Analogue Electronics 1

Diploma In Electrical and Electronics, Module 1

Course Content

1. Atomic Theory of Matter
2. Thermionic Emission
3. Semiconductor Theory
4. Semiconductor Diodes
5. Bipolar Junction Transistors
6. Field Effect Transistor
7. Power Supply Units

Prepared by *Kelvin Mugweru*

Contents

Chapter 1: Atomic Theory of Matter

The theory explains several concepts that are relevant in the observable world: the composition of a pure gold necklace, what makes the pure gold necklace different than a pure silver necklace, and what occurs when pure gold is mixed with pure copper. This section explains the theories that Dalton used as a basis for his theory: (1) Law of Conservation of Mass (2) Law of Definite Proportions, and (3) Law of Multiple Proportions

Law 1: The Conservation of Mass

"Nothing comes from nothing" is an important idea in ancient Greek philosophy that argues that what exists **now** has always **existed**, since no new matter can come into existence where there was none before. Antoine Lavoisier (1743-1794) restated this principle for chemistry with the law of conservation of mass, which "means that the atoms of an object cannot be created or destroyed, but can be moved around and be changed into different particles." This law says that when a chemical reaction rearranges atoms into a new product, the mass of the reactants (chemicals before the chemical reaction) is the same as the mass of the products (the new chemicals made). More simply, whatever you do, you will still have the same amount of stuff (however, certain nuclear reactions like fusion and fission can convert a small part of the mass into energy.

The law of conservation of mass states that the total mass present before a chemical reaction is the same as the total mass present after the chemical reaction; in other words, **mass is conserved**. The law of conservation of mass was formulated by Lavoisier as a result of his combustion experiment, in which he observed that the mass of his original substance—a glass vessel, tin, and air—was equal to the mass of the produced substance—the glass vessel, “tin calx”, and the remaining air.

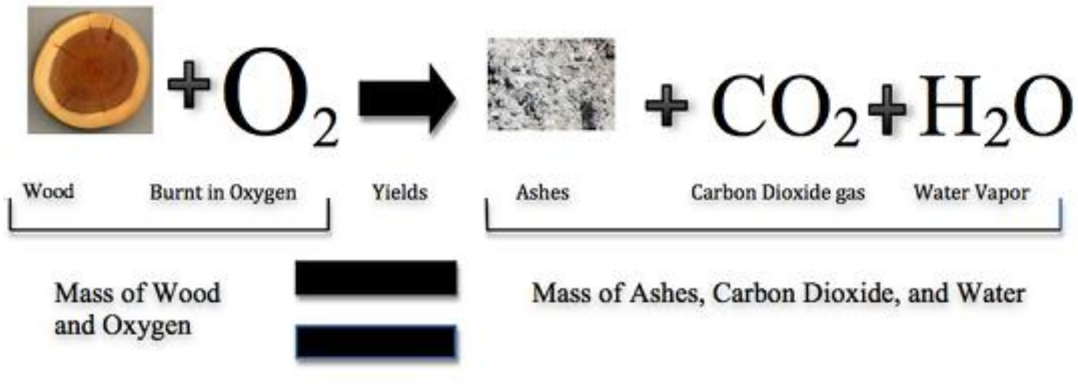


Figure 2.1.12.1.1: Image of the wood courtesy of Ehamberg and Stannered on Wikimedia Commons, available under Creative Commons Attribution 2.5 Generic license. Image of ashes courtesy of Walter Siegmund. Image as a whole constructed by Jessica Thornton (UCD).

Historically, this was a difficult concept for scientists to grasp. If this law was true, then how could a large piece of wood be reduced to a small pile of ashes? The wood clearly has a greater mass than the ashes. From this observation scientists concluded that mass had been lost. However, Figure 2.1.12.1.1 shows that the burning of word does follow the law of conservation of mass. Scientists did not account for the gases that play a critical role in this reaction.

The law of conservation of mass states that the total mass present before a chemical reaction is the same as the total mass present after the chemical reaction.

Law 2: Definite Proportions

Joseph Proust (1754-1826) formulated the *law of definite proportions* (also called the **Law of Constant *Composition*** or **Proust's Law**). This law states that if a compound is broken down into its constituent elements, the masses of the constituents will always have the same proportions, regardless of the quantity or source of the original substance. Joseph Proust based this law primarily on his experiments with basic copper carbonate. The illustration below depicts this law in action.

