einstein's theory of photoelectric effect

When a photon of energy $h\upsilon$ is incident on the surface of the metal, a part of energy Φ is used in liberating the electron from the metal. This energy is known as the work function of the metal. The rest of energy is given to the electron so that is acquires kinetic energy ½ mv2. Thus, a photon of energy $h\upsilon$ is completely absorbed by the emitter.

Energy of photon = Energy needed to liberate the electron + Maximum K.E of the liberated electron
$h\upsilon = \Phi + KEmax$

$$h\upsilon = \Phi + \frac{1}{2} mv2 \text{ max}$$

The above equation is called Einstein's photoelectric equation. This equation can explain all the features of the photoelectric effect.

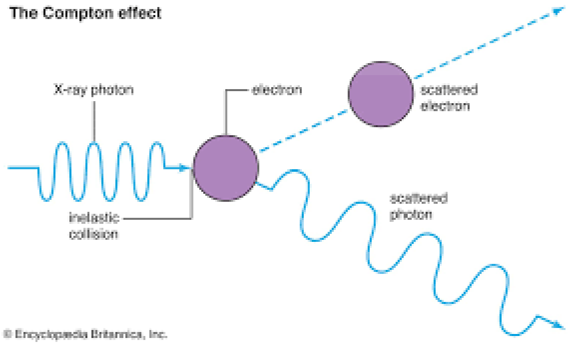
Compton Effect

When a beam of high frequency radiation (x-ray or gammaray) is scattered by the loosely bound electrons present in the scatterer, there are also radiations of longer wavelength along with original wavelength in the scattered radiation. This phenomenon is known as Compton Effect.

When a photon of energy $h\nu$ collides with the electron, some of the energy is given to this electron. Due to this energy, the electron gains kinetic energy and photon loses energy. Hence scattered photon will have lower energy $h\nu'$ that is longer wavelength than the incident one.

$(\lambda' - \lambda) = h/mc\ [1-\cos\Phi]$ where $h/mc = \lambda C$ = Compton wavelength = 0.02424Å

The Compton effect



© Encyclopaedia Britannica, Inc.

De Broglie hypothesis

Louis De Broglie a French Physicist put forward his bold ideas like this

“Since nature loves symmetry, if the radiation behaves as a particle under certain circumstances and waves under other circumstances, then one can even expect that entities which ordinarily behave as particles also exhibit properties attributable to waves under appropriate circumstance and those types of waves are termed as matter waves.

All matter can exhibit wave-like behavior. For example, a beam of electrons can be diffracted just like a beam of light or a water wave. The concept that matter behaves like a wave was proposed by Louis de Broglie in 1924. It is also referred to as the de Broglie hypothesis of matter waves. On the other hand de Broglie hypothesis is the combination of wave nature and particle nature.

If ‘ E ’ is the energy of a photon of radiation and the same energy can be written for a wave as follows

E = mc2 ---(1) (particle nature) and E = hv = hc/λ ---(2) (wave nature)

Comparing eqns (1) & (2) we get

mc 2 = hc/λ or λ = h/mc = h/p λ = h/p ;

where λ = De Broglie wavelength

Particles of the matter also exhibit wavelike properties and those waves are known as matter waves.

Expression for de Broglie wavelength of an accelerated electron

De Broglie wavelength for a matter wave is given by

λ = h/p ; where λ = De Broglie wavelength -------------(1)

From eqn. (1) we find that,

if the particles like electrons are accelerated to various velocities, we can produce waves of various wavelengths. Thus higher the electron velocity, smaller will be the de-Broglie wavelength. If velocity v is given to an electron by accelerating it through a potential difference V, then the work done on the electron is eV. This work done is converted to kinetic energy of electron. Hence, we can write

$$\frac{1}{2} mv2 = eV$$

$$mv = (2meV)1/2 \text{ -------------}(2)$$

But eqn.(1) can be written as $\lambda = h/mv$ -------------------(3)

Substituting eqn.(2) in eqn.(3) we get

$$\lambda = h/(2meV)^{1/2}$$

PROPERTIES OF MATTER WAVES:

1. The wavelength of a matter wave is inversely related to its particles momentum.
2. Matter wave can be reflected, refracted, diffracted and undergo interference
3. The position and momentum of the material particles cannot be determined accurately and simultaneously.
4. The amplitude of the matter waves at a particular region and time depends on the probability of finding the particle at the same region and time.

Wave packet:

Two or more waves of slightly different wavelengths alternately interfere and reinforce so that an infinite succession of groups of waves or wave packets are produced.

The velocity of the individual wave in a wave packet is called phase velocity of the wave and is represented by Vp.

Phase and Group velocities:

According to de Broglie each particle of matter (like electron, proton, neutron etc) is associated with a de Broglie wave; this de Broglie wave may be regarded as a wave packet, consisting of a group of waves. A number of frequencies mixed so that the resultant wave has a beginning and an end forms the group. Each of the component waves propagates with a definite velocity called wave velocity or phase velocity.

Expression for Phase velocity:

A wave can be represented by $Y = A \sin(\omega t - kx)$ ---------- (1)

Where $k = \omega/v$ = wave number (rad/m) ;

ω = Angular frequency (rad/s)

When a particle moves around a circle ν times/s, sweeps out $2\pi\nu$ rad/s

In eqn.(1) the term $(\omega t - kx)$ gives the phase of the oscillating mass

$(\omega t - kx)$ = constant for a periodic wave

$d(\omega t - kx)/dt = 0$ or $\omega - k(dx/dt) = 0$ or $dx/dt = \omega/k$

$$v_p = \omega/k$$

When a wave packet or group consists of a number of component waves each traveling with slightly different velocity, the wave packet (group) travels with a velocity different from the velocities of component waves of the group; this velocity is called Group velocity.

Expression for Group velocity

A wave group can be mathematically represented by the superposition of individual waves of different wavelengths. The interference between these individual waves results in the variation of amplitude that defines the shape of the group. If all the waves that constitute a group travel with the same velocity, the group will also travel with the same velocity.

If however the wave velocity is dependent on the wavelength the group, velocity will be different from the velocity of the individual waves.

The simplest wave group is one in which two continuous waves are superimposed. Let the two waves be represented by

$y_1 = a \cos(\omega_1 t - k_1 x)$ and $y_2 = a \cos(\omega_2 t - k_2 x)$

The resultant

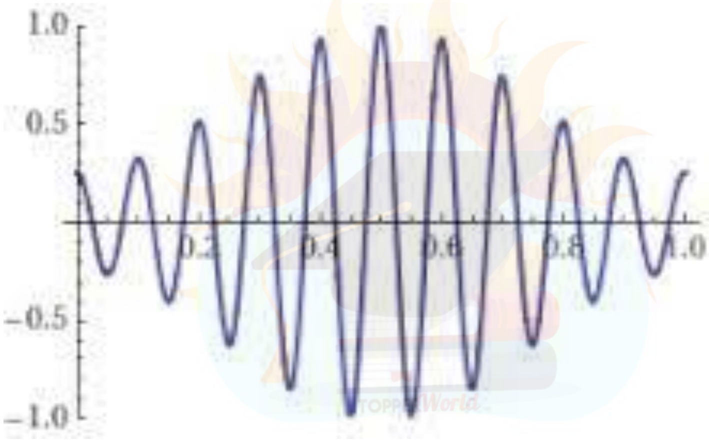
$y = y_1 + y_2 = a \cos(\omega_1 t - k_1 x) + a \cos(\omega_2 t - k_2 x)$

$$v_g = \frac{\partial \omega}{\partial k}$$

$$\omega = \omega_m \sin\left(\frac{ka}{2}\right); \quad \frac{\partial \omega}{\partial k} = \omega_m \frac{a}{2} \cos\left(\frac{ka}{2}\right)$$

In the long wavelength limit (k→0)

$$v_g = \frac{\partial \omega}{\partial k} = \omega_m \frac{a}{2} = v_p$$



Two or more waves of slightly different wavelengths alternately interfere and reinforce so that an infinite succession of groups of waves or wave packets are produced. The de Broglie wave group associated with a particle travel with a velocity equal to the particle velocity.

Principle of complimentarity

The experiment of Davisson & Germer demonstrated the diffraction of electron beams. The wave nature of electrons can also be demonstrated by interference with a double slit.

But it is an extremely difficult task to prepare a suitable double slit that can transmit an electron beam. But the experiment was done by Jönson in 1961. He passed a 50,000eV beam of electron through a double slit.

The pattern obtained by him was very similar to the interference pattern obtained by Young with visible light. In an experiment of the above type it is rather tempting to try to find out through which slit an electron has passed.

If we design a suitable device for detecting the passage of an electron through one of the slits, the interference pattern is found to vanish. If the electron is to behave like a classical particle, it has to pass through one of the two slits.

On the other hand, if it is a wave, it can pass through both the slits! When we try observing the passage of electron through one of the slits, we are examining its particle aspect.

However, when we observe the interference pattern, we are investigating the wave aspect of electron. At a given moment and under given circumstances the electron will behave either as a particle or as a wave but not as both.

In other words, the particle and wave nature of a physical entity cannot be observed simultaneously.

Heisenberg's Uncertainty principle.

Physical quantities like position, momentum, time, energy etc. can be measured accurately in macroscopic systems (i.e., classical mechanics). However, in the case of microscopic systems, the measurement of physical quantities for particles like electrons, protons, neutrons, photons etc are not accurate. If the measurement of one is certain and that of other will be uncertain.

A wave packet that represents and symbolizes all about the particle and moves with a group velocity describes a de Broglie wave. According to Bohr'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 the wave packet. As the wave packet has a velocity spread, there is an uncertainty about the momentum of the particle. Thus, according to uncertainty principle states that the position and the momentum of a particle in an atomic system cannot be determined simultaneously and accurately. If Δx is the uncertainty associated with the position of a particle and Δpx the uncertainty associated with its momentum, then the product of these uncertainties will always be equal or greater than h/4π. That is

$$\Delta x \ \Delta px \geq h/4\pi$$

Different forms of uncertainty principle

$$\Delta E \ \Delta t \geq h/4\pi$$

$$\Delta \omega \ \Delta \theta \geq h/4\pi$$

Applications Heisenberg's Uncertainty principle (Nonexistence of electron in the nucleus)

The radius 'r' of the nucleus of any atom is of the order of 10-14m so that if an electron is confined in the nucleus, the uncertainty in its position will be of the order of 2r = Δx (say) i.e diameter of the nucleus

But according to HUP

Δx Δp ≥ h/4π (Δp = uncertainty in momentum)

Δx ~ 2x10-14 m

Therefore,

Δp = h/(4π Δx) = 6.625E-34 / (4π x 2x10-14) = 2.63 x 10-21 kg-m/s

Taking Δp ~ p we can calculate energy using the formula

E ^2 = c ^2 (p2 + mo ^2 c ^2) = (3x108) 2 x [(2.63 x 10-21) 2 + (9.1x10-31) 2 x (3x108) 2] = 7.932x10-13J = 4957745 eV ~ 5 MeV

However, the experimental investigations on beta decay reveal that the kinetic energies of electrons must be equal to 4MeV. Since there is a disagreement between theoretical and experimental energy values we can conclude that electrons cannot be found inside the nucleus.

Wave function (ψ)

Water waves ------------ height of water surfaces varies

Light waves -------------- electric & magnetic fields vary

Matter waves -------------- wave function (ψ)

- Ψ is related to the probability of finding the particle. Max Born put these ideas forward for the first time
- The wave function ψ indicates the state of the particle. However it has no direct physical significance. There is a simple reason why ψ cannot be interpreted in terms of an experiment. The probability that something be in a certain place at a given time must lie between 0 & 1 i.e. the object is definitely not there and the object is definitely there respectively.
- An intermediate probability, say 0.2, means that there is a 20% chance of finding the object. However, the amplitude of a wave can be negative as well as positive and a negative probability -0.2 is meaningless. Hence ψ by itself cannot be an observable quantity.
- Because of this the square of the absolute value of the wave function ψ is considered and is known as probability density denoted by | ψ |^2
- The probability of experimentally finding the body described by the wave function ψ at the point x, y, z at the time t is proportional to the value of | ψ |^2 .

A particle of mass 'm' traveling in the increasing x- direction with no force acting on it is called a free particle.

According to Schrodinger the wave function ψ(x,t) for a free particle moving in the positive x direction is given by

$\psi(x,t) = \psi o\ e\ i(kx - \omega t)$, here ψo = amplitude and $\psi(x,t)$ = complex

Schrodinger Wave Equation

Schrodinger wave equation is a mathematical expression describing the energy and position of the electron in space and time, taking into account the matter wave nature of the electron inside an atom.

