**CLASSIFICATION OF MATERIALS**

**Matters are classified into three broad categories according to their electrical properties as-conductors, insulators and semiconductors.**

**Conductors** are materials that have very low values of resistivity, usually in the micro-ohms per metre. This low value allows them to easily pass an electrical current due to there being plenty of free electrons floating about within their basic atom structure. But these electrons will only flow through a conductor if there is something to spur their movement, and that something is an electrical voltage.

When a positive voltage potential is applied to the material these “free electrons” leave their parent atom and travel together through the material forming an electron drift, more commonly known as a current. How “freely” these electrons can move through a conductor depends on how easily they can break free from their constituent atoms when a voltage is applied. Then the amount of electrons that flow depends on the amount of resistivity the conductor has.

Examples of good conductors are generally metals such as Copper, Aluminium, Silver or non metals such as Carbon because these materials have very few electrons in their outer “Valence Shell” or ring, resulting in them being easily knocked out of the atom’s orbit.

This allows them to flow freely through the material until they join up with other atoms, producing a “Domino Effect” through the material thereby creating an electrical current. Copper and Aluminium is the main conductor used in electrical cables as shown.

Generally speaking, most metals are good conductors of electricity, as they have very small resistance values, usually in the region of micro-ohms per metre, (μΩ.m).

While metals such as copper and aluminium are very good conducts of electricity, they still have some resistance to the flow of electrons and consequently do not conduct perfectly.

The energy which is lost in the process of passing an electrical current, appears in the form of heat which is why conductors and especially resistors become hot as the resistivity of conductors increases with ambient temperature.

**Insulators**

**Insulators** on the other hand are the exact opposite of conductors. They are made of materials, generally non-metals, that have very few or no “free electrons” floating about within their basic atom structure because the electrons in the outer valence shell are strongly attracted by the positively charged inner nucleus.

In other words, the electrons are stuck to the parent atom and can not move around freely so if a potential voltage is applied to the material no current will flow as there are no “free electrons” available to move and which gives these materials their insulating properties.

Insulators also have very high resistances, millions of ohms per metre, and are generally not affected by normal temperature changes (although at very high temperatures wood becomes charcoal and changes from an insulator to a conductor). Examples of good insulators are marble, fused quartz, PVC plastics, rubber etc.

Insulators play a very important role within electrical and electronic circuits, because without them electrical circuits would short together and not work. For example, insulators made of glass or porcelain are used for insulating and supporting overhead transmission cables while epoxy-glass resin materials are used to make printed circuit boards, PCB’s etc. while PVC is used to insulate electrical cables as shown.

**Semiconductor Basics**

**Semiconductors** materials such as silicon (Si), germanium (Ge) and gallium arsenide (GaAs), have electrical properties somewhere in the middle, between those of a “conductor” and an “insulator”. They are not good conductors nor good insulators (hence their name “semi”-conductors). They have very few “free electrons” because their atoms are closely grouped together in a crystalline pattern called a “crystal lattice” but electrons are still able to flow, but only under special conditions.

The ability of semiconductors to conduct electricity can be greatly improved by replacing or adding certain donor or acceptor atoms to this crystalline structure thereby, producing more free electrons than holes or vice versa. That is by adding a small percentage of another element to the base material, either silicon or germanium.

On their own Silicon and Germanium are classed as intrinsic semiconductors, that is they are chemically pure, containing nothing but semi-conductive material. But by controlling the amount of impurities added to this intrinsic semiconductor material it is possible to control its conductivity. Various impurities called donors or acceptors can be added to this intrinsic material to produce free electrons or holes respectively.

This process of adding donor or acceptor atoms to semiconductor atoms (the order of 1 impurity atom per 10 million (or more) atoms of the semiconductor) is called **Doping**. The as the doped silicon is no longer pure, these donor and acceptor atoms are collectively referred to as “impurities”, and by doping these silicon material with a sufficient number of impurities, we can turn it into an N-type or P-type semi-conductor material.

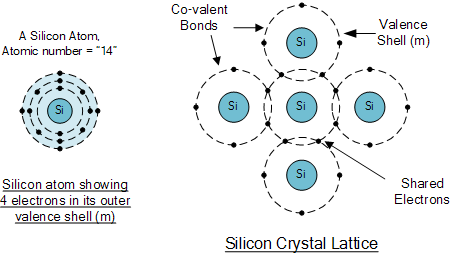
The most commonly used semiconductor basics material by far is **silicon**. Silicon has four valence electrons in its outermost shell which it shares with its neighbouring silicon atoms to form full orbital’s of eight electrons. The structure of the bond between the two silicon atoms is such that each atom shares one electron with its neighbour making the bond very stable.

As there are very few free electrons available to move around the silicon crystal, crystals of pure silicon (or germanium) are therefore good insulators, or at the very least very high value resistors.

Silicon atoms are arranged in a definite symmetrical pattern making them a crystalline solid structure. A crystal of pure silica (silicon dioxide or glass) is generally said to be an intrinsic crystal (it has no impurities) and therefore has no free electrons.

But simply connecting a silicon crystal to a battery supply is not enough to extract an electric current from it. To do that we need to create a “positive” and a “negative” pole within the silicon allowing electrons and therefore electric current to flow out of the silicon. These poles are created by doping the silicon with certain impurities.

**A Silicon Atom Structure**

****

The diagram above shows the structure and lattice of a ‘normal’ pure crystal of Silicon.

## N-type Semiconductor Basics

In order for our silicon crystal to conduct electricity, we need to introduce an impurity atom such as Arsenic, Antimony or Phosphorus into the crystalline structure making it extrinsic (impurities are added). These atoms have five outer electrons in their outermost orbital to share with neighbouring atoms and are commonly called “Pentavalent” impurities.

This allows four out of the five orbital electrons to bond with its neighbouring silicon atoms leaving one “free electron” to become mobile when an electrical voltage is applied (electron flow). As each impurity atom “donates” one electron, pentavalent atoms are generally known as “donors”.

**Antimony** (symbol Sb) as well as **Phosphorus** (symbol P), are frequently used as a pentavalent additive to silicon. Antimony has 51 electrons arranged in five shells around its nucleus with the outermost orbital having five electrons. The resulting semiconductor basics material has an excess of current-carrying electrons, each with a negative charge, and is therefore referred to as an **N-type** material with the electrons called “Majority Carriers” while the resulting holes are called “Minority Carriers”.

When stimulated by an external power source, the electrons freed from the silicon atoms by this stimulation are quickly replaced by the free electrons available from the doped Antimony atoms. But this action still leaves an extra electron (the freed electron) floating around the doped crystal making it negatively charged.

Then a semiconductor material is classed as N-type when its donor density is greater than its acceptor density, in other words, it has more electrons than holes thereby creating a negative pole as shown.

### Antimony Atom and Doping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode2.gif

The diagram above shows the structure and lattice of the donor impurity atom Antimony.

## P-Type Semiconductor Basics

If we go the other way, and introduce a “Trivalent” (3-electron) impurity into the crystalline structure, such as Aluminium, Boron or Indium, which have only three valence electrons available in their outermost orbital, the fourth closed bond cannot be formed. Therefore, a complete connection is not possible, giving the semiconductor material an abundance of positively charged carriers known as holes in the structure of the crystal where electrons are effectively missing.

As there is now a hole in the silicon crystal, a neighbouring electron is attracted to it and will try to move into the hole to fill it. However, the electron filling the hole leaves another hole behind it as it moves. This in turn attracts another electron which in turn creates another hole behind it, and so forth giving the appearance that the holes are moving as a positive charge through the crystal structure (conventional current flow).

This movement of holes results in a shortage of electrons in the silicon turning the entire doped crystal into a positive pole. As each impurity atom generates a hole, trivalent impurities are generally known as “**Acceptors**” as they are continually “accepting” extra or free electrons.

**Boron** (symbol B) is commonly used as a trivalent additive as it has only five electrons arranged in three shells around its nucleus with the outermost orbital having only three electrons. The doping of Boron atoms causes conduction to consist mainly of positive charge carriers resulting in a **P-type** material with the positive holes being called “Majority Carriers” while the free electrons are called “Minority Carriers”.

Then a semiconductor basics material is classed as P-type when its acceptor density is greater than its donor density. Therefore, a P-type semiconductor has more holes than electrons.

### Boron Atom and Doping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode3.gif

The diagram above shows the structure and lattice of the acceptor impurity atom Boron.

## Semiconductor Basics Summary

### N-type (e.g. doped with Antimony)

These are materials which have **Pentavalent** impurity atoms (Donors) added and conduct by “electron” movement and are therefore called, **N-type Semiconductors**.

In N-type semiconductors there are:

1. The Donors are positively charged.

2. There are a large number of free electrons.

3. A small number of holes in relation to the number of free electrons.

4. Doping gives:

* + positively charged donors.
  + negatively charged free electrons.

5. Supply of energy gives:

* + negatively charged free electrons.
  + positively charged holes.

### P-type (e.g. doped with Boron)

These are materials which have **Trivalent** impurity atoms (Acceptors) added and conduct by “hole” movement and are therefore called, **P-type Semiconductors**.

In these types of materials are:

1. The Acceptors are negatively charged.

2. There are a large number of holes.

3. A small number of free electrons in relation to the number of holes.

4. Doping gives:

* + negatively charged acceptors.
  + positively charged holes.