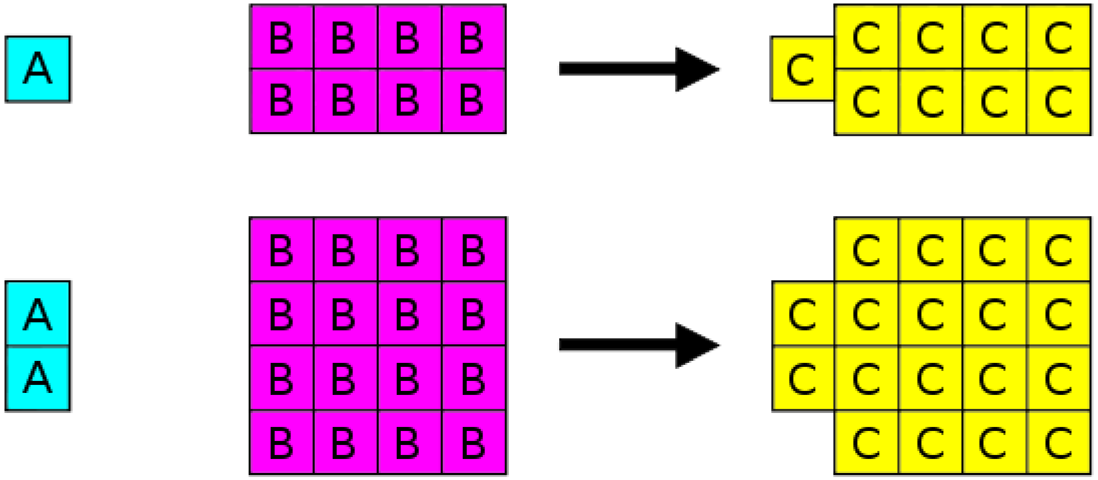


Figure 2.1.22.1.2: If 1 gram of A reacts with 8 grams of B, then by the Law of Definite Proportions, 2 grams of A must react with 16 grams of B. If 1 gram of A reacts with 8 grams of B, then by the Law of Conservation of Mass, they must produce 9 grams of C. Similarly, when 2 grams of A react with 16 grams of B, they must produce 18 grams of C.

Law of Definite Proportions states that in a given type of chemical substance, the elements are always combined in the same proportions by mass.

The Law of Definite Proportions applies when elements are reacted together to form *the same* product. Therefore, while the Law of Definite Proportions can be used to compare two experiments in which hydrogen and oxygen react to form water, the Law of Definite Proportions can *not* be used to compare one experiment in which hydrogen and oxygen react to form water, and another experiment in which hydrogen and oxygen react to form hydrogen peroxide (peroxide is another material that can be made from hydrogen and oxygen).

Example 2.1.12.1.1: water

Oxygen makes up 88.8% of the mass of any sample of pure water, while hydrogen makes up the remaining 11.2% of the mass. You can get water by melting ice or snow, by condensing steam, from river, sea, pond, etc. It can be from different places: USA, UK, Australia, or anywhere. It can be made by chemical reactions like burning hydrogen in oxygen.

However, if the water is **pure**, it will **always** consist of 88.8 % oxygen by mass and 11.2 % hydrogen by mass, irrespective of its source or method of preparation.

Law 3: Multiple Proportions

Many combinations of elements can react to form more than one compound. In such cases, this law states that the weights of one element that combine with a fixed weight of another of these elements are integer multiples of one another. It's easy to say this, but please make sure that you understand how it works. Nitrogen forms a very large number of oxides, five of which are shown here.

