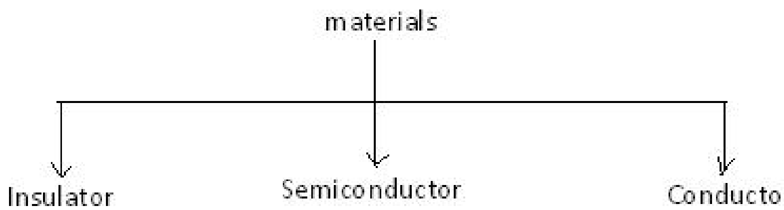
UNIT-I

PN JUNCTION DIODE

1.0 INTRODUCTON

1.0.1. Based on the electrical conductivity all the materials in nature are classified as insulators, semiconductors, and conductors.

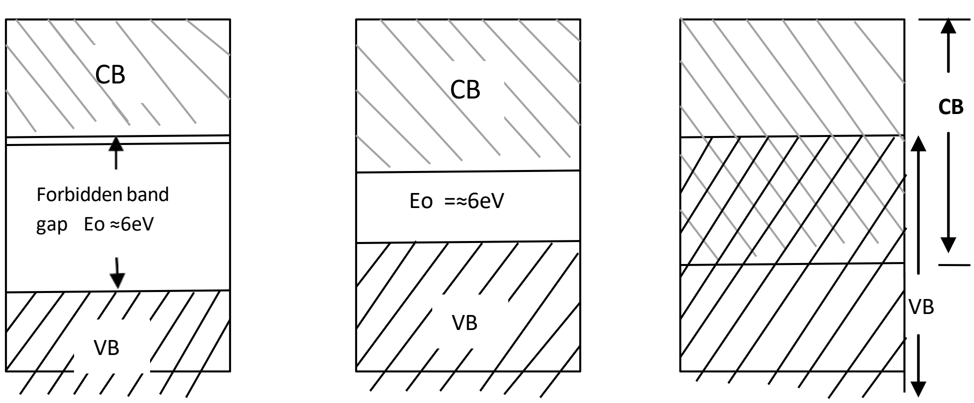


Insulator: An insulator is a material that offers a very low level (or negligible) of conductivity when voltage is applied. Eg: Paper, Mica, glass, quartz. Typical resistivity level of an insulator is of the order of 10^{10} to 10^{12} Ω-cm. The energy band structure of an insulator is shown in the fig.1.1. Band structure of a material defines the band of energy levels that an electron can occupy. Valance band is the range of electron energy where the electron remain bended too the atom and do not contribute to the electric current. Conduction bend is the range of electron energies higher than valance band where electrons are free to accelerate under the influence of external voltage source resulting in the flow of charge.

The energy band between the valance band and conduction band is called as forbidden band gap. It is the energy required by an electron to move from balance band to conduction band i.e. the energy required for a valance electron to become a free electron.

$$1 \text{ eV} = 1.6 \times 10^{-19} \text{ J}$$

For an insulator, as shown in the fig.1.1 there is a large forbidden band gap of greater than 5Ev. Because of this large gap there a very few electrons in the CB and hence the conductivity of insulator is poor. Even an increase in temperature or applied electric field is insufficient to transfer electrons from VB to CB.



Insulator Semiconductor Conductor

FiG:1.1 Energy band diagrams insulator, semiconductor and conductor

Conductors: A conductor is a material which supports a generous flow of charge when a voltage is applied across its terminals. i.e. it has very high conductivity. Eg: Copper, Aluminum, Silver, Gold. The resistivity of a conductor is in the order of 10^{-4} and 10^{-6} Ω-cm. The Valance and conduction bands overlap (fig1.1) and there is no energy gap for the electrons to move from valance band to conduction band. This implies that there are free electrons in CB even at absolute zero temperature (0K). Therefore at room temperature when electric field is applied large current flows through the conductor.

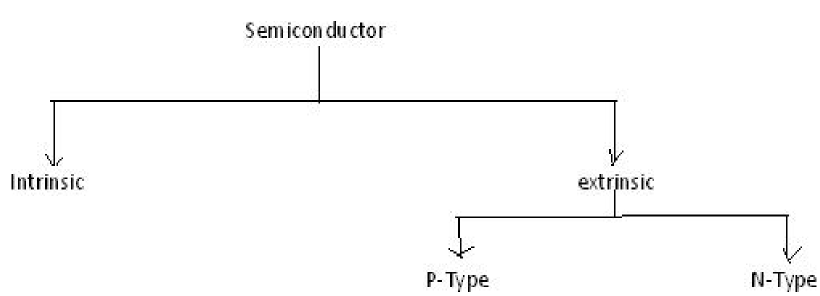
Semiconductor: A semiconductor is a material that has its conductivity somewhere between the insulator and conductor. The resistivity level is in the range of 10 and 10^4 Ω-cm. Two of the most commonly used are Silicon (Si=14 atomic no.) and germanium (Ge=32 atomic no.). Both have 4 valance electrons. The forbidden band gap is in the order of 1eV. For eg., the band gap energy for Si, Ge and GaAs is 1.21, 0.785 and 1.42 eV, respectively at absolute zero temperature (0K). At 0K and at low temperatures, the valance band electrons do not have sufficient energy to move from V to CB. Thus semiconductors act a insulators at 0K. as the temperature increases, a large number of valance electrons acquire sufficient energy to leave the VB, cross the forbidden bandgap and reach CB. These are now free electrons as they can move freely under the influence of electric field. At room temperature there are sufficient electrons in the CB and hence the semiconductor is capable of conducting some current at room temperature.

Inversely related to the conductivity of a material is its resistance to the flow of charge or current. Typical resistivity values for various materials' are given as follows.

Insulator	Semiconductor	Conductor
10^{-6} Ω-cm (Cu)	50Ω-cm (Ge)	10^{12} Ω-cm (mica)
	$50x10^3$ Ω-cm (Si)	

Typical resistivity values

1.0.1 Semiconductor Types



A pure form of semiconductors is called as intrinsic semiconductor. Conduction in intrinsic sc is either due to thermal excitation or crystal defects. Si and Ge are the two most important semiconductors used. Other examples include Gallium arsenide GaAs, Indium Antimonide (InSb) etc.

Let us consider the structure of Si. A Si atomic no. is 14 and it has 4 valance electrons. These 4 electrons are shared by four neighboring atoms in the crystal structure by means of covalent bond. Fig. 1.2a shows the crystal structure of Si at absolute zero temperature (0K). Hence a pure SC acts has poor conductivity (due to lack of free electrons) at low or absolute zero temperature.

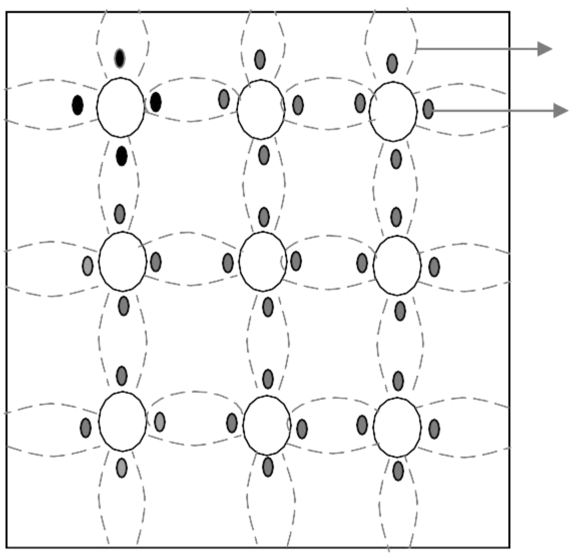


Fig. 1.2a crystal structure of Si at 0K

At room temperature some of the covalent bonds break up to thermal energy as shown in fig 1.2b. The valance electrons that jump into conduction band are called as free electrons that are available for conduction.

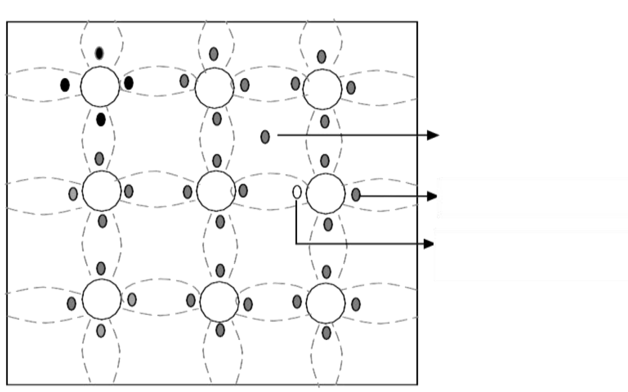


Fig. 1.2b crystal structure of Si at room temperature0K

The absence of electrons in covalent bond is represented by a small circle usually referred to as hole which is of positive charge. Even a hole serves as carrier of electricity in a manner similar to that of free electron.

