



**MODULE-4**

**Semiconductor Physics and Devices:**

MODULE-4	<b>Semiconductor Physics and Devices:</b>	
	L1	Formation of energy bands of solids
	L2	Definition and significance of Fermi-Dirac distribution function
	L3	Description of Fermi level and derivation of variation of its dependence on temperature and carrier concentration
	L4	Definition of Density of States, Derivation of carrier concentration in intrinsic and extrinsic semiconductors (procedure & expressions)
	L5	Expression of Carrier generation – recombination, mobility, drift-diffusion current
	L6	Demonstration of Hall effect, Derivation for Hall Voltage & Applications
	L7	Construction & Working of P-N Junctions and explanation of I-V characteristics
	L8	Construction & Working of Bipolar Junction Transistors and explanation of I-V characteristics

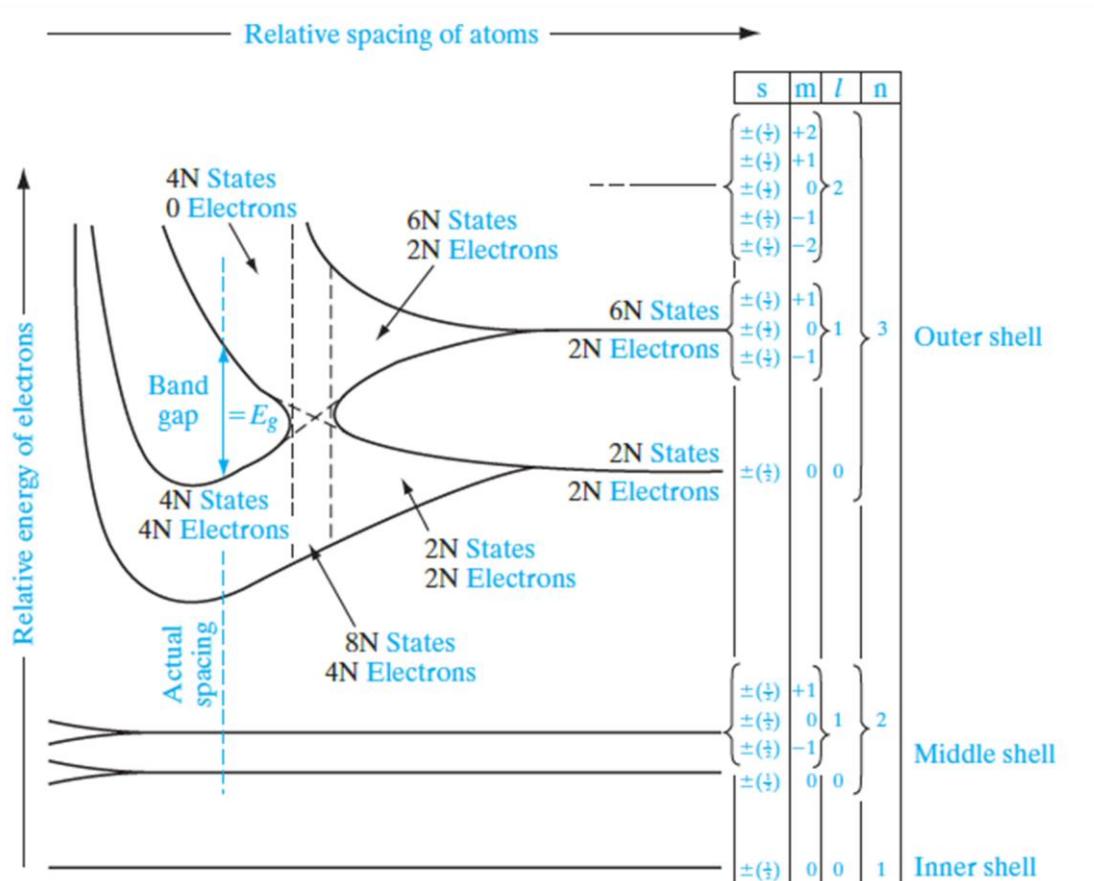


## Lecture 1

### Band Theory of solids:

According to the Bohr atomic model, in an isolated atom the energy of any of its electrons is decided by the orbit in which it revolves. But when the atoms come together to form a solid they are close to each other. So the outer orbits of electrons from neighbouring atoms would come very close or could even overlap. This would make the nature of electron motion in a solid very different from that in an isolated atom.

Consider that the Si or Ge crystal contains N atoms. Electrons of each atom will have discrete energies in different orbits. The electron energy will be same if all the atoms are isolated, i.e., separated from each other by a large distance. However, in a crystal, the atoms are close to each other ( $2 \text{ to } 3\text{\AA}$ ) and therefore the electrons interact with each other and also with the neighbouring atomic cores. The overlap (or interaction) will be more felt by the electrons in the outer most orbit while the inner orbit or core electron energies may remain unaffected.



**Figure 1:** Energy levels in Si as a function of inter atomic spacing. The core levels ( $n = 1, 2$ ) in Si are completely filled with electrons. At the actual atomic spacing of the crystal, the  $2N$  electrons in the  $3s$  sub shell and the  $2N$  electrons in the  $3p$  sub shell undergo  $sp^3$  hybridization, and all end up in the lower  $4N$  states (valence band), while the higher-lying  $4N$  states (conduction band) are empty, separated by a band gap.



For Si, the outer most orbit is the third orbit ( $n=3$ ). The number of electrons in the outer most orbit is 4 (2s and 2p electrons). Hence, the total number of outer electrons in the crystal is  $4N$ . The maximum possible number of outer electrons in the orbit is 8 (2s + 6p electrons). So, out of the  $4N$  electrons,  $2N$  electrons are in the  $2N$  s-states and  $2N$  electrons are in the available  $6N$  p-states. Obviously, some p-electron states are empty as shown in the extreme right of Figure. This is the case of well separated or isolated atoms.

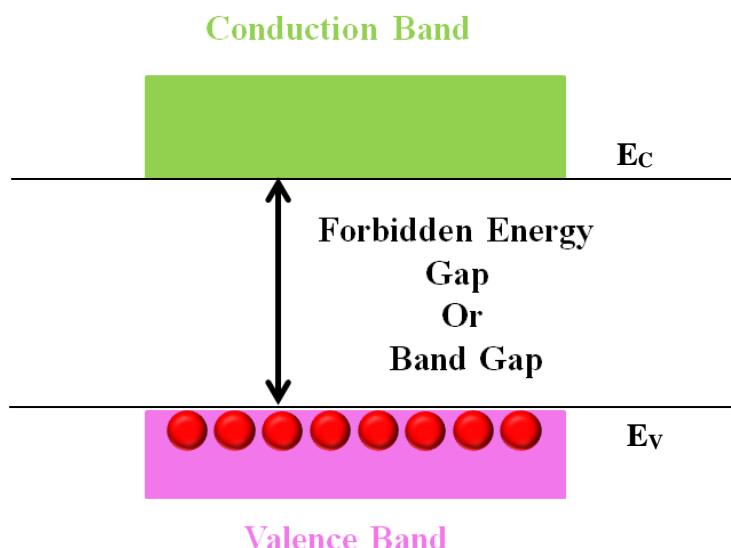
Suppose these atoms start coming nearer to each other to form a solid. The energies of these electrons in the outer most orbits may change (both increase and decrease) due to the interaction between the electrons of different atoms. The  $6N$  states, which originally had identical energies in the isolated atoms, spread out and form energy band. Similarly, the  $2N$  states having identical energies in the isolated atoms, split into a second band separated from the first one by an energy gap.

At still smaller spacing, however, there comes a region in which the bands merge with each other. The lowest energy state that is a split from the upper atomic level appears to drop below the upper state that has come from the lower atomic level. In this region, no energy gap exists where the upper and lower energy states get mixed. Finally, if the distance between the atoms further decreases, the energy bands again split apart and are separated by an energy gap  $E_g$ . The total number of available energy states  $8N$  has been re-apportioned between the two bands (4N states each in the lower and upper energy bands).

- The lower energy band in which all the available valence electrons from the atoms (4N) are present is known as valance band. The valance band is completely filled

- The upper band is completely empty and is called the conduction band.

The lowest energy level in the conduction band is called as  $E_C$  and highest energy level in the valence band is called as  $E_V$ . Above  $E_C$  and below  $E_V$  there are a large number of closely spaced energy levels, as shown in Fig. The gap between the top of the valence band and bottom of the conduction band is called the energy band gap  $E_g$ .



**Figure 2:** The energy band gap or Forbidden energy gap of materials.

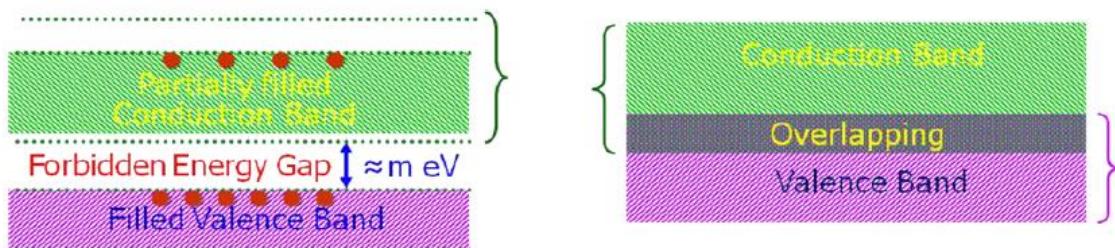


## Classification of Solids:

In the basis of band theory, solids can be broadly classified into three categories, viz,

### Metals:

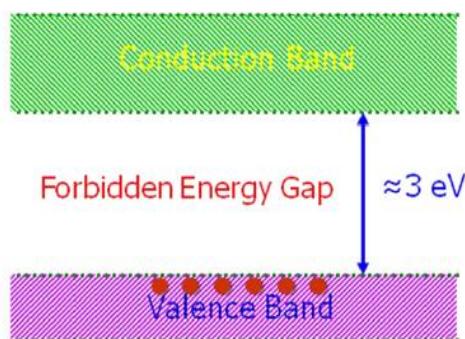
In case of metal either when the conduction band is partially filled and the balanced band is partially empty or when the conduction and valance bands overlap. When there is overlap electrons from valence band can easily move into the conduction band. This situation makes a large number of electrons available for electrical conduction. When the valence band is partially empty, electrons from its lower level can move to higher level making conduction possible. Therefore, the resistance of such materials is low or the conductivity is high.



**Figure 3:** The energy band gap or Forbidden energy gap of metals.

### Insulators:

In case of insulators, a large band gap  $E_g$  exists ( $E_g > 3\text{eV}$ ). There are no electrons in the conduction band, and therefore no electrical conduction is possible. Note that the energy gap is so large that electrons cannot be excited from the valence band to the conduction band by thermal excitation. This is the case of insulators.



**Figure 4:** The energy band gap or Forbidden energy gap of insulators.



## Semiconductor:

In case of semiconductor, a finite but small band gap ( $E_g < 3\text{ eV}$ ) exists. Because of the small band gap, at room temperature some electrons from valence band can acquire enough energy to cross the energy gap and enter the conduction band. These electrons can move in the conduction band. Hence, the resistance of semiconductors is not as high as that of the insulators.

