Chapter - 16

Electronics

The success of science and technology in this era is due to a big contribution from electronics. Electronic devices and equipments are used in telecommunications, satellite communication, entertainment and areas like computers, weather forecasting, nuclear physics, etc. The working basis of electronic devices is the controlled flow of electrons. The period of electronics started after the construction of devices based on vacuum tubes. But today devices based on semiconductors have replaced the equipments using vacuum tubes. Semiconductor devices have not only made the electronic equipments smaller in size and increased their work efficiency but their low cost has also made them available for the usage of common man.

For example consider personal computers which are very small in size compared to ancient computers which were based on vacuum tubes having the size of a big room and were of limited use only for normal calculations. The device solar cell which converts solar energy into electric energy is also made from semiconductors. For understanding the behavior of semiconductors devices, it is utmost necessary to have knowledge of theoretical aspects of semiconductors. From this point of view we will study about the nature, conduction process in semiconductors and also gain knowledge about the uses of some common semiconductor devices like diode and transistor, in this chapter.

16.1 Energy Bands in Solids

From the study of atomic structure we know that in an isolated atom, the electrons are confined to well defined energy levels. Every electron in an atom stays in one of these discrete energy levels. The maximum number of electrons present in any energy level is decided by the Pauli exclusion principle. The outermost energy level in which the electrons are present in unexcited state of atom is called the valence energy level, and such electrons are called valence electrons. For example, the electronic configuration for sodium (atomic number 11) is $1s^2 2s^2 2p^2 3s^1$. Here, in 3s energy level there is one electron which is the valence electron.

Generally, most of the solid substances including metals are of crystalline nature. In crystalline solids the atoms are arranged in a regular periodic arrangement and the nearby atoms are apart by a very small distance called the 'lattice constant'. The value of lattice constant is different for different crystalline substances and is of the order of atomic size (Å). Clearly, it is seen that due to such small distance between the nearby atoms the electrons of any atom are not only affected by the coulomb force of its nuclei but nucleus and electrons of nearby atoms would also exert some coulomb force them. This can be summarised as the mutual interaction between the atoms of solids. This interaction is responsible for the bonding of various atoms leading to the formation of crystalline structure.

When the atoms in a crystal are interacting then they are not isolated. Hence in crystals the energy levels of the electrons in atoms are not same as the energy level of isolated atoms. Therefore, due to the interaction of the atoms of a crystal, energy levels for atoms in crystals are modified. To understand this transformation first, we discuss the interaction between two identical atoms. Initially, it is assumed that these atoms are so apart that there is no interaction between them (if the distance between two atoms is more (~ 50Å) than its linear dimension (~ 10Å) then it is possible to assume so). In this situation both atom can be treated as isolated. Therefore their energy levels will be same as the energy levels of isolated atoms. This is shown in the diagram 16.1 (a). When these two atoms are close enough to interact. Due to this interaction each energy level of both the atoms is splited into two energy levels, out of which one is a little higher and the other is a little lower than the original energy level. [figure (16.1(b)]. In other words, it can be said that for the system formed by the two atoms there are two energy levels corresponding to each level of an isolated atom. This change in energy levels is in accordance with the 'Pauli exclusion Principle'.

Now we will see the process of crystal formation. For convenience we assume a one dimensional crystal made up of N-atoms. When these atoms are brought closer then due to their mutual interaction energy level of every atom gets splited. This split in the energy levels is directly proportional to the number of interacting atoms. As a result, in a system of N atoms there would be 2 N energy levels in place of a single energy level. If the value of N is very high then these N energy levels will be so closely spaced that they can be considered as to form nearly continuous energy groups theses continuous energy groups are called energy bands. In an actual crystalline solid the value of N is very high approximately from 10^{22}-10^{23} atom/cm^3. Hence, in each energy bands same number of energy levels will be there. There would be a very small separation between such energy levels. For example if there is a difference of leV between the minimum and maximum energy of a band and there are 10^{22} energy levels then there would be 10^{-22}eV interval between two consecutive energy levels. For such a smalls energy interval it can be considered that in an energy band, energy is continuous. This process of energy band formation is shown in the figure 16.1 (c) It is worth noting that in the process of getting the atoms closer the electrons in the outermost orbit of all the atoms will be affected first and and the most. Due to this the energy bands associated with them will be of higher width. The internal orbits of the atoms are not much affected by the interaction hence the energy bands will be of smaller width. This is also shown in the figure 16.1(c). It is also clear from the figure that corresponding to 1s, 2s, 2p, 3s..........energy levels for atoms there are associated energy bands, in a crystal which are called 1s band, 2s band, 2p band 3s band etc.

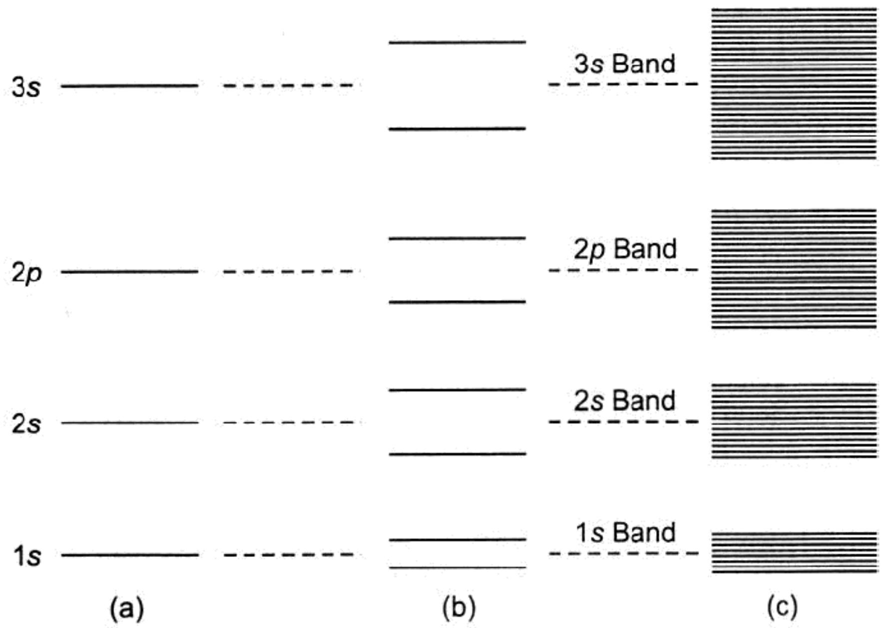


Fig. 16.1 : (a) Engergy level of single isolated atom (b) Split in energy levels due to the interaction of two atoms (c) N atoms (Band formation in crystal structure)

In an isolated atom the electrons first fill the minimum energy level, then to higher energy level then to still higher energy level; upto the valence levels similarly the electrons are filled in these bands. The band in which valence electrons are present is called the valence band. In an isolated atomt here exists an energy gap between discrete energy levels. Similarly in solids there are forbidden energy gaps between the energy bands with the gap representing a range of energies that no electron can have. Here it should be mentioned that the energy band formation is different for different soilds and the energy span of bands and value of forbidden energy gaps is different for different solids.

As in case of isolated atoms the electrons can move from lower energy level to higher energy level by gaining necessary energy. Similar transition of electrons is possible in solids. In solids these transition are of two types : (1) transition with in the same band of energy levels. (2) transition between of energy level belonging to different bands. For both the transition it is necessary that higher energy level should be empty. Since there is very little difference between the two consecutive energy levels of the same band, for such type of transitions very small energy is required. Whereas for transition between two consecutive bands the electrons will require energy equal to the forbidden energy gaps.

16.2 Classification of Soilds as Conductors, Insulators and Semiconductors

Electrical conductivity is a physical quantity whose span is very wide. On one hand we know about metals having very high electrical conductivity on the other hand there are insulators like quartz and mica whose electrical conductivity is very low. Substances are also known, whose conductivity is very small compared to metals at normal temperatures but very large than that of insulators. These are called semiconductors, examples being silicon and germenium.

The conductivity of semiconductors is not only intermediate to conductors and insulators but also the change in the conductivity with temperature is very interesting. Near absolute zero their behaviour is similar to insulators but as the temperature increases the conductivity also increases which is opposite to the observed behaviour for metals. The following questions are not answered by free electron theory for metallic conduction which you have read earlier:

(1) Why are the conductivities of solid substances different?

(2) Why does a substance show the behaviour of a semiconductor?

(3) Why is the change in conductivity with temperature different for metals and semi conductor?

The theory of energy bands in solids provides answer to these questions, and on the basis of their band structures materials are divided into conductors, insulators and semiconductors. Every solid substance has its own band structure which defines its electrical behaviour.

16.2.1 Conductors

Conductors are such solid substances in which, the velence band is partially filled or this band overless with next higher band to from a new band which is again partially filled. In both the situations there are empty energy levels available with in the band to which electrons can make transitions, provided they get energy from external electric field and get exicted.

For example we consider sodium which is a monovalent metal. For this the band structure is such that 1s, 2s and 2p bands are filled to their capacities with electrons and 3s band is half filled.

The reasons for this is that in an isolated sodium atom1s, 2s and 2p energy levels are completely filled but 3s energy level, whose capacity is to take two electrons, is filled only with one electron. Completely filled inner energy bands are not useful for electrical conductivity because the transition of electrons is not possible between electronic energy levels within such bands.

The transition of electrons is not possible between internal bands from 1s to 2s or 2s to 2p because in both the situations the empty energy levels are not available to the electrons. Opposite to this is 3s band which is half filled so half of the energy levels in this band are available for transitions of electrons.

Therefore, the electrical conductivity of sodium is due to this partially filled band. This band is shown in the figure 16.2 (a). The lower half part of this band is called the valence band and upper half part is called the conduction band, because after getting the energy from the electric field the electrons reach this part to start conduction.

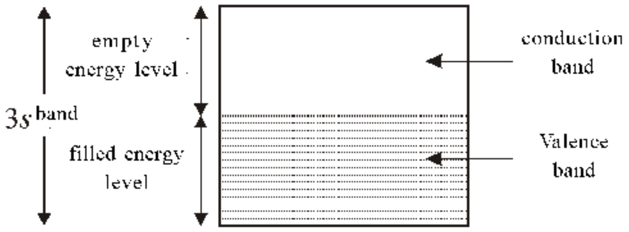


Fig. 16.2 (A) Half filled 3s band for sodium

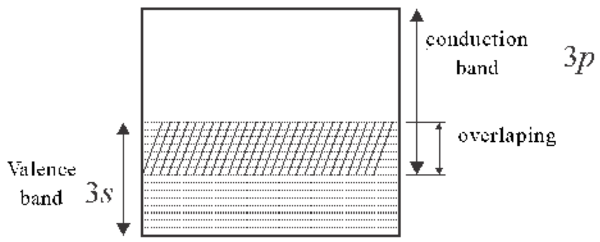


Fig. 16.2 (B) Overlaping of bands in magnesium

Like sodium other monovalent metals also have partially filled bands.

Magnesium, zinc and other substances like them, which are the members of the second group of the periodic table are called divalent substances and they also come in the category of metals. In these types of substances (in the solid state) there is overlapping of bands. For example magnesium's (atomic number 12) electrnic configuration is $1s^2\, 2s^2\, 2p^6\, 3s^2$, and in atomic state there is energy gap between 3s and 3p. But in the process of crystal formation the spliting in the energy levels is such that the fully filled 3s band gets overlapped with fully empty 3p band. Now electrons get the necessary empty levels for transition in this overlapped band. In this state if 3s band is called the valence band, then 3p is the conduction band, thus for such metals the valence band and the conduction band overlap. (figure 16.2(b)).

To conclude, in both of the above metals there is no interval between the maximum energy of valence band and minimum energy of conduction band because of this, the forbidden energy gap for a metal is zero.

Generally, the energy provided by conventionally used sources of electric current is in the range 10^{-4} to 10^{-8}eV; which is sufficient to move the electrons to the empty energy levels of this partially filled band. As mentioned already, there is very small energy difference between the energy level of any band. Thus if empty energy levels are available with in a band, the electrons after absorbing a very small energy reach to such states like free electrons. Hence these are also called free electrons. When electrons after getting a very small amount of energy, reach the empty energy levels then there is a drift of electrons due to which conduction is possible.

In metals, both the free electrons and the necessary energy levels for their transition are available in abundance. Due to this the electrical and heat conductivites of metals are very large. For normal temperatures, the value of electrical conductivity of metals is of the order of 10^6 to 10^8 mho/meter clearly supporting this fact. There is no significant change in the number of free electrons, due to thermal energy in a metal. The effect of heat energy will be to increase in number of collisions between the free electrons and the vibrating ions of the solids, due to which their conductivity will reduce with the increase in temperature.

16.2.2 Insulators

Insulator are solid substances in which the configuration of energy bands is such that the valence band is completely filled and the forbidden energy gap between the maximum energy E_v of valence band and minimum energy E_c of conduction band is very large. The forbidden energy gap is given by; (figure 16.3).

$$E_g = E_c - E_v$$

For insulators, the value of E_g is around 3 - 7 eV. As there are no electrons in fully empty band so it does not take part in electrical conduction. Similarly, in fully filled band the electrons are there but there are no empty energy levels available for the transition so again no electrical conduction is possible for such bands.

As it is clear that very less energy is gained from commonly employed energy sources, hence the electrons do not get enough energy to move from valence band to conduction band. Similarly at normal and even at high temperatures the electrons of valence band do not get energy equal to the forbidden energy gap. Due to this the electrons are not able to reach from valence band to conduction band. Hence, such solids are called insulators.

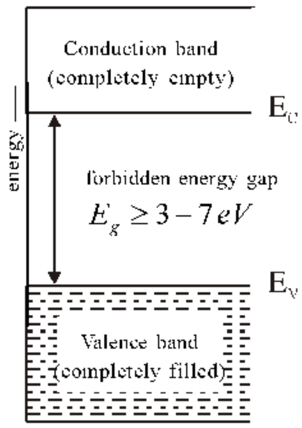


Fig. 16.3 Energy band diagram for insulator

For example, diamond (forbidden energy gap (~ 6eV), comes in the category of insulators. Generally, the conductivity of insulators is between 10^{-12} mho/meter to 10^{-18} mho/meter (meaning resitivity 10^{12} ohm-meter to 10^{18} ohm-meter).

16.2.3 Semiconductors

Semiconductors are such solid substances having band structure similar to that of insulators. The only difference is that the value of forbidden energy gap is less in comparison to that of insulators and is approximately of 1eV order. At absolute zero, the valence band is completely filled and the conduction band is fully empty. Due to this both these bands do not participate in conduction. This is the same bahaviour as that of insulators, thus, at absolute zero the insulators and semiconductors behave in the same way.

At room temperature or a temperature higher than that, the electrons in the energy levels of the upper part of the valence band get thermal energy and move to conduction band, to take part in the process of conduction. Same number of energy levels get emptied in the valence band in which the electrons of the same band make transition to take part in the process of conduction. (In valence band the process of conduction is explained in terms of a positive free charge (hole) which is explained in next section). Thus at room temperature the conductivity of semiconductors is more than that of insulators. However, as only a very small number of the electrons from the valence band reach to the conduction band due to thermal energy for participation in the process of conduction. In the case of conductors a large number of electrons are present in the conduction band. Thus the conductivity of semiconductors is much less, in comparison to conductors. The above discussion clearly suggests that the conductivity of semiconductors is intermediate between the conductivity of conductos and insulators. This is the reason why such solids are named as 'semiconductors'.

The conductivity of intrinsic semiconductors increases as the temperature increases, this is explained in next section. Silicon, germanium and gallium, arsenide are examples of some useful semiconductors. Apart from this lead sulphide indium antimonide, gallium phosphilde and silicon carbide are also semiconductors. The forbidden energy gaps for some semiconductors are given in the table (16.1) below.

Table 16.1

Semiconductors	Forbidden Energy Gap
Silicon (Si)	1.12 eV
Germanium (Ge)	0.7 eV
Indium antimonide (InSb)	0.17 eV
Gallium arsenide (GaAs)	0.33 eV

16.3 Intrinsic Semiconductors

The semiconductors having no impurity are called intrinsic semiconductors. In ideal state in this type of intrinsic semiconductors, there should be atoms of only that semiconductor. But in reality it is not possible to get such type of crystals. Therefore, if in a semiconductor material the number of inpurity atoms and the number of semiconductor atoms is in the ratio $1:10^8$ or less, then it is considered as an intrinsic semiconductor. To study the properties of intrinsic semiconductors, here we consider silicon and germanium.

Silicon and germanium both are the members of fourth group of the periodic table and their valency is 4. Their electronic configurations are as follows:

Ge (32) = $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^6\ 3d^{10}\ 4s^2\ 4p^2$

Si (14) = $1s^2\ 2s^2\ 2p^6\ 3s^2\ 3p^2$

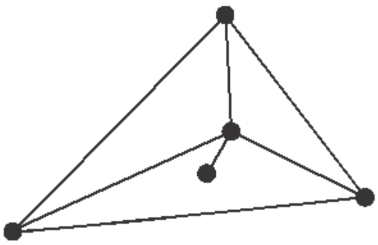


Fig. 16.4 : Tetrahadron structure for silicon (or germanium)

In the crystal of both these elements the atoms are in an ordered array in such a way that each atom is inside a regular tetrahedron and on the four corners there are other nearby atoms. Figure (16.4) shows one such tetrahedron unit. Every valence electron of the atom makes a covalent bond with the valence electron of the nearby atom. In this way every atom is connected with the nearby four atoms through covalent bonds. In every covalent bond there are two electrons. For convenience in study ; figure (16.5) shows the two dimensional form of crystal structure of germanium; which is also applicable for silicon.

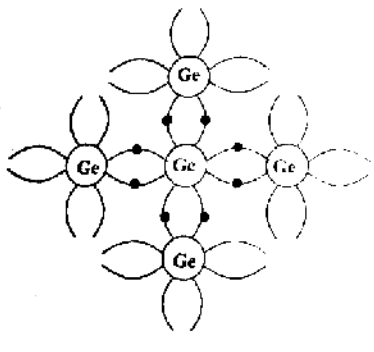


Fig. 16.5 Crystal structure of Germenium at 0 K

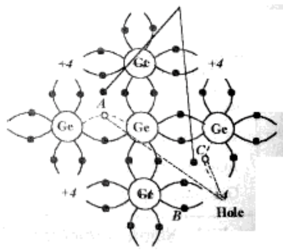


Fig. 16.6 Electron-hole pair in germenium

At absolute zero (0 K) the valence electrons are bonded in covalent bonds hence there is no free electron available for electrical conduction. Due to this, at absolute zero temperature, intrinsic semiconductors behave as insulators. When the temperature of the crystal is increased, some valence electron gain enough thermal energy and break their covalent bonds to become free. These free electrons move in the crystal freely and participate in electrical conduction, when external electric field is applied

When one electron moves out of the covalent bond a vacancy is created. This vacancy is called a hole. The absence of electron is equal to the presence of same magnitude of positive charge. Hence, the hole can be assumed as a positive charge having the same magnitude as that of an electron. As explained later, holes also participate in electrical conduction in semiconductors. In a semiconductor when a covalent bond breaks then an electrons hole pair is generated. At room temperature (300K) or a temperature close to it there are many such electron hole pairs available. The production of electron hole pair is shown in figure (16.6)

At normal temperatures, electrons and holes have random motion in semiconductors. The random motion of hole can be understood in figure (16.6).