The mechanism by which a hole contributes to conductivity is explained as follows:

When a bond is in complete so that a hole exists, it is relatively easy for a valance electron in the neighboring atom to leave its covalent bond to fill this hole. An electron moving from a bond to fill a hole moves in a direction opposite to that of the electron. This hole, in its new position may now be filled by an electron from another covalent bond and the hole will correspondingly move one more step in the direction opposite to the motion of electron. Here we have a mechanism for conduction of electricity which does not involve free electrons. This phenomenon is illustrated in fig1.3

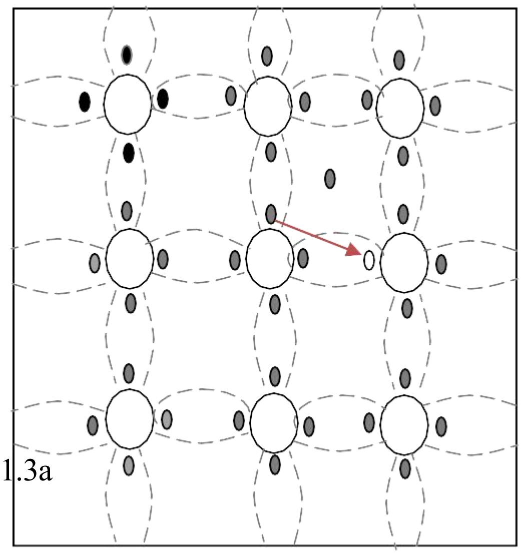
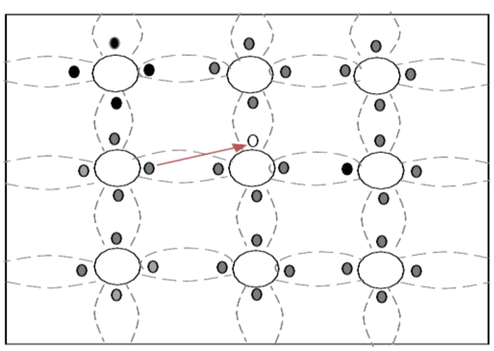


Fig. 1.3a

→ Electron

movement ← Hole

movement



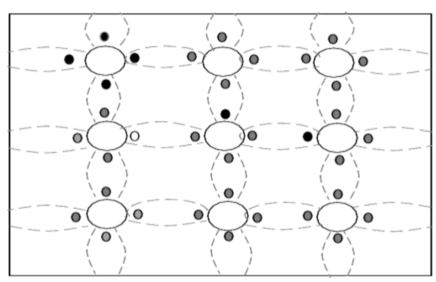


Fig. 1.3c

Fig 1.3a show that there is a hole at ion 6.Imagine that an electron from ion 5 moves into the hole at ion 6 so that the configuration of 1.3b results. If we compare both fig1.3a &fig 1.3b, it appears as if the hole has moved towards the left from ion6 to ion 5. Further if we compare fig 1.3b and fig 1.3c, the hole moves from ion5 to ion 4. This discussion indicates the motion of hole is in a direction opposite to that of motion of electron. Hence we consider holes as physical entities whose movement constitutes flow of current.

In a pure semiconductor, the number of holes is equal to the number of free electrons.

EXTRINSIC SEMICONDUCTOR

Intrinsic semiconductor has very limited applications as they conduct very small amounts of current at room temperature. The current conduction capability of intrinsic semiconductor can be increased significantly by adding a small amounts impurity to the intrinsic semiconductor. By adding impurities it becomes impure or extrinsic semiconductor. This process of adding impurities is called as doping. The amount of impurity added is 1 part in 10^6 atoms.

N type semiconductor: If the added impurity is a pentavalent atom then the resultant semiconductor is called N-type semiconductor. Examples of pentavalent impurities are Phosphorus, Arsenic, Bismuth, Antimony etc.

A pentavalent impurity has five valance electrons. Fig 1.4a shows the crystal structure of N-type semiconductor material where four out of five valance electrons of the impurity atom(antimony) forms covalent bond with the four intrinsic semiconductor atoms. The fifth electron is loosely bound to the impurity atom. This loosely bound electron can be easily

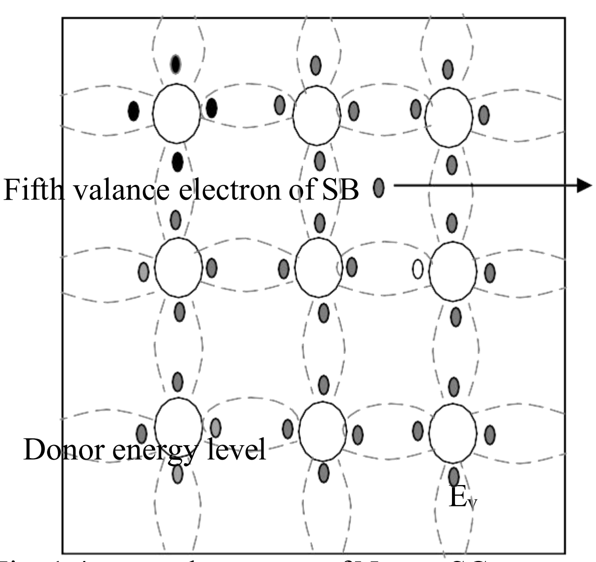


Fig. 1.4a crystal structure of N type SC

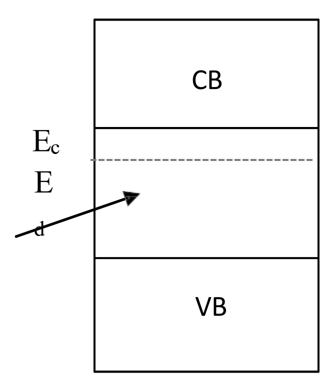


Fig. 1.4bEnergy band diagram of N type

Excited from the valance band to the conduction band by the application of electric field or increasing the thermal energy. The energy required to detach the fifth electron form the impurity atom is very small of the order of 0.01ev for Ge and 0.05 eV for Si.

The effect of doping creates a discrete energy level called donor energy level in the forbidden band gap with energy level E_d slightly less than the conduction band (fig 1.4b). The difference between the energy levels of the conducting band and the donor energy level is the energy required to free the fifth valance electron (0.01 eV for Ge and 0.05 eV for Si). At room temperature almost all the fifth electrons from the donor impurity atom are raised to conduction band and hence the number of electrons in the conduction band increases significantly. Thus every antimony atom contributes to one conduction electron without creating a hole.

In the N-type sc the no. of electrons increases and the no. of holes decreases compared to those available in an intrinsic sc. The reason for decrease in the no. of holes is that the larger no. of electrons present increases the recombination of electrons with holes. Thus current in N type sc is dominated by electrons which are referred to as majority carriers. Holes are the minority carriers in N type sc

P type semiconductor: If the added impurity is a trivalent atom then the resultant semiconductor is called P-type semiconductor. Examples of trivalent impurities are Boron, Gallium , indium etc.

The crystal structure of p type sc is shown in the fig1.5a. The three valance electrons of the impurity (boon) forms three covalent bonds with the neighboring atoms and a vacancy exists in the fourth bond giving rise to the holes. The hole is ready to accept an electron from the neighboring atoms. Each trivalent atom contributes to one hole generation and thus introduces a large no. of holes in the valance band. At the same time the no. electrons are decreased compared to those available in intrinsic sc because of increased recombination due to creation of additional holes.

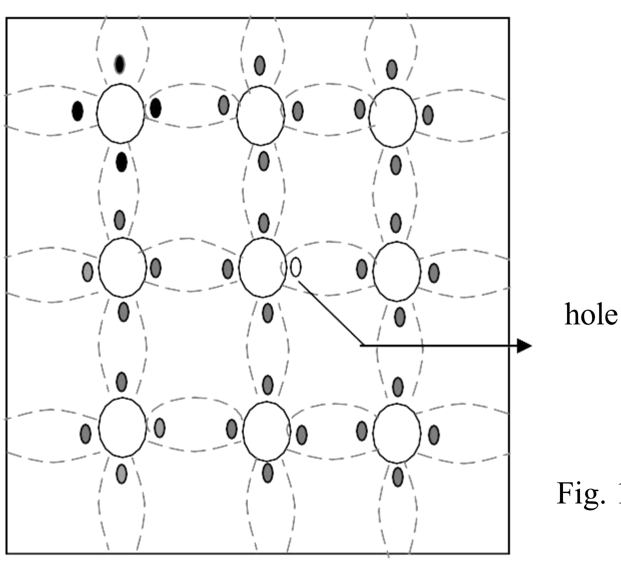


Fig. 1.5a crystal structure of P type sc

Thus in P type sc , holes are majority carriers and electrons are minority carriers. Since each trivalent impurity atoms are capable accepting an electron, these are called as acceptor atoms. The following fig 1.5b shows the pictorial representation of P type sc

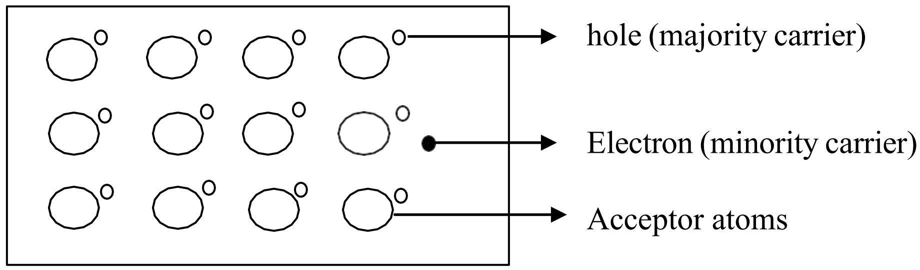


Fig. 1.5b crystal structure of P type sc

- The conductivity of N type sc is greater than that of P type sc as the mobility of electron is greater than that of hole.
- For the same level of doping in N type sc and P type sc, the conductivity of an Ntype sc is around twice that of a P type sc

1.1 QUANTITATIVE THEORY OF PN JUNCTION DIODE

1.1.1 PN JUNCTION WITH NO APPLIED VOLTAGE OR OPEN CIRCUIT CONDITION:

In a piece of sc, if one half is doped by p type impurity and the other half is doped by n type impurity, a PN junction is formed. The plane dividing the two halves or zones is called PN junction. As shown in the fig the n type material has high concentration of free electrons, while p type material has high concentration of holes. Therefore at the junction there is a tendency of free electrons to diffuse over to the P side and the holes to the N side. This process is called diffusion. As the free electrons move across the junction from N type to P type, the donor atoms become positively charged. Hence a positive charge is built on the N-side of the junction. The free electrons that cross the junction uncover the negative acceptor ions by filing the holes. Therefore a negative charge is developed on the p –side of the junction..This net negative charge on the p side prevents further diffusion of electrons into the p side. Similarly the net positive charge on the N side repels the hole crossing from p side to N side. Thus a barrier sis set up near the junction which prevents the further movement of charge carriers i.e. electrons and holes. As a consequence of induced electric field across the depletion layer, an electrostatic potential difference is established between P and N regions, which are called the potential barrier, junction barrier, diffusion potential or contact potential, Vo. The magnitude of the contact potential Vo varies with doping levels and temperature. Vo is 0.3V for Ge and 0.72 V for Si.

