UNIT 2

PHYSICS OF SOLIDS

INTRODUCTION

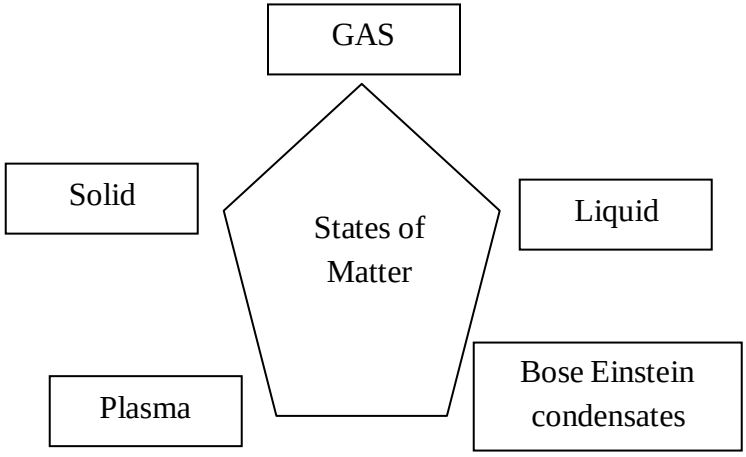


Fig. 2.1 States of Matter

There are five known phases in states of matter are, Gas, Liquid, Solids, Plasma and Nose Einstein condensates.

Gas has no definite volume and no definite shape. Gas particles have a great deal of space between them. Hence the atoms are more free to move. Gas particles have high kinetic energy.

The increasing pressure reduces the space between the particles and the pressure exerted by their collisions increases. Liquids have an indefinite shape. The liquid

particles are not held in a regular arrangement, but are still very close to each other so liquids have a definite volume.

Hence the atoms are free to move than solids. In the liquid phase, the particles of a substance have more kinetic energy than those in a solid. Liquids, cannot be compressed under pressure.

Solids have a definite shape. The atoms are tightly packed in a solid. Hence the atoms are unable to move freely. Particles of a solid have very low kinetic energy. The increasing pressure will not compress the solid to a smaller volume. Based on the arrangement of atoms, the solids are further classified into three major categories are (i) Crystalline solids (ii) Semi crystalline solids and (iii) amorphous solid.

Plasma is a charged gas, (neutral molecules and atoms) with strong Coulomb interactions**.** Plasma is not a common state of matter on Earth, but the most common state of matter in the universe. Plasma consists of highly charged particles with extremely high kinetic energy. The nobel gases such as helium, neon, argon, krypton, xenon and radon are often used to make glowing signs by using electricity to ionize them to the plasma state. Stars are essentially superheated balls of plasma.

In 1995, technology enabled scientists to create a new state of matter, the Bose-Einstein condensate (BEC). Using a combination of lasers and magnets, Eric Cornell and Carl Weiman cooled a sample of **Rubidium** to within a few degrees of absolute zero. At this extremely low temperature molecular motion comes very close tostopping altogether. Since there is almost no kinetic energy being transferred from one atom to another, the atoms begin to clump together. There are no longer thousands of separate atoms, just one “super atom”. BECs are also used to simulate conditions that might apply in black holes.

Let us briefly discuss about solids in our further discussions.

STRUCTURE OF SOLIDS

Solids are classified into following categories based on the arrangement of atoms. There are:

- **Amorphous solids**

An amorphous solid is a substance whose constituents do not possess an orderly arrangement and isotropic in nature. Important examples of amorphous solids are glass and plastics. Although amorphous solids consist of microcrystalline substance but the orderly arrangement is restricted to very short distances. They do not have definite heat of fusion and sharp melting point.

- **Semi (Poly) crystalline solids**

Semi-crystalline materials have a highly ordered molecular structure with sharp melt points. They have excellent chemical resistance.

- **Crystalline solids**

A crystalline solid is a substance whose constituents possess an orderly arrangement in a definite geometric pattern. Some very common examples of crystalline substances are sodium chloride, sugar and diamond. Crystalline solids exhibit definite heats of fusion.

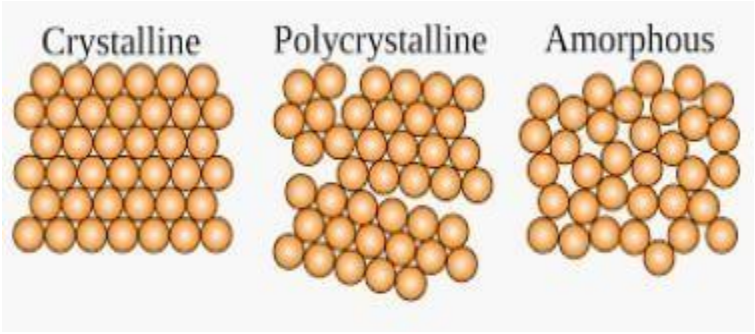
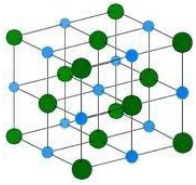
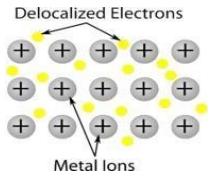
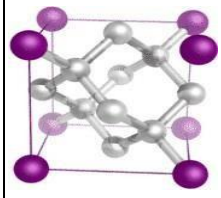
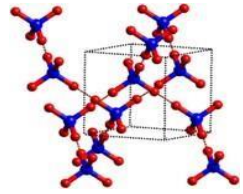


Fig. 2.2 Types of solids

Crystalline substances can be described by the types of particles in them and the types of chemical bonding that takes place between the particles. Based on chemical bonding the types of solids as follows:

2.1.1 Types of Crystalline Solids

There are four types of crystals: (1) **ionic** (2) **metallic** (3) **covalent** network, and (4) **molecular**. Properties and several examples of each type are listed in the following table and are described in the table below.

	Ionic Solid	**Metallic Solid**	**Covalent Solid**	**Molecular Solid**
1	The ionic crystal structure consists of alternating positively-charged cations and negatively-charged anions The ions may either be monatomic or polyatomic.	Metallic crystals consist of metal cations surrounded by a "sea" of mobile valence electrons These electrons, also referred to as delocalized electrons, do not belong to any one atom, but are capable of moving through the entire crystal.	A covalent network crystal consists of atoms at the lattice points of the crystal, with each atom being covalently bonded to its nearest neighbor atoms.	Molecular crystals typically consist of molecules at the lattice points of the crystal, held together by relatively weak intermolecular forces. Molecular solids are three types (i) Polar (ii) Non polar (iii)Hydrogen bonded
2	They have high melting points	They have wide range of melting points	They have high melting and boiling points.	They have low melting and boiling points.
3	Hard but brittle; shatter under stress	soft	very hard and brittle	Easily deformed under stress; ductile and malleable
4	Poor conductors of heat and electricity	Good conductors of heat and electricity than ionic crystal	Poor conductors of heat and electricity	Good conductors of heat and electricity
5	Relatively dense	low density	low density	high density
6	 NaCl crystal	 Iron	 Diamond	 Ice crystal structure

Several theories are successfully explained the various properties of solids and it is explained as follows.

ELECTRON THEORY OF SOLIDS

The electron theory of solids explains the structures and electrical, thermal properties of solids through their electronic structure. This theory also explains the binding in solids, behavior of conductors, semiconductors and insulators, Ferromagnetism, electrical and thermal conductivities of solids, elasticity, cohesive and repulsive forces in solids etc., the theory has been developed in three main stages, are.

(I) Classical free electron theory of metals

(II) Quantum free electron theory of metals

(III) Band theory of metals (or) Brillouin zone theory

The brief discussion about these theories as follows;

Classical free electron theory of metals

The classical free electron theory was introduced by P. Drude in 1900 and developed by Lorentz in 1909 to explain electrical conduction in metals. This theory has some assumptions / Postulates; they are:

1. The valence electrons of metallic atoms are free to move in the spaces between ions from one place to another place within the metallic specimen similar to gaseous molecules so that these electrons are called free electron gas. These electrons also participate in electrical conduction; hence they are called conduction electrons.
2. There is no interaction between these conduction electrons.
3. The interaction of free electrons with ion cores is negligible.
4. The free electrons find uniform electric field of positive ions and that of electrons in the metal.

5. Since the electrons are assumed to be a perfect gas they obey classical kinetic theory of gasses.
6. Classical free electrons in the metal obey Maxwell-Boltzmann statistics.

Postulates of free electron theory in absence of electric field:

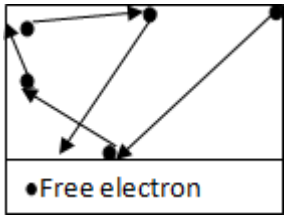


Fig. 2a Free motion of electron

The free electrons are moving randomly shown in Fig.2a within the boundaries of a metal similar to the gas molecules. The total energy of the electron is assumed to be purely about its kinetic energy because of its zero potential energy

Due to the free motion of free electrons the electrons are collides with each other elastically without losing their energy, called as free electrons.

In presence of electrical field:

To study electrical conductivity

In the absence of an external electric field, the free electrons move in random directions, similar to gaseous molecules in a container so that there is no net resultant motion of electrons along any direction.

When an electric field is applied to the conductor, the electrons are subjected to acceleration. The drift velocity of the electrons will depend upon the applied field, greater is the drift velocity and consequently, higher would be the number of electrons per second through unit area, i.e., the current passing through the conductor will be high.

To obtain an expression for electrical conductivity, consider a metallic rod of length *l* and area of cross section '*A*' as shown in Fig. 2b.

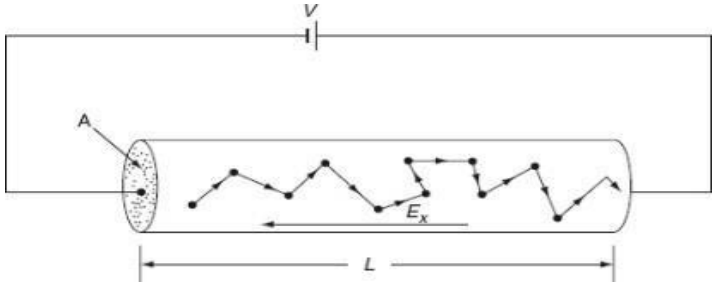


Fig. 2b Electron drifting in electric field (E_x)

The root mean square velocity of electrons is obtained by applying kinetic theory of gases and equipartition law of energy. The pressure '*P*' of the gas is:

$$P = \frac{1}{3}\rho \bar{C}^2 = \frac{1}{3}mn\bar{C}^2 \qquad (1)$$

where *m* is mass of an electron, '*n*' is the number of free electrons per unit volume, 'ρ' is the density of an electron gas and $\bar{C}$ is root mean square velocity.

For molar volume (V_m) of the metal,

$$P = \frac{1}{3}\rho \bar{C}^2 = \frac{1}{3}m\frac{N_A}{V_m}\bar{C}^2 \qquad (2)$$

where N_A is Avogadro's number

From Equation (2)

$$PV_m = \frac{1}{3}mN_A\bar{C}^2 = RT_u \text{ (or)} \quad mn\bar{C}^2 = \frac{3R_uT}{N_A} = 3KT_B \qquad (3)$$

where R_u is universal gas constant, *T* is absolute temperature and K_B is Boltzmann's constant.

From Equation (3)

$$\bar{C}^2 = \sqrt{\frac{3K_BT}{m}} \qquad (4)$$

Now a potential difference (V) is applied across the ends of the rod, kept parallel to X-direction, then an electric field, E_x acts on the electrons, then the electrons acquirea constant velocity v_x in the opposite direction to that of electric field. This constant velocity is called drift velocity, this velocity is superimposed on the thermal velocity ($\bar{C}$) of the electrons.

If 'R_x' is the resistance of the rod, then the current I_x through it is:

$$I_x = \frac{V_x}{R_x} = AJ_x \qquad (5)$$

where J_x is the current density along X-direction

(or) $$J_x = \frac{V_x}{AR_x} = \frac{E_x l}{A} \times \frac{A}{\rho_x l} = \frac{E_x}{\rho_x} = \sigma_x E_x \qquad (6)$$

where E_x = intensity of electric field; ρ_x = resistivity and σ_x = electrical conductivity of the metal.