Suppose due to thermal energy an electron gets free from a covalent bond at position A, hence at this place a hole is generated. From a nearby atom, the valence electron breaks the covalent bond (at place B) and have a transition to hole at position A. The reason for this is that electron is having transition from one bond to the other and all the electrons present in the bond are approximately are of same energies. As shown in the figure due to the movement of electron from C to B the hole will move from B To C, etc. In this way in any semiconductor the electrons and holes both do random motion. Since, hole can be taken as a free positive charge +e, hence in a semiconductor both holes and electrons act as charge carriers and participate in electrical conduction.

This process of hole electron pair production in intrinsic semiconductors can also be explained by band theory. The forbidden energy gap in semiconductors is of 1eV order. At absolute zero temperature, the valence band is completely filled and conduction band is fully empty. Hence, semiconductors behave like insulators. When the temperature of the crystal is increased electrons get enough thermal energy and are capable of crossing the forbidden energy gap. Such electrons reach the conduction band from the valence band and holes are generated in the valence band in place of electrons (figure 16.7) Due to transition of one electron from valence band to conduction band a hole is generated and here also the presence of electron hole pair can be understood. The electrons present in conduction band are called free electrons and do random motion, similarly in valence band holes do random motion.

In fact, the band describtion and covalent bond describtion for a semiconductor are equivalent. It can also be said that due to thermal energy the covalent bond breaks and the electron reaches the conduction band from the valence band. Therefore, for this process the necessary minimum energy would be equal to band interval energy E_g. Here this questions is logical that at room temperature ie. 300 **K** the average kinetic energy of the electron is kT = 1/40 = 0.025 eV, whereas the necessary energy for transition is about 1eV. Then, how does an electron reaches the conduction band from the valence band? Here it is worth noting that kT is the value of average kinetic energy and not that all electrons are of this energy. Very few electrons are there having energy of 1eV order capable of making transitions from valence band to conductrion band.

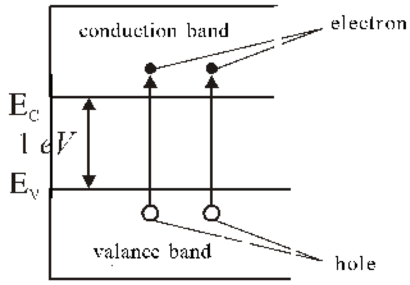


Fig. 16.7 : Effect of temprature on intrinsic semi conductor

It is clear from above discussion that in any intrinsic semiconductor at a finite temperatures the number of electrons and holes is same. If in an intrinsic semiconductor the concentrations of free electrons and holes are n and p respectively then

$$n = p = n_i$$

$$np = n_i^2 \qquad ......(16.1)$$

where n_i is called as the intrinsic charge carrier density.

When the temperature of an intrinsic semiconductor is increased compared to prior temperature, more covalent bonds break, meaning more electrons reach the conduction band from valence band and the number of electron hole pair increases. This means that n_i is dependent on temperature and increases with temperature. Like wise of two different semiconductors, (for example, silicon and germanium) are at same temperatures then since the value of forbidden energy gap E_g for them is different, hence, the semiconductor for which the value E_g is less will have more covalents bonds to break or relatively more electrons will make transition from valence band to conduction band. Clearly at same temperature, for two semiconductors for which the forbidden energy gap E_g are not same; the value of intrinsic charge carrier density, n_i will not be same but would be more for the semiconductor having less value of E_g. In this way in an intrinsic semiconductor, the value of intrinsic charge carrier density, n_i is dependent on (i) temperature (ii) and nature of semiconductor material. In mathematical form this dependency is shown by the following formula;

$$n_i \approx AT^{3/2} \exp-\left[\frac{E_g}{2kT}\right] \qquad \text{........(16.2)}$$

Where T is absolute temperature, k is Boltzmann constant and A is another constant.

At same temperature, comparing intrinsic silicon (E_g = 1.1 eV) and germanium (E_g = 0.7 eV) there are more number electron hole pairs in intrinsic germanium than in intrinsic silicon.

16.3.1. Conduction in Semiconductors

If some electric potential is applied across a semiconductor, then drift motion of elctron and holes takes place. Electrons drift towards the positive terminal from the negative terminal whereas the holes drift from positive terminal to negative terminal. (figure 16.8). In other words, in a semiconductor the electrons move opposite to the direction of applied electric field and holes move in the direction of electric field. The current in the metallic wires in external circuit is due to the flow of electrons.

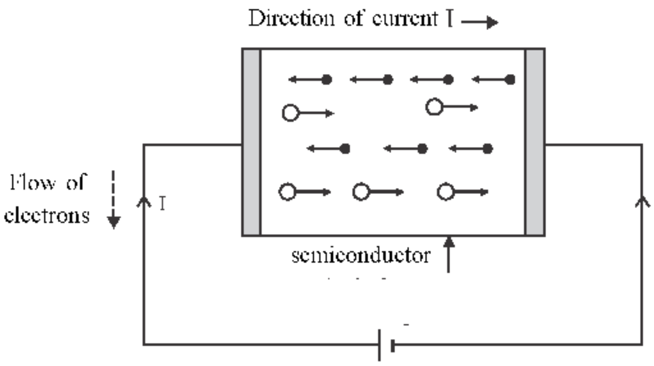


Fig. 16.8 : Current in intrinsic semiconductor

Generally, the drift velocity of electrons and holes for commonly employed electric fields is directly proportional to the applied electic field. If v_n and v_p are the drift velocities for electrons and holes respectively

then ; $$v_n \propto E \ and \ v_n = \mu_n E \qquad \text{....(16.3)}$$

and $$v_p \propto E \ or \ v_p = \mu_p E \qquad \text{.....(16.4)}$$

Where μ_n and μ_p are the mobilities for electrons and holes respectively. The mobility is measured in $m^2 v^{-1} s^{-1}$. Drift velocites in vector form are given as

$$\vec{v}_n = -\mu_n \vec{E}$$

$$\vec{v}_p = +\mu_p \vec{E} \qquad \text{.......(16.5)}$$

beacause electrons move opposite to the field $\vec{E}$. In presence of some electric field if the current density due to electrons is denoted by $\vec{J}_n$ and that due to holes by $\vec{J}_p$ then total current density due to drift in a semiconductor is;

$$\vec{J} = \vec{J}_n + \vec{J}_p \qquad \text{...........(16.6)}$$

By the diffinition of current density ;

Current Density = The number density of number of charges $\times$ charge $\times$ drift velocity, therefore

$$\vec{J}_n = n(-e)\vec{v}_n = n(-e)(-\mu_n \vec{E})$$

or $$\vec{J}_n = ne\mu_n \vec{E} \qquad \text{...........(16.7)}$$

and $$\vec{J}_p = p(e)\vec{v}_p = pe\mu_p \vec{E} \qquad \text{...........(16.8)}$$

Hence, it is clear that although the drift velocities of both the electrons and holes are in opposite direction but the corresponding current densities are in the same direction (direction of external electric field). Therefore total current density;

$$\vec{J} = \vec{J}_n + \vec{J}_p$$

$$\vec{J} = ne\mu_n \vec{E} + pe\mu_p \vec{E}$$

$$\vec{J} = (n\mu_n + p\mu_p)e\vec{E} \qquad \text{...........(16.9(a))}$$

Since, in the above formula $\vec{J}$ and $\vec{E}$ are in the same direction. Hence, in scalar form:

$$J = (n\mu_n + p\mu_p)eE \qquad \text{...........(16.9(b))}$$

By difinition the electrical conductivity of semiconductor.

$$\sigma = \frac{J}{E} = e(n\mu_n + p\mu_p) \qquad ....(16.10)$$

And resistivity of semiconductor;

$$\rho = \frac{1}{\sigma} = \frac{1}{e(n\mu_n + p\mu_p)} \qquad ....(16.11)$$

The above discussion is true for both intrinsic and extrinsic semiconductors for an intrinsic semiconuctor $n - p - n_i$ respctively.

$$\sigma = n_i e(\mu_n + \mu_p) \qquad ...(16.12)$$

$$\rho = \frac{1}{n_i e(\mu_n + \mu_p)} \qquad ...(16.13)$$

Here, it is worth noting that the electron and hole have same charge magnitude but their mobilities are not same. In electric current, holes behaves as a positive charge having some effective mass. This mass is slightly more than the effective mass of an electron. Hence, the mobilities of electron and hole are different. (You will study about the effective mass in detailed in higher classes). The effective mass of electron and hole are different for different substances therefore mobilities of electrons and holes are different for different semiconductors. Table (16.2) shows the values of μ_n and μ_p for silicon and germanium at temperature T= 300 K.

Table 16.2

Semiconductor	μ_n(m^2/Vs)	μ_p(m^2/Vs)	$n_i(/m^{-3})$
Si	0.13	0.048	9.65×10^{15}
Ge	0.39	0.19	2.5×10^{19}

Since, in the flow of current in a semiconductor both negative and positive free charges (electron and hole) participate; the conduction in semiconductor is called bipolar conduction. Opposite to this is metals where only electrons participate in the flow of current. hence they have unipolar conduction.

16.3.2 Effect of Temperature on Electrical Conductivity of Interinsic Semiconductors

The equation (16.12) suggest that the conductivity of any intrinsic semiconductor is dependent on the number of charge carriers (n_i) present. Since, n_i is dependent on temperature and forbidden energy gap, hence, σ also dependent on them. For any given semiconductor (E_g = constant) if the temperature T increases then σ increases exponentially (or for resistivity, when temperature increases it decreases exponentially.) Also with the temperature increase the collisions between the free electrons and holes and vibrating atoms of semiconductors will be more frequent due to which the value of μ_n and μ_p decreases. However, this reduction is not that effective as the increases in the value of n_i due to increase in temperature and the increases in the value of σ. The increase in the conductivity of intrinsic semiconductors is a special property of intrinsic semiconductors which is not seen in metals. In metals the conductivity is reduced (or the resistivity is increased) with the increase in temperature, exhibiting $\sigma \propto T^{-1} (or\, \rho \propto T)$ behaviour at normal temperatures.

The reason for the difference in two is that in metals the free electron density n does not change with temperature but with the increase in temperature the collisions between the electrons and vibrating ions are more frequent in metals so conductivity decreases with temperature.

On the basis of above description, it can also be said that the temperature coefficient of resistance for intrinsic semiconductors is negative. For silicon its average value is $\approx -0.07/K$. Here it is important to note that at absolute temperature $n_i = 0$ for semiconductors hence $\sigma = 0$ (insulator behaviour).

Since the conductivity of intrinsic semiconductors is small and limited, there is no direct use of these in semiconductor devices. But their resistivity changes due to temperature or incident light energy hence these are used in as radiation sensitive resistance. or as LDR.

Example 16.1 : Calculate the resistivity of intrinsic germanium at 300 K temperature. Mobility of electrons and holes and n_i value for germanium are according to the table (16.2).

Solution : Resistivity formula for intrinsic semiconductor

$$\rho = \frac{1}{n_i e(\mu_n + \mu_p)}$$

Where $e = 1.6 \times 10^{-19} C$ and from table (16.2)

$$\mu_n = 0.39 m^2 / Vs.$$

$$\mu_p = 0.19 m^2 / Vs.$$

$$n_i = 2.5 \times 10^{19} / m^3$$

Substituting the value in above formula;

$$\rho = \frac{1}{1.6 \times 10^{-19} \times 1.5 \times 10^{19} (0.39 + 0.19)}$$

$$= \frac{1}{4 \times 0.58} = 0.43 \Omega m$$

Example 16.2 : For some intrinsic semiconductor the forbidden energy gap is E_g electrons volt. What maximum wavelength of light can be absorbed by this semiconductor.

Solution : If ν is the frequency of incident light photon then the energy of this photon will be $h\nu$. If the energy of incident photons is higher than the forbidden energy gap of the semiconductor then the electrons present in the valence band will absorb these photons and reach the conduction band. The process is called photo electron hole pair production. Hence, minimum frequency of photon for absorption ν_{min}, is given by -

$$h\nu_{min} = E_g.$$

If λ_{max} is the maximum wavelength for absorption of photon then;

$$\frac{hc}{\lambda_{max}} = E_g \quad \text{or} \quad \lambda_{max} = \frac{hc}{E_g}$$

As hc = 1242 eV nm and E_g when taken in eV unit, we have

$$\lambda_{max} = \frac{1242}{E_g} nm.$$

16.4 Extrinsic Semiconductors

As described earlier that intrinsic semiconductors have limited conductivity, hence they are not used as such. If in any intrinsic semiconductor impurities of suitable material are mixed in very small proportion then the conductivity of semiconductor so obtained is many times more than that of intrinsic semiconductors. To increase the conductivity of a semiconductor impurities of suitable kinds are added to it in very small quantity. This process is called doping, and the semiconductors so obtained is called extrinsic semiconductors. The suitable impurities for silicon and Germanium are either the elements of Vth group of the Periodic Table (like arsenic (As), antimony (Sb), phosphorous (P), etc). or the element of group (III) (like, aluminium (Al), gallium (Ga), indium (In), boron (B) etc). Due to the impurities of group V element in Si or Ge, the number of free electrons increases while for the impurities of Group III elements the number of holes increases.

Therefore the extrinsic semiconductors are of two types.

(i) N-type extrinsic semiconductors.

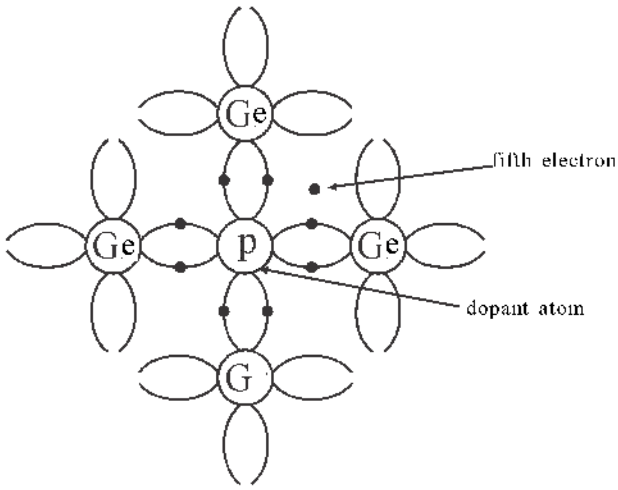
(ii) P-type extrinsic semiconductors.

The conductivity of any extrinsic semicounductor is controlled by the quantity of impurity mixed. The quantity of impurity element is very small. One atom of impurity is mixed with approximately 10^6 atoms of an intrinsic semiconductor. Due to this there is no significant change in the original crystal lattice but the increase in conductivity is very large. For example, if 1 impurity atom is mixed in 10^9 germanium atoms then its conductivity increases 10^3 times in comparison to intrinsic semiconductor. Generally, all semiconductor devices are made form extrinsic semiconductors.

16.4.1 N-Type Semiconductors

When a very small amount of impurity of pentavalent element is mixed in in tetravalent intrinsic semiconductor then N-type of extrinsic semiconductors is obtained. To understand the effect of such kind of impurity element on the semiconductors, we take the example of mixing phosphorus (impurity) in intrinsic silicon. In the process some silicon atoms are replaced by the phosphorous atoms. Since, the impure materials are mixed in very small quantity, hence, they are surrounded by silicon atoms. Phosphorous atom has five valence electrons to make the bonds with neighbouring Silicon atoms having only four valence electrons. Hence, pentavalent phosphorous makes covalent bonds with the four nearby silicon atoms but its fifth electron remains unbound, [figure (16.9)]. In comparison to the four other electrons, this fifth electrons is lousely bound its bonding energy is approximately 0.05 eV. This energy is equal to the mean thermal energy of the atoms, hence the electron gains this energy and is freed. At room temperature (300 K) each impurity atom can provide one free electron to the crystal. In other words, every impurity atom gives a free electron to the semiconductor. Due to this, the impurity atom is called donor atom and such type of impurities are called donor impurities. The number density of electrons so obtained depends on the number of donor atoms.

On the other hand the energy required for getting free electron by breaking the covalent bonds of silicon is 1.1eV. Hence, a very small number of electrons get free. Therefore, at normal temperature only 1 atom out of 10^{10} silicon atoms is able to generate electron hole pair. Now if one impurity phosphorous atom is present per 10^6 silicon atoms, then every phosphorous atom denotes one electron to the crystal and per 10^{10} atoms one electron is produced by breaking of covalent bond. Hence, when comparison is done then a ratio of $10^4 : 1$ is obtained from impurity and thermal generation respectively. Clearly, with a very small impurity the number of free electrons and conductivity increases by a multiple of 10^4. In such type of semiconductors the number of free electrons is more than the number of holes. Therefore here electrons are majority charge carriers and holes are minority charge carriers. Such type of semiconductors are called N-type of semiconductors which clearly illustrates the fact that here the negative free charge carriers electrons are in abundance.



e

Fig. 16.9 : Dopping of phospores in germanium

Although in a N-type semiconductor, the negative charge carriers are in abundance as compared to holes but over all, the N-type semiconductors is electrically neutral. Donor atom gives an electrons and becomes a positive ion. The number of free electrons given to the crystal by the donor atoms, is equal to the number of positive donors ions so produced. Thus electrical neutrality of crystal is maintained. Positive ions remains fixed at their place in the crystal, because their remaining electrons are bonded with the other atoms. Hence, they do not contribute in electrical conduction.

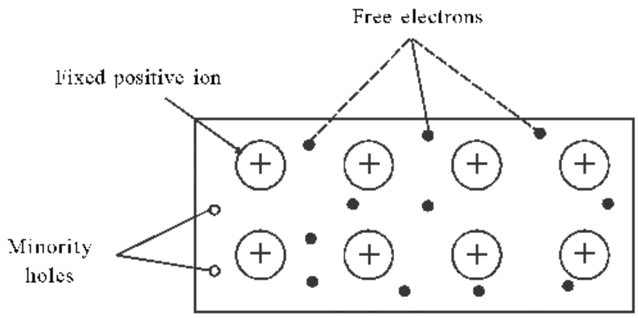


Fig. 16.10 N type semiconduction

16.4.1 Energy Band description for N-type Semiconductors

When donor impurities are mixed in a intrinsic semiconductors then such impurity atom create energy levels a little below the bottom of conduction band in band gap. These levels are called impurity levels. (Fig 16.11). The difference between minimum energy of conduction band E_c and the energy of these energy levels E_c - E_d, is very small (for phosphorus in Si it is ~ 0.05 eV and for Germanium it is ~ 0.01 eV). Hence, the electrons of these donor levels, gain energy easily form the crystal's thermal vibrations and reach the conduction band to participate in conduction.

Apart from this, due to thermal energy some electrons from valence band also reach conduction band as in the case of intrinsic semiconductor. In such a process holes are generated in valence band. But now the number of holes in valence band is less then the number of electrons in conduction band; because the conduction band is getting electrons both from thermally generated electrons and from donor energy levels.