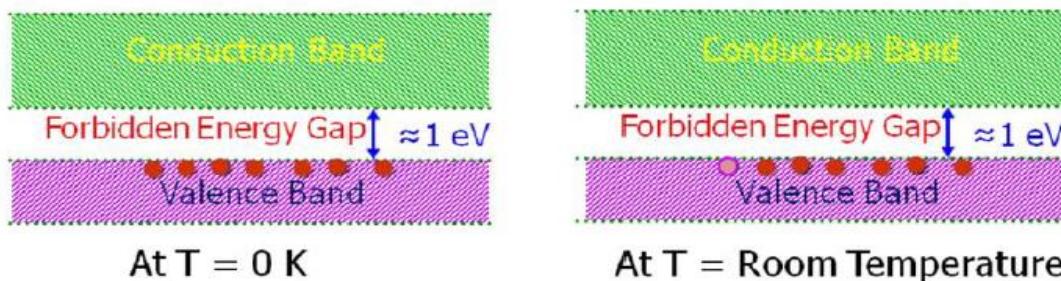
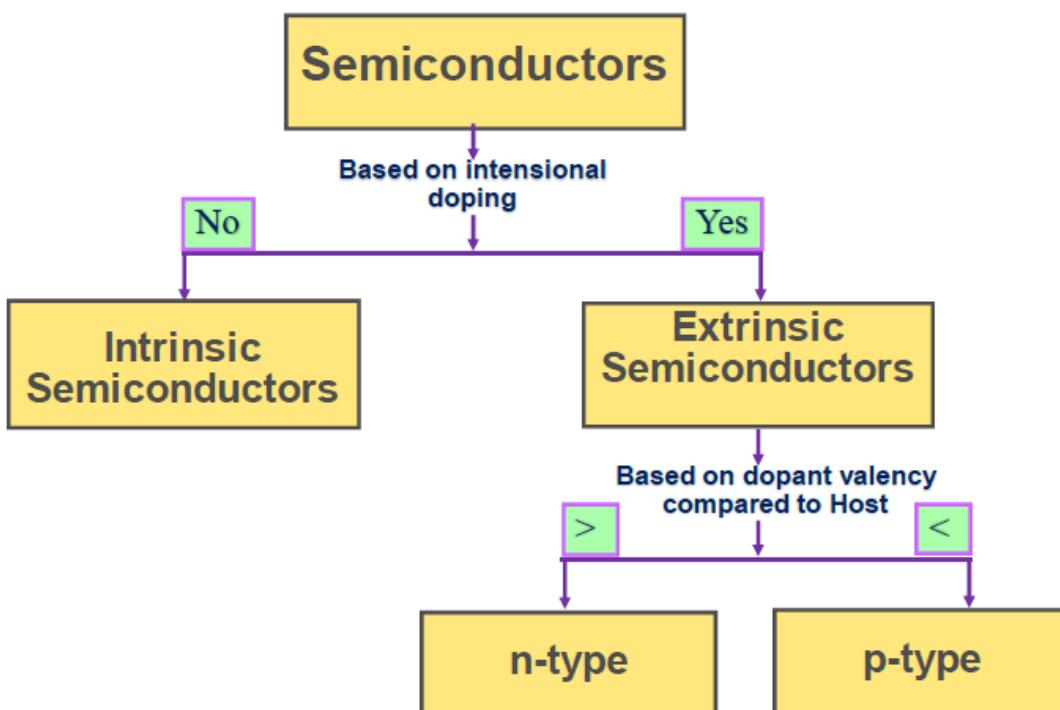


Figure 5: The energy band gap or Forbidden energy gap of semiconductors.

## Classification of Semiconductors:

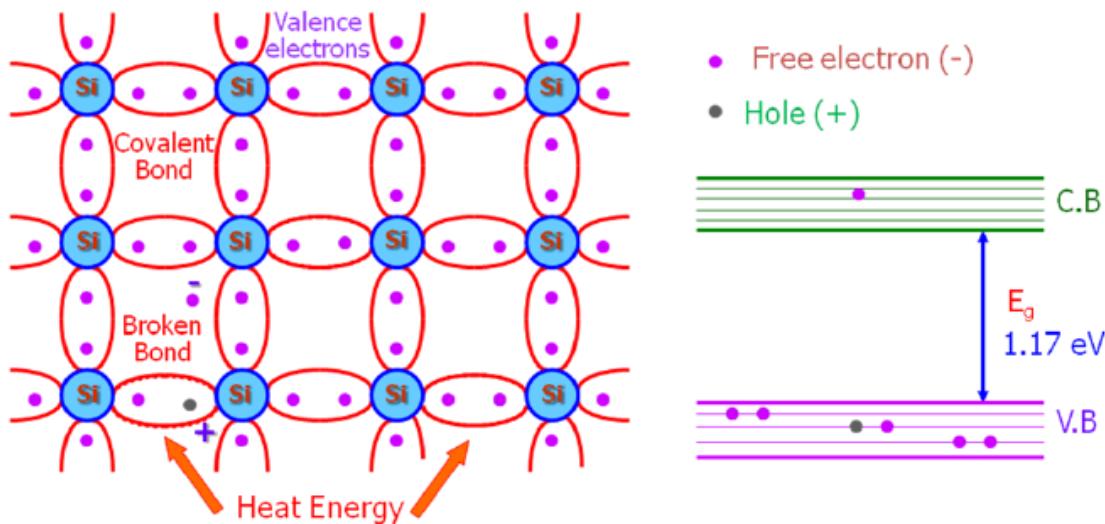




## Intrinsic semiconductors:

Semiconductor in pure form is known as Intrinsic Semiconductor.

Example: Pure Si and Pure Ge



**Figure 6:** Electrons and Holes configuration in a Silicon atom of Intrinsic Semiconductor.

Each Si atom has four valence electrons which are shared to form covalent bonds with the four surrounding Si atoms.

- **At T = 0K:** No electrons are available for current conduction, i.e. valence band is completely filled and conduction band is completely empty, and the material behaves as insulator.

### • **At Room Temperature:**

1. On receiving energy, one of the electrons from a covalent band breaks and is free to move in the crystal lattice, and contribute to current conduction.
2. While coming out of the covalent bond, electron leaves behind a vacancy named ‘hole’.
3. An electron from the neighbouring atom can break away and come to the place of the missing electron (or hole) for completing the covalent bond, which creates a hole at another place, resulting random movement of holes.
4. The completion of a bond may not be necessarily due to an electron from a bond of a neighbouring atom. The bond may be completed by a conduction band electron (free electron), which is referred as ‘electron -hole recombination’.
5. Thermal excitation results number of electron in conduction band is equal to the number of holes in valence band.

$$n_i = p_i$$

where  $n_i$  and  $p_i$  are the intrinsic carrier concentration.



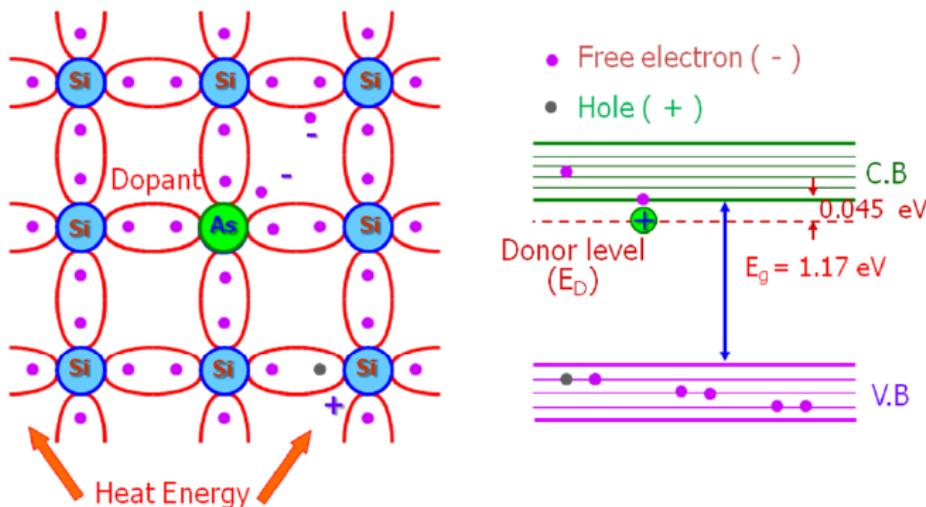
## Extrinsic semiconductors:

When a small amount, say, a few parts per million (ppm), of a suitable impurity is added to the pure semiconductor, the conductivity of the semiconductor is increased manifold. Such materials are known as extrinsic semiconductors or impurity semiconductors. The deliberate addition of a desirable impurity is called doping and the impurity atoms are called dopants. Such a material is also called a doped semiconductor. The dopant has to be such that it does not distort the original pure semiconductor lattice. It occupies only a very few of the original semiconductor atom sites in the crystal. A necessary condition to attain this is that the sizes of the dopant and the semiconductor atoms should be nearly the same.

The extrinsic semiconductor is of two types

### N-type semiconductors:

N-type Semiconductor is formed when pentavalent impurity (As or P) is doped to pure semiconductor of tetravalent element (Si or Ge).



**Figure 7:** Electrons and Holes configuration in a Silicon atom of N-type extrinsic semiconductors.

When Silicon (Si) is doped with arsenic (As), the four valence electrons of As form covalent bonds with four Si atoms and the fifth electron of As atom is loosely bound.

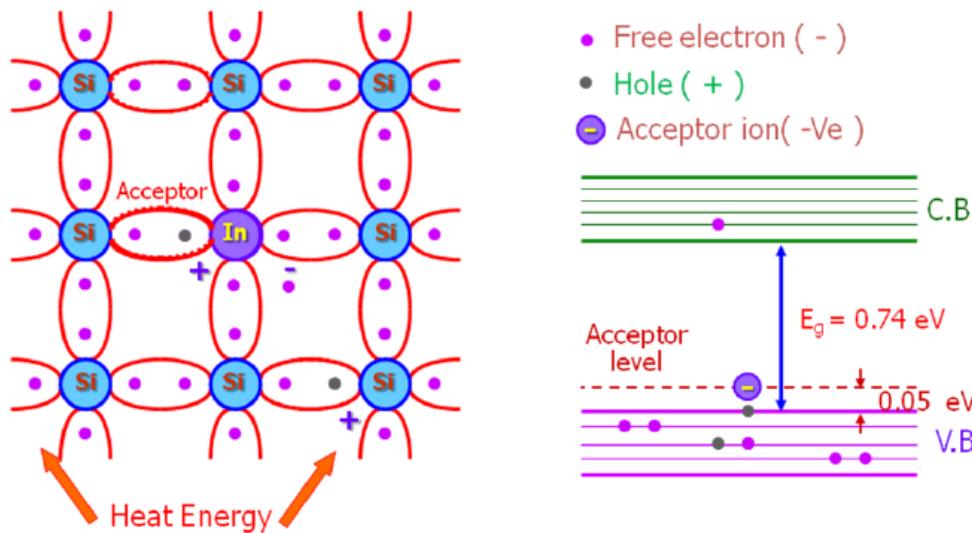
- The energy required to detach the fifth loosely bound electron is only of the order of 0.054 eV for Si.
- The energy required to detach the fifth loosely bound electron is only of the order of 0.054 eV for Si, which is provided by thermal agitation.
- The force of attraction between this mobile electron and the positively charged (+5) impurity ion is weakened by the dielectric constant of the medium.
- Such electrons from impurity atoms will have energies slightly less than the energies of the electrons in the conduction band.



- The energy state corresponding to the fifth electron is in the forbidden gap and slightly below the lower level of the conduction band, and the energy level is referred as donor level.

