

Module - VII

Semiconductor electronics: Materials, Devices and Simple Circuits

Lectures: 03

Objective:

In this module we will show the evolution of electronics from semiconductors. The students will learn regarding different semiconducting materials, electronic devices and simple circuits.

Pre requisite:

The students may be asked following questions before starting this module.

1. What are the differences between semiconductors and metals?
2. How the resistance changes with temperature in semiconductor?
3. How the resistance changes with temperature in metals?
4. What are the majority and minority charge carriers in n -type and p -type superconductors and why?
5. How are energy bands formed in solids?
6. How do you classify metals, semiconductors and insulators? How does one can understand them via electron energy band diagram?
7. Define an intrinsic and extrinsic semiconductor. What role does donor and acceptor impurities play in forming extrinsic semiconductor?
8. Describe how a p - n junction is formed and its characteristics in forward and reverse biased mode. How one can exploit these properties to use it for rectification of ac voltages?
9. Describe in brief the concept of filtering a time varying voltage signal to obtain a steady dc output voltage.

1. Introduction:

Controlled flow of electrons is the feature required in any device forming the building blocks of the electronic circuits. Prior to the discovery of transistor in 1948 such devices mostly were vacuum tubes containing many electrodes like cathode, anode, grids etc. The applied voltages at each electrode gave a large degree of control over the flow of electrons and hence the desired operation of the device can be ensured. Some examples of such devices are valves, triodes, tetrodes etc. each one with its specific function. These devices mostly require a vacuum inside a tube which houses the various electrodes hence they were bulky in size and difficult to operate comfortably. Moreover they consume high power, need high voltages for their operation (~ 100 eV) and have a limited life. With the discovery of modern solid state based semiconductors in 1930's and its applications in device manufacturing offered significant advantages over vacuum tubes in many of the areas mentioned above. Hence gradually the electronic industry replaced vacuum tubes with semiconductor devices extensively.

The figures and examples in this presentation have been obtained from NCERT Physics textbook for class 12th course of the central board of secondary education (CBSE), New Delhi.

2. Classification of metals, conductors and semiconductors:

2.1 On the basis of conductivity:

Based on values of resistivity ρ and conductivity σ ($=1/\rho$);

(i) *Metals*: Very low resistivity (or high conductivity).

$$\rho \sim 10^{-2} - 10^{-8} \Omega \text{ m}$$

$$\sigma \sim 10^2 - 10^8 \text{ S m}^{-1}$$

(ii) *Semiconductors*: Resistivity or conductivity intermediate to metals and insulators.

$$\rho \sim 10^{-5} - 10^6 \Omega \text{ m}$$

$$\sigma \sim 10^5 - 10^{-6} \text{ S m}^{-1}$$

(iii) *Insulators*: High resistivity (or low conductivity).

$$\rho \sim 10^{11} - 10^{19} \Omega \text{ m}$$

$$\sigma \sim 10^{-11} - 10^{-19} \text{ S m}^{-1}$$

Semiconductors are mainly of two types:

(i) Elemental Semiconductors: e.g. Si and Ge.

(ii) Compound Semiconductors: CdS, GaAs, CdSe, InP, anthracene, doped phthalocyanines, polypyrrole, polyaniline, polythiophene etc.

In this module we will restrict ourselves to elementary semiconductors particularly. The general concepts introduced here for discussing the elemental semiconductors, by-and-large, apply to most of the compound semiconductors as well.

2.2 On the basis of energy bands:

The overlap of electronic orbitals of atoms with each other leads to a continuous distribution of their energies in solids giving rise to band formation. Thus every atomic orbital of the solid becomes a part of some band inside the solid. The valence electrons of atoms form a band known as the valence band of the solid. Then by definition we have the valence band as being completely occupied at 0 K. The lowest unoccupied energy band is called the conduction band. The relative position of valence band and conduction band varies for distinguishing metals, semiconductors and insulators. An electron in a conduction band must be considered as free contributing to electrical conduction.