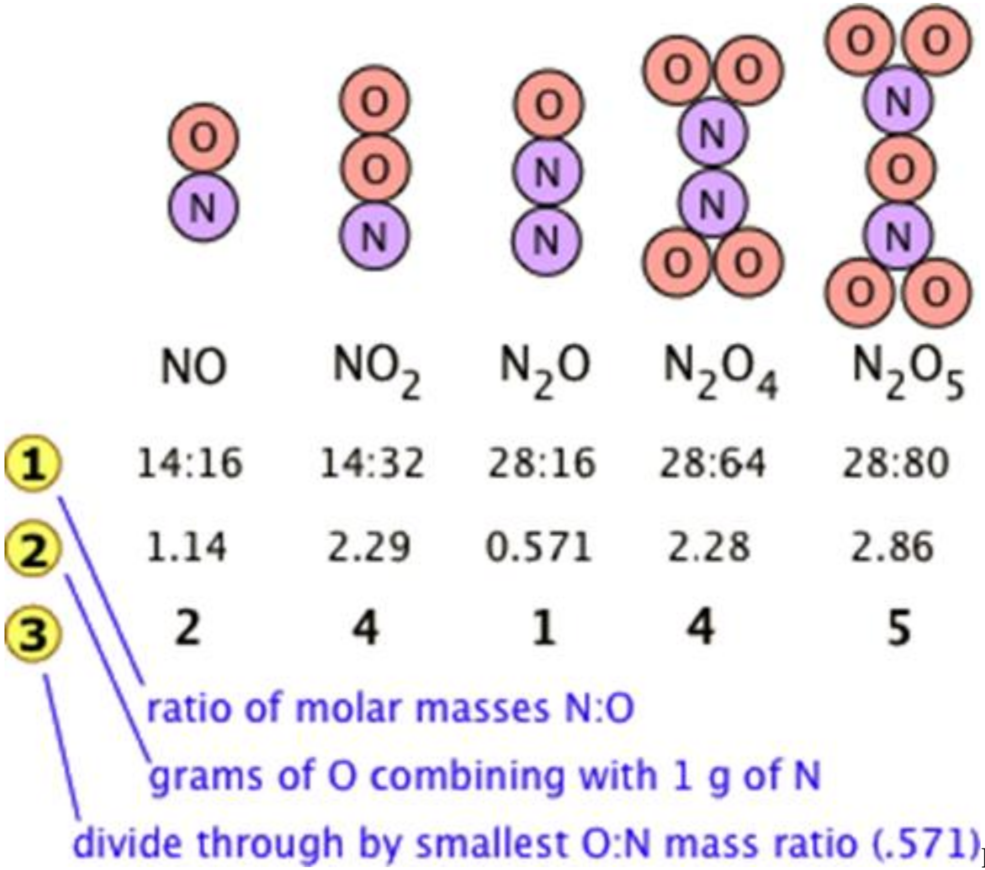


Figure 2.1.32.1.3: Law of Multiple Proportions applied to nitrogen oxides (NO_xNOx) compounds. (CC-BY; Stephen Lower)

- Line 1 shows the ratio of the relative weights of the two elements in each compound. These ratios were calculated by simply taking the molar mass of each element, and multiplying by the number of atoms of that element per mole of the compound. Thus for NO_2, we have (1 × 14) : (2 × 16) = 13:32. (These numbers were not known in the early days of Chemistry because atomic weights (i.e., molar masses) of most elements were not reliably known.)
- The numbers in Line 2 are just the mass ratios of O:N, found by dividing the corresponding ratios in line 1. But someone who depends solely on experiment would work these out by finding the mass of O that combines with unit mass (1 g) of nitrogen.
- Line is obtained by dividing the figures the previous line by the smallest O:N ratio in the line above, which is the one for N_2O. Note that just as the law of multiple proportions says, the weight of oxygen that combines with unit weight of nitrogen work out to small integers.
- Of course we just as easily could have illustrated the law by considering the mass of nitrogen that combines with one gram of oxygen; it works both ways!

The law of multiple proportions states that if two elements form more than one compound between them, the masses of one element combined with a fixed mass of the second element form in ratios of small integers.

Example 2.1.22.1.2: Oxides of Carbon

Consider two separate compounds are formed by only carbon and oxygen. The first compound contains 42.9% carbon and 57.1% oxygen (by mass) and the second compound contains 27.3% carbon and 72.7% oxygen (again by mass). Is this consistent with the law of multiple proportions?

Solution

The *Law of Multiple Proportions* states that the masses of one element which combine with a fixed mass of the second element are in a ratio of **whole** numbers. Hence, the masses of oxygen in the two compounds that combine with a fixed mass of carbon should be in a whole-number ratio.

Thus for every 1 g of the first compound there are 0.57 g of oxygen and 0.429 g of carbon. The mass of oxygen per gram carbon is:

$$\frac{0.571\text{g oxygen}}{0.429\text{g carbon}} = 1.33 \frac{\text{g oxygen}}{\text{g carbon}}$$

Similarly, for 1 g of the second compound, there are 0.727 g oxygen and 0.273 g of carbon. The ration of mass of oxygen per gram of carbon is

$$\frac{0.727\text{g oxygen}}{0.273\text{g carbon}} = 2.66 \frac{\text{g oxygen}}{\text{g carbon}}$$

Dividing the mass of oxygen per g of carbon of the second compound:

$$\frac{2.66}{1.33} = 2$$

Hence the masses of oxygen combine with carbon in a 2:1 ratio which s consistent with the Law of Multiple Proportions since they are whole numbers.

