Chapter 4

Semiconductor Physics

Semiconductors are materials having electrical conductivity considerably greater than that of an insulator but significantly lower than that of a conductor. Of all the elements in the periodic table, eleven elements are semiconductors. Germanium and silicon are the most widely used semiconductors in device manufacturing applications. They are known as elemental semiconductors. Besides these, there are certain compound semiconductors such as Gallium Arsenide, Indium Phosphide etc which are formed from the combinations of the elements of group III and IV. The unique and interesting feature of semiconductor is that they are polar and the two charge carriers, namely electrons and holes, transport current in these materials.

Semiconductor has electrical resistance which is lesser than an insulator but more than that of a conductor. Its electrical resistivity is in the order of 10^{-4} to 0.5 ohm metre.

A semiconductor has nearly an empty conduction band and almost filled valence band with a very small energy gap ($\approx$ 1eV)

When a suitable impurity is added to a pure semiconductor, its electrical conductivity changes. Based on this property, the semiconductors are classified into two types They are

- Intrinsic semiconductor or pure semiconductor

- Extrinsic semiconductor or doped semiconductor

4.1 Intrinsic Semiconductor

An *intrinsic semiconductor* is a pure form of semiconductor material without any significant impurities or dopants. Its electrical conductivity lies between that of conductors and insulators. It is changed only by thermal excitation.

The common examples for intrinsic semiconductors are pure silicon (Si) and germanium (Ge). They belong to fourth group elements in the periodic table. They are tetravalent atoms since they have four valence electrons. The neighbouring atoms form covalent bonds by sharing four electrons with each other so as to form a stable structure.

The Figure 4.1 shows a two dimensional crystal structure of Silicon and energy band representation of the semiconductor at very low temperature.

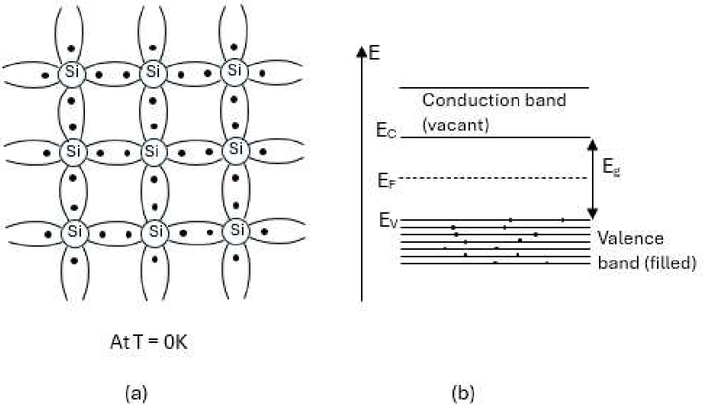


Figure 4.1: (a)Silicon crystal at 0K (b) Corresponding energy band diagram

4.1.1 Carrier Concentration

With an increase in temperature, covalent bonds are broken in an intrinsic semiconductor and electron - hole pairs are generated. Electrons and holes are charged particles. They are collectively called as charge carriers. The number of electrons in the conduction band per unit volume n and the number of holes in the valence band per unit volume p of the material is known as carrier concentration. Carrier concentration is also known as density of charge carriers.

4.1.2 Density of Electrons in Conduction Band

Let dn be the number of electrons whose energy lies in the energy interval E and $E + dE$ in the conduction band. Then

$$dn = Z(E)f(E)dE \tag{4.1}$$

where $Z(E)dE$ is the density of states in the energy interval E and $E+dE$ and $f(E)$ is the probability that a state of energy is occupied by an electron.

The electron density in the conduction band is given by integrating equation (4.1) between the limits E_c and ∞.

E_c is the energy corresponding to the bottom edge of the conduction band and ∞ is the energy corresponding to the top edge of the conduction band.

As the probability of electrons occupying upper levels of conduction band $f(E)$ readily approaches zero for higher energies, the upper limit namely the top of the conduction band is taken as ∞ . Thus

$$n = \int_{E_c}^{\infty} dn = \int_{E_c}^{\infty} Z(E)f(E)dE \tag{4.2}$$

The density of states in conduction band between the energy range E and $E + dE$ is given by,

$$Z(E)dE = \frac{4\pi}{h^3}(2m_e^*)^{3/2}E^{1/2}dE \qquad \text{for } E > E_c \tag{4.3}$$

The bottom edge of conduction band E_c corresponds to the potential energy of an electron at rest.

Therefore, $(E - E_c)$ will be the kinetic energy of conduction electrons at higher energy levels.

$$\text{Hence,} \quad Z(E)dE = \frac{4\pi}{h^3}\,(2m_e^*)^{3/2}\,(E - E_c)^{1/2}dE \tag{4.4}$$

The probability of an electron occupying an energy level is given by Fermi distribution function.

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

When the number of particles is very small compared to the available energy levels, the probability of an energy state being occupied by more than one electron is very small. Such a situation is valid when $(E - E_f)$ » 3kT.

Under this circumstance, the number of available states in the conduction band is far larger than the number of electrons in the band. Then Fermi - Dirac function can be approximated to Boltzmann function.

$$f(E) = e^{-\left(\frac{E-E_F}{kT}\right)} \tag{4.5}$$

Using the equations (4.4) and equation (4.5) in equation (4.2), we have

$$n = \frac{4\pi}{h^3}(2m_e^*)^{3/2}\int_{E_c}^{\infty}(E - E_c)^{1/2}e^{-(E-E_F)/kT}dE \tag{4.6}$$

$$\text{Since, } e^{-(E-E_F)/kT} = e^{(E_F-E_c)/kT} \times e^{-(E-E_c)/kT}$$

We have,

$$n = \frac{4\pi}{h^3}(2m_e^*)^{3/2}e^{(E_F-E_c)/kT}\int_{E_c}^{\infty}(E - E_c)^{1/2}e^{-(E-E_c)/kT}dE \tag{4.7}$$

The integral in equation (4.7) is of the standard form which has a solution of the following form

$$\int_0^{\infty} x^{1/2}e^{-ax}\,dx = \frac{\sqrt{\pi}}{2a\sqrt{a}}$$

$$\text{Here} \quad a = \frac{1}{kT} \quad \text{and} \quad x = (E - E_c).$$

$$\text{Then} \quad n = \frac{4\pi}{h^3}(2m_e^*)^{3/2}e^{(E_F-E_c)/kT}\left[\frac{\sqrt{\pi}(kT)^{3/2}}{2}\right] \tag{4.8}$$

$$\text{Rearranging the terms, we get } n = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} e^{-(E_c-E_F)/kT} \tag{4.9}$$

This is the expression for the *electron concentration* in the conduction band of an intrinsic semiconductor.

In equation (4.9), designating $N_c = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2}$

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

Here N_c is temperature dependent material constant known as the *effective density of states* in conduction band. In Si, at 300K, $N_c = 2.8 \times 10^{25}/m^3$.