5. Supply of energy gives:

* + positively charged holes.
  + negatively charged free electrons.

and both P and N-types as a whole, are electrically neutral on their own.

Antimony (Sb) and Boron (B) are two of the most commonly used doping agents as they are more feely available compared to other types of materials. They are also classed as “metalloids”. However, the periodic table groups together a number of other different chemical elements all with either three, or five electrons in their outermost orbital shell making them suitable as a doping material.

These other chemical elements can also be used as doping agents to a base material of either Silicon (Si) or Germanium (Ge) to produce different types of basic semiconductor materials for use in electronic semiconductor components, microprocessor and solar cell applications. These additional semiconductor materials are given below.

### Periodic Table of Semiconductors

|  |  |  |
| --- | --- | --- |
| **Elements Group 13** | **Elements Group 14** | **Elements Group 15** |
| **3-Electrons in Outer Shell (Positively Charged)** | **4-Electrons in Outer Shell (Neutrally Charged)** | **5-Electrons in Outer Shell (Negatively Charged)** |
| **(5)Boron  ( B )** | **(6)Carbon  ( C )** |  |
| **(13)Aluminium  ( Al )** | **(14)Silicon  ( Si )** | **(15)Phosphorus  ( P )** |
| **(31)Gallium  ( Ga )** | **(32)Germanium  ( Ge )** | **(33)Arsenic  ( As )** |
|  |  | **(51)Antimony  ( Sb )** |

In the next tutorial about semiconductors and diodes, we will look at joining the two semiconductor basics materials, the P-type and the N-type materials to form a PN Junction which can be used to produce diodes.

# PN Junction Theory

A PN-junction is formed when an N-type material is fused together with a P-type material creating a semiconductor diode

In the previous tutorial we saw how to make an N-type semiconductor material by doping a silicon atom with small amounts of Antimony and also how to make a P-type semiconductor material by doping another silicon atom with Boron.

This is all well and good, but these newly doped N-type and P-type semiconductor materials do very little on their own as they are electrically neutral. However, if we join (or fuse) these two semiconductor materials together they behave in a very different way merging together and producing what is generally known as a “**PN Junction**“.

When the N-type semiconductor and P-type semiconductor materials are first joined together a very large density gradient exists between both sides of the PN junction. The result is that some of the free electrons from the donor impurity atoms begin to migrate across this newly formed junction to fill up the holes in the P-type material producing negative ions.

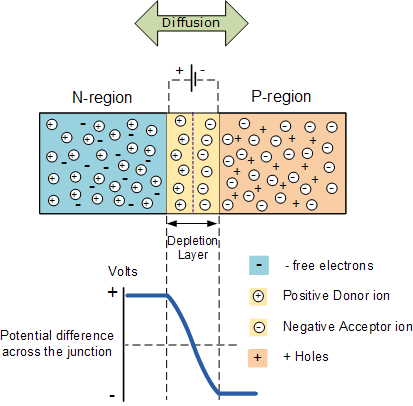
However, because the electrons have moved across the PN junction from the N-type silicon to the P-type silicon, they leave behind positively charged donor ions ( ND ) on the negative side and now the holes from the acceptor impurity migrate across the junction in the opposite direction into the region where there are large numbers of free electrons.

As a result, the charge density of the P-type along the junction is filled with negatively charged acceptor ions ( NA ), and the charge density of the N-type along the junction becomes positive. This charge transfer of electrons and holes across the PN junction is known as **diffusion**. The width of these P and N layers depends on how heavily each side is doped with acceptor density NA, and donor density ND, respectively.

This process continues back and forth until the number of electrons which have crossed the junction have a large enough electrical charge to repel or prevent any more charge carriers from crossing over the junction. Eventually a state of equilibrium (electrically neutral situation) will occur producing a “potential barrier” zone around the area of the junction as the donor atoms repel the holes and the acceptor atoms repel the electrons.

Since no free charge carriers can rest in a position where there is a potential barrier, the regions on either sides of the junction now become completely depleted of any more free carriers in comparison to the N and P type materials further away from the junction. This area around the **PN Junction** is now called the **Depletion Layer**.

**The PN junction**

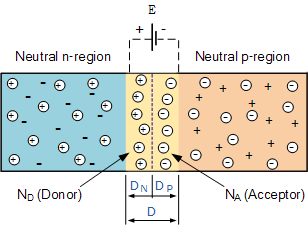
****

The total charge on each side of a PN Junction must be equal and opposite to maintain a neutral charge condition around the junction. If the depletion layer region has a distance D, it therefore must therefore penetrate into the silicon by a distance of Dp for the positive side, and a distance of Dn for the negative side giving a relationship between the two of:  Dp\*NA = Dn\*ND  in order to maintain charge neutrality also called equilibrium.

### PN Junction Distance

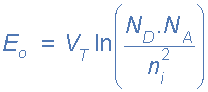
The total charge on each side of a PN Junction must be equal and opposite to maintain a neutral charge condition around the junction. If the depletion layer region has a distance D, it therefore must therefore penetrate into the silicon by a distance of Dp for the positive side, and a distance of Dn for the negative side giving a relationship between the two of:  Dp\*NA = Dn\*ND  in order to maintain charge neutrality also called equilibrium.

### PN Junction Distance

****

As the N-type material has lost electrons and the P-type has lost holes, the N-type material has become positive with respect to the P-type. Then the presences of impurity ions on both sides of the junction cause an electric field to be established across this region with the N-side at a positive voltage relative to the P-side. The problem now is that a free charge requires some extra energy to overcome the barrier that now exists for it to be able to cross the depletion region junction.

This electric field created by the diffusion process has created a “built-in potential difference” across the junction with an open-circuit (zero bias) potential of:

****

Where: Eo is the zero bias junction voltage, VT the thermal voltage of 26mV at room temperature,  ND and NA are the impurity concentrations and ni is the intrinsic concentration.

A suitable positive voltage (forward bias) applied between the two ends of the PN junction can supply the free electrons and holes with the extra energy. The external voltage required to overcome this potential barrier that now exists is very much dependent upon the type of semiconductor material used and its actual temperature.

Typically at room temperature the voltage across the depletion layer for silicon is about 0.6 – 0.7 volts and for germanium is about 0.3 – 0.35 volts. This potential barrier will always exist even if the device is not connected to any external power source, as seen in diodes.

The significance of this built-in potential across the junction, is that it opposes both the flow of holes and electrons across the junction and is why it is called the potential barrier. In practice, a **PN junction** is formed within a single crystal of material rather than just simply joining or fusing together two separate pieces.

The result of this process is that the PN junction has rectifying current–voltage (IV or I–V) characteristics. Electrical contacts are fused onto either side of the semiconductor to enable an electrical connection to be made to an external circuit. The resulting electronic device that has been made is commonly called a PN junction Diode or simply Signal Diode.

Then we have seen here that a PN junction can be made by joining or diffusing together differently doped semiconductor materials to produce an electronic device called a diode which can be used as the basic semiconductor structure of rectifiers, all types of transistors, LED’s, solar cells, and many more such solid state devices.

In the next tutorial about the PN junction, we will look at one of the most interesting applications of the **PN junction** is its use in circuits as a diode. By adding connections to each end of the P-type and the N-type materials we can produce a two terminal device called a PN Junction Diode which can be biased by an external voltage to either block or allow the flow of current through it.

The effect described in the previous tutorial is achieved without any external voltage being applied to the actual PN junction resulting in the junction being in a state of equilibrium.

However, if we were to make electrical connections at the ends of both the N-type and the P-type materials and then connect them to a battery source, an additional energy source now exists to overcome the potential barrier.

The effect of adding this additional energy source results in the free electrons being able to cross the depletion region from one side to the other. The behaviour of the PN junction with regards to the potential barrier’s width produces an asymmetrical conducting two terminal device, better known as the **PN Junction Diode**.

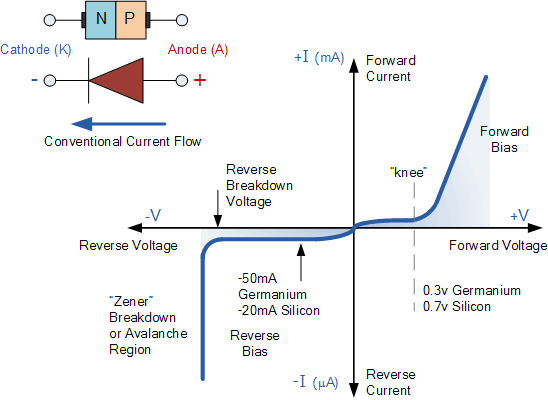
A PN Junction Diode is one of the simplest semiconductor devices around, and which has the characteristic of passing current in only one direction only. However, unlike a resistor, a diode does not behave linearly with respect to the applied voltage as the diode has an exponential current-voltage ( I-V ) relationship and therefore we can not described its operation by simply using an equation such as Ohm’s law.

If a suitable positive voltage (forward bias) is applied between the two ends of the PN junction, it can supply free electrons and holes with the extra energy they require to cross the junction as the width of the depletion layer around the PN junction is decreased.

By applying a negative voltage (reverse bias) results in the free charges being pulled away from the junction resulting in the depletion layer width being increased. This has the effect of increasing or decreasing the effective resistance of the junction itself allowing or blocking current flow through the diode.