It is based on three considerations. They are;

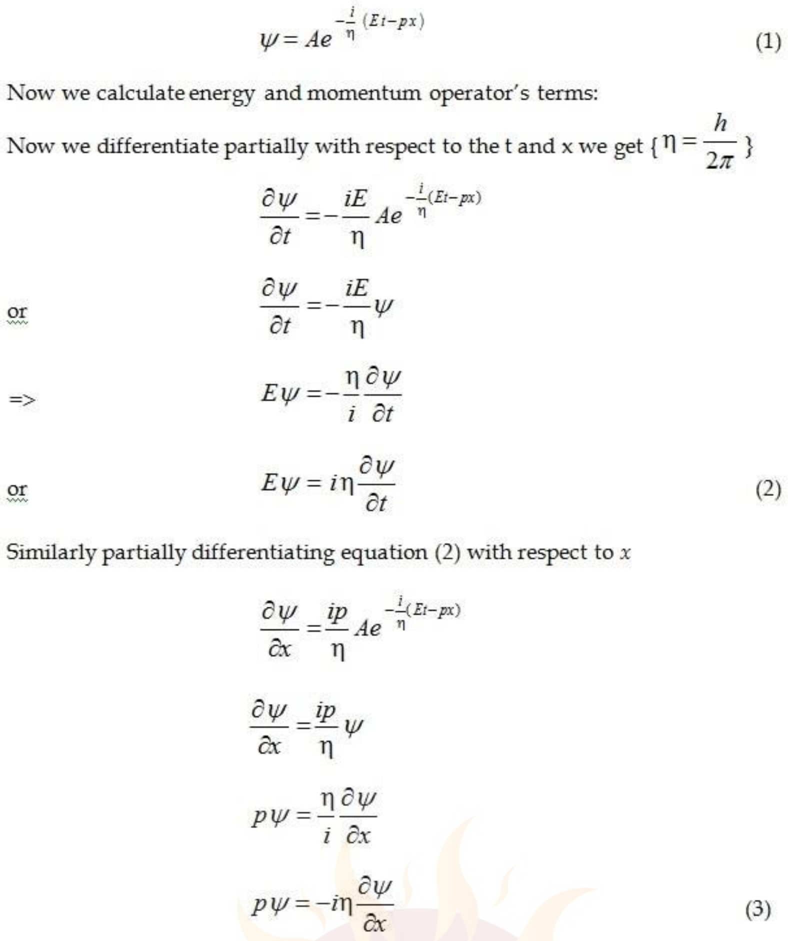
- Classical plane wave equation,
- Broglie's Hypothesis of matter-wave, and
- Conservation of Energy.

Schrodinger equation gives us a detailed account of the form of the wave functions or probability waves that control the motion of some smaller particles. The equation also describes how these waves are influenced by external factors. Moreover, the equation makes use of the energy conservation concept that offers details about the behaviour of an electron that is attached to the nucleus.

There are two equations, which are time-dependent Schrödinger equation and a time-independent Schrödinger equation.

Time independent one dimensional Schrodinger wave equation

consider a particle of mass *m* moving in the positive *x-direction*. The potential energy of the particle is *V*, momentum is *p* and total energy is *E*. So the *free particle wave equation* is



$$\psi = Ae^{-\frac{i}{\eta}(Et-px)} \qquad (1)$$

Now we calculate energy and momentum operator's terms:

Now we differentiate partially with respect to the t and x we get $\{\eta = \frac{h}{2\pi}\}$

$$\frac{\partial\psi}{\partial t} = -\frac{iE}{\eta} Ae^{-\frac{i}{\eta}(Et-px)}$$

or

$$\frac{\partial\psi}{\partial t} = -\frac{iE}{\eta}\psi$$

=>

$$E\psi = -\frac{\eta}{i}\frac{\partial\psi}{\partial t}$$

or

$$E\psi = i\eta\frac{\partial\psi}{\partial t} \qquad (2)$$

Similarly partially differentiating equation (2) with respect to x

$$\frac{\partial\psi}{\partial x} = \frac{ip}{\eta} Ae^{-\frac{i}{\eta}(Et-px)}$$

$$\frac{\partial\psi}{\partial x} = \frac{ip}{\eta}\psi$$

$$p\psi = \frac{\eta}{i}\frac{\partial\psi}{\partial x}$$

$$p\psi = -i\eta\frac{\partial\psi}{\partial x} \qquad (3)$$

The total energy is the sum of kinetic energy and potential energy; so the total energy of the particle is

$$E = K.E + P.E$$

$$E = \frac{1}{2}mv^2 + V$$

or
$$E = \frac{1}{2}\frac{m^2}{m}v^2 + V$$

=>
$$E = \frac{P^2}{2m} + V$$

on multiplying both sides by ψ

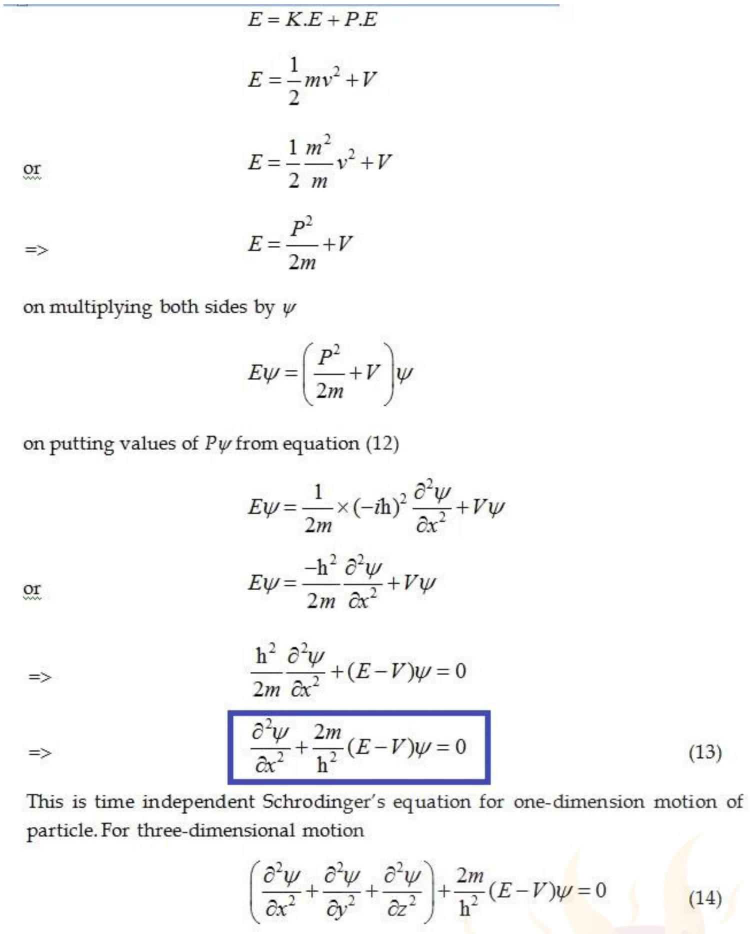
$$E\psi = \left(\frac{P^2}{2m} + V\right)\psi$$

on putting values of $E\psi$ and $P\psi$ from equation (7) and (8)

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\times\frac{\partial^2\psi}{\partial x^2} + V\psi \qquad (9)$$

In three-dimensional form above equation can be written as

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + V\psi \qquad (10)$$



$$E = K.E + P.E$$

$$E = \frac{1}{2}mv^2 + V$$

or $$E = \frac{1}{2}\frac{m^2}{m}v^2 + V$$

=> $$E = \frac{P^2}{2m} + V$$

on multiplying both sides by ψ

$$E\psi = \left(\frac{P^2}{2m} + V\right)\psi$$

on putting values of $P\psi$ from equation (12)

$$E\psi = \frac{1}{2m} \times (-i\text{h})^2 \frac{\partial^2\psi}{\partial x^2} + V\psi$$

or $$E\psi = \frac{-\text{h}^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V\psi$$

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

=> $$\boxed{\frac{\partial^2\psi}{\partial x^2} + \frac{2m}{\text{h}^2}(E - V)\psi = 0} \qquad (13)$$

This is time independent Schrodinger's equation for one-dimension motion of particle. For three-dimensional motion

$$\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + \frac{2m}{\text{h}^2}(E - V)\psi = 0 \qquad (14)$$

This is a time-independent Schrodinger equation for the one-dimension motion of a particle.

Probability density:

If ψ is a complex no. then its complex conjugate is obtained by replacing i by –i, ψ alone don't have any meaning but only ψψ* gives the probability of finding the particle. In quantum mechanics we cannot assert where exactly a particle is. We cannot say where it is likely to be

P (x) = ψψ* = [ψo e i(kx – ωt)] [ψo e -i(kx – ωt)] = |ψo| 2

Large value of |ψ|2 ----- Strong possibility of presence of particle

Small value of |ψ|2 ----- Less possibility of presence of particle

Normalization of wave function

The probability of finding the particle between any two coordinates x1 & x2 is determined by summing the probabilities in each interval dx.

Therefore there exists a particle between x1 & x2 in any interval dx. This situation can be mathematically represented by

$P_{x\in-\infty:\infty}(t)=\int_{\infty-\infty}|\psi(x,t)|_{2}dx.$

Properties of wave function.

Ψ should

- satisfy the law of conservation of energy i.e Total energy = PE + KE
- be consistent with de Broglie hypothesis i.e $\lambda = h/p$
- be single valued (because probability is unique)
- be continuous
- be finite
- be linear so that de Broglie waves have the important superposition property

Eigen value & Eigen function

A wave function Ψ, which satisfies all the properties is said to be Eigen function (Eigen = proper)

An operator O^ is a mathematical operator (differentiation, integration, addition, multiplication, division etc.) which may be applied on a function Ψ(x), which changes the function to another function Φ(x). This can be represented as

$$\hat{O}\,\Psi(x) = \Phi(x)$$

If a function is Eigen function, then by result of operation with an operator O , we get the same function as

$$\hat{O}\,\Psi(x) = \text{lamda}\Psi(x)$$

Unit – 3

FREE ELECTRON THEORY

Free Electron means unbound electron of the atom of any metal which can move freely. Many properties of metal-like electrical conductivity, magnetic susceptibility, Thermal conductivity, Mechanical strength, optoelectronic properties, etc. depend on the free electrons.

Free electron theory are of two types

1. Classical or Drude-Lorentz Theory
2. Quantum or Sommerfield Theory

This theory was first proposed by Lorentz and Drude with the help of classical laws of physics. The classical theory of free electron explains the properties like ohm's law, metallic luster, etc. but it could not explain heat capacity and the paramagnetic susceptibility of the conduction electrons. The Quantum theory removed the limitations of classical theory by using statistics developed by Sommerfield and succeeded in explaining many experimental facts.

Drude-Lorentz or Classical Theory

In 1900, P. Drude suggested that the metal crystals consist of positive ion cores with their valence electrons free to move among these positive metal ions. These valence electrons are also known as conduction electrons since they are responsible for the conduction of electricity in metals. The Coulomb's force of attraction between these positive ions and negatively charged electrons do not permit these free electrons to leave the metal surface. The potential field due to these ion's cores is supposed to be uniform and hence potential energy of electrons may be taken as constants (which is taken as zero for convenience). Hence we have to deal only with the kinetic energy of the electrons. The mutual repulsion among the electrons is neglected.

It has been suggested that the free electrons I the metal behave like these of atoms or molecules in a perfect gas. Hence, these electrons are sometimes known as the free-electron cloud. Fermi gas is constituted by electrons which are charged particles while the atoms or molecules which constitute ordinary gas are neutral The concentration of electrons in Fermi gas is large 10^{29}per m^3 as compared with the concentration of atoms or molecules of ordinary gas 10^{25} per m^3. That’s why Drude-Lorentz classical theory also known as Free Electron Gas Model.

In 1909, Lorentz suggested that this free electron gas behaves like perfect gas obeys Maxwell-Boltzmann statistics. Hence, the classical theory is also known as the Drude-Lorentz theory.

Success of Free Electron Theory

The free-electron theory successfully explained

1. Electrical conductivity
2. Ohm’s Law
3. Thermal conductivity
4. Weidmann-Franz law
5. Complete opacity of metals and their high luster

1. *Electrical conductivity*

It is defined as the amount of electricity that flows in unit time per unit area of cross-section of the conductor per unit potential gradient.