4.1.3 Density of Holes in Valence Band

We know that if an electron is transferred from valence band to conduction band, a hole is created in valence band.

Let dp be the number of holes whose energy lies between E and $E+dE$ in valence band.

$$dp = Z(E)[1 - f(E)]dE \tag{4.10}$$

where $Z(E)dE$ is the density of states in the energy interval E and $E+dE$; $[1-f(E)]$ is the probability that a state of energy is vacant and not occupied by an electron.

If $f(E)$ is the probability for occupancy of an energy state at E by an electron, then the probability that the energy state is vacant is $[1 - f(E)]$.

Since a hole represents a vacant state in valence band, the probability for occupancy of a state at E by a hole is equal to the probability of absence of electrons at that state.

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

$$\text{Hence } [1 - f(E)] \approx e^{-(E_F-E)/kT} \tag{4.11}$$

The density of states in the valence band is given by

$$Z(E)dE = \frac{4\pi}{h^3}(2m_h^*)^{3/2}E^{1/2}\,dE \tag{4.12}$$

The top edge of the valence band E_v corresponds to the potential energy of a hole at rest. Therefore, $(E_v - E)$ will be the kinetic energy of the hole at lower energy levels. Hence the equation (4.12) is to be modified as follows.

$$Z(E)dE = \frac{4\pi}{h^3}(2m_h^*)^{3/2}(E_v - E)^{1/2}dE \tag{4.13}$$

Therefore, the number of holes in the energy interval E and $E + dE$ is

$$dp = \frac{4\pi}{h^3}(2m_h^*)^{3/2}(E_v - E)^{1/2}e^{-(E_F-E)/kT}\,dE \tag{4.14}$$

In order to calculate the number of holes in the valence band, equation (4.14) is to be integrated between the limits $-\infty$ and E_v.

The hole density in the valence band is therefore given by

$$p = \frac{4\pi}{h^3}(2m_h^*)^{3/2}\int_{-\infty}^{E_v}(E_v - E)^{1/2}e^{-(E_F-E)/kT}\,dE \tag{4.15}$$

$$= \frac{4\pi}{h^3}(2m_h^*)^{3/2}e^{-(E_F-E_v)/kT}\int_{-\infty}^{E_v}(E_v - E)^{1/2}e^{-(E_v-E)/kT}dE \tag{4.16}$$

The integral equation (4.16) is of the standard form which has a solution of the following form.

$$\int_0^{\infty} x^{1/2}e^{-ax}dx = \frac{\sqrt{\pi}}{2a\sqrt{a}}$$

$$\text{Here} \quad a = \frac{1}{kT} \quad \text{and} \quad x = (E_v - E)$$

Therefore,

$$p = \frac{4\pi}{h^3}(2m_h^*)^{3/2}e^{-(E_F-E_v)/kT}\left[\frac{\sqrt{\pi}(kT)^{3/2}}{2}\right]$$

Rearranging the term, we get,

$$p = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2}e^{-(E_F-E_v)/kT}$$

The above equation is the expression for the *hole concentration in the valence band* of an intrinsic semiconductor.

$$\text{Designating,} \quad N_v = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2} \tag{4.17}$$

$$\text{We get} \quad p = N_v e^{-(E_F-E_v)/kT} \tag{4.18}$$

Here N_v is the *effective density of states in the valence band*. For silicon at 300K, $N_v = 10^{25}/m^3$.

It is seen that $N_c/N_v = 2.8$.

4.1.4 Intrinsic Carrier Concentration

A single event of bond breaking in a pure semiconductor leads to generation of an electron-hole pair. At any temperature T, the number of electrons generated will be equal to the number of holes generated.

As the two charge carrier concentrations are equal, they are denoted by the common symbol n_i which is called intrinsic density or intrinsic concentration .

Thus

$$n = p = n_i$$

$$n_i^2 = np \tag{4.19}$$

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

$$n_i^2 = N_c N_v e^{-(E_c - E_v)kT} \tag{4.20}$$

The term $(E_c - E_v)$ stands for the difference in energy between the top level of valence band and the bottom level of conduction band.

Thus, it represents the separation between the valence band and conduction bands, which is the band gap E_g.

$$(E_c - E_v) = E_g$$

Therefore,

$$n_i^2 = N_c N_v e^{-E_g/kT} \tag{4.21}$$

Substituting the values of N_c and N_v into the above equation, we obtain

$$n_i^2 = 4\left[\frac{2\pi kT}{h^2}\right]^3 \left(m_e^* m_h^*\right)^{3/2} e^{-E_g/kT}$$

$$n_i = 2\left[\frac{2\pi kT}{h^2}\right]^{3/2} \left(m_e^* m_h^*\right)^{3/4} e^{-E_g/2kT} \tag{4.22}$$

This is the expression for intrinsic carrier concentration.

4.1.5 Variation of Carrier Concentration with Temperature

The expression for intrinsic carrier concentration is given as

$$n_i = 2\left[\frac{2\pi kT}{h^2}\right]^{3/2} \left(m_e^* m_h^*\right)^{3/4} e^{-E_g/2kT}.$$

This is written as

$$n_i = AT^{3/2}e^{-E_g/2kT} \approx 10^{21.7}T^{3/2}e^{-2500E_g/T}$$

where A is a material dependent constant.

This shows that the intrinsic carrier concentration in a semiconductor varies with temperature due to the increase in thermal energy which affects the generation of electron-hole pairs.

The following important points may be inferred form the above equation.

(i) The intrinsic concentration is independent of Fermi level.

(ii) The intrinsic concentration has an exponential dependence on the band gap value E_g. As the band gap increases, the intrinsic carrier concentration decreases exponentially, meaning that materials with larger band gaps are generally less conductive at a given temperature. Materials with a smaller bandgap will have a higher intrinsic carrier concentration at a given temperature compared to those with a larger band gap.

(iii) It strongly depends on the temperature. The $T^{3/2}$ term comes from the temperature dependence of the effective density of states N_c and N_v.

(iv) The factor of 2 in the exponent of the intrinsic carrier concentration equation indicates that breaking one covalent bond in a semiconductor generates two charge carriers: one electron and one hole (the energy required per carrier is $\frac{E_g}{2}$).

4.1.6 Fermi Level in Intrinsic Semiconductor

In a pure semiconductor, the electrons in the conduction band gather close to the bottom edge of the band. Similarly, we assume that the holes are at the top edge of the valence band.

The electron concentration in the conduction band is given by

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

The hole concentration in the valence band is given by

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

In intrinsic semiconductor, the electron and hole concentrations are equal.