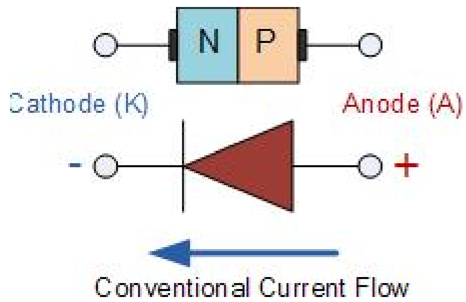


Fig 1.6: Symbol of PN Junction Diode

The electrostatic field across the junction caused by the positively charged N-Type region tends to drive the holes away from the junction and negatively charged p type regions tend to drive the electrons away from the junction. The majority holes diffusing out of the P region leave behind negatively charged acceptor atoms bound to the lattice, thus exposing a negatives pace charge in a previously neutral region. Similarly electrons diffusing from the N region expose positively ionized donor atoms and a double space charge builds up at the junction as shown in the fig. 1.7a

1.1.2 FORWARD BIASED JUNCTION DIODE

When a diode is connected in a **Forward Bias** condition, a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material. If this external voltage becomes greater than the value of the potential barrier, approx. 0.7 volts for silicon and 0.3 volts for germanium, the potential barriers opposition will be overcome and current will start to flow. This is because the negative voltage pushes or repels electrons towards the junction giving them the energy to cross over and combine with the holes being pushed in the opposite direction towards the junction by the positive voltage. This results in a characteristics curve of zero current flowing up to this voltage point, called the "knee" on the static curves and then a high current flow through the diode with little increase in the external voltage as shown below.

Forward Characteristics Curve for a Junction Diode

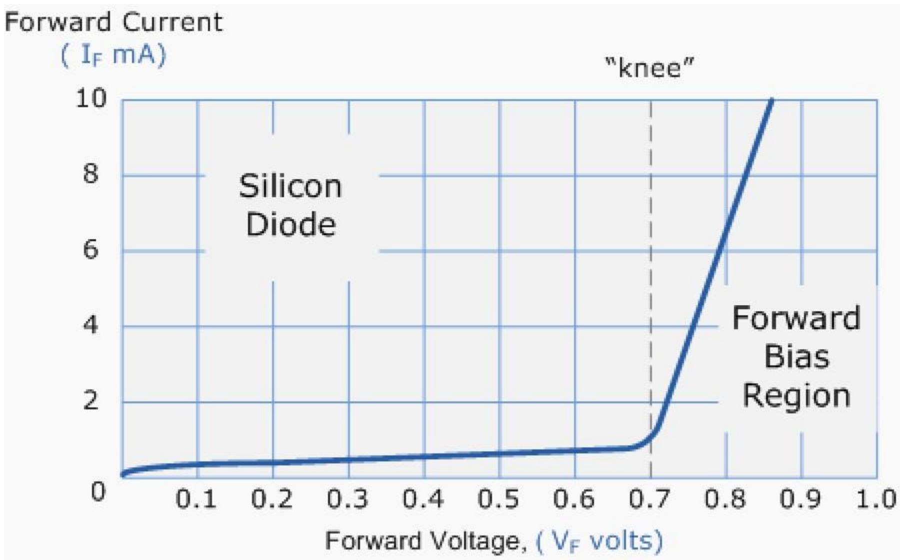


Fig 1.8a: Diode Forward Characteristics

The application of a forward biasing voltage on the junction diode results in the depletion layer becoming very thin and narrow which represents a low impedance path through the junction thereby allowing high currents to flow. The point at which this sudden increase in current takes place is represented on the static I-V characteristics curve above as the "knee" point.

Forward Biased Junction Diode showing a Reduction in the Depletion Layer

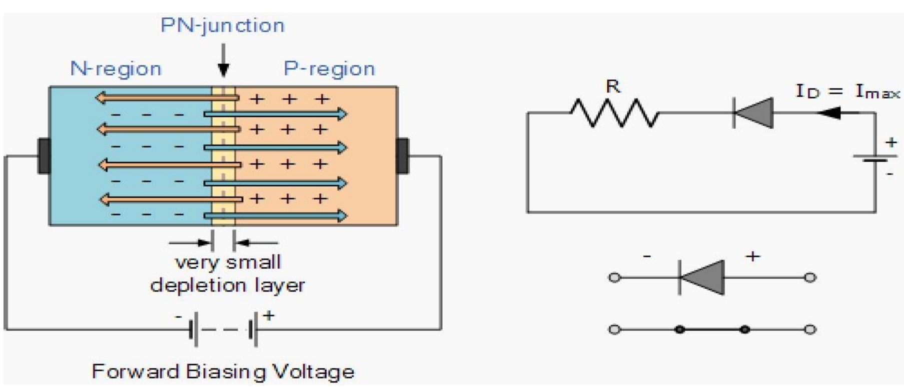


Fig 1.8b: Diode Forward Bias

This condition represents the low resistance path through the PN junction allowing very large currents to flow through the diode with only a small increase in bias voltage. The actual potential difference across the junction or diode is kept constant by the action of the depletion layer at approximately 0.3v for germanium and approximately 0.7v for silicon junction diodes. Since the diode can conduct "infinite" current above this knee point as it effectively becomes a short circuit, therefore resistors are used in series with the diode to limit its current flow. Exceeding its maximum forward current specification causes the device to dissipate more power in the form of heat than it was designed for resulting in a very quick failure of the device.

1.1.2 PN JUNCTION UNDER REVERSE BIAS CONDITION:

Reverse Biased Junction Diode

When a diode is connected in a **Reverse Bias** condition, a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material. The positive voltage applied to the N- type material attracts electrons towards the positive electrode and away from the junction, while the holes in the P-type end are also attracted away from the junction towards the negative electrode. The net result is that the depletion layer grows wider due to a lack of electrons and holes and presents a high impedance path, almost an insulator. The result is that a high potential barrier is created thus preventing current from flowing through the semiconductor material.

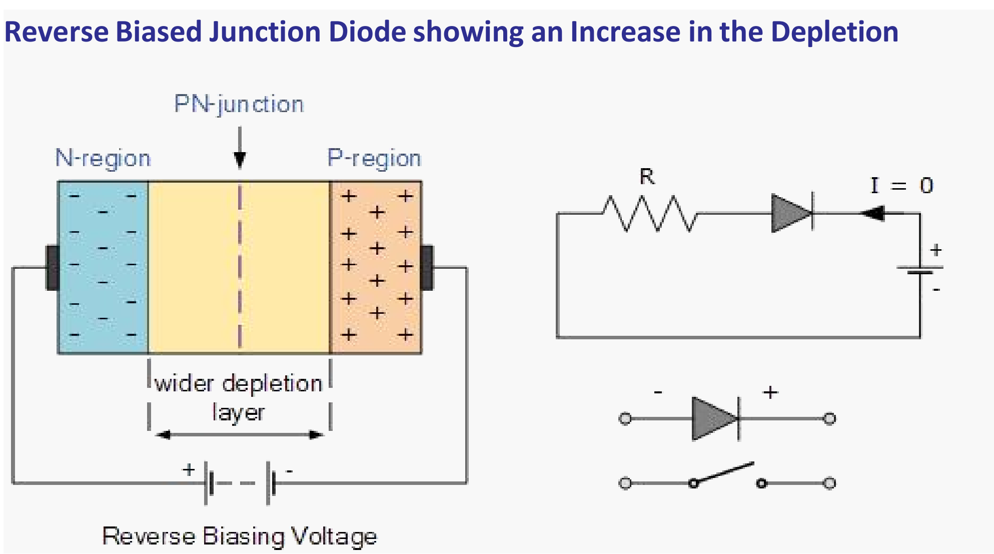


Fig 1.9a: Diode Reverse Bias

This condition represents a high resistance value to the PN junction and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small **leakage current** does flow through the junction which can be measured in microamperes, (μA). One final point, if the reverse bias voltage Vr applied to the diode is increased to a sufficiently high enough value, it will cause the PN junction to overheat and fail due to the avalanche effect around the junction. This may cause the diode to become shorted and will result in the flow of maximum circuit current, and this shown as a step downward slope in the reverse static characteristics curve below.

Reverse Characteristics Curve for a Junction Diode

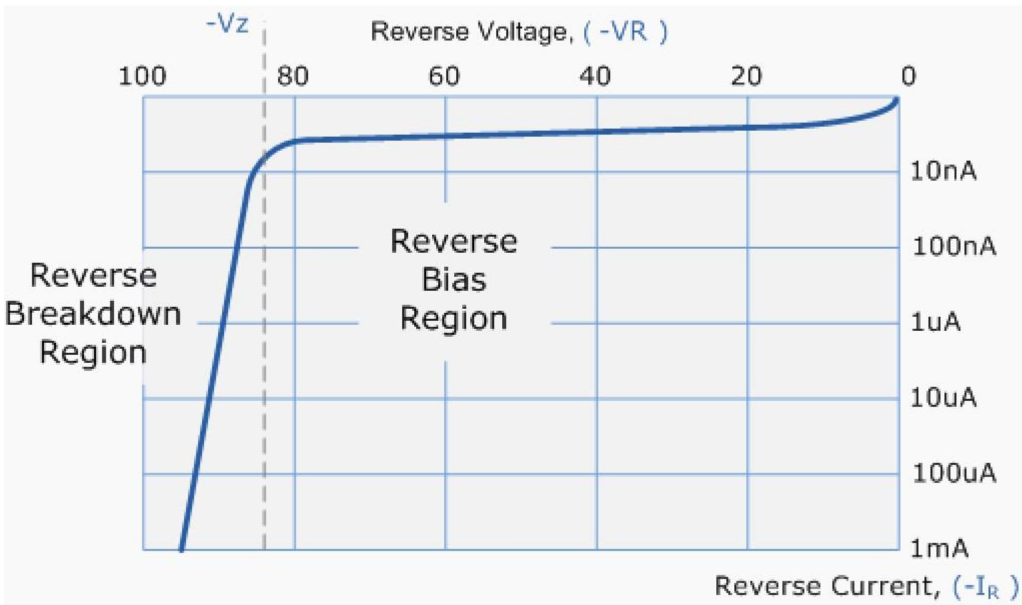


Fig 1.9b: Diode Reverse Characteristics

Sometimes this avalanche effect has practical applications in voltage stabilizing circuits where a series limiting resistor is used with the diode to limit this reverse breakdown current to a preset maximum value thereby producing a fixed voltage output across the diode. These types of diodes are commonly known as **Zener Diodes**