### P-type semiconductors:

P-type Semiconductor is formed when trivalent impurity (In or Ga) is doped to pure semiconductor of tetravalent element (Si or Ge).



**Figure 7:** Electrons and Holes configuration in a Silicon atom of P-type extrinsic semiconductors.

When Silicon (Si) is doped with indium (In), the three valence electrons of In form three covalent bonds with three Si atoms. The vacancy that exists with the fourth covalent bond with fourth Si atom constitutes a hole.

- The hole which is deliberately created may be filled with an electron from neighbouring atom, creating a hole in that position from where the electron jumped.
- Since the hole is associated with a positive charge moving from one position to another, the semiconductor is called p – type semiconductor.
- The trivalent impurity atom is called ‘acceptor’, which produces an energy level just above the valence band, and is called ‘acceptor level’.
- The energy difference between the acceptor energy level and the top of the valence band is of the order of meV, which is much smaller than the band gap.
- Electrons from the valence band can easily move into the acceptor level by thermal agitation.
- In a p-type semiconductor, holes are the majority charge carriers and the electrons are the minority charge carriers.
- If  $n$  and  $p$  represent the electron and hole concentrations respectively in p-type semiconductor, then
$$np = n_i p_i = n_i^2$$
where,  $n_i$  and  $p_i$  are intrinsic carrier concentration.



## Lecture-2

### Fermi-Dirac Probability distribution function

The Fermi-Dirac distribution function  $f(E)$  specifies how many of the existing states at the energy  $E$  will be filled with electrons. The function  $f(E)$  specifies, under equilibrium conditions, the probability that an available state at an energy  $E$  will be occupied by an electron. It is a probability distribution function and expressed as

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}$$

Where  $E_F$ = Fermi energy or Fermi level,

$k$ = Boltzmann constant =  $1.38 \times 10^{-23} \text{ J/K} = 8.6 \times 10^{-5} \text{ eV/K}$ ,

$T$ = absolute temperature in K.

**The highest energy level that can be occupied by an electron at 0K is called Fermi energy level. It is denoted by  $E_F$ .**

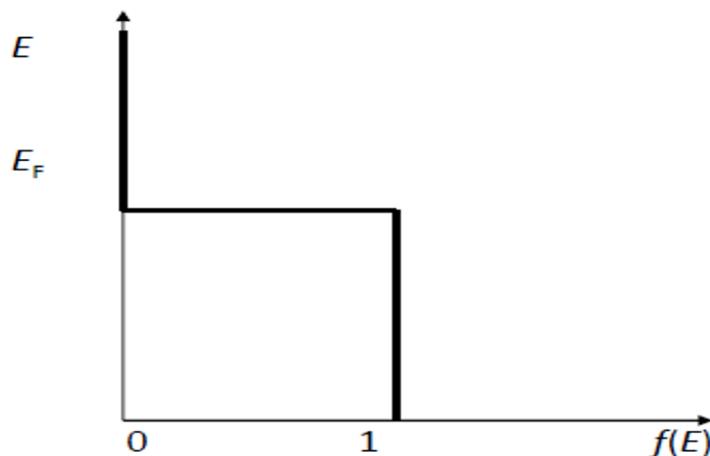
**Case I: Consider  $T = 0 \text{ K}$**

For  $E > E_F$ :

$$f(E > E_F) = \frac{1}{1 + e^{(+\infty)}} = 0$$

For  $E < E_F$ :

$$f(E > E_F) = \frac{1}{1 + e^{(-\infty)}} = 1$$



**Figure 8:** Fermi-Dirac distribution function at  $T=0\text{K}$



### Case II: Consider $T > 0$ K

If  $E = E_F$  then  $f(E_F) = \frac{1}{2}$

If  $E \geq E_F + 3kT$  then  $e^{\left(\frac{E-E_F}{kT}\right)} \gg 1$

Thus the following approximation is valid:

$$f(E) = e^{\left(\frac{-(E-E_F)}{kT}\right)}$$

i.e., most states at energies  $3kT$  above  $E_F$  are empty.

If then

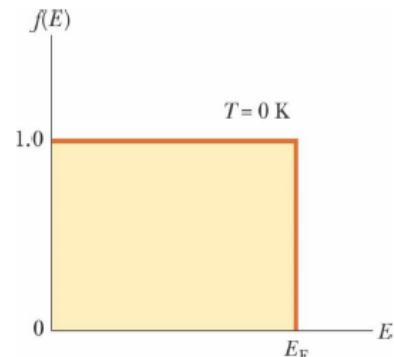
$$E \leq E_F - 3k \quad \text{then } e^{\left(\frac{E-E_F}{kT}\right)} \ll 1$$

Thus the following approximation is valid:

So,  $1-f(E) = \text{Probability that a state is empty, decays to zero.}$

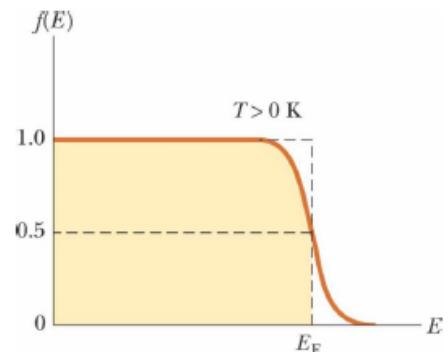
So, most states will be filled.

$kT$  (at 300 K) = 0.025eV,  $Eg(\text{Si}) = 1.1\text{eV}$ , so  $3kT$  is very small in comparison.



### Fermi-Dirac Distribution Function at T= 0

- At  $T=0$ , all states having energies less than the Fermi energy are occupied.
- All states having energies greater than the Fermi energy are vacant.



### Fermi-Dirac Distribution Function at T> 0

- As  $T$  increases, the distribution rounds off slightly
- States near and below  $E_F$  lose population
- States near and above  $E_F$  gain population

Figure 9: Fermi-Dirac distribution function at  $T=0\text{K}$  and  $T \neq 0\text{ K}$ .



## Lecture-3

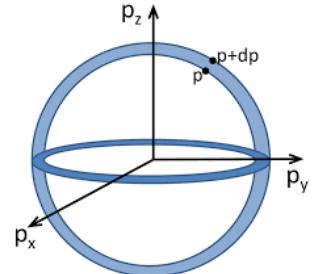
**Density of State:** The function  $D(E)$  gives the number of states per unit energy range and is referred to as the density of states.

The number of quantum states  $D(E)dE$  within the energy values

$E$  and  $E+dE$  or within the volume of the momentum space

between concentric spheres of radii  $p$  and  $p+dp$ , can be written as

$$D(E)dE = \frac{V}{h^3} 4\pi p^2 dp; \quad E = \frac{p^2}{2m}$$



$$D(E)dE = \frac{2\pi V}{h^3} (2m)^{3/2} E^{1/2} dE$$

$$D(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2}$$

Note:  $D(E)dE = 1/h^3 \times \text{volume in space coordinate} \times \text{Volume in Momentum space}$

## Thermal Equilibrium:

Thermal equilibrium is referred to a steady-state condition at a given temperature without any external excitation such as light, pressure and electric field.

• The semiconductor is in thermal equilibrium implies that:

1. No external excitation other than temperature.
2. No net motion of charge or energy

## Equilibrium electron concentration:

The density of states for electron at  $E$  is given by

$$D_e(E) = \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} E^{1/2}$$

For electron at  $E$  in conduction band:

$$E = E_c + \frac{\hbar^2 k^2}{2m}$$



$$\Rightarrow \frac{\hbar^2 k^2}{2m} = E_k = E - E_c$$

Density of states for electron at E in conduction band is given by:

$$D_e(E) = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2}$$

Probability occupancy of electron at E in conduction band is given:

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}}$$

For  $E - E_F \gg kT$

$$f(E) \approx e^{-(E-E_F)/kT}$$

Thermal-equilibrium concentration of electrons in conduction band may be obtained by integrating the product of the density of allowed energy states for electron and probability occupancy of electrons in the conduction band.

$$n = \int_{E_c}^{\infty} D_e(E) f(E) dE$$

From the above equations, we can write

$$n = \int_{E_c}^{\infty} \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} (E - E_c)^{1/2} e^{-(E-E_F)/kT} dE$$

$$\text{Let } x = \frac{E - E_c}{kT} \quad \Rightarrow E - E_c = xkT \quad \Rightarrow (E - E_c)^{1/2} = (xkT)^{1/2}$$

and  $dE = kT dx$

At  $E = E_c$ ,  $x = 0$  and  $E = \infty$ ,  $x = \infty$

So, we can get

$$\begin{aligned} n &= \int_0^{\infty} \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} (xkT)^{1/2} e^{-(E-E_F)/kT} kT dx \\ &= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} \int_0^{\infty} (x)^{1/2} e^{-(E-E_c-E_F+E_c)/kT} dx \\ &= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} \int_0^{\infty} (x)^{1/2} e^{-(E-E_c)/kT} e^{-(E_c-E_F)/kT} dx \end{aligned}$$



$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_c - E_F)/kT} \int_0^\infty (x)^{1/2} e^{-x} dx$$

$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_c - E_F)/kT} \Gamma \left( \frac{3}{2} \right)$$

$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_c - E_F)/kT} \left( \frac{\sqrt{\pi}}{2} \right)$$

$$= \frac{1}{4\pi^{3/2}} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_c - E_F)/kT}$$

$$= 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2} e^{-(E_c - E_F)/kT}$$

$$n = N_c e^{-(E_c - E_F)/kT}$$

Where  $N_c = 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2}$ ; effective density of states in CB

$N_c: 2.86 \times 10^{19} \text{ cm}^{-3}$  for Si at 300 K

$N_c: 4.7 \times 10^{19} \text{ cm}^{-3}$  for GaAs at 300 K

## Equilibrium hole concentration:

Density of states for hole at E is given by:

$$D_h(E) = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} E^{1/2}$$

For hole at E in valence band:

$$E = E_V - \frac{\hbar^2 k^2}{2m}$$

$$\Rightarrow \frac{\hbar^2 k^2}{2m} = E_k = E_V - E$$

Density of states for hole at E in valence band is given by:

$$D_h(E) = \frac{1}{2\pi^2} \left( \frac{2m}{\hbar^2} \right)^{3/2} (E_V - E)^{1/2}$$

Probability occupancy of hole at E in valence band is given:



$$\begin{aligned}
 1 - f(E) &= 1 - \frac{1}{1 + e^{(E-E_F)/kT}} \\
 &= \frac{1 + e^{(E-E_F)/kT} - 1}{1 + e^{(E-E_F)/kT}} \\
 &= \frac{e^{(E-E_F)/kT}}{1 + e^{(E-E_F)/kT}} \\
 &= \frac{e^{(E-E_F)/kT}}{e^{(E-E_F)/kT} [e^{-(E-E_F)/kT} + 1]} \\
 &= \frac{1}{e^{(E_F-E)/kT} + 1}
 \end{aligned}$$

For  $E_F - E \gg kT$

$$f(E) \approx e^{-(E_F - E)/kT}$$

Thermal-equilibrium concentration of hole in valence band may be obtained by integrating the product of the density of allowed energy states for hole and probability occupancy of hole in the valence band.