Thus $n = p$.

$$N_c e^{-(E_c - E_F)/kT} = N_v e^{-(E_F - E_v)/kT} \tag{4.23}$$

Taking logarithm on both sides, we get

$$\begin{aligned} lnN_c \times \frac{-(E_c - E_F)}{kT} &= lnN_v \times \frac{-(E_F - E_v)}{kT} \\ \frac{-(E_c - E_F)}{kT} &= ln\frac{N_v}{N_c} \times \frac{-(E_F - E_v)}{kT} \\ -E_c + E_F &= kT\ ln\frac{N_v}{N_c} - E_F + E_v \\ 2E_F &= (E_c + E_v) + kTln\frac{N_v}{N_c} \end{aligned}$$

$$E_F = \frac{E_c + E_v}{2} + \frac{1}{2}kT\ ln\frac{N_v}{N_c} \tag{4.24}$$

$$\text{But} \quad N_c = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} \quad \text{and} \quad N_v = 2\left[\frac{2\pi m_h^* kT}{h^2}\right]^{3/2}$$

$$\frac{N_v}{N_c} = \left(\frac{m_h^*}{m_e^*}\right)^{3/2}$$

$$ln\left(\frac{N_v}{N_c}\right) = \frac{3}{2}\ ln\left(\frac{m_h^*}{m_e^*}\right) \tag{4.25}$$

Substituting equation (4.25) in equation (4.24),

$$E_F = \frac{E_c + E_v}{2} + \frac{3}{4}kT\ ln\left(\frac{m_h^*}{m_e^*}\right) \tag{4.26}$$

We can also write the above equation as

$$E_F = \frac{E_c + E_v}{2} - \frac{3}{4}kT\ ln\left(\frac{m_e^*}{m_h^*}\right) \tag{4.27}$$

If the effective mass of the free electron is assumed to be equal to the effective mass of a hole, then , $m_h^* = m_e^*$

$$ln\left(\frac{m_h^*}{m_e^*}\right) = 0$$

$$E_F = \frac{E_c + E_v}{2} \tag{4.28}$$

To make the meaning of the equation (4.28) more explicit, we write,

$$E_F = \frac{E_c - E_v}{2} + E_v$$
$$E_F = \frac{E_g}{2} + E_v$$

If we denote the top of the valence band E_v as zero level, E_v = 0.

$$E_F = \frac{E_g}{2} \tag{4.29}$$

This equation shows that in an intrinsic semiconductor, the Fermi level lies in the middle of the forbidden gap as shown in the figure (4.2).

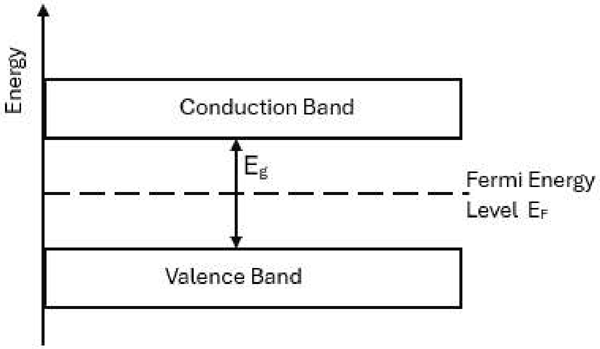


Figure 4.2: Fermi level in intrinsic semiconductor

The Fermi level in an intrinsic semiconductor may be considered as independent of temperature.

4.2 Extrinsic Semiconductor

The conductivity of semiconductor can be improved by adding a small amount of suitable impurity. The impurity semiconductor so obtained are called *extrinsic semiconductor* or doped semiconductor.

The process of adding impurity atoms to an intrinsic semiconductor to improve its conductivity is called *doping*. The impurity atoms that are added to the semiconductors are called dopants. The size of the impurity atoms should be nearly same as that of the semiconductor atoms.

The process of doping increases either the number of free electrons or holes in a semiconductor. Generally, two types of impurities are available.

(i) Impurity atoms having five valence electrons known as pentavalent impurity atoms like Arsenic (As) , Antimony (Sb), Phosphorous (P) etc.

(ii) Impurity atoms having three valence electrons known as trivalent impurity atoms, like Indium (In), Boron (B), Aluminium (Al) etc

Depending on the type of impurities used, extrinsic semiconductors are classified into

1. n-type semiconductor
2. p-type semiconductor

4.2.1 n-type Semiconductor

Semiconductors formed by doping pentavalent impurity atoms are called*n-type semiconductor*. When semiconductors like Si or Ge is doped with a pentavalent impurity such as phosphorous, four valence electrons of the pentavalent atom share electrons with four neighbouring Si or Ge atoms. The fifth valence electrons remains very weakly bound to its parent pentavalent atom and when detached from the parent atom, will be available as a carrier of the current. Thus each pentavalent impurity donates one free electron to semiconductor. At room temperature, each pentavalent impurity atom donates one electron to the conduction band. Hence they are called *donar impurities*.

After donating electron, the impurity atom becomes positively charged ion (*donar ion*) which is fixed in the crystal lattice . Current conduction in n-type semiconductor is due to the movement of electrons and holes.

Electrons in the n-type semiconductors are considered as majority carriers and holes as minority carries.

A n-type semiconductor with its energy band diagram is shown in Figure 4.3.

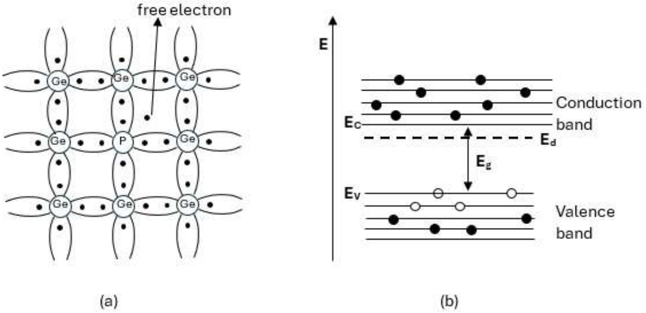


Figure 4.3: (a) n-type semiconductor (b) Energy band diagram of n-type semiconductor at T>0K

4.2.2 p-type Semiconductor

A p-type semiconductor is produced when a pure semiconductor is doped with a trivalent impurity such as boron. Boron atom has three valence electrons. Therefore, it has a shortage of one electron for completing the four covalent bonds with its neighbours. This vacancy of electron constitutes *hole*.

When an electron from a neighbouring atom acquires energy and jumps into the vacancy to form the fourth bond, it leaves behind a hole. Thus the boron atom having acquired an additional electron becomes a negative ion. The hole can move freely in the valence band whereas the impurity ion is fixed in position by the covalent bonds. As the boron atom accepted an electron, it becomes a negatively charged ion and is known as *acceptor ion*. So the acceptor ions do not contribute to the conduction of current in a p-type semiconductor.

Thus in a p-type semiconductor, current flows due to the movement of holes and electrons. But a major part of the current flows due to the movement of holes. Therefore, holes in a p-type semiconductor are majority carriers and electrons are minority carriers.

A n-type semiconductor with its energy band diagram is shown in Figure 4.4.