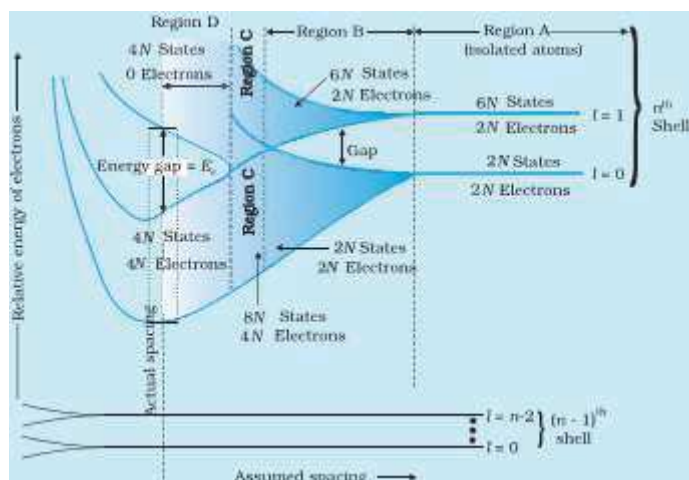


FIG. 1 Illustrating the band formation in solids from atoms coming close to each other.

Consider the band formation in Si or Ge crystal having N atoms. Each atom has a $2s^2 2p^2$ valence electronic configuration hence we have $4N$ valence electrons in the crystal in total. However every atom has 4 p -electronic states

unoccupied hence the crystal has $4N$ unoccupied electronic states too. When the atoms come together during the formation of a solid the valence electronic states of neighbouring atoms overlap with each other giving rise to slight shift of their energies with respect to that inside the atom. It causes both downward and upward shift of their mutual energies. *When the number of atoms N is very large, ideally infinity, then the energy distribution for those electrons becomes continuous leading to the formation of energy bands.* These energy bands are most prominent for the valence electrons since they occupy outermost orbits and hence they strongly interact with each other inside the solid. The core electrons are well shielded from the electrostatic potential of the neighbouring atoms by the outermost electrons which keeps their wave function confined close to the nuclei of the atoms and hence their wave functions' overlap with those of the neighbouring ones is negligible; hence the band formation among them is insignificant. *The valence electrons form valence band and the unoccupied electronic states form the conduction band.*

The formation of the band is nicely illustrated in fig. 1. At large values of the separation between the atoms (region A) its electronic states are mostly influenced by the electrostatic potential of the nucleus of that atom itself and hence the overlap of the valence states is weak giving rise to well defined valence electronic energies corresponding to that atom. As one slowly decrease the separation the overlap of the valence states increases gradually leading to the spread of the electronic energy of the atom (region B). This amounts to the formation of a band from those states. Further decrease in the separation leads to the further downward (upward) shift of the lower (upper) end of the conduction (valence) band such that they both cross each other (region C). With continuous decrease in the separation leads to the reappearance of the energy gap E_g between valence and conduction band (region D).

It must be noted that the no. of occupied (unoccupied) states in a band are same as the no. of occupied (unoccupied) energy states of the atoms from where they were created.

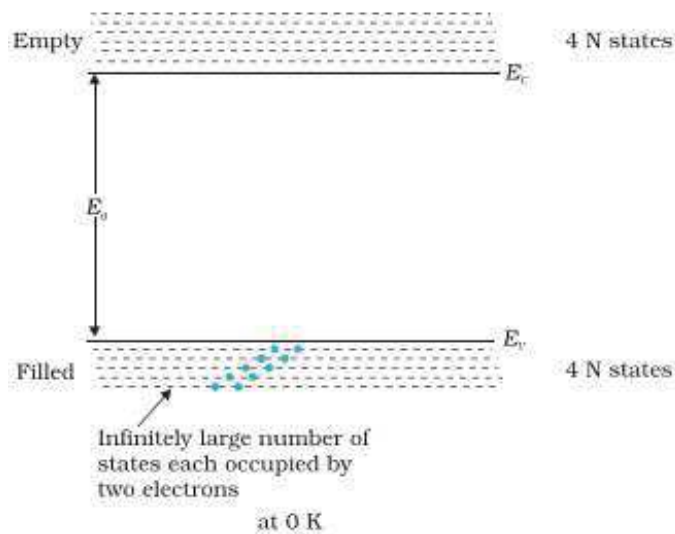


FIG.2 The energy band position in a semiconductor at 0 K. The upper band, called the conduction band, consists of infinitely large number of closely spaced energy states. The lower band, called the valence band, consists of closely spaced completely filled energy states.

The lowest energy level in the conduction band is denoted as E_C and the highest energy level in the valence band is denoted as E_V . Above E_C and below E_V there are a large number of closely spaced energy levels, as shown in fig. 2. The gap between E_V and E_C is called the energy band gap (Energy gap E_g).

It may be large, small, or zero, depending upon the material. These cases are discussed in the following figure.