$$p = \int_{-\infty}^{E_V} D_h(E) f(E) dE$$

From the above equations, we can write

$$p = \int_{-\infty}^{E_V} \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} (E_V - E)^{1/2} e^{-(E_F - E)/kT} dE$$

$$\text{Let } x = \frac{E_V - E}{kT} \quad \Rightarrow E_V - E = xkT \quad \Rightarrow (E_V - E)^{1/2} = (xkT)^{1/2}$$

$$\text{and} \quad dE = -kT dx$$

$$\text{At } E = E_V, x = 0 \text{ and } E = -\infty, x = -\infty$$

So, we can get

$$\begin{aligned}
 p &= \int_{-\infty}^{E_V} \frac{1}{2\pi^2} \left(\frac{2m}{\hbar^2}\right)^{3/2} (xkT)^{1/2} e^{-(E_F - E)/kT} (-kT dx) \\
 &= \frac{1}{2\pi^2} \left(\frac{2mkT}{\hbar^2}\right)^{3/2} \int_0^{\infty} (x)^{1/2} e^{-(E_F - E_V + E_V - E)/kT} dx \\
 &= \frac{1}{2\pi^2} \left(\frac{2mkT}{\hbar^2}\right)^{3/2} \int_0^{\infty} (x)^{1/2} e^{-(E_V - E)/kT} e^{-(E_F - E_V)/kT} dx
 \end{aligned}$$



$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_F - E_V)/kT} \int_0^\infty (x)^{1/2} e^{-x} dx$$

$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_F - E_V)/kT} \Gamma \left( \frac{3}{2} \right)$$

$$= \frac{1}{2\pi^2} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_F - E_V)/kT} \left( \frac{\sqrt{\pi}}{2} \right)$$

$$= \frac{1}{4\pi^{3/2}} \left( \frac{2mkT}{\hbar^2} \right)^{3/2} e^{-(E_F - E_V)/kT}$$

$$= 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2} e^{-(E_F - E_V)/kT}$$

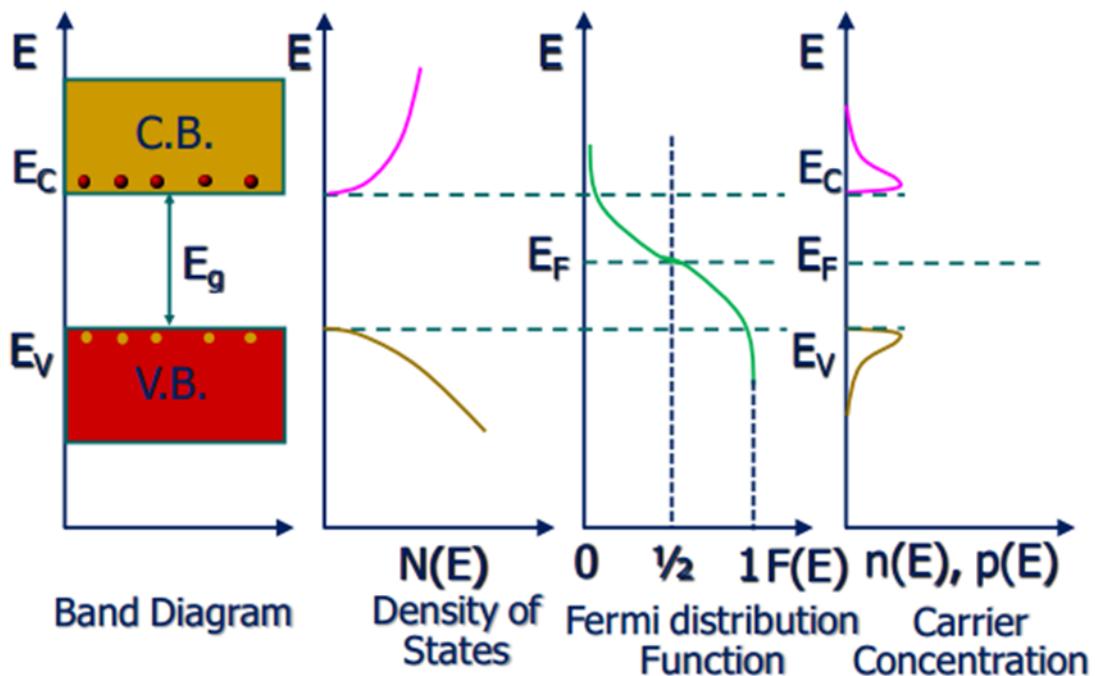
$$p = N_V e^{-(E_F - E_V)/kT}$$

Where  $N_V = 2 \left( \frac{mkT}{2\pi\hbar^2} \right)^{3/2}$ ; effective density of states in VB

$N_V: 2.66 \times 10^{19} \text{ cm}^{-3}$  for Si at 300 K

$N_V: 7.0 \times 10^{18} \text{ cm}^{-3}$  for GaAs at 300 K

### Graphical Representation





**Figure 10:** Band Diagram, density of states, fermi distribution function and carrier concentration in conduction band and valence band.

## Lecture-4

### Intrinsic Carrier Concentration:

For intrinsic Semiconductor:

Electron concentration in CB = hole concentration in VB

$$n = p = n_i$$

Where  $n_i$  is intrinsic carrier concentration.

Electron concentration in CB:

$$n = n_i = N_c e^{-(E_c - E_F)/kT}$$

Hole concentration in VB:

$$p = n_i = N_v e^{-(E_F - E_V)/kT}$$

Product of equation

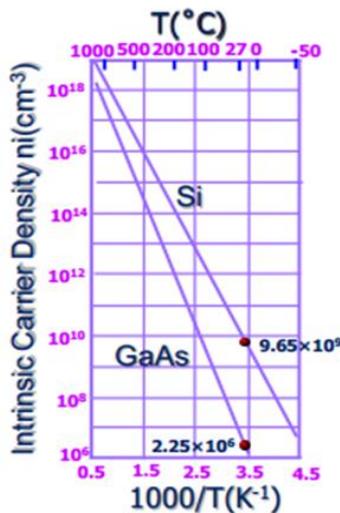
$$np = n_i^2 = N_c N_v e^{-(E_c - E_F)/kT} e^{-(E_F - E_V)/kT}$$

$$n_i^2 = N_c N_v e^{-(E_c - E_F + E_F - E_V)/kT}$$

$$n_i^2 = N_c N_v e^{-(E_c - E_V)/kT}$$

$$n_i = \sqrt{N_c N_v} e^{-(E_g)/2kT}$$

For a given semiconductor at a constant temperature the intrinsic carrier concentration ( $n_i$ ) does not depend on the Fermi energy.



At room temperature (300K):

$$n_i: 9.65 \times 10^9 \text{ cm}^{-3} \text{ for Si}$$

$$n_i: 2.25 \times 10^6 \text{ cm}^{-3} \text{ for GaAs}$$

**Figure 11:** Temperature dependence intrinsic carrier density of different materials.

## Intrinsic Fermi Level Position:

For intrinsic Semiconductor:

Electron concentration in CB = hole concentration in VB

$$n_i = p_i$$

$$N_c e^{-(E_c - E_F)/kT} = N_v e^{-(E_F - E_v)/kT}$$

$$\Rightarrow \frac{e^{-(E_c - E_F)/kT}}{e^{-(E_F - E_v)/kT}} = \frac{N_v}{N_c}$$

$$\Rightarrow e^{-(E_c - E_F - E_F + E_v)/kT} = \frac{N_v}{N_c}$$

$$\Rightarrow e^{-(E_c + E_v) + 2E_F/kT} = \frac{N_v}{N_c}$$

$$\Rightarrow [-(E_c + E_v) + 2E_F]/kT = \ln \left[ \frac{N_v}{N_c} \right]$$

$$\Rightarrow \frac{2E_F}{kT} = \frac{(E_c + E_v)}{kT} + \ln \left[ \frac{N_v}{N_c} \right]$$

$$\Rightarrow E_F = \frac{(E_c + E_v)}{2} + \frac{kT}{2} \ln \left[ \frac{N_v}{N_c} \right]$$



$$\Rightarrow E_F = \frac{(E_C + E_V)}{2} + \frac{kT}{2} \ln \left[ \frac{2 \left( \frac{m_h kT}{2\pi\hbar^2} \right)^{3/2}}{2 \left( \frac{m_e kT}{2\pi\hbar^2} \right)^{3/2}} \right]$$

$$\Rightarrow E_F = \frac{(E_C + E_V)}{2} + \frac{kT}{2} \ln \left[ \frac{m_h}{m_e} \right]^{3/2}$$

$$\Rightarrow E_F = \frac{(E_C + E_V)}{2} + \frac{3kT}{4} \ln \left[ \frac{m_h}{m_e} \right]$$

For  $m_h = m_e$  :

$$E_F = \frac{(E_C + E_V)}{2} + \frac{3kT}{4} \ln[1]$$

$$E_F = \frac{(E_C + E_V)}{2} \quad (\text{E}_F \text{ at the centre of Bandgap})$$

For  $m_h > m_e$  :

$$E_F = \frac{(E_C + E_V)}{2} + \frac{3kT}{4} \times x \quad (\text{E}_F \text{ above the centre of Bandgap})$$

Where  $x = \ln \left[ \frac{m_h}{m_e} \right]$  is non zero value.

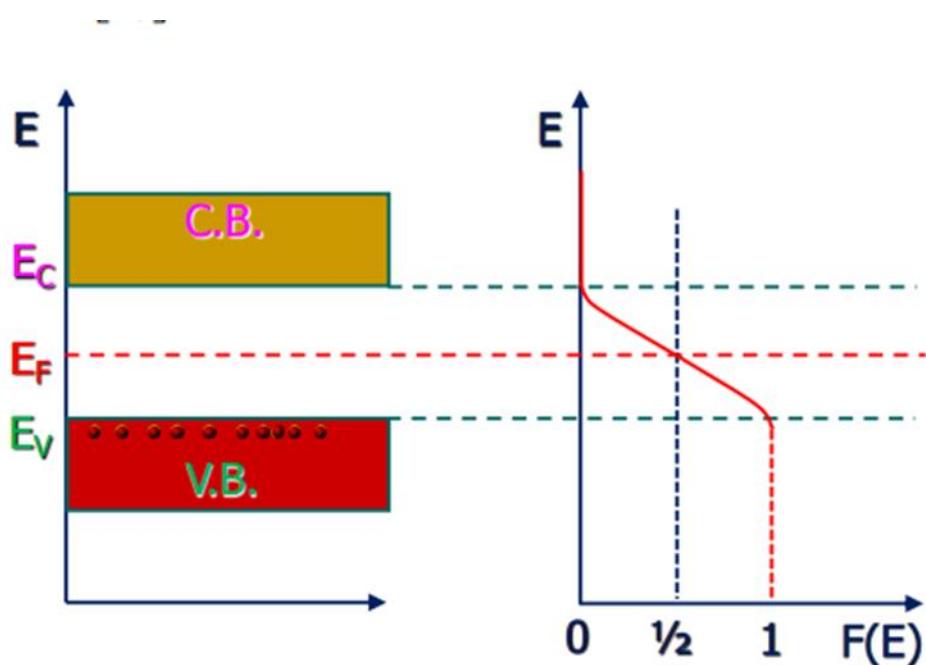




Figure 12: Variation of fermi energy level in an intrinsic semiconductor.

## Nondegenerate Semiconductor:

A nondegenerate semiconductor is a semiconductor with such a low level of doping that the Fermi energy is at least  $3kT$  away from either band edge.