Dalton's Atomic Theory

The modern atomic theory, proposed about 1803 by the English chemist John Dalton (Figure 2.1.42.1.4), is a fundamental concept that states that all elements are composed of atoms. Previously, an atom was defined as the smallest part of an element that maintains the identity of that element. Individual atoms are extremely small; even the largest atom has an approximate diameter of only 5.4×10^{-10} m. With that size, it takes over 18 million of these atoms, lined up side by side, to equal the width of the human pinkie (about 1 cm).

Dalton's ideas are called the *modern* atomic theory because the concept of atoms is very old. The Greek philosophers Leucippus and Democritus originally introduced atomic concepts in the fifth century BC. (The word *atom* comes from the Greek word *atomos*, which means "indivisible" or "uncuttable.") Dalton had something that the ancient Greek philosophers didn't have, however; he had experimental evidence, such as the formulas of simple chemicals and the behavior of gases. In the 150 years or so before Dalton, natural philosophy had been maturing into modern science, and the scientific method was being used to study nature. When Dalton announced a modern atomic theory, he was proposing a fundamental theory to describe many previous observations of the natural world; he was not just participating in a philosophical discussion.

Dalton's Theory was a powerful development as it explained the three laws of chemical combination (above) and recognized a workable distinction between the fundamental particle of an element (atom) and that of a compound (molecule). Six postulates are involved in Dalton's Atomic Theory:

1. All matter consists of indivisible particles called atoms.
2. Atoms of the same element are similar in shape and mass, but differ from the atoms of other elements.
3. Atoms cannot be created or destroyed.
4. Atoms of different elements may combine with each other in a fixed, simple, whole number ratios to form compound atoms.

5. Atoms of same element can combine in more than one ratio to form two or more compounds.
6. The atom is the smallest unit of matter that can take part in a chemical reaction.

In light of the current state of knowledge in the field of Chemistry, Dalton's theory had a few drawbacks. According to Dalton's postulates,

1. The indivisibility of an atom was proved wrong: an atom can be further subdivided into protons, neutrons and electrons. However an atom is the smallest particle that takes part in chemical reactions.
2. According to Dalton, the atoms of same element are similar in all respects. However, atoms of some elements vary in their masses and densities. These atoms of different masses are called isotopes. For example, chlorine has two isotopes with mass numbers 35 and 37.
3. Dalton also claimed that atoms of different elements are different in all respects. This has been proven wrong in certain cases: argon and calcium atoms each have an same atomic mass (40 amu).
4. According to Dalton, atoms of different elements combine in simple whole number ratios to form compounds. This is not observed in complex organic compounds like sugar ($C_{12}H_{22}O_{11}$C12H22O11).
5. The theory fails to explain the existence of allotropes (different forms of pure elements); it does not account for differences in properties of charcoal, graphite, diamond.

Chapter 2: Thermionic emission

Thermionic emission definition

The process by which free electrons are emitted from the surface of a metal when external heat energy is applied is called thermionic emission.

Thermionic emission occurs in metals that are heated to a very high temperature. In other words, thermionic emission occurs, when large amount of external energy in the form of heat is supplied to the free electrons in the metals.

Metals under normal temperature

When a small amount of heat energy is applied to the metal, the valence electrons gain enough energy and break the bonding with the parent atom. The valence electron, which breaks the bonding with the parent atom, becomes free. This electron, which breaks the bonding with the parent atom, is called as the free electron.

The free electrons in the metal have some kinetic energy. However, they do not have enough energy to escape from the metal. The attractive force of the atomic nuclei opposes the free electrons, which try to escape from the metal.

Free electrons in the metal have less energy compared to the free electrons in vacuum. Hence, free electrons require extra energy from the outside source in order to jump into the vacuum.

Metals under high temperature

When heat energy applied to the metal is increased to a higher value, the free electrons gain enough energy and overcome the attractive force of the atomic nucleus, which holds the free electrons in the metal. The free electrons, which overcome the attractive force of the nuclei, break the bonding with the metal and jumps into the vacuum.