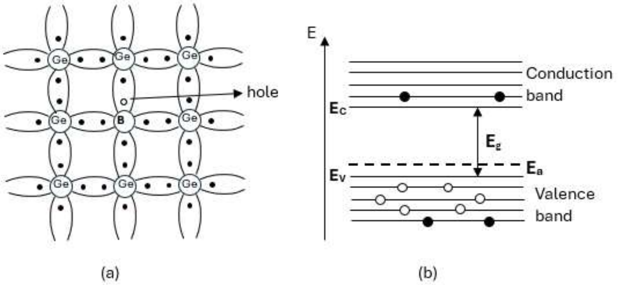


Figure 4.4: (a) p-type semiconductor (b) Energy band diagram of p-type semiconductor at T>0K

SL.No	n-type semiconductor	p-type semiconductor
1	When pentavalent impurity is doped to intrinsic semiconductor, n-type semiconductor is formed	When a trivalent impurity is doped to intrinsic semiconductor, p-type semiconductor is formed
2	n-type impurity is called donor impurity since it donates electron.	p-type impurity is called acceptor impurity since it accepts electron.
3	Majority charge carriers are electrons	Majority charge carriers are holes
4	Minority charge carriers are holes.	Minority charge carriers are electrons.
5	The donor energy level is very close to the bottom of the conduction band.	The acceptor energy level is very close to the top of the valence band.
6	Fermi energy decreases with increase of temperature.	Fermi energy increases with increase of temperature.

Table 4.1: Differences between n-type and p-type semiconductors

4.2.3 Fermi Level in Extrinsic Semiconductor

Fermi Level

Fermi level gives the probability of finding an electron at a given energy value. If Fermi level lies exactly at the middle of the two levels, then the probability of finding an electron is half. That is the case as in an intrinsic semiconductor.

In extrinsic semiconductor, Fermi level strongly depends on temperature as well as the nature of doping and doping concentration.

In an n-type semiconductor, the Fermi level lies in the upper half of the gap E_g, as the majority carriers reside in the conduction band and their average energy is more than E_{Fn}.

In p-type semiconductor, the Fermi level lies in the lower half of the gap, as the majority carriers reside in the valence band and their average energy is less than E_{Fp}.

Variation of Fermi Level with Temperature in an n-type Semiconductor

The donor atom in the n-type semiconductor at low temperature are ionized and supply electrons to the conduction band. Electrons in the conduction band are only from the donor levels.

- At T = 0K, E_{Fn} lies midway between the donor levels E_d and the bottom of the conduction band E_c. It is thus

$$E_{Fn} = \frac{E_c + E_d}{2} \quad \text{at}\, T = 0K \tag{4.30}$$

- As the temperature increases the donor levels gradually get depleted and the Fermi level moves downward. At a temperature T_d, the Fermi level coincides with the donor level E_d. Thus

$$E_{Fn} = E_d \quad \text{at} \quad T = T_d \tag{4.31}$$

- As the temperature grows further above T_d, the Fermi level shifts downward in an approximately linear fashion.

- At a temperature T_i, which is sufficiently high temperature, the intrinsic carrier concentration increases exponentially because more electron-hole pairs are thermally generated. At high enough temperatures, the intrinsic electron concentration can become comparable to or even exceed the extrinsic doping levels, causing the n-type semiconductor to behave more like an intrinsic semiconductor. The Fermi level then moves downward and approaches the Fermi level of intrinsic semiconductor. Thus

$$E_{Fn} = E_{Fi} = \frac{E_g}{2} \quad \text{at} T \geq T_i \tag{4.32}$$

The variation of Fermi level E_{Fn} in an n-type semiconductor with temperature is given in the Figure (4.5).

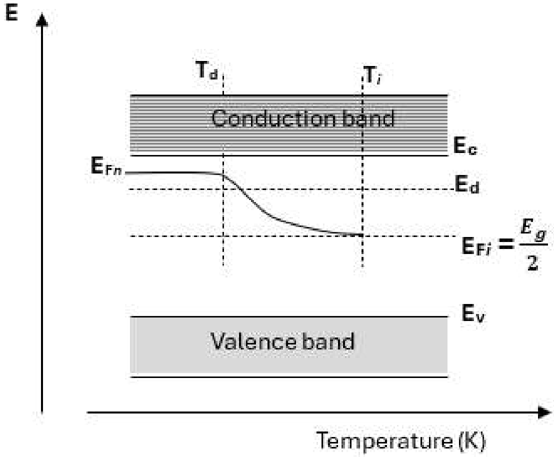


Figure 4.5: Qualitative dependence of Fermi level on temperature in an n-type semiconductor

Variation of Fermi Level with Temperature in an p-type Semiconductor

In case of p-type semiconductor, in the low temperature region, holes in the valence band are only due to the transitions of electrons from the valence band to the acceptor levels. As the valence band is the source of electrons and the acceptor levels are the recipients for them, the Fermi level must lie between the top of the valence band E_v and the impurity acceptor levels E_a.

- At T = 0K, Fermi level lies midway between the acceptor levels E_a and the top of the valence band E_v. Thus,

  $$E_{Fp} = \frac{E_v + E_a}{2} \quad \text{at} \quad T = 0K \tag{4.33}$$

  As the temperature increases, more and more acceptor atoms are ionised. The acceptor levels E_a gradually get filled and the Fermi level E_{Fp} moves upward.

- At the temperature of saturation T_s, the Fermi level E_{Fp} coincides with the acceptor level E_a. Thus

  $$E_{Fp} = E_a \quad \text{at} \quad T = T_s \tag{4.34}$$

- As the temperature grows above T_s, the Fermi level shifts upward in an approximately linear fashion.

- At higher temperature, p-type semiconductor loses its extrinsic character and behaves as an intrinsic semiconductor. Here, the hole concentration in the valence band increases exponentially and the Fermi level approaches the intrinsic value. Thus

  $$E_{Fp} = E_{Fi} = \frac{E_g}{2} \quad \text{at} \quad T = T_i \tag{4.35}$$

The variation of Fermi level E_{Fp} in an p-type semiconductor with temperature is given in the Figure (4.6).

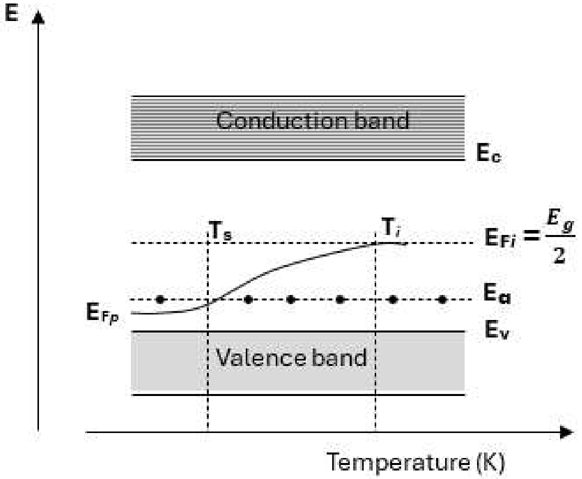


Figure 4.6: Qualitative dependence of Fermi level on temperature in an p-type semiconductor

4.3 p-n Junction

If we join a piece of p-type semiconductor to a piece of n-type semiconductor, a p-n junction is formed. It is the basic building block of many semiconductor devices and is called a *semiconductor diode or p-n juction diode*.