- For shallow donors in Si and GaAs, there usually the enough thermal energy to supply the energy  $E_D$  to ionize all donor impurities.
- Under complete ionization condition, the loosely bound valence electrons of donor impurities move to conduction band leaving positive ion at donor level.
- The number of electron in conduction band is equal to the donor concentration ( $N_D$ ).

$$n = N_D$$

We know,

$$N_c e^{-(E_c - E_F)/kT} = N_D$$

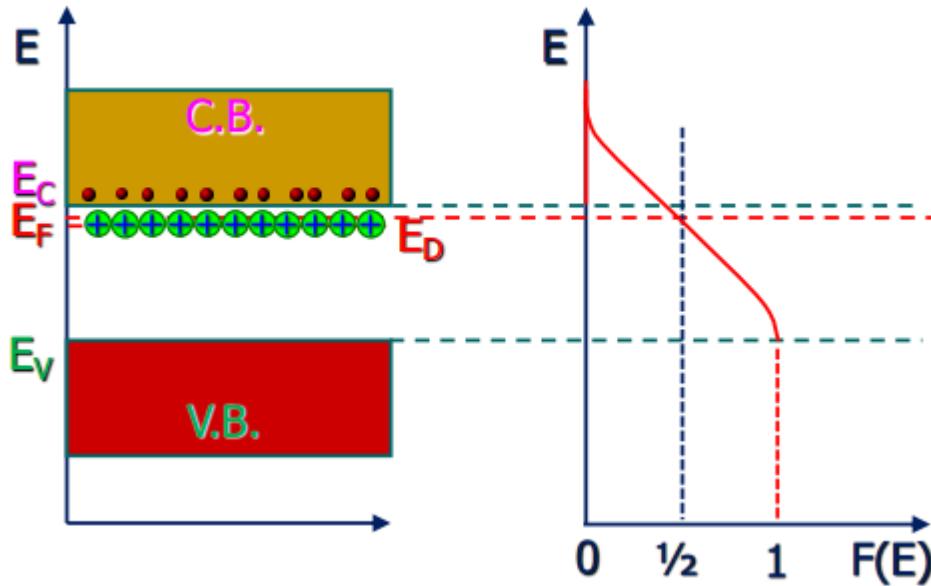
$$e^{-(E_c - E_F)/kT} = \frac{N_D}{N_c}$$

$$\ln[e^{-(E_c - E_F)/kT}] = \ln\left[\frac{N_D}{N_c}\right]$$

$$-(E_c - E_F)/kT = \ln\left[\frac{N_D}{N_c}\right]$$

$$(E_c - E_F) = kT \ln\left[\frac{N_c}{N_D}\right]$$

Thus  $(E_c - E_F)$  decreases with increasing the donor impurities, i.e Fermi level move towards CB.



**Figure 13:** Variation of EF and  $f(E)$  with donor concentration in N-type semiconductor.

For shallow acceptors in Si and GaAs, there usually the enough thermal energy to supply the energy  $EA$  to ionize all acceptor impurities.

Under complete ionization condition, the acceptor impurities accepted the additional electron of host atoms, which resulted creation of positive charge hole in valence band.

The number of holes in the valence band is equal to the acceptor concentration ( $N_A$ ).

$$p = N_A$$

We know,

$$N_V e^{-(E_F - E_c)/kT} = N_A$$

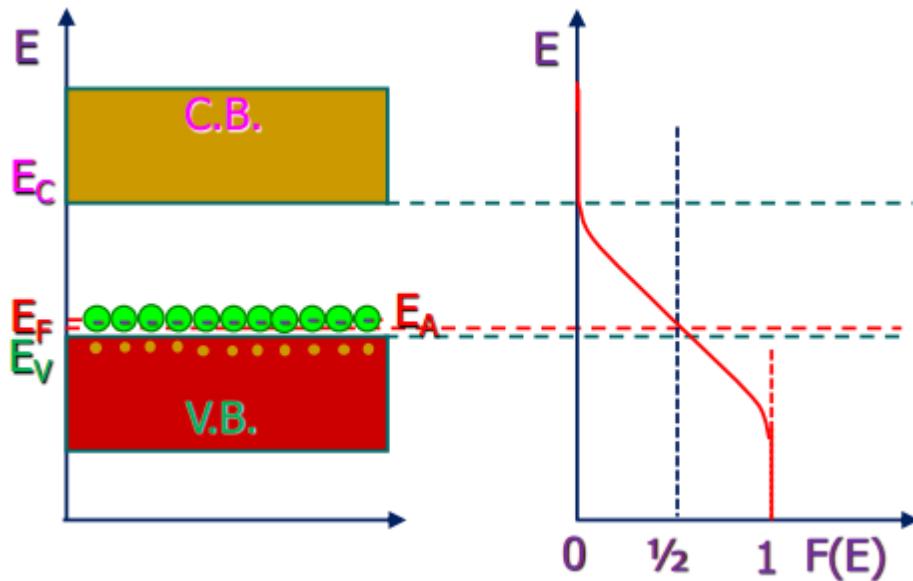
$$e^{-(E_F - E_c)/kT} = \frac{N_A}{N_V}$$

$$\ln[e^{-(E_F - E_c)/kT}] = \ln\left[\frac{N_A}{N_V}\right]$$

$$-(E_F - E_c)/kT = \ln\left[\frac{N_A}{N_V}\right]$$

$$(E_F - E_c) = kT \ln\left[\frac{N_V}{N_A}\right]$$

Thus  $(E_F - E_V)$  decreases with increasing the acceptor impurities, i.e. Fermi level move towards VB.



**Figure 14:** Variation of  $E_F$  and  $f(E)$  with acceptor concentration in P-type semiconductor.

## Lecture -5

### Electron and hole mobility

#### Electron mobility

The ability of an electron to move through a metal or semiconductor, in the presence of applied electric field is called electron mobility.

It is mathematically written as

$$V_n = \mu_n E$$

$$\mu_n = \frac{V_n}{E}$$

Where  $V_n$  = drift velocity of electrons

$\mu_n$  = mobility of electrons

$E$  = applied electric field

Let us consider a semiconductor that consists of large number of free electrons. When there is no voltage or electric field applied across the semiconductor, the free electrons moves randomly.



However, when the voltage or electric field is applied across the semiconductor, each free electron starts to move more quickly in particular direction. Electrons move very fast in vacuum. However, in metals or semiconductors, free electrons do not move very fast instead they move with a finite average velocity, called drift velocity.

Drift velocity is directly proportional to electric field. Hence, when the electric field increases drift velocity also increases. However, mobility of electrons is independent of applied electric field i.e. change in electric field does not change the mobility of electrons.

The SI unit of electric field is V/m, and the SI unit of velocity is m/s. Thus, the SI unit of mobility is m<sup>2</sup>/ (V.s).

### Hole mobility

The ability of a hole to move through a metal or semiconductor, in the presence of applied electric field is called hole mobility.

It is mathematically written as

$$V_p = \mu_p E$$

$$\mu_p = \frac{V_p}{E}$$

Where  $v_p$  = drift velocity of holes

$\mu_p$  = mobility of holes

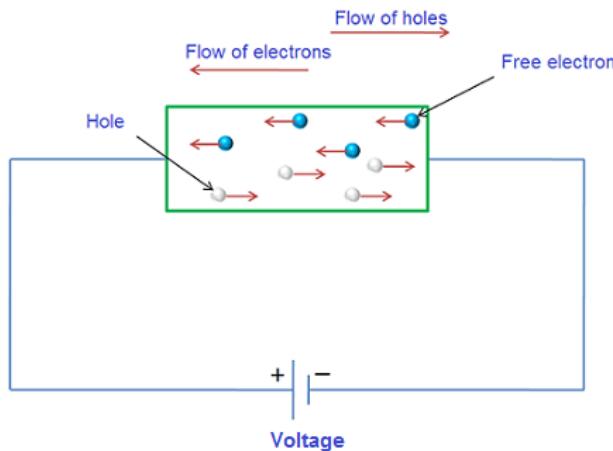
E = applied electric field

### Drift current

The flow of charge carriers, which is due to the applied voltage or electric field is called drift current.

In a semiconductor, there are two types of charge carriers, they are electrons and holes. When the voltage is applied to a semiconductor, the free electrons move towards the positive terminal of a battery and holes move towards the negative terminal of a battery.

Electrons are the negatively charged particles and holes are the positively charged particles. As we already discussed that like charges repel each other and unlike charges attract each other. Hence, the electrons (negatively charged particle) are attracted towards the positive terminal of a battery and holes (positively charged particle) are attracted towards the negative terminal.



**Figure 15:** Schematic diagram to drift current in the materials.

In a semiconductor, the electrons always try to move in a straight line towards the positive terminal of the battery. But, due to continuous collision with the atoms they change the direction of flow. Each time the electron strikes an atom it bounces back in a random direction. The applied voltage does not stop the collision and random motion of electrons, but it causes the electrons to drift towards the positive terminal. The average velocity that an electron or hole achieved due to the applied voltage or electric field is called drift velocity.

The drift velocity of electrons is given by

$$V_n = \mu_n E$$

The drift velocity of holes is given by

$$V_p = \mu_p E$$

Where  $v_n$  = drift velocity of electrons

$v_p$  = drift velocity of holes

$\mu_n$  = mobility of electrons

$\mu_p$  = mobility of holes

E = applied electric field

The drift current density due to free electrons is given by

$$J_n = e n \mu_n E$$

and the drift current density due to holes is given by

$$J_p = e p \mu_p E$$

Where  $J_n$  = drift current density due to electrons

$J_p$  = drift current density due to holes

$e$  = charge of an electron =  $1.602 \times 10^{-19}$  Coulombs (C).

n = number of electrons

p = number of holes

Then the total drift current density is

$$\begin{aligned} J &= J_n + J_p \\ &= e n \mu_n E + e p \mu_p E \\ J &= e(n \mu_n + p \mu_p) E \end{aligned}$$

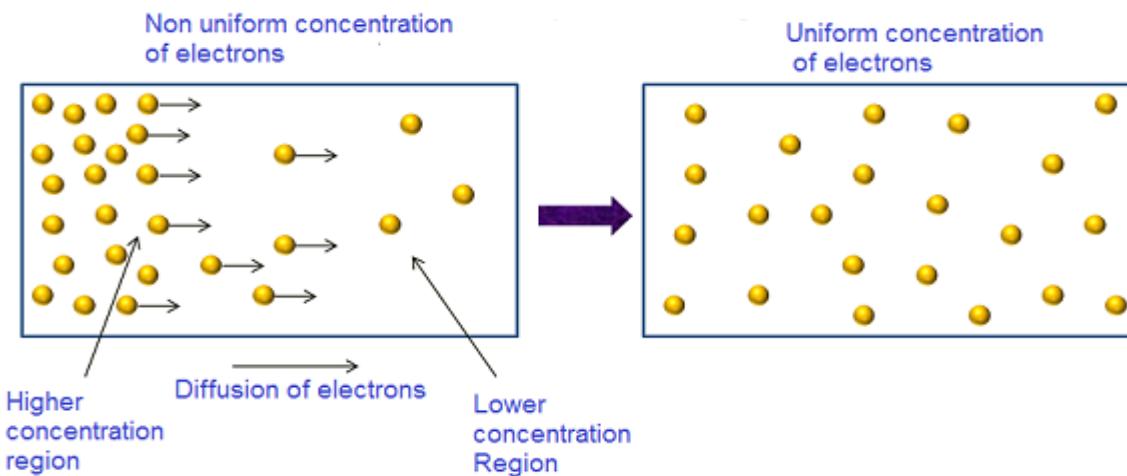


## Diffusion current

The process by which, charge carriers (electrons or holes) in a semiconductor moves from a region of higher concentration to a region of lower concentration is called diffusion.

