Chapter- 14

SEMICONDUCTOR ELECTRONICS: MATERIALS, DEVICES AND SIMPLE CIRCUITS

BANDS IN CONDUCTORS, SEMICONDUCTORS, AND INSULATORS (QUALITATIVE IDEA ONLY)

CLASSIFICATION OF METALS, CONDUCTORS, AND SEMICONDUCTORS

(1)Based on conductivity

Based on the relative values of electrical conductivity or resistivity, the solids are broadly classified as:

(i) Metals: They possess very low resistivity (or high conductivity).

$\rho \sim 10^{-2} - 10^{-8}\,\Omega\cdot\mathrm{m}$

$\sigma \sim 10^{2} - 10^{8}\,\mathrm{S}\cdot\mathrm{m}^{-1}$

(ii) Semiconductors: They have resistivity or conductivity intermediate to metals and insulators.

$\rho \sim 10^{-5} - 10^{6}\,\Omega\cdot\mathrm{m}$

$\sigma \sim 10^{5} - 10^{-6}\,\mathrm{S}\cdot\mathrm{m}^{-1}$

(iii)Insulators: They have high resistivity (or low conductivity).

$\rho \sim 10^{11} - 10^{19}\,\Omega\cdot\mathrm{m}$

$\sigma \,\square\, 10^{-11} - 10^{-19}\,\mathrm{S}\cdot\mathrm{m}^{-1}$

(2)Based on energy bands

Inside the crystal each electron has a unique position and no two electrons see exactly the same pattern of surrounding charges. Because of this, each electron will have a different energy level. These different energy levels with continuous energy variation form energy bands. The energy band which includes the energy levels of the valence electrons is called the valence band. The energy band above the valence band is called the conduction band. With no external energy, all the valence electrons will reside normally in the valence band. Normally the conduction band is empty. But when it overlaps on the valence band electrons can move freely into it. This is the case with metallic conductors.

If there is some gap between the conduction band and the valence band, electrons in the valence band all remain bound and no free electrons are available in the conduction band. This makes the material an insulator. But some of the electrons from the valence band may gain external energy to cross the gap between the conduction band and the valence band. Then these electrons will move into the conduction band. At the same time, they will create vacant energy levels in the valence band where other valence electrons can move. Thus the process creates the possibility of conduction due to electrons in the conduction band as well as due to vacancies in the valence band.

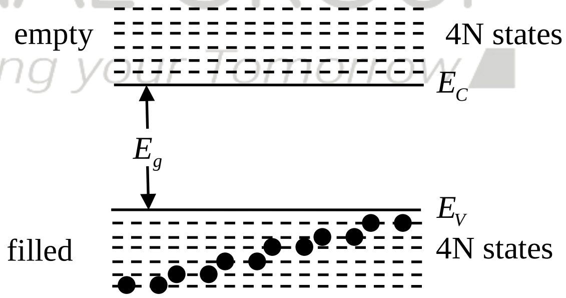


fig.(1) The energy band positions in a semiconductor at 0 K

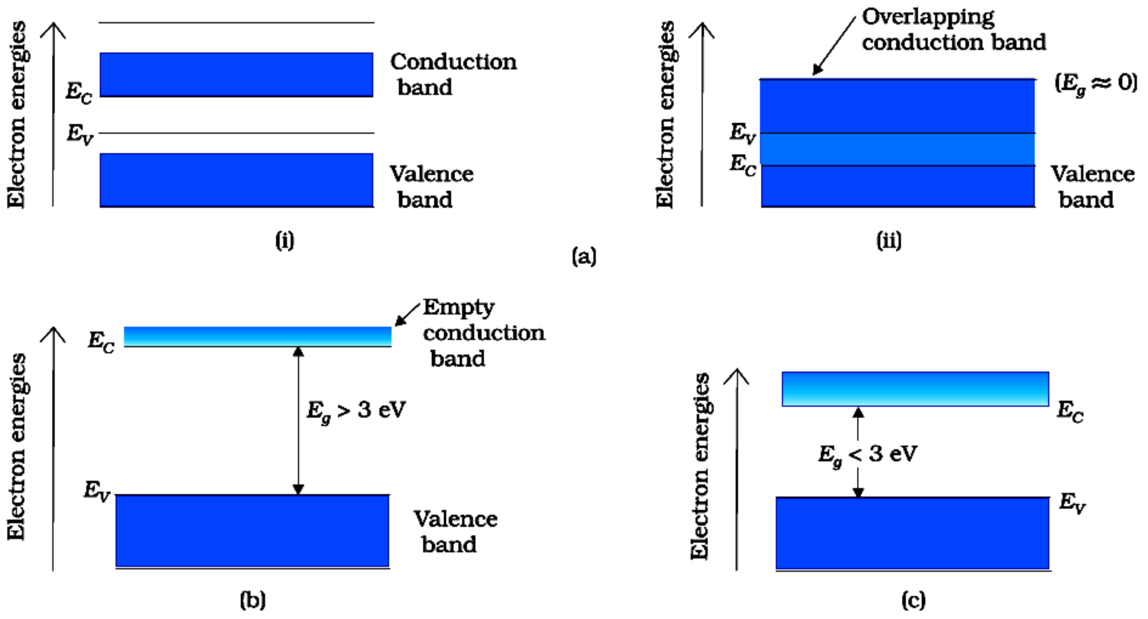
Let us consider what happens in the case of Si or Ge crystal containing N atoms. For Si, the outermost orbit is the third orbit (n = 3), while for Ge it is the fourth orbit (n = 4). The number of electrons in the outermost orbit is 4 (2s and 2p electrons). Hence, the total number of outer electrons in the crystal is 4N. So, for the 4N valence electrons, there are 8N available energy states. At the distance between the atoms in the crystal lattices of Si and Ge, the energy band of these 8N states is split apart into two which are separated by an energy gap E_g (Fig. 1). The lower band which is completely occupied by the 4N valence electrons at a temperature of absolute zero is the valence band. The other band consisting of 4N energy states, called the conduction band, is empty at absolute zero.