Then the depletion layer widens with an increase in the application of a reverse voltage and narrows with an increase in the application of a forward voltage. This is due to the differences in the electrical properties on the two sides of the PN junction resulting in physical changes taking place. One of the results produces rectification as seen in the PN junction diodes static I-V (current-voltage) characteristics. Rectification is shown by an asymmetrical current flow when the polarity of bias voltage is altered as shown below.

### Junction Diode Symbol and Static I-V Characteristics

****

But before we can use the PN junction as a practical device or as a rectifying device we need to firstly **bias** the junction, ie connect a voltage potential across it. On the voltage axis above, “Reverse Bias” refers to an external voltage potential which increases the potential barrier. An external voltage which decreases the potential barrier is said to act in the “Forward Bias” direction.

There are two operating regions and three possible “biasing” conditions for the standard **Junction Diode** and these are:

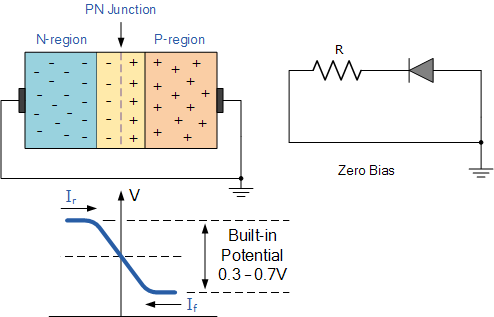
* 1. Zero Bias – No external voltage potential is applied to the PN junction diode.
* 2. Reverse Bias – The voltage potential is connected negative, (-ve) to the P-type material and positive, (+ve) to the N-type material across the diode which has the effect of **Increasing** the PN junction diode’s width.
* 3. Forward Bias – The voltage potential is connected positive, (+ve) to the P-type material and negative, (-ve) to the N-type material across the diode which has the effect of **Decreasing** the PN junction diodes width.

**Zero Biased Junction Diode**

When a diode is connected in a **Zero Bias** condition, no external potential energy is applied to the PN junction. However if the diodes terminals are shorted together, a few holes (majority carriers) in the P-type material with enough energy to overcome the potential barrier will move across the junction against this barrier potential. This is known as the “**Forward Current**” and is referenced as IF

Likewise, holes generated in the N-type material (minority carriers), find this situation favourable and move across the junction in the opposite direction. This is known as the “**Reverse Current**” and is referenced as IR. This transfer of electrons and holes back and forth across the PN junction is known as diffusion, as shown below.

**Zero Biased PN Junction Diode**

****

The potential barrier that now exists discourages the diffusion of any more majority carriers across the junction. However, the potential barrier helps minority carriers (few free electrons in the P-region and few holes in the N-region) to drift across the junction.

Then an “Equilibrium” or balance will be established when the majority carriers are equal and both moving in opposite directions, so that the net result is zero current flowing in the circuit. When this occurs the junction is said to be in a state of “**Dynamic Equilibrium**“.

The minority carriers are constantly generated due to thermal energy so this state of equilibrium can be broken by raising the temperature of the PN junction causing an increase in the generation of minority carriers, thereby resulting in an increase in leakage current but an electric current cannot flow since no circuit has been connected to the PN junction.

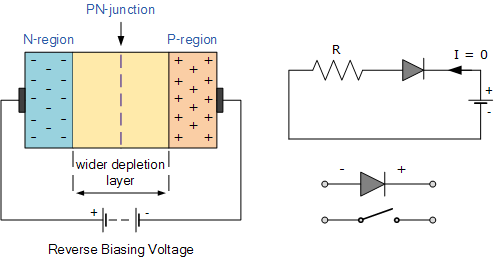
## Reverse Biased PN Junction Diode

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.

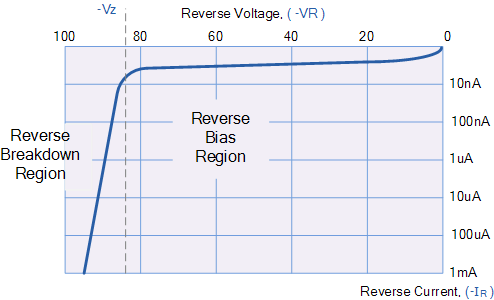
### Increase in the Depletion Layer due to Reverse Bias

****

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 micro-amperes, ( 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 diode’s 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

****

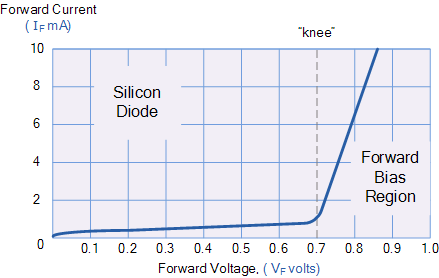
Sometimes this avalanche effect has practical applications in voltage stabilising circuits where a series limiting resistor is used with the diode to limit this reverse breakdown current to a preset maximum value thereby producing a fixed voltage output across the diode. These types of diodes are commonly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

## Forward Biased PN 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.

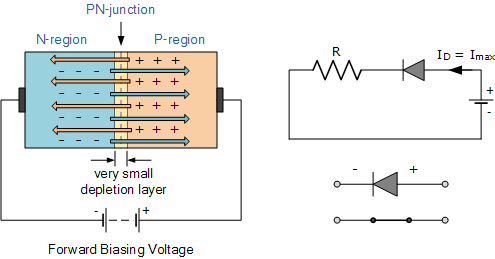
This is because the negative voltage pushes or repels electrons towards the junction giving them the energy to cross over and combine with the holes being pushed in the opposite direction towards the junction by the positive voltage. This results in a characteristics curve of zero current flowing up 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

****

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.

### Reduction in the Depletion Layer due to 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.

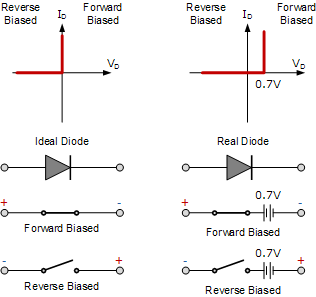
**Junction Diode Summary**

The PN junction region of a **Junction Diode** has the following important characteristics:

* Semiconductors contain two types of mobile charge carriers, “Holes” and “Electrons”.
* The holes are positively charged while the electrons negatively charged.
* A semiconductor may be doped with donor impurities such as Antimony (N-type doping), so that it contains mobile charges which are primarily electrons.
* A semiconductor may be doped with acceptor impurities such as Boron (P-type doping), so that it contains mobile charges which are mainly holes.
* The junction region itself has no charge carriers and is known as the depletion region.
* The junction (depletion) region has a physical thickness that varies with the applied voltage.
* When a diode is **Zero Biased** no external energy source is applied and a natural **Potential Barrier** is developed across a depletion layer which is approximately 0.5 to 0.7v for silicon diodes and approximately 0.3 of a volt for germanium diodes.
* When a junction diode is **Forward Biased** the thickness of the depletion region reduces and the diode acts like a short circuit allowing full current to flow.
* When a junction diode is **Reverse Biased** the thickness of the depletion region increases and the diode acts like an open circuit blocking any current flow, (only a very small leakage current).

We have also seen above that the diode is two terminal non-linear device whose I-V characteristic are polarity dependent as depending upon the polarity of the applied voltage, VD the diode is either *Forward Biased*, VD > 0 or *Reverse Biased*, VD < 0. Either way we can model these current-voltage characteristics for both an ideal diode and for a real silicon diode as shown:

**Junction Diode Ideal and Real Characteristics**

****

In the next tutorial about diodes, we will look at the small signal diode sometimes called a switching diode which is used in general electronic circuits. As its name implies, the signal diode is designed for low-voltage or high frequency signal applications such as in radio or digital switching circuits.

Signal diodes, such as the 1N4148 only pass very small electrical currents as opposed to the high-current mains rectification diodes in which silicon diodes are usually used. Also in the next tutorial we will examine the Signal Diode static current-voltage characteristics curve and parameters.

# The Signal Diode

Signal Diodes are small two-terminal which conducts current when forward biased and blocks current flow when reverse biased

The semiconductor **Signal Diode** is a small non-linear semiconductor devices generally used in electronic circuits, where small currents or high frequencies are involved such as in radio, television and digital logic circuits.

Signal diodes, also sometimes known by its older name of the **Point Contact Diode** or the **Glass Passivated Diode**, are physically very small in size compared to their larger Power Diode cousins.

Generally, the PN junction of a small signal diode is encapsulated in glass to protect the PN junction, and usually have a red or black band at one end of their body to help identify which end is the cathode terminal. The most widely used of all the glass encapsulated signal diodes is the very common *1N4148* and its equivalent *1N914* signal diode.

Small signal and switching diodes have much lower power and current ratings, around 150mA, 500mW maximum compared to rectifier diodes, but they can function better in high frequency applications or in clipping and switching applications that deal with short-duration pulse waveforms.

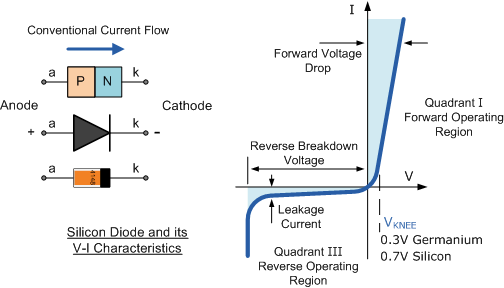
The characteristics of a signal point contact diode are different for both germanium and silicon types and are given as:

1. Germanium Signal Diodes – These have a low reverse resistance value giving a lower forward volt drop across the junction, typically only about 0.2 to 0.3v, but have a higher forward resistance value because of their small junction area.