The field E_x produces a force $-eE_x$ on each electron, due to this force the acceleration on ith electron is:

$$a_{ix} = \frac{dv_{ix}}{dt} = -\frac{eE_x}{m} \qquad (7)$$

where v_{ix} is the drift velocity of ith electron along X-direction. Since the electrical force eE_x is the same on all electrons, so the average acceleration is:

$$\frac{d}{dt}\langle v_x \rangle = -\frac{eE_x}{m} \qquad (8)$$

where $\langle v_x \rangle$ is the average velocity of *n* free electrons in unit volume of metal, given by:

$$\langle v_x \rangle = \frac{1}{n} \sum_{i=1}^{n} v_{ix} \qquad (9)$$

The current density along X-direction is:

$$J_x = n \langle v_x \rangle (-e) = -ne \langle v_x \rangle \qquad (10)$$

The minus sign shows that J_x is in the opposite direction to that of $\langle v_x \rangle$. The average velocity $\langle v_x \rangle$ can be obtained by considering the acceleration and retardation in steady state.

In addition to acceleration, $[eE_x/m]$ the electrons get retarded due to collisions of electrons with lattice ions. This is referred to as electron lattice scattering.

Where,

$$\langle v_x \rangle = -\left(\frac{eT_x}{m}\right) E_x \qquad (11)$$

where τ_x is known as relaxation time. It is defined as the time taken by an electron to decay its drift velocity to 1/*e* of its initial value.

Equation (11) shows that the drift velocity acquired by the electron is proportional to applied electric field and the proportionality constant ($e\tau_x/m$) is known as their mobility (μ_x). It is defined as the drift velocity produced per unit applied electric field and it is given by:

$$\mu_x = \frac{\langle v_x \rangle}{E_x} = \frac{eT_x}{m} \qquad (12)$$

Finally, comparing Equations $J_x = \sigma_x E_x$ and $J_x = -ne \langle v_x \rangle$, we have:

$$\sigma_x E_x = -ne \langle v_x \rangle \qquad (13)$$

Substituting Equation (11) in (13)

$$\sigma_x E_x = -ne\left(-\frac{eT_x}{m}\right)E_x$$

$$\therefore \sigma_x = \frac{ne^2}{m}\tau_x \qquad (14)$$

Substituting Equation (12) in (14) gives:

$$\sigma_x = ne\mu_x \qquad (15)$$

(or) resistivity, $\rho_x = \frac{1}{\sigma_x} = \frac{1}{ne\mu_x}$ (16)

Equations (14), (15) and (16) represent electrical conductivity and resistivity of a material.

Basic definitions:

- Drift velocity (v_d): It is defined as the average velocity acquired by the free electron in a particular direction due to the application of electric field.
- Relaxation time (τ): It is the time taken by the free electron to reach its equilibrium position from its disturbed position in the presence of applied field.
- Collision time (τ_c): It is the average time taken by a free electron between two successive collisions.
- Mean free path (λ): The average distance travelled between two successive collisions is called mean free path

Thermal Conductivity (K)

Definition:

Thermal conductivity of the material is defined as the amount of heat conducted per unit area per unit time maintained at unit temperature gradient.

$$Q = -K \frac{dT}{dx}$$

where, dT/dx = temperature gradient

Q = heat flux or heat density

K = the coefficient of thermal conductivity of the material.

Expression for Thermal Conductivity of a metal:

On the basis of kinetic theory, if two temperatures are equal, if T1 = T2 , then there is no transfer of energy. If T1 > T2 , there is transfer of energy from A to B (Fig.)

Consider a uniform rod AB with the temperature T_1 at end A and T_2 at the other end.

The two cross sections A and B are separated by a distance λ. Heat flows from hot end A to the cold end b by the electrons. The amount of heat (Q) conducted by the rod from the end A to be of length λ is given by,

$$Q = \frac{KA(T_1 - T_2)t}{\lambda} \qquad (17)$$

∴ The coefficient of conductivity per unit area per unit time

$$K = \frac{Q}{(T_1 - T_2)} \lambda \qquad (18)$$

where, K = coefficient of thermal conductivity

A = Area of cross section of the rod In Collision, the electrons near A loses their kinetic energy while the electrons near B gains the energy.

The average kinetic energy of the electron near A

$$A = \frac{3}{2} K T_B \qquad (19)$$

The average kinetic energy of the electron near B

$$B = \frac{3}{2} K T_{B\,2} \qquad (20)$$

Where, K_B = Boltzmann constant

∴Excess kinetic energy carried by the electron from A to B (eqn. 19 - eqn.20)

$$= \frac{3}{2} K T_{B\,1} - \frac{3}{2} K T_{B\,2}$$

$$= \frac{3}{2} K_B (T_1 - T_2) \qquad (21)$$

Let us assume that there is equal probability for the electrons to move in all the 6 directions as shown in below Fig.2c. Since each electron travels with thermal velocity 'v' , if n is the free electron density, then on an average 1/6 nv.

Electrons will travel in any one direction. Then, The number of electrons crossing unit area per second in one direction from A to B

$$= \frac{1}{6} nv$$

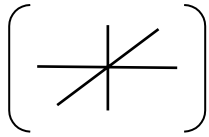


Fig.2c (6

Excess of energy transferred from A to B for unit area in unit time

$$= \frac{1}{6} nv[\text{K. Eat A} - \text{K.Eat B}] \qquad (23)$$

We know that Kinetic energy K.E

Therefore the total energy transferred

$$= \frac{1}{2} mv^2 = \frac{3}{2} K_B T$$

$$= \frac{1}{6} nv \left[\frac{3}{2} K_B T_1 - \frac{3}{2} K_B T_2 \right]$$

$$= \frac{1}{6} nv \frac{3}{2} K_B [T_1 - T_2]$$

$$= \frac{nvK_B (T_1 - T_2)}{4} \qquad (24)$$

Similarly, deficiency of energy carried from B to A for unit area in unit time

$$= \frac{nvK_B (T_1 - T_2)}{4} \qquad (25)$$

Hence net energy transferred from A to B per unit area per unit time (or) heat flux Q

is given by If K be the thermal conductivity of the material, then the transfer of energy per unit area per unit time is given by

$$Q = \frac{nvK_B(T_1 - T_2)}{4} - \left(-\frac{nvK_B(T_1 - T_2)}{4}\right)$$

$$Q = \frac{nvK_B(T_1 - T_2)}{4} + \frac{nvK_B(T_1 - T_2)}{4}$$

i.e. $$Q = \frac{nvK_B(T_1 - T_2)}{4} \quad (26)$$

From Eqns.(26) and (27)

$$-\frac{K(T_1 - T_2)}{\lambda} = \frac{nvK_B(T_1 - T_2)}{2} \quad (27)$$

$$\therefore K = \frac{nvK_B\lambda}{2} \quad (28)$$

Where λ is the mean free path of an electron.

We know that for metals, Relaxation time = Collision time

i.e., $\tau = \tau_c = \frac{\lambda}{v}$

(or) $$\tau v = \lambda \quad (29)$$

Substituting eqn.(29) in equation (28), we get

$$K = \frac{nvK_B\tau}{2} \quad (30)$$

Equation (30) is the classical expression for thermal conductivity.

Wiedemann – Franz law (Derivation) Statement

The ratio between the thermal conductivity and the electrical conductivity of a metal is directly proportional to the absolute temperature of the metal. This ratio is

constant for all metals.

i.e., $\frac{K}{\sigma} \propto T$ (or) $\frac{K}{\sigma} = LT$ (A)

Where L is known as Lorentz number whose value is, $2.44 \times 10^{-8}\ W\Omega K^{-2}$ at 293 K

Derivation

Wiedemann-Franz law can be derived using the expressions of thermal conductivity and electrical conductivity of metal. Electrical conductivity of metals is given by,

$$\sigma = \frac{ne^2\tau}{m} \quad (31)$$

From eqn (14)

Thermal conductivity is given by,

$$K = \frac{nv^2 K_B \tau}{4} \quad (32)$$

From eqn (30)

From eqn (31) and eqn (32) and (A)

$$\frac{\text{Thermal conductivity}}{\text{Electrical conductivity}} = \frac{K}{\sigma} = 2\frac{\frac{1}{4} nv^2 K_B \tau}{\frac{ne^2\tau}{m}}$$

$$\frac{K}{\sigma} = \frac{1}{2}\frac{mv^2 K_B}{e^2} \quad (33)$$

We know that the kinetic energy of an electron

$$\frac{1}{2}mv^2 = \frac{3}{2}K_B T \qquad (34)$$

From eqn (33) and eqn (34) we get,

$$\frac{K}{\sigma} = \frac{3}{2}\left(\frac{K_B}{e}\right)^2 T \qquad (35)$$

(or) $$\frac{K}{\sigma} = LT \qquad (36)$$

Where , L is a constant known as Lorentz Number. By comparing (35) and (36);

$$L = \frac{3}{2}\left(\frac{K_B}{e}\right)^2$$

(or) $$\frac{K}{\sigma} \propto T \qquad (37)$$

Thus the ratio of the thermal conductivity to electrical conductivity of a metal is directly proportional to the absolute temperature of the metal.

This proves Wiedemann-Franz law for a conducting material.

Lorentz Number: Lorentz number is a constant given by

$$L = \frac{3}{2}\left(\frac{K_B}{e^2}\right)^2$$

$K_B = 1.39 \times 10^{-23}\ JK^{-1}$

$e = 1.6021 \times 10^{-19}$ Coulomb

$$\therefore L = \frac{3}{2}\left(\frac{1.38\times10^{-23}}{1.6021\times10^{-19}}\right)^2 = 1.12 \times 10^{-8}\ W\Omega K^{-2}$$

2（1.6021×10　）

i.e. $L = 1.12 \times 10^{-8}\ W\Omega K^{-2}$

where, It is found that the classical value of Lorentz number is only half of the experimental value i.e., $2.44 \times 10^{-8}\ W\Omega K^{-2}$

This discrepancy in the experimental and theoretical value of Lorentz number is the failure of classical theory. This discrepancy is rectified by quantum theory.

Merits of classical free electron theory:

- Classical theory successfully explains electrical and thermal conductivities of metals, the increase of resistance with temperature, opacity, luster, etc.,
- It verifies Ohm's law
- It is used to derive Wiedemann Franz law
- It is also used to explain optical properties

Drawbacks of classical free electron theory:

- It is a macroscopic theory
- There is an controversy in Classical theory and quantum theory; Classical theory states that all free electrons will absorb energy but quantum theory states that only few electron will absorb energy.
- This theory cannot explain the Compton effect , photo electric effect and about magnetism.
- The theoretical and experimental electronic specific heat value is not agreed well.
- According to classical theory, the electrical conductivity (σ) of a metal is inversely proportional to its square root of its absolute temperature (T). i.e.,

$\sigma \propto \frac{1}{\sqrt{T}}$. But practically, it has been observed that the electrical conductivity is inversely proportional to its absolute temperature. i.e., $\sigma \propto \frac{1}{\sqrt{T}}$

- The Lorentz number is mismatched in classical theory and it is rectified in quantum theory.

Quantum free electron theory

- This theory is proposed by Sommerfeld in 1928 and It is a Microscopic theory.
- It is explained that the electrons move in a constant potential and it obeys quantum laws
- In this theory it is assumed that the electron is moving with a constant potential
- The mass of an electron is constant, when it moves in constant potential
- This theory states that the electrons are moving with higher velocities the potential energy of an electron becomes zero.
- This theory successfully explains electrical conductivity, specific heat, thermionic emission and paramagnetism.
- The quantum theory verified Wiedemann Franz law and proved the value of Lorentz number by using the concept of effective mass of an electron.
- This theory failed to explain some of the physical properties are (i) it fails to explain the difference between conductors, insulators and semiconductors, (ii) positive Hall coefficient of metals and (iii) lower conductivities of divalent metals than monovalent metals.

  Wiedemann Franz law by quantum theory:

  In quantum theory the mass of an electron is replaced by effective mass of an electron m*,

  Hence the expression of electrical conductivity eqn. (14) becomes,

$$\sigma = \frac{ne^2\tau}{m^*} \qquad (1)$$

The thermal conductivity expression eqn (30) becomes,

$K = 2\,(\pi/3)^2 \times 3/2\,[nK_B^2T\tau/m]$

$[1/2mv^2 = 3/2\ K_BT;$ Hence $v^2 = 3\ K_BT/m]$

$K = \pi^2nK_B^2T\tau/3m^*$ (2)

By using (A), (1) and (2)

$K/\sigma = \pi^2/3\,(K_B/e)^2\,T$ (3)

Hence $L = \pi^2/3\,(K_B/e)^2 = 2.44 \times 10^{-8}\ W\Omega K^{-2}$

This shows that, quantum theory verified Wiedemann Franz law and good agreement with Lorentz number.