The lowest energy level in the conduction band is shown as E_C and the highest energy level in the valence band is shown as E_V. Above E_C and below E_V there are a large number of closely spaced energy levels, as shown in (Fig. 1). The gap between the top of the valence band and the bottom of the conduction band is called the energy band gap (energy gap E_g). It may be large, small, or zero, depending upon the material. These different situations, are depicted in (Fig.2) and discussed below:

The case I: One can have a metal either when the conduction band is partially filled and the valance band is partially empty fig. (2a)(i) or when the conduction and valance bands overlap fig. (2a)(ii).

Case II: In the case of an insulator, as shown in (Fig.2b), a large band gap E_g exists (E_g > 3 eV). There are no electrons in the conduction band, and therefore no electrical conduction is possible.

Case III: In the case of semiconductor, as shown in (Fig.2c), a small band gap E_g exists (E_g < 3 eV). At room temperature, some electrons from the valence band can acquire enough energy to cross the energy gap and enter the conduction band.



[fig.(2)]

NOTE-semiconductors can be:

(i) Elemental semiconductors: Si and Ge

(ii) Compound semiconductors: Examples are:

a) **Inorganic**: CdS, GaAs, CdSe, InP, etc.

b) **Organic**: anthracene, doped phthalocyanines, etc.

c) **Organic polymers**: polypyrrole, polyaniline, polythiophene, etc.

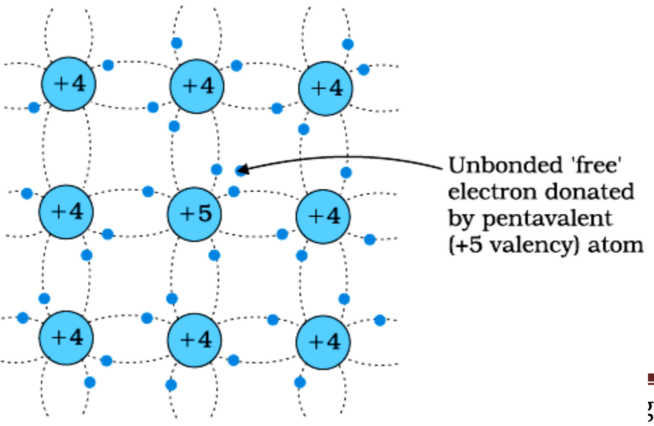
Most of the currently available semiconductor devices are based on elemental semiconductors Si or Ge and compound inorganic semiconductors. However, after 1990, a few semiconductor devices using organic semiconductors and semiconducting polymers have been developed signaling the birth of a futuristic technology of polymer electronics and molecular-electronics.

EXTRINSIC SEMICONDUCTOR

When a small amount, say, a few parts per million (ppm), of a suitable impurity is added to the pure or intrinsic semiconductor, the conductivity of the semiconductor is increased manifold. Such materials are known as extrinsic semiconductors or impurity semiconductors. The deliberate addition of a desirable impurity is called doping and the impurity atoms are called dopants. Such a material is also called a doped semiconductor. The dopant has to be such that it does not distort the original pure semiconductor lattice.

There are two types of dopants used in doping the tetravalent Si or Ge and accordingly, two types of extrinsic semiconductors are prepared:

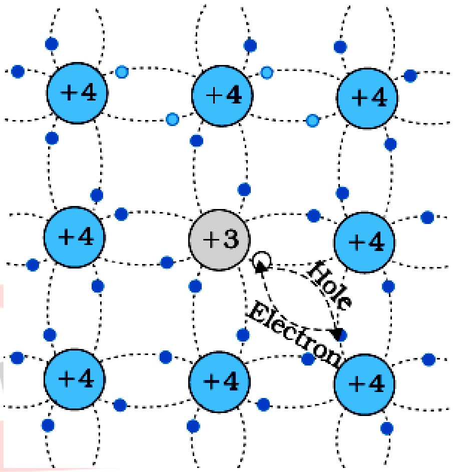
(i)n-type semiconductor This is obtained when Si or Ge is doped with pentavalent (valency 5) impurity like Arsenic (As), Antimony (Sb), Phosphorous (P), etc. The



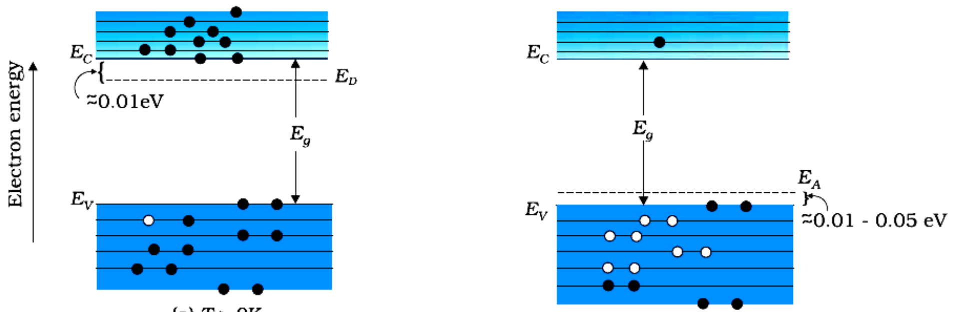
n-type semiconductor

pentavalent dopant is donating one extra electron for conduction and hence is known as a donor impurity. In n-type semiconductors, the total number of conduction electrons ne is due to the electrons contributed by donors and those generated intrinsically, while the total number of holes nh is only due to the holes from the intrinsic source. So electrons become the majority carriers and holes the minority carriers. $n_e > n_h$.