The region in which more number of electrons is present is called higher concentration region and the region in which less number of electrons is present is called lower concentration region. Current produced due to motion of charge carriers from a region of higher concentration to a region of lower concentration is called diffusion current. Diffusion process occurs in a semiconductor that is non-uniformly doped.

Consider an n-type semiconductor that is non-uniformly doped as shown in below figure. Due to the non-uniform doping, more number of electrons is present at left side whereas lesser number of electrons is present at right side of the semiconductor material. The number of electrons present at left side of semiconductor material is more. So, these electrons will experience a repulsive force from each other.



**Figure 16:** Schematic diagram to understand diffusion current in the materials.

The electrons present at left side of the semiconductor material will move to right side, to reach the uniform concentration of electrons. Thus, the semiconductor material achieves equal concentration of electrons. Electrons that moves from left side to right side will constitute current. This current is called diffusion current. In p-type semiconductor, the diffusion process occurs in the similar manner.

Both drift and diffusion current occurs in semiconductor devices. Diffusion current occurs without an external voltage or electric field applied. Diffusion current does not occur in a conductor. The direction of diffusion current is same or opposite to that of the drift current.

### Concentration gradient

The diffusion current density is directly proportional to the concentration gradient. Concentration gradient is the difference in concentration of electrons or holes in a given area. If the concentration gradient is high, then the diffusion current density is also high. Similarly, if the concentration gradient is low, then the diffusion current density is also low.

The concentration gradient for n-type semiconductor is given by

$$J_n \propto \frac{dn}{dx}$$



The concentration gradient for p-type semiconductor is given by

$$J_p \propto \frac{dp}{dx}$$

Where  $J_n$  = diffusion current density due to electrons

$J_p$  = diffusion current density due to holes

### Diffusion current density

The diffusion current density due to electrons is given by

$$J_n = +e D_n \frac{dn}{dx}$$

Where  $D_n$  is the diffusion coefficient of electrons

The diffusion current density due to holes is given by

$$J_p = -e D_p \frac{dp}{dx}$$

Where  $D_p$  is the diffusion coefficient of holes

The total current density due to electrons is the sum of drift and diffusion currents.

$J_n$  = Drift current + Diffusion current

$$J_n = en\mu_n E + e D_n \frac{dn}{dx} \quad (\text{read topic : drift current})$$

The total current density due to holes is the sum of drift and diffusion currents.

$J_p$  = Drift current + Diffusion current

$$J_p = ep\mu_p E - e D_p \frac{dp}{dx}$$

The total current density due to electrons and holes is given by

$$J = J_n + J_p$$

## Lecture - 6

When an electrical current passes through a sample placed in a magnetic field, a voltage develops across the sample in a direction perpendicular to both the current and the magnetic field. This is known as Hall effect.

The basic experimental setup for study of Hall effect is shown in Fig. 17. A rectangular slab of a semiconducting sample with its *width* ( $w$ ) along  $y$ -direction and *thickness* ( $d$ ) along  $z$ -direction is placed in a magnetic field of strength  $B$  directed along the  $z$ -direction. Now an electric current,  $I_H$  is made to pass through the sample along its length by maintaining a potential difference along  $x$ -direction. The corresponding current density is

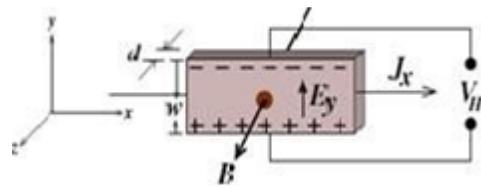


$$J_x = \frac{I_H}{wd}$$

Suppose that the charge carriers are positive, each having charge  $+q$ , and are moving along  $+x$  direction with velocity  $\vec{v}$ . Then the Lorentz force experienced by the carriers due to the magnetic field is

$$\vec{F}_B = q(\vec{v} \times \vec{B}) = -(qvB)\hat{y}$$

This force  $\vec{F}_B$  along  $-y$  direction deflects the positive charge carriers towards the *bottom surface* of the sample. This makes the *bottom surface* positively charged while leaving the *top surface* negatively charged.



**Figure 17:** Schematic diagram to Hall Effect.

This accumulation of charges near the bottom and top surfaces of the sample leads to the development of a transverse electric field along the  $y$ -direction. Force due to this electric field, opposed the Lorentz force and prevents further charge accumulation. In the steady state condition, these two forces balance out each other and we get,

$$qE_y = qvB$$

Now we define a quantity called *Hall coefficient*  $R_H$ , as the ratio of the electric field  $E_y$  to the current density  $J_x$  multiplied by magnetic field  $B$ , that is

$$R_H = \frac{E_y}{J_x B} = \frac{1}{nq}$$

where we have that  $J_x = nqv$ ,  $n$  being the number density ( $\text{m}^{-3}$ ) of charge carriers. In order to determine  $R_H$ , we can be writing  $v = J_x/nq$  and multiplying both sides of above equation by  $wd$ , we get

$$E_y wd = \frac{J_x wd B}{nq}$$

But  $E_y = V_H$ , the voltage across the top and bottom surfaces called the Hall voltage and  $J_x wd = I_H$ . This gives,

$$V_H = \left( \frac{R_H B}{d} \right) I_H$$

Therefore, if we measure the Hall voltage  $V_H$  against Hall current  $I_H$  for a fixed magnetic field  $B$  and plot  $V_H$  versus  $I_H$ , the curve will be a straight line with the slope  $m$  being,

$$m = \frac{R_H B}{d}$$

The Hall coefficient  $R_H$  can be calculated from the value of this slope,  $m$  if the thickness  $d$  of the sample is known. Once  $R_H$  is determined, the carrier density  $n$  can also be calculated. Now assume the situation where the charge carriers are negative with  $q = -e$ . In that case, for current direction along  $+x$ ,



the charges will be moving with velocity  $v$ .

The Lorentz force,  $\vec{F}_B = q(\vec{v} \times \vec{B}) = -(eBv)\hat{y}$  will still be along negative  $y$  direction as before. However, this time, the bottom surface acquires negative polarity and consequently, the sign of the Hall voltage  $V_H$  will be opposite to what was observed in case of positive charge. Thus for given directions of the Hall current and the magnetic field, we can determine the type of charge carriers (whether +ve or -ve) by looking at the sign of the Hall voltage  $V_H$ .

## Lecture-7

### The PN-junction diode:

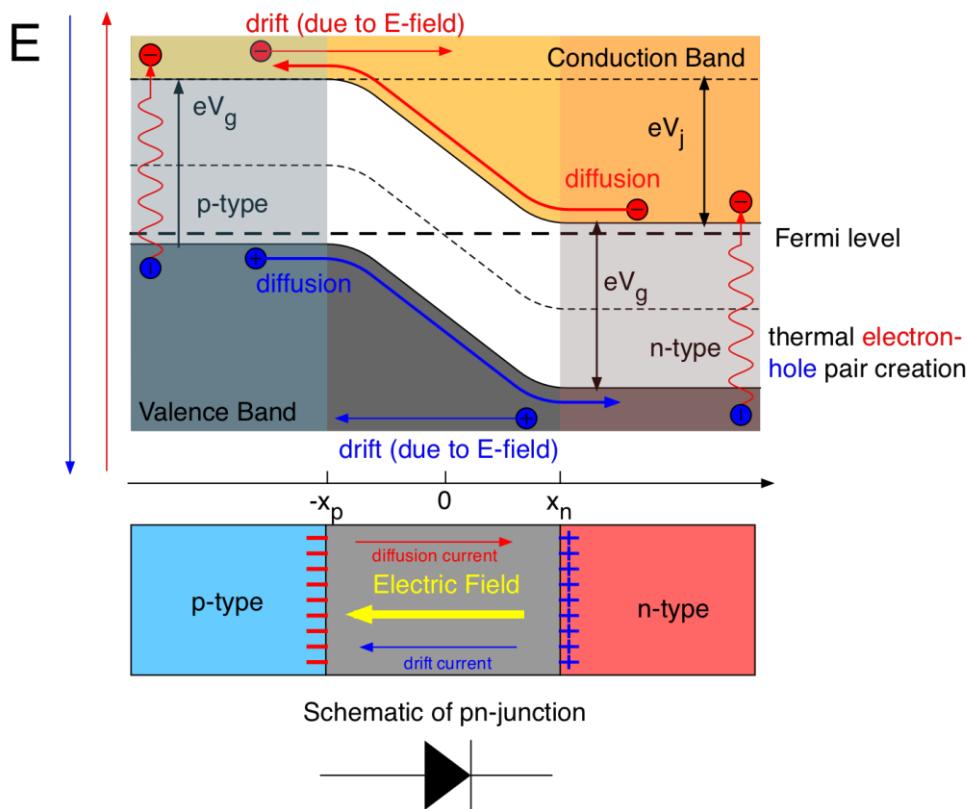
A pn junction at equilibrium is characterized by a depletion region where there are no charge carriers (except for those created and annihilated dynamically) and a contact potential. The contact potential is related to the dopant concentration in the p and n sides with higher concentrations leading to larger contact potentials. This, in turn, is related to the position of the Fermi levels in the p and n sides since a higher dopant concentration pushes the Fermi level closer to the valence or conduction band. The pn junction can



be biased by connecting to an external circuit and there are two types of biasing (similar to the arguments for the metal-semiconductor Schottky junction)

When a layer of P-type semiconducting material is placed with a layer of N-type semiconducting material in such a way that the atoms of P-type can combine with the atoms of N-type across the surface of contact, such a surface junction where combination has occurred is known as PN junction.

The energy band diagram of unbiased P-N Junction is as shown in the figure. The main concept of the formation of the P-N junction is that, the Fermi level of n type and P type will lie in same line.



**Figure 18:** Schematic diagram band structure of PN junction diode.

## Basic Principle:

Two important processes occur during the formation of a p-n junction: *diffusion and drift*. We know that in an n-type semiconductor, the concentration of electrons (number of electrons per unit volume) is more compared to the concentration of holes. Similarly, in a p-type semiconductor, the concentration of holes is more than the concentration of electrons. During the formation of p-n junction, and due to the concentration gradient across p-, and n-sides, holes diffuse from p-side to n-side and electrons diffuse from n-side to p-side. This motion of charge carries gives rise to diffusion current across the junction.



When an electron diffuses from n top, it leaves behind an ionised donor on n-side. This ionised donor (positive charge) is immobile as it is bonded to the surrounding atoms. As the electrons continue to diffuse from n top, a layer of positive charge (or positive space-charge region) on n-side of the junction is developed.

- Similarly, when a hole diffuses from p to n due to the concentration gradient, it leaves behind an ionised acceptor (negative charge) which is immobile. As the holes continue to diffuse, a layer of negative charge (or negative space-charge region) on the p-side of the junction is developed. This space-charge region on either side of the junction together is known as **depletion region** as the electrons and holes taking part in the initial movement across the junction depleted the region of its free charges. The thickness of depletion region is of the order of one-tenth of a micrometre.
- Due to the positive space-charge region on n-side of the junction and negative space charge region on p-side of the junction, an electric field directed from positive charge towards negative charge develops. Due to this field, an electron on p-side of the junction moves to n-side and a hole on n side of the junction move stop-side. The motion of charge carriers due to the electric field is called **drift**. Thus a drift current, which is opposite in direction to the diffusion current starts.