According to free-electron theory, in a solid, the electron moves freely. If E is the applied electric field, then the acceleration of an electron having charge is given by

$$dt^2 \qquad m \quad ....(1.1)$$

Integrating eq. (1.1), we get

$$dt \qquad m$$

At t=0,

dt , so C=0

Hence

dt m

If λ is the mean free path of electrons, then the relaxation time τ between two successive collisions is given by

v ...(1.2)

So average velocity between two successive collisions

$\bar{0}$ $\bar{0}$

$= \tau m \quad 2$

$2m$

Putting the values of τ from eq. (1.2), we get

$2mv$ ...(1.3)

Since

$$\frac{1}{2}mv^2 = \frac{3}{2}KT$$

(T is the absolute temperature and K is Boltzmann constant)

$$mv = \frac{3KT}{v}$$

Putting this value of (mv) in eq. (1.3), we get

$$\bar{\upsilon} = \frac{eE\lambda v}{6KT}$$ ...(1.4)

If n is the number density of electron in the conductor, then-current density *i* is given by

$$i = ne\bar{\upsilon}$$

Putting eq. (1.4) in above eq., we get

$$i = \frac{ne^2 E\lambda v}{6KT} \quad ...(1.5)$$

If q charge is flowing through a conductor of cross-section area A in time t, then

$$q = \sigma AEt$$

Or,

$$\frac{q}{t} = \sigma AE$$

$$i = \sigma AE\,(i = q/t)$$

$$\sigma = \frac{i}{AE} \quad ...(1.6)$$

For the unit area of cross-section

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

Using eq. (1.5), we get

$$\sigma = \frac{ne^2 \lambda v}{6KTA} \quad ...(1.7)$$

This expression show that conductivities of different materials depending on the number of free electron, T, and A.

2. Ohm’s law

From eq. (1.6) we have

$$\sigma E = \frac{i}{A}$$

$$\sigma E = J \quad \text{...(1.8)}$$

This is microscopic form of ohm’s law.

3. Thermal conductivity

There is no transfer of energy, if there is no temperature difference between two points in a system. So to discuss the thermal conductivity of metals, we suppose that a temperature gradient exists across the system instead of the voltage gradient, hence the transport of thermal energy takes place due to this gradient.

Suppose A and B be the two ends of the system in the form of a metallic rod and end A is at a higher temperature than B, then the thermal conductivity from A to B takes place by electrons. In a collision, the electron near A loose their kinetic energy while the electrons near B gain energy

The amount of heat Q passing through a cross-section of the rod per unit area per second is given by

$$Q = \frac{1}{3} n v \lambda \frac{dE}{dt} \quad \text{...(1.9)}$$

Where λ is mean free path, v is the velocity of electrons and n is the number density of free electron.

From Kinetic theory of gas

$$E = \frac{3}{2} KT$$

$$\frac{dE}{dt} = \frac{3}{2} K \frac{dT}{dt}$$

Hence

$$Q = \frac{1}{2} n v \lambda K \frac{dT}{dt}$$

Or

$$K\frac{dT}{dt} = \frac{1}{2}nv\lambda K\frac{dT}{dt}\left(Q = K\frac{dT}{dt}\right)$$

$$K = \frac{1}{2}nv\lambda K \quad \text{...(1.10)}$$

This value of K is verified experimentally and the theory of the free electron is found to be successful to explain thermal conductivity.

4. Wiedemann-Franz relation

In 1853, Weidmann and Franz discovered that all good electrical conductors are also good thermal conductors and the ratio of thermal conductivity to the electrical conductivity at any temperature is constant for all metals.

i.e.,

$$\frac{K}{\sigma} = constant$$

using eqns. (1.7) and (1.10), we get

$$\frac{K}{\sigma} = \frac{\frac{1}{2}nv\lambda K.6KT}{ne^2\lambda v}$$

$$= 3\left(\frac{K}{e}\right)^2 T$$

$$\frac{K}{\sigma} \propto T$$

This is Weidmann-Franz relation.

5. Lustre and opacity of metals

When electromagnetic radiations fall on metal, it produces forced oscillations in the free electrons having the same velocity as that of electromagnetic radiations. Thus the energy of incident radiations is absorbed by free electrons and the metal appears opaque. The excited electron on returning to its initial state emits a photon having the same energy as is absorbed initially. This energy is given out in the form of visible light in all directions, but only the light rays directed towards the metal surface can get through. Hence, the metal appears to reflect virtually all the light that is incident on it, giving it the characteristic metallic luster.

Failure of the theory

1. It fails to explain the heat capacity of materials. According to this theory

$$C_v \propto T$$

But experimentally,

$$C_v \propto \text{independent of } T$$

2. It does not explain the paramagnetic susceptibility of the conduction electrons, Experimentally χ is independent of T, but by this theory

$$\hat{} \quad T$$

3. It does not explain why some crystals are metallic.

4. It is unable to explain why metals prefer certain structures.

5. It does not explain the temperature variation of electrical resistivity because this theory predicts the variation of resistivity as $\sqrt{T}$ which actually is linear.

6. it is unable to differentiate insulator, semiconductor and conductor.

Quantum Free Electron Theory (Sommerfeld Theory)

To overcome the drawbacks of classical free electron theory, Sommerfeld proposed quantum free electron theory. He treated electron as a quantum particle. He retains the vital features of classical free electron theory and included the Pauli Exclusion

Principle & Fermi-Dirac statistics. The following are the assumptions of quantum free electron theory.

1. The free electrons in a metal can have only discrete energy values. Thus the energies are quantized.

2. The electrons obey Pauli's Exclusion Principle, which states that there cannot be more than two electrons in any energy level.

3. The distribution of electrons in various energy levels obey the Fermi-Dirac quantum statistics.

4. Free electrons have the same potential energy everywhere within the metal, because the potential due to ionic cores is uniform throughout the metal.

5. The force of attraction between electrons & lattice ions and the force of repulsion between electrons can be neglected.

6. Electrons are treated as wave-like particles.

Free Electron Gas in one-dimensional Box

Consider an electron having mass m confined to move in a one-dimensional rectangular box of length L i.e., The electron moves along a straight line say along the x-axis, let the wall of the box be rigid and non-penetrable. Let the electron can travel along the x-axis between x=0 and x= L.

The potential V is defined as

$$\{\infty \; for \; x < 0 \; and \; x > L$$

The Schrodinger's wave equation for the wave-function of the electron moving along the x-axis in a straight line may be written as

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

For free electron inside the box V=0, so the above eq. becomes

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

Where E is the Kinetic-energy of the electron

Or

$$\frac{d^2\Psi}{dx} + K^2\Psi = 0 \quad \text{...(1)}$$

Where

$$\hbar^2 \quad \text{...(2)}$$

The general solution of Eq. (1) may be written as

$$\Psi(x) = A\sin Kx + \cos Kx \quad \text{...(3)}$$

Where A and B are the constants to be determined by the boundary conditions.

Boundary condition

The wave function Ψ=0 at x=0 and x=L

1. At x=0, Ψ=0, so eq. (3) becomes

0= 0+B

B=0

So eq. (3) becomes

Ψ= A Sin Kx …(4)

2. At x=L, Ψ=0, so eq. (4) becomes

0 = A Sin Kx (A is constant so can't equal to 0)

$$\sin KL = \sin n\pi$$

$$KL = n\pi$$

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

$$K^2 = \frac{n^2\pi^2}{L^2} \quad \text{...(5)}$$

For nth state eq. (4) becomes

$$\Psi = A\sin\frac{n\pi x}{L} \quad \text{...(6)}$$

Using eq. (2), eq. (5) can be written as

$$\frac{2mE}{ħ^2} = \frac{n^2\pi^2}{L^2}$$

Or

$$- \qquad 2mL^2$$

Since E depends on ‘n’, so let us denote the energy of the electron by E_n

$$E_n = \frac{n^2\pi^2ħ^2}{2mL^2} \quad \text{...(7)}$$

The number n is called the quantum number. We find the lowest energy is obtained when n=1

From eq. (7), we have

$$E_1 = \frac{ħ^2\pi^2}{2mL^2}$$

i.e.,

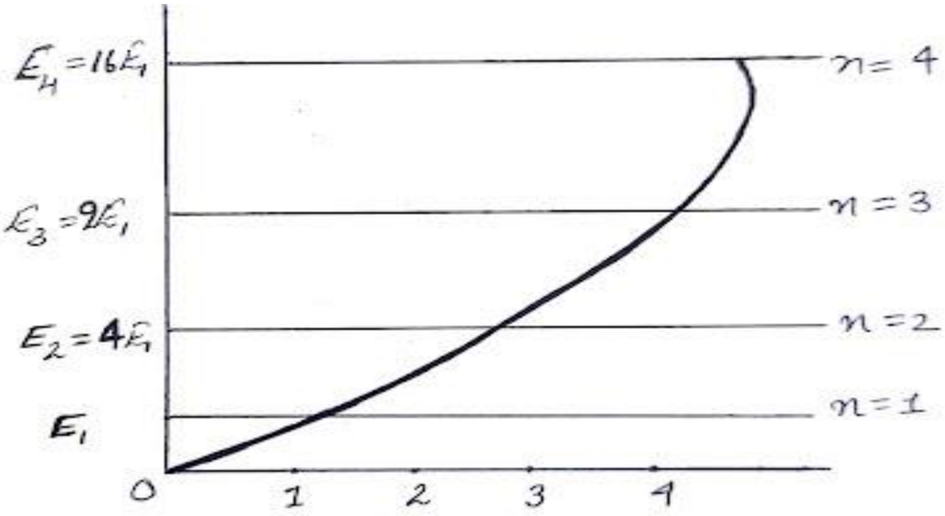
$$E_n = n^2E_1 \quad \text{...(8)}$$

Now, the spacing between the energy levels can be found by

$$E_{n+1} - E_n = (n+1)^2E_1 + n^2(n+1)^2E_1$$

$$= (2n+1)E_1$$

The Energy level diagram is shown as -



That’s mean energy level separation increases with n.

Fermi Energy

At T=0K, there will be no thermal excitation i.e., all the electrons remain at ground level. For a large value of n, there will be the uppermost filled level at T=0K called the Fermi level. This is the last filled level. The energy of the Fermi level is called Fermi Energy.

$$E_f = \frac{{n_f}^2\pi^2\hbar^2}{2mL^2} \quad \ldots(9)$$

In a linear solid, there are two quantum numbers of a conduction electron orbital i.e., n (=1,2,3…) and spin s (1/2). This means that for each value of n,s can have two values depending upon the orientation of the electron. Electrons are indistinguishable particles and obey Pauli’s exclusion principle states that no two electrons can have all four quantum numbers identical. Hence every orbital or quantum state n can accommodate two electrons one with spin up and the other with spin down. In other words, each energy level is doubly degenerate.

Let n_F denotes the uppermost filled energy level, and if there are N electrons we can write $N = 2n_F$

Hence eq. (9) give

$$E_f = \frac{\hbar^2}{2m}\left(\frac{N\pi}{2L}\right)^2 \quad \ldots(10)$$

This is the Fermi energy in 1-Dim.

Total Energy or Ground state Energy

At T=0K, all the electrons are in ground state i.e., there is no thermal excitation. Sum of all the energy states will be total ground state energy.

$$E_0 = 2\sum_{n=1}^{n=n_F} \frac{n^2\pi^2\hbar^2}{2mL^2}$$

Here factor 2 comes into the picture, since each level accommodates two-electron with equal energies.

Above eq. can be written as

$$E_0 = \frac{\pi^2\hbar^2}{2mL^2}\sum_{n=1}^{n=n_F} n^2 \qquad \ldots(11)$$

As we know

$$\sum_{n=1}^{S} n^2 = \frac{S(S+1)(2S+1)}{6}$$

$$= \frac{2S^3}{6} + \frac{3S^2}{6} + \frac{S}{6}$$

In a solid, total no. of electrons are very large, suppose S=1000, then we can neglect S^2 and S as compare to S^3

$n=1$

So

$$E_0 = \frac{\pi^2\hbar^2}{2mL^2}\frac{1}{3}n_f^3$$

$$E_0 = \frac{\pi^2\hbar^2}{2mL^2}\frac{1}{3}\left(\frac{N}{2}\right)^3 \left(n_f = \frac{N}{2}\right)$$

$$E_0 = \frac{N}{3} \times \frac{\hbar^2}{2m}\left(\frac{N\pi}{2L}\right)^2$$

$$E_0 = \frac{N}{3}E_F \qquad \text{....(12)}$$

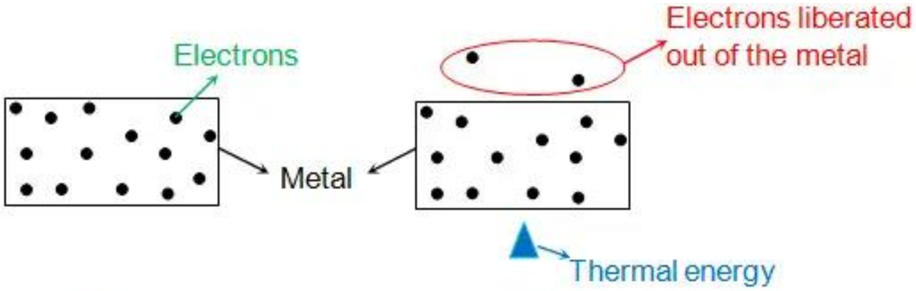
Thus, we find that for one-dimension, the average energy in the ground state is one-third of that of the Fermi energy.

Thermionic Emission

Thermionic emission is defined as the electron emission caused by a sufficiently high level of thermal energy. When a metal is heated sufficiently, the thermal energy supplied to the free electrons causes the emission of electrons from the metal surface. This occurs because the thermal energy given to the carrier overcomes the work function of the material. At average room temperature, the energy possessed by free electrons in a metal is insufficient to initiate thermionic emission.