1.2 VI CHARACTERISTICS AND THEIR TEMPERATURE DEPENDENCE

Diode terminal characteristics equation for diode junction current:

$$I_D = I_o\left(e^{V_D/V_T} - 1\right)$$

Where V_T = KT/q;

V_D _ diode terminal voltage, Volts

I_o _ temperature-dependent saturation current, µA T

_ absolute temperature of p-n junction, K

K _ Boltzmann's constant 1.38x 10 -23J/K)

q _ electron charge 1.6x10-19 C

 = empirical constant, 1 for Ge and 2 for Si

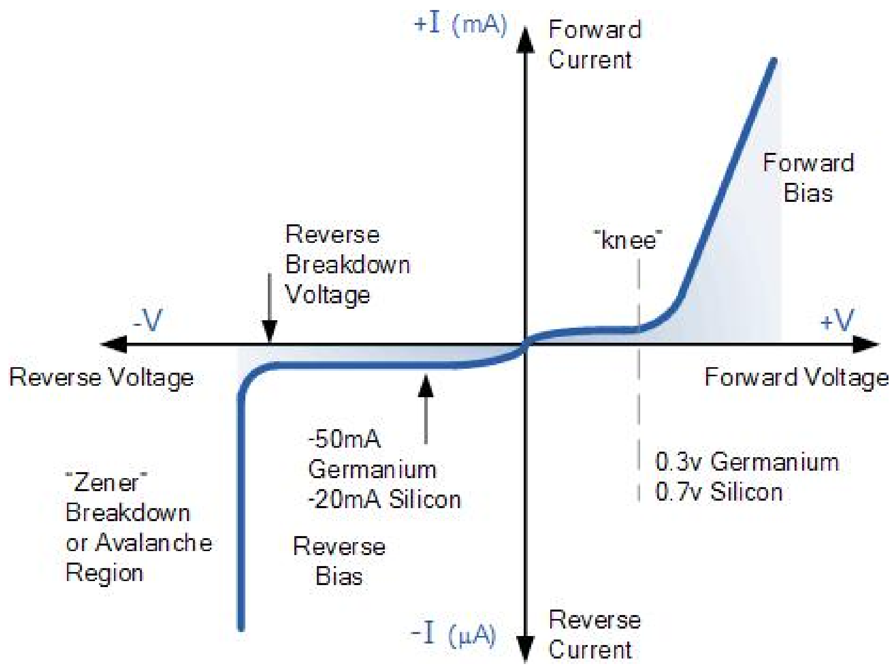


Fig 1.10: Diode Characteristics

1.3 IDEAL VERSUS PRACTICAL RESISTANCE LEVELS

DC or Static Resistance

The application of a dc voltage to a circuit containing a semiconductor diode will result in an operating point on the characteristic curve that will not change with time. The resistance of the diode at the operating point can be found simply by finding the corresponding levels of VD and ID as shown in Fig. 1.12 and applying the following Equation:

$$R_D = \frac{V_D}{I_D}$$

The dc resistance levels at the knee and below will be greater than the resistance levels obtained for the vertical rise section of the characteristics. The resistance levels in the reverse-bias region will naturally be quite high. Since ohmmeters typically employ a relatively constant-current source, the resistance determined will be at a preset current level (typically, a few mill amperes).

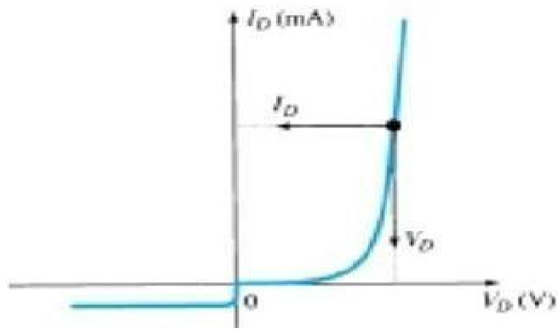


Fig 1.12 Determining the dc resistance of a diode at a particular operating point.

AC or Dynamic Resistance

It is obvious from Eq. 1.3 that the dc resistance of a diode is independent of the shape of the characteristic in the region surrounding the point of interest. If a sinusoidal rather than dc input is applied, the situation will change completely. The varying input will move the instantaneous operating point up and down a region of the characteristics and thus defines a specific change in current and voltage as shown in Fig. 1.13. With no applied varying signal, the point of operation would be the Q- point appearing on Fig. 1.13 determined by the applied dc levels. The designation Q-point is derived from the word quiescent, which means "still or unvarying." A straight-line drawn tangent to the curve through the Q-point as shown in Fig. 1.13 will define a particular change in voltage and current that can be used to determine the ac or dynamic resistance for this region of the diode characteristics. In equation form,

$$r_d = \frac{\Delta V_d}{\Delta I_d}$$

Where Δ Signifies a finite change in the quantity

Fig 1.13: Determining the ac resistance of a diode at a particular operating point.

1.4 DIODE EQUIVALENT CIRCUITS

An equivalent circuit is a combination of elements properly chosen to best represent the actual terminal characteristics of a device, system, or such in a particular operating region. In other words, once the equivalent circuit is defined, the device symbol can be removed from a schematic and the equivalent circuit inserted in its place without severely affecting the actual behavior of the system. The result is often a network that can be solved using traditional circuit analysis techniques.

Piecewise-Linear Equivalent Circuit

One technique for obtaining an equivalent circuit for a diode is to approximate the characteristics of the device by straight-line segments, as shown in Fig. 1.31. The resulting equivalent circuit is naturally called the piecewise-linear equivalent circuit. It should be obvious from Fig. 1.31 that the straight-line segments do not result in an exact duplication of the actual characteristics, especially in the knee region. However, the resulting segments are sufficiently close to the actual curve to establish an equivalent circuit that will provide an excellent first approximation to the actual behaviour of the device. The ideal diode is included to establish that there is only one direction of conduction through the device, and a reverse-bias condition will result in the open- circuit state for the device. Since a silicon semiconductor, diode does not reach the conduction state until VD reaches 0.7 V with a forward bias (as shown in Fig. 1.14a), a battery V_T opposing the conduction direction must appear in the equivalent circuit as shown in Fig. 1.14b. The battery simply specifies that the voltage across the device must be greater than the threshold battery voltage before conduction through the device in the direction dictated by the ideal diode can be established. When conduction is established, the resistance of the diode will be the specified value of r_{av}.

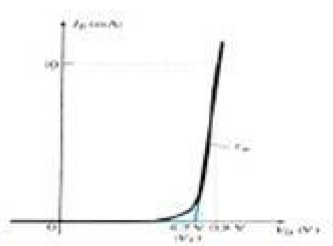


Fig: 1.14aDiode piecewise-linear model characteristics

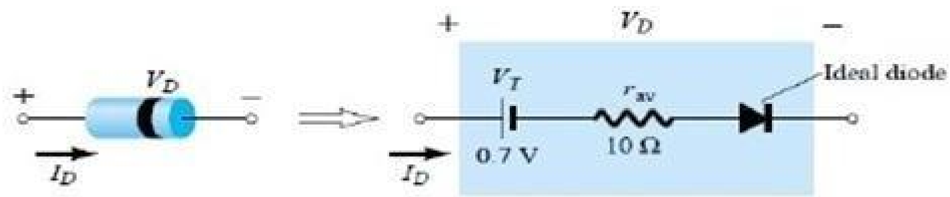


Fig: 1.14b Diode piecewise-linear model equivalent circuit

The approximate level of r_{av} can usually be determined from a specified operating point on the specification sheet. For instance, for a silicon semiconductor diode, if IF _ 10 mA (a forward conduction current for the diode) at VD _ 0.8 V, we know for silicon that a shift of 0.7 V is required before the characteristics rise.

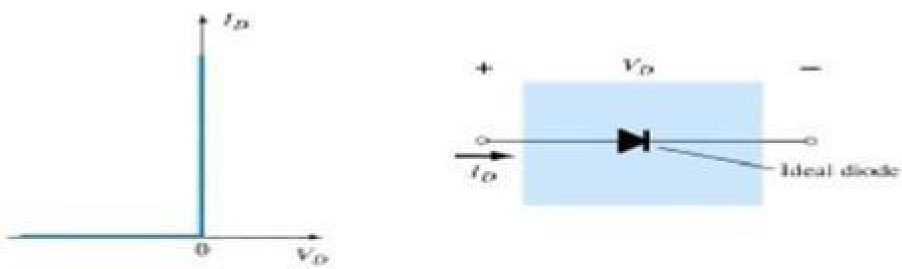


Fig 1.15 Ideal Diode and its characteristics

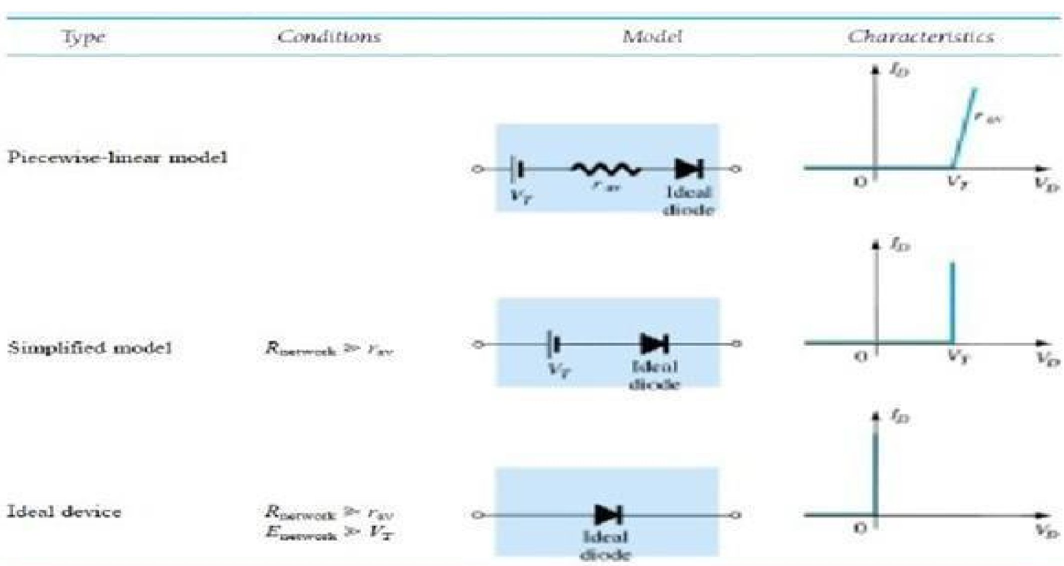


Figure 1.16: Diode equivalent circuits(models)