(ii)p-type semiconductor This is obtained when Si or Ge is doped with trivalent (valency 3) impurity like Al, B, In, etc. The trivalent dopant is donating one extra hole or accepting one electron for conduction and hence is known as accepter impurity. In a p-type semiconductor, the total number of conduction holes nh is due to the holes contributed by accepter and those generated intrinsically, while the total number of electrons ne is only due to the electrons from the intrinsic source. So holes become the majority carriers and electrons the minority carriers. $n_e < n_h$.



p-type semiconductor



(a) $T > 0K$
one thermally generated electron-hole pair + 9 electrons from donor atoms

(b) $T > 0K$

PROBLEM

Suppose a pure Si crystal has 5×10^{28} atoms m^{-3}. It is doped by 1 ppm concentration of pentavalent As. Calculate the number of electrons and holes. Given that $n_i = 1.5\times10^{16}\,\mathrm{m}^{-3}$.

SOL-

$$n_e n_h = (n_i)^2$$

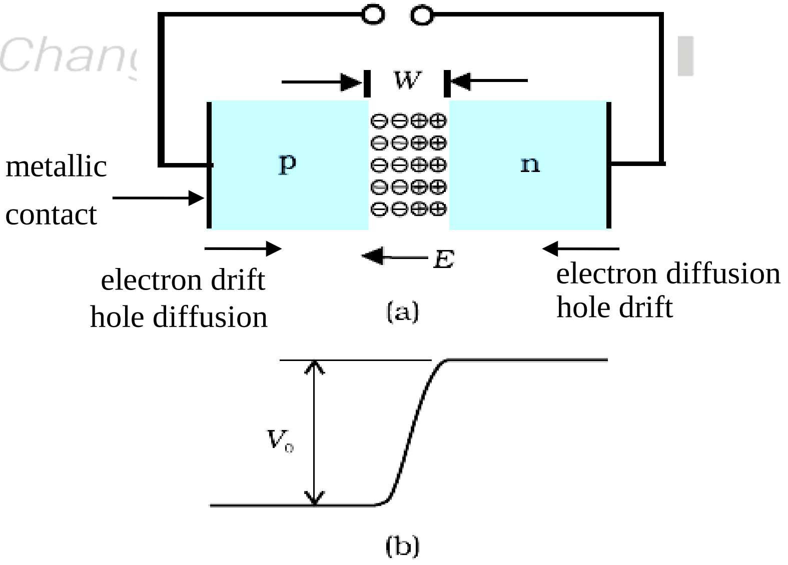
$$\Rightarrow n_h = \frac{(n_i)^2}{n_e} = \frac{(1.5\times10^{16}\,\mathrm{m}^{-3})^2}{5\times10^{22}\,\mathrm{m}^{-3}}$$

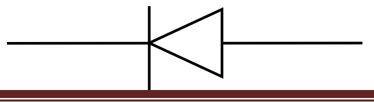
$$\Rightarrow n_h = \frac{(1.5\times10^{16}\,\mathrm{m}^{-3})^2}{5\times10^{22}\,\mathrm{m}^{-3}}$$

$$= 4.5\times10^{9}\,\mathrm{m}^{-3}$$

P-N JUNCTION FORMATION

Consider thin p-type silicon (p-Si) semiconductor wafer. By adding precisely a small quantity of pentavalent impurity, part of the p-Si wafer can be converted into n-Si. The





symbol for diode

wafer now contains p-region and n-region and a metallurgical junction between p-, and n-region. Two important processes occur during the formation of a p-n junction: diffusion and drift. Due to the concentration gradient across p-, and n- sides, holes diffuse from p-side to n-side, and electrons diffuse from n-side to p-side. This motion of charge carries gives rise to diffusion current across the junction. A layer of positive charge (or positive space-charge region) on n-side of the junction is developed and a layer of negative charge (or negative space-charge region) on the p-side of the junction is developed. This space-charge region on either side of the junction together is known as the depletion region. An electric field directed from n-region towards p-region develops. Due to this field, an electron on the p-side of the junction moves to n-side, and a hole on n-side of the junction moves to the p-side. This motion of charge carriers due to the electric field constitutes drift current. Initially, the diffusion current is large, and the drift current is small. In a p-n junction under equilibrium, the diffusion current equals the drift current and there is no net current. In equilibrium V_0 is the barrier potential.

QUESTION

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

ANS

No! Any slab, howsoever flat, will have roughness much larger than the inter-atomic crystal spacing hence continuous contact at the atomic level will not be possible. The junction will behave as a discontinuity for the flowing charge carriers.

P-N JUNCTION UNDER FORWARD BIAS

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 it is said to be forward-biased. The direction of the applied voltage (V) is opposite to the built-in potential V_0. As a result, the depletion layer width decreases, and the effective barrier height under forward bias is ($V_0 - V$). This allows for more diffusion to take place. The drift current remains almost unchanged. Thus the diffusion current exceeds the drift current and there is a net current from p-side to n-side across the junction.