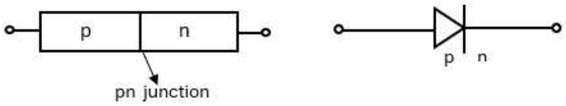


Figure 4.7: p-n junction diode and its symbol

The terminal connected to p-layer is called anode and the terminal connected to n-layer is called cathode. The p-n junction allows the current to flow in one direction and resists it in the opposite direction. The formation and operation of a diode are explained below.

4.3.1 p-n Junction with No External Voltage

A p-n junction when just formed is shown in Fig.4.8. The p-region has holes and negatively charged impurity ions (acceptor ions). The n-region has free electrons and positively charged impurity ions (donor ions). Holes and electrons are the mobile charge carriers while the ions are immobile. There are also minority carriers present which are thermally produced. When temperature increases, more covalent bonds are broken and hence electron-hole pairs are created which constitute minority carriers.

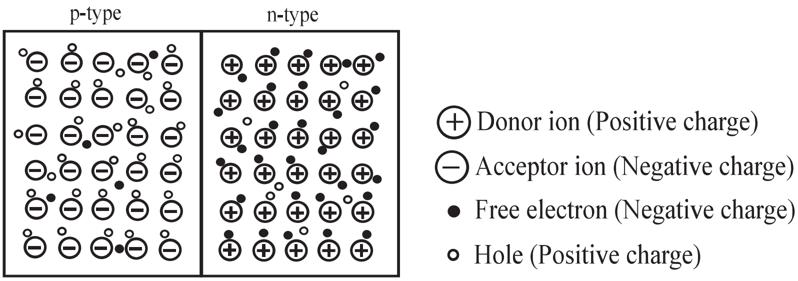


Figure 4.8: A p-n junction just formed

As soon as the p-n junction is formed, the following processes happen.

- Holes from the p-side diffuse into n-side where they combine with free electrons.
- Free electrons from n-side diffuse into p-side where they combine with holes.
- The combining of electrons and holes depletes the holes in the p-region and the electrons in the n-region near the junction. (Each recombination eliminates a hole and a free electron). This region in a p-n junction which contains *only immobile ions* and devoid of free carriers is called *depletion region.*

Since depletion region contains immobile ions which are electrically charged, it is also called the space charge region. This space charge region is as shown in the Figure 4.9. These immobile ions create an electric field that provides a force opposing the further diffusion of charge carriers. The electric field creates a potential difference across the junction which is called space charge potential or barrier potential. This potential is 0.7 V for Si and 0.3 V for Ge at room temperature.

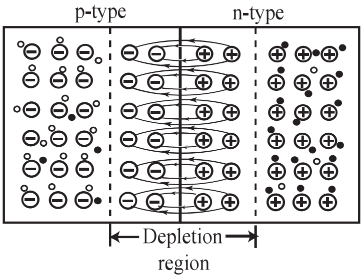


Figure 4.9: Space charge formed at the junction

The barrier potential prevents the further diffusion of majority carriers through the junction but helps the minority carriers (a few electrons in p-region and a few holes in n-region) to drift across the junction. We know minority carriers are generated due to thermal energy. But the drift of minority carriers across the junction is counterbalanced by the diffusion of the same number of majority carriers across the junction. These few majority carriers have sufficiently high kinetic energy to overcome the barrier and cross the junction.

Thus in a p-n junction without external bias, a barrier voltage is developed across the junction and no current flows through the junction.

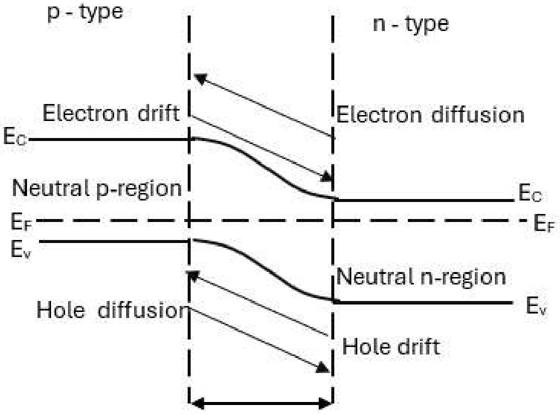


Figure 4.10: Energy band diagram of pn junction

4.3.2 Forward Biased p-n Junction

Consider a voltage source connected across a p-n junction such that the positive terminal of a voltage source is connected to the p-region and negative terminal to the n-region. In this condition the p-n junction is said to be forward biased. Figure 4.11 shows a forward biased p-n junction.

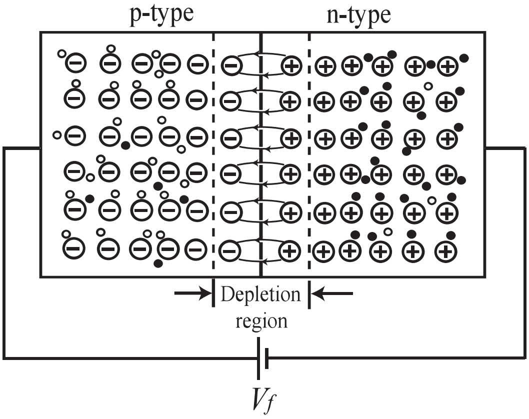


Figure 4.11: Forward biased p-n junction

When a p-n junction is forward biased, the holes which are repelled from the positive terminal of the source move towards the junction. Similarly the electrons which are repelled from the negative terminal of the battery move towards the junction. Because of the acquired energy from the voltage source, some of the holes and electrons penetrate the depletion region and recombine themselves. This reduces the width of depletion region as well as the barrier potential. As a result, more majority carriers diffuse across the junction and hence a large current flows through the p-n junction.

Thus in a forward biased p-n junction

- The width of depletion region is reduced.
- The potential barrier is reduced and at some forward voltage, it is eliminated completely.
- The junction offers low resistance to the current flow. The magnitude of current depends on the applied voltage.

4.3.3 Reverse Biased p-n Junction

If a voltage source is connected across a p-n junction such that the positive terminal of the voltage source is connected to the n-region and negative to the p-region, then the p-n junction is said to be reverse biased. Figure 4.12 shows a reverse biased p-n junction.

When a p-n junction is reverse biased, the holes in the p-region are attracted towards the negative terminal of the voltage source. Similarly, the electrons in the n-region are attracted towards the positive terminal. Thus the majority carriers move away from the junction. This widens the depletion region and increases the barrier potential. The increased barrier potential prevents the diffusion of majority carriers through the junction. Thus there is no current due to majority carriers through a reverse biased p-n junction. The junction offers a very high resistance under reverse biased condition.