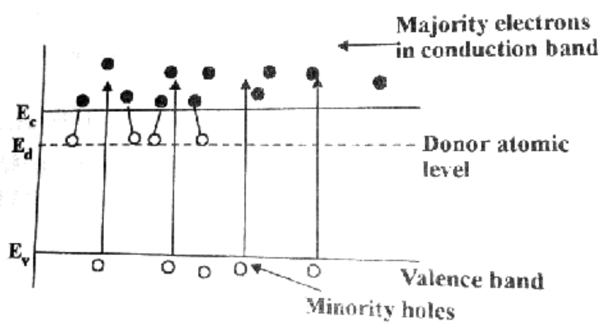


Fig. 16.11 : Band model for N-type semiconductor

16.4.2 P-Type Semiconductors

When a trivalent impurity (element) like (aluminium, boron, indium or gallium) is mixed in very small quantity with tetravalent intrinsic semiconductors, then P-type semiconductors are obtained. For example, if indium is mixed as an impurity element in intrinsic silicon then indium atom replace silicon atom in the crystal lattice. Its there valence electrons form covalent bonds with the nearby silicon atoms but fourth silicon atom's covalent bond lackes one electron . Hence, a vacancy is created which generates a hole. At room temperature due to thermal energy, an electron from a near by silicon atom may move towards the vacant position of this impurity atom to complete the bond. [Figure 16.12 (a)]. In this process, the impurity atom now becomes a negatively charged ion. On the other hand a hole is generated in this atom. Here, every impurity atom gains an electrons and provides a hole to the semiconductor. These holes are in addition to the thermally generated holes. The impurity atoms are called acceptors because they recieve electron from the bonds of intrinsic semiconductors. The number density of the holes depend mainly on the quantity of impurity.

In such type of semiconductors the number of holes is much more than the number of free electrons. Hence, here the holes are the majority charge carriers and the electons are the minority charge carriers. Since, holes behave as positive particles, such type of semiconductor is called P-type semiconductors.

P-type semiconductors can also be understood, on the basis of band structure. Due to the doping of acceptor atoms, acceptor energy levels are generated in forbidden energy gaps a little above the maximum free energy E_V of the valence band [Figure 16.12(b)].

If E_a denotes the energy of these energy levels then in Silicon E_a E_V ~ 0.05 eV and for germanium E_a $-E_V$ ~ 0.01 eV. This energy is easily acquired by the thermal vibrations of the crystal and the electron make transition form the valence band to acceptor energy levels; due to which a hole is generated in the valence band.

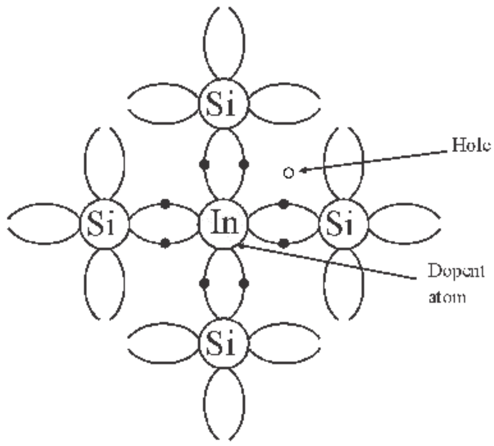


Fig. 16.12 (A) Bond diagram for P-type semicondutor

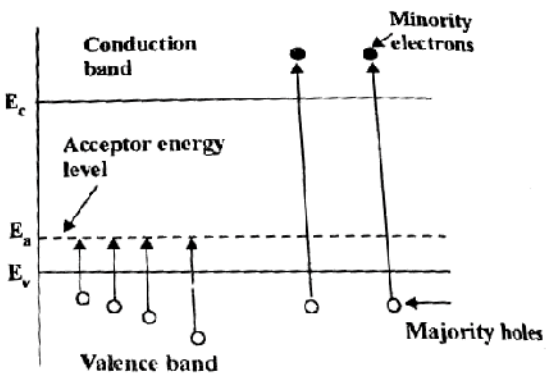


Fig. 16.12 (B) Energy band model for P-type semiconductor

Similar to N-type semiconductor, P-type semiconductors is also electrically neutral. Here as many holes are generated from impurty atoms, same number of negative acceptor ions are also produced. Also the total sum of positive and negative charges is zero. Negative ions are fixed in the lattice hence, they do not contribute to electric current.

16.5 P-N Junction

When a P-type semiconductor is joined atomically to a N-type semiconductor so that the continuity of crystal structure is maintained at their contact surface (interface) ; then this contact surface is called P-N junction and the device so formed is called PN junction diode. Generally, in a single crystal of silicon or germanium, the impurities are so added by special techniques that its one side becomes P-type semiconductor and the other side becomes N-type semiconductor, the boundary between these two regions is called P-N Junction. A PN junction cannot be formed by pressing a P-type semiconductor over a N-type semiconductor because in this case, the crystal structure is not continous across the contact surface.

As soon as a junction is formed, as there is a large concentration of holes on P side and a very small concentration of electrons on N side there will be a diffusion of holes from P side to N side. Similarly because of difference in concentration electrons diffuse from N side to P side. This motion of charge gives rise to diffusion current across the junction. Note that diffusion current due to electrons and holes is from P side to N side. When an electron diffuses from N side to P side it leaves behind an ionised donor on N-side. This ionised donor (positive) charge is immobile as it is fixed in crystal lattice. As the electrons continue to diffuse form N to P, a layer of positive charge (or positive space-charge-region) is formed on N side of the junction. Similarly, when a hole difuses form P to N, it leaves behind an ionised acceptor (negative charge) which is also immobile. As the holes continue to diffuse, a layer of negatively charged ions (or negative space-charge region) is formed on the P-side of the junction . Thus a space charge layer is formed near the junction having negative immbolie ions on P-side and positive immbolie ions on N-side. This region is devoided of free charge carries and is called the deplection layer. (Fig. 16.13)

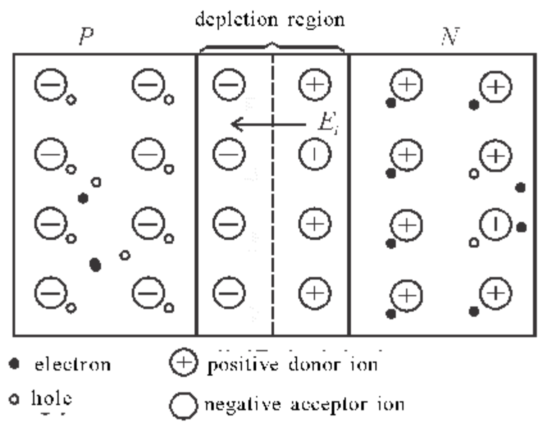


Fig. 16.13 P-N junction

Due to the ions present in the depletion layer, an electric field is generated in this layer directed from N-side to P-side. Any hole near the junction is pushed by this electric field into P-side. Likewise, any free electron near the junction is pushed by electric field into N-side. Thus diffusion current is opposed. However, this electric field supports the drift of minority electron son P-side into N-side and drift of minority holes from N-side to P-side. This electric field is called barrier electric field. Thus a drift current which is opposite in direction to the diffusion current starts. Initially, diffusion current is large and drift current is small. As the diffusion process continues, the space-charge regions on either side of the junction extend, thus increasing the electric field strength and hence drift current. The process continues until the diffusion current equals the drift current. In a P-N Junction under equilibrium there is no net current, but electric field still exists, this built in electric field causes a difference of potential across the junction of two regions. N-side is at a higher potential than P. Since, this potential tends to prevent the movement of electrons from the N-region into the P-region, it is often called a barrier potential. In general for silicon P-N Junction it is approximately 0.7 volt and for germanium PN-junction it is 0.3 Volt.

The width of the depletion layer (d) depends on the amount of impurities present in the, P and N sides. If the amount is more the width is small. This width is of the order of one micron (10^{-6} meter). Potential barrier V_B and barrier field E_i are related by the following formula.

$$E_i = \frac{V_B}{d}$$

If the width of the depletion layer of silicon P-N junction is assumed to be 1 micrometer;

$$E_i = \frac{0.7}{1 \times 10^{-6}} = 7 \times 10^5 \quad \text{V/m}$$

Any P-N junction can be connected in two ways in a circuit.

(i) When P-side is at a higher potential than N-side.

(ii) When N-side is at a higher potential than P-side.

The above situations are called forward biasing and reverse biasing respectively. In both these situations the electrical behaviour of P-N junction in the circuits is very different. We will study this in detail.

16.5.1 Forwad Biasing

When an external voltage 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 (figure 16.14); then P side is at higher potential than N-side. This is called forward biasing. In this situation, an external electric field, E is established at the junction which is directed from P to N. It is opposite to the internal electric field, E_i in the junction. In forward bias as barrier potential at the junction reduces it allow more diffusion to take place, while drift current remains unchaged. As a result, more number of holes and electrons reach the depletion layer and its width decreases. In this situation, diffusion current is more than drift current.

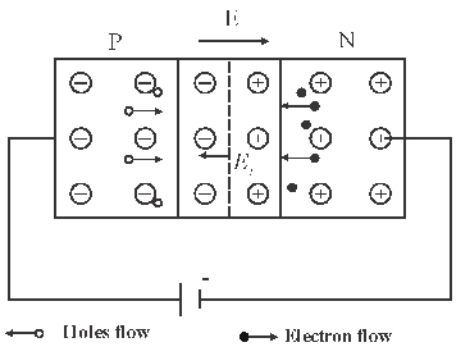


Fig. 16.14 : Forward biased P-N juction

In the P-region near the positive terminal of the battery if an electons reach this place by breaking a covalent bond, as a result of this process a hole is generated which moves towards the junction region while the electon moves through the connecting wires to reach the positive terminal of the battery. At the same time an electron is released form the negative terminal of the battery and enters N-region to move towards the junction. In this way, there is flow of current in the junction. In P-region it is due to holes and in N regions due to electrons which are the majority-charge carriers in these regions respectively. Therefore, this current is mainly due to the diffusion of majority charge carriers. It is obvious that such a bias helps in current build up and is called forward bias. The current in external circuit is due to electrons.

If the voltage applied by the battery is increased the potential barrier will reduce, more majority charge carriers will diffuse in the junction regions and the value of electric current will increase. Since, the flow of current is easy through the junction, hence, the resistance of the junction will be very small in this case.

16.5.2 Reverse Biasing

When the P-terminal of a P-N Junction is connected to the negative and N-terminal is connected to the positive terminal of the battary; then this is called reverse biasing. (figure 16.15). In this situation the external electric field E and internal electric field E_i both are in the same direction (form N side to P side). As a result, the barrier potential at the junction increases. Due to this the diffusion of majority charge-carriers from the junction is not possible as the hole in the P-region and the electrons in the N-region both move away from the junction; this increases the width of the depletion layer.

In this situation a very small current flows through the junction due to drift of minority-charge-carriers. For germanium P-N junction this current is microampere (10^{-6} A) order and for silicon P-N junction this is nanoampere (10^{-9}A) order. At normal applied voltages there is no change in the number of minority charge carriers, hence this current is constant for such potential differennce and is called reverse saturation current.

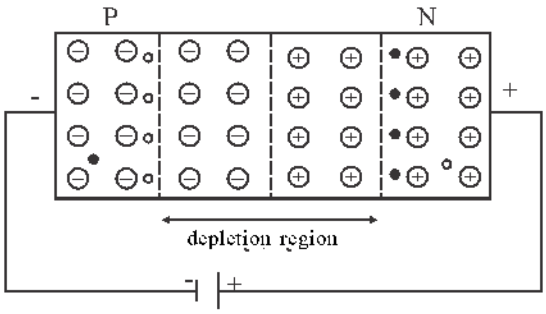


Fig. 16.15 : Reverse biased P-N junction

As a very small current exists in reverse bias condition, the resistance of PN junction is very large. If the temperature of the junction is increased then more covalent bonds break and the number of minority-charge-carriers increases. As a result reverse bias saturation current depends on temperature and increases with the increase in temperature of the junction. If the value of potential difference increases more than a limiting value then the current increase rapidly. This process is called breakdown which is discussed later.

16.6 Junction Diode and Its voltage current Characteristics

To connect a P-N junction in the circuit metallic electrodes are made at the P and N ends of the device. Hence, the device is called diode. Diode word is formed by di-electrode which means two electrodes. P-N junction diode is also called semiconductor diode.

The symbol used to represent diode in the circuit is shown in the figure (16.16).

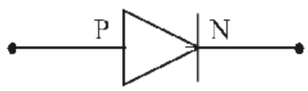


Fig. 16.16 : Circuit symbol for P-N junction diode

Here the arrow head represent P-region and bar represents N-region. The direction of arrow points from P to N representing direction of current flow under forward bias. The curves describing the variation in current through a diode with change in bias voltage are called V-I characteristics of diode. PN junction exhibits different behaviour under forward and reverse biased states. Therefore two types of characteristic curves are drawn :

(i) forward bias characteristics.

(ii) Reverse bias characteristics.

16.6.1 Forward Bias Characteristics

For obtaining the V-I characteristics of a forward biased P-N junction diode, the experimental circuit arrangement is shown in the figure (16.17).

Here the applied voltage. V on the diode is changed by the potential divider arrangement, and can be easily read by the voltmeter connected in parallel to the diode. The current I flowing in the diode corresponding to different applied voltages are noted by a milliammeter. The curve obtained for different values of V and I is shown in the figure (16.18). It is also called as the P-N junction diode forward characteristic curve.

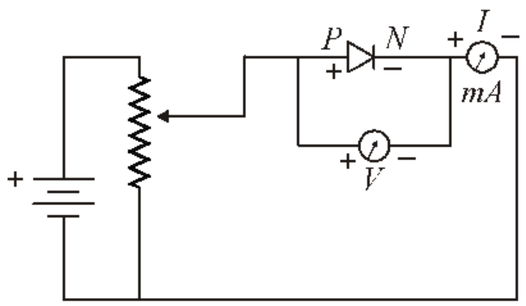


Fig. 16.17 : Experimental setup for characterstics of P-N diode in forward bias

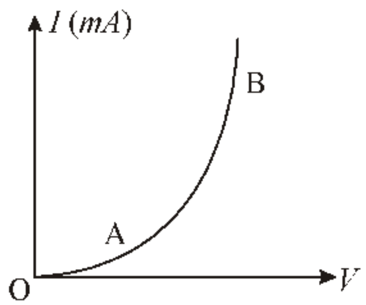


Fig. 16.18 : forward bias characterstics for P-N Junction

In forward bias, for very small values of voltage. (for Ge is 0.3 V and for Si is 0.7 V) the forward current is very small. The reason for this is that the value of applied voltage is less than the potential barrier due to which the current is very small. This behaviour of the diode is represented by the OA part of the curve. When the value of applied potential is increased the current in the diode increases exponentially. This behaviour is shown by part AB. The potential at which the current value increases rapidly is called the diode's knee voltage or cut in voltage. For germanium diode its value is approximately 0.3 V and for silicon diode its value is 0.7V.

16.6.2 Reverse Bias Characteristics

For obtaining the V-I characteristics of a reverse biased P-N junction diode the experimental circuit arrangement is shown in the figure 16.19 (a). Here with the help of a potential divider arrangement the P and N terminals of the diode are connected to the negative and positive terminals of the battery. Since reverse current is very small a miroammeter is used in place of milliammeter in this circuit. At different reverse voltages the corresponding reverse current values are noted and a curve is drawn as shown in the figure 16.19 (b).

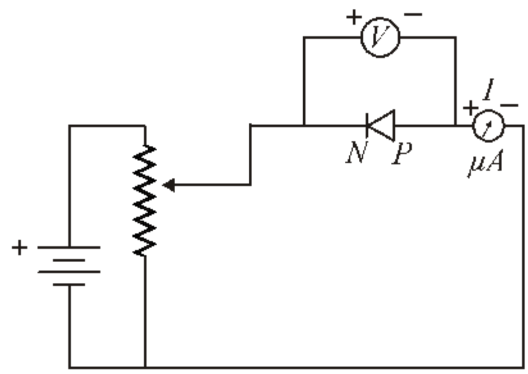


Fig. 16.19 (A) Experimental setup for characterstics of P-N diode in reverse bias

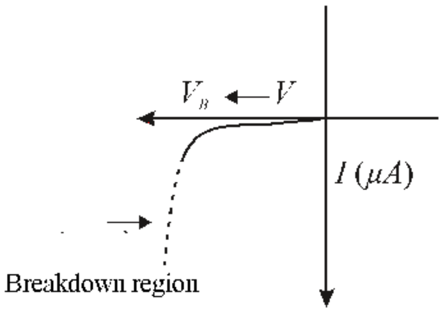


Fig. 16.19 (B) Reverse bias characterstics for P-N junction

As explained earlier in reverse bais, current is very small, as it is due to minority-charge-carriers, and remains constant till the reverse potential reaches the breakdown voltage (V_B) there after the reverse current increases very rapidly even for a slight increase in the breakdown voltage.

For a semiconductor diode the complete forward and reverse bias characteristic curve is shown in the figure 16.19.(c). If we do not consider the break down regions of this then in forward bias large current (~ mA) flows while in reverse bias small current (μA or nA) flows. As a result P-N junction diode is a unidirectional device. The characteristic curve of the diode also shows that this device does not follow Ohm's law and hence is a non-linear device.

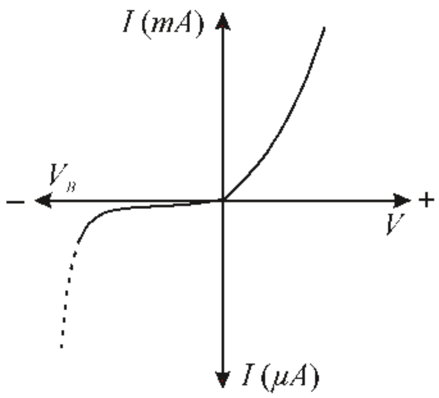


Fig. 16.19(C) : characteristic curve for P-N juction

For a pN junction diode the variation of current with voltage is given by -

$$I = I_s\left[\exp\frac{qV}{kT} - 1\right]$$

Where $q = 1.6\times10^{-19}$ C, k is Boltzman constant, T is the temperature of the junction and V is the potential difference across the junction, I_s is reverse saturation current

In forward biasing as V is positive so when;

$$\exp\frac{qV}{kT} >> 1$$

and $$I = I_s\exp\frac{qV}{kT}$$

and current increase exponentially with the voltage. In reverse biasing where V is negative then;

$$\exp\frac{qV}{kT} << 1$$

$$\therefore \quad I = I_s$$

thus under reverse bias current is almost constant.

The above equations are quite valid to expalin I-V characteristics of Ge-PN diode. The diode made of silicon follows this equation only qualitatively. Also for both the types of diodes, the diode equation is not true for reverse voltages higher than breakdown voltage.

16.6.3 Diode resistance

Diode is a nonlinear device and it does not follow the Ohm's law thus the resistance of a diode is not constant. For such types of devices dynamic resistance is a more useful quantity rather than static or d.c. resistance.

(i) Forward dynamic resistance : By definition, forward dynamic resistance is given as;

$$r_f = \frac{\text{small change in the forward bias potnetial}}{\text{Corresponding change in forward current}} = \frac{\Delta V}{\Delta I}$$

In figure 16.20 for the calculation of forward resistance, ΔV and ΔI are shown near point C If ΔV and ΔI are very small then dynamic resistance at point C would be inversely proportional to the slope of the tangent drawn at this point on the curve. Forward dynamic resistance is generally very small $(1-100\Omega)$.