Brillouin Zone theory / Band Theory

- This theory is proposed by Bloch in 1928.
- It is explained the concept of electron moves in a periodic potential.
- According to the zone theory, the mass of an electron varies when it moves through the periodic potential and it is called as effective mass of an electron.

This theory also explains the mechanism of Conductors, semiconductors and insulators based on bands and hence called as band theory.

Let us discuss about formation of bands in solids as follows.

ORIGIN OF ENERGY BANDS

Generally solid consists of more number of atoms and each atom consists of the central nucleus of positive charge around which small negatively charged particles called electrons revolve in different paths or orbits. An Electrostatic force ofattraction between electrons and the nucleus holds up electrons in different orbits.

An isolated atom possesses discrete energies of different electrons. If two isolated atoms are brought to very close proximity, then the electrons in the orbits of two atoms interact with each other. So, that in the combined system, the energies of electrons will not be in the same level but changes and the energies will be slightly lower and larger than the original value. So, at the place of each energy level, a closely spaced two energy levels exists. If 'N' number of atoms are brought together to form a solid and if these atoms' electrons interact and give 'N' number of closely spaced energy levels in the place of discrete energy levels, it is known as bands of **allowed energies/band.** Between the bands of allowed energies, a small region does not allow any energy levels. Such regions are called **forbidden energy band/ energygap** (E_g). The allowed bands of energies and forbidden bands of energies are present alternatively one after another for the electrons of a solid (Fig.2.3).

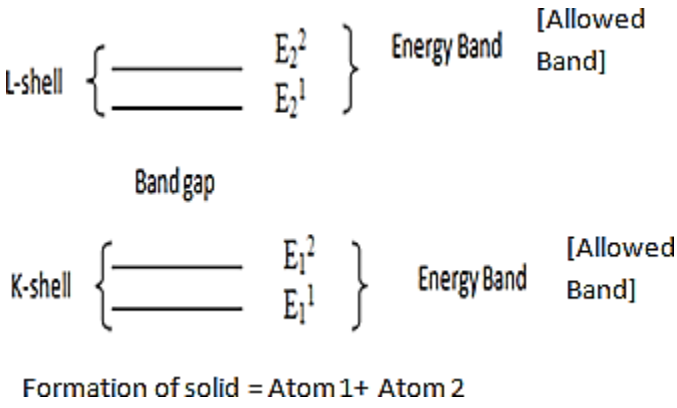


Fig. 2.3 Band formation in solids

When two atoms of equal energy levels are brought closer together, the k shell energy levels of individual atoms lie in the energy level E_1 and splits into E_1^1 and E_1^2 become as **energy band.** Similarly for L-shell energy levels lies in the energylevel E_2 and splits into E_2^1 and E_2^2. Similarly when three atoms are brought together, the energy levels are in K – shell and L-shell are E_1^1 , E_1^2 and E_2^1, E_2^2 respectively.

Similarly for K, L,M and N shells are E_1^1, E_1^2, E_1^3, E_1^4 and E_2^1, E_2^2, E_2^3, E_2^4 respectively. These type of transformations from the original energy levels into two (or) more energy levels is known as energy level splitting.

In example Lithium atom consists of 3 electrons. Among them the two electrons are in the same energy state (1s) and the other one is in the higher energy state (2s). In a solid of lithium contains N number of atoms, the lower energy level forms a band of 2N electrons occupying N different energy levels are shown in Fig. 2.4. The higher energy level can form a band of 2N electrons (completely filled). But only N numbers of electrons are present in the next energy level. Therefore the band is half-filled band. The unoccupied band of the solid corresponds to the un occupied excitation level of the isolated atom.

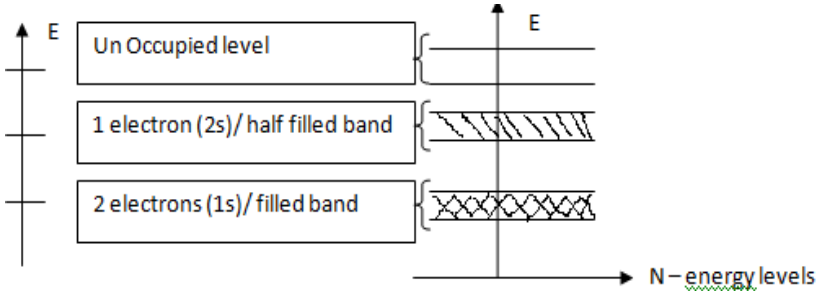


Fig.2.4 (a) Isolated atom and (b) Solid of N atoms

The width of the energy band depends on the relative freedom of electrons in the crystal. The electrons in lower filled energy (closely bounded with nucleus) band do not have any freedom to move in the crystal. The electrons in half filled energy level (loosely bound with nucleus) are free to move inside the crystal. Generally the half-filled band is called **valence band** and the upper unoccupied (unfilled) band is called **conduction band**. In between the valence band and conduction the gap is known as energy gap

The differences between metals, semiconductors and insulators can be made based on (i) the width of the energy gap (ii) number of effective electrons (The

availability of number of free electrons that participate in electrical conduction per unit volume when an electric field is applied. These free electrons are called **effective electrons)**.

BAND STRUCTURE OF CONDUCTOR, SEMICONDUCTORS, INSULATORS, HALF METALS AND SEMI METALS

Conductors:

In case of conductors, there is no forbidden band and the valence band and conduction band overlap each other (Fig.2.5). Low resistive materials and it is about $10^{-8}\Omega$m are generally called as conducting materials. These materials have high electrical and thermal conductivity. The conducting property of a solid is not a function of total number of the electrons in a metal and it is about the number of free/valence electrons called as conduction electrons. Hence in metals the electrical conductivity depends on the number of free electrons. Here, plenty of free electrons are available for electric conduction. The electrons from valence band freely enter in the conduction band due to overlapping of bands.

Examples: Iron, Copper etc.,

Semiconductors:

In semiconductors, the forbidden band is very small (Fig. 2.5). Germanium and silicon are the best examples of semiconductors. In Germanium the forbidden band is of the order of 0.7eV while in case of silicon, the forbidden gap is of the order of 1.1eV. Actually, a semi-conductor material is one whose electrical properties lie between insulators and good conductors. At 0°K, there are no electrons in conductionband and the valence band is completely filled. When a small amount of energy is supplied, the electrons can easily jump from valence band to conduction band due to minimum energy gap. For example, when the temperature is increased, the forbidden band is decreased so that some electrons are liberated into the conduction band. In semi-conductors, the conductivities are of the order of 10^2 ohm-meter.

Examples: Silicon, Germanium etc.,

The detailed description about the types of semiconductor as follows in section 2.7.

Insulator

In case of insulators, the forbidden energy band is wide (Fig.2.5). Due to this fact, electrons cannot jump from valence band to conduction band. In insulators, the valence electrons are bound very tightly to their parent atoms. For example, in case of materials like glass, the valence band is completely filled at 0K and the energy gap between valence band and conduction band is of the order of 6 eV. Even in the presence of high electric field, the electrons do not move from valence band to conduction band. When a very large energy is supplied, an electron may be able to jump across the forbidden gap. Increase in temperature enables some electrons to go to the conduction band. This explains why certain materials which are insulator become conductors at high temperature. The resistively of insulators is of the order of 10^7 ohm-meter.

Examples: Wood, Glass etc.,

At absolute zero (0 K) the energy levels are completely occupied by electrons up to a certain level. Above that the energy levels are completely unoccupied by electrons. This highest level which is completely filled by electrons is called **Fermi level**. In conductor the Fermi level lies in the permitted band (since the valence and conduction band overlap without an energy gap). In Semiconductor the Fermi level lies in the small energy gap. In Insulator the Fermi level lies exactly in the middle of energy gap.

The **Fermi energy** (E_F) is described as the highest energy that the electrons assumes at a temperature of 0 K (Fig.2.5). The Fermi energy is the difference in energy, mostly kinetic. To put this into perspective one can imagine a cup of coffee and the cup shape is the electron band; as one fills the cup with the liquid the top surface increases. This can be compared to the Fermi energy because as electrons enter the electron band the Fermi energy increases. The Semiconductors act differently than the above stated analogy.

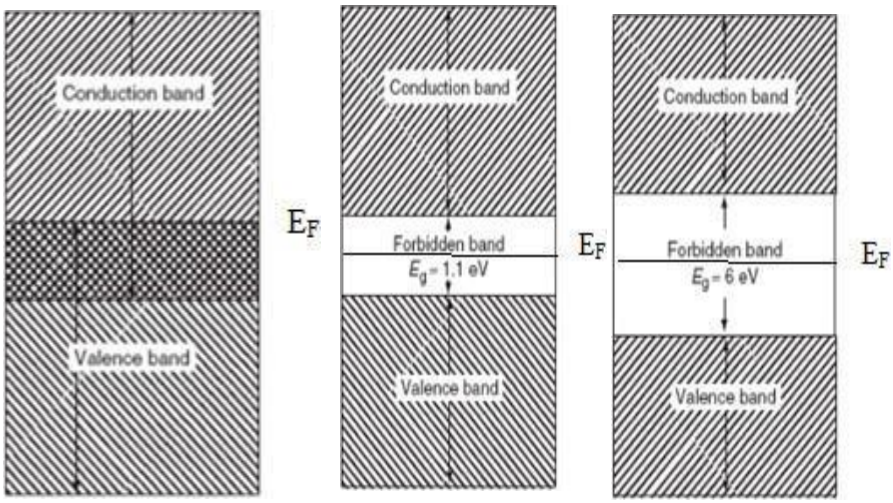


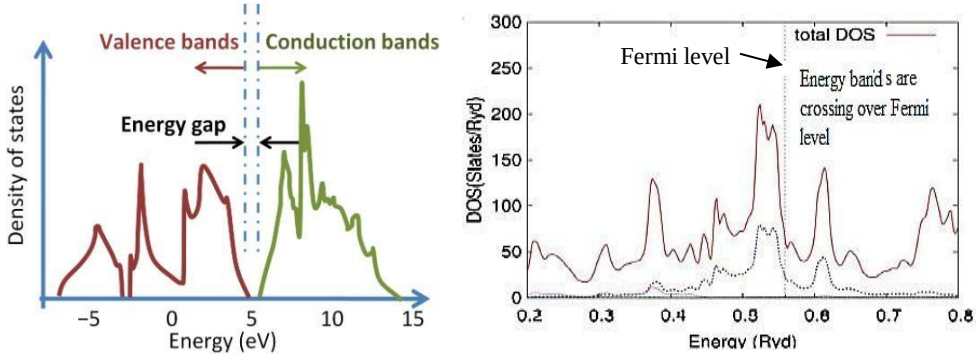
Fig. 2.5 (a) Conductor (b) Semiconductor (c) Insulator

Because in an n-type semiconductor will have a Fermi energy closes to the conduction band, whereas a p-type semiconductor will have Fermi energy close to the valence band. In metals this means that it gives us the velocity of the electrons during conduction. So during the conduction process, only electrons that have an energy that is close to that of the Fermi energy can be involved in the process. The concept of the Fermi energy is vital key to understanding electrical and thermal properties of solids. Another factor that the Fermi energy plays is in the role of understanding specific heat of solids at room temperatures.

The Fermi level describes the probability of electrons occupying a certain energy state, but in order to correctly associate the energy level the number of available energy states need to be determined is known as **Density of States (N (E_F)).**It will be useful to study about the conducting and physical properties of the materials. It gives the number of allowed electron (or) hole states per volume at a

given energy. In the energy vs Density of States profile if the electron states are crosses the Fermi level are considered as metals or else it is considered as insulators [refer section 2.4.4 and 2.4.5].

Example:



Density of states (DoS) histogram for insulator (Left) Conductor (Right)

Fermi Surface

At absolute zero, the Fermi surface (constant energy surface in k-space) is the highest filled energy surface in k or wave vector space (The k-space is the relation of particle energy (E) versus the momentum as a wave number (k)). Alternatively it can be defined as the boundary between the filled and empty states in ground state of the crystal. The Fermi surface is a useful tool to predict thermal, electrical, magnetic, and optical properties of materials.