2. Silicon Signal Diodes – These have a very high value of reverse resistance and give a forward volt drop of about 0.6 to 0.7v across the junction. They have fairly low values of forward resistance giving them high peak values of forward current and reverse voltage.

The electronic symbol given for any type of diode is that of an arrow with a bar or line at its end and this is illustrated below along with the Steady State V-I Characteristics Curve.

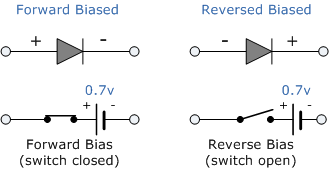
**Silicon Diode V-I Characteristic Curve**

****

The arrow always points in the direction of conventional current flow through the diode meaning that the diode will only conduct if a positive supply is connected to the Anode, ( a ) terminal and a negative supply is connected to the Cathode ( k ) terminal thus only allowing current to flow through it in one direction only, acting more like a one way electrical valve, ( Forward Biased Condition ).

However, we know from the previous tutorial that if we connect the external energy source in the other direction the diode will block any current flowing through it and instead will act like an open switch, ( Reversed Biased Condition ) as shown below.

### Forward and Reversed Biased Diode

****

Then we can say that an ideal small signal diode conducts current in one direction ( forward-conducting ) and blocks current in the other direction ( reverse-blocking ). Signal Diodes are used in a wide variety of applications such as a switch in rectifiers, current limiters, voltage snubbers or in wave-shaping circuits.

**Signal Diode Parameters**

**Signal Diodes** are manufactured in a range of voltage and current ratings and care must be taken when choosing a diode for a certain application. There are a bewildering array of static characteristics associated with the humble signal diode but the more important ones are.

**1. Maximum Forward Current**

The **Maximum Forward Current** ( IF(max) ) is as its name implies the *maximum forward current* allowed to flow through the device. When the diode is conducting in the forward bias condition, it has a very small “ON” resistance across the PN junction and therefore, power is dissipated across this junction ( Ohm´s Law ) in the form of heat.

Then, exceeding its ( IF(max) ) value will cause more heat to be generated across the junction and the diode will fail due to thermal overload, usually with destructive consequences. When operating diodes around their maximum current ratings it is always best to provide additional cooling to dissipate the heat produced by the diode.

For example, our small 1N4148 signal diode has a maximum current rating of about 150mA with a power dissipation of 500mW at 25oC. Then a resistor must be used in series with the diode to limit the forward current, ( IF(max) ) through it to below this value.

**2. Peak Inverse Voltage**

The **Peak Inverse Voltage** (PIV) or *Maximum Reverse Voltage* ( VR(max) ), is the maximum allowable **Reverse** operating voltage that can be applied across the diode without reverse breakdown and damage occurring to the device. This rating therefore, is usually less than the “avalanche breakdown” level on the reverse bias characteristic curve. Typical values of VR(max) range from a few volts to thousands of volts and must be considered when replacing a diode.

The peak inverse voltage is an important parameter and is mainly used for rectifying diodes in AC rectifier circuits with reference to the amplitude of the voltage were the sinusoidal waveform changes from a positive to a negative value on each and every cycle.

**3. Total Power Dissipation**

Signal diodes have a **Total Power Dissipation**, ( PD(max) ) rating. This rating is the maximum possible power dissipation of the diode when it is forward biased (conducting). When current flows through the signal diode the biasing of the PN junction is not perfect and offers some resistance to the flow of current resulting in power being dissipated (lost) in the diode in the form of heat.

As small signal diodes are non-linear devices the resistance of the PN junction is not constant, it is a dynamic property then we cannot use Ohms Law to define the power in terms of current and resistance or voltage and resistance as we can for resistors. Then to find the power that will be dissipated by the diode we must multiply the voltage drop across it times the current flowing through it: PD = V\*I

**4. Maximum Operating Temperature**

The **Maximum Operating Temperature** actually relates to the *Junction Temperature* ( TJ ) of the diode and is related to maximum power dissipation. It is the maximum temperature allowable before the structure of the diode deteriorates and is expressed in units of degrees centigrade per Watt, ( oC/W ).

This value is linked closely to the maximum forward current of the device so that at this value the temperature of the junction is not exceeded. However, the maximum forward current will also depend upon the ambient temperature in which the device is operating so the maximum forward current is usually quoted for two or more ambient temperature values such as 25oC or 70oC.

Then there are three main parameters that must be considered when either selecting or replacing a signal diode and these are:

* The Reverse Voltage Rating
* The Forward Current Rating
* The Forward Power Dissipation Rating

Other types of specialized diodes not included here are Photo-Diodes, Zener Diodes, PIN Diodes, Tunnel Diodes, Light Emitting Diodes(LED) and Schottky Barrier Diodes. By adding more PN junctions to the basic two layer diode structure other types of semiconductor devices can be made.

For example a three layer semiconductor device becomes a Transistor, a four layer semiconductor device becomes a Thyristor or Silicon Controlled Rectifier and five layer devices known as Triac’s are also available.

# Power Diodes and Rectifiers

Power Diodes are semiconductor pn-junctions capable of passing large currents at high voltage values for use in rectifier circuits.

a semiconductor signal diode will only conduct current in one direction from its anode to its cathode (forward direction), but not in the reverse direction acting a bit like an electrical one way valve.

A widely used application of this feature and diodes in general is in the conversion of an alternating voltage (AC) into a continuous voltage (Unidirectional). In other words, Rectification.

But small signal diodes can also be used as rectifiers in low-power, low current (less than 1-amp) rectifiers or applications, but where larger forward bias currents or higher reverse bias blocking voltages are involved the PN junction of a small signal diode would eventually overheat and melt so larger more robust **Power Diodes** are used instead.

The power semiconductor diode, known simply as the **Power Diode**, has a much larger PN junction area compared to its smaller signal diode cousin, resulting in a high forward current capability of up to several hundred amps (KA) and a reverse blocking voltage of up to several thousand volts (KV).

Since the power diode has a large PN junction, it is not suitable for high frequency applications above 1MHz, but special and expensive high frequency, high current diodes are available. For high frequency rectifier applications Sochottky Diodes are generally used because of their short reverse recovery time and low voltage drop in their forward bias condition.

Power diodes provide uncontrolled rectification of power and are used in applications such as battery charging and DC power supplies as well as AC rectifiers and inverters. Due to their high current and voltage characteristics they can also be used as free-wheeling diodes and snubber networks.

Power diodes are designed to have a forward “ON” resistance of fractions of an Ohm while their reverse blocking resistance is in the mega-Ohms range. Some of the larger value power diodes are designed to be “stud mounted” onto heat sinks reducing their thermal resistance to between 0.1 to 1oC/Watt.

If an alternating voltage is applied across a power diode, during the positive half cycle the diode will conduct passing current and during the negative half cycle the diode will not conduct blocking the flow of current. Then conduction through the power diode only occurs during the positive half cycle and is therefore unidirectional i.e. DC as shown.

### Power Diode Rectifier

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode39.gif

Power diodes can be used individually as above or connected together to produce a variety of rectifier circuits such as “Half-Wave”, “Full-Wave” or as “Bridge Rectifiers”. Each type of rectifier circuit can be classed as either uncontrolled, half-controlled or fully controlled where an uncontrolled rectifier uses only power diodes, a fully controlled rectifier uses thyristors (SCRs) and a half controlled rectifier is a mixture of both diodes and thyristors.

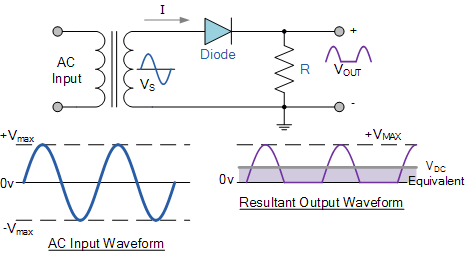
The most commonly used individual power diode for basic electronics applications is the general purpose 1N400x Series Glass Passivated type rectifying diode with standard ratings of continuous forward rectified current of about 1.0 ampere and reverse blocking voltage ratings from 50v for the 1N4001 up to 1000v for the 1N4007, with the small 1N4007GP being the most popular for general purpose mains voltage rectification.

## Half Wave Rectification

A rectifier is a circuit which converts the Alternating Current (AC) input power into a Direct Current (DC) output power. The input power supply may be either a single-phase or a multi-phase supply with the simplest of all the rectifier circuits being that of the **Half Wave Rectifier**.

The power diode in a half wave rectifier circuit passes just one half of each complete sine wave of the AC supply in order to convert it into a DC supply. Then this type of circuit is called a “half-wave” rectifier because it passes only half of the incoming AC power supply as shown below.

### Half Wave Rectifier Circuit



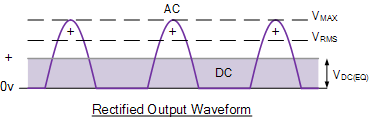
During each “positive” half cycle of the AC sine wave, the diode is *forward biased* as the anode is positive with respect to the cathode resulting in current flowing through the diode.

Since the DC load is resistive (resistor, R), the current flowing in the load resistor is therefore proportional to the voltage (Ohm´s Law), and the voltage across the load resistor will therefore be the same as the supply voltage, Vs (minus Vƒ), that is the “DC” voltage across the load is sinusoidal for the first half cycle only so Vout = Vs.