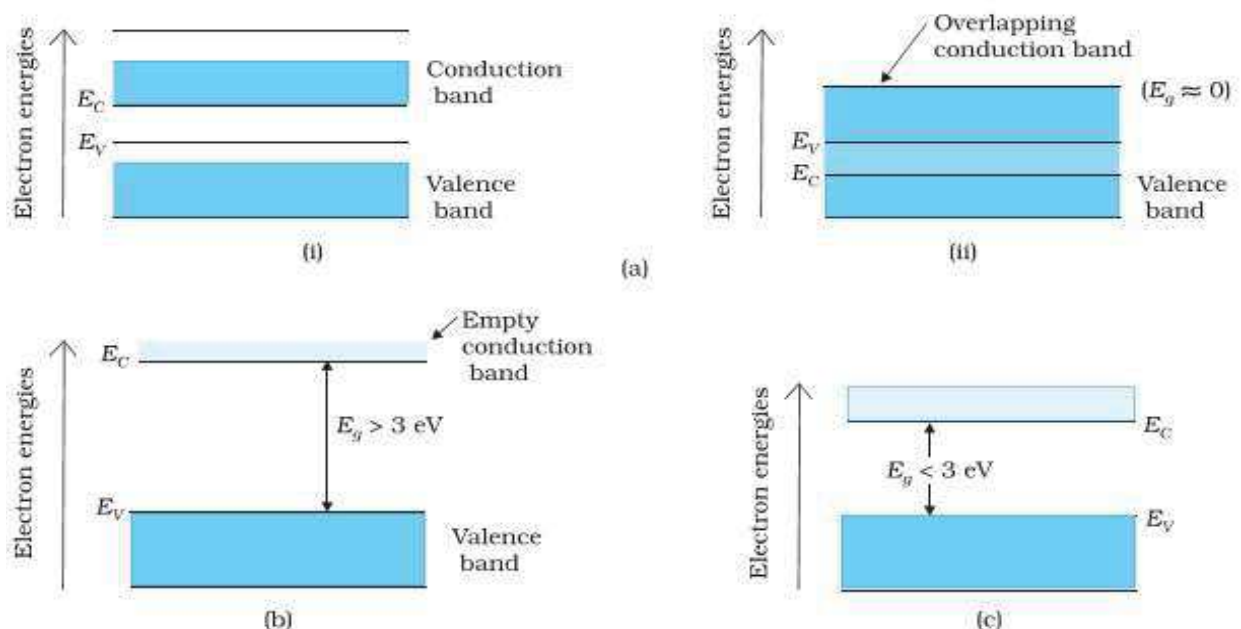


FIG. 3 Difference between energy bands of (a) metals, (b) insulators and (c) semiconductors.

Case I: This refers to a situation, as shown in (a) of fig. 3. Metals form either when the conduction band is partially filled and the valence band is partially empty or when the conduction and valence bands overlap. When one has either of the above situations an extremely small amount of energy given to the sample can easily excite large no. of electrons into the unoccupied states hence such materials easily conduct electricity and they are called as metals. The metals have high conductivity and low resistivity.

Case II: In this case, as shown in (b), a large band gap E_g exists ($E_g > 3 \text{ eV}$). The valence band is completely filled and the conduction band is completely empty. One has to supply energy of the order of 3 eV in order to excite an electron from the valence band to the conduction band in order to generate a free electron. Such a high value of energy cannot be supplied under ambient conditions hence these materials do not conduct electricity and are called as insulators. A natural source of energy for causing excitation of the electrons in solids is the atmospheric temperature which can cause thermal excitations of the electrons, however, at practical temperatures e.g. room temperature $T = 300 \text{ K}$ the available energy to cause thermal excitations is $\sim 25 \text{ meV} \ll E_g$. Thus thermal excitation is inadequate in causing electrical conduction in insulators. The insulators have low conductivity and high resistivity.

Case III: This situation is shown in (c). Here 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 (though small in numbers) can move in the conduction band and contribute to electrical conduction. Hence, the resistance of semiconductors is not as high as that of the insulators but also not low as that of metals. Semiconductors sometimes demonstrate metallic like behavior at high temperatures due to that existence of free electrons at high temperatures.

***Fermi Level**

Top of the collection of electron energy levels at absolute zero temperature. This concept comes from Fermi-Dirac statistics. Electrons are fermions and by the Pauli exclusion principle cannot exist in identical energy states.