Characteristics:

First Brillouin

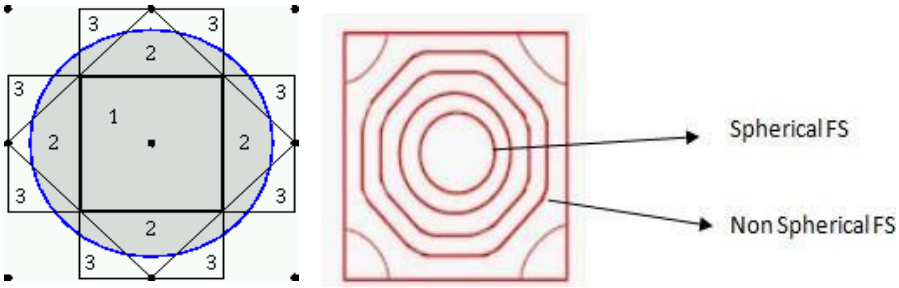


Fig. 2.6 (a) Fermi surface (FS) in Brillouin Zone (b) Shapes of FS

- The volume of the Fermi Surface represents the number of conduction electrons.
- The shape of the Fermi surface (Fig. 2.6) is derived from the periodicity and symmetry of the crystalline lattice and energy bands.

The Fermi surface is spherical within the first Brillouin zone, when the velocity (v) in k-space is proportional and parallel to the wave vector K.

The Fermi surface is non-spherical in higher zone, when the velocity is proportional to the gradient of the energy in k-space.

- The degree of distortion depends (a) how near the FS to the zone boundaries (b) The magnitude of the effective potential
- When an electrical field is applied, the Fermi surface shifts in the opposite direction of the field direction. This is shown below in Fig. 2.7. Even though the electrons are cancelled out due to opposite direction of motion, someelectrons remain un compensated resulting in a net current.

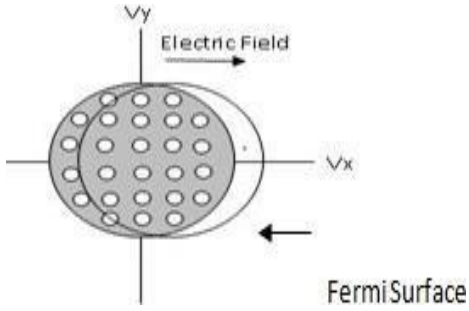


Fig.2.7 Due to electric field there is displacement in the Fermi surface

- It is used to know about the heat capacity, Pauli's paramgnetism and electrical conductivity etc.,

Semi metals

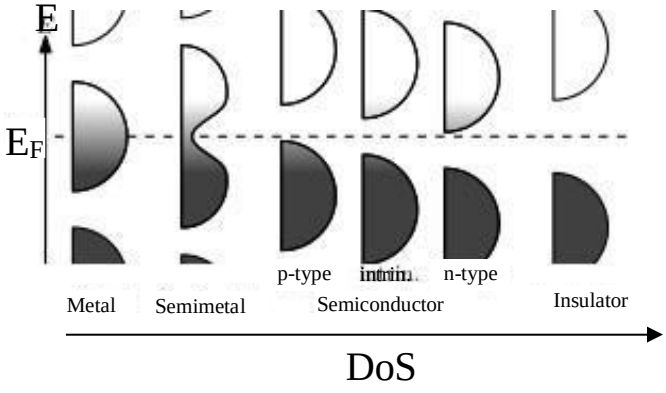


Fig. 2.8 Comparison of Band structure (Energy vs DoS)

One additional type of electron energy spectrum for solids is a "semimetal," two distinct energy bands overlap slightly at the Fermi level (shown in Fig.2.8). The semimetals are metals. Semimetals have properties intermediate between metals and semiconductors. Indeed they do allow conduction at absolute zero temperature. How-

ever, their further behavior is noticeably different from true metals because the overlap of the two bands is only small. One difference is that the electrical conduction of semimetals increases with temperature, unlike that of metals. Likely as semiconductors, semimetals at a higher temperature means that there are more electrons in the upper band and more holes in the lower band.

Examples: Arsenic, Antimony, Bismuth and a single sheet of carbon, called graphene, is right on the boundary between semimetal and semiconductor.

Half Metals

Half-metals were first described in 1983, as an explanation for the electrical properties of Mn-based Heusler alloys. Half-metals have attracted some interest for their potential use in spintronics.

A **half-metal** is any substance that acts as a conductor to electrons of one spin orientation, but as an insulator or semiconductor to those of the opposite orientation. Although all half-metals are ferromagnetic types but all ferromagnets are not half-metals. A ferromagnetic material with high spin polarization above room temperature is most important in the spintronics area.

In half-metals, the valence band for one spin orientation is partially filled while there is a gap in the density of states for the other spin orientation. This results in conducting behavior for only electrons in the first spin orientation. In some half-metals, the majority spin channel is the conducting while in others are in the minority channel.

The spin polarization (P) is defined as the ratio of the density of states (D) of up-spin and down-spin electrons at a Fermi level, $P=(D_{up}-D_{down})/(D_{up}+D_{down})$, asshown in Fig. 2.9. Since the density of states of up-spin and down-spin electrons are equal in paramagnetic materials, P=0 for paramagnetic materials. On the other hand, since the density of state of up-spin and down-spins are different in ferromagnetic materials, P is larger than 0, but smaller than 1 for ferromagnetic materials.

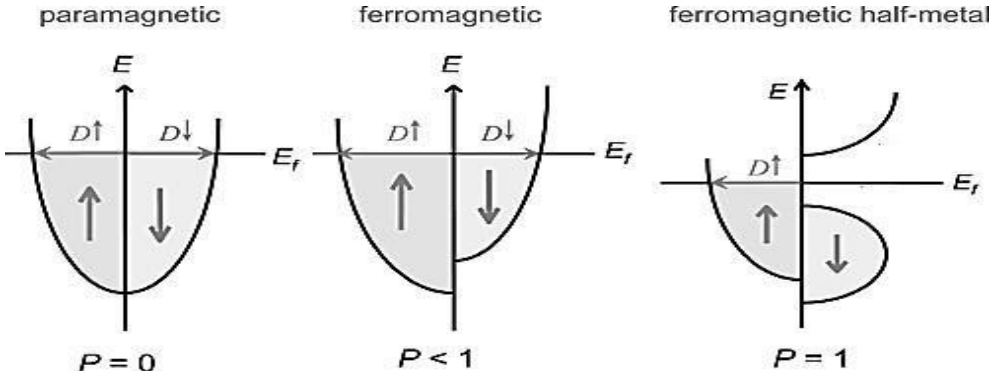


Fig. 2.9 Density of states of paramagnetic, ferromagnetic and ferromagnetic half-metal materials

The P values of Fe and Co are known to be around 0.5. If a material has a band gap in the minority band (semiconducting) at a Fermi level and exhibits metallic behavior in the majority band, the density of state of the minority band is zero at the Fermi level. In this case, since only up-spin electrons are present at the Fermi level, P=1. This type of material is called "half-metal" because it exhibits both metallic and semiconducting behaviors.

Examples: chromium (IV) oxide, magnetite, lanthanum strontium manganite (LSMO), chromium arsenide. Many of the known examples of half-metalsare oxides, sulfides, or Heusler alloys.

SEMICONDUCTING MATERIALS

Classification of Semiconductors based on Energy band theory

Semiconductors are classified into two types. Intrinsic semiconductors and Extrinsic semiconductors.

a) Intrinsic Semiconductors: A semiconductor in an extremely pure form is known as Intrinsic semiconductor. The vacant conduction band is separated by a forbidden energy gap E_g from the filled valence band. So at absolute zero (0^0k), there is no electric

conduction. Fermi level is at the middle of valance band and conduction band.

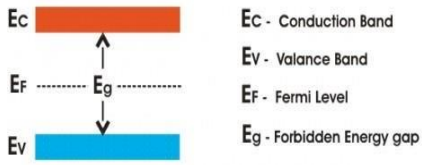


Fig. 2.11. Intrinsic semiconductor

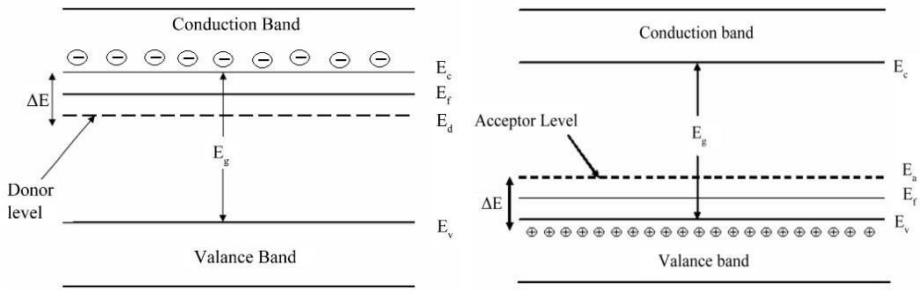
b) Extrinsic Semiconductors: When an impurity is added to an intrinsic semiconductor its conductivity changes. This process of adding impurity to a semiconductor is called Doping and the impure semiconductor is called extrinsic semiconductor. Depending on the type of impurity added, extrinsic semiconductors are further classified as n-type and p-type semiconductor.

n-type semiconductor

When an impurity (electron) is added to an intrinsic semiconductor its conductivity changes and is called n-type semiconductor. This extra electron is forming a new energy level called **donor energy level** and is situated just below the conduction band. Fermi level is at the middle of donor energy level and conduction band.

p-type semiconductor

When an impurity (hole) is added to an intrinsic semiconductor its conductivity changes and is called p-type semiconductor. This extra hole is forming a new energy level called **acceptor energy level** and is situated just above the valance band. Fermi



level is at the middle of acceptor energy level and conduction band.

Fig. 2.12 (a) n-type semiconductor (b) p-type semiconductor

The detailed description about intrinsic and extrinsic semiconductor as follows

(a) Intrinsic or Pure Semiconductor:

A semiconductor in an extremely pure form is known as Intrinsic semiconductor. Example: Silicon, germanium. Both silicon and Germanium are tetravalent (having 4 valence electrons). Each atom forms a covalent bond or electron pair bond with the electrons of neighbouring atom. The structure is shown below.

At absolute zero temperature 0^0 k, all the valence electrons are tightly bounded the nucleus hence no free electrons are available for conduction. The semiconductor therefore behaves as an Insulator at absolute zero temperature. The vacant conduction band is separated by a forbidden energy gap E_g from the filled valence band.

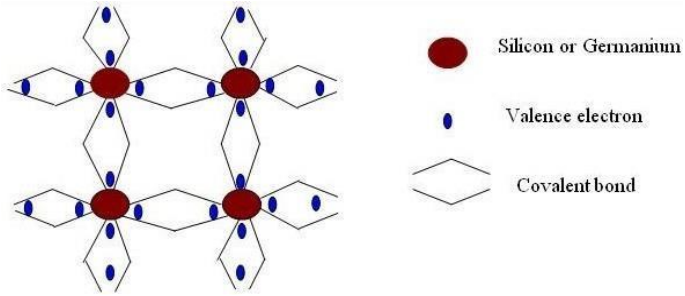


Fig. 2.13 Intrinsic semiconductor

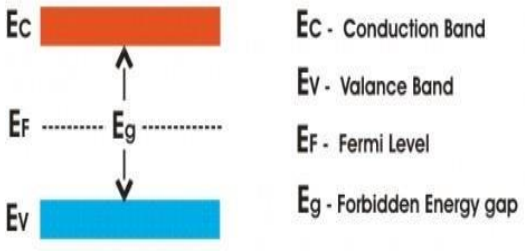
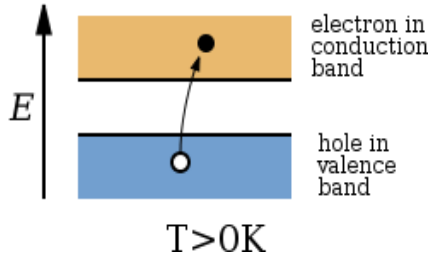


Fig. 2.14 (a) Band Structure of an Intrinsic Semiconductor at 0°K



(b) Intrinsic Semiconductor – Holes and Electrons

When the temperature is increased, the electrons are thermally excited from the valence band to the conduction band. So they (electrons) become free in the conduction band. Therefore, electric conduction is possible in the conduction band. Due to this transition, the number of electrons in the conduction band will be more, but the same numbers of electrons are missing in the valence band. This gives the same number of 'holes' in the valence band as shown in figure. These holes act as a positive charge carrier. The electrons in the conduction band and holes in the valence band are responsible for electric conduction.