TRANSITION AND DIFFUSION CAPACITANCE

Electronic devices are inherently sensitive to very high frequencies. Most shunt capacitive effects that can be ignored at lower frequencies because the reactance XC=1/2πfC is very large (open- circuit equivalent). This, however, cannot be ignored at very high frequencies. XC will become sufficiently small due to the high value of f to introduce a low-reactance "shorting" path. In the p-n semiconductor diode, there are two capacitive effects to be considered. In the reverse-bias region we have the transition- or depletion region capacitance (CT), while in the forward-bias region we have the diffusion (CD) or storage capacitance. Recall that the basic equation for the capacitance of a parallel- plate capacitor is defined by C=€A/d, where € is the permittivity of the dielectric (insulator) between the plates of area A separated by a distance d. In the reverse-, bias region there is a depletion region (free of carriers) that behaves essentially like an insulator between the layers of opposite charge. Since the depletion width (d) will increase with increased reverse-bias potential, the resulting transition capacitance will decrease. The fact that the capacitance is dependent on the applied reverse-bias potential has application in a number of electronic systems. Although the effect described above will also be present in the forward-bias region, it is over shadowed by a capacitance effect directly dependent on the rate at which charge is injected into the regions just outside the depletion region. The capacitive effects described above are represented by a capacitor in parallel with the ideal diode, as shown in Fig. 1.38. For low- or mid-frequency applications (except in the power area), however, the capacitor is normally not included in the diode symbol.

Fig 1.17: Including the effect of the transition or diffusion capacitance on the semiconductor diode

Diode capacitances: The diode exhibits two types of capacitances transition capacitance and diffusion capacitance.

- Transition capacitance: The capacitance which appears between positive ion layer in n-region and negative ion layer in p-region.
- Diffusion capacitance: This capacitance originates due to diffusion of charge carriers in the opposite regions.

The transition capacitance is very small as compared to the diffusion capacitance. In reverse bias transition, the capacitance is the dominant and is given by:

$$C_T = \epsilon\, A/W$$

where C_T - transition capacitance

A - diode cross sectional area

W - depletion region width

In forward bias, the diffusion capacitance is the dominant and is given by:

$$C_D = dQ/dV = \tau * dI/dV = \tau * g = \tau/r \text{ (general)}$$

where C_D - diffusion capacitance

dQ - change in charge stored in depletion region V

- change in applied voltage

τ- time interval for change in voltage

g - diode conductance

r - diode resistance

The diffusion capacitance at low frequencies is given by the formula:

$$C_D = \tau * g/2 \text{ (low frequency)}$$

The diffusion capacitance at high frequencies is inversely proportional to the frequency and is given by the formula:

$$C_D = g(\tau/2\omega)^{1/2}$$

Note: The variation of diffusion capacitance with applied voltage is used in the design of varactor.

1.5 BREAK DOWN MECHANISMS

When an ordinary **P-N junction diode** is reverse biased, normally only very small reverse saturation current flows. This current is due to movement of minority carriers. It is almost independent of the voltage applied. However, if the reverse bias is increased, a point is reached when the junction breaks down and the reverse current increases abruptly. This current could be large enough to destroy the junction. If the reverse current is limited by means of a suitable series resistor, the power dissipation at the junction will not be excessive, and the device may be operated continuously in its breakdown region to its normal (reverse saturation) level. It is found that for a suitably designed diode, the breakdown voltage is very stable over a wide range of reverse currents. This quality gives the breakdown diode many useful applications as a voltage reference source.

The critical value of the voltage, at which the breakdown of a P-N junction diode occurs, is called the *breakdown voltage.* The breakdown voltage depends on the width of the depletion region, which, in turn, depends on the doping level. The junction offers almost zero resistance at the breakdown point.

There are two mechanisms by which breakdown can occur at a reverse biased P-N junction:

1. ***avalanche breakdown* and**
2. ***Zener breakdown.***

Avalanche breakdown

The minority carriers, under reverse biased conditions, flowing through the junction acquire a kinetic energy which increases with the increase in reverse voltage. At a sufficiently high reverse voltage (say 5 V or more), the kinetic energy of minority carriers becomes so large that they knock out electrons from the covalent bonds of the semiconductor material. As a result of collision, the liberated electrons in turn liberate more electrons and the current becomes very large leading to the breakdown of the crystal structure itself. This phenomenon is called the avalanche breakdown. The breakdown region is the knee of the characteristic curve. Now the current is not controlled by the junction voltage but rather by the external circuit.

Zener breakdown

Under a very high reverse voltage, the depletion region expands and the potential barrier increases leading to a very high electric field across the junction. The electric field will break some of the covalent bonds of the semiconductor atoms leading to a large number of free minority carriers, which suddenly increase the reverse current. This is called the Zener effect. The breakdown occurs at a particular and constant value of reverse voltage called the breakdown voltage, it is found that Zener breakdown occurs at electric field intensity of about 3×10^7 V/m.

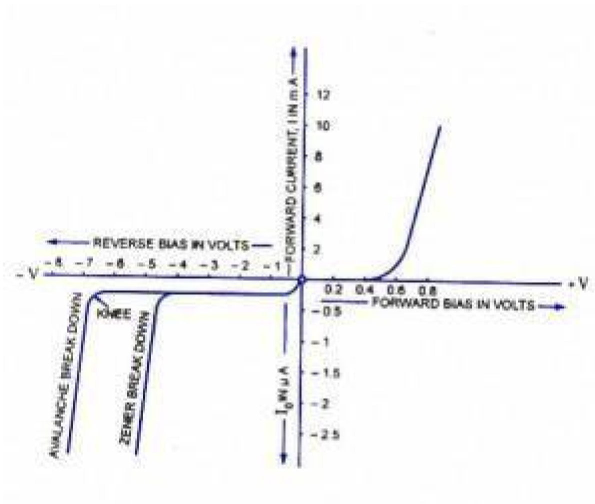


Figure 1.18: Diode characteristics with breakdown

Either of the two (Zener breakdown or avalanche breakdown) may occur independently, or both of these may occur simultaneously. Diode junctions that breakdown below 5 V are caused by Zener effect. Junctions that experience breakdown above 5 V are caused by avalanche effect. Junctions that breakdown around 5 V are usually caused by combination of two effects. The Zener breakdown occurs in heavily doped junctions (P-type semiconductor moderately doped and N-type heavily doped), which produce narrow depletion layers. The avalanche breakdown occurs in lightly doped junctions, which produce wide depletion layers. With the increase in junction temperature Zener breakdown voltage is reduced while the avalanche breakdown voltage increases. The Zener diodes have a negative temperature coefficient while avalanche diodes have a positive temperature coefficient. Diodes that have breakdown voltages around 5 V have zero temperature coefficient. The breakdown phenomenon is reversible and harmless so long as the safe operating temperature is maintained.

1.6 ZENER DIODES

The **Zener diode** is like a general-purpose signal diode consisting of a silicon PN junction. When biased in the forward direction it behaves just like a normal signal diode passing the rated current, but as soon as a reverse voltage applied across the zener diode exceeds the rated voltage of the device, the diodes breakdown voltage V_B is reached at which point a process called *Avalanche Breakdown* occurs in the semiconductor depletion layer and a current starts to flow through the diode to limit this increase in voltage.

The current now flowing through the zener diode increases dramatically to the maximum circuit value (which is usually limited by a series resistor) and once achieved this reverse saturation current remains fairly constant over a wide range of applied voltages. This breakdown voltage point, V_B is called the "zener voltage" for zener diodes and can range from less than one volt to hundreds of volts.

The point at which the zener voltage triggers the current to flow through the diode can be very accurately controlled (to less than 1% tolerance) in the doping stage of the diodes semiconductor construction giving the diode a specific *zener breakdown voltage*, (Vz) for example, 4.3V or 7.5V. This zener breakdown voltage on the I-V curve is almost a vertical straight line.

Zener Diode I-V Characteristics

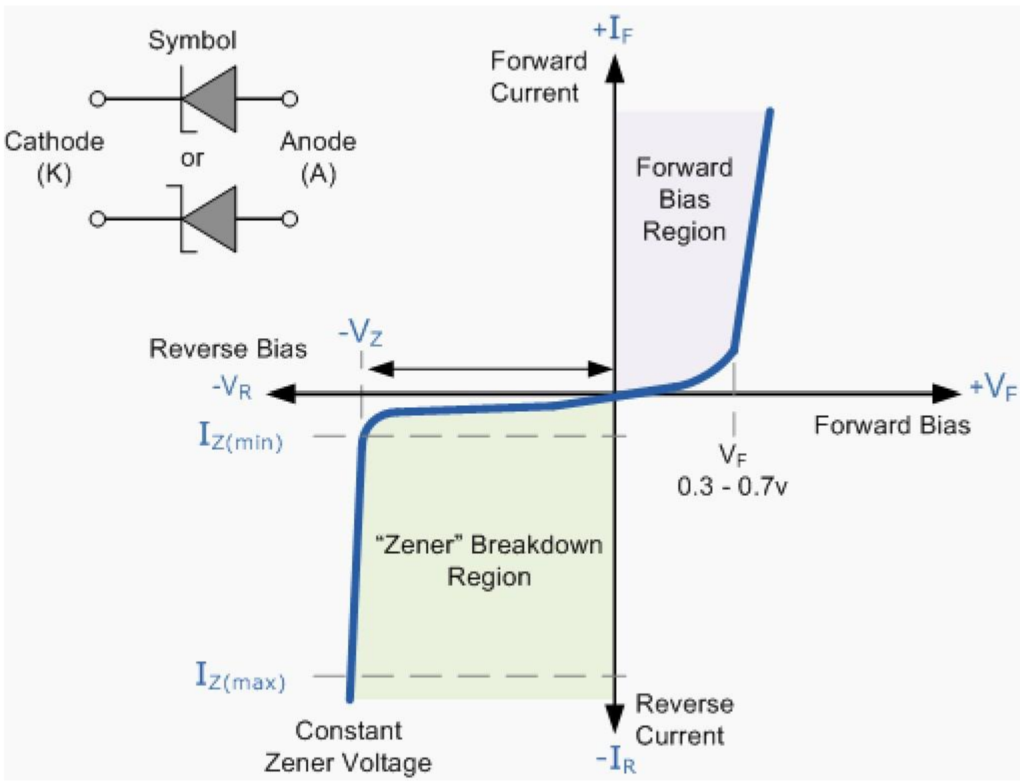


Fig 1.19 : Zener diode characteristics

The **Zener Diode** is used in its "reverse bias" or reverse breakdown mode, i.e. the diodes anode connects to the negative supply. From the I-V characteristics curve above, we can see that the zener diode has a region in its reverse bias characteristics of almost a constant negative voltage regardless of the value of the current flowing through the diode and remains nearly constant even with large changes in current as long as the zener diodes current remains between the breakdown current $I_{Z(min)}$ and the maximum current rating $I_{Z(max)}$.