Intrinsic Semiconductor:

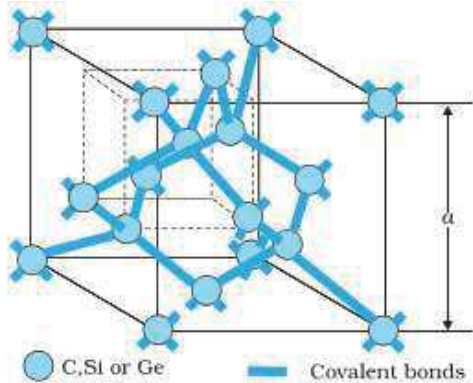


FIG. 4 Three dimensional diamond like crystal structure for Carbon, Silicon or Germanium with respective lattice spacing 'a' equal to 3.56, 5.43 and 5.66 Å.

Common examples of intrinsic semiconductors are elements Si and Ge which have 4 valence electrons in the outermost shell in $2s^2 2p^2$ electronic configuration. From fig. 4 we observe that each atom is coordinated by 4 nearest neighbours forming covalent bonds with each other. In each covalent bond a valence electron from one atom is shared with the corresponding one of the neighbouring atom thus the 4 neighbouring atoms will 'add' 4 additional electrons to the valence shell of the central atom completing its inert shell configuration.

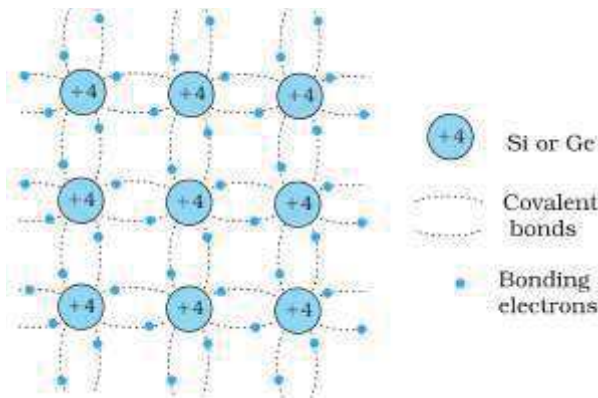


FIG. 5 Schematic two-dimensional representation of Si or Ge structure showing covalent bonds at low temperature (all bonds intact). +4 symbol indicates inner cores of Si or Ge.

A two dimensional schematic representation of the covalent bonding is represented in fig. 5. The covalently bonded electrons are depicted as shuttling between the sharing atoms. The figure shows an idealized scenario of all covalent bonds being intact which is expected only at low temperatures and in an ideal crystal without lattice defects. In real crystals there will be imperfections in the crystalline lattice during crystal growth. Similarly at high temperatures there will be thermal excitations of the electrons which will force some of the electrons to break away from the bonds. However in order to study the effect of intrinsic electronic structure in giving rise to semi-conductivity we imagine an idealized case when we do not have any lattice defects and assume that we have a perfect crystal produced for us. For such a case at low temperatures all the bonds are intact and we do not have any thermal excitation induced breaking of bonds.

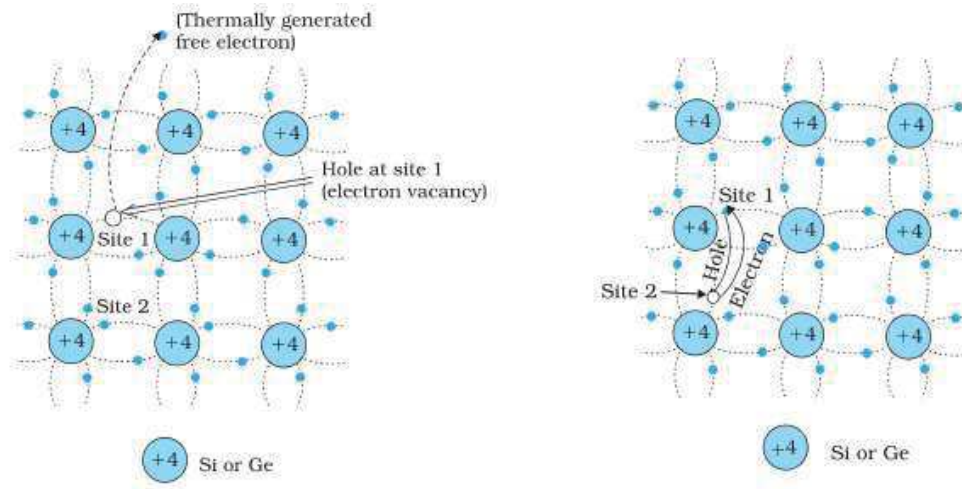


FIG. 6 (a) Schematic model of generation of hole at site 1 and conduction electron due to thermal energy at moderate temperatures. (b) Simplified representation of

possible thermal motion of a hole. The electron from the lower left hand covalent bond (site 2) goes to the earlier hole site 1, leaving a hole at its site indicating an

apparent movement of the hole from site 1 to site 2.