All materials are composed of atoms which in turn consists of a nucleus, made of protons and neutrons, surrounded by electrons. These electrons are distributed at various levels around the nucleus and thus possess different levels of energy. Now, imagine that we start heating a particular material. The thermal energy so supplied increases the kinetic energy of the electrons within the material. This causes them to overcome the force of attraction which exists between them and the protons within the respective nuclei.

As a result, they get knocked-out from the material and will be liberated into space surrounding the material (Figure 1). As more heat is supplied, more are the number of electrons ejected. This phenomenon is known as thermionic emission i.e. emission of ions called thermions due to the thermal energy supplied. Thermionic emission was first observed by Thomas Alva Edison in 1883.



(a) (b)

Figure1 Electrons in the case of (a) metal (b) heated metal

Rate of Thermionic Emission

The number of thermions emitted per second from a substance is known as the rate of thermionic emission. This value depends on the:

1. Nature of the Material
   In general, every element can be characterized by its electronic configuration i.e. by the distribution of electrons surrounding its nucleus. When we speak of thermionic emission, our particular interest is in the valence electrons (electrons in the outermost shell). This is because these are the electrons which can be easily freed from the force of attraction so as to enable conduction. However, the energy which must be supplied differs from element to element and is regarded to be its threshold energy or work function.
2. Surface Temperature
   Higher is the temperature of the substance, greater is the rate of thermionic emission.
3. Surface Area
   If the surface area of the material considered is larger, then there will be a greater number of thermions emitted. This means that the rate of thermionic emission is directly proportional to the surface area of the material.

Thermionic Current

The flow of thermions gives rise to the flow of current known as thermionic current. Mathematically the thermionic equation which gives the current density of electrons is expressed as:

$$j = AT^2 e^{-e\Phi_W / k_B T}$$

Where:

- T is the absolute temperature,
- k_B is the Boltzmann Constant,
- Φ_W is the work function,
- e is the electron charge
- A is a constant.

Applications of Thermionic Emission

Thermionic emission forms the basic principle on which many of the devices used in the field of electronics and communication operates. Example applications of thermionic emission include vacuum tubes, diode valves, cathode ray tube, electron tubes, electron microscopes, X-ray tubes, thermionic converters, and electrodynamic tethers.

BAND THEORY OF SOLID

Bloch Theorem

The band theory of solid assume that electron move in a periodic potential of the period ‘a’ (lattice constant).

i.e

$$V(x) = V(x+a) \quad ...(1)$$

The Schrodinger's equation for free-electron moving in a constant potential V_0 is given as

$$\frac{d^2\Psi}{dx} + \frac{2m}{\hbar^2}(E - V_0)\Psi = 0 \quad ...(2)$$

So the Schrodinger's equation for an electron moving in periodic potential V(x) is written as

$$\frac{d^2\Psi}{dx} + \frac{2m}{\hbar^2}(E - V(x)\Psi = 0 \quad ...(3)$$

The solution of Eq. (3) are of the type

$$\Psi(x) = \exp(\pm iKx)\, u_K(x) \quad ...(4)$$

Where $u_K(x)$ is known as Bloch function and which is again periodic with the same periodicity of the lattice.

i.e.

$$u_K(x) = u_K(x+a) \quad ...(5)$$

Hence

$$\Psi(x+a) = \exp[iK(x+a)]\, u_K(x+a)$$

i.e.,

$$\Psi(x+a) = \exp(iKx)\exp(iKa)\, u_K(x)$$

or

$$\Psi(x+a) = \Psi(x)\exp(iKa)$$

$$\Psi(x+a) = \gamma\,\Psi(x) \quad ...(6)$$

Where

$$\gamma = \exp(ika)$$

Thus Bloch Theorem is a mathematical statement regarding the form of the one-electron wave function for a perfectly periodic potential.

The Kronig-Penney model

The Kronig-Penney model demonstrates that a simple one-dimensional periodic potential yields energy bands as well as energy band gaps. While it is an oversimplification of the three-dimensional potential and band structure in an actual crystal, it is an instructive tool to demonstrate how the band structure can be calculated for a periodic potential, and how allowed and forbidden energies are obtained when solving the Schrodinger equation.

The Kronig-Penney model is represented by the one-dimensional periodic potential as :

A virtue of this model is that it is possible to analytically determine the energy eigenvalues and eigenfunctions. It is also possible to find analytic expressions for the dispersion relation (*E* vs. *k*) and the electron density of states. [2] Derivations are given below.

This form can be used to plot the dispersion relation and the density of states for the Kronig Penney model.

Conclusion from Kronig –Penny Model:

1). The Energy spectrum of E consists of an infinite number of allowed energy bonds separated by intervals in which, there are no allowed energy levels. These are known as forbidden regions.

2). When α a increases, the first term of e on LHS decreases, so that the width of the allowed energy bonds is increased, and forbidden energy regions become narrow.

3). The width of the allowed bond decreases with the increase of p value. When p → α, the allowed energy regions become infinity narrow and the energy spectrum becomes line spectrum.

Effective mass

An electron in a crystal interacts with the lattice potential, hence its response to external force is different from a free electron. Hence the deviation of electrons behavior is taken into account by considering electrons to have effective mass m^*.

It can be negative and positive or infinite. It is given by

$$m^* = \frac{\hbar^2}{\left[\frac{d^2E}{dK^2}\right]}$$

Let us consider an electron in external electric field E. when this field acts on electron for a small time dt, electron gains some energy

$$dE = eEdx = eEvdt \quad ...(1)$$

According to de-Broglie theory, the velocity of an electron with wave vector K is the same as group velocity of wave representing the electron i.e.

$$v = \frac{d\omega}{dK}$$

Where ω is the angular frequency and is related to the electron energy E by the relation $E = \hbar\omega$. In term of energy, the velocity is expressed as

$$v = \frac{1}{\hbar}\left(\frac{dE}{dK}\right) \quad ...(2)$$

Therefore by using eq. (2) in eq. (1), we get

$$dE = \frac{eE}{\hbar}\frac{dE}{dK}dt$$

$$\frac{dK}{dt} = \frac{eE}{\hbar} \quad ...(3)$$

Now

$$v = \frac{1}{\hbar}\frac{dE}{dK}$$

Hence to get acceleration, differentiating this eq. w.r.t

$$a = \frac{dv}{dt} = \frac{1}{\hbar}\frac{d^2E}{dK^2}\frac{dK}{dt}$$ ...(4)

From eq. (3) and (4), we get

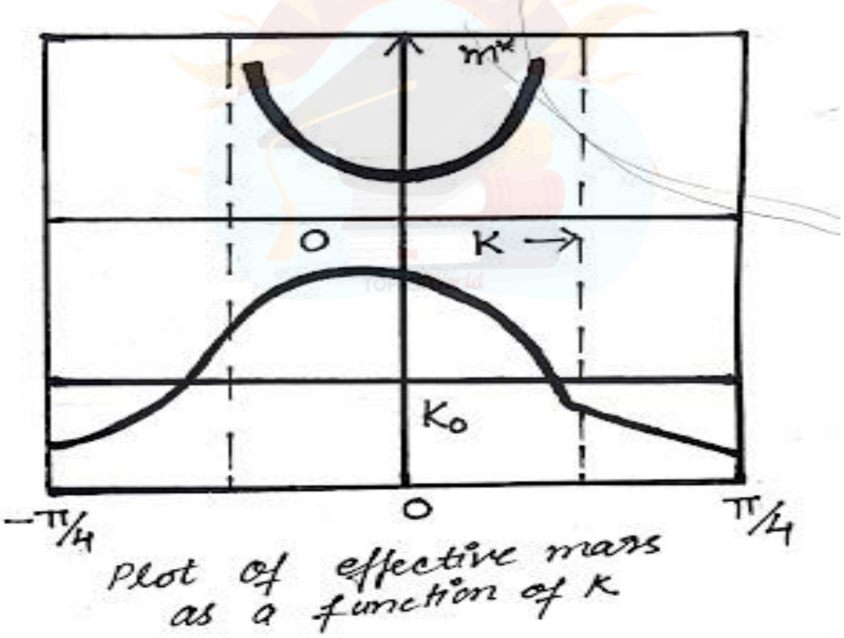
$$a = \frac{eE}{\hbar^2}\left(\frac{d^2E}{dK^2}\right)$$ ...(5)

Comparing this result with classical result

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

It follows that an electron behaves as if it has an effective mass m^* is given by

$$m^* = \frac{\hbar^2}{\left[\frac{d^2E}{dK^2}\right]}$$



Plot of effective mass as a function of K

In fig.m^* is represented as a function of . An interesting fact we observe here is that m^* is positive in the lower half and negative in the upper half of the energy band and m^* become infinite at the inflection point $K = K_0$. This means that in the upper half the electron behaves as a positively charged particle which is usually referred to as a hole. In most conductors $m^* = m$ because the band is only partially filled. But in semiconductors and insulators having almost filled valence bands, the effective mass play as a important role in the conduction process.

Energy Bands

In gaseous substances, the arrangement of molecules is spread apart and are not so close to each other. In liquids, the molecules are closer to each other. But, in solids, the molecules are closely arranged together, due to this atoms of molecules tend to move into the orbitals of neighbouring atoms. Hence, the electron orbitals overlap when atoms come together.
In solids, several bands of energy levels are formed due to the intermixing of atoms in solids. We call these set of energy levels as **energy bands**.

Formation of Energy Bands

In an isolated atom, the electrons in each orbit possess definite energy. But, in the case of solids, the energy level of the outermost orbit electrons is affected by the neighbouring atoms.

When two isolated charges are brought close to each other, the electrons in the outermost orbit experience an attractive force from the nearest or neighbouring atomic nucleus. Due to this reason, the energies of the electrons will not be at the same level, the energy levels of electrons are changed to a value which is higher or lower than that of the original energy level of the electron.

The electrons in the same orbit exhibit different energy levels. The grouping of these different energy levels is known as the energy band.

However, the energy of the inner orbit electrons is not much affected by the presence of neighbouring atoms.

Classification of Energy Bands

Valence Band

The electrons in the outermost shell are known as valence electrons. These valence electrons contain a series of energy levels and form an energy band known as the valence band. The valence band has the highest occupied energy.

Conduction Band

The valence electrons are not tightly held to the nucleus due to which a few of these valence electrons leave the outermost orbit even at room temperature and become free electrons. The free electrons conduct current in conductors and are therefore known as conduction electrons. The conduction band is one that contains conduction electrons and has the lowest occupied energy levels.

Forbidden Energy Gap

The gap between the valence band and the conduction band is referred to as the forbidden gap. As the name suggests, the forbidden gap doesn’t have any energy and no electrons stay in this band. If the forbidden energy gap is greater, then the valence band electrons are tightly bound or firmly attached to the nucleus. We require some amount of external energy that is equal to the forbidden energy gap.

Conductors

Gold, Aluminium, Silver, Copper, all these metals allow an electric current to flow through them.

There is no forbidden gap between the valence band and conduction band which results in the overlapping of both the bands. The number of free electrons available at room temperature is large.

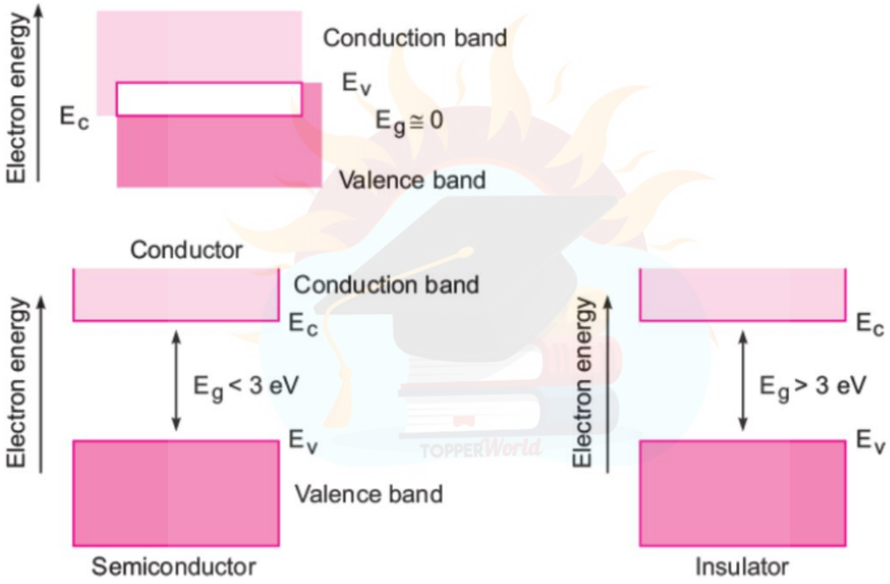
Insulators

Glass and wood are examples of the insulator. These substances do not allow electricity to pass through them. They have high resistivity and very low conductivity.