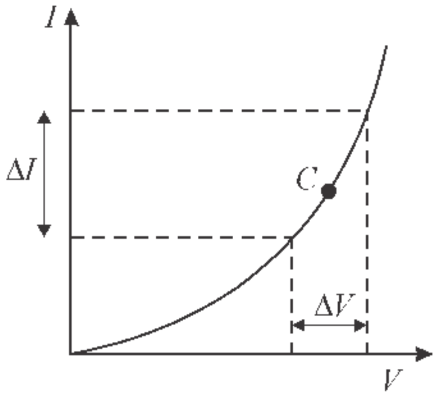


Fig. 16.20 : Forward dynemic resistant

(ii) Reverse Resistance : By definition, reverse resistance is;

$$r_r = \frac{\text{Small change in reverse voltage}}{\text{Corresponding change in reverse current}}$$

Since under reverse bias the current is very small, hence, reverse resistance is very high. This is of the order of mega ohm $(10^6\,\Omega)$.

16.6.4 Reverse Breakdown

As explained earlier if the reverse bias voltage across a PN diode is increased beyond a certain limit V_B the reverse current increases rapidly to a large value. This phenomenon is called reverse breakdown and the potential V_B at which it starts is called breakdown voltage. This breakdown can be due to one or both the mechanisms given blow-

(i) Avalanche Breakdown

(ii) Zener Breakdown

(i) Avalanche Breakdown : At large reverse voltages the minority charge carriers of P and N region get accelerated by high electric field and cross the junction. If these minority carriers acquire enough kinetic energy, from their collision with the covalent bonds new electron-hole-pair are generated. These charge carriers also get accelerated and produce new electron hole pairs. This process is cumulative due to which the number of minority charge carriers and hence reverse current increase rapily. This process is called avalanche breakdown. If the current is not controlled then due to the heat generated the diode may get damaged. If the potential is decreased before the diode gets damaged then diode current reduces and the diode again shows its normal reverse behaviour.

(ii) Zener Breakdown : This process occurs in a diode which is highly doped due to which the depletion layer is very thin. In this situation even for a small potential, the electric field at the junction is very strong. It exerts a force on the electrons of the covalent bonds, close to the junction and these bonds break. As a result minority charge carriers increase in number and the value of current increase rapidly. This process is called Zener breakdown.

The breakdown can be by due to any process but it is customary to called a diode meant to operate in break down region as Zener diode. We will study the use of zener diodes later.

Example 16.3 : For a P-N junction diode forward bias is increased form 2.0V to 2.5 V so forward current is changed from 16.5 mA to 26.5 mA. For same diode reverse bias is increased form 5V to 10 V so the reverse current changes from 20 microampere to 30 microampere. Calculate the dynamic resistance for this diode in both the situations.

Solution : According to question in, forward biasing ;

$$\Delta V_f = 2.5 - 2.0 = 0.5\ V$$

and change in forward current

$$\Delta I_f = 26.5 - 16.5 = 10mA$$

Therefore, forward dynamic resistance

$$r_f = \frac{\Delta V_f}{\Delta I_f} = \frac{0.5V}{10mA}$$

$$= \frac{0.5V}{10 \times 10^{-3} A} = 0.5 \times 10^2 = 50\Omega$$

(ii) For reverse biasing

$$\Delta V_r = 10 - 5 = 5V$$

and change in current is ;

$$\Delta I_r = 30 - 20 = 10\mu A$$

Therefore, reverse dynamic resistance.

$$r_r = \frac{\Delta V_r}{\Delta I_r} = \frac{5V}{10\mu A}$$

$$= 0.5 \times 10^6 \Omega = 0.5\, M\Omega$$

Example 16.4 : For a P-N junction diode the dynamic resistance under forward bias is 25Ω. How much change should be made in forward bias potential so that there is 1 mA change in forward current?

Solution : Under forward bies

$$r_f = \frac{\Delta V_f}{\Delta I_f}$$

$$\Delta V_f = r_f \Delta I_f$$

$$= 25 \times (1 \times 10^{-3})$$

$$= 25 \times 10^{-3} mV$$

$$= 25 mV$$

Hence, there should be a change of 25 mV.

16.7 Use of a P-N Diode as Rectifier

We know that the generation and transmission of alternating current is more simple and less expensive than the generation and transmission of direct current. Due to this generally alternating current is used. But in many electronic instruments direct current or potential is also used. The direct current is obtained form cell or battery but in practice alternating current is converted into direct current and used in electric devices.

Rectification is the process by which alternating current is converted to direct current. The device used for this purpose is called rectifier. As stated earlier, a P-N junction diode allows current in forward bias and no current (negligible current) flows in reverse biasing. This unidirectional property of diode is used for rectification.

16.7.1 Half-Wave Rectifier

Figure (16.21) Shows the circuit of a half wave rectifier. In this an input alternating voltage source is connected to the primary coil of a transformer. The secondary coil of the transformer, is connected to a diode and load resistor R_L in series between point A and B. From the secondary coil of the transformer a desirable alternating voltage is obtained between point A and B. This alternating voltage can be more or less than the input alternating voltage. In case of a step up transfer V_S is more than in put V_{in} while for step down transformer V_S is less than V_{in} step up nature then V_S.

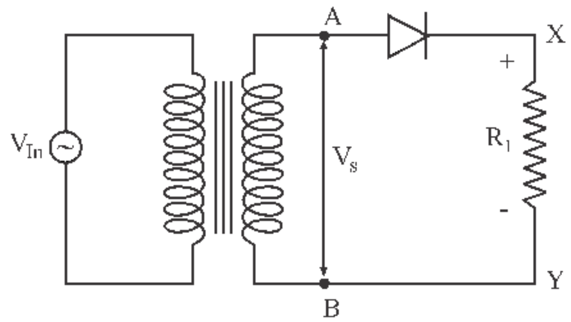


Fig. 16.21 : Half wave rectifier circuit

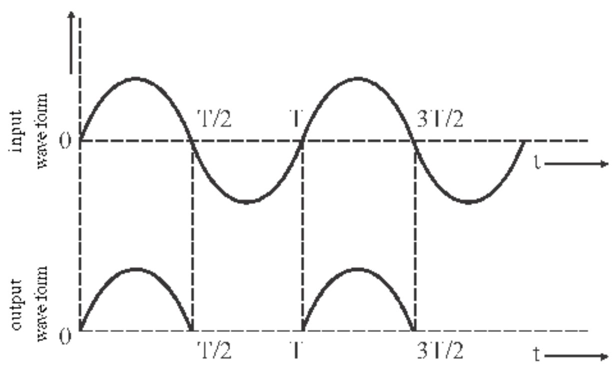


Fig. 16.22 : Wave form of output & input half wave rectifier

For positive half cycles of alternating potential V_i (or V_s) the end A of the secondary coil is at positive potenitial and end B is at to negative potential due to which the P terminal of the diode is at a higher potential than N terminal. Thus this diode will be conduction and allow current. In this case there will be a potential drop across the load resistor R_L. In the negative half cycle of input ac end A is at negative and end B is at positive potential, due to which diode is reverse biased. In reverse biasing the current is negligible hence diode is in non conducting state there will be no current in the circuit. In load resistor R_L the current flows only in one direction that is from A to B hence, potential drop at the ends of R_L is of dc nature.

However this potential drop across R_L is pulsating. It changes in magnitude but not in direction. In the above precess the negative half cycle of input ac is supressed by diode, and only the positive half cycle is obtained as dc. Hence, this circuit is called half wave rectifier. The above process is repeated for consecutive cycles of alternative voltage. Input ac voltage and output voltage waveform from the rectifier are shown in the figure (16.22).

In half wave rectifier only half cycle of input a.c if used remaining half cycle remains unused. In full wave rectifier the full cycle of the input ac is utilised. Hence, due to this the efficiency of half wave rectifier is half of the full wave rectifier. Therefore generally half wave rectifier is not used. Here, it should also be noted that in half wave rectifier the use of transformer is optional. The input voltage can be directly applied to diode load resistor combination.

16.7.2 Full Wave Rectifier

In a full wave rectifier circuit the output voltage is obtained for both the half cycles of input alternating voltage. Here two junction diodes are used such that one diode allows the positive half cycles of input ac and the other allows the negative half cycles. Figures (16.23) shows a circuit of full wave rectifier.

Input ac is applied ascross the terminals of the primary coil of a centre taped transformer.

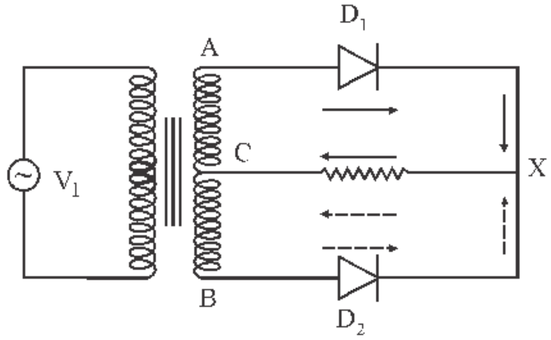


Fig. 16.23 : Full wave rectifier circuit

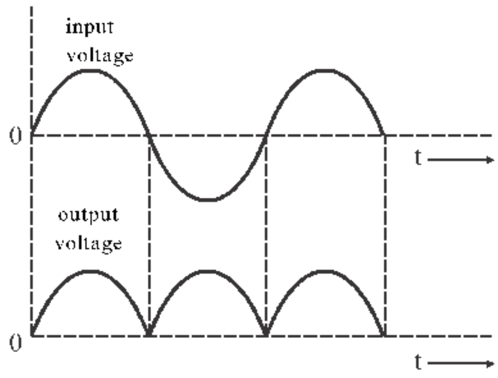


Fig. 16.24 : Wave form for input & output voltage for full wave rectifier

The P-terminals of the diodes is connected to the terminals A and B of the secondary of transformer. The N-side of the diodes are connected together and the output is taken across load register R_L connected between this common point of diode and the centre tap C of the secondary of the transforemer. The point C is at reference zero potential. Thus if a given instant if one terminal of secondary is positive with respect to point C, then other terminal will be negative. Let for the positive half of input ac terminal A is positive and B is negative. So diode D_1 gets forward biased and conducts (while D_2 being reverse biased is not conducting). Hence, during the positive half cycle we get an output current (and a output voltage across the load resistor R_L). During negative half cycle the voltage at A becomes negative with respect to centre tap, the voltage at B would be positive. In this part of the cycle diode D_1 would not conduct but diode D_2 would give an output current and output voltage (across R_L) ac. Note that for both halves of input, current in R_L is directed from X to C i.e. it is unidirectional. Thus, we get output voltage during both the positive as well as negative half of the cycle. Obviously, this is more efficient circuit for getting rectified voltage or current than the half wave rectifier current in R_L is from X to C.

For a full wave rectifies a centre tap transformer is necessary. Here an ordinary transformer cannot be used. If the current is to be obtained at high voltage then this transformer should be step up transforms and if the current is to be obtained at low voltage then the transformer should be step down transformer.

16.7.3 Full Wave Bridge Rectifier

In such type of full wave rectifier a centre tap transformer is not necessary but 4 (four) diodes are used instead of two diodes as shown in the figure (16.25) in the form of a bridge in the circuit. The input alternating voltage is obtained from the secondary coil of the transformer.

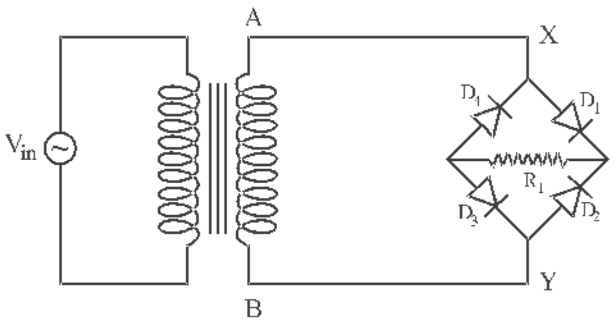


Fig. 16.25 : Bridge rectifier circuit

In the positive half cycle of the input voltage the terminal A of the secondary coil is positive and B is negative. Hence, diodes D_1 and D_3 are forward biased and diodes D_2 and D_4 are reverse biased. Current flows in AD_1XYD_3BA direction. There is no conduction is diodes D_2 and D_4.

In the negative half cycle of input voltage terminal A of the secondary coil is negative and B is positive. then diodes D_2 and D_4 are forward biased and diodes D_1 and D_3 are reverse biased. Now the current flows in directions BD_2XYD_4AB/D_2 and D_2 do no conduct.

In this way, in bridge rectifier at any instant only two junction diodes help in the flow of current and remaining two diodes do no conduct. But for every cycle the current flows from X to Y in load resister. Hence, this is unidirectional and the out voltage on R_L is rectified. As described earlier in place of centre tap transformer an ordinary transformer is used hence in comparison to full wave rectifier the bridge rectifier is less expensive which is an advantage.

The diodes used in rectifier circuit should be such that when they are not conducting i.e. when they are reverse biased then there should be no breakdown of diodes otherwise they would also be conducting and the process of rectification will not be possible. Hence, the rated potential of diodes should be higher than that of the maximum value of alternating voltage of secondary coil. These diodes are also called power diodes.

By the analysis of the waveform of output voltage (current) obtained from full wave rectifier it is clear that this potential is unidirectional but not constant, it is called fluctuating rectified potential whereas by definition pure dc voltage is constant in magnitude. From mathematical analysis of waveform of full wave rectifier, it is known that this type of changing potential is superposition of various harmonic frequencies and a pure dc component.

Using some special circuits, which are called smoothing filters; pure dc compoment is obtained by isolating the rectifying potential from alternating components. In such type of filters capacitors, inductors or their combination are used are because their behaviour is different for direct (rectified) potential and alternating potential. One such filter is shown in figure (16.26).

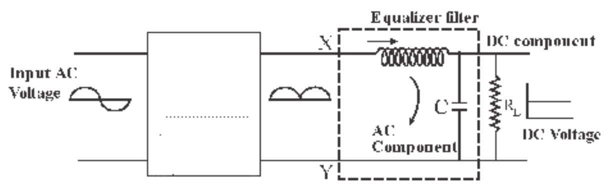


Fig. 16.26 : Full wave rectifier using equalizer filter

When rectifier's output voltage is applied between the terminals X and Y of the filter circuit then the due to the reactance offered by inductor L ac component of rectified output is opposed while dc component is not affeated ac component is further by passed by capacitors C connected in parallel with load resistor R_L. It is so because capacitor afters low reactance to ac components compared to the resitance offered by R also capacitor blacks dc so, it will not pass through C. Thus nearly pure dc potential will be obtained across R_L.

16.8 Special purpose diodes

Apart from rectification, P-N Junction diodes have other uses also. The diodes used for some specific purposes are different from the diodes used for some other specific purpose, in properties, like semiconductor, used used, amount of impurities and construction, etc. Now we will study about some special purpose diodes.

16.8.1 Zener Diode

We have seen that for a PN diode at reverse breakdown voltage the current rises Sharply however the potential difference across the diode remains practically constant at its breakdown voltage. The breakdown voltage for a diode depends on amount of doping. Thus depending upon the doping levels, diodes having specified breakdown voltage can be fabricated. Such diodes are called Zener diodes. Nowadays diodes having breakdown voltage from 1 V to several hundred volts are available. The circuit symbol for a Zener diode is shown in fig. (16.27).

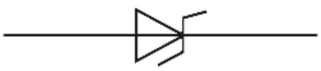


Fig. 16.27 : Symbol of zener Diode

As the voltage across a diode remains constant at its breakdown while current increases. This property makes it usefull for voltage stabilization and voltage regulation. We know that in an ordinary dc power supply consisting of rectifier and filters, ac is changed to dc. In such type of power supplies when there is a change in load current the value of dc voltage also changes. Therefore, in such situations where current is to be obtained at constant dc voltage such a power supply cannot be used. With the help of a zener diode the output voltage of a power supply can be kept constant.

Figure (16.28) shows a power supply whose output voltage is assumed to be V_i and a resistance R_s and a zener diode are connected across its output terminals. The load resistor R_L is parallel to the zener diode. The zener diode is so selected whose breakdown voltage V_Z is equal to the constant value of that dc voltage to be obtained at load resistor R_L. Zener diode is connected in the circuit in reverse biased condition. If the value of V_i is more than V_Z then the diode comes in breakdown region and the potential across its terminals remains V_Z. Since R_L is in parallel to the diode, hence potential across is also V_z. Even if the value of R_L is changed potential difference across it is still V_Z. The resistance R_S is so selected that the diode is not damaged by the thermal energy due to the increase in the zener current. This value of current is known from the data sheet given by the manufacturing company.

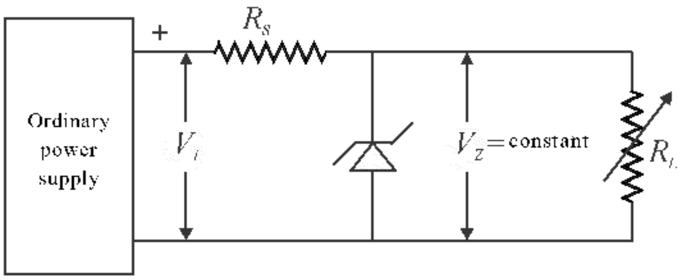


Fig. 16.28 : Voltage regulation by zener diode

16.8.2 Photo Diode

When light (electromagnetic radiation) of suitable frequency is incident on a semiconductor, it can be absorbed by semiconductor. The energy so recieved may cause transition of an electron from the valence band to conduction band, resulting in generation of an-electron hole pair. These electron-hole pairs are in addition to thermally generated electron-hole pairs. Obviously in the process conductivity of semiconductor increases. This increase in conductivity is called photo conductivity and phenomenon is called photo conductive effect. If the frequency of incident photos is v and their energy h_v is greater than or equal to the band gap energy Eg (i.e. $h_v > E_g$) then photo conductivity can be observed.

Photo diode is a PN junction whose function depends on the phenomenon of photo conductive effect. In such diodes one of the regions either P or N is made relatively thin so that the light photons incident on it are able to reach junction region.

Normally photo diodes are operated under reverse bias. The circuit symbol for photo diode is shown in Fig. (16.29). In the absence of light current through a reverse biased photo diode is small ($\sim N\mu A$). When light is incident photons absorbed near the junction create new-electron-hole pairs. These carries get separated due to electric field present in deplection region and cross the junction. Thus reverse current increases. The increase in reverse current is much larger than current under dark (no light) condition. While keeping the frequency constant if the intensity of light is increased reverse current increases more. Thus current is circuit is controlled by the intensity of the incident light. The effect of intensity level (ϕ) on reverse current is shown in fig. 16.31.Current in presence of light may be upto few hundred micro ampere. It is necessary to keep reverse voltage well below the breakdown voltage. Measuring the change in reverse current due to change in intensity of light, can be use to measure the intensity of light.