During each “negative” half cycle of the AC sinusoidal input waveform, the diode is *reverse biased* as the anode is negative with respect to the cathode. Therefore, NO current flows through the diode or circuit. Then in the negative half cycle of the supply, no current flows in the load resistor as no voltage appears across it so therefore, Vout = 0.

The current on the DC side of the circuit flows in one direction only making the circuit **Unidirectional**. As the load resistor receives from the diode a positive half of the waveform, zero volts, a positive half of the waveform, zero volts, etc, the value of this irregular voltage would be equal in value to an equivalent DC voltage of 0.318\*Vmax of the input sinusoidal waveform or 0.45\*Vrms of the input sinusoidal waveform.

Then the equivalent DC voltage, VDC across the load resistor is calculated as follows.



### Half-wave Rectifier with Smoothing Capacitor

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode17.gif

When rectification is used to provide a direct voltage (DC) power supply from an alternating (AC) source, the amount of ripple voltage can be further reduced by using larger value capacitors but there are limits both on cost and size to the types of smoothing capacitors used.

For a given capacitor value, a greater load current (smaller load resistance) will discharge the capacitor more quickly ( RC Time Constant ) and so increases the ripple obtained. Then for single phase, half-wave rectifier circuit using a power diode it is not very practical to try and reduce the ripple voltage by capacitor smoothing alone. In this instance it would be more practical to use “Full-wave Rectification” instead.

In practice, the half-wave rectifier is used most often in low-power applications because of their major disadvantages being. The output amplitude is less than the input amplitude, there is no output during the negative half cycle so half the power is wasted and the output is pulsed DC resulting in excessive ripple.

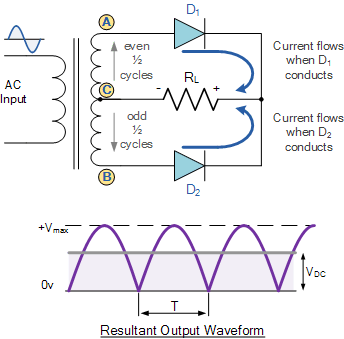
To overcome these disadvantages a number of **Power Diode** are connected together to produce a Full Wave Rectifier as discussed in the next tutorial.

**Full Wave Rectifier**

Full wave rectifiers have some fundamental advantages over their half wave rectifier counterparts. The average (DC) output voltage is higher than for half wave, the output of the full wave rectifier has much less ripple than that of the half wave rectifier producing a smoother output waveform.

In a **Full Wave Rectifier** circuit two diodes are now used, one for each half of the cycle. A multiple winding transformer is used whose secondary winding is split equally into two halves with a common centre tapped connection, (C). This configuration results in each diode conducting in turn when its anode terminal is positive with respect to the transformer centre point C producing an output during both half-cycles, twice that for the half wave rectifier so it is 100% efficient as shown below.

### Full Wave Rectifier Circuit

****

The full wave rectifier circuit consists of two *power diodes* connected to a single load resistance (RL) with each diode taking it in turn to supply current to the load. When point A of the transformer is positive with respect to point C, diode D1 conducts in the forward direction as indicated by the arrows.

When point B is positive (in the negative half of the cycle) with respect to point C, diode D2 conducts in the forward direction and the current flowing through resistor R is in the same direction for both half-cycles. As the output voltage across the resistor R is the phasor sum of the two waveforms combined, this type of full wave rectifier circuit is also known as a “bi-phase” circuit.

The main disadvantage of this type of full wave rectifier circuit is that a larger transformer for a given power output is required with two separate but identical secondary windings making this type of full wave rectifying circuit costly compared to the “Full Wave Bridge Rectifier” circuit equivalent.

## The Full Wave 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

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode20.gif

The four diodes labeled D1 to D4 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.

### The Positive Half-cycle

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode21.gif

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 0.637Vmax.

### The Negative Half-cycle



**Typical Bridge Rectifier**

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\*0.7 = 1.4V ) less than the input VMAX amplitude. The ripple frequency is now twice the supply frequency (e.g. 100Hz for a 50Hz supply or 120Hz for a 60Hz supply.)

Although we can use four individual power diodes to make a full wave bridge rectifier, pre-made bridge rectifier components are available “off-the-shelf” in a range of different voltage and current sizes that can be soldered directly into a PCB circuit board or be connected by spade connectors.

The image to the right shows a typical single phase bridge rectifier with one corner cut off. This cut-off corner indicates that the terminal nearest to the corner is the positive or +ve output terminal or lead with the opposite (diagonal) lead being the negative or -ve output lead. The other two connecting leads are for the input alternating voltage from a transformer secondary winding.

## The Smoothing Capacitor

We saw in the previous section that the single phase half-wave rectifier produces an output wave every half cycle and that it was not practical to use this type of circuit to produce a steady DC supply. The full-wave bridge rectifier however, gives us a greater mean DC value (0.637 Vmax) with less superimposed ripple while the output waveform is twice that of the frequency of the input supply frequency.

We can improve the average DC output of the rectifier while at the same time reducing the AC variation of the rectified output by using smoothing capacitors to filter the output waveform. Smoothing or reservoir capacitors connected in parallel with the load across the output of the full wave bridge rectifier circuit increases the average DC output level even higher as the capacitor acts like a storage device as shown below.

### Full-wave Rectifier with Smoothing Capacitor

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode23.gif

### The smoothing capacitor converts the full-wave rippled output of the rectifier into a more smooth DC output voltage.

# The Zener Diode

A Semiconductor Diode blocks current in the reverse direction, but will suffer from premature breakdown or damage if the reverse voltage applied across becomes too high.

However, the **Zener Diode** or “Breakdown Diode”, as they are sometimes referred too, are basically the same as the standard PN junction diode but they are specially designed to have a low and specified **Reverse Breakdown Voltage** which takes advantage of any reverse voltage applied to it.

The **Zener diode** behaves just like a normal general-purpose diode consisting of a silicon PN junction and when biased in the forward direction, that is Anode positive with respect to its Cathode, it behaves just like a normal signal diode passing the rated current.

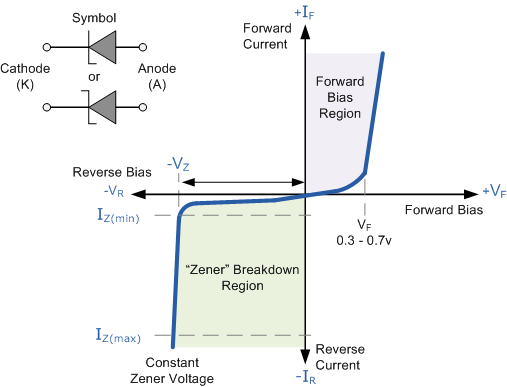
However, unlike a conventional diode that blocks any flow of current through itself when reverse biased, that is the Cathode becomes more positive than the Anode, as soon as the reverse voltage reaches a pre-determined value, the zener diode begins to conduct in the reverse direction.

This is because when the reverse voltage applied across the zener diode exceeds the rated voltage of the device 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 reverse voltages. The voltage point at which the voltage across the zener diode becomes stable is called the “zener voltage”, ( Vz ) and for zener diodes this voltage can range from less than one volt to a few hundred 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



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.

This voltage remains almost constant even with large changes in current providing the zener diodes current remains between the breakdown current IZ(min) and its maximum current rating IZ(max).

This ability of the zener diode to control itself can be used to great effect to regulate or stabilise 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 characteristic of the zener diode as it can be used in the simplest types of voltage regulator applications.

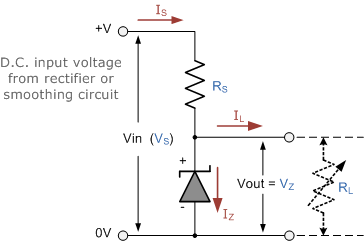
The function of a voltage regulator is to provide a constant output voltage to a load connected in parallel with it in spite of the ripples in the supply voltage or variations in the load current. A zener diode will continue to regulate its voltage until the diodes holding current falls below the minimum IZ(min) value in the reverse breakdown region.

## The Zener Diode Regulator

**Zener Diodes** can be used to produce a stabilised voltage output with low ripple under varying load current conditions. By passing a small current through the diode from a voltage source, via a suitable current limiting resistor (RS), the zener diode will conduct sufficient current to maintain a voltage drop of Vout.

We remember from the previous tutorials that the DC output voltage from the half or full-wave rectifiers contains ripple superimposed onto the DC voltage and that as the load value changes so to does the average output voltage. By connecting a simple zener stabiliser circuit as shown below across the output of the rectifier, a more stable output voltage can be produced.

### Zener Diode Regulator



Resistor, RS is connected in series with the zener diode to limit the current flow through the diode with the voltage source, VS being connected across the combination. The stabilised output voltage Vout is taken from across the zener diode.

The zener diode is connected with its cathode terminal connected to the positive rail of the DC supply so it is reverse biased and will be operating in its breakdown condition. Resistor RS is selected so to limit the maximum current flowing in the circuit.

With no load connected to the circuit, the load current will be zero, ( IL = 0 ), and all the circuit current passes through the zener diode which in turn dissipates its maximum power. Also a small value of the series resistor RS will result in a greater diode current when the load resistance RL is connected and large as this will increase the power dissipation requirement of the diode so care must be taken when selecting the appropriate value of series resistance so that the zener’s maximum power rating is not exceeded under this no-load or high-impedance condition.