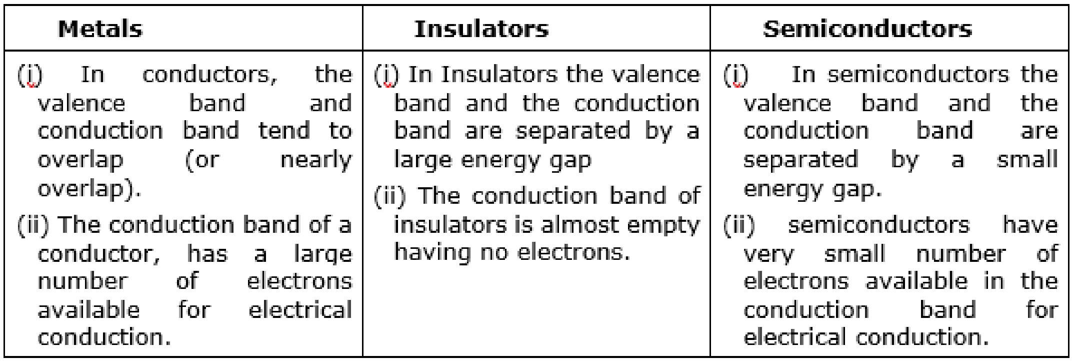
The energy gap in the insulator is very high up to 7eV. The material cannot conduct because the movement of the electrons from the valence band to the conduction band is not possible.

Semiconductors

Germanium and Silicon are the most preferable material whose electrical properties lie in between semiconductors and insulators. The energy band diagram of semiconductors is shown where the conduction band is empty and the valence band is completely filled but the forbidden gap between the two bands is very small that is about 1eV. For Germanium, the forbidden gap is 0.72eV and for Silicon, it is 1.1eV. Thus, semiconductor requires small conductivity.



Distinction between Metals, Insulators and Semiconductors



Metals	Insulators	Semiconductors
(i) In conductors, the valence band and conduction band tend to overlap (or nearly overlap). (ii) The conduction band of a conductor, has a large number of electrons available for electrical conduction.	(i) In Insulators the valence band and the conduction band are separated by a large energy gap (ii) The conduction band of insulators is almost empty having no electrons.	(i) In semiconductors the valence band and the conduction band are separated by a small energy gap. (ii) semiconductors have very small number of electrons available in the conduction band for electrical conduction.

Hall Effect

Hall effect is a process in which a transverse electric field is developed in a solid material when the material carrying an electric current is placed in a magnetic field that is perpendicular to the current. The Hall effect was discovered by Edwin Herbert Hall in 1879. In this article, let us learn about the Hall effect in detail.

Principle of Hall effect

The principle of the Hall effect states that when a current-carrying conductor or a semiconductor is introduced to a perpendicular magnetic field, a voltage can be measured at the right angle to the current path. This effect of obtaining a measurable voltage is known as the Hall effect.

Theory

When a conductive plate is connected to a circuit with a battery, then a current starts flowing. The charge carriers will follow a linear path from one end of the plate to the other end. The motion of charge carriers results in the production of magnetic fields. When a magnet is placed near the plate, the magnetic field of the charge carriers is distorted. This upsets the straight flow of the charge carriers. The force which upsets the direction of flow of charge carriers is known as Lorentz force.

Due to the distortion in the magnetic field of the charge carriers, the negatively charged electrons will be deflected to one side of the plate and positively charged holes to the other side. A potential difference, known as the Hall voltage will be generated between both sides of the plate which can be measured using a metre.

The Hall voltage represented as V_H is given by the formula:

$$V_H = \frac{I B}{q n d}$$

Here,

I is the current flowing through the sensor

B is the magnetic field strength

q is the charge

n is the number of charge carriers per unit volume

d is the thickness of the sensor.

Applications of Hall Effect

Hall effect principle is employed in the following cases:

- Magnetic field sensing equipment
- For the measurement of direct current, Hall effect Tong Tester is used.
- It is used in phase angle measurement
- Proximity detectors
- Hall effect Sensors and Probes
- Linear or Angular displacement transducers
- For detecting wheel speed and accordingly assist the anti-lock braking system.

Unit – 4

SEMICONDUCTORS

Semiconductors are the materials which have a **conductivity between conductors** (generally metals) and non-conductors or **insulators** (such as ceramics). Semiconductors can be compounds such as gallium arsenide or pure elements, such as germanium or silicon. Physics explains the theories, properties and mathematical approach governing semiconductors.

Examples of Semiconductors:

Gallium arsenide, germanium, and silicon are some of the most **commonly used semiconductors**. Silicon is used in electronic circuit fabrication and gallium arsenide is used in solar cells, laser diodes, etc.

Holes and Electrons in Semiconductors

Holes and electrons are the types of charge carriers accountable for the flow of current in semiconductors. **Holes** (valence electrons) are the positively charged electric charge carrier whereas **electrons** are the negatively charged particles. Both electrons and holes are equal in magnitude but opposite in polarity.

Mobility of Electrons and Holes

In a semiconductor, the **mobility of electrons is higher than that of the holes**. It is mainly because of their different band structures and scattering mechanisms.

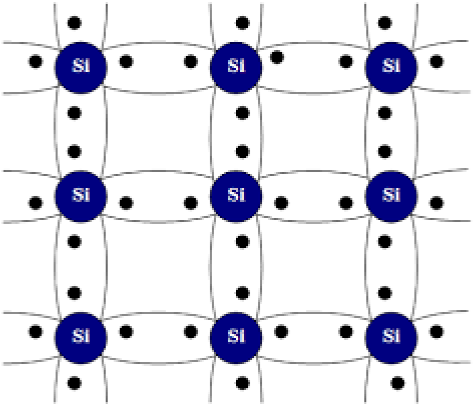
Electrons travel in the conduction band whereas holes travel in the valence band. When an electric field is applied, holes cannot move as freely as electrons due to their restricted movent. The elevation of electrons from their inner shells to higher shells results in the creation of holes in semiconductors. Since the holes experience stronger atomic force by the nucleus than electrons, holes have lower mobility.

The mobility of a particle in a semiconductor is more if;

- Effective mass of particles is lesser
- Time between scattering events is more

For intrinsic silicon at 300 K, the mobility of electrons is 1500 cm^2 (V·s)$^{-1}$ and the mobility of holes is 475 cm^2 (V·s)$^{-1}$.

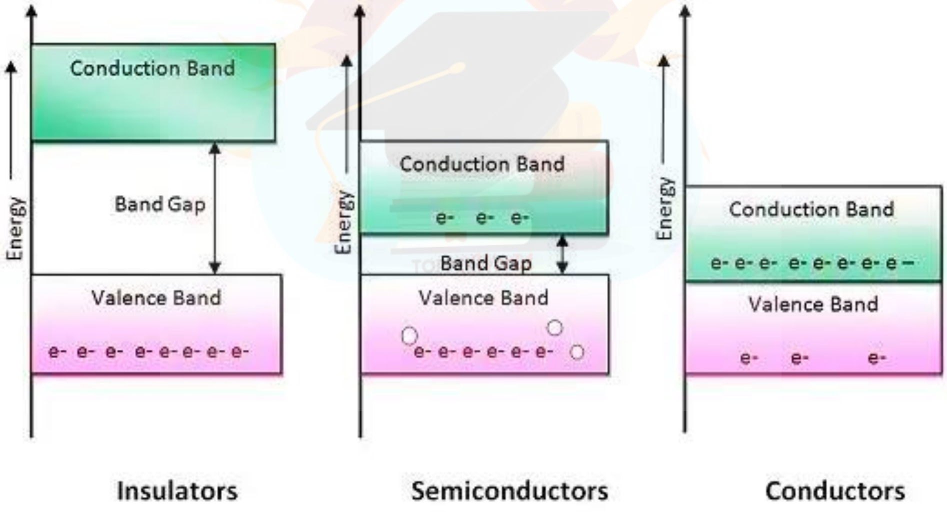
The **bond model** of electrons in silicon of valency 4 is shown below. Here, when one of the free electrons (blue dots) leaves the lattice position, it creates a hole (grey dots). This hole thus created takes the opposite charge of the electron and can be imagined as positive charge carriers moving in the lattice.



Band Theory of Semiconductors

The introduction of band theory happened during the quantum revolution in science. Walter Heitler and Fritz London discovered the energy bands.

We know that the electrons in an atom are present in different energy levels. When we try to assemble a lattice of a solid with N atoms, then each level of an atom must split up into N levels in the solid. This splitting up of sharp and tightly packed energy levels forms **Energy Bands**. The gap between adjacent bands representing a range of energies that possess no electron is called a **Band Gap**.



Conduction Band and Valence Band in Semiconductors

Valence Band:

The energy band involving the energy levels of valence electrons is known as the valence band. It is the highest occupied energy band. When compared with insulators, the bandgap in semiconductors is smaller. It allows the electrons in the valence band to jump into the conduction band on receiving any external energy.

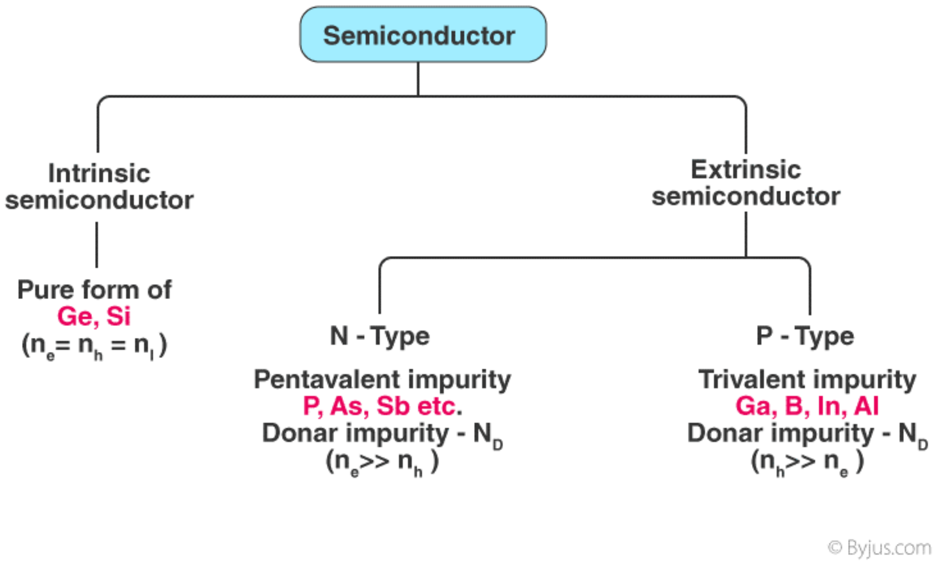
Conduction Band:

It is the lowest unoccupied band that includes the energy levels of positive (holes) or negative (free electrons) charge carriers. It has conducting electrons resulting in the flow of current. The conduction band possess high energy level and are generally empty. The conduction band in semiconductors accepts the electrons from the valence band.

Types of Semiconductors

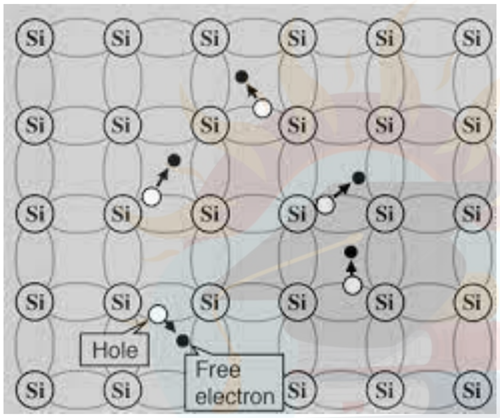
Semiconductors can be classified as:

- **Intrinsic Semiconductor**
- **Extrinsic Semiconductor**



Intrinsic Semiconductor

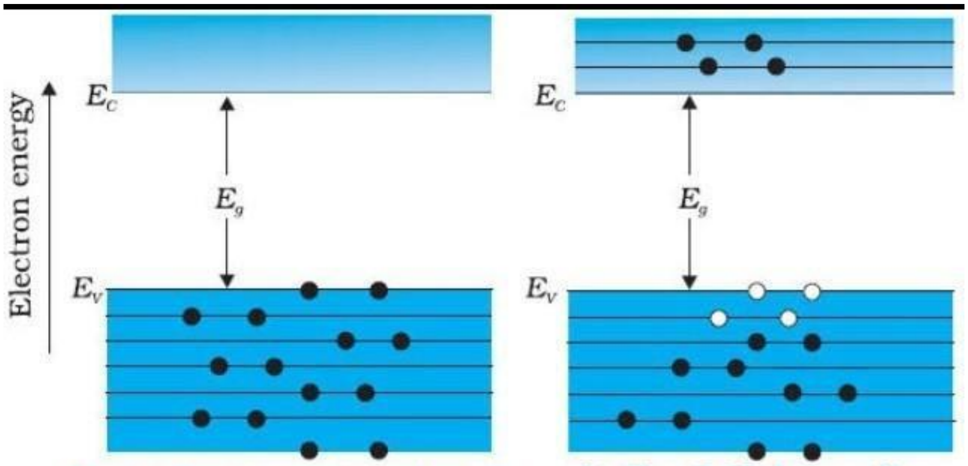
An intrinsic type of semiconductor material is made to be very pure chemically. It is made up of only a single type of element.