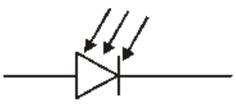


Fig. 16.29 : Circuit symbol for photo diode

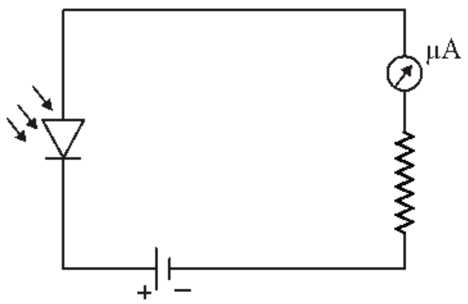


Fig. 16.30 : Photo diode for light detection

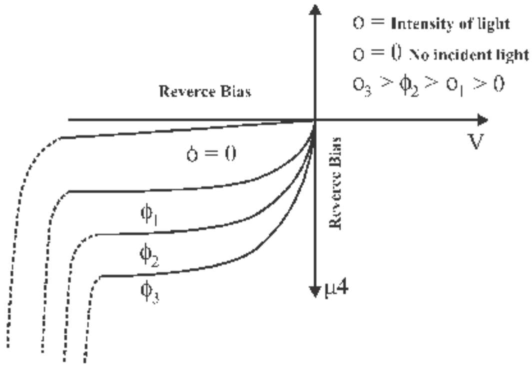


Fig. 16.31 : Effect of light intensity on reverse current voltage of a photo diode

Photo diode is used in following devices :

(i) Light detection.

(ii) Light operated switch.

(iii) Reproduction of sound in films.

(iv) For reading computer tapes and computer cards.

16.8.3 Light Emitting Diode (LED)

Light emitting diodes are those P-N junction diodes, which under forward bias, emit light. When an electron present in the conduction band of a semiconductor make transition to the hole present in the valence band, recombination of electron hole take place. In this process energy is released. Generally the electrons are at the minimum energy level of the conduction band (E_C) and the holes are at the maximum energy level of the valence band (E_V). Hence, energy relased in the process is E_C E_V − E_g equal to the forbidden energy gap. In many semiconductor like silicon and germinaum, this energy is released in the form of heat (thermal energy) but in gallium arsenide phosphide (GaAsP) and gallium phosphide (GaP) like semiconductors, this energy is released in the form of visible light.

To obtain visible light of good intenaity electron-hole recombination events must occur in large number. But for intrinsic and extrinsic both types of semiconductors such recombination events occur in limited number. This is due to the reason that in intrinsic semiconductors the number of electrons hole pairs is quite small and in extrinsic semiconductors one type of charge carrier are in majority and the other type of charge carriers are in minority.

In semiconductors like GaP if the P and N region of the junction are heavily doped, so the depletion region is very thin. If the junction is forward bias then the large number of holes of P region move towards N region and the large number of electrons of N region move towards P region. Hence, near the depletion layer there would be a large number of recombinations of electrons and holes. As a result the emitted light will be of high intensity. The figure (16.32) shows the symbol for a light emitting diode.

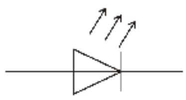


Fig. 16.32 : Symbol of LED

The semiconductor for LED is chosen as per the desired colour of light. At present the LED's of red, green, yellow, orange and blue colour are available.

The LED is used for following purposes :

1. In the form of light indicator.

2. In the form of seven segment display unit.

3. LED generating high intensity light are used in optical fibre communication.

4. High intensity light using LED bulbs.

16.9 Transistor

A transistor is a semiconduce device having three layers and two P-N junction with three terminals for connection to an external circuit. A transistor is capable of amplification of ac signals. The transistor was invented in the year 1948 in USA (United states of America) at Bell telephone laboretory by Bardeen, Brattain and Shockley. For this they were honoured by the nobel prize in the year 1956 in physics. The semiconductor devices like detectors and rectifiers were being used before 1948 but the invention of transistor has laid the foundation of present era of semiconductors electronics. Today many types of transistors for example, junction transistor, field effect transistor or MOSFET are available. Here we will study only about junction transistors. You will study about other transistors in higher classes.

16.9.1 Junction Transistor

A simple junction transistor is generally a single crystal of an extrinsic semiconductor (silicon or germanium) in which three regions of different conductivities are present. The width of the middle region is less then the other two regions and the type (N or P) of the semiconductor in this region is different than the other two regions. In this way we obtain two type of junction transistors which are called PNP and NPN transistors respectively. In a PNP transistor, there are two P-types semiconductors a thin layer of N type semiconductor is sandwiched between them. Similarly in an NPN transistor there are two N types semiconductor regions and a very thin layer P-type semiconductor is sandwiched between these two.

The middle region in both the types of transistors is called the base (B) and out of the other two regions one is called the emitter (E) and the other is called the collectors (C). [Figure (16.33)]. Connected to all the three regions there is a metallic electrode and a lead through which the transistor can be connected to external electric circuit. The leads are named as E lead, B lead and C lead.

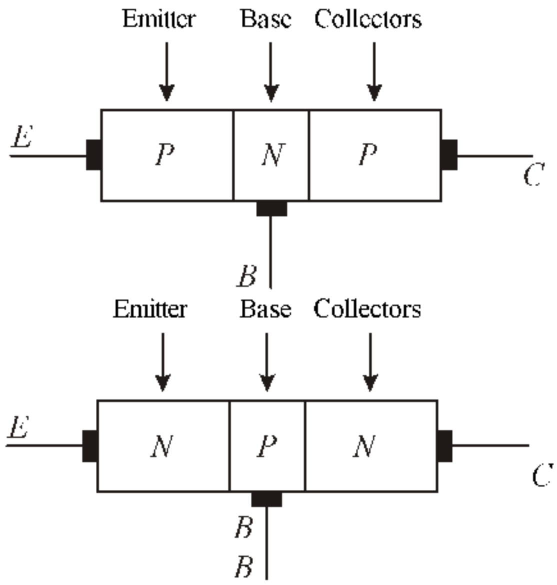


Fig. 16.33 : Construction of PNP & NPN transitor

The emitter and collector are made of the same type of semiconductor (P or N) but the amount of doping in them is different; also the collector region is bigger in size than the emitter. Due to differences in physical and electrical properties and because each region has its own specific purpose emitter and collector cannot be interchanged with each other while using a transitor.

Emitter is heavily doped because its purpose is to provide majority charge carrier in large numbers to the base. There is very little doping in base; also it is of very small width so that it does not provide opposite type of charge carrier in large number for recombination to the majority charge carriers coming from emitter. The purpose of the collector is to collect the majority charge carriers which pass through the base. Here the doping is less than the emitter but more than the base i.e it is moderatly doped.

The area of collector-base junction is larger than the that of emitter-base. Due to this the collection of the charge carriers take place very well and the transistor when used as an amplifier, this large area helps in quick dissipation of the generated heat. There are two P-N junctions in these transistors which are called E-B junction (emitter-base junction) and B-C junction (base-collector junction). Theoretically, both these junction can be considered to be connected back to back. (Figure 16.34).

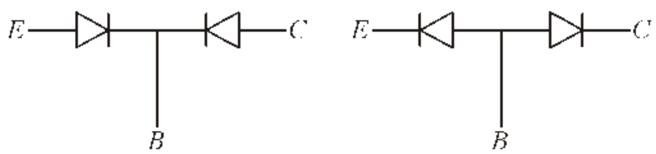


Fig. 16.34 : Assumption of transistor diode junction

However, it is worth noting that when two P-N junction are joined according to the figure (16.34) then a transitor is not obtained. In transistor, both these junctions are in present in the same crystal, whereas two diodes can not be joined to make a single crystal. The width of the base of the transistors is of micrometer order, hence both these junctions are very close.

When Transistor is used as an amplifier the emitter-base junction is always forward biased and base-collector junction is always reverse biased. Due to this the majority charge carriers always flow from the emitter towards the base, but the current flowing in this junction can be from emitter to base (E to B) or base to emitter (B to E) depending upon the nature of majority carriers. These two different directions of the electric current are used to differentiate between the symbols of PNP and NPN transistor. These symbols are shown in the figure (16.35).

The line segment having an arrow sign represents the emitter, the middle line segment represents the base and third line segment represents the collector. The arrow head shows the direction of current. Since in forward bias the emitter base junctionof the PNP transistor the majority holes from the P-type emitter move towards the N-type base hence, the current will flow in the junction from E to B. Therefore, in the symbol of PNP transistor the arrow head is pointing from E to B. Similarly, in NPN transistor in forward bias the majority electrons move form N-type emitter to P-type base hence the current will be in the direction from B to E. Therefore, in the symbol of NPN transistor the arrow head is pointing from B to E.

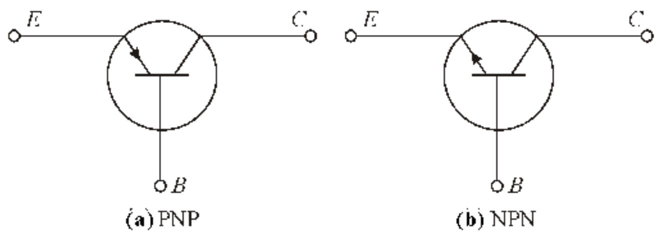


Fig. 16.35 : Symbol for junction diode

For biasing of two P-N junctions of a transistor following four possibilities are there:

1. When emitter-base junction is forward biased and base-collector junction is reverse biased; for such biasing, transistor operates in active region. Generally, transistor are biased in the this way.

2. When emitter-base junction and base-collector junction both are forward biased in such situation the transistors is in saturation region.

3. When emitter-base junction is reverse biased and base-collector junction is forward biased then transistor is in cut off region.

4. When emitter base junction in reversed baised on a collector base junction is reverse bias, transistor is considered in inverted mode.

From the above biasing possibilities only first i.e active bias is used for the working of a transistor as an amplifier. Hence, we make use of only this possibility to understand the transistor action. The other possibilities are used in switching and other circuits, which are not studied here.

16.9.2 Operation of Transitor

For any transistor to work properly E-B junction is forward biased and B-C junction is reverse biased. In this case the transistor is said to be in active state. Firgure 16.36 (a) and (b) shows active bias of PNP and NPN transistors. In both the figures the depletion regions corresponding to both the junctions are shown.

Since, E-B junction in forward bias and the doping in emitter is high, E-B junction will be narrow. Whereas B-C junction is reverse biased hence it is broader. The forward bias voltage V_{EB} at E - B Junction is small (0.5–1V) and the reverse bias voltage V_{CB} is more for B – C junction (5 to 15 V).

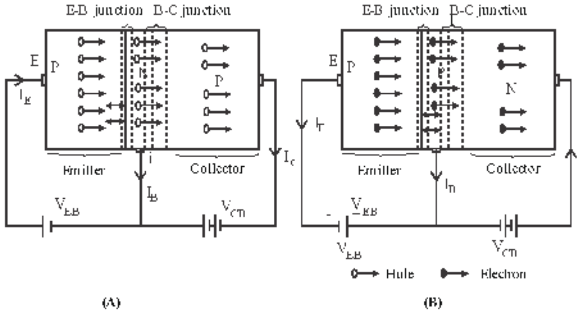


Fig. 16.36 Biasing of PNP and NPN transistor for its active operation

Now we consider the PNP transistor shown in the figure 16.36 (a). Since, E - B junction is forward biased hence majority charge carriers 'holes' from emitter will diffuse through this junction to base in large numbers. This is known as the injection of the holes from the emitter to the base. Similarly from the N-type base the electons cross the junction to reach emitter. The motion of both charge carriers are in opposite direction but the current related to them flows from E to B i.e from E to B. This current is called emitter current I_E which is due to both the holes and electrons. But due to less doping in base region of transistors this current is basically due to holes, for a PNP transitor.

The tendancy of the injected holes from the emitter to the base is of recombination with the electrons present in the base. But as the base is thin and lightly dopped very small (less than 5%) recombination will take place, and majority of holes pass base collector junction are able to reach the collector. Because the collector terminal is negative, these holes reach easily to this terminal to constitute collector current I_C.

Some holes injected form the emitter to the base, recombine with the electrons in the base and generate base current I_B. For every electron lost in the process of recombination an electron is supplied by negative terminal of battery V_{EB} connected to base. Therefore, here base current I_B flows outward from the base terminal B. For a PNP transistor the direction of currents I_E, I_B and I_C are show in the figure 16.36 (a).

If Kirchhoff's current law is used for transitor as a whole then, it is clear that emitter current will be equal to the sum of base current and collector current,i.e

$$I_E = I_B + I_C \qquad ....(16.14)$$

Here $I_B << I_E$ and $I_B << I_C$

Since, E - B junction is forward biased hence its forward resistance, is very small, and B - C junction is reverse biased hence its reverse resistance is large. Due to this it is may appear that emitter junction current I_E would be much more than the current collector I_C. But from equation (16.14) as I_B is very small hence; I_C

$\approx I_E$. So from the view of operation, transistor is a device which transfers the current I_E from a very low resistance (forward biased EB junction) to the high resistance (reverse biased B_C junction) keeping the same value I_C ($I_C \approx I_E$). For this transfer process of current this device is named by the conjuction of two words transfer + resistor and abbreviated as transistor.

In other words, it can also be said that in for current conduction of transistor operation, the effect of voltage at B - E junction is very high on the collector current. If the value of V_{EB} is high then both emitter current and collector current are more.

The above description can also be used for the active operation of NPN transistor. In this transistor majority charge carriers are electron; which when injected form the emitter of N-type, diffuse through the P type base and reach the N type collector to produce collector current I_C. Figure 16.36 (b) shows the directions for I_E, I_B and I_C for an NPN transistor.

The emitter current in a NPN transistor is mainly due to electrons. Although both PNP and NPN transistors are used but due to the high mobility of electrons. **NPN transistors are more effective** in high frequency circuits.

16.10. Transistor Circuit Configurations :

Generally, the electronic circuits are four terminal networks in which two terminals are used for input signals and remaining two terminals are used for output signals. In a junction transistor there are only three terminals emitter (E), base (B) and collector (C). Hence, in such type of circuits, the transistor is so connected that out of the terminals E, B and C one terminal is common for both input and output. In this way to connect the transistor in the circuit following there configurations, are used.

1. Common Base (CB) Configuration.
2. Common Emitter (CE) Configuration.
3. Common Collector (CC) Configuration.

In any transistor circuit, all voltage can be taken relative to the earthed common terminal.

Figure (16.37) shows circuit diagrams for these configurations of transistor. In each circuit, the emitter base junction is forward biased and base collector junction is reverse biased.

16.10.1. Transistor Characteristic Curves

The graph representing the change in the current in the input and output circuits of the transistor with the applied voltages are called the characteristic curves of the transistor. When only direct current flows in the circuit and there is no load resistance connected between the output terminals then such curves are called static characteristics curves. When there are alternating currents in the circuit and there is load resistance connected between the output terminals then such curves are called dynamic characteristic curves. With the help of characteristic curves, the construction of various circuits of transistors is possible. Here we will study only about static characteristic curves. Generally, two types of

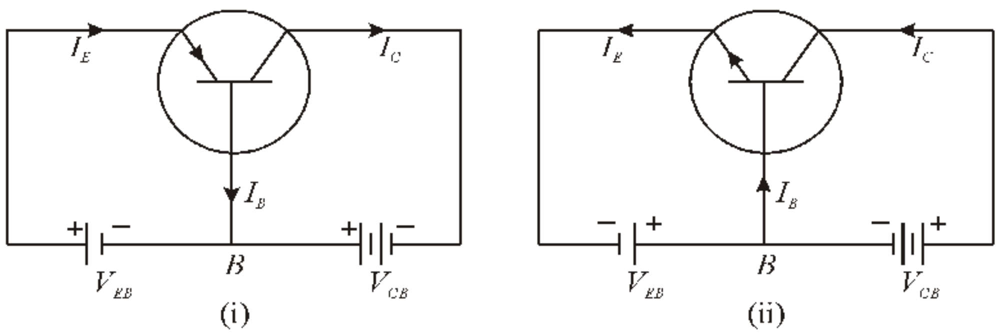


Fig. 16.37 (a) : CB circuit (i) PNP (ii) NPN

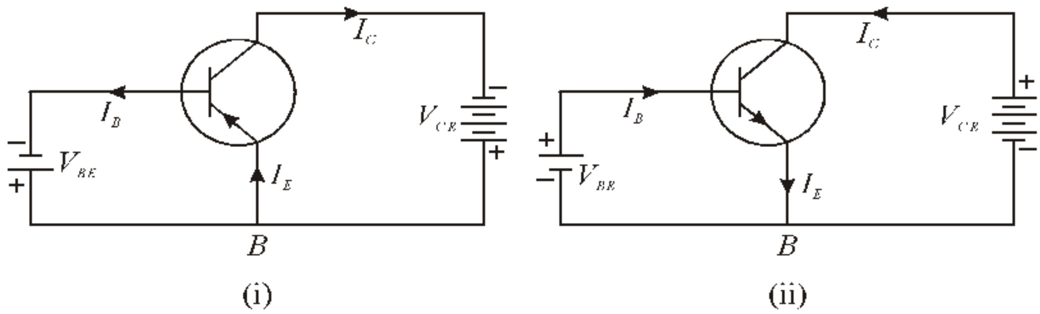


Fig. 16.37 (b) : CE circuit (i) PNP (ii) NPN

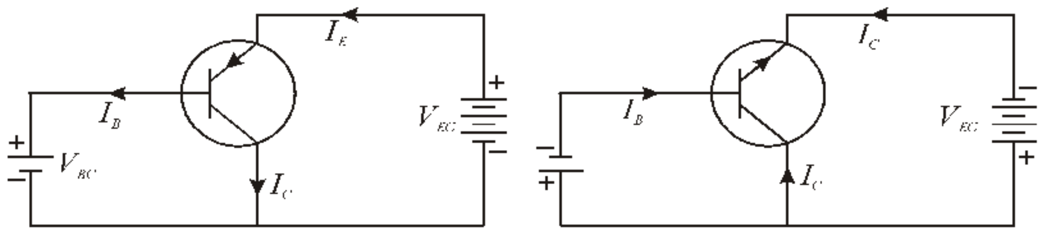


Fig. 16.37 (c) : CC circuit (i) PNP (ii) NPN

characteristic curves are useful.

Input Characteristic : By keeping the output voltage of a transistor constant, a graph drawn between input current and input potential difference is called input characteristic curve. For various constant input voltages many such curves are drawn. A family of these curves is called characteristics.

Output Characteristic : For constant input current a graph drawn between output voltage and output current is called the output characteristics curve. For various constant input currents a family of such curves is obtained called output characteristics.

In all the three configurations of the transistor there is difference between the characteristics and operation of a transistor. Hence, they are studied separately. Here, we will study only about common emitter and common base configurations.

16.10.2 Common Base Configuration

In this configuration (figure 16.37(a)) the base terminal of the transistor is common for input and output. The potential difference between the emiter and base is called the input voltage. Where as the potential difference between collector and base is called output voltage. Emitter current I_E is called input current and collector current I_C is called the output current. The arrangement shown in the figure (16.37 (a)) is the basic arrangement of this configuration. Fig (16.38) shows the circuit used for obtaining the characteristic curves of the transistor experimentally. In the figure shown a PNP transistor is used. By making a proper change in biasing arrangement a similar circuit can be made for obtaining NPN transistor characteristic curves.

In the figure (16.38) B - E junction is forward biased by battery V_{BB} and B - C junction is reverse biased by battery V_{CC}. Since, we have to see the effect on I_E and I_C by changing the applied voltage V_{EB} and V_{CB}. Hence, the potential divider R_1 and R_2 are employed with V_{BB} and V_{CC}.