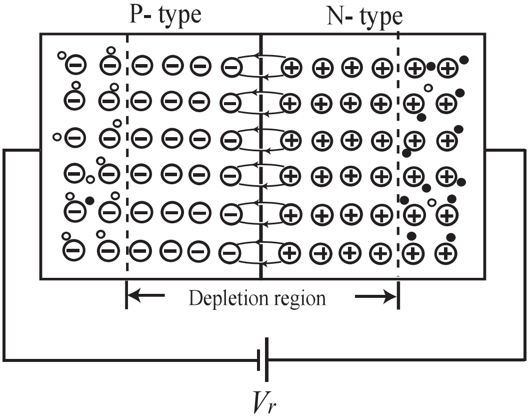


Figure 4.12: Reverse biased p-n junction

The barrier potential helps the minority carriers to cross the junction. Hence a small amount of current flows through a reverse biased p-n junction due to these minority carriers. The current due to minority carriers which is flowing through a reverse biased p-n junction is called *reverse saturation current*. This current is very small since the number of minority carriers is less. It is of the order of nano-amperes in silicon diodes and micro-amperes in germanium diodes. Reverse saturation current increases with increase in temperature.

Thus in a reverse biased p-n junction

- The width of depletion region is increased.
- The potential barrier is increased and the junction offers high resistance.
- No current flows due to majority carriers.
- A small current due to minority carriers known as reverse saturation current flows through the junction.

4.3.4 V-I characteristics of a p-n Junction Diode

The performance of a p-n junction diode when connected in an electric circuit is obtained from its V-I (volt-ampere) characteristics. It is a graph between the voltage applied across the terminals of the p-n junction and the current flowing through it.

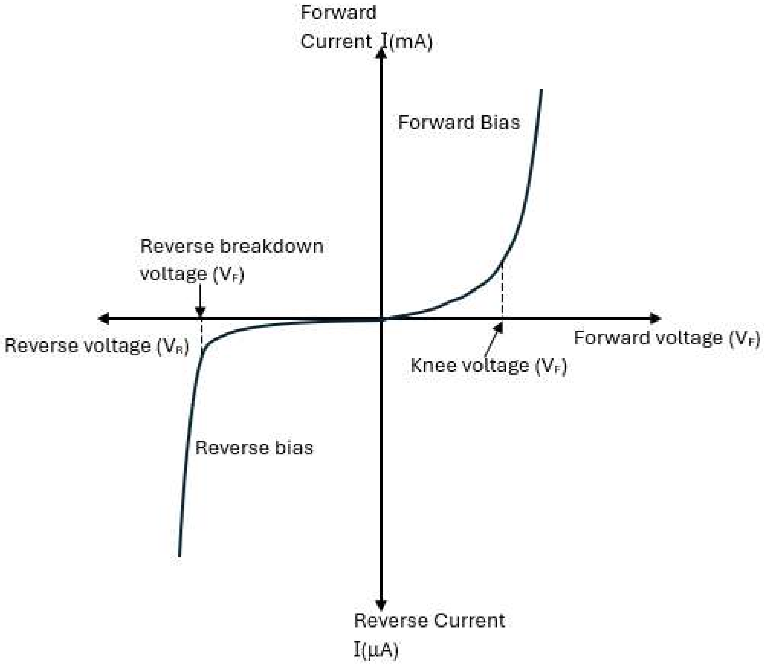


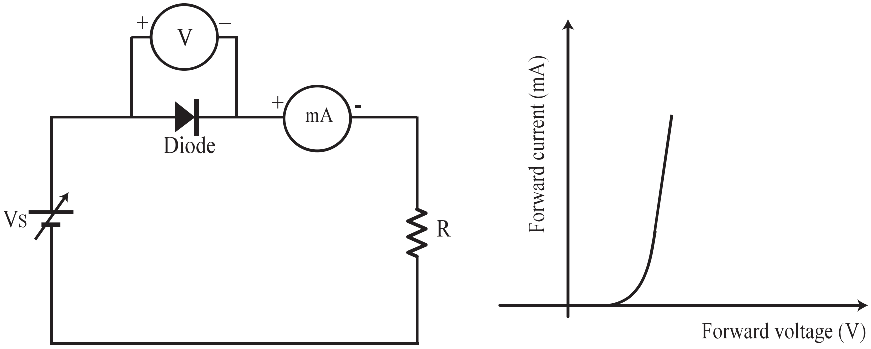
Figure 4.13: V-I characteristics of diode

Figure 4.13 shows the V- I characteristics of a typical p-n junction. It is clear that the graph has two parts namely forward characteristics and reverse characteristics.

4.3.5 Forward Characteristics

Figure 4.14(a) shows the circuit arrangement for obtaining the forward characteristics of a diode. The diode is connected to a variable dc voltage source V_S and a resistor R. The purpose of resistor R is to limit the current through the diode so that the diode will not be damaged due to the flow of excessive current. A voltmeter is connected across the diode to measure the voltage, while a milliammeter measures the current in the circuit.

The positive terminal of the voltage source is connected to the anode (p-region) of the diode and negative terminal to the cathode (n-region). Thus the diode is forward biased. The voltage is gradually increased in small steps (say 0.1V) and the corresponding values of diode current is noted. Now if we plot a graph with voltage along the horizontal axis and diode current along the vertical axis, a curve as shown in Figure 4.14(b) is obtained. It is clear that the diode current is very small for a small applied voltage. The diode does not conduct until the external voltage overcomes the barrier potential (0.7 V for silicon and 0.3 V for germanium at room temperature).



(a) Circuit to obtain forward characteristics

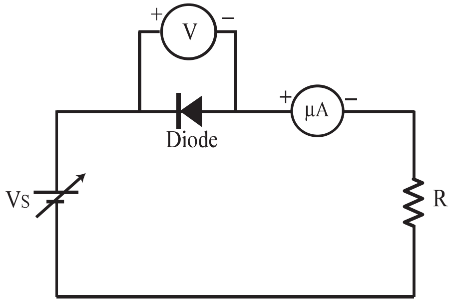
(b) Forward characteristics

Figure 4.14: Circuit and forward characteristics

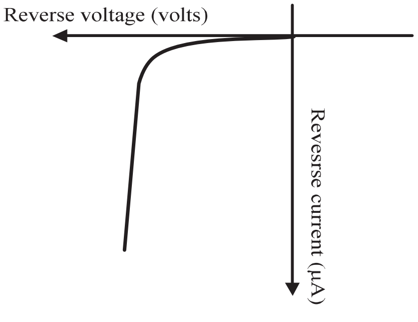
As the applied voltage exceeds this barrier potential, larger number of free electrons and holes start crossing the junction. Above the barrier potential even a small increase in the voltage produces a sharp increase in current. *The voltage at which the current starts to increase rapidly is called the cut-in or knee voltage* (V_k) *of the diode.*

4.3.6 Reverse Characteristics

The circuit arrangement for obtaining the reverse characteristics of a diode is shown in Figure 4.15(a). The negative terminal of the voltage source is connected to the anode of the diode and positive terminal to the cathode to reverse bias the diode. Figure 4.15(b) shows the reverse characteristics of a diode. When the applied reverse voltage is below a certain voltage (called breakdown voltage), the diode current is very small and remains constant.