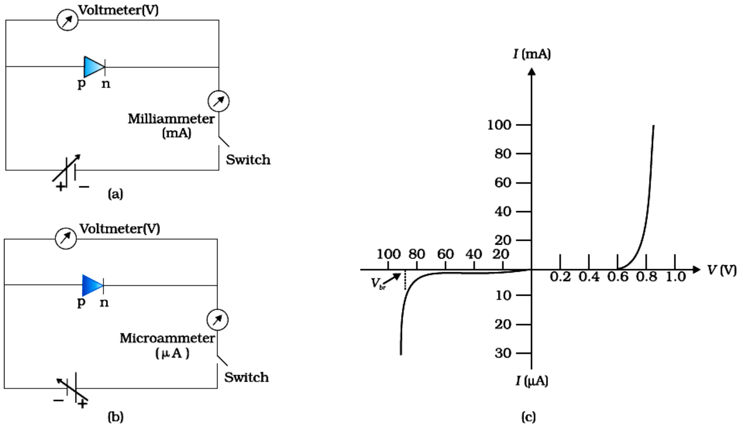
P-N JUNCTION UNDER REVERSE BIAS

When an external voltage V is applied across a semiconductor diode such that p-side is connected to the negative terminal of the battery and n-side to the positive terminal it is said to be reverse biased. The direction of the applied voltage (V) is in the same direction as the built-in potential V_0. As a result, the depletion layer width increases, and the effective barrier height under forward bias is ($V_0 + V$). Diffusion becomes more difficult. The drift current remains almost unchanged. Thus the drift current exceeds the diffusion current and there is a net current from n-side to p-side across the junction. However, this current is very small.

forward-biased

reverse-biased

I-V CHARACTERISTIC OF DIODE



In forward bias, the current first increases very slowly, almost negligibly, till the voltage across the diode crosses a certain value. After the characteristic voltage, the diode current increases significantly (exponentially), even for a very small increase in the diode bias voltage. This voltage is called the threshold voltage or cut-in voltage (~0.2V for germanium diode and ~0.7 V for a silicon diode).

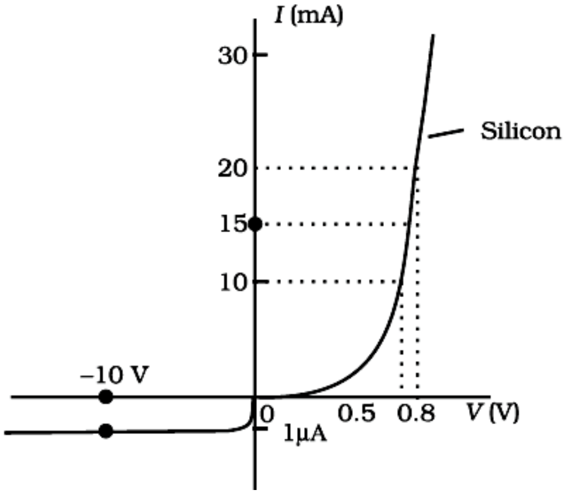
For the diode in reverse bias, the current is very small and almost remains constant with a change in bias. It is called reverse saturation current. However, for special cases, at very high reverse bias (break down voltage), the current suddenly increases.

For diodes, we define a quantity called dynamic resistance as the ratio of small change in voltage ΔV to a small change in the current ΔI :

$$r_d = \frac{\Delta V}{\Delta I}$$

PROBLEM-

The V-I characteristic of a silicon diode is shown in the following figure. Calculate the



dynamic resistance at (a) $I_D = 15\text{A}$ (b) $V_D = -10\text{V}$.

SOL-

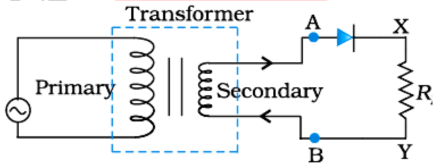
a) $r_d = \frac{\Delta V}{\Delta I} = \frac{0.8-0.7}{10} 10^3 \Omega = 10\Omega$

a) $r_d = \frac{\Delta V}{\Delta I} = \frac{-10}{-1} 10^6 \Omega = 10^7 \Omega$

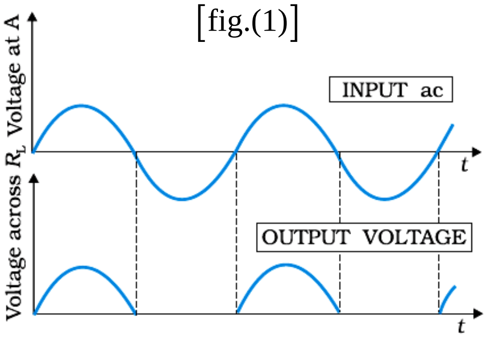
DIODE AS A RECTIFIER

The diode allows current to pass only when it is forward biased. This property is used to rectify alternating voltages and the circuit used for this purpose is called a rectifier.

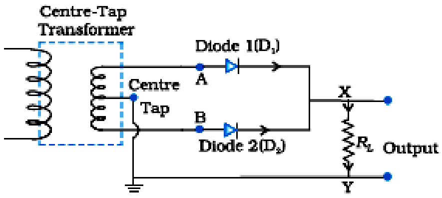
A half-wave rectifier circuit is shown in fig(1). The secondary of a transformer supplies the desired ac voltage across terminals A and B. When the voltage at A is positive, the diode is forward biased and it conducts. When A is negative, the diode is reverse-biased and it does not conduct. Therefore we get an output voltage, as shown in fig. (2). Thus, the output voltage, though still varying, is restricted to only one direction and is said to be rectified. Since the rectified output of this circuit is only for half of the input ac wave it is called a half-wave rectifier.