The load is connected in parallel with the zener diode, so the voltage across RL is always the same as the zener voltage, ( VR = VZ ). There is a minimum zener current for which the stabilisation of the voltage is effective and the zener current must stay above this value operating under load within its breakdown region at all times. The upper limit of current is of course dependant upon the power rating of the device. The supply voltage VS must be greater than VZ.

One small problem with zener diode stabiliser circuits is that the diode can sometimes generate electrical noise on top of the DC supply as it tries to stabilise the voltage. Normally this is not a problem for most applications but the addition of a large value decoupling capacitor across the zener’s output may be required to give additional smoothing.

Then to summarise a little. A zener diode is always operated in its reverse biased condition. As such a simple voltage regulator circuit can be designed using a zener diode to maintain a constant DC output voltage across the load in spite of variations in the input voltage or changes in the load current.

The zener voltage regulator consists of a current limiting resistor RS connected in series with the input voltage VS with the zener diode connected in parallel with the load RL in this reverse biased condition. The stabilised output voltage is always selected to be the same as the breakdown voltage VZ of the diode.

The values of the individual Zener diodes can be chosen to suit the application while the silicon diode will always drop about 0.6 – 0.7V in the forward bias condition. The supply voltage, Vin must of course be higher than the largest output reference voltage and in our example above this is 19v.

A typical **zener diode** for general electronic circuits is the 500mW, *BZX55* series or the larger 1.3W, *BZX85* series were the zener voltage is given as, for example, *C7V5* for a 7.5V diode giving a diode reference number of *BZX55C7V5*.

The 500mW series of zener diodes are available from about 2.4 up to about 100 volts and typically have the same sequence of values as used for the 5% (E24) resistor series with the individual voltage ratings for these small but very useful diodes are given in the table below.

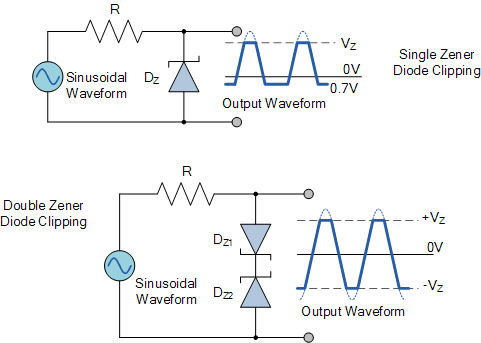
## Zener Diode Clipping Circuits

Thus far we have looked at how a zener diode can be used to regulate a constant DC source but what if the input signal was not steady state DC but an alternating AC waveform how would the zener diode react to a constantly changing signal.

Diode clipping and clamping circuits are circuits that are used to shape or modify an input AC waveform (or any sinusoid) producing a differently shape output waveform depending on the circuit arrangement. Diode clipper circuits are also called limiters because they limit or clip-off the positive (or negative) part of an input AC signal. As zener clipper circuits limit or cut-off part of the waveform across them, they are mainly used for circuit protection or in waveform shaping circuits.

For example, if we wanted to clip an output waveform at +7.5V, we would use a 7.5V zener diode. If the output waveform tries to exceed the 7.5V limit, the zener diode will “clip-off” the excess voltage from the input producing a waveform with a flat top still keeping the output constant at +7.5V. Note that in the forward bias condition a zener diode is still a diode and when the AC waveform output goes negative below -0.7V, the zener diode turns “ON” like any normal silicon diode would and clips the output at -0.7V as shown below.

### Square Wave Signal

****

The back to back connected zener diodes can be used as an AC regulator producing what is jokingly called a “poor man’s square wave generator”. Using this arrangement we can clip the waveform between a positive value of +8.2V and a negative value of -8.2V for a 7.5V zener diode.

So for example, if we wanted to clip an output waveform between two different minimum and maximum values of say, +8V and -6V, we would simply use two differently rated zener diodes. Note that the output will actually clip the AC waveform between +8.7V and -6.7V due to the addition of the forward biasing diode voltage.

In other words a peak-to-peak voltage of 15.4 volts instead of expected 14 volts, as the forward bias volt drop across the diode adds another 0.7 volts in each direction.

This type of clipper configuration is fairly common for protecting an electronic circuit from over voltage. The two zener’s are generally placed across the power supply input terminals and during normal operation, one of the zener diodes is “OFF” and the diodes have little or no affect. However, if the input voltage waveform exceeds its limit, then the zener’s turn “ON” and clip the input to protect the circuit.

In the next tutorial about diodes, we will look at using the forward biased PN junction of a diode to produce light. We know from the previous tutorials that when charge carriers move across the junction, electrons combine with holes and energy is lost in the form of heat, but also some of this energy is dissipated as photons but we can not see them.

If we place a translucent lens around the junction, visible light will be produced and the diode becomes a light source. This effect produces another type of diode known commonly as the Light Emitting Diode which takes advantage of this light producing characteristic to emit light (photons) in a variety of colours and wavelengths

# The Light Emitting Diode

They are the most visible type of diode, that emit a fairly narrow bandwidth of either visible light at different coloured wavelengths, invisible infra-red light for remote controls or laser type light when a forward current is passed through them.

The “**Light Emitting Diode**” or LED as it is more commonly called, is basically just a specialised type of diode as they have very similar electrical characteristics to a PN junction diode. This means that an LED will pass current in its forward direction but block the flow of current in the reverse direction.

Light emitting diodes are made from a very thin layer of fairly heavily doped semiconductor material and depending on the semiconductor material used and the amount of doping, when forward biased an LED will emit a coloured light at a particular spectral wavelength.

When the diode is forward biased, electrons from the semiconductors conduction band recombine with holes from the valence band releasing sufficient energy to produce photons which emit a monochromatic (single colour) of light. Because of this thin layer a reasonable number of these photons can leave the junction and radiate away producing a coloured light output.

**LED Construction**

Then we can say that when operated in a forward biased direction **Light Emitting Diodes** are semiconductor devices that convert electrical energy into light energy.

The construction of a Light Emitting Diode is very different from that of a normal signal diode. The PN junction of an LED is surrounded by a transparent, hard plastic epoxy resin hemispherical shaped shell or body which protects the LED from both vibration and shock.

Surprisingly, an LED junction does not actually emit that much light so the epoxy resin body is constructed in such a way that the photons of light emitted by the junction are reflected away from the surrounding substrate base to which the diode is attached and are focused upwards through the domed top of the LED, which itself acts like a lens concentrating the amount of light. This is why the emitted light appears to be brightest at the top of the LED.

However, not all LEDs are made with a hemispherical shaped dome for their epoxy shell. Some indication LEDs have a rectangular or cylindrical shaped construction that has a flat surface on top or their body is shaped into a bar or arrow. Generally, all LED’s are manufactured with two legs protruding from the bottom of the body.

Also, nearly all modern light emitting diodes have their cathode, ( – ) terminal identified by either a notch or flat spot on the body or by the cathode lead being shorter than the other as the anode ( + ) lead is longer than the cathode (k).

Unlike normal incandescent lamps and bulbs which generate large amounts of heat when illuminated, the light emitting diode produces a “cold” generation of light which leads to high efficiencies than the normal “light bulb” because most of the generated energy radiates away within the visible spectrum. Because LEDs are solid-state devices, they can be extremely small and durable and provide much longer lamp life than normal light sources.

**Light Emitting Diode Colours**

So how does a light emitting diode get its colour. Unlike normal signal diodes which are made for detection or power rectification, and which are made from either Germanium or Silicon semiconductor materials, **Light Emitting Diodes** are made from exotic semiconductor compounds such as Gallium Arsenide (GaAs), Gallium Phosphide (GaP), Gallium Arsenide Phosphide (GaAsP), Silicon Carbide (SiC) or Gallium Indium Nitride (GaInN) all mixed together at different ratios to produce a distinct wavelength of colour.

Different LED compounds emit light in specific regions of the visible light spectrum and therefore produce different intensity levels. The exact choice of the semiconductor material used will determine the overall wavelength of the photon light emissions and therefore the resulting colour of the light emitted.

**Light Emitting Diode Colours**

|  |  |  |  |
| --- | --- | --- | --- |
| **Typical LED Characteristics** | | | |
| **Semiconductor Material** | **Wavelength** | **Colour** | **VF @ 20mA** |
| **GaAs** | **850-940nm** | **Infra-Red** | **1.2v** |
| **GaAsP** | **630-660nm** | **Red** | **1.8v** |
| **GaAsP** | **605-620nm** | **Amber** | **2.0v** |
| **GaAsP:N** | **585-595nm** | **Yellow** | **2.2v** |
| **AlGaP** | **550-570nm** | **Green** | **3.5v** |
| **SiC** | **430-505nm** | **Blue** | **3.6v** |
| **GaInN** | **450nm** | **White** | **4.0v** |

Thus, the actual colour of a light emitting diode is determined by the wavelength of the light emitted, which in turn is determined by the actual semiconductor compound used in forming the PN junction during manufacture.

Therefore the colour of the light emitted by an LED is NOT determined by the colouring of the LED’s plastic body although these are slightly coloured to both enhance the light output and to indicate its colour when its not being illuminated by an electrical supply.

Light emitting diodes are available in a wide range of colours with the most common being **RED, AMBER,  YELLOW** and **GREEN** and are thus widely used as visual indicators and as moving light displays.