Voltmeters and milli ammeters are used for the measurement of V_{EB} and V_{CB} and I_E and I_C respectively in the circuit. With the help of this circuit the input and output characteristic curves are obtained as follows.

Input Characteristics : Here the output voltage V_{CB} is kept constant and change in input current I_E is measured relative to the input voltage V_{EB}. For this with the help of R_2; V_{CB} is kept constant at some value. After this with the help of R_1, V_{EB} is varied starting from zero in descrete step (for example in the range of 0-5 volt). The corresponding values of I_E are measured with the

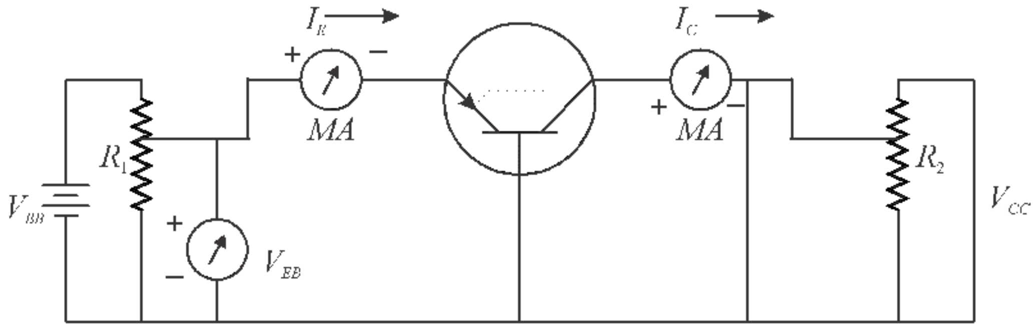


Fig. 16.38 : Circuit of obtaining transitor charactcrstics of a common base PNP transitor

help of a milliammeter. In this way at a constant value of V_{CB} a graphs is drawn for the various value of V_{EB} and corresponding I_E which is shown in the figure (16.39). The same process is repeated for other constant values of V_{CB}. It should be noticed that due to reverse biasing the value of V_{CB} is negative.

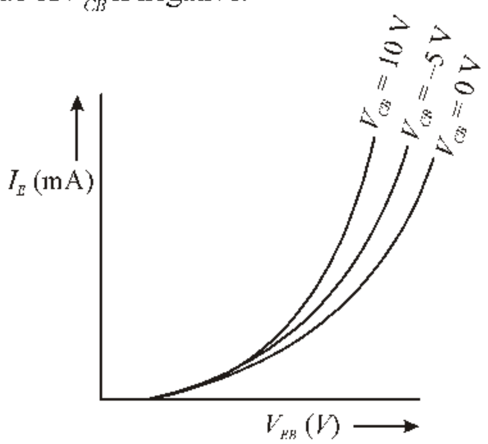


Fig. 16.39 : Input characteristic curve for CB transitor

The form of a input characteristic curve is similar to the graphs of forward characteristic of a P-N junction diode. Since, E - B junction is forward biased hence this is expected. Initially when V_{EB} is zero the current is also zero; hence the curve starts from origin. When the value of V_{EB} is initially increased then increases in I_E is negligible; but after a certain value of V_{EB} called the threshold voltage of the transistor; the current increases rapidly. The reason for this is also as explained for PN junction. According to the nature of the transistor its value is between 0.1 V to 0.5 V.

The dynamic input resistance for common base configuration of the transistor is given by -

$$R_{ib} = \left|\frac{\Delta V_{EB}}{\Delta I_E}\right|_{V_{CB}=\text{constant}}$$

The subscribt V_{CB} = constant indicates the curve for which the value of R_{ib} has been calculated.

The value of R_{ib} is between 50-100Ω.

Output Characteristics : Here input current I_E is kept constant and the changes in output current I_C is measured for various output voltage V_{CB}.

Initially, with the help of R_1, I_E is kept at a desired value. Now, with the help of R_2, V_{CB} is changed starting from zero in discrete steps, corresponding values of I_C are measured with the help of a milliammeter.

The graph drawn between V_{CB} and I_C is the output characteristic curve. The same process is repeated for other constant values of I_E and group of curves so obtained is called the output characteristic. Due to the reverse biasing of C-B junction for these graphs V_{CB} and I_C both are negative. By the study of these graphs the following facts are noticed :

(i) For $I_E = 0$, $V_{CB} = 0$, then I_C is also zero. But for $I_E = 0$ and non zero values of V_{CB}; a small I_C value is obtained. This is due to the reason that for $V_{CB} \neq 0$ in B-E junction is reversed biased and a negligible reverse current (~ μ A) flows. The graph obtained for $I_E = 0$ is similar to the graph for reverse biased diode.

(ii) For other value of I_E ($I_E \neq 0$) at $V_{CB} = 0$ value of I_C is not zero. When the value of V_{CB} is increased from zero; for very small values of V_{CB}, I_C first increases then becomes almost saturated. In this situation graph is parallel to V_{CB} axis.

(iii) When the increases I_C increase of I_E. In saturated state the value of I_C is little less than the of I_E.

(iv) At non-zero values of I_E to make I_C zero, it is necessary that the polarity of V_{CC} is changed for V_{CB} and C-B junction is made forward biased in place of reverse biased. In such a case V_{CB} is positive.

When V_{CB} is made positive from zero; then at very low value of V_{CB}, I_C will be come zero. For various values of I_E in this forward biased state of collector-base junction the transitor is said to be working in saturated region. Figure (16.40) shows this.

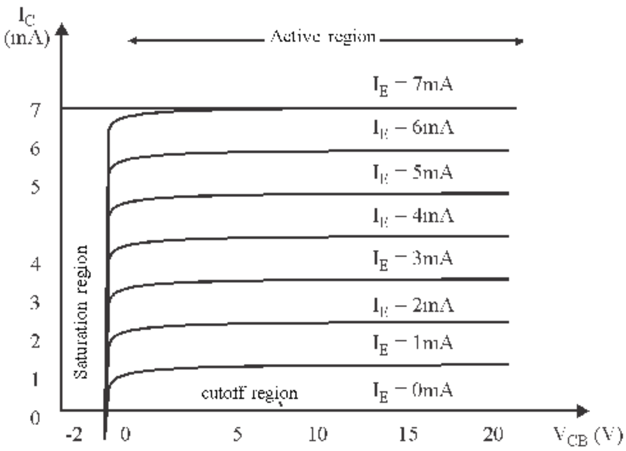


Fig. 16.40 : Output characterstics of CB transitor

The dynamic output resistance R_{ab} for common-base configuration is defined as follows:

$$R_{ab} = \left|\frac{\Delta V_{CB}}{\Delta I_C}\right|_{I_E=\text{constant}}$$

It is clear from output characteristic curve; that for a large change in V_{CB} there is small change in I_C. Due to this dynamic output resistance is high of the order of $10^6\Omega$.

Inverse of output resistance is called output conductance.

16.10.3 Current Gain or Current Amplification Factor for Transistor in CB Configuration

In Common base configuration, at constant collector base voltage, the ratio of collector current I_C (output current) to the emitter current I_E (input current) is called as static current gain or static current amplification factor. It is denoted by α_{dc}.

$$\alpha_{ac} = \left.\frac{I_C}{I_E}\right|_{V_{CB}=\text{constant}} \qquad ....(16.15)$$

At constant V_{CB}, if a small change ΔI_E in input current I_E and a small change ΔI_C in output current I_C, then dynamic current amplification factor is represented by α_{ac} or α

$$\alpha_{dc} = \left.\frac{\Delta I_C}{\Delta I_E}\right|_{V_{CB}=\text{constant}} \qquad ....(16.16)$$

Since, $I_C < I_E$

and $\Delta I_C < \Delta I_E$

So, α_{dc} and α_{ac} are nearby 1 but less than 1.

Usually, $0.9 \le \alpha_{ac} \le 0.99$

16.10.4 Common Emitter Configuration

In this configuration emitter (E) is common in both input and output circuits. The potential between base and emitter is called potential difference input voltage whereas potential difference between collector and emitter is called output voltage. The basic arrangement for this configuration is shown in the figure 16.37 (b). Here base current I_B is input current and collector current I_C is output current.

For a PNP transistor the circuit for is common emitter characteristic curves shown in the figure (16.41). With the help of battery V_{BB} and the potential devider arrangement R_1 forward bias V_{BE} is provided to the base-emitter junction. the measurement of V_{BE} is done by the voltmeter connected across base and emitter terminals.

Here, the value of base current I_B is very low (~ μ A) hence it is measured by micro-ammeter. The potential difference V_{CE} between the collector and emitter is pro vided through the battery V_{CC} and potential divider R_2 arrangement. It is measured by the voltmeter and collector current I_C is measured by milliammeter.

Here, it is natural to ask that how does the base-collector junction is reverse biased, since there is no battery connected between them? To answer this question see figure (16.41) carefully. Here, both base and collector are connected to the negative terminals of batteries of repective circuits (V_{BB} and V_{CC}). If the magnitude of V_{CE} is higher then V_{BE} then base (N-type) will be at negative potential compared to collector (P-type). In other words, N-type base is at a higher potential than P-type collector, clearly this junction will be reverse biased.

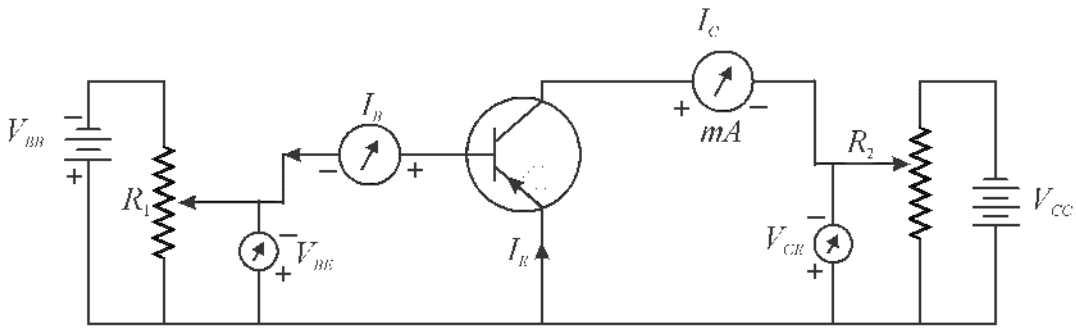


Fig. 16.41 : Circuit of obtaining transitor characterstics of a common emitter PNP transitor

Input Characteristics: Here output voltage V_{CE} is kept constant and the changes in input current I_B are studied related to the change in input voltage V_{BE}. For this with the help of R_2, V_{CE} is kept at a constant value. Now with the help of R_1, the value of V_{BE} is increased from zero in discrete steps and corresponding values of I_B are noted. A graph is drawn between I_B and V_{BE}, the same process is repeated for other values of V_{CE}. The graphs or curves so obtained are called the input characteristics of common emitter configuration. Figure (16.42) shows the characteristic curves for three constant values of V_{CE}.

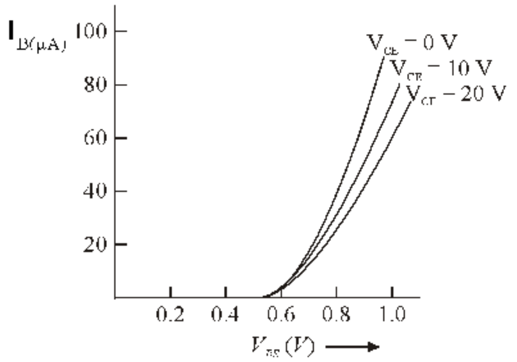


Fig. 16.42 : Input characterstics of CE transitor

Dynamic Input Resistance :

The dynamic input resistance for common emitter configuration of transistor is given by following relation :

$$R_{ie} = \left.\frac{\Delta V_{BE}}{\Delta I_B}\right|_{V_{CE}=\text{Constant}}$$

The value of this is approximately of the order of 100Ω. If this value is compared with common base circuit input resistance R_{ib} then, $R_{ie} > R_{ib}$.

Output Characteristics : Here input current I_B is kept constant and the change in output current I_C corresponding to output voltage V_{CE} is studied. The value of V_{CE} is changed with the help of R_2 and corresponding value of I_C is measured. A graph is drawn between V_{CE} and I_C. This process is repeated for other values of I_B. The group of curves so obtained are called output characteristics. These are shown in the figure (16.43).

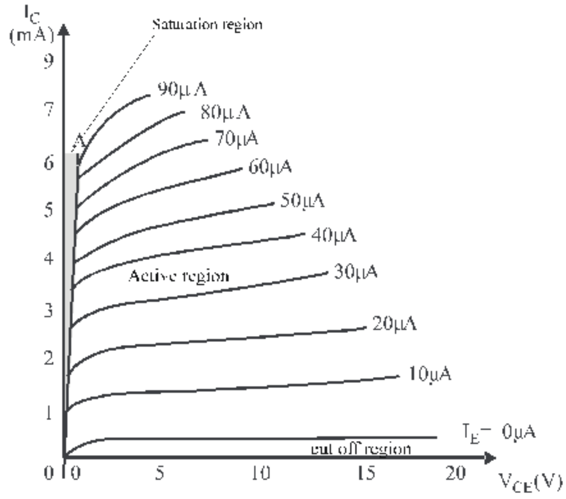


Fig. 16.43 : Output characterstics of CE transitor

At any constant value of I_B, when V_{CE} is increased from zero, I_C gets almost sturated rapily. After this the value of I_C increases slowly. The line *OA* in figure (1643) is called saturat line and the area between I_C axis and this line is called saturated region. When $I_B = 0$ but I_C is not zero, meaning when input current is zero still some output current flows. The region between I_B = 0 and V_{CE} axis is called cut off region.

The remaining region (apart from the cut of region and saturated region) is called the active region. Here, the emitter-base-junction is forward biased and base collector junction is reverse biased.

In active region I_C ia almost independent of V_{CE}.

The dynamic output resistance R_{oe} for this configuration is defined as follows :

$$R_{oe} = \left.\frac{\Delta V_{CE}}{\Delta I_C}\right|_{V_B=\text{Constant}}$$

Since, I_C does not change much with V_{CE} (leaving the initial values of V_{CE}) therefore, R_{oe} is charge of the order of 50-100kΩ.

Current Amplification Factor (β)

This is defined as the ratio of the change in collector current to the change in base current at a constant collector emitter voltage (V_{CE}) when the transistor is in active state.

The ratio of change in I_C (ΔI_C) to change in I_B (ΔI_B) at constant V_{CB} is called dynamic ampilication factor . It is denoted by β_{dc} ; $\beta_{dc} = \frac{I_c}{I_B}$

The ratio of change in I_C (ΔI_C) and change in I_B (ΔI_B) at constant V_{CB} is called dynamic amplificator factor β_{ac} or β

$$\beta_{ac} = \frac{\Delta I_C}{\Delta I_B}$$

Since, $I_C >> I_B$ and $\Delta I_C >> \Delta I_B$. Therefore, β_{dc} and β both are grater than 1 i.e β >> 1.

Relation between α and β

For any configuration of a transistor the emitter current I_E is equal to the sum of base current I_B and the collector current I_C.

$\therefore \quad I_E = I_B + I_C$ ....(16.18)

$\therefore$ for small changes in currents

$\Delta I_E = \Delta I_B + \Delta I_C$ ....(16.19)

Hence, it can be written;

or $\quad \frac{\Delta I_E}{\Delta I_C} = \frac{\Delta I_B}{\Delta I_C} + 1$ ....(16.20)

But $\quad \frac{\Delta I_C}{\Delta I_E} = \alpha$ and $\frac{\Delta I_C}{\Delta I_B} = \beta$

Hence, putting the values in equation (16.20)

$$\frac{1}{\alpha} = \frac{1}{\beta} + 1$$

Or $\quad \frac{1}{\alpha} = \frac{\beta + 1}{\beta}$

Or $\quad \alpha = \frac{\beta}{1+\beta}$ ....(16.22(a))

From equation (16.21);

$$\frac{1}{\beta} = \frac{1-\alpha}{\alpha}$$

$$\beta = \frac{\alpha}{1-\alpha} \qquad ....(16.22(b))$$

Since, the value of α is slightly less then 1 hence the value of β is very high in comparison to α.

16.11 Transistor Amplifier

Amplifier is an active electronic device in which the amplitude of output signal obtained is more than the amplitude of the input signal applied. Input signal is generally alternating current or voltage (ac). The process of increase in the amplitude of the signal is called amplification. In this there is no change in the shape and frequency of the signal. The necessary energy required for the increase in the amplitude of the signal is obtained from the dc power supply used in the amplifier circuit. Figure (16.44) shows the block diagram of an amplifier.

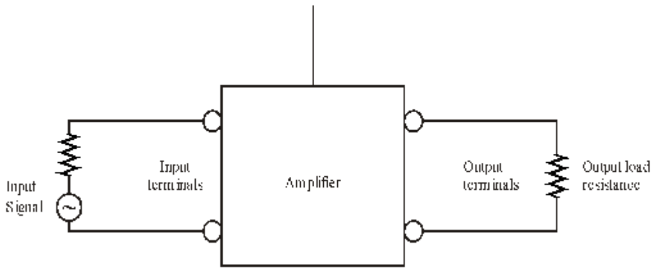


Fig. 16.44 : Block diagram of an amplifier

One type of amplifier, which is used in our daily life and which is known to us is the audio frequency amplifier. With the help of a microphone, the sound signals are converted into electrical signals of same frequency and given as input to the amplifier and output of high amplitude electrical signals are obtained which are again changed into sound signals by the loudspeaker. Hence, we hear a high intensity sound.

The ratio between output and input signals for an amplifier is known as amplification factor or gain. If the input signal voltage is V_i and output signal voltage is V_0 then;

Voltage amplification factor or voltage gain

$$A_i = \frac{\text{Output signal voltage}}{\text{Input signal voltage}} = \frac{V_o}{V_i}$$

Similarly, current amplification factor or current gain;

$$A_i = \frac{\text{Output signal current}}{\text{Input signal current}} = \frac{I_o}{I_i}$$

and Power amplification factor or power gain

$$A_p = \frac{\text{Output signal power}}{\text{Input signal power}} = \frac{P_o}{P_i}$$

$\because \qquad P_o = V_o \times I_o$

and $\qquad P_i = V_i \times I_i$

$\because \qquad A_p = \frac{V_o I_o}{V_i I_i}$

$$A_p = A_v A_i$$

Hence, for amplifier all the three amplification factors are correlated.

When a transistor is used as an amplifier its one terminal is common the both of input and output circuits. The signal whose amplification is to be done is used in input circuit, and output circuit is taken across a load resistor R_L connected in circuit.The amplification by a transistor is possible for both the common base and common emitter configurations. But the voltage, current and power gain for common emitter amplifier is more than the common-base amplifier. Hence, here we will study only about common-emitter amplifier.

16.11.1 Common-Emiter Amplifier

Figure 16.45 shows a common-emitter amplifier circuit for PNP transistor. Here, emitter is common for both the input and output circuits. With the help of battery

V_{BB} the emitter-base junction is forward is provided. Whereas with the help of battery $V_{CC}(>V_{BB})$ reverse bias is provided to the collector-emitter junction. V_i is signal voltage whose amplification is to be done (input signal).

A load resistor R_L is connected to output circuit across which amplified voltage V_O is obtained.