(a) Circuit to obtain reverse characteristics



(b) Reverse characteristics

Figure 4.15: Circuit and reverse characteristic

When the reverse voltage is increased to a sufficiently large value, the diode reverse current increases rapidly as shown in the Figure 4.15(b). The applied reverse voltage at which sharp increase in current takes place is called breakdown voltage.

Reverse breakdown

As discussed earlier, the current through a reverse biased p-n junction is very small. If the applied reverse voltage is made too large, the increased electric field causes an increase in velocities of minority carriers. These high energy carriers break covalent bonds, thereby generating more carriers. These generated carriers are also accelerated by the electric field. They break more covalent bonds during their movement. Thus a chain reaction is established creating a large number of carriers which give rise to a high reverse current. This phenomenon is called *avalanche breakdown*. There is another type of breakdown known as zener breakdown which will be discussed later.

4.4 Diode Equation

When the diode is forward biased, the potential barrier is lowered by an amount of energy eV_F.

The probability of a majority carrier crossing the junction is increased by a factor $e^{eV_F/kT}$.

Therefore, the diffusion current density decreases by a factor $e^{eV_F/kT}$.

Thus the diffusion current density components J^*_{hp} and J^*_{en} in a forward biased diode are given by

$$J^*_{hp} = J_{hp}e^{eV_F/kT} = J_{hn}e^{eV_F/kT}$$

.

Also,

$$J^*_{en} = J_{en}e^{eV_F/kT} = J_{ep}e^{eV_F/kT}$$

where J_{hp} and J_{en} are diffusion current densities in unbiased diode.

The drift current density components have not changed and have the same magnitude as in equilibrium case. Therefore, the net hole current density across the forward biased junction is

$$J_h = J^*_{hp} - J_{hn} = J_{hn}(e^{eV_F/kT} - 1)$$

Similarly, the net electron current density across the junction is

$$J_e = J^*_{en} - J_{ep} = J_{ep}(e^{eV_F/kT} - 1)$$

The total current density across the forward biased p-n junction is a sum of electrons and hole current density components. i.e.,

$$
\begin{aligned}
J &= J_e + J_h \\
&= (J_{hn} + J_{ep})(e^{eV_F/kT} - 1)
\end{aligned}
$$

$$J = J_0(e^{eV_F/kT} - 1) \quad \text{where} \quad J_0 = (J_{hn} + J_{ep}) \tag{4.36}$$

If the area of cross section of the junction is A, then the current is $I = JA$. Therefore,

$$I = I_0(e^{eV_F/kT} - 1) \quad \text{where} \quad I_0 = J_0 A$$

When the diode is reverse-biased, the connection from the voltage source is reversed and a voltage ($-V_R$) in place of V_F in the equation (4.36), we get

$$J = J_0(e^{-eV_R/kT} - 1) \tag{4.37}$$

For larger values of V_R,

$$e^{-eV_R/kT} \rightarrow 0$$

Therefore,

$$J = -J_0$$

or

$$I = -I_0$$

Equation (4.36) and Equation (4.37) can be combined into a single equation by denoting the forward and reverse voltages by a single symbol V. Thus , we write,

$$J = J_0(e^{eV/kT} - 1) \tag{4.38}$$

In terms of current,

$$I = I_0(e^{eV/kT} - 1) \tag{4.39}$$

where $V = V_F$ for forward bias and $V = -V_R$ for reverse bias.
Equation (4.39) is known as rectifier equation or *diode equation.*
The term I_0 is known as *reverse saturation current*

4.5 Solved Numerical Problems

Example 4.5.1 The forbidden gap in pure silicon is 1.1eV. Compare the number of conduction electrons at temperatures 37^oC and 27^oC.

Solution:

Let n_1 be the number of conduction electrons at 37^oC and n_2 at 27^oC.
E_g is 1.1eV.
The expression for intrinsic carrier concentration is

$$\begin{aligned} n_i &= 2\left[\frac{2\pi kT}{h^2}\right]^{3/2}\left(m_e^* m_h^*\right)^{3/4} e^{-E_g/2kT} \\ n_i &= 10^{21.7} T^{3/2} \times 10^{-2500E_g/T} \\ n_1 &= 10^{21.7}(300K)^{3/2} \times 10^{-2500\times 1.1eV/300T} \\ n_2 &= 10^{21.7}(310K)^{3/2} \times 10^{-2500\times 1.1eV/310T} \\ \frac{n_2}{n_1} &= \frac{(310K)^{3/2} \times 10^{-8.87}}{(300K)^{3/2} \times 10^{-9.2}} = 2.96 \end{aligned}$$

Thus we can see that, as the temperature increases from 27^oC to 37^oC, the number of electrons in the conduction band is increased by three times.

Example 4.5.2 Calculate the intrinsic carrier concentration in silicon at 400K and 500K. Given A for silicon is $7.3 \times 10^{15} cm^{-3} K^{-3/2}$. E_g for silicon = 1.12eV.

Solution:

Intrinsic charge carrier concentration ,

$$n_i = AT^{3/2} e^{-E_g/2kT} \approx 10^{21.7} T^{3/2} e^{-2500E_g/T}$$

$$\begin{aligned} \text{At 400K, } n_i &= 7.3 \times 10^{15} \times 400^{3/2} \times exp\left(\frac{-1.12}{2 \times 1.38 \times 10^{-23} \times 400}\right) \\ &= 7.3 \times 10^{15} \times 8000 \times 8.5 \times 10^{-15} \times 7.4 \times 10^{-12} \\ &= 4.94 \times 10^5 cm^{-3} \\ \text{At 500K, } n_i &= 7.3 \times 10^{15} \times 500^{3/2} \times exp\left(\frac{-1.12}{2 \times 1.38 \times 10^{-23} \times 500}\right) \end{aligned}$$

$$= 7.3 \times 10^{15} \times 11180 \times 7.4 \times 10^{-12}$$
$$= 6.07 \times 10^{5} cm^{-3}$$

Example 4.5.3 Compute the concentration of intrinsic charge carriers in a germanium crystal at 300K. Given that E_g = 0.72eV and m_e^* = m_e.