Germanium (Ge) and Silicon (Si) are the most common type of intrinsic semiconductor elements. They have four valence electrons (tetravalent). They are bound to the atom by covalent bond at absolute zero temperature.

When the temperature rises, due to collisions, few electrons are unbounded and become free to move through the lattice, thus creating an absence in its original position (hole). These free electrons and holes contribute to the conduction of electricity in the semiconductor. The negative and positive charge carriers are equal in number.

The thermal energy is capable of ionizing a few atoms in the lattice, and hence their conductivity is less.



An intrinsic semiconductor behaves like insulator at $T = 0$ K

At $T > 0$ K thermally generated electron-hole pairs

In intrinsic semiconductors, current flows due to the motion of free electrons as well as holes. The total current is the sum of the electron current I_e due to thermally generated electrons and the hole current I_h

Total Current (I) = $I_e + I_h$

For an intrinsic semiconductor, at finite temperature, the probability of electrons to exist in conduction band decreases exponentially with increasing bandgap (E_g)

$n = n_0 e^{-Eg/2.Kb.T}$

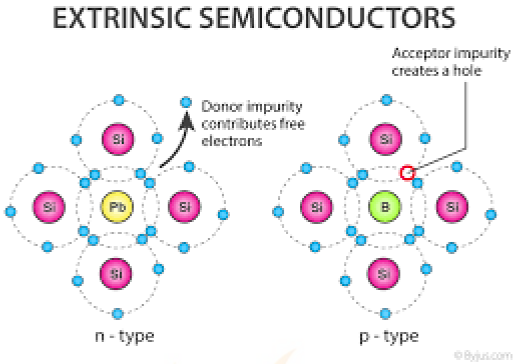
Where,

- Eg = Energy bandgap
- K_b = Boltzmann’s constants

Extrinsic Semiconductor

The conductivity of semiconductors can be greatly improved by introducing a small number of suitable replacement atoms called IMPURITIES. The process of adding impurity atoms to the pure semiconductor is called DOPING. Usually, only 1 atom in 10^7 is replaced by a dopant atom in the doped semiconductor. An extrinsic semiconductor can be further classified into:

- **N-type Semiconductor**
- **P-type Semiconductor**



N-Type Semiconductor

- Mainly due to electrons
- Entirely neutral
- $I = I_h$ and $n_h >> n_e$
- Majority – Electrons and Minority – Holes

When a pure semiconductor (Silicon or Germanium) is doped by pentavalent impurity (P, As, Sb, Bi) then, four electrons out of five valence electrons bonds with the four electrons of Ge or Si.

The fifth electron of the dopant is set free. Thus, the impurity atom donates a free electron for conduction in the lattice and is called "**Donar**".

Since the number of free electron increases by the addition of an impurity, the negative charge carriers increase. Hence, it is called n-type semiconductor.

Crystal as a whole is neutral, but the donor atom becomes an immobile positive ion. As conduction is due to a large number of free electrons, the electrons in the n-type semiconductor are the MAJORITY CARRIERS and holes are the MINORITY CARRIERS.

P-Type Semiconductor

- Mainly due to holes
- Entirely neutral
- $I = I_h$ and $n_h >> n_e$
- Majority – Holes and Minority – Electrons

When a pure semiconductor is doped with a trivalent impurity (B, Al, In, Ga) then, the three valence electrons of the impurity bonds with three of the four valence electrons of the semiconductor.

This leaves an absence of electron (hole) in the impurity. These impurity atoms which are ready to accept bonded electrons are called “**Acceptors**“.

With the increase in the number of impurities, holes (the positive charge carriers) are increased. Hence, it is called p-type semiconductor.

Crystal as a whole is neutral, but the acceptors become an immobile negative ion. As conduction is due to a large number of holes, the holes in the p-type semiconductor are MAJORITY CARRIERS and electrons are MINORITY CARRIERS.

Difference Between Intrinsic and Extrinsic Semiconductors

Intrinsic Semiconductor	**Extrinsic Semiconductor**
Pure semiconductor	Impure semiconductor
Density of electrons is equal to the density of holes	Density of electrons is not equal to the density of holes

Electrical conductivity is low	Electrical conductivity is high
Dependence on temperature only	Dependence on temperature as well as on the amount of impurity
No impurities	Trivalent impurity, pentavalent impurity

Industrial Uses of Semiconductors

The physical and chemical properties of semiconductors make them capable of designing technological wonders like microchips, transistors, LEDs, solar cells, etc.

The microprocessor used for controlling the operation of space vehicles, trains, robots, etc is made up of transistors and other controlling devices which are manufactured by semiconductor materials.

Importance of Semiconductors

Here we have discussed some advantages of semiconductors which makes them highly useful everywhere.

- They are highly portable due to the smaller size
- They require less input power
- Semiconductor devices are shockproof
- They have a longer lifespan
- They are noise-free while operating

SEMICONDUCTOR DEVICES

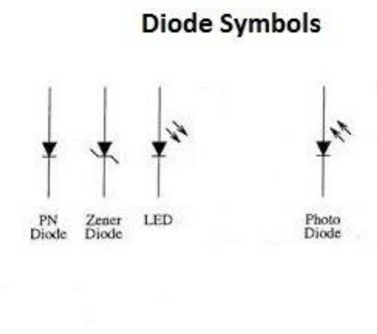
P-N Junction

A P-N junction is an interface or a boundary between two semiconductor material types, namely the p-type and the n-type, inside a semiconductor.

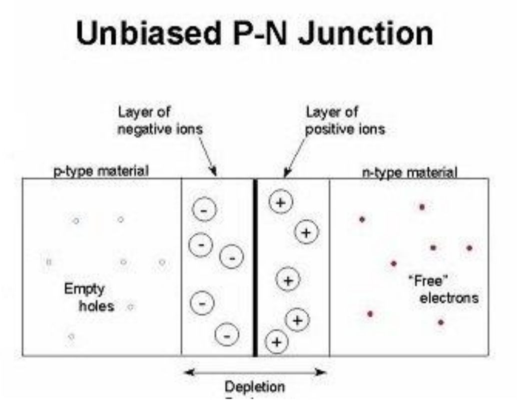
In a semiconductor, the P-N junction is created by the method of doping. The p-side or the positive side of the semiconductor has an excess of holes, and the n-side or the negative side has an excess of electrons. The process of doping is explained in further detail in the next section.

Formation of P-N Junction

As we know, if we use different semiconductor materials to make a P-N junction, there will be a grain boundary that would inhibit the movement of electrons from one side to the other by scattering the electrons and holes and thus, we use the process of doping. We will understand the process of doping with the help of this example. Let us consider a thin p-type silicon semiconductor sheet. If we add a small amount of pentavalent impurity to this, a part of the p-type Si will get converted to n-type silicon. This sheet will now contain both the p-type region and the n-type region and a junction between these two regions. The processes that follow after forming a P-N junction are of two types – diffusion and drift. There is a difference in the concentration of holes and electrons at the two sides of a junction. The holes from the p-side diffuse to the n-side, and the electrons from the n-side diffuse to the p-side. These give rise to a diffusion current across the junction.



Diode Symbols



Unbiased P-N Junction

when an electron diffuses from the n-side to the p-side, an ionised donor is left behind on the n-side, which is immobile. As the process goes on, a layer of positive charge is developed on the n-side of the junction. Similarly, when a hole goes from the p-side to the n-side, an ionized acceptor is left behind on the p-side, resulting in the formation of a layer of negative charges in the p-side of the junction. This region of positive charge and negative charge on either side of the junction is termed as the depletion region. Due to this positive space charge region on either side of the junction, an electric field with the direction from a positive charge towards the negative charge is developed. Due to this electric field, an electron on the p-side of the junction moves to the n-side of the junction. This motion is termed the drift. Here, we see that the direction of the drift current is opposite to that of the diffusion current.

Biasing Conditions for the P-N Junction Diode

There are two operating regions in the P-N junction diode:

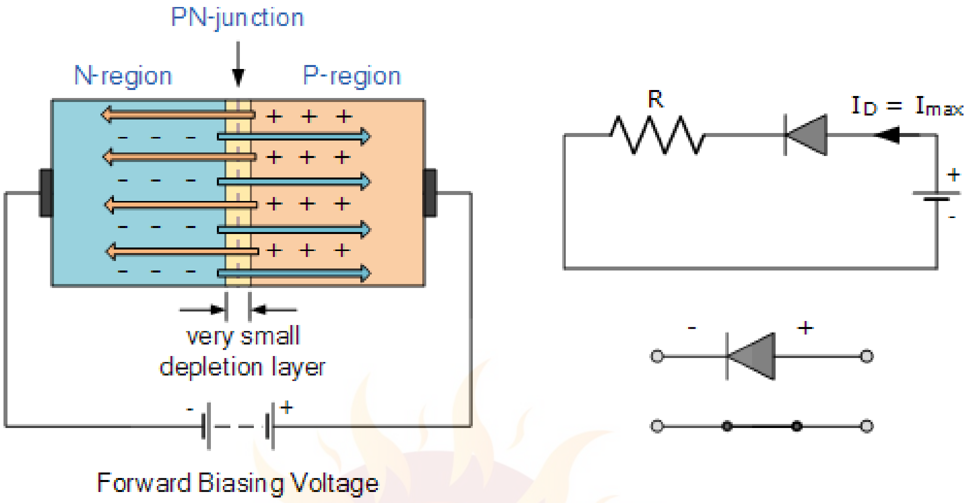
- P-type
- N-type

There are three biasing conditions for the P-N junction diode, and this is based on the voltage applied:

- Zero bias: No external voltage is applied to the P-N junction diode.

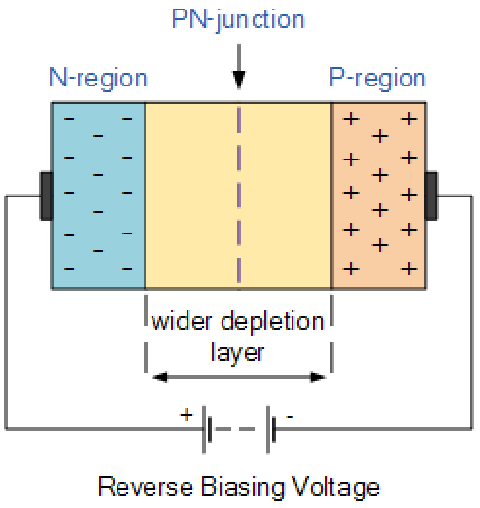
- Forward bias: The positive terminal of the voltage potential is connected to the p-type while the negative terminal is connected to the n-type.
- Reverse bias: The negative terminal of the voltage potential is connected to the p-type and the positive is connected to the n-type.

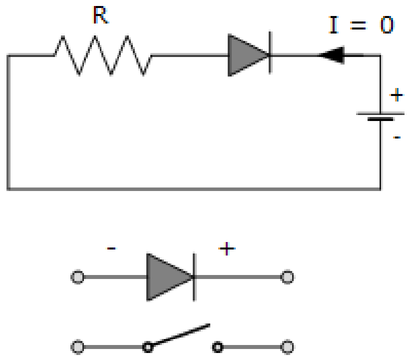
Forward Bias



When the p-type is connected to the battery's positive terminal and the n-type to the negative terminal, then the P-N junction is said to be forward-biased. When the P-N junction is forward biased, the built-in electric field at the P-N junction and the applied electric field are in opposite directions. When both the electric fields add up, the resultant electric field has a magnitude lesser than the built-in electric field. This results in a less resistive and thinner depletion region. The depletion region's resistance becomes negligible when the applied voltage is large. In silicon, at the voltage of 0.6 V, the resistance of the depletion region becomes completely negligible, and the current flows across it unimpeded.