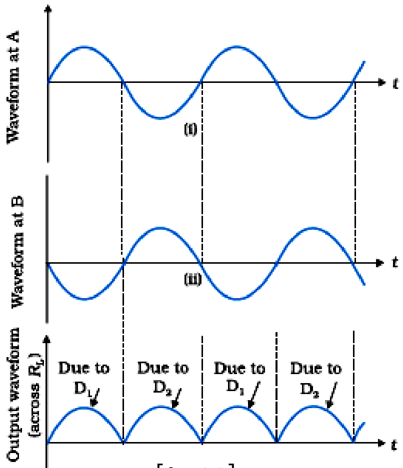
[fig.(1)]



[fig.(2)]



[fig.(3)]

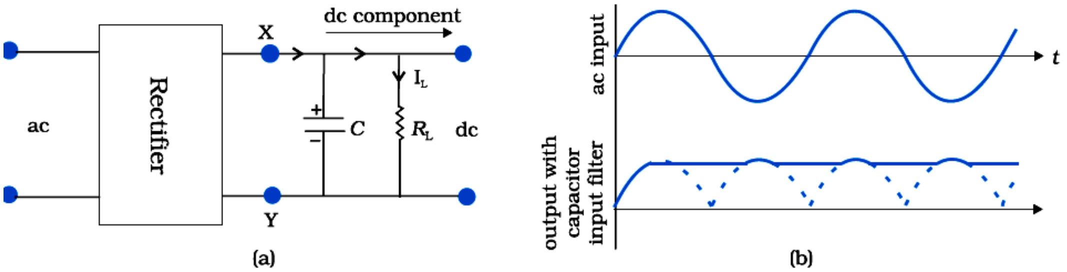


[fig.(4)]

A full-wave rectifier circuit is shown in fig(3). Here the p-side of the two diodes is connected to the ends of the secondary of the transformer. The n-side of the diodes are connected together and the output is taken between this common point of diodes and the midpoint of the secondary of the center-tap transformer. As can be seen from fig. (4) the voltage rectified by each diode is only half the total secondary voltage. Thus, the output between common terminals of the diodes and the center tap of the transformer becomes a full-wave rectifier output. Suppose the input voltage to A with respect to the center tap at any instant is positive. It is clear that, at that instant, the voltage at B being out of phase will be negative. So, diode D_1 gets forward biased and conducts while D_2 being reverse biased is not conducting. When the voltage at A becomes negative with respect to the center tap, the voltage at B would be positive. In this part of the cycle, the diode D_1 would not conduct but the diode D_2 would, giving an output current and output voltage (across R_L) during the negative half cycle of the input ac.

FILTER CIRCUITS

The rectified voltage is in the form of pulses of the shape of half sinusoids. Though it is unidirectional it does not have a steady value. To get steady dc output from the pulsating voltage normally a capacitor is connected across the output terminals (parallel to the load R_L). One can also use an inductor in series with R_L for the same purpose. Since these additional circuits appear to filter out the ac ripple and give a pure dc voltage, so they are called filters.



(a) A full-wave rectifier with capacitor filter, (b) Input and output voltage of rectifier in (a).

When the voltage across the capacitor is rising, it gets charged. It gets discharged through the load and the voltage across it begins to fall. The rate of fall of the voltage across the capacitor depends upon the inverse product of capacitor C and the effective resistance R_L used in the circuit and is called the time constant. To make the time constant large value of C should be large. So capacitor input filters use large capacitors.

LED

It is a heavily doped p-n junction that under forward bias emits spontaneous radiation. The diode is encapsulated with a transparent cover so that emitted light can come out. When the diode is forward biased, electrons are sent from n to p (where they are minority carriers) and holes are sent from p to n (where they are minority carriers). At the junction boundary, the concentration of minority carriers increases compared to the equilibrium concentration (i.e., when there is no bias). Thus at the junction boundary on either side of the junction, excess minority carriers are there which recombine with majority carriers near the junction. On recombination, the energy is released in the form of photons. Photons with energy equal to or slightly less than the bandgap are emitted. When the forward current of the diode is small, the intensity of light emitted is small. As the forward current increases, the intensity of light increases and reaches a maximum. Further, an increase in the forward current results in a decrease of light intensity. LEDs are biased such that the light-emitting efficiency is maximum. The V-I characteristics of a LED are similar to that of a Si junction diode. But the threshold voltages are much higher and slightly different for each color. The reverse breakdown voltages of LEDs are very low, typically around 5V. So care should be taken that high reverse voltages do not appear across them. LEDs that can emit red, yellow, orange, green, and blue light are commercially available. The semiconductor used for the fabrication of visible LEDs must at least have a bandgap of 1.8 eV (spectral range of visible light is from about $0.4\mu\text{m}$ to $0.7\mu\text{m}$, i.e., from about 3 eV to 1.8 eV). The compound semiconductor Gallium Arsenide – Phosphide ($GaAs_{1-x}P_x$) is used for making LEDs of different colors. $GaAs_{0.6}P_{0.4}$ (E_g ~ 1.9 eV) is used for the red LED. GaAs (E_g ~ 1.4 eV) is used for making infrared LED. These LEDs find extensive use in remote controls, burglar alarm systems, optical communication, etc. Extensive research is being done for developing white LEDs which can replace incandescent lamps. LEDs have the following advantages over conventional incandescent low power lamps:

(i) Low operational voltage and less power.