When a hole is created in the valence band, the electron from the nearby covalent bond (in the valence band) breaks its own bond and moves forward to fill

the previous hole. Therefore, a new covalent bond is again formed in the valance band and this creates another hole in the valence band. This hole is considered as particle similar to the electron and it have positive charge with definite mass. In the conduction band conduction is possible by electrons, since the electrons in the conduction band move freely inside the crystal. In the valence band, conduction is possible by holes since the movement of electrons in the valence band is restricted.

(b) Extrinsic semiconductors: When an impurity is added to an intrinsic semiconductor its conductivity changes. This process of adding impurity to a semiconductor is called Doping and the impure semiconductor is called extrinsic semiconductor. Depending on the type of impurity added, extrinsic semiconductors are further classified as n-type and p-type semiconductor.

n-type semiconductor: When a small amount of pentavalent impurity is added to a pure semiconductor (extrinsic semiconductor) is called as **n-type semiconductor.** Addition of Pentavalent impurity provides a large number of free electrons in a semiconductor crystal. Typical example for pentavalent impurities is Arsenic, Antimony and Phosphorus etc. Such impurities which produce n-type semiconductors are known as Donor impurities because they donate or provide free electrons to the semiconductor crystal.

To understand the formation of n-type semiconductor, consider a pure silicon crystal with impurity say arsenic added to it. We know that a silicon atom has 4 valence electrons and Arsenic has 5 valence electrons. When Arsenic is added as impurity to silicon, the 4 valence electrons of silicon make co-valent bond with covalent bond 4 valence electrons of Arsenic. The 5^{th} Valence electrons finds no place in the thus, it becomes free and travels to the conduction band as shown in figure. Therefore, for each arsenic atom added, one free electron will be available in the silicon crystal. Though each arsenic atom provides one free electrons yet an extremely small amount of arsenic impurity provides enough atoms to supply millions of free electrons.

Due to thermal energy, still hole election pairs are generated but the number of free electrons are very large in number when compared to holes. So in an n-type semiconductor electrons are majority charge carriers and holes are minority charge carriers. Since the current conduction is pre-dominantly by free electrons (-vely charges) it is called as n-type semiconductor (n- means –ve).

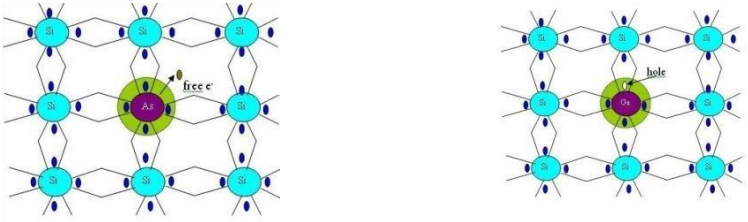


Fig. 2.15 (a) n-type semiconductor (b) p-type semiconductor

p-type semiconductor: When a small amount of trivalent impurity is added to a pure semiconductor it is called p-type semiconductor**.** The addition of trivalent impurity provides large number of holes in the valance band in semiconductor crystals. Example: Gallium, Indium or Boron etc. Such impurities which produce p- type semiconductors are known as acceptor impurities because the holes created can accept the electrons in the semiconductor crystal.

To understand the formation of p-type semiconductor, consider a pure silicon crystal with an impurity say gallium added to it. We know that silicon atom has 4 valence electrons and Gallium has 3 electrons. When Gallium is added as impurity to silicon, the 3 valence electrons of gallium make 3 covalent bonds with 3 valence electrons of silicon. The 4^{th} valence electrons of silicon cannot make a covalent bond with that of Gallium because of short of one electron as shown above. This absence of electron is called a hole. Therefore for each gallium atom added one hole is created, a small amount of Gallium provides millions of holes. Due to thermal energy, still hole-electron pairs are generated but the number of holes is very large compared to the number of electrons. Therefore, in a p-type semiconductor holes are majority carriers and electrons are minority carriers. Since the current conduction is predominantly by hole (+ charges) it p-type semiconductor (p means +ve).

Carrier concentration in an intrinsic semi conductor

The number of charge carriers (electrons and Holes) per unit volume in a material is called charge density or carrier concentration (n_i).

Density of Electrons in conduction band

The number of electrons in the conduction band per unit volume is called electron concentration or electron density (n_e).

Let us consider an energy band structure of an intrinsic semiconductor. Let E_c be the energy corresponding to the bottom edge of the conduction band and E_v the energy corresponding to the top edge of the valance band as shown in fig.

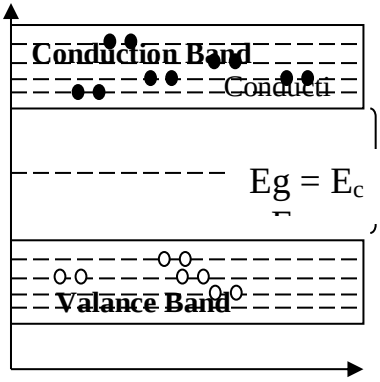
Let m_e^* be the effective mass of the electron

The number of electrons in the entire range within the energy limits of E_c and ∞

$\int dn_e = n_e \int_{E_c}^{\infty} Z(E) d\,E.F(E)$---(1)

Where $Z(E)d(E) - - - - - - \rightarrow$ Density of states

F(E) $- - - - - - \rightarrow$ Fermi function (probability of filling electron in a given state)



But we know quantim mechanically

$Z(E)\,dE = \frac{\pi}{2}\left(\frac{8m_{e^*}}{h^2}\right)^{3/2} (E - Ec)^{1/2}\,dE$ ---------(2)

We know that the probability of the electron distribution in a given energy state ie (Fermi function) according to Fermi Dirac statistics,

$F(E) = \frac{1}{1 + e^{(E-E_f)/K_BT}}$ --------(3)

Since to move an electron from valance band to conduction band the energy required is greater than 4 K_bT

i.e. $E - E_f >> K_BT$; ie, $e^{(E-E_f)/K_BT} >> 1$

i.e. $1 + e^{(E-E_f)/K_BT} \cong e^{(E-E_f)/K_BT}$

i.e. $F(E) = \frac{1}{e^{(E-E_f)/K_BT}}$

$\therefore$ F(E) = $= e^{(E_f - E)/K_B T}$ -------(4)

Substituting (4) and (2) in eqn (1)

$n_e = \int_{E_c}^{\infty} \frac{\pi}{2} \left(\frac{8m_e^*}{h^2}\right)^{3/2} (E - E_c)^{1/2}\, dE \,.\, e^{(E_f - E)/K_B T}$

$n_e = \frac{\pi}{2} \left(\frac{8m_e^*}{h^2}\right)^{3/2} \int_{E_c}^{\infty} (E - E_c)^{1/2} \,.\, e^{(E_f - E)/K_B T} dE$.-----(5)

To evaluate the integrals let us assume that

$$E - E_c = xK_BT \longrightarrow (a)$$

$$E = E_c + xK_BT \longrightarrow (b)$$

$$dE = K_BT.dx \longrightarrow (c)$$

To find the Limits:

When

$E = E_c$	$E = \infty$
Eqn (a) becomes	Eqn (a) becomes
$Ec - E_c = xK_BT$	$\infty - E_c = xK_BT$
ie $0 = xK_BT$	ie $\infty = xK_BT$
ie $x = 0$	$x = \infty$

$\therefore$ The limits are $x = 0$ and $x = \infty$

Applying limits in Eqn (5)

$n_e = \frac{\pi}{2} \left(\frac{8m_e^*}{h^2}\right)^{3/2} \int_0^{\infty} (xK_BT)^{1/2} \,.\, e^{(E_f - (E_c + xK_BT)/K_BT} . K_BT . dx$.

ie $n_e = \frac{\pi}{2} \left(\frac{8m_e^* K_BT}{h^2}\right)^{3/2} e^{(E_f - E_c)/K_BT} \int_0^{\infty} (x)^{1/2} . e^{(-xK_BT)/K_BT} . dx$

$n_e = \frac{\pi}{2} \left(\frac{8m_e^* K_BT}{h^2}\right)^{3/2} e^{(E_f - E_c)/K_BT} \int_0^{\infty} (x)^{1/2} . e^{(-x)} . dx$

We know from gamma function $\int_0^{\infty} (x)^{1/2} . e^{(-x)} . dx = \frac{\sqrt{\pi}}{2}$

$n_e = \frac{\pi}{2} \left(\frac{8m_e^* K_BT}{h^2}\right)^{3/2} e^{(E_f - E_c)/K_BT} . \left(\frac{\sqrt{\pi}}{2}\right)$

$n_e = \frac{1}{4} \left(\frac{8\pi m_e^* K_BT}{h^2}\right)^{3/2} e^{(E_f - E_c)/K_BT}$.

$$n_e = 2 \left(\frac{2\pi m_e^* K_BT}{h^2}\right)^{3/2} e^{(E_f - E_c)/K_BT} \text{----------(6)}$$

Equation (6) represents the carrier concentration of electrons in the conduction band.

Intrinsic Carrier Concentration (n_i) in terms of E_g :

For an intrinsic semi conductor ,

$$n_e = n_h$$

where n_e is the no. of electrons in the conduction band

n_h is the no. of holes in the valance band

Intrinsic carrier concentration is equal to the number of electrons in the conduction band or the number of electrons in the valance band.

i.e. $n_i = n_e = n_h \longrightarrow (1)$

i.e. $n_i^2 = n_e \times n_h \longrightarrow (2)$

substituting the expressions for n_e & n_h in eqn (2)

$$n_i^2 = 2\left(\frac{2\pi m_e * K_B T}{h^2}\right)^{3/2} e^{(E_f - Ec)/K_b T} \times 2\left(\frac{2\pi m_h * K_B T}{h^2}\right)^{3/2} e^{(E_v - E_f)/K_B T}$$

$$n_i^2 = 4\left(\frac{2\pi K_B T}{h^2}\right)^3 (m_e * m_h *)^{3/2} e^{(E_v - E_c)/K_B T}$$

$$n_i^2 = 4\left(\frac{2\pi K_B T}{h^2}\right)^3 (m_e * m_h *)^{3/2} . e^{(-E_G)/K_B T} \longrightarrow (3) \qquad (\therefore E_c - E_v = E_G)$$

Taking square root of eqn (3) on both sides

$$(n_i^2)^{1/2} = \left(4\left(\frac{2\pi K_B T}{h^2}\right)^3 (m_e * m_h *)^{3/2} . e^{(-E_G)/K_B T}\right)^{1/2}$$

$$n_i = 2\left(\frac{2\pi K_B T}{h^2}\right)^{3/2} (m_e * m_h *)^{3/4} . e^{(-E_G)/2K_B T} \longrightarrow (4)$$

Variation of Fermi level with temperature for an Intrinsic semiconductor

For an Intrinsic semiconductor

$n_e = n_h \longrightarrow (1)$ where

$n_e \longrightarrow$ No. of electrons in the conduction band $n_h \longrightarrow$ No. of holes in the valance band

Substituting the values of n_e and n_h in (1)

$$2\left(\frac{2\pi m_e * K_B T}{h^2}\right)^{3/2} e^{(E_f - Ec)/K_B T} = 2\left(\frac{2\pi m_h * K_B T}{h^2}\right)^{3/2} e^{(E_v - E_f)/K_B T}$$

$$(m_e *)^{3/2} . e^{(E_f - Ec)/K_B T} = (m_h *)^{3/2} . e^{(E_v - E_f)/K_B T}$$

$$\left(\frac{m_h *}{m_e *}\right)^{3/2} = \frac{e^{(E_f - E_c)/K_B T}}{e^{(E_v - E_f/K_B T}}$$

$$\left(\frac{m_h *}{m_e *}\right)^{3/2} = e^{(E_f - E_c - E_v + E_f)/K_B T}$$

$\left(\frac{m_{h^*}}{m_{e^*}}\right)^{3/2} = e^{(2E_f - (E_c + E_v))/K_B T} \longrightarrow (2)$

Taking log on both the sides

$\frac{3}{2} \log_e \left(\frac{m_{h^*}}{m_{e^*}}\right) = \frac{(2E_f - (E_c + E_v))}{K_B T}$

i.e. $2E_f = (E_c + E_v) + \frac{3}{2} K_B T \log_e \left(\frac{m_{h^*}}{m_{e^*}}\right)$

i.e. $E_f = (\frac{E_c + E_v}{2}) + \frac{3}{4} K_B T \log_e \left(\frac{m_{h^*}}{m_{e^*}}\right) \longrightarrow (3)$

The above equation represents the Fermi level in an intrinsic semiconductor

If $m_h * = m_e *$; then

$\log_e \left(\frac{m_{h^*}}{m_{e^*}}\right) = 0$

$\therefore E_f = (\frac{E_c + E_v}{2}) \longrightarrow (4)$

Eqn (4) indicates that Fermi level lies exactly in the midway between the valance band and the conduction band as shown in fig (1.3). its position is independent of temperature.