Solution:

Intrinsic charge carrier concentration ,

$$n_i = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} e^{-(E_c - E_F)/kT}$$

$$\text{Since, } E_c - E_F = \frac{E_g}{2}$$

$$n_i = 2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} e^{-E_g/2kT}$$

$$2\left[\frac{2\pi m_e^* kT}{h^2}\right]^{3/2} = 2\left[\frac{2 \times 3.14 \times 9.11 \times 10^{-31} \times 1.38 \times 10^{-23}}{(6.626 \times 10^{-34})^2}\right]^{3/2}$$
$$= 4.83 \times 10^{21}$$

$$T^{3/2} = (300)^{3/2} = 5196$$

$$exp\left(\frac{-E_g}{2kT}\right) = exp\left(\frac{-0.72}{2 \times (\frac{1.38 \times 10^{-23}}{1.6 \times 10^{-19}}) \times 300}\right) = 9.07 \times 10^{-7}$$

$$\text{Therefore, } n_i = 4.83 \times 10^{21} \times 5196 \times 9.07 \times 10^{-7}$$
$$= 2.276 \times 10^{19}/m^3$$

Example 4.5.4 What is the probability of an electron being thermally promoted to the conduction band in diamond at 27^oC, if the bandgap is 5.6eV wide?

Solution:

The probability that an electron being thermally promoted to the conduction band is given by

$$f(E_c) = \frac{1}{1 + exp\left(\dfrac{E_g}{2kT}\right)}$$

$$= \frac{1}{1 + exp\left(\dfrac{5.6eV}{2 \times (\frac{1.38 \times 10^{-23}}{1.6 \times 10^{-19}}) \times 300}\right)}$$

$$= \frac{1}{1 + exp\left(\dfrac{5.6}{2 \times 0.00008625 \times 300}\right)}$$

$$= \frac{1}{1 + e^{108.21}} = 1.011 \times 10^{-47}$$

Example 4.5.5 Estimate the fraction of electrons in conduction band at room temperature in Ge with $E_g = 0.72$eV and in diamond with $E_g = 5.6$eV.
Solution:
Fraction of electrons in conduction band

$$\frac{n}{N} = exp\left[-\frac{E_g}{2kT}\right]$$

For Ge, fraction of electrons in conduction band

$$\frac{n}{N} = exp\left(\frac{5.6eV}{2 \times (\frac{1.38 \times 10^{-23}}{1.6 \times 10^{-19}}) \times 300}\right)$$

$$= exp\left(-\frac{0.72eV}{2 \times 0.00008625 \times 300}\right) = 9.098 \times 10^{-7}$$

For diamond, fraction of electrons in conduction band

$$\frac{n}{N} = exp\left(-\frac{5.6eV}{2 \times 0.00008625 \times 300}\right) = 1.011 \times 10^{-47}$$

Example 4.5.6 If the effective mass of an electron is equal to twice the effective mass of hole, determine the position of Fermi level in an intrinsic semiconductor from the centre of forbidden gap at room temperature.
Solution:

$$\begin{aligned}
E_F &= \frac{E_g}{2} - \frac{3}{4}kT\ ln\left(\frac{m_e^*}{m_h^*}\right) \\
E_F &= \frac{E_g}{2} - \frac{3}{4}kT\ ln2 \\
&= \frac{E_g}{2} - \frac{3}{4}(0.026eV)(0.69) \\
E_F &= \frac{E_g}{2} - 0.0135eV
\end{aligned}$$

The Fermi level is below the centre of the forbidden gab by 0.014eV.

Example 4.5.7 A silicon diode is forward biased with a voltage of 0.5V at 298K. The reverse saturation current is found to be 10nA. Find the diode current.
Solution:
The current flowing through a junction is given by $I = I_0(e^{eV/kT} - 1)$
We have reverse saturation current $I_0 = 10 \times 10^{-9}A$; V = 0.5V ;
$k = 1.38 \times 10^{-23}$; T = 298K

$$I = 10 \times 10^{-9}\left[exp\left(\frac{1.6 \times 10^{-19} \times 0.5}{1.38 \times 10^{-23} \times 298}\right) - 1\right] = 2.78A$$

Example 4.5.8 In a rectifier circuit operating at 27^{o}c, the forward current is 6.63A for a forward voltage of 0.22V. Calculate the reverse saturation current.
Solution:
The current flowing through a junction is given by $I = I_0(e^{eV/kT} - 1)$
We have forward current I = 6.63A; V = 0.22V

$$I_0 = \frac{I}{(e^{eV/kT} - 1)}$$

$$I_0 = \frac{6.63}{\left[exp\left(\frac{1.6\times10^{-19}\times0.22}{1.38\times10^{-23}\times300}\right)-1\right]} = 1.34mA$$

Example 4.5.9 Current flowing in a p-n junction is 0.2μA at room temperature when a large reverse bias voltage is applied. Calculate the current when a forward bias of 0.1V is applied.

Solution:

The current flowing through a junction is given by $I = I_0(e^{eV/kT} - 1)$
We have $I_0 = 2\times10^{-7}A$; V = 0.1V ; $k = 1.38\times10^{-23}$;
T is the room temperature which is 27°C = 300K
Then,

$$I = 2\times10^{-7}\left[exp\left(\frac{1.6\times10^{-19}\times0.1}{1.38\times10^{-23}\times300}\right)-1\right] = 9.34\mu A$$

Example 4.5.10 A silicon diode operates at a temperature of 300 K, with a reverse saturation current $I_0 = 1\times10^{-12}A$. If the temperature is increased to 350 K, calculate the new current for a forward bias of 0.7 V.

Solution:

At T = 300K

$$I = 1\times10^{-12}\left[exp\left(\frac{1.6\times10^{-19}\times0.7}{1.38\times10^{-23}\times300}\right)-1\right] = 0.561A$$

At T = 350K,

$$I = 1\times10^{-12}\left[exp\left(\frac{1.6\times10^{-19}\times0.7}{1.38\times10^{-23}\times350}\right)-1\right] = 0.0117A$$

4.6 Excercise

1. The band gap of silicon is 1.12eV. At room temperature (300 K), the effective density of states for the conduction and valence bands are given as $N_c = 2.8 \times 10^{25}/m^3$ and $N_v = 1.04 \times 10^{25}/m^3$ Calculate the intrinsic carrier concentration.
Hint: $n_i \approx 6.74 \times 10^{15} m^{-3}$

2. Determine the fraction of electrons in conduction band in silicon at 27^oC and 227^oC. Given E_g = 1.1eV and k = 1.38×10^{-23}J/K
Hint: $5.7 \times 10^{-10}; 2.8 \times 10^{-6}$

3. Determine the position of Fermi level in silicon semiconductor at 300K. Given that the band gap is 1.12eV and m_e^*= 0.12m and m_h^* = 0.28m.
Hint: The Fermi level is 0.016eV above the centre of the forbidden gap or 0.576eV from the top of the valence band.

4.A silicon diode operates at a temperature of 300 K, with a reverse saturation current $I_0 = 1 \times 10^{-12} A$. If a forward voltage of 0.7 V is applied, calculate the diode current using the diode equation where $q = 1.6 \times 10^{-19} C$, $k = 1.38 \times 10^{-23} J/K$.
Hint: $I = 0.561 A$