At high temperatures the electrons which have broken these bonds become free and contribute to electrical conduction (fig. 6). When the electron is thermally excited it leaves a vacancy at the site of the bond from which it got excited which is called a 'hole'. A hole is an *effective* positive charge with the magnitude of the charge same as that of the outgoing electron. Under the influence of temperature the hole generated at one site can 'migrate' to another site thus holes inside the semiconductors are mobile. In semiconductors both these charge carriers i.e. the free electrons and the 'holes' contribute to the electrical conduction independently and hence the net electrical conduction in semiconductors is the sum of these individual contributions.

In intrinsic semiconductors, the number of free electrons, n_e is equal to the number of holes, n_h . That is $n_e = n_h = n_i$, where n_i is called intrinsic carrier concentration.

The thermally excited free electron moves completely independently as conduction electron and thus gives rise to an electron current, I_e under an applied electric field. Remember that the motion of hole is only a convenient way of describing the actual motion of *bound* electrons, whenever there is an empty bond anywhere in the crystal. Under the action of an electric field, these holes move towards negative potential giving the hole current, I_h . The total current, I is thus the sum of the electron current I_e and the hole current I_h : i.e. $I = I_e + I_h$.

Example: C, Si and Ge have same lattice structure. Why is C insulator while Si and Ge intrinsic semiconductors?

Solution: The 4 bonding electrons of C, Si or Ge lie, respectively, in the second, third and fourth orbit. Hence, energy required to take out an electron from these atoms (i.e., ionisation energy E) will be least for Ge, followed by Si and highest for C. Hence, number of free electrons for conduction in Ge and Si are significant but negligibly small for C. Note that the band gap of diamond (carbon) is 5.5 eV, of Si is 1.11 eV and of Ge is 0.67 eV

10. Extrinsic Semiconductor:

The intrinsic semiconductors require finite temperatures for its operation. Only then one can get electron-holes available for conduction and semiconducting operation. However technologically it is much more desirable to have the semi conductivity to be realized at all temperatures. Moreover the electrical conductivity of intrinsic semiconductors is quite low for the sake of their technological use. Hence it is desirable to improve the properties of intrinsic semiconductors in order to tailor them for better technological use. One idea in this direction is to use foreign atoms added to intrinsic semiconductor in very small quantities in parts per million (ppm) level (called impurities) which would enhance their physical properties for technological exploitation e.g. one can enhance the electrical conductivity manifold upon such impurity addition. Such materials are known as extrinsic semiconductors (also called doped semiconductors) since the semiconducting properties are primarily due to the impurity atoms than from the parent compound itself. The impurity atoms added are technically called as *dopants*. An important consideration which must go into while choosing a dopant for doping is that the size of the dopant atom should be close to the size of the parent atom so as not to distort the crystal lattice significantly. This will ensure that the dopant merely adds its promised physical properties to the parent compound without generating spurious features arising due to the distortion of the lattice.

Mainly two types of dopants are used in the doping tetravalent Si or Ge:

- (i) **Pentavalent (valency 5):** like Arsenic (As), Antimony (Sb), Phosphorous (P), etc.
- (ii) **Trivalent (valency 3):** like Indium (In), Boron (B), Aluminium (Al), etc.

You will see from above that the dopant atoms lie in either side of Si and Ge in the periodic table of elements. Thus they are expected to have atomic sizes close to Si and Ge hence their substitution in place of Si or Ge is not expected to create significant crystal structure modifications in the parent lattice. Both these substitutions give rise to entirely different types of semi conductivity.

(i) n-type semiconductor:

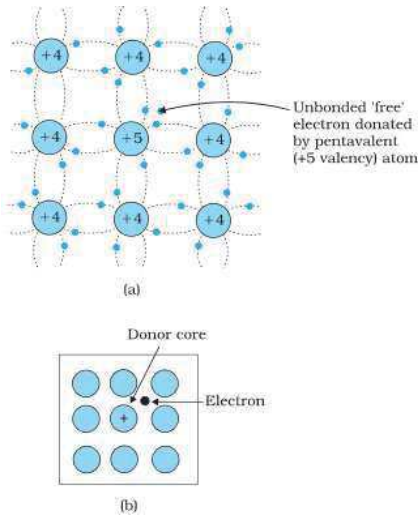


FIG. 7(a) Pentavalent donor atom (As, Sb, P, etc.) doped for tetravalent Si or Ge giving n-type semiconductor, and (b) Commonly used schematic representation of n-type material which shows only the fixed cores of the substituent donors with one additional effective positive charge and its associated extra electron.