But in realitly $m_h^* > m_e^*$ and so the Fermi energy level slightly increases with the rise in temperature as in fig (b).

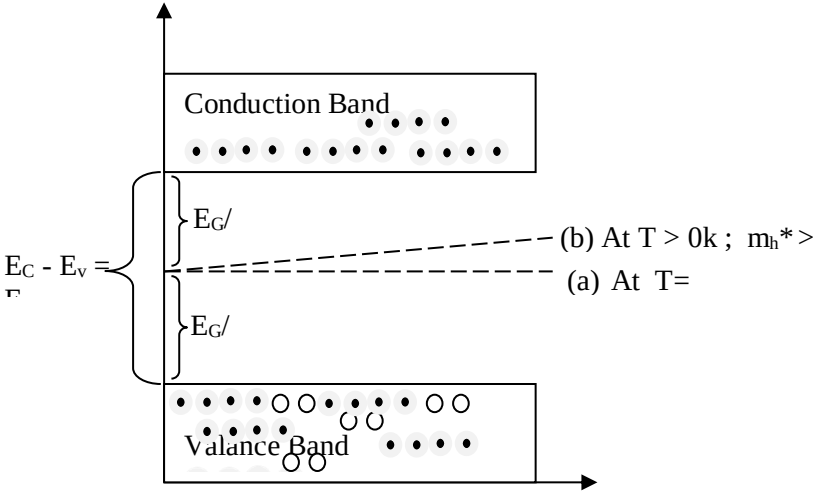


Fig : Variation of Fermi level in Intrinsic Semiconductor

Therefore, the concentration of electron in the conduction band and hole in the valence band depend upon the temperature of the intrinsic semiconductor.

2.9. EXPERIMENTAL DETERMINATION OF BAND GAP OF A SEMICONDUCTOR

The diode is immersed in an oil bath that in turn is kept in a heating mantle. A thermometer is also kept in the oil path such that its measure bulb is just at the height of diode. The diode is reversed biased with the help of DC voltage obtained from a DC power supply and the current that flows through the reverse diode is measured with a milliammeter. The Circuit diagram of this experiment is shown in the Fig.2.18(a).

The power supply is switched on and the voltage is adjusting say five volts. The current through the diode and the room temperature are noted. The power supplyin switched off and the heating mantle is switched on. The oil bath is heated up to 65^0 C. The oil is stirred well. The temperature of oil bath stabilized say at 75^0C.

The power supply is again switched on and the voltage is kept at 5V. The temperature (say 75^0C) and the corresponding current through the diode are noted.

Now the oil bath is allowed to cool slowly. As the temperature falls, the current through the diode decreases. The current falls for various temperatures (any convenient interval) are noted down.

The calculations are completed and a graph is plotted taking 1000/T on X-axis and log I_0 on Y-axis. A straight line is obtained as shown in the Fig.2.18 (b). The

slow of the straight line is determined and using it the band gap is calculated usingthe following equation.

$$E_g = 3.973 \times 10^{-4} \times \text{slope eV}$$

Care should be taken while performing the experiment the diode and thethermometer are placed at the same level in the oil bath. The maximum temperature of the diode should not be beyond 80^0C. Reading of the current and temperature to betaken simultaneously.

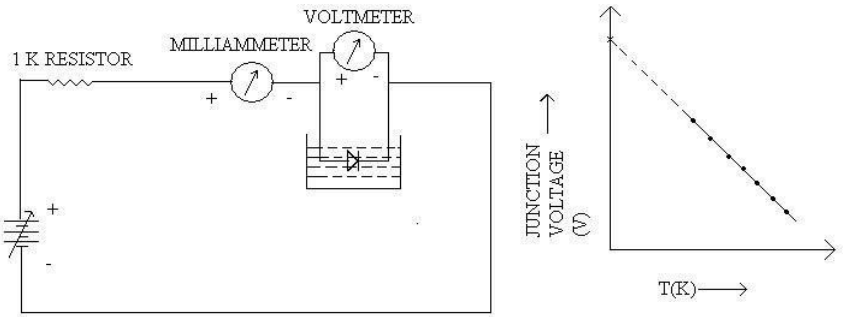


Fig. 2.18 (a) Circuit diagram **(b) Experimental graph**

2.10 Application of Semiconductors in fabrication of basic components

The IC fabrication starts with the wafer preparation, oxide layer deposition or a protective layer, introducing impurities into the wafer (n-type or p-type), implantation of ions to the surface of the crystal, CVD to produce thin films, photolithography (creating a mask pattern with the help of UV light), deposition of the metal for interconnections, and the packaging

The basic fabrication processes of the Integrated Circuits are as follows:

- Wafer Preparation
- Oxidation
- Diffusion
- Ion Implantation
- Chemical-Vapor Deposition
- Photolithography
- Metallization
- Packaging

2.10.1 Wafer Preparation

A wafer is a thin material used for making various **Integrated circuits** and **transistors.** Wafer acts as a base for such devices. The material of a wafer is the semiconductor, especially **crystalline silicon.** The silicon crystals used for the wafer manufacturing are highly pure. The process of extracting pure metal from the melt is known as a **boule.** The impurities are further added to the molten state of the material in a specific amount to make it n-type or p-type.

The wafer preparation is the first step for IC fabrication. It involves **cutting, shaping,** and **polishing** the wafer material to make it suitable for further fabrication. Some wafers are modified because of their sharp edges, irregular surface, and shape to convert

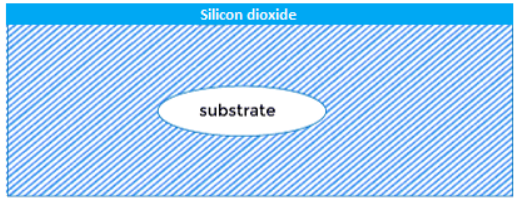
them to the required wafer.

2.10.2 Oxidation

The oxidation process is one of the essential parts of the IC fabrication process and other VLSI (Very Large Scale Integration) devices.

Oxidation is the process of adding **oxygen.** In a semiconductor, the oxygen and the silicon react to form silicon dioxide. The oxidation is carried out in furnaces at high temperatures upto 1250 degrees Celsius. Oxidation is classified as **wet** oxidation or **dry** oxidation. Both processes are widely used and have their own advantages and disadvantages. Wet oxidation is fast, while dry oxidation has good electrical properties.

Wet oxidation is also known as **steam.** Both types of oxidation have excellent electrical insulation properties. The deposition of silicon dioxide on the silicon wafer protects from many impurities. The dopants can be applied only to areas not covered with the SiO2. An example of oxide layer on the substrate is shown below:



2.10.3 Chemical-Vapor Deposition

CVD or Chemical Vapor deposition is used to produce **high-quality solid materials.** It produces thin films in the semiconductor industry. The process is carried out at a pressure below the atmospheric pressure, also known as **vacuum deposition.** The chemicals and vapours react to form solids on the surface of a substrate. The protective layers on the substrate, such as **silicon dioxide, polysilicon,** and **silicon nitride,** are deposited using the CVD process.

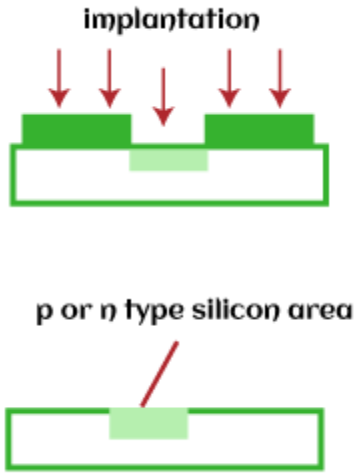
For example,

The reaction of silane gas, a molecule of silicon atom, with the oxygen on the surface of the silicon wafer results in the formation of SiO2 or silicon dioxide. The CVD process is not good enough compared to the thermally grown oxide but is enough to act as an insulator

on the substrate. Its main advantage is the faster deposition rate at a lower temperature.
CVD can be performed at **very low atmospheric** pressures, **sub-atomic** pressures, and **ultra high presures.** It depends on the chemical reactions involved in the process. Tungsten CVD is also used to create contacts or plugs on a semiconductor device.

2.10.4 Ion Implantation

The ion implantation is a process of accelerating ions from the element to the solid target. The accelerated ions on the solid can change its properties. The ions are applied at a low temperature and high energy. The low temperature allows it to work at room temperature. But, the excess energy can also damage the crystal structure of the solid.



Ion implantation introduces impurity atoms in the crystal (semiconductor material). The accelerated atom gets embedded when it strikes the surface of the crystal. The energy and accelerating field voltage determine the depth of the penetration of the ions. The quantity of dopants can be controlled by adjusting the amount of ion concentration. Hence, ion implantation is more accurate due to the controlled current, voltage, and energy.
The process is generally used when an accurate amount of impurity atoms is necessary for the device's operation.
Diffusion

Diffusion is a process of adding impurities atoms from a region with high concentration to a region of low concentration. The dopants or impurity atoms are added to the silicon (semiconductor material), which changes its resistivity. The process of diffusion is highly dependent on the temperature. It is carried out in high-temperature furnaces between 1000 degrees Celsius and 1200 degrees Celsius. The depth and width of the impurities depend on the temperature range and the timings. The high doping concentration improves the conductivity of a metal.

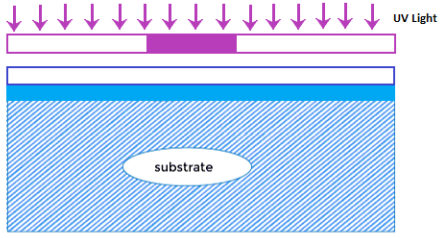
The dopants can be of any state, solid, liquid, or gas. The preferred dopants are pentavalent impurities or n-type, such as antimony, phosphorous pentoxide, and arsine (gas). The trivalent impurities or p-type are **gallium, indium, boron,** etc.

2.10.5 Photolithography

Photolithography is also known as **optical lithography,** using light to produce thin films. It is used to pattern thin films on the substrate, such as silicon wafers. It protects certain areas during successive fabrication processes (deposition, ion implantation, etc.). The types of light used to introduce mask patterns on the silicon wafer include **UV light, X-rays,** and **extreme UV** that emits at different frequencies. The most used type of light to produce the film is Ultra-Violet light.

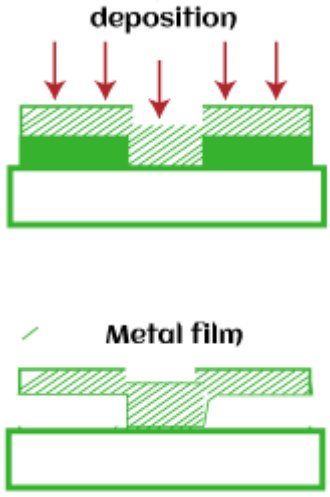
The silicon wafer or the substrate is first coated with a photoresist. On the photoresist layer, a mask pattern with the help of photolithography is applied to the silicon wafer. The exposed areas of the wafer become soft and can be removed. It creates a pattern on the wafer. The wavelength of each type of light determines the feature size impressed on the photoresist.

Photolithography is somewhat similar to photography because the exposed area is patterned with help of the light. The incident light can be direct or projected through the lens.