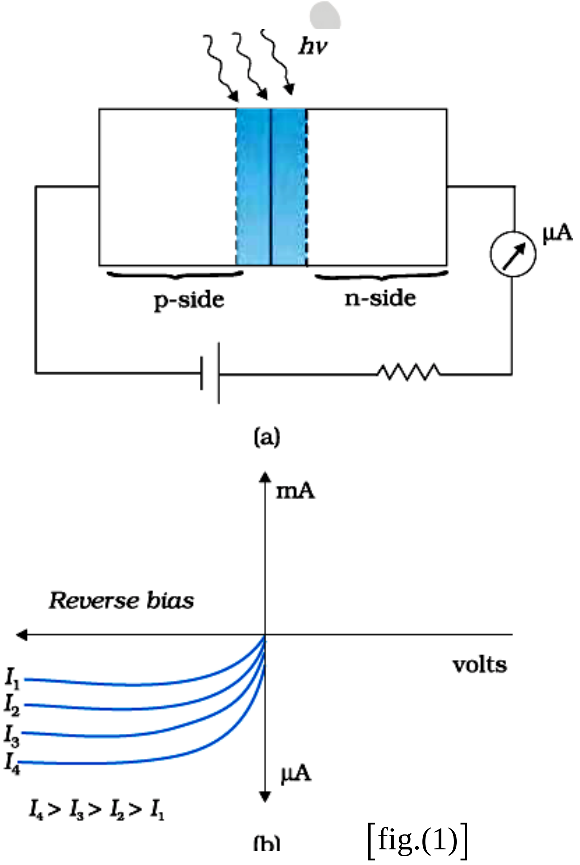
(ii) Fast action and no warm-up time required.

(iii) The bandwidth of emitted light is $100\,\mathring{A}$ to $500\,\mathring{A}$ or in other words it is nearly (but not exactly) monochromatic.

(iv) Long life and ruggedness.

(v) Fast on-off switching capability.

PHOTODIODE



[fig.(1)]

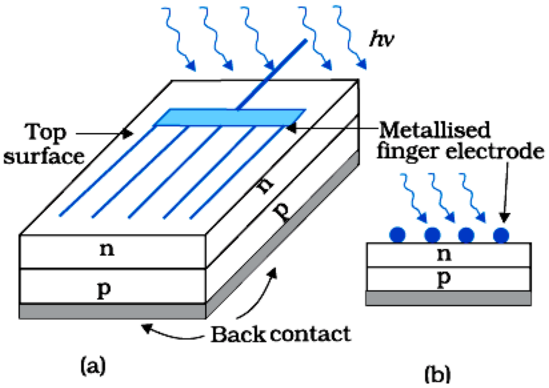
A Photodiode is again a special purpose p-n junction diode fabricated with a transparent window to allow light to fall on the diode. It is operated under reverse bias. When the photodiode is illuminated with light (photons) with energy ($h\nu$) greater than the energy gap (E_g) of the semiconductor, then electron-hole pairs are generated due to the absorption of photons. The diode is fabricated such that the generation of e-h pairs takes place in or near the depletion region of the diode. Due to the electric field of the junction, electrons and holes are separated before they recombine. The direction of the electric field is such that electrons reach the n-side and holes reach the p-side. Electrons are collected on the n-side and holes are collected on the p-side giving rise to an emf. When an external load is connected, current flows. The magnitude of the photocurrent depends on the intensity of incident light (photocurrent is proportional to incident light intensity). It is easier to observe the change in the current with the change in the light intensity if a reverse bias is applied. Thus photodiode can be used as a photodetector to detect optical signals. The circuit diagram used for the measurement of I-V characteristics of a photodiode is shown in fig. (1) and typical I-V characteristics in fig. (2).

QUESTION

The current in the forward bias is known to be more than the current in the reverse bias. What is the reason to operate the photodiodes in reverse bias?

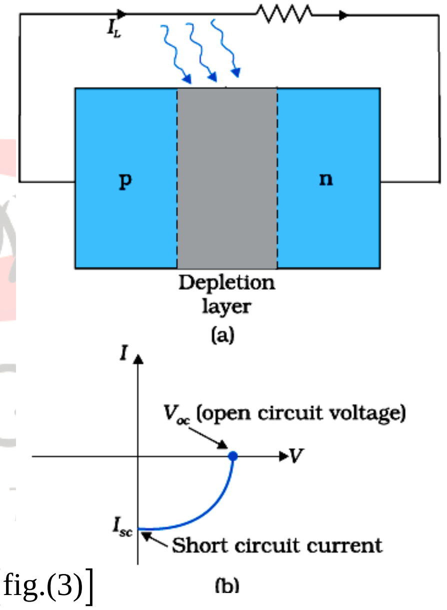
ANS On illumination the fractional change in the majority carriers would be much less than that in the minority carriers. In general, we can state that the fractional change due to the photo-effects on the minority carrier dominated reverse bias current is more easily measurable than the fractional change in the forward bias current. Hence, photodiodes are preferably used in the reverse bias condition for measuring light intensity.

SOLAR CELL



[fig.(2)]

A solar cell is a p-n junction that generates emf when solar radiation falls on the p-n junction. It works on the same principle (photovoltaic effect) as the photodiode, except that no external bias is applied and the junction area is kept much larger for solar radiation to be incident because we are interested in more power. A simple p-n junction solar cell is shown in fig. (2). A p-Si wafer of about 300μm is taken over which a thin layer □ 0.3μm of n-Si is grown on one-side by diffusion process. The other side of p-Si is coated with a metal (back contact). On the top of n-Si layer, a metal finger electrode (or metallic grid) is deposited. This acts as a front contact. The metallic grid occupies only a very small fraction of the cell area (<15%) so that light can be incident on the cell from the top. The generation of emf by a solar cell, when light falls on, it is due to the following three basic processes: generation, separation and collection



[fig.(3)]

(i) generation of e-h pairs due to light (with $h\nu > E_g$) close to the junction;

(ii) separation of electrons and holes due to the electric field of the depletion region. Electrons are swept to n-side and holes to p-side;

(iii) the electrons reaching the n-side are collected by the front contact and holes reaching the p-side are collected by the back contact. Thus the p-side becomes positive and the n-side becomes negative giving rise to photovoltage.