If we dope a pentavalent impurity (As, Sb, P) into Si or Ge, then the 4 valence electrons of the pentavalent atom will be involved in the formation of 4 covalent bonds with its neighboring Si or Ge atoms and the fifth valence electron of the dopant will find itself loosely attached to its parent nucleus (see fig.7). This is because the 4 covalent bonds will apparently complete the inert shell configuration for the dopant atom which will screen the positive nuclear charge from attracting the fifth valence electron. Hence that electron will be very loosely bound to the dopant atom and with very small amount of energy it will get excited to the conduction band and contribute to the conduction. E.g for pentavalent impurity in Si this energy ~ 0.05 eV and for Ge ~ 0.01 eV. Since this extra electron is donated by the pentavalent dopant hence it is also called as the donor impurity. Since the extra electron lies very close to the conduction band the free electron concentration in such cases is very weakly temperature dependent.

In such semiconductors the number of free electrons n_e is contributed largely by the no. of dopant atoms and weakly by the intrinsically generated free electrons by thermal excitation. Moreover the since the free electron concentration in doped semiconductors is enhanced in comparison with intrinsic semiconductors the rate of hole recombination also enhanced which further reduces the no. of available holes n_h at any temperature. Hence the electrical conduction in such compounds occurs primarily by the free conduction electrons and less by the holes in the valence band. Thus, with proper level of doping the number of conduction electrons can be made much larger than the number of holes. Hence in an extrinsic semiconductor doped with pentavalent impurity, electrons become the *majority* carriers and holes the *minority* carriers. These semiconductors are, therefore, known as n-type semiconductors. For n-type semiconductors, we have,

$$n_e \gg n_h.$$

(ii) p-type semiconductor:

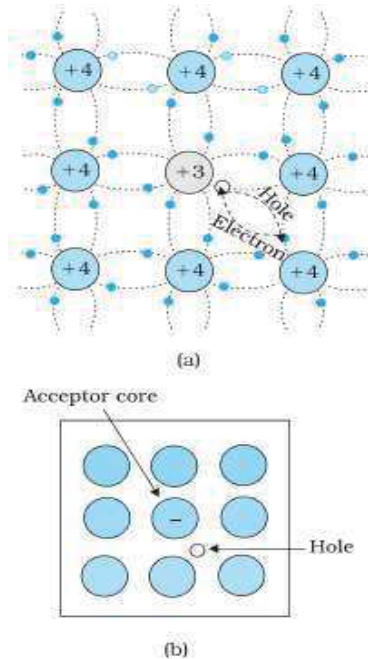


FIG. 8 (a) Trivalent acceptor atom (In, Al, B etc.) doped in tetravalent Si or Ge lattice giving p-type Semiconductor. (b) Commonly used schematic representation of p-type material which shows only the fixed core of the substituent acceptor with one effective additional negative charge and its associated hole.

This is obtained when we dope trivalent impurity like (Al, B, In etc.) into Si or Ge. The trivalent impurity will form 3 covalent bonds with neighbouring Si or Ge atoms while the last neighbour will not have any bonding with the impurity (see fig.8). When the impurity tries to attract an additional electron from the surroundings it will generate a hole at that place. This hole itself is not stable with time but will undergo annihilation with the electron coming from its neighbouring Si or Ge atom. This process continues over the whole lattice hence the hole is delocalized over the whole lattice and hence is available for electrical conduction. These holes are in addition to the intrinsically generated hole due to thermal excitation of electrons in Si or Ge atoms. Hence the holes are in excess to electrons in such semiconductors and these are called *p*-type semiconductors and the impurities are called acceptor impurities. The excess of holes further deplete the thermally generated free electrons via recombination. Hence we have for *p*-type semiconductors $n_h \gg n_e$.

Thus we see that the effect of dopant in extrinsic semiconductors is to enhance the majority charge carriers and reduce the minority charge carriers thus leading to the conductivity provided essentially by majority charge carriers.