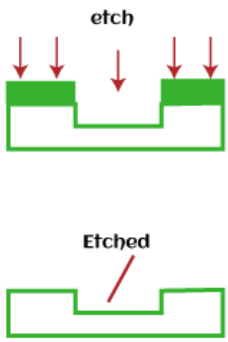


2.10.6 Metallization

Metallization is defined as the process of coating a metal layer on the metallic surface or non-metallic surface. The coating can be of **aluminum, zinc,** or **silver.** The metal coating in CMOS fabrication is aluminum, which protects the surface from external environmental factors, as dust, air, water, etc.



Metallization is also used to interconnect various components that form an Integrated circuit. The components can be resistors, capacitors, transistors, relays, etc. The metal layer is first deposited on the surface of the silicon wafer, as discussed above. After that, the required pattern or area for interconnected is etched, as shown below:



The metallization process uses chamber to apply metal layer. The wafer is placed inside the chamber, which coats the entire surface inside it. The thickness of the metal layer can vary depending on the requirements

2.10.7 Packaging

Packaging is the last stage of the IC fabrication process. The finished silicon wafer or chips of various sizes with number of small components are tested electrically to recheck their working. The testing is performed using an automatic probing station. It is a cost effective testing machine that includes microwave and radio-frequency testing. The default circuits are separated from all the circuits. The good circuits are sent for packaging or headers.

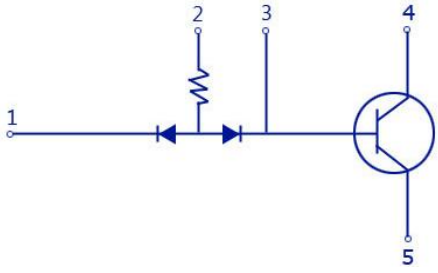
Packaging also assembles the ICs with other devices. IC is a component of a device. The external connections with the IC are packed together to form a device. We can also say that packaging is a connection between the manufacturing and end use of a device. It makes a product suitable for its end use.

The package is finally sealed with **epoxy** or **plastic** to prevent it from atmospheric dust. It is performed under vacuum or an inert atmosphere, which prevents the material from oxygen that may get trap in the packaging.

Monolithic IC's Fabrication

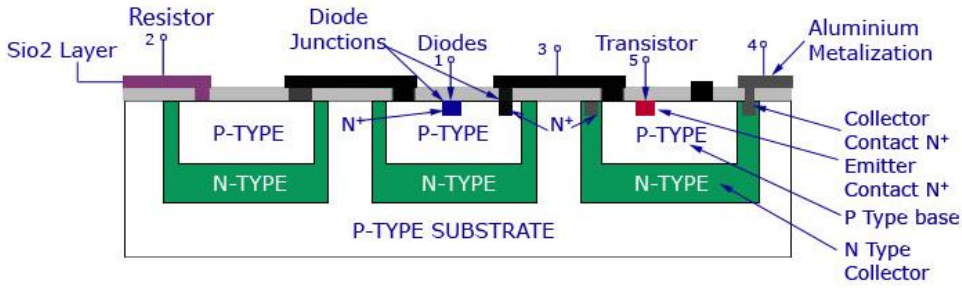
We have already discussed the basics of **Integrated Circuits** in our previous post. The concepts of a basic monolithic IC will be discussed here.

To know the basics a sample circuit must be considered to be converted to its monolithic form. With basic components like resistor, diode, and transistor a basic circuit is first made.



Basic Monolithic IC Circuit

With the basic circuit, the different layers for the monolithic IC are then considered. The basic structure of a monolithic IC will have 4 layers of different materials. The base layer will be a P-type silicon layer and is named as the substrate layer. This layer will have a typical thickness of 200 micrometer. Silicon is the preferred semiconductor for the P-type and N-type layer because of its favourable characteristics for the manufacturing of an IC.
The layer above the substrate P-type silicon layer is the N-type layer. All the active and passive components required for the circuit are fabricated onto this layer. This layer has a typical thickness of 25 micrometer. The N-type silicon material is grown as a single crystal extension of the P-layer and the components are required are fabricated using series of P-type and N-type impurity diffusions. The N-type layer becomes the collector for the transistor or an element for a diode or a capacitor.



Basic Monolithic IC

The layer above N-type is made of silicon dioxide (SiO2) material. Since there is a selective P-type and N-type impurity diffusion going on in the second layer, this layer acts as a barrier in the process. This layer is etched away from the region where diffusion is desired to be permitted with photolithographic process. The rest of the wafer remains protected against diffusion. This layer also protects the silicon layer from contamination.
The up-most layer is that made of aluminium. This metallic layer is used to provide interconnections between the different components used in the IC.

Monolithic IC Manufacturing Process

For the manufacture and production of the monolithic IC, all circuit components and their interconnections are to be formed in a single thin wafer. The different processes carried

out for achieving this are explained below.

1. P-layer Substrate Manufacture

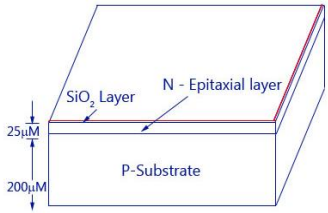
Being the base layer of the IC, the P-type is silicon is first built for the IC. A silicon crystal of P-type is grown in dimensions of 250mm length and 25mm diameter. The silicon is then cut into thin slices with high precision using a diamond saw. Each wafer will precisely have a thickness of 200 micrometer and a diameter of 25 mm. These thin slices are termed wafers. These wafers may be circular or rectangular in shape with respect to the shape of the IC. After cutting hundreds of them each wafer is polished and cleaned to form a P-type substrate layer.

2. N-type Epitaxial Growth

The epitaxial groth process of a low resistive N-type over a high resistive P-type is to be carried out. This is done by placing the n-type layer on top of the P-type and heating then inside a diffusion furnace at very high temperature (nearly 1200C). After heating, a gas mixture f Silicon atoms and pentavalent atoms are also passed over the layer. This forms the epitaxial layer on the substrate. All the components required for the circuit are built on top of this layer. The layer is then cooled down, polished and cleaned.

3. The Silicon Dioxide Insulation Layer

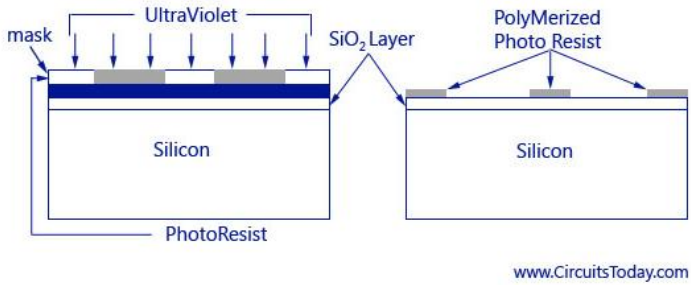
As explained above, this layer is required contamination of the N-layer epitaxy. This layer is only 1 micrometer thin and is grown by exposing the epitaxial layer to oxygen atmosphere at 1000C. A detailed image showing the P-type, N-type epitaxial layer and SiO2 layer is given below.



Monolithic IC-Substrates and Layers

4. Photolithographic Process for SiO2

To diffuse the impurities with the N-type epitaxial region, the silicon dioxide layer has to be etched in selected areas. Thus openings must be brought at these areas through photolithographic process. In this process, the SiO2 layer is coated with a thin layer of a photosensitive material called photoresist. A large black and white pattern is made in the desired patter, where the black pattern represents the area of opening and white represents the area that is left idle. This pattern is reduced in size and fit to the layer, above the photoresist. The whole layer is then exposed to ultraviolet light. Due to the exposure, the photoresist right below the white pattern becomes polymerized. The pattern is then removed and the wafer is developed using a chemical like trichloroethylene. The chemical dissolves the unpolymerized portion of the photoresist film and leaves the surface. The oxide not covered by polymerised photoresist is then removed by immersing the chip in an etching solution of HCl. Those portions of the Si02 which are protected by the photoresist remain unaffected by the acid. After the etching and diffusion process, with the help of chemical solvents like sulphuric acid, the resist mask is then removed by mechanical abrasion. The appropriate impurities are then diffused through oxide free windows.

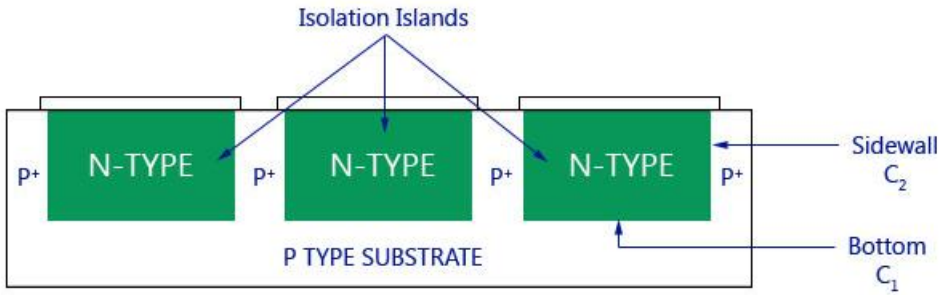


Monolithic IC - Photolithographic-Process

5. Isolation Diffusion

After photolithographic process the remaining SiO2 layer serves as a mask for the diffusion of acceptor impurities. To get a proper time period for allowing a P-type impurity to penetrate into the N-type epitaxial layer, isolation diffusion is to be carried out. By this process, the P-type impurity will travel through the openings in SiO2 layer, and the N-type layer and thus reach the P-type substrate, Isolation junctions are used to isolate between

various components of the IC. The temperature and time period of isolation diffusion should be carefully monitored and controlled. As a result of isolation diffusion, the formation of N-type region called Isolation Island occurs. Each isolated island is then chosen to grow each electrical component. From the figure below you can see that the isolation islands look like back-to-back P-N junctions. The main use if this is to allow electrical isolation between the different components inside the IC. Each electrical element is later on formed in a separate isolation island. The bottom of the N-type isolation island ultimately forms the collector of an N-P-N transistor. The P-type substrate is always kept negative with respect to the isolation islands and provided with reverse bias at P-N junctions. The isolation will disappear if the P-N junctions are forward biased.



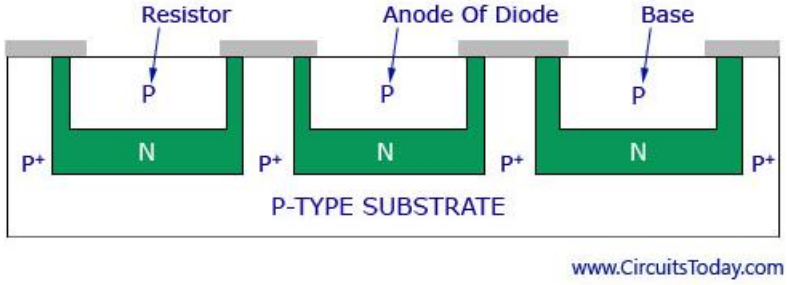
Monolithic IC - Isolation Diffusion

An effect of capacitance is produced in the region where the two adjoining isolation islands are connected to the P-type substrate. This is basically a parasitic capacitance that will affect the performance of the IC. This kind of capacitance is divided into two. As shown in the figure C1 is one kind of capacitance that forms from the bottom of the N-type region to the substrate and capacitance C2 from the sidewalls of the isolation islands to the P-region. The bottom component C1 is essentially due to step junction formed by epitaxial growth and, therefore, varies as the square root of the voltage V between the isolation region and substrate. The sidewall capacitance C2 is associated with a diffused graded junction and so varies as (-1/2) exponential of V. The total capacitance is of the order of a few picoFarads.

6. Base Diffusion

The working of base diffusion process is shown in the figure below. This process is done to create a new layer of SiO2 over the wafer. P-regions are formed under regulated

environments by diffusing P-type impurities like boron. This forms the base region of an N-P-N transistor or as well as resistors, the anode of diode, and junction capacitor. In this case, the diffusion time is so controlled that the P-type impurities do not reach the substrate. The resistivity of the base layer is usually much higher than that of the isolation regions.

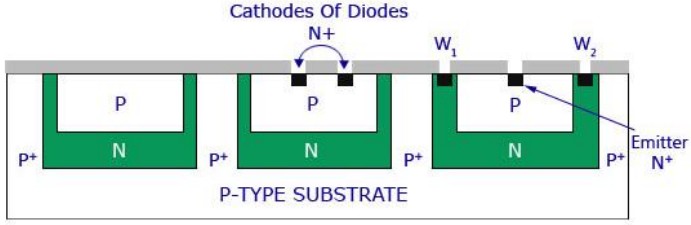


Monolithic IC - Base Diffusion

The isolation regions will have a lot lesser resistivity than that of the base layer.