Reverse Bias



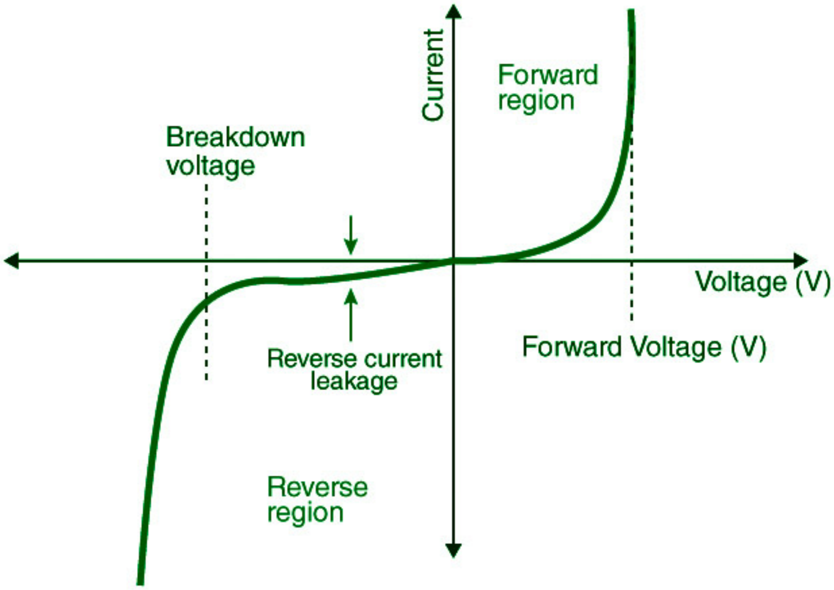


When the p-type is connected to the battery’s negative terminal and the n-type is connected to the positive side, the P-N junction is reverse biased. In this case, the built-in electric field and the applied electric field are in the same direction. When the two fields are added, the resultant electric field is in the same direction as the built-in electric field, creating a more resistive, thicker depletion region. The depletion region becomes more resistive and thicker if the applied voltage becomes larger.

current flow in the PN junction diode

The flow of electrons from the n-side towards the p-side of the junction takes place when there is an increase in the voltage. Similarly, the flow of holes from the p-side towards the n-side of the junction takes place along with the increase in the voltage. This results in the concentration gradient between both sides of the terminals. Due to the concentration gradient formation, charge carriers will flow from higher-concentration regions to lower-concentration regions. The movement of charge carriers inside the P-N junction is the reason behind the current flow in the circuit.

V-I Characteristics of P-N Junction Diode



VI characteristics of P-N junction diodes is a curve between the voltage and current through the circuit. Voltage is taken along the x-axis while the current is taken along the y-axis. The above graph is the V-I characteristics curve of the P-N junction diode. With the help of the curve, we can understand that there are three regions in which the diode works, and they are:

- Zero bias
- Forward bias
- Reverse bias

When the P-N junction diode is in zero bias condition, there is no external voltage applied and this means that the potential barrier at the junction does not allow the flow of current.

When the P-N junction diode is in forward bias condition, the p-type is connected to the positive terminal while the n-type is connected to the negative terminal of the external voltage. When the diode is arranged in this manner, there is a reduction in the potential barrier. For silicone diodes, when the voltage is 0.7 V and for germanium diodes, when the voltage is 0.3 V, the potential barriers decrease, and there is a flow of current.

When the diode is in forward bias, the current increases slowly, and the curve obtained is non-linear as the voltage applied to the diode overcomes the potential barrier. Once the diode overcomes the potential barrier, the diode behaves normally, and the curve rises sharply as the external voltage increases, and the curve obtained is linear.

When the P-N junction diode is in negative bias condition, the p-type is connected to the negative terminal while the n-type is connected to the positive terminal of the external voltage. This results in an increase in the potential barrier. Reverse saturation current flows in the beginning as minority carriers are present in the junction.

When the applied voltage is increased, the minority charges will have increased kinetic energy which affects the majority charges. This is the stage when the diode breaks down. This may also destroy the diode.

Applications of P-N Junction Diode

- P-N junction diode can be used as a photodiode as the diode is sensitive to the light when the configuration of the diode is reverse-biased.
- It can be used as a solar cell.
- When the diode is forward-biased, it can be used in LED lighting applications.
- It is used as rectifier in many electric circuits and as a voltage-controlled oscillator in varactors.

Transistor

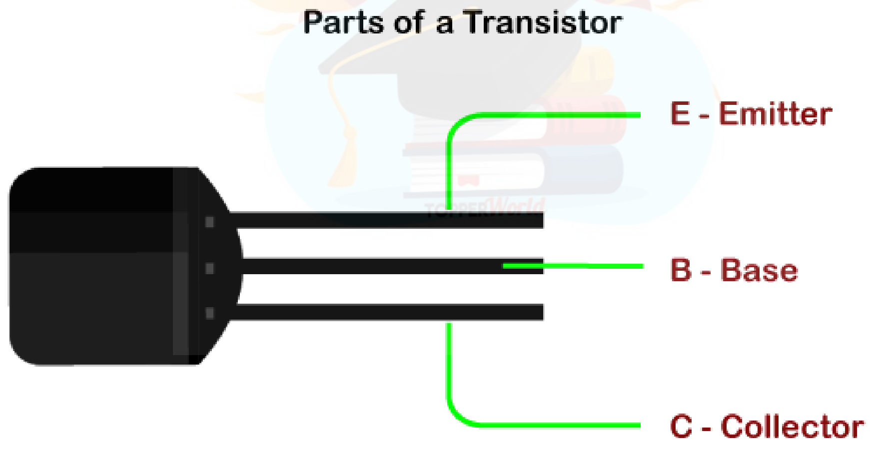
A transistor is a type of a semiconductor device that can be used to both conduct and insulate electric current or voltage. A transistor basically acts as a switch and an amplifier. In simple words, we can say that a transistor is a miniature device that is used to control or regulate the flow of electronic signals.

Transistors are one of the key components in most of the electronic devices that are present today. Developed in the year 1947 by three American physicists John Bardeen, Walter Brattain and William Shockley, the transistor is considered as one of the most important inventions in the history of science.

Parts of a Transistor

A typical transistor is composed of three layers of semiconductor materials or more specifically terminals which helps to make a connection to an external circuit and carry the current. A voltage or current that is applied to any one pair of the terminals of a transistor controls the current through the other pair of terminals. There are three terminals for a transistor. They are:

- Base: This is used to activate the transistor.
- Collector: It is the positive lead of the transistor.
- Emitter: It is the negative lead of the transistor



Types of Transistors

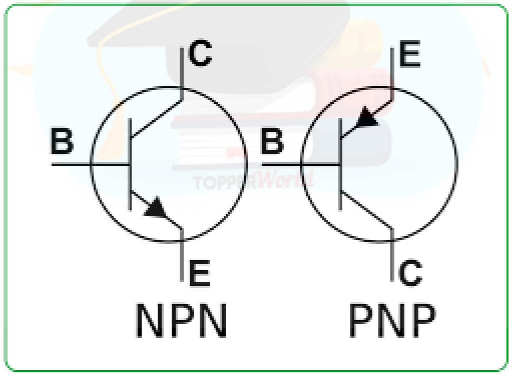
Based on how they are used in a circuit there are mainly two types of transistors.

Bipolar Junction Transistor (BJT)

The three terminals of BJT are base, emitter and collector. A very small current flowing between base and emitter can control a larger flow of current between the collector and emitter terminal.

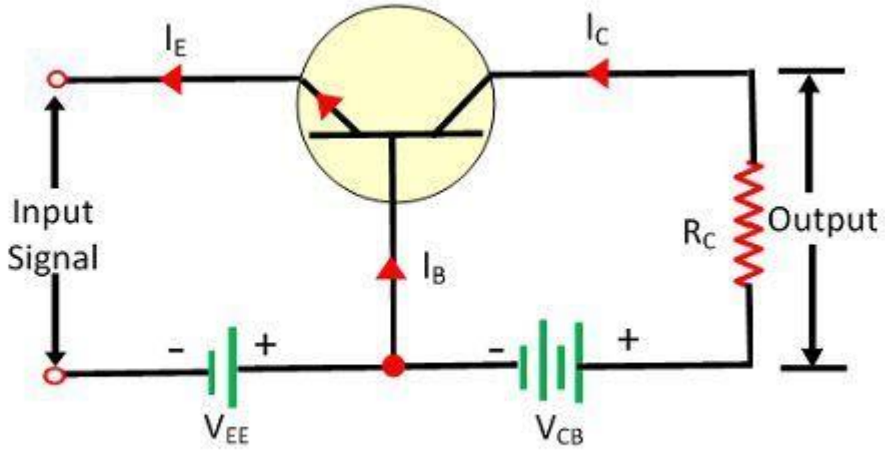
Furthermore, there are two types of BJT. These include;

- P-N-P Transistor: It is a type of BJT where one n-type material is introduced or placed between two p-type materials. In such a configuration, the device will control the flow of current. PNP transistor consists of 2 crystal diodes which are connected in series. The right side and left side of the diodes are known as the collector-base diode and emitter-base diode, respectively.
- N-P-N Transistor: In this transistor, we will find one p-type material that is present between two n-type materials. N-P-N transistor is basically used to amplify weak signals to strong signals. In NPN transistor, the electrons move from the emitter to collector region resulting in the formation of current in the transistor. This transistor is widely used in the circuit.



There are three types of configuration as a common base (CB), common collector (CC) and common emitter (CE).

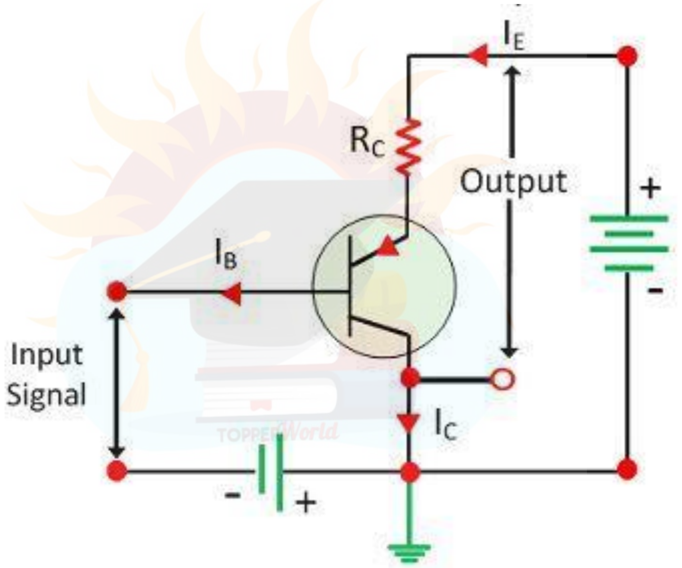
In Common Base (CB) configuration the base terminal of the transistor is common between input and output terminals.



Common Base Conection of NPN Transistor

Circuit Globe

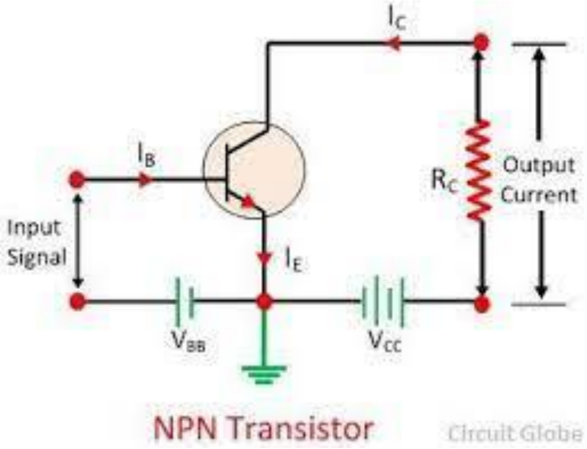
In Common Collector (CC) configuration the collector terminals are common between the input and output terminals.



NPN Transistor

Circuit Globe

In Common Emitter (CE) configuration the emitter terminal is common between the input and the output terminals.



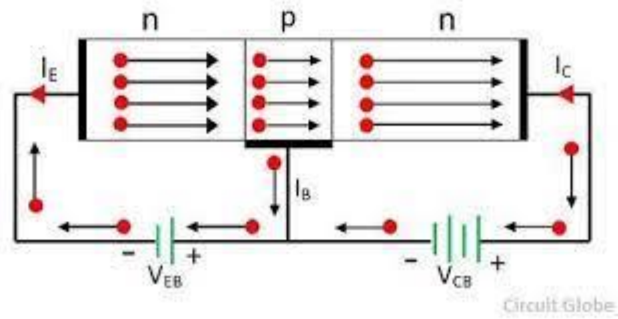
Field Effect Transistor (FET)

For FET, the three terminals are Gate, Source and Drain. The voltage at the gate terminal can control a current between source and drain. FET is a unipolar transistor in which N channel FET or P channel FET are used for conduction. The main applications of FETs are in low noise amplifier, buffer amplifier and an analogue switch.

Operation of NPN Transistor

The emitter of NPN device is made by n-type material, hence the majority carriers are electrons. When the base-emitter junction is forward biased the electrons will move from the n-type region towards the p-type region and the minority carriers holes moves towards the n-type region.

When they meet each other they will combine enabling a current to flow across the junction. When the junction is reverse biased the holes and electrons move away from the junction, and now the depletion region forms between the two areas and no current will flows through it.



When a current flows between base and emitter the electrons will leave the emitter and flow into the base as shown above. Normally the electrons will combine when they reach the depletion region.

But the doping level in this region is very low and the base is also very thin. This means that most of the electrons are able to travel across the region without recombining with holes. As a result, the electrons will drift towards the collector.