This ability to control itself can be used to great effect to regulate or stabilize a voltage source against supply or load variations. The fact that the voltage across the diode in the breakdown region is almost constant turns out to be an important application of the zener diode as a voltage regulator. The function of a regulator is to provide a constant output voltage to a load connected in parallel with it in spite of the ripples in the supply voltage or the variation in the load current and the zener diode will continue to regulate the voltage until the diodes current falls below the minimum $I_{Z(min)}$ value in the reverse breakdown region.

RECTIFIERS &FILTERS

2.0 INTRODUCTION

For the operation of most of the electronics devices and circuits, a d.c. source is required. So it is advantageous to convert domestic a.c. supply into d.c.voltages. The process of converting a.c. voltage into d.c. voltage is called as rectification. This is achieved with i) Step-down Transformer, ii) Rectifier, iii) Filter and iv) Voltage regulator circuits.

These elements constitute d.c. regulated power supply shown in the fig 1 below.

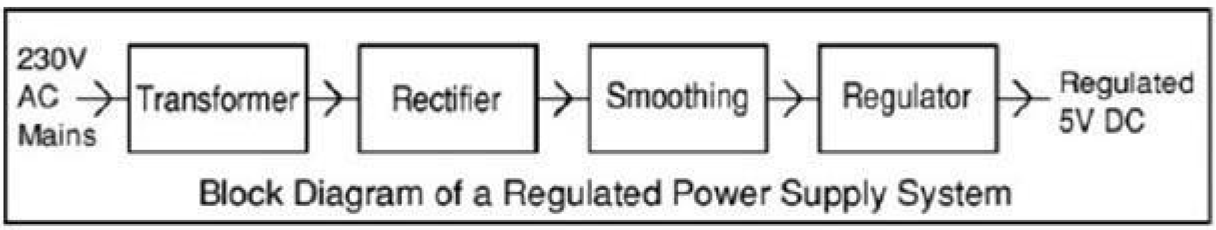


Fig 2.1: Block Diagram of regulated D.C Power Supply

- ✓ Transformer – steps down 230V AC mains to low voltage AC.
- ✓ Rectifier – converts AC to DC, but the DC output is varying.
- ✓ Smoothing – smooth the DC from varying greatly to a small ripple.
- ✓ Regulator – eliminates ripple by setting DC output to a fixed voltage.

The block diagram of a regulated D.C. power supply consists of step-down transformer, rectifier, filter, voltage regulator and load. An ideal regulated power supply is an electronics circuit designed to provide a predetermined d.c. voltage Vo which is independent of the load current and variations in the input voltage ad temperature. If the output of a regulator circuit is a AC voltage then it is termed as voltage stabilizer, whereas if the output is a DC voltage then it is termed as voltage regulator.

2.1 RECTIFIER

Any electrical device which offers a low resistance to the current in one direction but a high resistance to the current in the opposite direction is called rectifier. Such a device is capable of converting a sinusoidal input waveform, whose average value is zero, into a unidirectional Waveform, with a non- zero average component. A rectifier is a device, which converts a.c. voltage (bi-directional) to pulsating d.c. voltage (Unidirectional).

Characteristics of a Rectifier Circuit:

Any electrical device which offers a low resistance to the current in one direction but a high resistance to the current in the opposite direction is called rectifier. Such a device is capable of converting a sinusoidal input waveform, whose average value is zero, into a unidirectional waveform, with a non- zero average component.

A rectifier is a device, which converts a.c. voltage (bi-directional) to pulsating d.c..Load currents: They are two types of output current. They are average or d.c. current and RMS currents.

Average or DC current: The average current of a periodic function is defined as the area of one cycle of the curve divided by the base.

It is expressed mathematically as

i) Average value/dc value/mean value= $\dfrac{\textit{Area over one period}}{\textit{Total time period}}$

$$V_{dc} = \frac{1}{T}\int_{0}^{T} V d(wt)$$

ii) Effective (or) R.M.S current:

The effective (or) R.M.S. current squared ofa periodic function of time is given by the area of one cycle of the curve, which represents the square of the function divided by the base.

$$V_{rms} = \sqrt{\frac{1}{T}\int_{0}^{t} V^{2} d(wt)}$$

iii) Peak factor:

It is the ratio of peak value to Rms value

$$\text{Peak factor} = \frac{peakvalue}{rmsvalue}$$

iv) Form factor:

It is the ratio of Rms value to average value

$$\text{Form factor} = \frac{Rms value}{average value}$$

v) Ripple Factor (γ) :

It is defined as ration of R.M.S. value of a.c. component to the d.c. component in the output is known as "Ripple Factor".

$$\gamma = \frac{V_{ac}}{V_{dc}}$$

$$V_{ac} = \sqrt{V_{rms}^2 - V_{dc}^2}$$

vi) Efficiency (η):

It is the ratio of d.c output power to the a.c. input power. It signifies, how efficiently the rectifier circuit converts a.c. power into d.c. power.

$$\eta = \frac{o/p\ power}{i/p\ power}$$

vii) Peak Inverse Voltage (PIV):

It is defined as the maximum reverse voltage that a diode can withstand without destroying the junction.

viii) Transformer Utilization Factor (UTF):

The d.c. power to be delivered to the load in a rectifier circuit decides the rating of the Transformer used in the circuit. So, transformer utilization factor is defined as

$$TUF = \frac{P_{dc}}{p_{ac(rated)}}$$

ix) % Regulation:

The variation of the d.c. output voltage as a function of d.c. load current is called regulation. The percentage regulation is defined as

$$\%Regulation = \frac{V_{NL} - V_{FL}}{V_{FL}} * 100$$

For an ideal power supply, % Regulation is zero.

2.2 CLASSIFICATION OF RECTIFIERS

Using one or more diodes in the circuit, following rectifier circuits can be designed.

1) Half - Wave Rectifier
2) Full – Wave Rectifier
3) Bridge Rectifier

2.2.1 HALF-WAVE RECTIFIER:

A Half – wave rectifier as shown in **fig 1.2** is one, which converts a.c. voltage into a pulsating voltage using only one half cycle of the applied a.c. voltage.

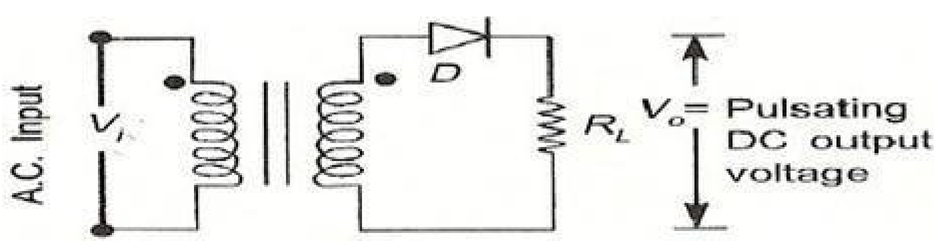


Fig 1.2: Basic structure of Half-Wave Rectifier

The a.c. voltage is applied to the rectifier circuit using step-down transformer-rectifying element i.e., p- n junction diode and the source of a.c. voltage, all connected is series. The a.c. voltage is applied to the rectifier circuit using step-down transformer

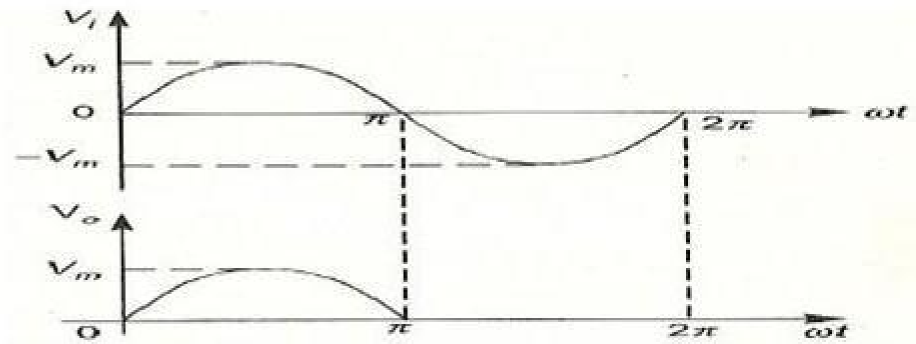


fig 3 Input and output waveforms of a Half wave rectifier

$V=V_m \sin(wt)$

The input to the rectifier circuit, Where V_m is the peak value of secondary a.c. voltage.

Operation:

For the positive half-cycle of input a.c. voltage, the diode D is forward biased and hence it conducts. Now a current flows in the circuit and there is a voltage drop across RL. The waveform of the diode current (or) load current is shown in **fig 3**.

For the negative half-cycle of input, the diode D is reverse biased and hence it does not Conduct. Now no current flows in the circuit i.e., i=0 and Vo=0. Thus for the negative half- cycle no power is delivered to the load.

Analysis:

In the analysis of a HWR, the following parameters are to be analyzed.