7. Emitter Diffusion

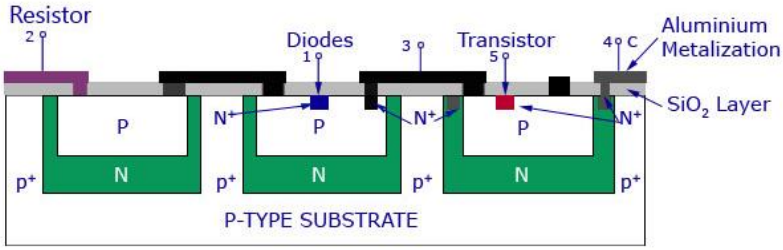
Masking and etching process is again carried out to form a layer of silicon dioxide over the entire surface and opening of the P-type region. The transistor emitters, the cathode regions for diodes, and junction capacitors are grown by diffusion using N-type impurities like phosphorus through the windows created through the process under controlled environmental process. As shown in the figure below there are two additional windows: W1 and W2. These windows are made in the N-region to carry an aluminium metallization process.



Emitter Diffusion

8. Aluminium Metallization

The windows made in the N-region after creating a silicon dioxide layer are then deposited with aluminium on the top surface. The same photoresist technique that was used in photolithographic process is also used here to etch away the unwanted aluminium areas. The structure then provides the connected strips to which the leads are attached. The process can be better understood by going through the figure below.



Aluminium Metalization

9. Scribing and Mounting

This is the final stage of the IC manufacturing process. After the metallization process, the silicon wafer is then scribed with a diamond tipped tool and separated into individual chips. Each chip is then mounted on a ceramic wafer and is attached to a suitable header. Next the package leads are connected to the IC chip by bonding of aluminium or gold wire from the terminal pad on the IC chip to the package lead. Thus the manufacturing process is complete. Thu, hundreds of IC's is manufactured simultaneously on a single silicon wafer.

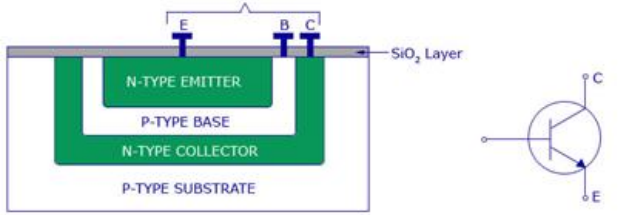
Monolithic IC – Component Fabrication

Now we shall discuss in detail how different circuit elements like capacitors, transistors, diodes, and resistors are fabricated into an IC. Please note that it is practically impossible to fabricate an inductor into an IC. It is thus added externally by connecting it to the corresponding IC pin as designed by the manufacturer.

Transistors

The fabrication process of a transistor is shown in the figure below. A P-type substrate is first grown and then the collector, emitter, and base regions are diffused on top of it as

shown in the figure. The surface terminals for these regions are also provided for connection.



Monoilithic IC - Transistor Fabrication

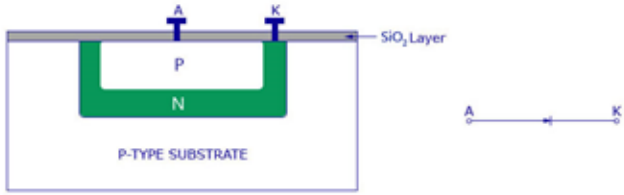
Both transistors and diodes are fabricated by using the epitaxial planar diffusion process that is explained earlier. In case of discrete transistors, the P-type substrate is considered as the collector. `But this is not possible in monolithic IC's, as all the transistors connected on one P-type substrate would have their collectors connected together. This is why separate collector regions are diffused into the substrate.

Even though separate collector regions are formed, they are not completely isolated from the substrate. For proper functioning of the circuit it is necessary that the P-type substrate is always kept negative with respect to the transistor collector. This is achieved by connecting the substrate to the most negative terminal of the circuit supply. The unwanted or parasitic junctions, even when reverse-biased, can still affect the circuit performance adversely. The junction reverse leakage current can cause a serious problem in circuits operating at very low current levels. The capacitance of the reverse-biased junction may affect the circuit high-frequency performance, and the junction break down voltage imposes limits on the usable level of supply voltage. All these adverse effects can be reduced to the minimum if highly resistive material is employed for the substrate. If the substrate is very lightly doped, it will behave almost as an insulator.

Diodes

They are also fabricated by the same diffusion process as transistors are. The only difference is that only two of the regions are used to form one P-N junction. In figure, collector-base junction of the transistor is used as a diode. Anode of the diode is formed during the base diffusion of the transistor and the collector region of the transistor becomes

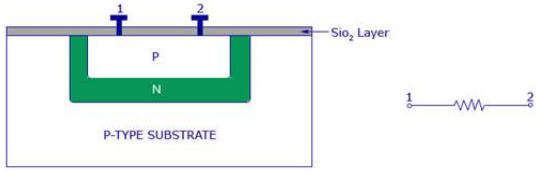
the cathode of the diode. For high speed switching emitter base junction is used as a diode.



Monolithic IC - Diode Fabrication

Resistors

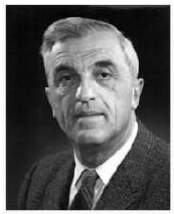
The resistors used in IC's are given their respective ohmic value by varying the concentration of doping impurity and depth of diffusion. The range of resistor values that may be produced by the diffusion process varies from ohms to hundreds of kilohms. The typical tolerance, however, may be no better than ± 5%, and may even be as high as ± 20%. On the other hand, if all the resistors are diffused at the same time, then the tolerance ratio may be good. Most resistors are formed during the base diffusion of the integrated transistor, as shown in figure below. This is because it is the highest resistivity region. For low resistance values, emitter region is used as it has much lower resistivity.

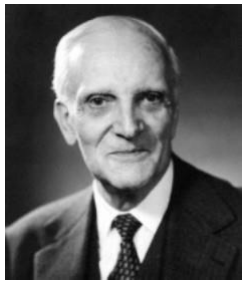


Monolithic IC - Resistor Fabrication

Another diffusion technique is also used for the growth of IC resistors. It is basically a thin-film technique. In this process a metal film is deposited on a glass or Si02 surface. The resistance value can be controlled by varying thickness, width and length of the film. Since diffused resistors can be processed while diffusing transistors. This technique is more economic and less time consuming and therefore, the most widely used.

Just to know

 **Dr.FELIX BLOCH** (1905-1983) Nobel Prize for Physics (1952)	 **HEIKE KAMMERLINGH ONNES**
His doctoral thesis established the quantum theory of solids, using Bloch waves todescribe electrons in periodic lattices. His additional excellent contributions are nuclearinduction and nuclear magnetic resonance, which are the underlying principles of MRI andMagnon.	Dutch scientist **Heike Kammerlingh Onnes** was awarded the 1913 Nobel Prize in Physics for his work on the properties of matter at low temperatures. He discovered superconductivity



LÉON BRILLOUIN
(1889-1969)

Concept of Brillouin Zone

The concept of a Brillouin zone was first developed by Léon Brillouin (1889-1969), a French physicist. During his work on the propagation of electron waves in a crystal lattice, he introduced the concept of Brillouin zone in 1930. Quantum mechanical perturbations techniquesby Brillouin and by Eugene Wigner resulted in what is known as the Brillouin-Wigner formula. He also made contributions to quantum mechanics, radio wave propagation in the atmosphere, solid state physics, and information theory. He also studied the propagation of monochromatic light waves and their interaction with acoustic waves, i.e., scattering of light with a frequency change, which became known as Brillouin scattering.

A Brillouin zone is defined as a Wigner-Seitz primitive cell in the reciprocal lattice. The first Brillouin zone is the smallest volume entirely enclosed by planes that are the perpendicular bisectors of the reciprocal lattice vectors drawn from the origin. The concept of Brillouin zone is particularly important in the consideration of the electronic structure of solids. There are also second, third, *etc*., Brillouin zones, corresponding to a sequence of disjoint regions (all with the same volume) at increasing distances from the origin, but these are used more rarely. As a result, the first Brillouin zone is often called simply the Brillouin zone.

Solved Examples

Problems:

1. Find the relaxation time of conduction electrons in a metal of resistivity $1.54 \times 10^{-8}\ \Omega$-m, if the metal has 5.8×10^{28} conductionelectrons/m3.

   Sol: Given data are:

   Resistivity of the metal, $\rho = 1.54 \times 10^{-8}\ \Omega$–m

   Number of conduction electrons, $n = 5.8 \times 10^{28}/m^3$

   Relaxation time, τ = ?

$$\sigma = \frac{ne^2\tau}{m} \text{ (or) } \tau = \frac{\sigma m}{ne^2} = \frac{m}{ne^2\rho}$$

$$= \frac{9.11 \times 10^{-31}}{5.8\times10^{28} \times [1.602 \times 10^{-19}]^2 \times 1.54 \times 10^{-8}}$$

$$= = \frac{9.11\times10^{-31}}{5.8\times(1.602)^2 \times 1.54 \times 10^{-18}} = \frac{9.11\times10^{-13}}{22.92}$$

$$= \frac{911\times10^{-15}}{22.92} = 39.747 \times 10^{-15}\ s$$

2. The electrical resistivity of copper at 20°C is $2x10^{-8}$ ohm-metre. Calculate itsthermal conductivity

Sol: Given data are

Resistivity of the metal,

$\rho = 2 \times 10^{-8}\ \Omega$–m

Temperature = 20°C = 20

+ 273 = 293 K Thermal

conductivity K ?

According to the Wiedemann Franz law

$K/\sigma = LT$ where $\sigma = 1/\rho$ and L=

$2.44\ 10^{-8}\ W\Omega/K^2$Hence $K= LT/\rho =$

$(2.44\ 10^{-8} * 293) / 2 \times 10^{-8}$

= 357.46 W/m/K

Part-A: Questions

1. Differentiate crystalline and non-crystalline solids
2. How the bands are formed in a solid
3. What is allowed band and forbidden band
4. Draw the band formation diagram
5. Differentiate metal, half metal and semi metal
6. Differentiate conductor and insulator
7. Explain types of semiconductor based on energy band theory
8. How Fermi level play a role in p-type and n-type semiconductor
9. Explain Fermi surface
10. Explain Fermi energy
11. Explain Fermi level
12. Explain density of states with respect to Fermi level

13. Explain the effect of external field on Fermi surface
14. Write some merits of free electron theory
15. Write few drawbacks of classical free electron theory
16. How quantum is better than classical free electron theory
17. Write about zone theory
18. State Wiedemann Franz law
19. Define drift velocity
20. Define collision time
21. Define mean free path
22. Define relaxation time
23. Compare classical and quantum free electron theory
24. Differentiate conductor, semiconductor and insulator.
25. Explain briefly about the types of semiconductor.
26. What do you meant by intrinsic semiconductor?
27. What do you meant by extrinsic semiconductor?
28. What do you meant by n-type semiconductor?
29. What do you meant by p-type semiconductor?
30. What is doping. Mention the use of doping in n-type and p-typesemiconductors.
31. Explain how the acceptor energy level is created
32. Explain how the donor energy level is created
33. What is Fermi level?
34. Define superconductivity.
35. What are superconductors?
36. Mention the advantages of integrated circuits.
37. Write down the various processes used to fabricate IC's using silicon planar technology.
38. What is the purpose of oxidation?

39. What do you mean by monolithic process?
40. List the advantages of integrated circuits over discrete component circuit
41. What is meant by epitaxial growth?
42. What is ion implantation?

Part – B: Questions

1. Explain the band structure of conductor, semiconductor, insulator, semimetalsand Half metals
2. State and derive Wiedemann Franz law
3. Explain band theory. Differentiate conductor, semiconductor and insulator based on band theory. Also differentiate intrinsic and extrinsic semiconductors based on band theory.
4. Define intrinsic semiconductor. Derive an equation of electron and hole concentration of an intrinsic semiconductor.
5. Define Band gap. Explain the determination of band gap of a semiconductor.
6. Explain the various steps involved in the process of fabricating IC
7. Explain the process of fabricating basic components using monolithic IC