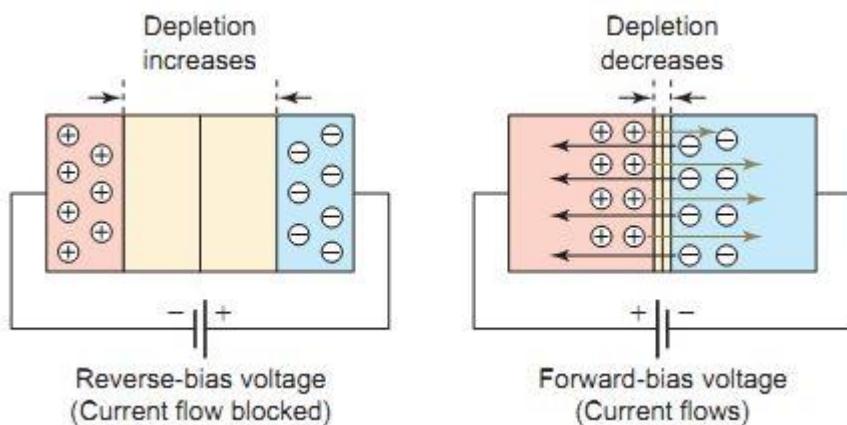
## Biasing of P-N Junction

When the P-N junction is subject to applied voltage across the terminal, it said to be biased PN Junction. The biasing is of 2 types:

- **Forward biased**
- **Reverse biased**

### Forward biased:

When an external V is applied across a semiconductor diode such that p-side is connected to the positive terminal of the battery and n-side to the negative terminal, it is said to be forward biased. The applied voltage mostly drops across the depletion region and the voltage drop across the p-side and n-side of the junction is negligible. (This is because the resistance of the depletion region—a region where there are no charges—is very high compared to the resistance of n-side and p-side.) The direction of the applied voltage (V) is opposite to the built-in potential  $V_0$ . As a result, the depletion layer width decreases and the barrier height is reduced. The effective barrier height under forward bias is  $(V_0 - V)$ . As the barrier height reduces and more number of carriers will have the required energy so the current increases.





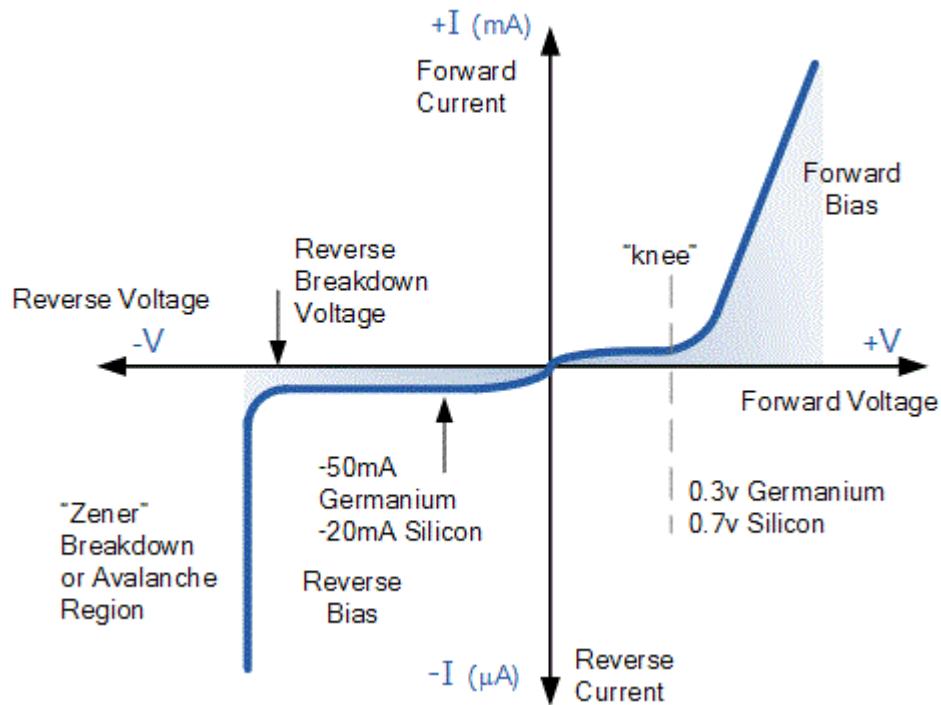
**Figure 19:** Schematic diagram of Reverse and Forward bias connection of PN junction diode.

### Reverse biased:

When an external voltage ( $V$ ) is applied across the diode such that n-side is positive and p-side is negative, it is said to be *reverse biased*. The applied voltage mostly drops across the depletion region. The direction of applied voltage is same as the direction of barrier potential. As a result, the barrier height increases and the depletion region widens due to the change in the electric field. The effective barrier height under reverse bias is ( $V_0 + V$ ). This suppresses the flow of electrons from n to p and holes from p to n. Thus, diffusion current, decreases enormously compared to the diode under forward bias. The drift current is also very small under forward bias but it is negligible ( $\mu\text{A}$ ) when compared with current due to injected carriers which is usually in mA.

### I-V Characteristics of P-N Junction:

The circuit arrangement for studying the V-I characteristics of a diode, (i.e., the variation of current as a function of applied voltage) are shown in Fig. 20. The battery is connected to the diode through a potentiometer that the applied voltage across the diode can be changed. For different values of voltages, the value of the current is noted. A graph between  $V$  and  $I$  is obtained as in Fig. 20. In forward bias, the current first increases very slowly, almost negligibly, till the voltage across the diode crosses a certain value. After the characteristic voltage, the diode current increases significantly even for a very small increase in the diode bias voltage. This voltage is called the **threshold voltage or cut-in voltage**. For the diode in reverse bias, the current is very small ( $\sim \mu\text{A}$ ) and almost remains constant with change in bias. It is called reverse **saturation current**. However, for special cases, at very high reverse bias (break down voltage), the current suddenly increases.



**Figure 20:** IV characteristic of PN junction diode.



## BIPOLAR JUNCTION TRANSISTOR

### What is a Bipolar Junction Transistor (BJT)?

The Bipolar junction transistor is a solid-state device and, in these transistors, the current flow in two terminals are emitter and collector, and the flow of current controlled by the third terminal is the base terminal. It is different from the other type of transistor, i.e. Field-effect transistor which is the output current controlled by the input voltage. The basic symbol of the BJTs n-type and p-type is shown below.

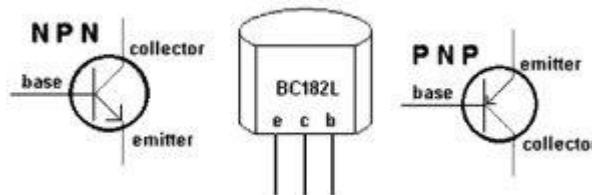


Figure 21: BJT or Bipolar Junction Transistor Symbol

### Why BJT is Called Bipolar?

A BJT is a 3-terminal semiconductor device, as the name suggests, the term bipolar is taken from the truth that this kind of transistor includes two kinds of semiconductor materials like P-type (positive type) and n-type (negative type) where the current flows from these regions. Usually, these transistors include Silicon.

### Construction of Bipolar Junction Transistor

The construction of BJT will determine its working characteristics. So, the construction of BJT can be done through three doped semiconductor sections which are separated through two PN-junctions. This transistor includes three layers namely base, emitter, and collector. These transistors are available in two types namely PNP and NPN and their physical structure is shown below.

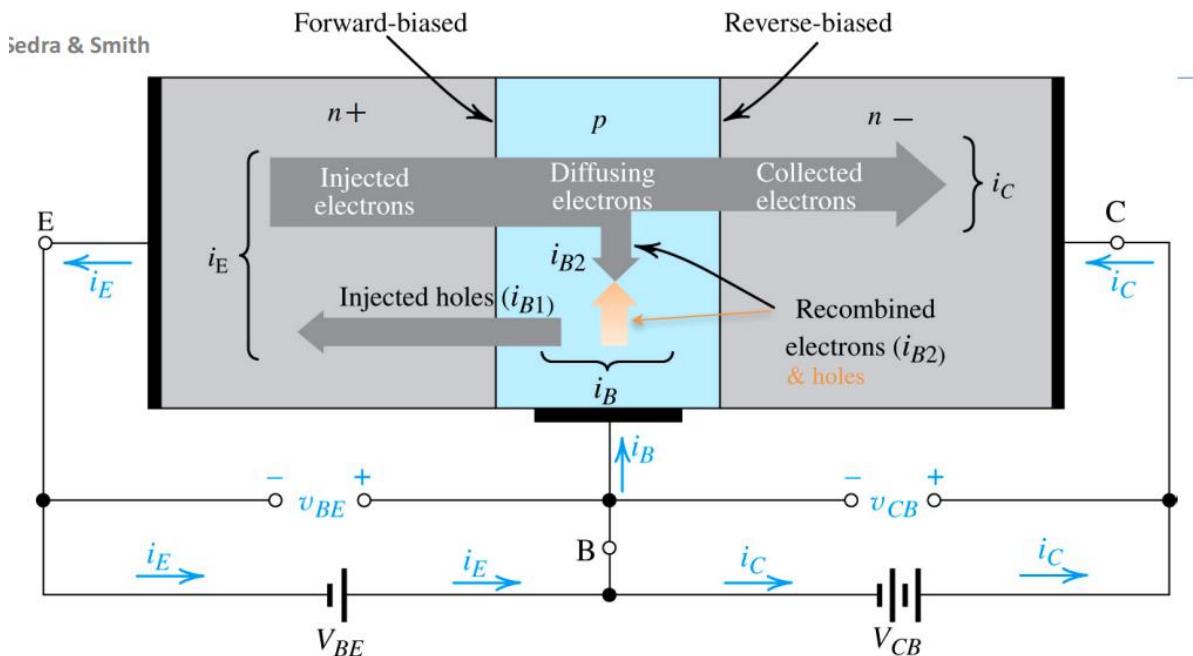


Figure 22: Construction of Bipolar Junction Transistor

In NPN, it includes two N-regions separated by one P-region whereas, in PNP, it includes two P-regions and separates with one N-region. In BJT, the term bipolar refers to the utilization of both the charge carriers like electrons and holes within the structure.



When a PN-junction connects the base and emitter region then it is known as a BE-junction. Similarly, once the pn-junction connects the base & the collector region then it is known as the BC-junction. A wire lead unites to every region and these leads are branded with C, B & E for the collector, base, and emitter respectively.

The base (B) – region is slightly doped & it is extremely thin as compared to the emitter terminal that is heavily doped & the somewhat doped collector region. The schematic symbols of the NPN and PNP BJTs include the following.

### Terminals of BJT

The BJT includes three terminals like base, Emitter, and Collector which are discussed in brief. The NPN and PNP transistors symbol representation can be done like the following. The arrow symbol on the emitter terminal is always there and the arrow direction will represent the current flow because of the charge carriers.

#### *Emitter Terminal*

The emitter terminal emits the charge carriers like holes or electrons to the other two terminals. The base terminal is always reverse biased with respect to other terminals like emitter so that it can produce huge majority carriers. In BJT, the emitter is the most heavily doped region.

The emitter-base junction must be connected in forwarding bias within both the transistors like PNP & NPN. The emitter terminal provides electrons toward the EB-junction within NPN transistor whereas it supplies charge carriers like holes to the same junction within the PNP transistor.