When an external load is connected as shown in fig. (3)(a) a photocurrent I_L flows through the load. A typical I-V characteristic of a solar cell is shown in fig. (3)(b). Semiconductors with bandgap close to 1.5 eV are ideal materials for solar cell fabrication. Solar cells are made with semiconductors like Si (E_g = 1.1 eV), GaAs

(E_g = 1.43 eV), CdTe (E_g = 1.45 eV), CuInSe2 (E_g = 1.04 eV), etc.

The important criteria for the selection of a material for solar cell fabrication are

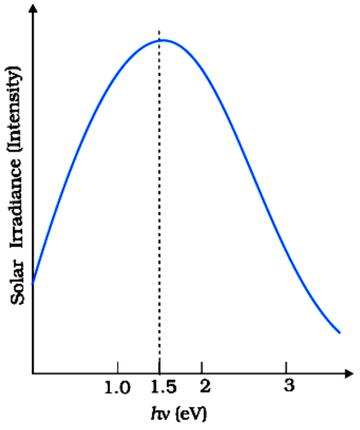
(i) the bandgap (~1.0 to 1.8 eV),

(ii) high optical absorption (□ $10^4\,\text{cm}^{-1}$),

(iii) electrical conductivity,

(iv) availability of the raw material, and

(v) cost.



[fig.(4)]

Note that sunlight is not always required for a solar cell. Any light with photon energies greater than the bandgap will do. Solar cells are used to power electronic devices in satellites and space vehicles and also as the power supply to some calculators. Production of low-cost photovoltaic cells for large-scale solar energy is a topic for research.

QUESTION

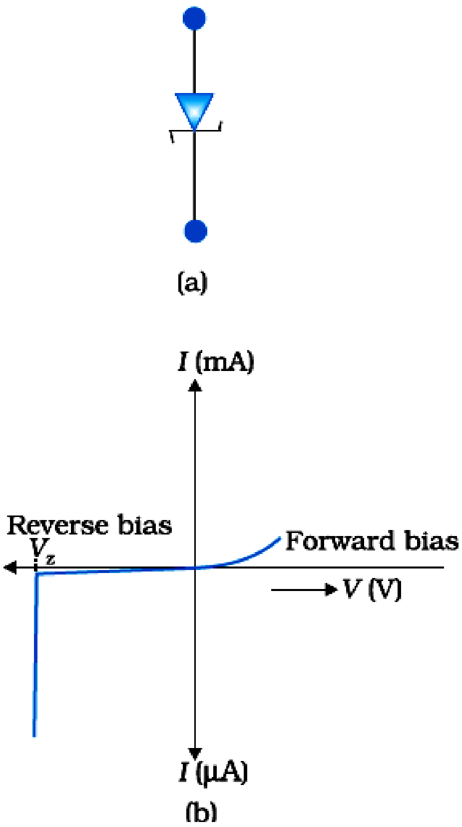
Why are Si and GaAs are preferred materials for solar cells?

ANS

The solar radiation spectrum received by us is shown in fig. (4). The maxima are near 1.5 eV. For photo-excitation, $h\nu > E_g$. Hence, a semiconductor with a bandgap ~1.5 eV or lower is likely to give better solar conversion efficiency. Silicon has Eg ~ 1.1 eV while for GaAs it is ~1.53 eV. GaAs is better (despite its higher bandgap) than Si because of its relatively higher absorption coefficient. If we choose materials like CdS or CdSe (Eg ~ 2.4 eV), we can use only the high energy component of the solar energy for photo-conversion and a significant part of the energy will be of no use. The question arises: why we do not use a material like PbS (E_g ~ 0.4 eV) which satisfies the condition $h\nu > E_g$ for ν maxima corresponding to the solar radiation spectra? If we do so, most of the solar radiation will be absorbed on the top layer of the solar cell and will not reach in or near the depletion region. For effective electron-hole separation, due to the junction field, we want the photo-generation to occur in the junction region only.

ZENER DIODE

It is a special purpose semiconductor diode, named after its inventor C. Zener. It is designed to operate under reverse bias in the breakdown region and used as a voltage regulator. The symbol for the Zener diode is shown in fig. (5)(a). Zener diode is fabricated by heavily doping both p- and n- sides of the junction. Due to this, the depletion region formed is very thin ($< 10^{-6}\,\mathrm{m}$) and the electric field of the junction is extremely high (□ $5\times10^{6}\,\mathrm{V/m}$) even for a small reverse bias voltage of about 5V. The I-V characteristics of a Zener diode are shown in fig. (5)(b). It is seen that when the applied reverse bias voltage(V) reaches the breakdown voltage (Vz) of the Zener diode, there is a large change in the current. Note that after the breakdown voltage Vz, a large change in the current can be produced by almost insignificant change in the reverse bias voltage. This property of the Zener diode is used for regulating supply voltages so that they are constant.

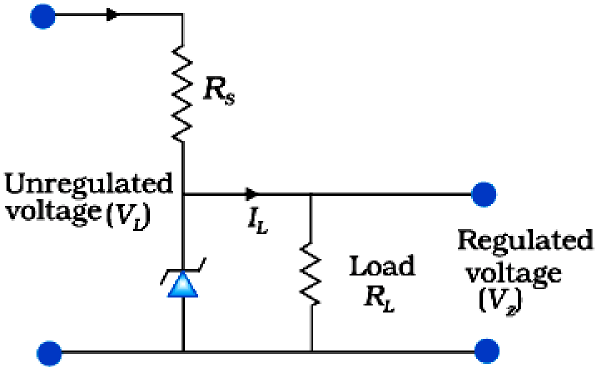


[fig.(5)]

Let us understand how reverse current suddenly increases at the breakdown voltage. We know that reverse current is due to the flow of electrons (minority carriers) from p to n and holes from n to p. As the reverse bias voltage is increased, the electric field at the junction becomes significant. When the reverse bias voltage V = Vz, then the electric field strength is high enough to pull valence electrons from the host atoms on the p-side which are accelerated to the n-side. These electrons account for the high current observed at the breakdown. The emission of electrons from the host atoms due to the high electric field is known as internal field emission or field ionization. The electric field required for field ionization is of the order of $10^6\,\text{V/m}$.