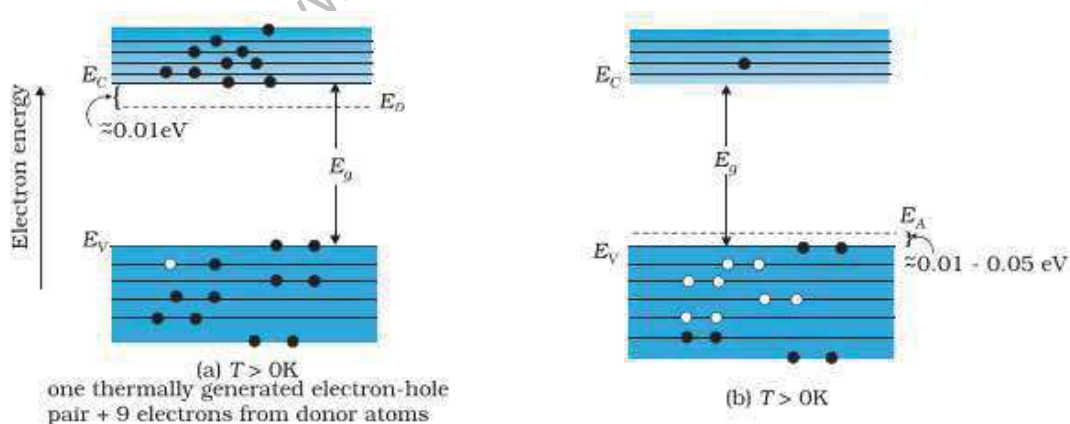


FIG. 9 Energy bands of (a) n-type semiconductor at $T > 0K$, (b) p-type semiconductor at $T > 0K$.

The energetics arising due to the donor and acceptor action of the impurity is schematically represented in fig. 9. The donor (E_D) and acceptor (E_A) impurity states are formed just below the bottom of the conduction band E_C and just above top of the valence band E_V respectively as shown in the figure. The energy separation between the impurity states and the corresponding edges of the bands are very small $\sim 0.01-0.05$ eV. As a result, at room temperature all the impurity states are populated and hence we have large no. of electrons and holes in the conduction and valence band respectively as shown in the figure. The electron and hole concentration in a semiconductor in thermal equilibrium is given by $n_h n_e = n_i^2$, where n_i is the intrinsic carrier concentration.

11. *p-n junction*:

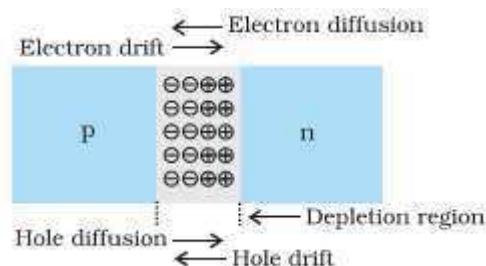


FIG.10 *p-n junction formation process.*

When we dope small amount of donor impurity into *p*-type semiconductor wafer on its side then we get a junction between *p*- and *n*-types of semiconductors called as a *p-n* junction. This device holds both types of majority charge carriers, electrons and holes together, but which occupy different regions in the device. Two main charge transport mechanisms operate within the junction namely diffusion and drift. Diffusion is the process in which the majority charge carriers on the *n*-side of the junction i.e. electrons migrate to the *p*-side of the junction as a result of the electron concentration gradient across the junction and vice versa for the majority charge carriers on the *p*-side of the junction i.e. holes. As a consequence we have a small region (\sim tenths of μm thick) around the junction from which all the majority charge carriers have been removed due to diffusion called as a depletion region (see fig. 10). The diffusion of charge carriers creates local positive or negative charge at the donor or acceptor impurity site respectively. Hence the depletion region is actually a space-charge region with the *n*-side of it acquiring a positive charge and the *p*-side of it acquiring a negative charge. Such a space-charge development leads to the formation of electric field across the depletion region from the positive charge to the negative charge which attracts the minority charge carriers across both *n*- and *p*-sides towards opposite direction giving rise to a drift current. Thus the diffusion and drift currents are in opposite direction to each other. Initially as the junction is formed the diffusion dominates the drift and hence the depletion region starts developing. After it gets well developed the wide depletion region gives rise to significant amount of electric field to increase the drift current. This happens till the diffusion current equals the drift current at which an equilibrium condition is reached and no net current flows across the junction.