Recently developed blue and white coloured LEDs are also available but these tend to be much more expensive than the normal standard colours due to the production costs of mixing together two or more complementary colours at an exact ratio within the semiconductor compound and also by injecting nitrogen atoms into the crystal structure during the doping process.

From the table above we can see that the main P-type dopant used in the manufacture of **Light Emitting Diodes** is Gallium (Ga, atomic number 31) and that the main N-type dopant used is Arsenic (As, atomic number 33) giving the resulting compound of Gallium Arsenide (GaAs) crystalline structure.

The problem with using Gallium Arsenide on its own as the semiconductor compound is that it radiates large amounts of low brightness infra-red radiation (850nm-940nm approx.) from its junction when a forward current is flowing through it.

The amount of infra-red light it produces is okay for television remote controls but not very useful if we want to use the LED as an indicating light. But by adding Phosphorus (P, atomic number 15), as a third dopant the overall wavelength of the emitted radiation is reduced to below 680nm giving visible red light to the human eye. Further refinements in the doping process of the PN junction have resulted in a range of colours spanning the spectrum of visible light as we have seen above as well as infra-red and ultra-violet wavelengths.

By mixing together a variety of semiconductor, metal and gas compounds the following list of LEDs can be produced.

**Types of Light Emitting Diode**

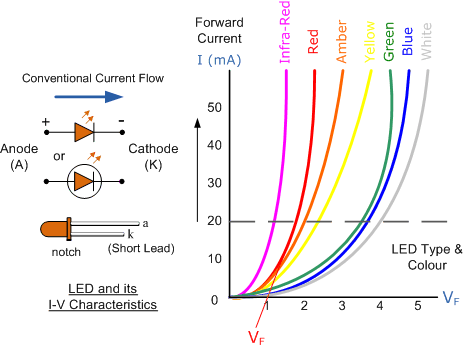
* Gallium Arsenide (GaAs) – infra-red
* Gallium Arsenide Phosphide (GaAsP) – red to infra-red, orange
* Aluminium Gallium Arsenide Phosphide (AlGaAsP) – high-brightness red, orange-red, orange, and yellow
* Gallium Phosphide (GaP) – red, yellow and green
* Aluminium Gallium Phosphide (AlGaP) – green
* Gallium Nitride (GaN) – green, emerald green
* Gallium Indium Nitride (GaInN) – near ultraviolet, bluish-green and blue
* Silicon Carbide (SiC) – blue as a substrate
* Zinc Selenide (ZnSe) – blue
* Aluminium Gallium Nitride (AlGaN) – ultraviolet

Like conventional PN junction diodes, light emitting diodes are current-dependent devices with its forward voltage drop VF, depending on the semiconductor compound (its light colour) and on the forward biased LED current. Most common LED’s require a forward operating voltage of between approximately 1.2 to 3.6 volts with a forward current rating of about 10 to 30 mA, with 12 to 20 mA being the most common range.

Both the forward operating voltage and forward current vary depending on the semiconductor material used but the point where conduction begins and light is produced is about 1.2V for a standard red LED to about 3.6V for a blue LED.

The exact voltage drop will of course depend on the manufacturer because of the different dopant materials and wavelengths used. The voltage drop across the LED at a particular current value, for example 20mA, will also depend on the initial conduction VF point. As an LED is effectively a diode, its forward current to voltage characteristics curves can be plotted for each diode colour as shown below.

**Light Emitting Diodes I-V Characteristics.**

****

Light Emitting Diode (LED) Schematic symbol and I-V Characteristics Curves  
showing the different colours available.

Before a light emitting diode can “emit” any form of light it needs a current to flow through it, as it is a current dependant device with their light output intensity being directly proportional to the forward current flowing through the LED.

As the LED is to be connected in a forward bias condition across a power supply it should be current limited using a series resistor to protect it from excessive current flow. Never connect an LED directly to a battery or power supply as it will be destroyed almost instantly because too much current will pass through and burn it out.

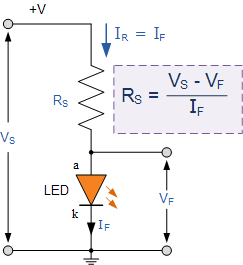
From the table above we can see that each LED has its own forward voltage drop across the PN junction and this parameter which is determined by the semiconductor material used, is the forward voltage drop for a specified amount of forward conduction current, typically for a forward current of 20mA.

In most cases LEDs are operated from a low voltage DC supply, with a series resistor, RS used to limit the forward current to a safe value from say 5mA for a simple LED indicator to 30mA or more where a high brightness light output is needed.

## LED Series Resistance.

The series resistor value RS is calculated by simply using Ohm´s Law, by knowing the required forward current IF of the LED, the supply voltage VS across the combination and the expected forward voltage drop of the LED, VF at the required current level, the current limiting resistor is calculated as:

### LED Series Resistor Circuit

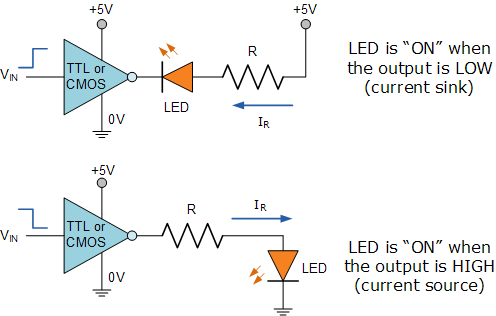
****

## LED Driver Circuits

Now that we know what is an LED, we need some way of controlling it by switching it “ON” and “OFF”. The output stages of both TTL and CMOS logic gates can both source and sink useful amounts of current therefore can be used to drive an LED. Normal integrated circuits (ICs) have an output drive current of up to 50mA in the sink mode configuration, but have an internally limited output current of about 30mA in the source mode configuration.

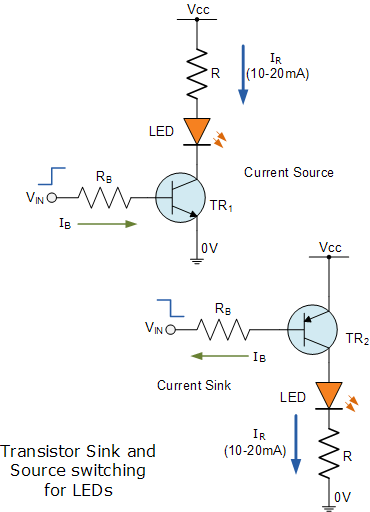
Either way the LED current must be limited to a safe value using a series resistor as we have already seen. Below are some examples of driving light emitting diodes using inverting ICs but the idea is the same for any type of integrated circuit output whether combinational or sequential.

### IC Driver Circuit

****

If more than one LED requires driving at the same time, such as in large LED arrays, or the load current is to high for the integrated circuit or we may just want to use discrete components instead of ICs, then an alternative way of driving the LEDs using either bipolar NPN or PNP transistors as switches is given below. Again as before, a series resistor, RS is required to limit the LED current.

### Transistor Driver Circuit

****

The brightness of a light emitting diode cannot be controlled by simply varying the current flowing through it. Allowing more current to flow through the LED will make it glow brighter but will also cause it to dissipate more heat. LEDs are designed to produce a set amount of light operating at a specific forward current ranging from about 10 to 20mA.

In situations where power savings are important, less current may be possible. However, reducing the current to below say 5mA may dim its light output too much or even turn the LED “OFF” completely. A much better way to control the brightness of LEDs is to use a control process known as “Pulse Width Modulation” or PWM, in which the LED is repeatedly turned “ON” and “OFF” at varying frequencies depending upon the required light intensity of the LED.

## Multi-coloured Light Emitting Diode

LEDs are available in a wide range of shapes, colours and various sizes with different light output intensities available, with the most common (and cheapest to produce) being the standard 5mm Red Gallium Arsenide Phosphide (GaAsP) LED.

LED’s are also available in various “packages” arranged to produce both letters and numbers with the most common being that of the “seven segment display” arrangement.

Nowadays, full colour flat screen LED displays, hand held devices and TV’s are available which use a vast number of multicoloured LED’s all been driven directly by their own dedicated IC.

Most light emitting diodes produce just a single output of coloured light however, multi-coloured LEDs are now available that can produce a range of different colours from within a single device. Most of these are actually two or three LEDs fabricated within a single package.

### Bi-colour Light Emitting Diodes

A bi-colour light emitting diode has two LEDs chips connected together in “inverse parallel” (one forwards, one backwards) combined in one single package. Bi-colour LEDs can produce any one of three colours for example, a red colour is emitted when the device is connected with current flowing in one direction and a green colour is emitted when it is biased in the other direction.

This type of bi-directional arrangement is useful for giving polarity indication, for example, the correct connection of batteries or power supplies etc. Also, a bi-directional current produces both colours mixed together as the two LEDs would take it in turn to illuminate if the device was connected (via a suitable resistor) to a low voltage, low frequency AC supply.

### A Bi-colour LED

### 

**LED Displays**

As well as individual colour or multi-colour LEDs, several light emitting diodes can be combined together within a single package to produce displays such as bargraphs, strips, arrays and seven segment displays.

A 7-segment LED display provides a very convenient way when decoded properly of displaying information or digital data in the form of numbers, letters or even alpha-numerical characters and as their name suggests, they consist of seven individual LEDs (the segments), within one single display package.