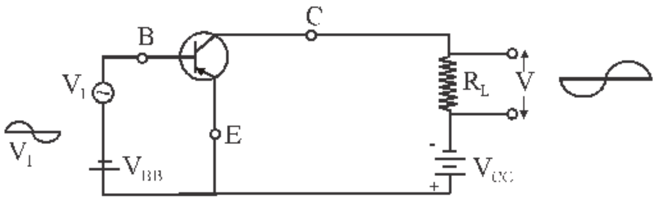


Fig. 16.45 : Common emitter amplifier

If alternative signal V_i is absent then potential difference V_{BE} across base-emitter junction is constant. Due to this the input current I_B and the output current I_C related to it are constants. In the presence of V_i, the potential difference across base-emitter junction changes. If for a small V_i changes in base and collector currents are i_b and i_c then, by definition :-

$$A_{ie} = \frac{\text{Output signal current}}{\text{Input signal current}}$$

$$= \frac{i_c}{i_b} = \beta$$

Since, β >> 1, therefore output symbol current is more than input current. Hence, in common emitter configuration the current is amplified. [In sign A_{ie} the subscript *e* represents common emitter configuration.]

Voltage Amplification :

By definition voltage amplification or voltage gain;

$$A_{ie} = \frac{\text{Output signal voltage}}{\text{Input signal voltage}} = \frac{V_o}{V_i}$$

Here, V_o =alternating voltage generated across load resistor R_L.

$$= i_c R_L$$

and $\quad V_i = i_b R_{ie}$

where R_{ie} is the input resistance in common-emitter configuration.

$$\therefore \quad A_{ve} = \frac{i_c R_L}{i_b R_{ie}} = \beta \times \frac{R_L}{R_{ie}}$$

$\frac{R_L}{R_{ie}}$ is called the resistance gain for common-emitter configuration.

As, $\beta >> 1$, and if R_L and R_{ie} are comparable;

$A_{ve} \approx \beta >> 1$ then voltage amplification is possible. If $R_L > R_{ie}$ then A_{ve} > β (>>1) more voltage amplification is obtained.

Power Amplification

By definition power amplification

$$A_{pe} - A_{ie} A_{ve}$$

$$= \beta\beta \frac{R_L}{R_{ie}} = \beta^2 \frac{R_L}{R_{ie}}$$

As β >> 1 then β^2 >> 1 hence A_{pe} >> 1 means that in this configuration high power amplification is possible.

Phase Relationship

In common emitter amplifier there is 180° phase difference (opposite phase) between input signal and output signal.

Example 16.5 : The current gain is 0.99 in common-base configuration of a transistor. What will be the current gain for the same transistor in common-emitter configuration?

Solution : The current gains for common-base configuration and common-emitter configuration are α and β respectively.

Since, $\quad \beta = \frac{\alpha}{1-\alpha}$

$\because \quad \alpha = 0.99$

$$\therefore \quad \beta = \frac{0.99}{1-0.99} = \frac{0.99}{0.01} = 99$$

Example 16.6 : In a common base circuit collector resistance is 2.0 kΩ and the potential difference across it ends is 2.0 V. For the transistor α = 0.95. Calculate base current I_B.

Solution : The collector resistance is 2.0 kΩ and potential difference across it is 2 V.

$$\therefore \quad I_C = \frac{2V}{2k\Omega} = 1mA$$

$$\therefore \quad I_E = \frac{I_C}{\alpha} = \frac{1}{0.95} = 1.05mA$$

Therefore , $I_B - I_E \;\; I_C = 1.05 - 1.0 = 0.5$ mA

Example 16.7 : In a junction transistor when its collector voltage V_{CB} is kept constant and emitter voltage V_{EB} is changed by 5 mV then its emitter current value changes by 0.15 mA. Calculate the input resistance of the transistor.

Solution : According to the question the transistor is in common-base configuration where its input resistance is :

$$R_{ib} = \frac{\Delta V_{EB}}{\Delta I_E}$$

According to the question;

$\Delta V_{EB} = 5$ mV

$\Delta I_{EB} = 0.15$ mV

$$\therefore \quad R_{ib} = \frac{5\times10^{-3}}{0.15\times10^{-3}} = 33.33\Omega$$

Example 16.8 : In an amplifier circuit a transistor is used in common-emitter configuration. A 20 μA change in base current brings 1 mA changes in collector current, and base-emitter voltage change by 0.04 V. Calculate :

(i) Input resistance,

(ii) Current amplification factor,

If a load resistor of 6 kΩ is used in the collector circuit also calculate the voltage gain of the amplifier.

Solution : In common-emitter configuration :

(i) Input resistance

$$R_{ie} = \frac{\Delta V_{BE}}{\Delta I_B} = \frac{0.04}{20A} = \frac{0.04V}{20\times10^{-6}A}$$

$$= 2000\Omega = 2k\Omega$$

(ii) Current amplification factor

$$B = \frac{\Delta I_C}{\Delta I_B} = \frac{1mA}{20A} = \frac{1\times10^{-3}}{20\times10^{-6}}\frac{A}{A}$$

$$= 50$$

Voltage gain;

$$A_{ve} = \beta\frac{R_L}{R_{ie}} = 50\times\frac{6k\Omega}{6k\Omega} = 150$$

16.12 Digital Electronics

In electronics, we use basically two types of signals;

(i) Analog

(ii) Digital

When we consider a current (or voltage) as analog, we mean having currents or voltages varying continueously with time (Fig. 16.46). In such a signal current (voltage) has continuous values in some range. Rectifier and amplifier circuits considered in this chapter are called analog circuits as input and output signals are analog in nature.

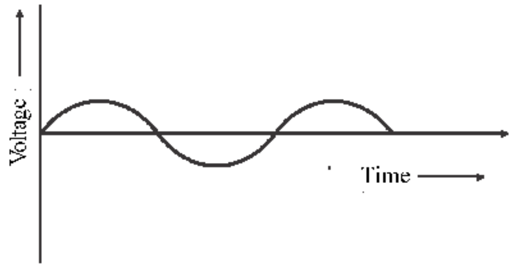


Fig. 16.46 : Analog signal

A digital signal is a signal in which the voltage (current) does not change continuously with time, here the voltage has a discrete value at each sampling point. In other words, the signals are obtained in the form of a pulse. For example, if we take a bulb connected to a switch, then in this bulb for voltage only two states are possible 0 when switch is off and maximum value 220 V when switch is ON. In this state a bulb is a device which can be assumed to be based on binary variable. If zero voltage is represented by 0 and maximum voltage 220 V is represented by 1. Then if switch is ON and OFF continuously then the voltage present on the bulb can be represented as in figure (16.47). Hence, it is an example of digital signal.

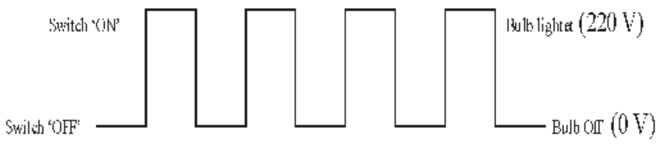


Fig. 16.47 : Digital Signal

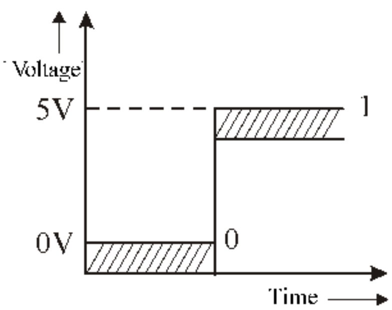


Fig. 16.48 : Representing digital signal

From the above example, it is clear that the digital signals can be represented by binary numbers 0 and 1 (bits). In reality 0 and 1 do not represent 0 V and 1 V. Bit 0 represents absence of signal and 1 represents presence of signal. Digital watches, modern computers, etc. are all examples of digital devices.

In digital circuits generally the high voltage level is V_{high} (5 ± 0.5) V which is represented by 1 and low voltage level V_{low} (0 ± 0.5) V which is represented by 0. Knowledge of exaxct values of V_{high} and V_{low} is not necessary, because the interval between these levels. A representative digital signal is shown in figure (16.48)

16.13 Logic Gates

A logic gate is an elementary building block of a digital circuit. Most logic gates have two inputs and one output. A logic gate is a digital circuit that follows certain logical relationship between the input and output voltages. The output signal is available only when some conditions are satisfied at input i.e. a logical relation exists between input signals. The five common logic gates used are OR gate, AND gate, NOT gate, NOR gate and NAND gate.

Each logic gate is indicated by a symbol and its function is defined by a truth table that shows all the possible input logic level combination with their respective output logic levels. Truth tables help to understand the behaviour of logic gates. These logic gates can be realised using semiconductor devices.

16.13.1 OR Gate

An OR gate has two or more inputs with one output. The output Y is 1 when either input A or input B or both are 1s, that is; if any of the input is high, the output is high.

Since logic gates are based on such variables for which only two values 0 and 1 are possible, the algebra for such variables is different than ordinary algebra. This type of algebra is known as Boolean algebra and the equations representing these relationships between variables are called Boolean expressions.

For a two input OR gate if inputs are given as A and B respectively and output is Y then boolean expression is

$$Y = A + B$$

Here the + sign between A and B represents OR operation. Y = A + B means Y is A or B.

where A and B can have values O or 1 A = 0, 1 B = 0, 1.

The truth table for this operation is as follows;

A	B	Y = A + B
0	0	0
1	0	1
0	1	1
1	1	1

Thus when either A or B or both are 1 then outputY is also 1. When Both A and B are zero then output Y is also zero.

OR operate can be explained by electrical switches connected in parallel. As shown in figure (16.49) the switches A and B are connected in parallel and are connected with a battery and a bulb Y.

If closed state (ON state) of A and B is represented by 1 and open state (off state) is represented by 0 and the illuminated and non illuminated situation of the bulb is indicated by 1 and 0 then for this circuit there are four possibilities.

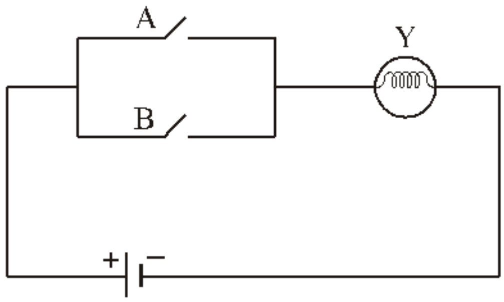


Fig. 16.49 : Representing OR operation

(i) If A and B both are open then there is no current in Y, hence bulb is not illuminated. In mathematical form A = 0, B = 0 then Y = A + B = 0.

(ii) If switch A is closed (A = 1) and B is open (B = 0) then the current will flow through switch A and bulb Y, illuminated hence Y = 1 : ∴ A= 1, B = 0 then Y = A + B = 1.

(iii) If A is open and B is closed then also bulb will be illuminated. ∴ A = 0, B = 1, Y = A + B = 1.

(iv) If A and B both are closed (A = B = 1) then bulb will be illuminated (Y = 1). ∴ A = B = 1, Y = A + B =1.

All the above four possibilities are according to the truth table. Figure (16.50) shows the symbol for OR gate.

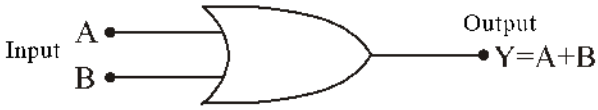


Fig. 16.50 : Symbol for two input OR gate

In practice an OR gate is constructed by circuit formed by two diodes D_1 and D_2 as shown in Fig. (16.51).

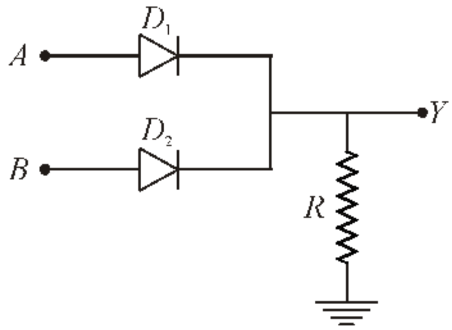


Fig. 16.51 : OR gate based on diode

The value of 0 V or 5 V is provided to the inputs A and B. 0 V is represented by 0 and 5 V represented by 1. R is the resistance at output Y and the other end of resistance R is earthed.

(v) When A = B = 0 meaning there is no signal then no diode will be operated and there will be no potential drop at R or Y = 0. This is equivalent to the first row of the truth table.

If input A is 5 V (1) or B = 0 V (0) then diode D_1 will be forward biased and behave as closed switch hence is in on state. If D_1 is ideal diode then there is no potential drop across it and a potential drop of 5 V (with respect to ground) will be available across R or will get Y = 1 condition. This will be similar to second row of truth table.

(vi) If A = 0 and B = 5 V (1) then in place of D_1, D_2 will be in on state. Still Y = 1. This is similar to third row of truth table.

(vii) If A and B both the diodes are (5 V) then both diodes will be in operation again Y = 1. This will be similar to fourth row of truth table.

16.13.2 AND Gate

An AND gate has two or more inputs and one output. The output Y of AND gate is 1 only when input A and input B are both 1. The boolean expression for the AND operation is represented as; Y = A•B

Here dot (•) represents AND operation

The above boolean equation means, Y = A and B. The truth table for this;

A	B	Y = A•B
0	0	0
1	0	0
0	1	0
1	1	1

Hence, in AND gate when both inputs are is 1 then only output is 1.

The AND gate can be explained by electrical switches; in series (Figure 16.52). When both switches A and B are open (A = B = 0) then there is no current in circuit. Hence, bulb Y is not illuminated. This is equivalent to the first line of the truth table. If any of the two switches is open then also Y = 0. When both switches are closed (A = B = 1), the Y = 1, bulb is illuminated. Figure (16.53) shows symbol for a two input AND gate.

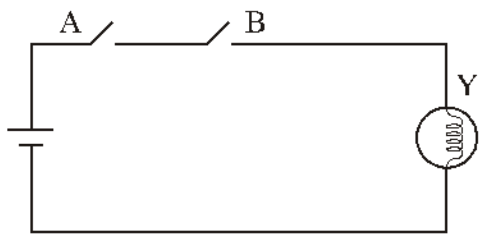


Fig. 16.52 : Representing AND gate

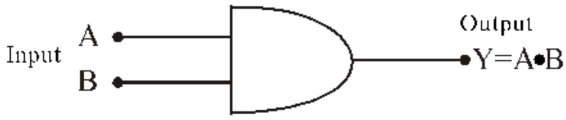


Fig. 16.53 : Symbol for AND gate

AND gate is also obtained with the help of diodes. Figure (16.54) shows a two input AND gate formed by diodes.

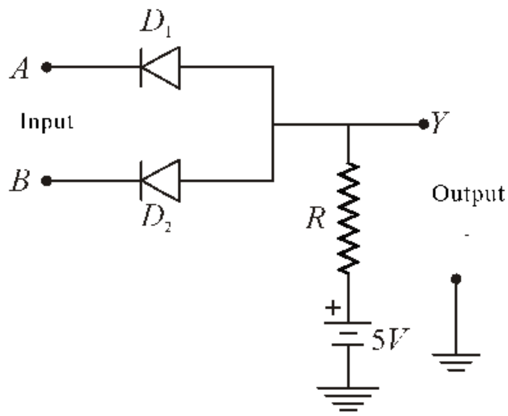


Fig. 16.54 : AND gate using diode

Here resistance R is connected to the positive terminal of a battery of potential difference 5 V and the negative terminal is earthed. Input terminals A and B are given 0 or 5 V level.

If A and B both are at zero potential (A = B = 0) then diodes D_1 and D_2 both are forward biased. As diodes are ideal there will not be any potential drop across any of the diodes and 5 V potential drop is available across R its end Y will be at zero potential with respect to ground (Y = 0). This represents the first line of truth table.

If A and B any of the two has value 5 V and the other has 0 V, then diode A is in operation and B is in non-operation mode. If diode A is ideal the there will be no potential drop across it. Again 5 V potential drop is there across R and Y = 0 the same is true A = 0 and B = 1. This represents the third line. If A = B = 5 V, then (A=B=1) then both diodes are non conducting and there is no current in R so its upper end is at same potential as its lower end i.e at + 5 V with respect to earth so Y=1 hence this represents the fourth line of truth table.

16.13.3 NOT Gate

This is a logic gate, with one input and one output. It produces '1' output if the input is '0' and vice-versa. That is it produces an inverted version of the input at its output. This is why it is also known as an inverter.

The Boolean expression is given by;

$Y = \overline{A}$.

$\overline{A}$ means NOT A meaning which is not A. ($\therefore$ A=0, Y = $\overline{0}$, =1) and (A = 1, Y = $\overline{1}$=0). The truth Table is givien as follows :

A	Y
0	1
1	0

Not operation can be understood by electrical circuit shown in Fig. 16.55. When switch is on (1) then Y (bulb) is off (0). When switch is off (0) then Y is illuminated (1). Figure (16.56) shows NOT gate symbol.

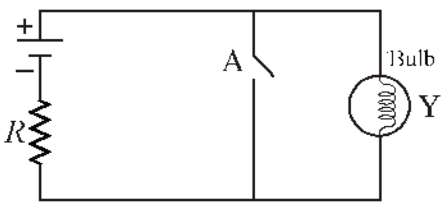


Fig. 16.55 : Circuit representing NOT gate

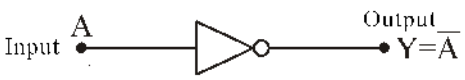


Fig. 16.56 : Symbol for NOT gate

For NOT gate transistor is used. A simple circuit is shown in the figure (16.57), in which an NPN transistor is used. The base B of the transistor is connected with the terminal A through a resistor R_B. Emitter E is earthed and collector C is connected to the positive end of dc supply V_{CC} = (5 V) through a resistor R_C. Negative terminal of the supply is grounded. Y is the potential difference of collector C relative to ground.

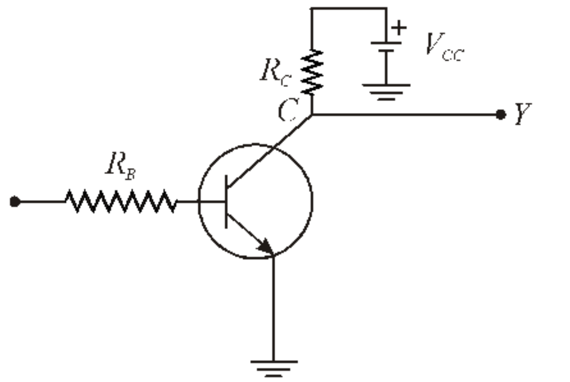


Fig. 16.57 : Transitor NOT gate

When, input terminal A is earthed then $V_A = 0$ base is also earthed. So there is no bias at EB junction by BC junction. In this the base current is zero, emitter current is zero the collector current is also zero, and the transistor is in cut off state, so collector is at + 5 V potential with respect to ground i.e at high potential and Y=1.

Now If relative to earth + 5 V voltage is used at base (meaning A = 1) then base-emitter junction is forward biased. Now base current, emitter current and collector current all are present. If the values of R_B and R_C are chosen in such a way the collector current is high then transistor is in saturation state. In this state, the potential drop across R_C is + 5 V which is equal to voltage V_{CC} so voltage at C is zero and Y = 0.

16.13.4 NOR Gate

It has two or more inputs and one output. A NOT operation applied after OR gate gives a NOT-OR gate or simply NOR gate. Its output Y is 1 only when both inputs A and B are 0. This logic is shown in the figure (16.58).