#### *Collector Terminal*

The part on the reverse side of the Emitter terminal will collect the emitted charge carriers called a collector. So, this terminal is heavily doped however the doping level of this terminal is among the base terminal which is a lightly doped & heavily doped emitter terminal. CB-junction must be reversely biased in both the transistors.

The main reason for this biasing is to eliminate charge carriers from the CB-junction. The collector terminal of the NPN transistor is to collects electrons that are emitted through the emitter terminal whereas, in the PNP transistor, it gathers holes that are emitted through the emitter terminal.

#### *Base Terminal*

The base terminal is the center part of collector & emitter terminals which forms two PN junctions among them. The base terminal is the most lightly doped segment in the transistor. So, being the central part of the BJT will allow it to manage the charge carriers flow among emitter & collector terminals. The BE junction shows high resistance as this junction can be connected in reversed bias.

### **Working of BJT**

Bipolar junction transistor is classified into two types namely PNP and NPN based on types of doping of the terminals. In NPN transistor, two semiconductor junctions are there which have a thin anode region doped with 'P' whereas PNP transistor includes two semiconductor junctions that include a thin cathode region doped with 'N'.

The flow of charge within a transistor is because of the charge carrier's diffusion among the two sections which belongs to different concentrations of charge carriers. The emitter section is doped highly as compared to the remaining layers.



Both the layers like base and collector include the same concentrations of charge carriers. So, between these three junctions, the BE junction can be connected in forwarding bias & the BC junction is reverse biased. The operating of BJT can be done in three different regions like active, saturation, and cut-off.

#### *Active Region*

In this region, one junction is connected in a forward bias whereas the other one is connected in reverse bias. Here, the base current ( $I_B$ ) is used to control the quantity of  $I_C$  (collector current). So, the active region can be mainly used for amplification wherever this transistor performs like an amplifier through a gain ' $\beta$ ' with the following equation;

$$I_C = \beta \times I_B$$

The active region is also called a linear region which lies in between the two regions like the cutoff & the saturation. The typical operation of this transistor happens within this region.

#### *Saturation Region*

In this region, the transistors both the junctions are connected in forwarding bias. So this region is mainly used for the ON-state for a switch wherever;

$$I_C = I_{sat}$$

Here, ' $I_{sat}$ ' is the saturation current and it is the highest quantity of current flow among the two terminals like emitter as well as collector once this transistor is connected in the saturation region. As these junctions are connected in forwarding bias then the transistor works like a short circuit.

#### *Cutoff Region*

In this region, both the transistor junctions are connected in reverse bias. Here this transistor works like the off condition of a switch wherever

$$I_C = 0$$

In the cut-off region, the operation is totally reverse as compared to the saturation region. So, there are no exterior supplies connected. If there is no collector current then there is no emitter current. In this method, the transistor works like an off-state of the switch, and this mode can be achieved by decreasing base voltage to below the voltage of both the emitter as well as collector.  $V_{BE} < 0.7$

### Types of Bipolar Junction Transistor

As we have seen a semiconductor offer less resistance to flow current in one direction and high resistance is another direction and we can call transistor as the device mode of the semiconductor. The Bipolar junction transistor consists of two types of transistors. Which, given us

- Point contact
- Junction transistor

By comparing two transistors the junction transistors are used more than point type transistors. Further, the junction transistors are classified into two types which are given below. There are three electrodes for each junction transistor they are emitter, collector, and base

- PNP junction transistors
- NPN junction transistors

#### **PNP Junction Transistor**



In the PNP transistors, the emitter is more positive with the base and also with respect to the collector. The PNP transistor is a three-terminal device that is made from semiconductor material. The three terminals are collector, base, and emitter, and the transistor is used for switching and amplifying applications. The operation of the PNP transistor is shown below.

Generally, the collector terminal is connected to the positive terminal and the emitter to a negative supply with a resistor either the emitter or collector circuit. To the base terminal, the voltage is applied and it operates transistor as an ON/OFF state. The transistor is in the OFF state when the base voltage is the same as the emitter voltage. The transistor mode is in ON state when the base voltage decreases with respect to the emitter. By using this property the transistor can act on both applications like switch and amplifier. The basic diagram of the PNP transistor is shown below.

### NPN Junction Transistor

The NPN transistor is exactly opposite to the PNP transistor. The NPN transistor contains three terminals which are the same as the PNP transistor which are emitter, collector, and base. The operation of the NPN transistor is

Generally, the positive supply is given to the collector terminal and the negative supply to the emitter terminal with a resistor either the emitter or collector or emitter circuit. To the base terminal, the voltage is applied and it operated as an ON/OFF state of a transistor. The transistor is in an OFF state when the base voltage is the same as the emitter. If the base voltage is increased with respect to the emitter then the transistor mode is in an ON state. By using this condition the transistor can act like both applications which are amplifier and switch. The basic symbol and the NPN configuration diagram as shown below.

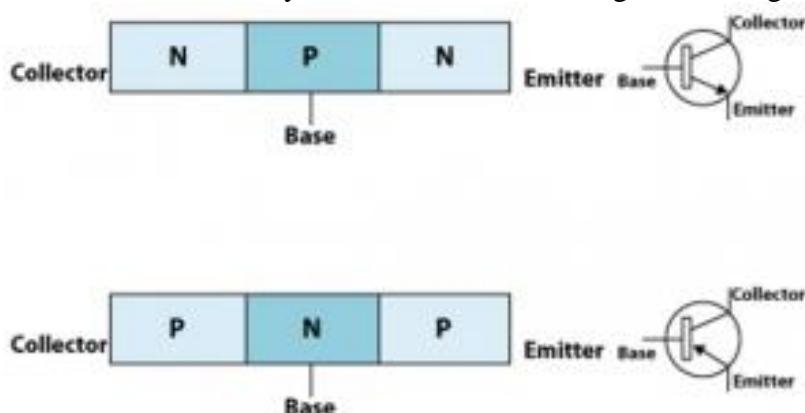


Figure 23: PNP & NPN Bipolar Junction Transistor

### Hetero Bipolar Junction

The Hetero bipolar junction transistor is also a type is the bipolar junction transistor. It uses different semiconductor materials to the emitter and base region and produces heterojunction. The HBT can handle the singles of very high frequencies of several hundred GHz generally it is used in ultrafast circuits and mostly used in radiofrequency. Its applications are used in cellular phones, and RF power amplifiers.



### Working Principle of BJT

The BE junction is forward bias and the CB is a reverse bias junction. The width of the depletion region of the CB junction is higher than the BE junction. The forward bias at the BE junction decreases the barrier potential and produces electrons to flow from the emitter to the base.

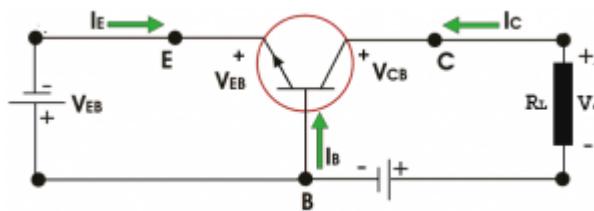
The base is thin and lightly doped, it has very few holes and less amount of electrons from the emitter about 2% it recombine in the base region with holes, and from the base terminal it will flow out. This initiates the base current flow due to the combination of electrons and holes. The leftover large number of electrons will pass the reverse bias collector junction to initiate the collector current. By using KCL we can observe the mathematical equation

$$I_E = I_B + I_C$$

The base current is very less as compared to emitter and collector current

$$I_E \sim I_C$$

Here the operation of the PNP transistor is the same as the NPN transistor the only difference is only holes instead of electrons. The below diagram shows the PNP transistor of the active mode region.



**Figure 24:** Working Principle of BJT

### BJT Configuration

A bipolar junction transistor includes a three-terminal device so it can be connected to a circuit in three possible ways through one terminal being common between others which means one terminal in between input as well as output is common. Every connection simply responds in a different way to the input signal.

#### Common Base Configuration

In the CB configuration of the transistor, the base terminal is common among the input & output signals. Here, the input signal can be given among the two terminals like base as well as emitter whereas the output can be obtained among the two terminals like base & collector. At the collector terminal, the output signal is low as compared to the input signal at the emitter terminal.

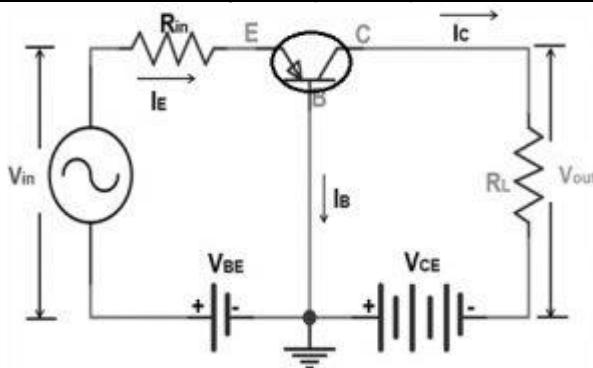


Figure 25: CB Configuration

Thus, its gain is less than 1, so, it attenuates the signal. This output of this configuration is non-inverting, so both inputs, as well as output signals, will be in phase. This common base configuration is not commonly used due to its high voltage gain. Because of its extremely high-frequency response, this kind of configuration is used only for 1- phase amplifier. These 1-phase amplifiers can be used like RF amplifiers and microphone pre-amplifier.

#### Gains of CB Configuration

$$\text{Voltage Gain} = A_v = V_{\text{out}}/V_{\text{in}} = I_C \times R_L / I_E \times R_{\text{in}}$$

$$\text{Current Gain} = I_C/I_E$$

$$\text{Resistance Gain} = R_L/R_{\text{in}}$$

#### Common Emitter Configuration

In this configuration, the emitter terminal is common among input & output. The input is given among the two terminals like base and emitter whereas the output can be obtained among collector and emitter terminals.

It can be simply identified through observing the circuit. When the emitter terminal is connected to GND then input & output are obtained from the base & collector terminals respectively. The CE configuration includes the maximum current & power gain between all types of configurations.

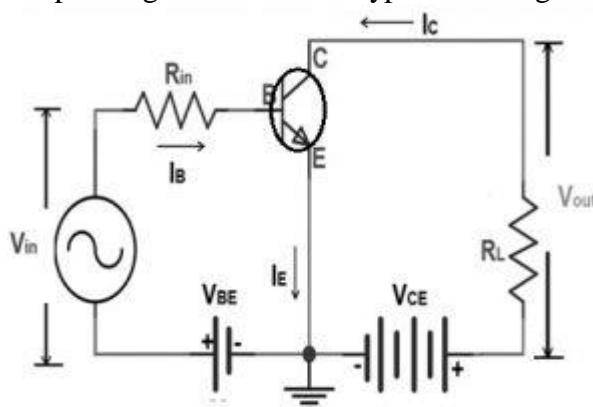


Figure 26: CE Configuration

The main reason is because of the input at the junction of forwarding bias, so its input impedance is extremely less whereas the output can be received from reverse bias junction. Thus, its output impedance



is extremely high. In CE configuration, the emitter current is equivalent to the amount of base & collector currents. So the equation is

$$I_E = I_C + I_B$$

From the above equation, ' $I_E$ ' is the emitter current. So, this configuration has maximum current gain like  $I_C/I_B$  because the load resistance can be connected within series through the collector terminal.

It can be observed from the above equation that a small increase within the base current will effect in very high current at the output region.

This CE configuration works like an inverting amplifier wherever the output signal is totally opposite within polarity toward the input signal. Thus, the output signal can be shifted at  $180^\circ$  with respect to its input signal.

Common collector configuration is missing