In this way, they are able to flow across what is effectively reverse-biased junction and the current flows in the collector circuit.

Characteristics of Transistor

Characteristics of the transistor are the plots which can represent the relation between the current and the voltage of a transistor in a particular configuration.

There are two types of characteristics.

- Input characteristics: It will give us the details about the change in input current with the variation in input voltage by keeping output voltage constant.

- Output characteristics: It is a plot of output current with output voltage by keeping input current constant.
- Current transfer Characteristics: This plot shows the variation of output current with the input current by keeping the voltage constant.

Metal-Semiconductor Junction

Metal-Semiconductor-Junction is also called hetero-junction because the material on each side of the junction is not the same. These junctions can also produce diodes.

There are two possible types of metal-semiconductor junctions:

1. Schottky Junction $(\varphi_m > \varphi_s)$ - These are the junctions in which work-function of metal is greater than the work-function of a semiconductor.

2. Ohmic Junction $(\varphi_m < \varphi_s)$ – These are the junctions in which work-function of metal is less than the work-function of a semiconductor.

Now let us discuss these two types of junctions in detail-

Schottky Junction

One of the first practical semiconductor devices used in the early 1900s was metal-semiconductor diodes. This diode also called a point contact whisper to an exposed semiconductor surface. These metal-semiconductor diodes were not easily reproduced or mechanically reliable and were replaced by the PN junction in the 1950s.

However, semiconductor and vacuum technology is now used to fabricate reproducible and reliable metal-semiconductor contacts. In most cases, the rectifying contacts are made on n-type semiconductor for this reason we will concentrate on this type of diode.

Metal-semiconductor ohmic Junction

Contacts must be made between any semiconductor device and the outside world. These contacts are made via ohmic contacts. Ohmic contacts are metal-to-semiconductor contacts, but in this case, they are not rectifying contacts. An ohmic contact is low-resistance junction providing conduction in both directions between the metal and the semiconductor. Two general types of ohmic contacts are possible: The first type is the ideal non-rectifying barrier, and the second is the tunneling barrier.

Semiconductor Laser

A semiconductor laser is a specially fabricated pn junction device (both the p and n regions are highly doped) which emits coherent light when it is forward biased. It is made from Gallium Arsenide (GaAs) which operated at low temperature and emits light in near IR region.

Semiconductor lasers are basically PN junction diode. When a p-type semiconductor (has an excess of holes) is in intimate contact with n-type semiconductor (has an excess of the electron), PN junction is formed at the interface.

When the PN junction is forward biased with an external supply voltage, electrons from the n- region and hole from the p- region are forced into the junction.

These electrons and holes attract each other. When they collide, they neutralize each other and as a result, emit recombination radiation.

The electrons in n-type semiconductors are in the **conduction band** (i.e.higher energy) and the hole in p-type semiconductors are in the **valence band** (i.e lower energy).

The energy difference between the conduction band and valence bands is called the **band gap** or **Forbidden energy gap** of the given material.

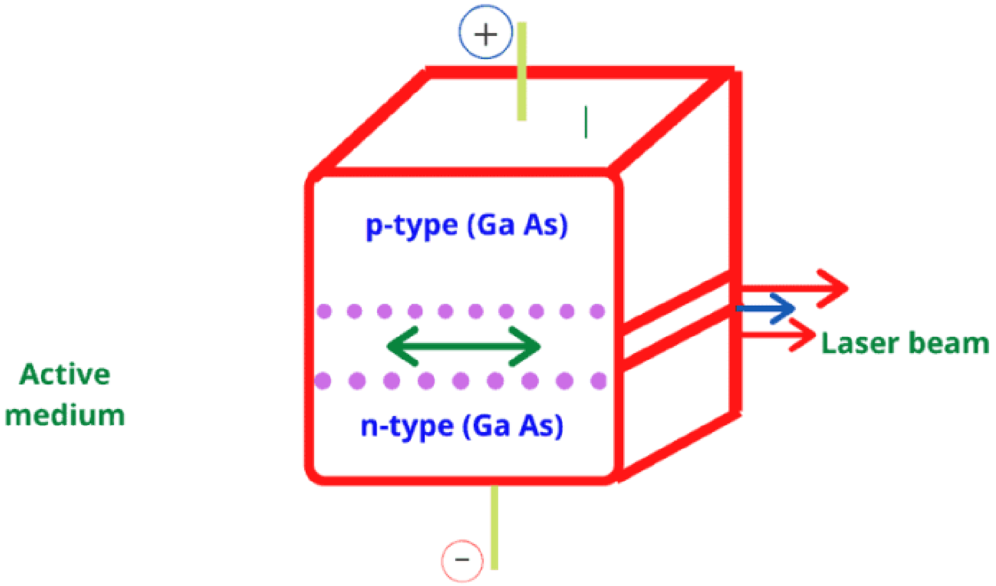
The energy of the photon emitted as recombination radiation is equal to the band gap of the material for a semiconductor laser.

Construction of Semiconductor Diode Laser

The semiconductor laser is made of different materials like gallium arsenide (Ga As), Indium phosphide (InP), gallium nitride (GaN), etc. The band gap of the semiconductor laser is different and hence light of different wavelengths is emitted by this laser.

For example band gap of GaAs is 1.42 eV at room temperature and laser light of a wide range of wavelengths from near-infrared region to the visible region of e.m. spectrum is emitted by a laser made of GaAs.

The band gap of InP is 1.35 eV and this material is used to produce laser light of wavelength 1.5 **µ**m. Similarly, GaN has a band gap equal to 3.36 eV. A laser made of GaN is used to Amit blue light and **Ultraviolet rays**.



Working of Semiconductor Diode Laser

The active medium of the semiconductor laser is a PN junction. In this laser, mirror is not used as in other resonator or cavity for optical feedback to sustain laser oscillation.

In this case, the reflectivity due to the refractive indices of two layers of a semiconductor laser is used for optical feedback. The end faces of two types of semiconductors (i.e. P-type and n-type) are cleaved and are perfectly parallel to each other for achieving optical feedback.

If the active medium or junction is made of a single type of semiconductor material, then the semiconductor laser is also known as **homojunction laser.**

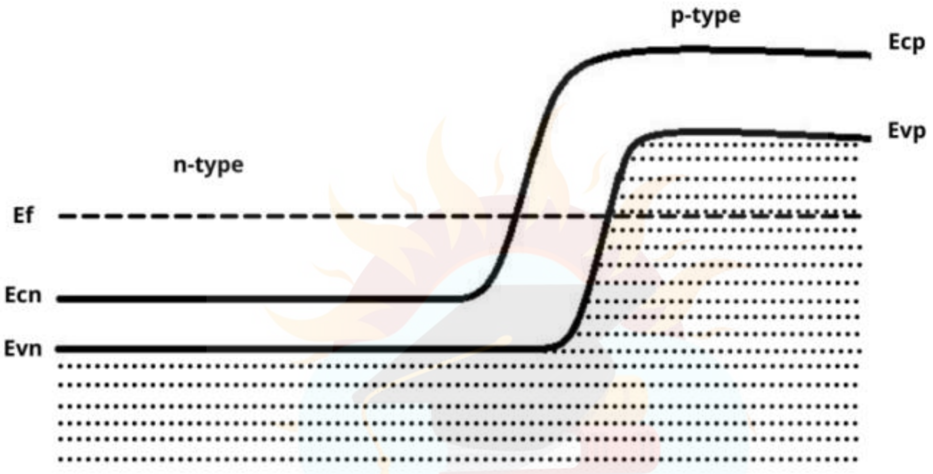
On the other hand, if the junction is made of different types of semiconductor materials, then the semiconductor laser is known as a **heterojunction laser**.

The population inversion in semiconductor diode (p-n junction diode) is achieved by forward biasing of PN junction diode.

The forward Bias Voltage causes the carrier pairs (i.e. electron in n region and hole in p region) to inject into the junction religion, where they recombine by means of stimulated emission.

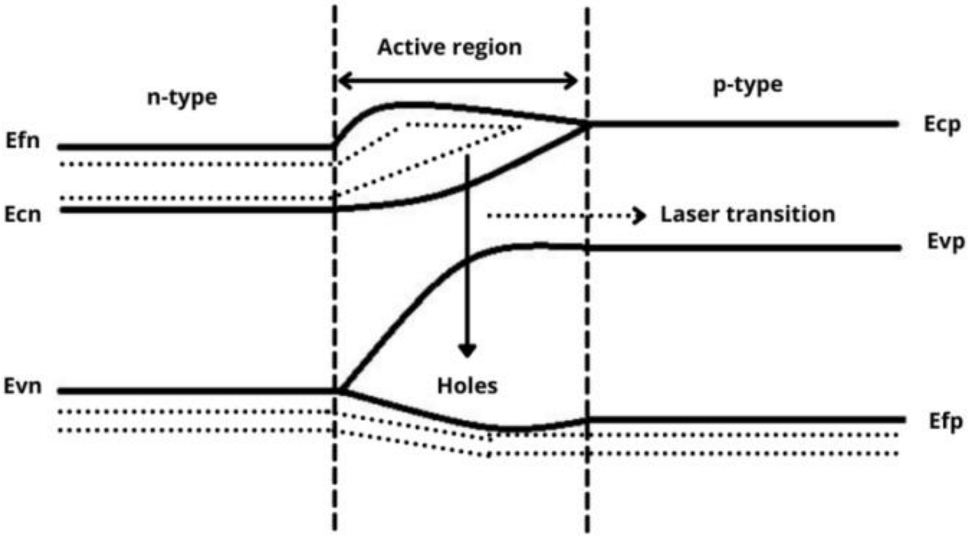
The process can be explained as follows: The valence band and conduction band of a p-n junction diode are shown in the figure.

At the equilibrium, the Fermi level is inside the conduction band of n-type semiconductor and it is inside the balance band of P-type semiconductor.



When a p-n junction is forward biased, the electrons will be injected into conduction band along and side, and the number of holes are produced in the valence band along the p-side of the junction.

Thus, there will be more electrons in the conduction band than that in the valence band. Hence, population inversion is achieved.

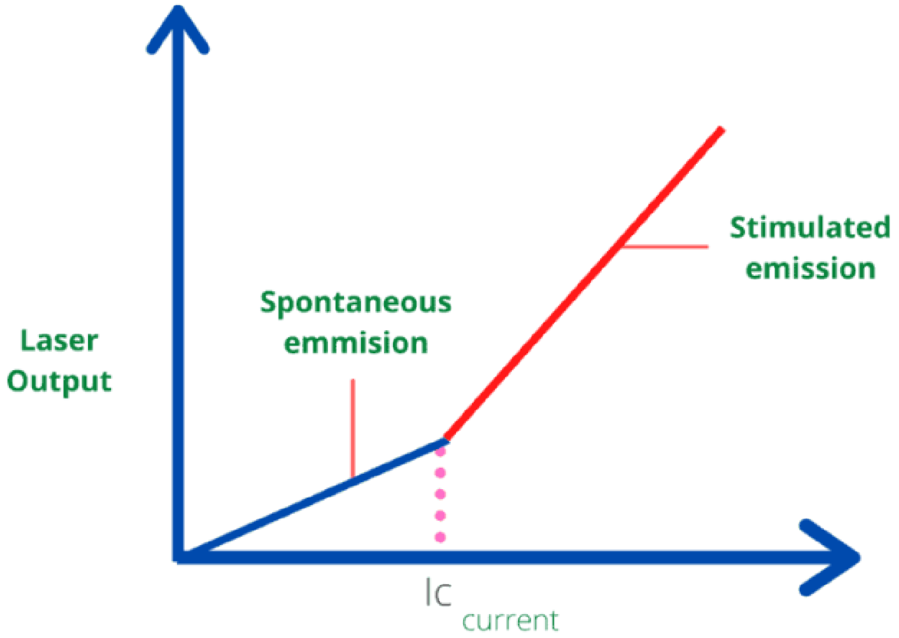


If the forward voltage is low i.e. forward current is small, the electron jumps from the conduction band to the valence band where they recombine with hole and emit incoherent light. This is the function of the light-emitting diode(LCD).

However, if the forward voltage is high i.e. forward current is large, the electron jumps from the conduction band to the valence band, then due to the recombination of electron and holes, a photon of energy equal to the forbidden energy gap (Eg) is emitted by **spontaneous emission** in the junction region of GaAs semiconductor.

This Photon stimulated the laser action near the junction. The laser action takes place in the narrow region and the **laser oscillation (i.e. optical feedback)** is done due to the repeated reflection between the cleaved end surfaces.

The variation of a typical output of a semiconductor laser with the current flowing through the laser shown in the figure.



When the forward current is low, the population inversion built compensate losers in the system i.e. absorption exceeds the gain and hence no lasing action takes place. In this case, laser output varies linearly with current.

When the current exceeds the critical value (Ic) called **threshold current,** the gain exceeds the losses in the system and lasing action takes place. In this case, laser output increases dramatically with an increase in the current.