ZENER DIODE AS VOLTAGE REGULATOR

We know that when the ac input voltage of a rectifier fluctuates, its rectified output also fluctuates. To get a constant dc voltage from the dc unregulated output of a rectifier, we use a Zener diode. The circuit diagram of a voltage regulator using a Zener diode is shown in fig. (1).



[fig.(1)]

The unregulated dc voltage (filtered output of a rectifier) is connected to the Zener diode through a series resistance R_S such that the Zener diode is reverse biased. If the input voltage increases, the current through R_S and Zener diode also increases. This increases the voltage drop across R_S without any change in the voltage across the Zener diode. This is because, in the breakdown region, Zener voltage remains constant even though the current through the Zener diode changes. Similarly, if the input voltage decreases, the current through R_S and Zener diode also decreases. The voltage drop across R_S decreases without any change in the voltage across the Zener diode. Thus any increase or decrease in the input voltage results in, increase or decrease of the voltage drop across R_S without any change in voltage across the Zener diode. Thus the Zener diode acts as a voltage regulator.

We have to select the Zener diode according to the required output voltage and accordingly the series resistance R_S.

PROBLEM

In a Zener regulated power supply, a Zener diode $V_Z = 6\text{V}$ is used for regulation. The load current is to be 4.0 mA and the unregulated input is 10.0 V. What should be the value of the series resistor R_S?

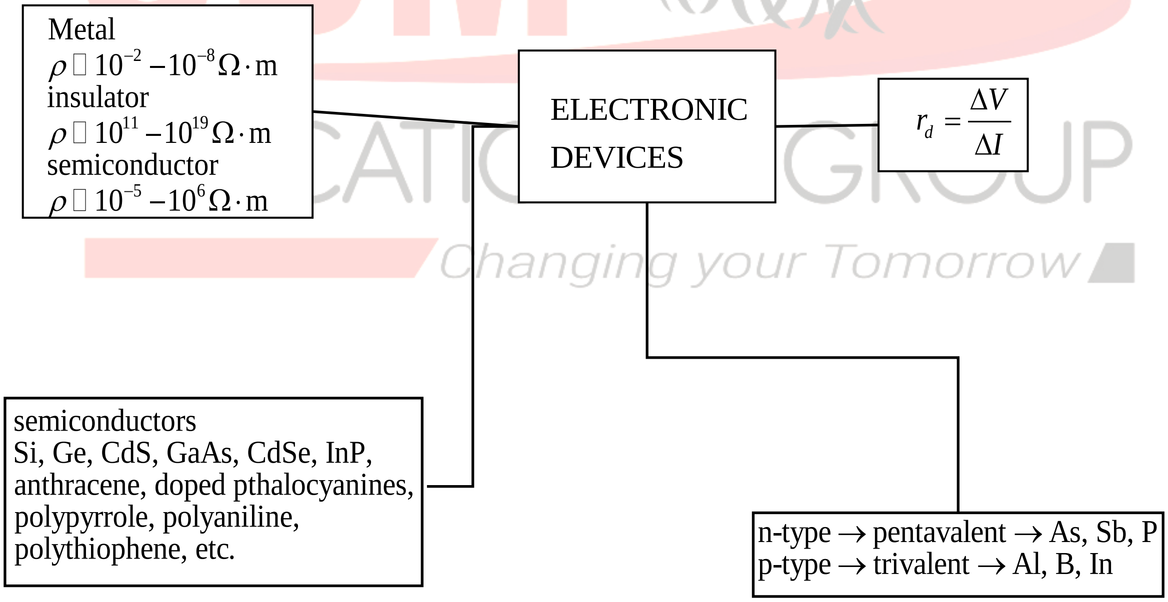
SOL-

The value of R_S should be such that the current through the Zener diode is much larger than the load current. This is to have good load regulation. Choose Zener current as five times the load current, i.e., $I_Z = 20\text{mA}$. The total current through R_S is, therefore, 24 mA. The voltage drop across R_S is $(10-6)\text{V} = 4\text{V}$. This gives

$$R_S = \frac{4\text{V}}{24\text{mA}} = 167\Omega.$$

The nearest value of carbon resistor is 150Ω. So, a series resistor of 150Ω is appropriate. Note that a slight variation in the value of the resistor does not matter, what is important is that the current I_Z should be sufficiently larger than I_L.

mind map



ELECTRONIC DEVICES

Metal
$\rho \sim 10^{-2} - 10^{-8}\,\Omega \cdot \mathrm{m}$
insulator
$\rho \sim 10^{11} - 10^{19}\,\Omega \cdot \mathrm{m}$
semiconductor
$\rho \sim 10^{-5} - 10^{6}\,\Omega \cdot \mathrm{m}$

$r_d = \frac{\Delta V}{\Delta I}$

semiconductors
Si, Ge, CdS, GaAs, CdSe, InP, anthracene, doped pthalocyanines, polypyrrole, polyaniline, polythiophene, etc.

n-type → pentavalent → As, Sb, P
p-type → trivalent → Al, B, In