1. DC output current
2. DC Output voltage
3. R.M.S. Current
4. R.M.S. voltage
5. Rectifier Efficiency (η)
6. Ripple factor (γ)

7. Peak Factor
8. % Regulation
9. Transformer Utilization Factor (TUF)
10. form factor
11. o/p frequency

vi) Transformer Utilization Factor (TUF):

The d.c. power to be delivered to the load in a rectifier circuit decides the rating of the transformer used in the circuit. Therefore, transformer utilization factor is defined as

$$TUF = \frac{p_{dc}}{P_{ac(rated)}}$$

TUF =0.286.

The value of TUF is low which shows that in half-wave circuit, the transformer is not fully utilized. If the transformer rating is 1 KVA (1000VA) then the half-wave rectifier can deliver

1000 X 0.287 = 287 watts to resistance load.

vii) Peak Inverse Voltage (PIV):

It is defined as the maximum reverse voltage that a diode can withstand without destroying the junction. The peak inverse voltage across a diode is the peak of the negative half- cycle. For half-wave rectifier, PIV is Vm.

DISADVANTAGES OF HALF-WAVE RECTIFIER:

1. The ripple factor is high.
2. The efficiency is low.
3. The Transformer Utilization factor is low.

Because of all these disadvantages, the half-wave rectifier circuit is normally not used as a power rectifier circuit.

2.2.2) FULL WAVE RECTIFIER:

A full-wave rectifier converts an ac voltage into a pulsating dc voltage using both half cycles of the applied ac voltage. In order to rectify both the half cycles of ac input, two diodes are used in this circuit. The diodes feed a common load RL with the help of a center-tap transformer. A center-tap transformer is the one, which produces two sinusoidal waveforms of same magnitude and frequency but out of phase with respect to the ground in the secondary winding of the transformer. The full wave rectifier is shown in the **fig 4** below

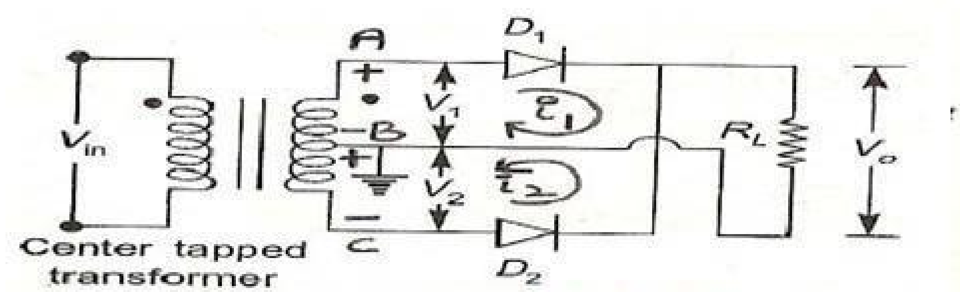


fig 4 Full-Wave Rectifier

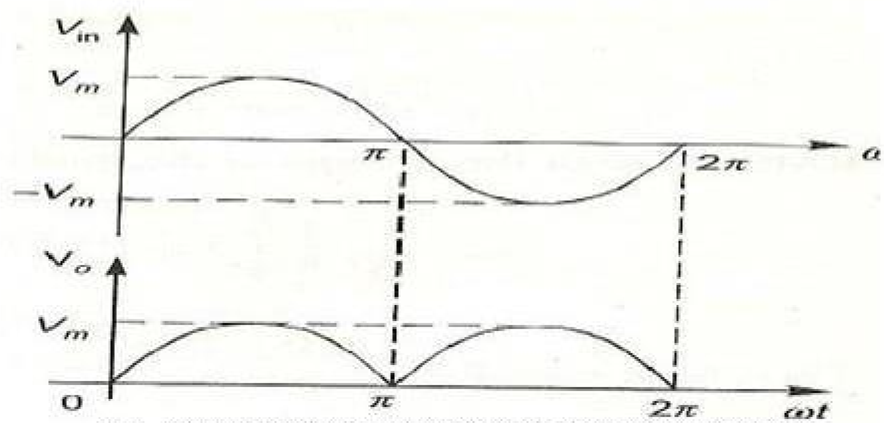


Fig. 5 input and output waveforms of Fullwave rectifier

Fig. 5 shows the input and output wave forms of the ckt.

During positive half of the input signal, anode of diode D1 becomes positive and at the same time the anode of diode D2 becomes negative. Hence D1 conducts and D2 does not conduct. The load current flows through D1 and the voltage drop across RL will be equal to the input voltage.During the negative half cycle of the input, the anode of D1 becomes negative and the anode of D2 becomes positive. Hence, D1 does not conduct and D2 conducts. The load current flows through D2 and the voltage drop across RL will be equal to the input voltage. It is noted that the load current flows in

the both the half cycles of ac voltage and in the same direction through the load resistance.

i) **AVERAGEVOLTAGE**

$$V_{dc} = I_{dc}.R_L. = \frac{2\,I_m}{\pi}.R_L \quad \text{We know } I_m = \frac{V_m}{R_S + R_f + R_L}$$

$$\therefore V_{dc} = \frac{2.V_m R_L}{\pi(R_S + R_f + R_L}$$

$$If\,(R_s + R_f) << R_L$$

$$V_{dc} = \frac{2V_m}{\pi} = 0.637V_{m.}$$

ii) **_AVERAGE CURRENT**

$$I_{dc} \quad \frac{1}{2\pi}\int_0^{2\pi} i\,d\theta = \frac{1}{2\pi}\int_0^{2\pi} I_m \sin\theta\,d\theta$$

$$= \frac{I_m}{2\pi}\left[\int_0^{\pi} \sin\theta\,d\theta - \int_{\pi}^{2\pi} \sin\theta\,d\theta\right]$$

$$= \frac{I_m}{2\pi}[(-2)(-2)]$$

$$= \frac{I_m}{2\pi}.4 = \frac{2\,I_m}{\pi} = 0.637\,I_m.$$

$$\boxed{I_{dc} = 0.637\,I_{m.}}$$

iii) $\therefore$ I_{DC} FWR $= 2\ I_{DC}$ HWR.

iv) RMS VOLTAGE:

$$V_{rms} = \sqrt{\frac{1}{T}\int_{0}^{T} V^{2} d(wt)}$$

$$V_{rms} = \sqrt{\frac{1}{2\Pi}\int_{0}^{2\Pi} (V_m sim(wt))^{2} d(wt)}$$

$$V_{rms} = \frac{V_m}{\sqrt{2}}$$

IV) RMS CURRENT

$$I_{rms} = \frac{2I_m}{\Pi}$$

V) PEAK FACTOR

factor = Peak $\frac{peakvalue}{rmsvalue}$

vi) Ripple Factor:

$$\gamma = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1}$$

for FWR,

$$I_{rms} = \frac{I_m}{\sqrt{2}} \quad \& \quad I_{DC} = \frac{2I_m}{\pi}$$

$$\therefore \gamma_{FWR} = \sqrt{\left(\frac{I_m}{\sqrt{2}} / \frac{2I_m}{\pi}\right)^2 - 1}$$

$$= \sqrt{\left(\frac{\pi}{2\sqrt{2}}\right)^2 - 1}$$

$$= \sqrt{\left(\frac{3.1416}{2\times1.414}\right)^2 - 1} = 0.483$$

vii) Efficiency ( ):

$$\square \; \square \frac{o / ppower}{i / ppower} *100$$

viii) Transformer Utilization Factor (TUF):

The d.c. power to be delivered to the load in a rectifier circuit decides the rating of the transformer used in the circuit. So, transformer utilization factor is defined as

a) TUF (Secondary) = $\frac{P_{dc} \text{ delivered to load}}{AC \text{ power rating of transformer secondary}}$

b) Since both the windings are used TUF $_{FWR}$ = 2 TUF $_{HWR}$
= 2 x 0.287 = 0.574

c) TUF primary = Rated efficiency = $\frac{P_{dc}}{P_{ac}} \times 100 = 81.2\%$

d) Average = $\frac{0.812 + 0.574}{2} = 0.693$

Advantages

1) Ripple factor = 0.482 (against 1.21 for HWR)

2) Rectification efficiency is 0.812 (against 0.405 for HWR)

3) Better TUF (secondary) is 0.574 (0.287 for HWR)

4) No core saturation problem

Disadvantages:

1) Requires center tapped transformer.

2.2.3) BRIDGE RECTIFIER.

Another type of circuit that produces the same output waveform as the full wave rectifier circuit above, is that of the **Full Wave Bridge Rectifier**. This type of single phase rectifier uses four individual rectifying diodes connected in a closed loop "bridge" configuration to produce the desired output. The main advantage of this bridge circuit is that it does not require a special centre tapped transformer, thereby reducing its size and cost. The single secondary winding is connected to one side of the diode bridge network and the load to the other side as shown below.

The Diode Bridge Rectifier

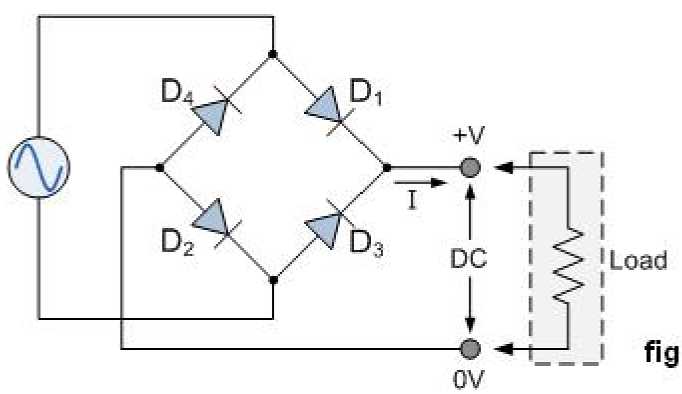


fig 6

The four diodes labelled D_1 to D_4 are arranged in "series pairs" with only two diodes conducting current during each half cycle. During the positive half cycle of the supply, diodes D1 and D2 conduct in series while diodes D3 and D4 are reverse biased and the current flows through the load as shown below (fig 7).