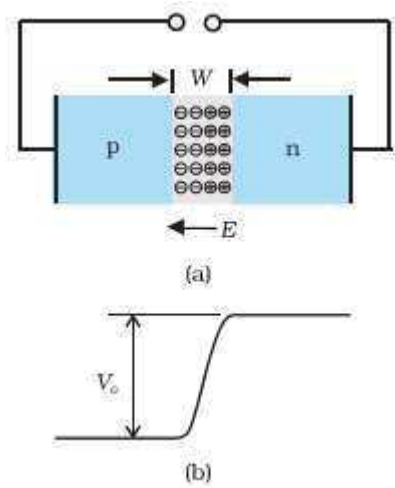


FIG. 11 (a) Diode under equilibrium ($V = 0$), (b) Barrier potential under no bias.

The formation of such an electric field across the junction gives rise a potential barrier across the junction which stops further flow of charge carriers across the junction. Such a potential barrier is represented in fig. 11 which describes a case where no external voltage has been applied to the junction.

Example: Can we take one slab of p-type semiconductor and physically join it to another n-type semiconductor to get p-n junction?

Solution: No! Any slab, howsoever flat, will have roughness much larger than the inter-atomic crystal spacing (~ 2 to 3 \AA) and hence continuous contact at the atomic level will not be possible. The junction will behave as a discontinuity for the flowing charge carriers.

12. Semiconductor Diode:

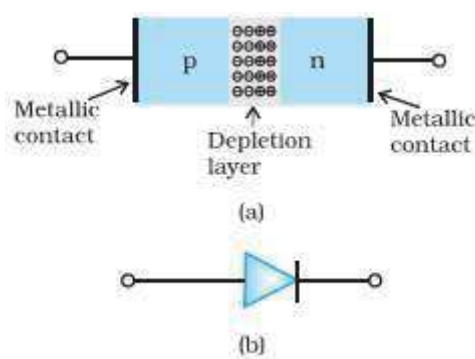


FIG. 12 (a) Semiconductor diode, (b) Symbol for p-n junction diode.

A semiconductor diode is merely a p-n junction with metallic leads at the ends for electrical connection to be made as shown in the fig. 12. The direction of arrow in the symbol used for diode represents the conventional direction in which the current through the diode will flow when it is operated in forward bias mode.

12.1 p-n junction diode under forward bias:

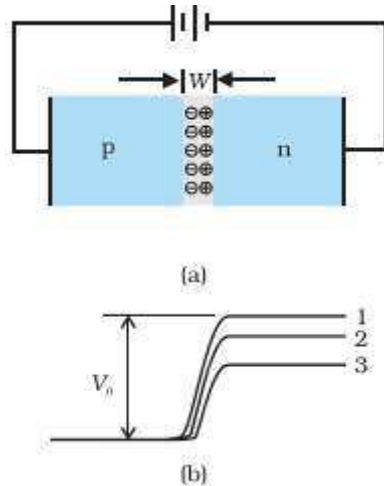


FIG. 13 (a) p-n junction diode under forward bias, (b) Barrier potential (1) without battery, (2) Low battery voltage, and (3) High voltage battery.

When we apply positive potential to the p -side and negative potential to n -side then the diode is said to operate under forward bias. The applied voltage (V) leads to the reduction in the height of the barrier potential since the applied voltage is in reverse direction to the barrier potential (see fig. 13). In that case the depletion region reduces in width since the holes are injected into the n -side from the p -side by the applied potential and vice versa for the electrons thus reducing the size of the depletion region. The injected holes into the n -side constitute minority charge carriers and the injected electrons into the p -side constitute minority charge carriers there hence the process is also called minority carrier injection. This leads to the establishment of net current in the direction of applied voltage the magnitude of which increases with the applied voltage. The effective barrier height under forward bias is $(V_0 - V)$. The magnitude of the net current flow depends on the height of the barrier. Higher the barrier height lesser will be the current since the resistance to the current flow will be higher in that case.

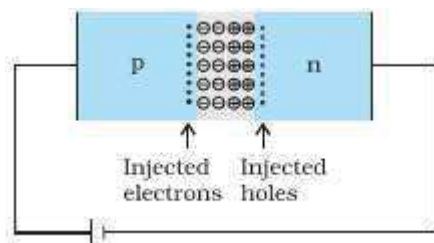


FIG. 14 Forward bias minority carrier injection.

At the junction boundary, on each side, the minority carrier concentration increases significantly compared to the locations far from the junction. Due to this concentration gradient, the injected electrons on p -side diffuse from the junction edge of p -side to the other end of p -side. Likewise, the injected holes on n -side diffuse from the junction edge of n -side to the other end of n -side (see fig. 14). This motion of charged carriers on either side gives rise to an electric current. The total diode forward current is sum of hole diffusion current and electron diffusion current. The total current is usually in mA.