Fig. 16.58 : Action of NOR gate

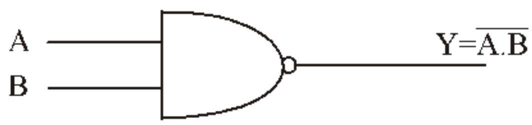


Fig. 16.59 : Symbol for NOR gate

The symbol is shown for NOR gets in figure (16.59), in which a bubble is placed at the OR gate output.

Boolean expression for this gate is given by

$$Y = \overline{A + B}$$

The truth table for this is;

A	B	A+B	$Y = \overline{A + B}$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

It is clear from this table that when all the inputs are zero, then only output is 1.

For three inputs the NOR gate is shown in the figure 16.60.

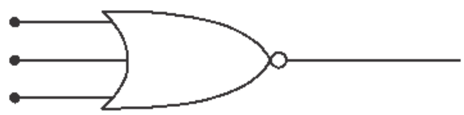


Fig. 16.60 : Symbol of NOR gate with three inputs

16.13.5 NAND Gate

This is an AND gate followed by a NOT gate as shown in Fig. 16.61. If inputs A and B are both 1 the output Y is 0. Its circuit symbol is shown in Fig. 16.62 NOT AND behaviour provides NAND gate.

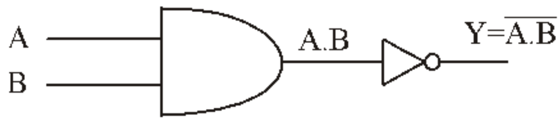


Fig. 16.61 : Action of NAND gate

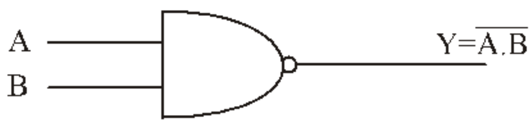


Fig. 16.62 : Symbol of two input NAND gate

The truth table for this gate is shown below.

A	B	A . B	$Y = \overline{A \cdot B}$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

16.13.6 XOR Gate

The Boolean expression for XOR operation is :

$$Y = A \oplus B$$

When any one of the input signals A or B has value 1 then only Y = 1. If both the two inputs are 0 or 1 then output is 0. Hence, truth table for this gate is :

A	B	$Y = A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

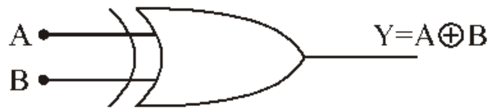


Fig. 16.63 : Symbol for XOR gate

The XOR gate symbol is shown in the figure 16.63. By making following truth table one can verify that -

$$A \oplus B = A\overline{B} + \overline{A}B$$

A	B	$\overline{A}$	$\overline{B}$	$A\overline{B}$	$\overline{A}B$	$A\overline{B} + \overline{A}B$
0	0	1	1	0	0	0
1	0	0	1	1	0	1
0	1	1	0	0	1	1
1	1	0	0	0	0	0

Hence, by using boolean expression $A\overline{B} + \overline{A}B$ XOR gate can be constructed. This is shown in the figure (16.64) in which we have used AND, OR and NOT gates.

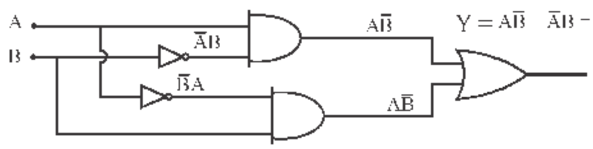


Fig. 16.64 : Boolian circuit for XOR gate

Special Notes

NAND gates are also called universal gates since by using these gates you can also realise other basic gates like OR, AND and NOT. Figure 16.65 shows how various gates are realised by NAND gate.

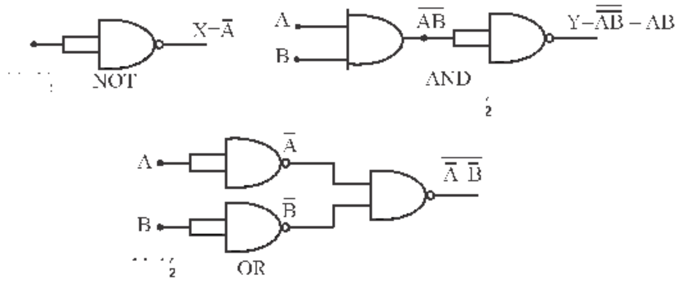


Fig. 16.65 : Different logic operations using NAND gate

In Fig. 16.55 (A) both the inputs are A so output from NAND operation is Y = $\overline{A}$. $\overline{A}$ = $\overline{A}$ which is NOT operation. In fig. 16.55 (B) output of first NAND gate is $\overline{AB}$ which is input for second NAND gate, so output for second NANO gate is Y = $\overline{\overline{AB}}$ = AB which is AND operation.

In Fig. 16.55 (c) construction of OR gate is shown. Input A and B are inverted through separate NAND gates and then signals $\overline{A}$ and $\overline{B}$ have been applied to third NAND gate. The output so obtained is Y = $\overline{\overline{A}\,\overline{B}}$. From the following truth table it can be seen that $\overline{\overline{A}\,\overline{B}}$ = A + B.

A	B	$\overline{A}$	$\overline{B}$	$\overline{A}\,\overline{B}$	$\overline{\overline{A}\,\overline{B}}$
0	0	1	1	1	0
1	0	0	1	0	1
0	1	1	0	0	1
1	1	0	0	0	1

NOR gates are also called universal gate since by using these gates we can also make other basic gates like OR, AND and Not gate Figure 16.66 shows how various gates are made by using NOR gates.

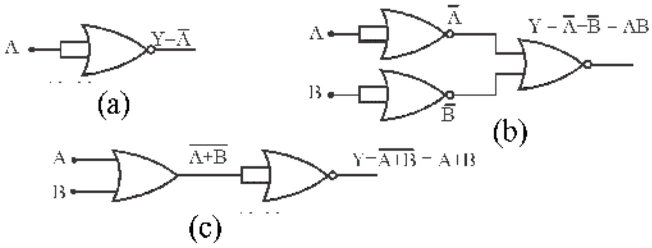


Fig. 16.66 : Different logic operations using NOR gate

In Fig. 16.66 (a) both the inputs for NOR gate are A so output X = $\overline{\overline{AB}}$ = A + B which is NOT operation.

In fig. 16.66 (b) inputs A and B applied through two separate NOR gates respectively. The respective outputs $\overline{A}$ and $\overline{B}$ now act as input for another NOR gate the output for which us Y = $\overline{A}$ + $\overline{B}$. Using a truth table it can be shown that Y = $\overline{A}$ + $\overline{B}$ = AB so AND operation is achieved. Students are advised to verity the same.

In figure 16.66 (c) the output of $\overline{A+B}$ of first NOR gate is input for the second NOR gate. The output of second NOR gate is Y = $\overline{\overline{A+B}}$ = A + B so OR operation is achieved.

Important Points

1. In isolated atoms energy levels are discrete. In a solid because of interaction between atoms energy level of atoms splits and energy bands are formed in place of energy levels. Every electron in an atom stays in one of these pre-defined energy levels. the continuous energy groups are called energy bands.
2. These are forbidden energy gaps between the energy bands. No electron is relevant to these energy intervals in solids.
3. On the basis of electrical conductance and energy band structure substances are divided as (i) conductors (ii) insulators and (iii) semiconductors. Completely empty and fully filled bands do not participate in conduction.
4. In conductors, either the band related to the valence electrons is partially filled or it over laps with next band so the new band is partly filled and hence transition is possible. Partially filled bands help in conduction.
5. In insulators valence band is completely filled and higher band (conduction band) is fully empty. The energy gap between these bands is called forbidden energy gap (E_g). In insulators this forbidden energy gap is 3 to 6 eV.
6. In intrinsic semiconductors the forbidden energy gap is very small in comparison to insulators and is of 1 eV range. At absolute zero temperature their behaviour is like insulators. At room temperature some electons of valence band get thermal energy and reach conduction band. In valence band in place of these electrons holes are generated.
7. In intrinsic conductor both electrons and holes participate in conduction and their conductivity is in between the conductivity of conductors and insulators. When temperature increases the conductivity of intrinsic conductors increases.
8. In intrinsic semiconductor if the impurity of suitable kind is mixed in very small quantity their conductivity increases very rapidly. Such conductors are called extrinsic semiconductors.
9. Extrinsic semiconductors are of two types - (i) N-type (ii) P type.
10. Tetravalent intrinsic semiconductors like silicon and germanium when doped with pentavalent element like arsenic, phosphorous in very small quantity then N-type extrinsic semiconductors are obtained. The impurity is called donor impurity.
11. When a trivalent impurity (element) like aluminium, boron, indium is doped in very smallquantity with intrinsic semiconductor then P-type semiconductor are formed, such impurities are acceptor impurities.
12. In N-type semiconductors electrons are majority charge carriers and holes are minority charge carriers. In P-type semiconductors reverse is the case. Both types of semiconductors are electrically neutral.
13. When a P type semiconductor is connected to a N-types semiconductor at atomic level then their contact surface is called P-N junction. Almost all semiconductor devices use P-N junction. A depletion layer is formed close to the P-N junction in which there are bounded positive and negative ions and the number of free electrons and holes reduces (negligible).
14. The device based on P-N junction is called P-N diode or semiconductors diode. In forward biasing of junction, the P-terminal is at higher potential than N-terminal. In reverse biasing it is opposite.
15. In forward biasing there is conduction of current by the diode whereas in reverse biasing there is negligible

current flow hence diode is in non conducting state.

16. The P-N junction diode is basically used for rectification in which ac is changed to dc.

17. Semiconductor diodes are also used for some special purposes. Out of these Zener diode is operates in reverse biased state to help in voltage regulation. On the basis of optical properties of the semi conductors there are P-N devices like photo diode, light emitting diode, etc.

18. Junction transistor is a very important device in which there are two P-N junction. There are of two types (i) PNP (ii) NPN. Their middle part 'base' is very thin and very lightly dopped, and the other two parts are emitter and collector, emitter has higher doping than collector but the nature of impurity is same.

19. For active operation of junction transistor, base-emitter junction is forward biased and base-collector junction is reverse biased.

20. In circuit combination the transistor is so joined that out of base (B), collector (C) and emitter (E) any one is common for both input and output circuits. Hence, three types of circuit configurations are possible for a transistor but they are different in electrical properties.

21. the current amplification factor for common base and common emitter configurations are α and β respectively.

$$\beta = \frac{\alpha}{1-\alpha} \text{ and } \alpha < 1 \text{ and } \beta >> 1$$

22. In common emitter transistor amplifier high voltage gain, high current gain and high power gain are obtained.

23. In digital electronics binary number system is used. In this we use only 0 and 1.

24. A logic gate is an elementary building block of a digital circuit. The gates are OR gate, AND gate, NOT gate, XOR gate, NAND gate and NOR gate. NAND and NOR gates are called universal gates; because with the help of these other gates can be obtained.

Questions For Practice

Multiple Choice Type Questions

1. At absolute zero temperature intrinsic Germanium and intrinsic Silicone; are:

   (a) Super conductor (b) Good semiconductors

   (c) Ideal insulator (d) Conductors

2. In insulators the forbidden energy gap between valence band and conduction band is of ......

   (a) 1 eV (b) 6 eV

   (c) 0.1 eV (d) 0.01 eV

3. At room temperature in intrinsic Silicon the number of charge carriers per unit volume is $1.6 \times 10^{16}/m^3$. If mobility of electons is $0.150\ m^2V^{-1}s^{-1}$ and mobility of holes is $0.05\ m^2V^{-1}s^{-1}$. Then conductivity of silicon is ($\Omega^{\ 1}\ m^{\ 1}$) is :

   (a) 1.28×10^{-4} (b) 3.84×10^{-4}

   (c) $5.12 \times 10^{\ 4}$ (d) $2.14 \times 10^{\ 4}$

4. If a NPN transistor is used as an amplifier then;

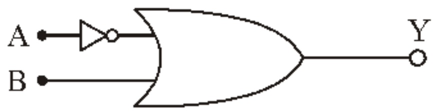
   (a) Electrons move from base to collector.

   (b) Holes move from emitter to base.

   (c) Holes move from base to emitter.

   (d) Electrons move from emitter to base.

5. the boolean equation for given circuit will be :



(a) $Y = A + \overline{B}$ (b) $Y = \overline{A + B}$

(c) $Y = \overline{A} + B$ (d) $Y = \overline{A}.B$

6. For some 'AND GATE' the three inputs are A, B and C then the output Y will be;

(a) Y = A.B + C (b) Y = A + B + C

(c) Y = A + B.C (d) Y = A.B.C.

7. The current amplification factor for common base circuit of a transistor is 0.95. When emitter current is 1mA then base current is :

(a) 0.1 mA (b) 0.2 mA

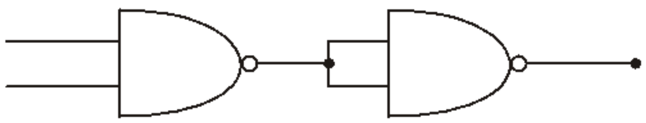
(c) 0.19 mA (d) 1.9 mA

8. the forbidden energy gap in Germanium is 0.7 eV. The wavelength at which its absorption is starts by Germanium is :

(a) 35000 Å (b) 17700 Å

(c) 25000 Å (d) 51600 Å

9. The logic gate obtained by two NAND gates shown in figure is :



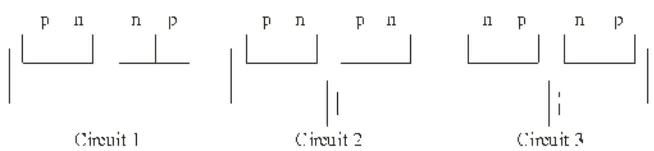
(a) AND gate (b) OR gate

(c) XOR gate (d) NOR gate

10. Two identical PN junction are joined in series with a battery (fig.). For which potential drop is same.

(a) Circuit 1 and 2 (b) Circuit 2 and 3

(c) Circuit 3 and 1 (d) Only circuit 1.



Answers (Multiple Choice Type Questions)

1. (c) 2. (b) 3. (c) 4. (d) 5. (c)

6. (d) 7. (a) 8. (b) 9. (a) 10. (b)

Very Short Answer type Questions

1. What is the direction of diffusion current in junction diode?

2. Write down the relation between α and β (current amplification factor) of the transistor.

3. Can the barrier potential of a forward biased pn junction be measured by a voltmeter connected at the ends of the device?

4. Make truth table for OR gate.

5. Write down the name of that logic gate in which the output is 1 when all the input are 1.

6. To use the transistor as an amplifier which junction is reverse biased?

7. What will be the value of α for a transistor for which β = 19?

8. The diode shown in the figure is in which biasing?

Short Answer Type Question

1. What is rectification? Draw the figure of bridge full wave rectifier.

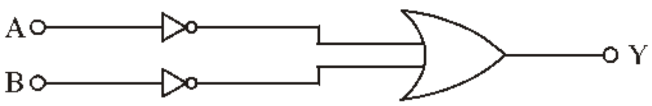
2. In a transistor why is the base made thin in comparison to the emitter and collector?

3. Make complete I-V characteristic curve for ideal PN junction diode. Define dynamic resistance in forward biased state.

4. What do you understand by logic gate? Draw the symbol for XOR gate and also write the truth table.

5. Draw the transistor based circuit diagram for NOT gate and also give its truth table.

6. Write down the boolean expression for given logic circuit. And also write truth table for it.



7. Draw the circuit used for voltage regulation by zener diode and also explain process in short.

Essay Type Questions

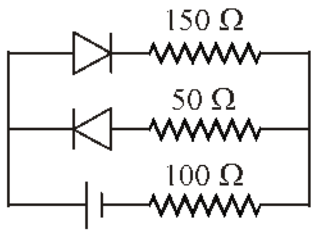
1. Differentiate between conductors, insulators and semiconductors on the basis of energy band theory. Explain the process of electrical conductance in intrinsic semiconductors.
2. What is PN junction? Explain the process occuring during its formation on the junction surface. If this junction is forward biased then explain the effect on depletion layer.
3. Draw the figure of full wave rectifier used to change ac to dc and explain its working.
4. Draw and explain the circuit arrangement for obtaining characteristic curve for forward and reverse biased PN-junction diode. Also draw the these curves.
5. What is junction transistor? Make the necessary diagram and explain the working of a PNP transistor.
6. Draw the circuit arrangement and also explain the characteristics for a transitor common emitter configuration. Also draw the curves and write the formula for voltage gain and current gain.
7. What do you understand by amplification? Make the diagram of a PNP transistor common amplifier and explain the process of amplification and calculate the formula for voltage gain.
8. Write down the names of some diodes used for special purposes and make their circuit diagrams. Write about their working and uses in short.
9. Draw circuit diagram for two input diode OR gate and AND gate. Also explain their working and give truth table.

Numerical Questions

1. Calculate the electic current generated in an intrinsic germanium plate at room temperature whose area is 2×10^{-4} m^2 and width is 1.2×10^{-3} m and a potential difference of 5 V is applied across its faces. Intrinsic charge carrier density is 1.6×10^{6}/ m^3 for germanium at room temperature. the mobility of electrons and holes is 0.4 $m^2\ V^{-1}s^{-1}$ and $0.2\ m^2v\ ^{1}s\ ^{1}$.

   [**Ans :** 1.28×10^{-13} A]

2. As shown in the circuit below the forward resistance of both diodes is 50 Ω and reverse resistance in infinite. If emf of the battery is 6 V then calculate the current flowing through 100Ω.



   [**Ans :** 0.02 A]

3. The current amplification is 0.99 for amy transistor in common-base configuration. Calculate the change in collector current when there is 5.0 millampere change in emitter current. What will be the change in base current?

   [**Ans :** 0.9 mA, 0.05 mA]

4. For a PN junction the average value of the potential barrier is 0.1 V and the electric field 10^5 V/m is present at junction region. What will be the width of depletion layer for this junction?

   [**Ans :** $10\ ^{6}$m]

5. A transistor is connected in common emitter configuration. A power supply of 8V is there in the collector circuit and the potential drop of 0.5 V is on the resistance of 800Ω connected in series with the collector. If current amplification factor is $\alpha = 0.96$. Then calculate base current.

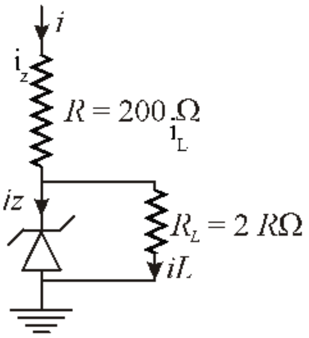
   [**Ans :** 26×10^{-3} Amp]

6. In a common emitter amplifier an increase of 50 μA in base current causes a 1.0 mA increase

in collector current. Calculate current gain β. What will be the change in emitter current? Calculate α with the help of β.

[**Ans :** $\beta = 20$, $\Delta I_E = 1050$ A, $\alpha = 0.95$]

7. Calculate the current and potential difference across ends of zener diode in the given circuit. If load resistor $R_L = 2\ k\Omega$ has a potential difference of 15V across its ends. The lowest working current of zener diode is 10mA.



[**Ans :** 17.5 mA, 15 V]