In order to produce the required numbers or characters from 0 to 9 and A to F respectively, on the display the correct combination of LED segments need to be illuminated. A standard seven segment LED display generally has eight input connections, one for each LED segment and one that acts as a common terminal or connection for all the internal segments.

* The Common Cathode Display (CCD) – In the common cathode display, all the cathode connections of the LEDs are joined together and the individual segments are illuminated by application of a HIGH, logic “1” signal.
* The Common Anode Display (CAD) – In the common anode display, all the anode connections of the LEDs are joined together and the individual segments are illuminated by connecting the terminals to a LOW, logic “0” signal.

## Opto-coupler

Finally, another useful application of light emitting diodes is in **Opto-coupling**. An opto-coupler or opto-isolator as it is also called, is a single electronic device that consists of a light emitting diode combined with either a photo-diode, photo-transistor or photo-triac to provide an optical signal path between an input connection and an output connection while maintaining electrical isolation between two circuits.

An opto-isolator consists of a light proof plastic body that has a typical breakdown voltages between the input (photo-diode) and the output (photo-transistor) circuit of up to 5000 volts. This electrical isolation is especially useful where the signal from a low voltage circuit such as a battery powered circuit, computer or microcontroller, is required to operate or control another external circuit operating at a potentially dangerous mains voltage.

### Photo-diode and Photo-transistor Opto-couplers

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode47.gif

The two components used in an opto-isolator, an optical transmitter such as an infra-red emitting Gallium Arsenide LED and an optical receiver such as a photo-transistor are closely optically coupled and use light to send signals and/or information between its input and output. This allows information to be transferred between circuits without an electrical connection or common ground potential.

Opto-isolators are digital or switching devices, so they transfer either “ON-OFF” control signals or digital data. Analogue signals can be transferred by means of frequency or pulse-width modulation.

# Diode Clipping Circuits

The **Diode Clipper**, also known as a Diode Limiter, is a wave shaping circuit that takes an input waveform and clips or cuts off its top half, bottom half or both halves together.

This clipping of the input signal produces an output waveform that resembles a flattened version of the input. For example, the half-wave rectifier is a clipper circuit, since all voltages below zero are eliminated.

But **Diode Clipping Circuits** can be used a variety of applications to modify an input waveform using signal and Schottky diodes or to provide over-voltage protection using zener diodes to ensure that the output voltage never exceeds a certain level protecting the circuit from high voltage spikes. Then diode clipping circuits can be used in voltage limiting applications.

We saw in the Signal Diodes tutorial that when a diode is forward biased it allows current to pass through itself clamping the voltage. When the diode is reverse biased, no current flows through it and the voltage across its terminals is unaffected, and this is the basic operation of the diode clipping circuit.

Although the input voltage to diode clipping circuits can have any waveform shape, we will assume here that the input voltage is sinusoidal. Consider the circuits below.

### Positive Diode Clipping Circuits

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode60.gif

In this diode clipping circuit, the diode is forward biased (anode more positive than cathode) during the positive half cycle of the sinusoidal input waveform. For the diode to become forward biased, it must have the input voltage magnitude greater than +0.7 volts (0.3 volts for a germanium diode).

When this happens the diodes begins to conduct and holds the voltage across itself constant at 0.7V until the sinusoidal waveform falls below this value. Thus the output voltage which is taken across the diode can never exceed 0.7 volts during the positive half cycle.

During the negative half cycle, the diode is reverse biased (cathode more positive than anode) blocking current flow through itself and as a result has no effect on the negative half of the sinusoidal voltage which passes to the load unaltered. Thus the diode limits the positive half of the input waveform and is known as a positive clipper circuit.

### Negative Diode Clipping Circuits

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode61.gif

Here the reverse is true. The diode is forward biased during the negative half cycle of the sinusoidal waveform and limits or clips it to –0.7 volts while allowing the positive half cycle to pass unaltered when reverse biased. As the diode limits the negative half cycle of the input voltage it is therefore called a negative clipper circuit.

### Clipping of Both Half Cycles

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode62.gif

If we connected two diodes in inverse parallel as shown, then both the positive and negative half cycles would be clipped as diode D1 clips the positive half cycle of the sinusoidal input waveform while diode D2 clips the negative half cycle. Then diode clipping circuits can be used to clip the positive half cycle, the negative half cycle or both.

For ideal diodes the output waveform above would be zero. However, due to the forward bias voltage drop across the diodes the actual clipping point occurs at +0.7 volts and –0.7 volts respectively. But we can increase this ±0.7V threshold to any value we want up to the maximum value, (VPEAK) of the sinusoidal waveform either by connecting together more diodes in series creating multiples of 0.7 volts, or by adding a voltage bias to the diodes.

## Biased Diode Clipping Circuits

To produce diode clipping circuits for voltage waveforms at different levels, a bias voltage, VBIAS is added in series with the diode to produce a combination clipper as shown. The voltage across the series combination must be greater than VBIAS + 0.7V before the diode becomes sufficiently forward biased to conduct. For example, if the VBIAS level is set at 4.0 volts, then the sinusoidal voltage at the diode’s anode terminal must be greater than **4.0 + 0.7 = 4.7 volts** for it to become forward biased. Any anode voltage levels above this bias point are clipped off.

### Positive Bias Diode Clipping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode63.gif

Likewise, by reversing the diode and the battery bias voltage, when a diode conducts the negative half cycle of the output waveform is held to a level –VBIAS – 0.7V as shown.

### Negative Bias Diode Clipping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode64.gif

A variable diode clipping or diode limiting level can be achieved by varying the bias voltage of the diodes. If both the positive and the negative half cycles are to be clipped, then two biased clipping diodes are used. But for both positive and negative diode clipping, the bias voltage need not be the same. The positive bias voltage could be at one level, for example 4 volts, and the negative bias voltage at another, for example 6 volts as shown.

### Diode Clipping of Different Bias levels

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode65.gif

When the voltage of the positive half cycle reaches +4.7 V, diode D1 conducts and limits the waveform at +4.7 V. Diode D2 does not conduct until the voltage reaches –6.7 V. Therefore, all positive voltages above +4.7 V and negative voltages below –6.7 V are automatically clipped.

The advantage of biased diode clipping circuits is that it prevents the output signal from exceeding preset voltage limits for both half cycles of the input waveform, which could be an input from a noisy sensor or the positive and negative supply rails of a power supply.

If the diode clipping levels are set too low or the input waveform is too great then the elimination of both waveform peaks could end up with a square-wave shaped waveform.

## Zener Diode Clipping Circuits

The use of a bias voltage means that the amount of the voltage waveform that is clipped off can be accurately controlled. But one of the main disadvantages of using voltage biased diode clipping circuits, is that they need an additional emf battery source which may or may not be a problem.

One easy way of creating biased diode clipping circuits without the need for an additional emf supply is to use Zener Diodes.

As we know, the zener diode is a another type of diode that has been specially manufactured to operate in its reverse biased breakdown region and as such can be used for voltage regulation or zener diode clipping applications. In the forward region, the zener acts just like an ordinary silicon diode with a forward voltage drop of 0.7V (700mV) when conducting, the same as above.

However, in the reverse bias region, the voltage is blocked until the zener diodes breakdown voltage is reached. At this point, the reverse current through the zener increases sharply but the zener voltage, VZ across the device remains constant even if the zener current, IZ varies.

Then we can put this zener action to good effect by using them for clipping a waveform as shown.

### Zener Diode Clipping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode66.gif

The zener diode is acting like a biased diode clipping circuit with the bias voltage being equal to the zener breakdown voltage. In this circuit during the positive half of the waveform the zener diode is reverse biased so the waveform is clipped at the zener voltage, VZD1. During the negative half cycle the zener acts like a normal diode with its usual 0.7V junction value.

We can develop this idea further by using the zener diodes reverse-voltage characteristics to clip both halves of a waveform using series connected back-to-back zener diodes as shown.

### Full-wave Zener Diode Clipping

### F:\BSc_CBCS_SYLLAB112Bsc_CBCS_SYLLABUS_\diode-diode67.gif

The output waveform from full wave zener diode clipping circuits resembles that of the previous voltage biased diode clipping circuit. The output waveform will be clipped at the zener voltage plus the 0.7V forward volt drop of the other diode. So for example, the positive half cycle will be clipped at the sum of zener diode, ZD1 plus 0.7V from ZD2 and vice versa for the negative half cycle.

Zener diodes are manufactured with a wide range of voltages and can be used to give different voltage references on each half cycle, the same as above. Zener diodes are available with zener breakdown voltages, VZ ranging from 2.4 to 33 volts, with a typical tolerance of 1 or 5%. Note that once conducting in the reverse breakdown region, full current will flow through the zener diode so a suitable current limiting resistor, R1 must be chosen.

## Diode Clipping Summary

As well as being used as rectifiers, diodes can also be used to clip the top, or bottom, or both of a waveform at a particular dc level and pass it to the output without distortion,. In or examples above we have assumed that the waveform is sinusoidal but in theory any shaped input waveform can be used.

**Diode Clipping Circuits** are used to eliminate amplitude noise or voltage spikes, voltage regulation or to produce new waveforms from an existing signal such as squaring off the peaks of a sinusoidal waveform to obtain a rectangular waveform as seen above.

The most common application of a “diode clipping” is as a flywheel or free-wheeling diode connected in parallel across an inductive load to protect the switching transistor form reverse voltage transients.