The Positive Half-cycle

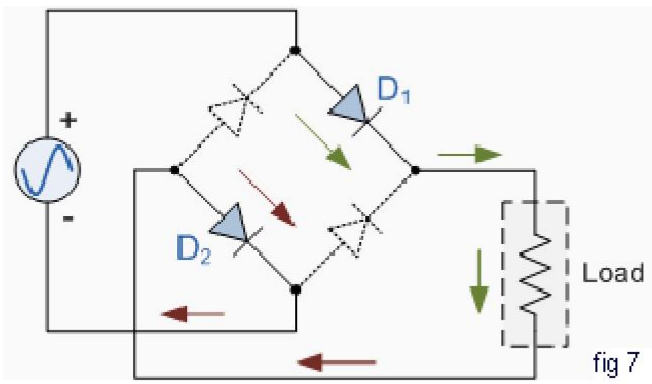


fig 7

The Negative Half-cycle

During the negative half cycle of the supply, diodes D3 and D4 conduct in series (fig 8), but diodes D1 and D2 switch "OFF" as they are now reverse biased. The current flowing through the load is the same direction as before.

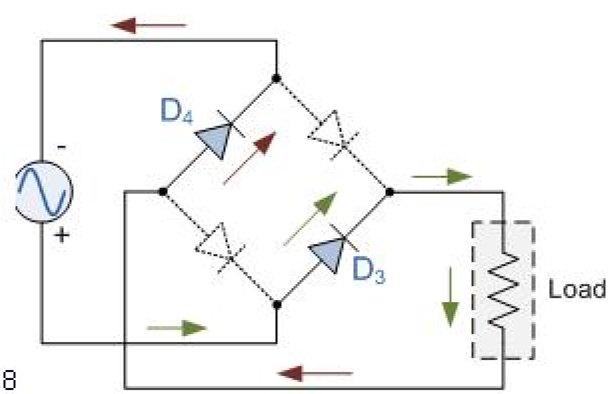


fig 8

As the current flowing through the load is unidirectional, so the voltage developed across the load is also unidirectional the same as for the previous two diode full-wave rectifier, therefore the average DC voltage across the load is $0.637V_{max}$. However in reality, during each half cycle the current flows through two diodes instead of just one so the amplitude of the output voltage is two voltage drops (2 x 0.7 = 1.4V) less than the input V_{MAX} amplitude. The ripple frequency is now twice the supply frequency (e.g. 100Hz for a 50Hz supply)

Therefore, the following expressions are same as that of full wave rectifier.

a) Average current $I_{dc} = \frac{2I_m}{\pi}$

b) RMS current $I_{rms} = \frac{I_m}{\sqrt{2}}$

c) DC output voltage (no.load) V_{DC} $\frac{2V_m}{\pi}$

d) Ripple factor γ=0.482

e) Rectification efficiency = $\eta = 0.812$

f) DC output voltage full load.

$= \quad V_{DCFL} = \frac{2V_m}{\pi} - I_{dc}(R_S + 2R_f)$; i.e., less by one diode loss.

TUF of both primary & secondary are 0.812 therefore TUF overall is 0.812 (better than FWR with 0.693)

Comparison:

Sl. No.	Parameter	HWR	FWR	BR
1	No. of diodes	1	2	4
2	PIV of diodes	V_m	2 V_m	V_m
3	Secondary voltage (rms)	V	V-0-V	V
4	DC output voltage at no load	$\frac{V_m}{\pi}$ =0.318 V_m	$\frac{2V_m}{\pi}$ =0.636 V_m	$\frac{2V_m}{\pi}$ =0.636 V_m
5	Ripple factor γ	1.21	0.482	0.482
6	Ripple frequency	f	2f	2f
7	Rectification efficiency η	0.406	0.812	0.812
8	TUF	0.287	0.693	0.812

2.3 FILTERS

The output of a rectifier contains dc component as well as ac component. Filters are used to minimize the

undesirable ac i.e., ripple leaving only the dc component to appear at the output
Some important filters are:

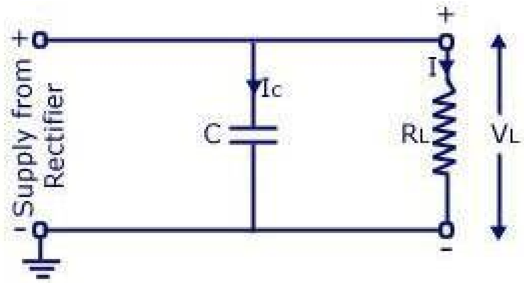
1. Inductor filter
2. Capacitor filter
3. LC or L section filter
4. CLC or Π-type filter

2.3.1 CAPACITOR FILTER

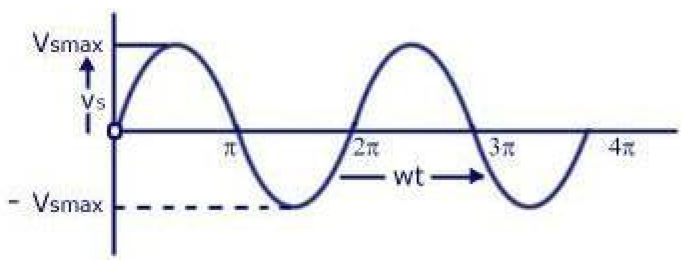
This is the most simple form of the **filter circuit** and in this arrangement a high value capacitor C is placed directly across the output terminals, as shown in figure. During the conduction period it gets charged and stores up energy to it during non-conduction period. Through this process, the time duration during which Ft is to be noted here that the capacitor C gets charged to the peak because there is no resistance (except the negligible forward resistance of diode) in the charging path. But the discharging time is quite large (roughly 100 times more than the charging time depending upon the value of RL) because it discharges through load resistance RL.

The function of the capacitor filter may be viewed in terms of impedances. The large value capacitor C offers a low impedance shunt path to the ac components or ripples but offers high impedance to the dc component. Thus ripples get bypassed through capacitor C and only dc component flows through the load resistance RL

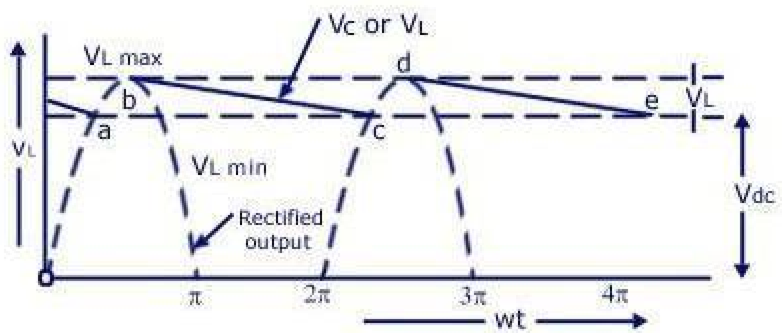
Capacitor filter is very popular because of its low cost, small size, light weight and good characteristics.



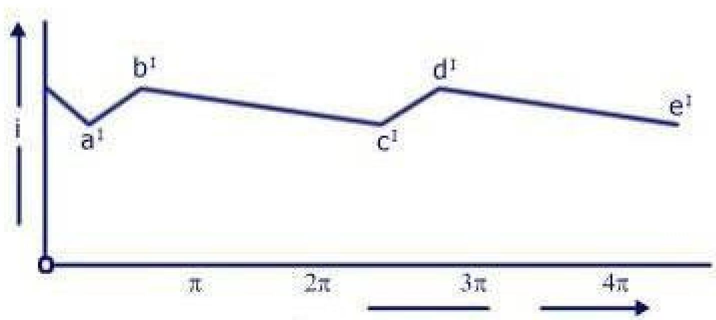
Circuit Diagram



Input voltage Waveform to Rectifier



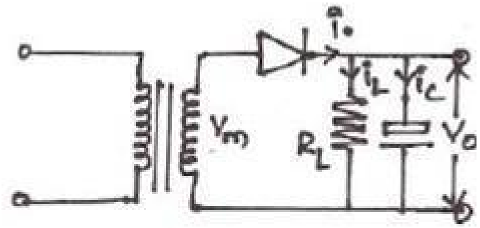
Rectified and filtered Output Voltage Waveform



Load Current Waveform

Half-wave Rectifier With Shunt Capacitor Filter

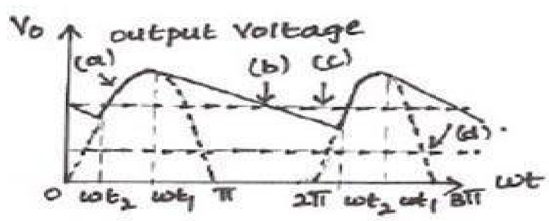
CAPACITOR FILTER WITH HWR



Cut In angle – wt_2

Cut out angle = Wt_1

$Wt_1 = \pi - \tan^{-1} WCR_L$



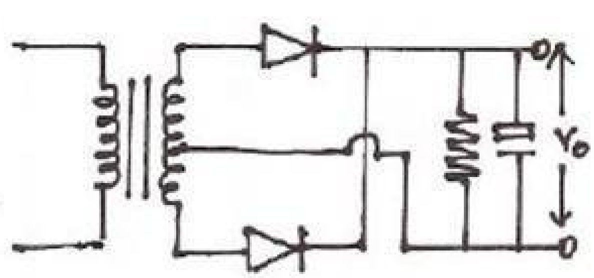
(a) Capacitor charging through diode ($Wt_2 - Wt_1$)

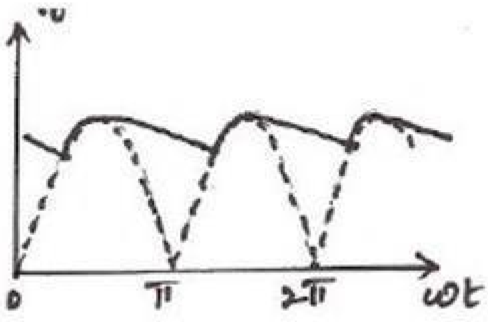
(b) Capacitor discharging through R_L (Wt_1 to Wt_2)

(c) Average (DC) voltage with fitter

(d) Average (DC) voltage without fitter.

CAPACITOR FILTER WITH FWR





$$\text{Ripple factor } r = \frac{1}{4\sqrt{3} fCR_L}$$

Ripple freq $_{FWR}$ = 2 ripple freq $_{HWR}$.