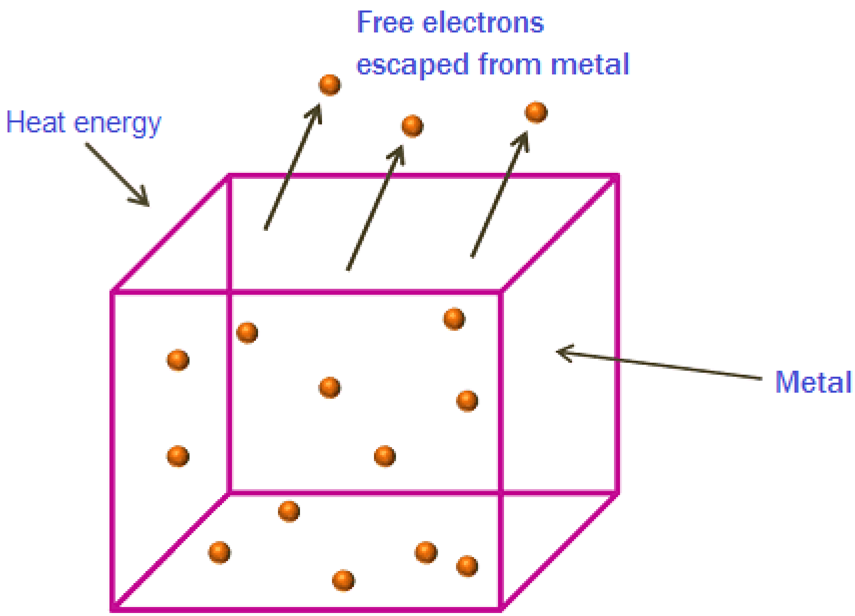


Fig: Electron emission

The free electrons, which are escaped from the surface of a metal when heat energy is supplied, are called thermions. Thermionic emission process plays a major role in the operation of electronic devices.

Thermionic emission depends on heat applied to the metal and work function of the metal

The number of free electrons escaped from the metal is depends on the amount of heat applied to the metal and the work function of the metal.

Heat applied to the metal:

If large amount of heat is applied to the metal, large number of free electrons gains enough energy and breaks the bonding with the metal. The free electron, which breaks the bonding with the metal, jumps into the vacuum.

On the other hand, if less amount of heat is applied to the metal, less number of free electrons gains enough energy and breaks the bonding with the metal. The free electron, which breaks the bonding with the metal will jumps into the vacuum.

Hence, the number of free electrons emitted from the metal increases with increase in heat. Thus, the free electrons emitted from the surface of metals are directly proportional to the temperature of the metals.

The minimum temperature at which the metal starts emitting the free electrons is called threshold temperature.

If the temperature of the metal is below the threshold temperature, the metal does not emit the free electrons. On the other hand, if the temperature of the metal is equal to the threshold temperature or greater than the threshold temperature, the metal emits the free electrons.

Work function of the metal:

The amount of external heat energy required to remove the free electron from the metal is called work function or threshold energy. The work function of metals is measured in electron volts (eV).

Metals that have low work function will require less amount of heat energy to cause the free electrons to escape from the metal. Hence, the metals with low work function emit large number of free electrons at high temperature.

On the other hand, metals that have high work function will require more amount of heat energy to cause the free electrons to escape from the metal. Hence, the metals with high work function emit less number of free electrons at high temperature.

Thus, the emission of free electrons from the metal is inversely proportional to the work function of a metal.

Applications of thermionic emission

The components, which are made by the process of thermionic emission are used in the electronic devices such as cathode ray tube, radio etc.

Rate of Thermionic Emission

The number of thermions emitted per second from a substance is known as the rate of thermionic emission. This value depends on the:

1. Nature of the Material
   In general, every element can be characterized by its electronic configuration i.e. by the distribution of electrons surrounding its nucleus. When we speak of thermionic emission, our particular interest is in the valence electrons (electrons in the outermost shell). This is because these are the electrons which can be easily freed from the force of attraction so as to enable conduction. However, the energy which must be supplied differs from element to element and is regarded to be its threshold energy or work function.
2. Surface Temperature
   Higher is the temperature of the substance, greater is the rate of thermionic emission.
3. Surface Area
   If the surface area of the material considered is larger, then there will be a greater number of thermions emitted. This means that the rate of thermionic emission is directly proportional to the surface area of the material.

By analyzing these factors, it can be concluded that the substance chosen to be a thermionic emitter should have a low work function, larger surface area, and high melting point. A few examples of this kind are metals like tungsten, thoriated tungsten, tantalum, etc and coated metals like barium oxide, strontium oxide, etc.

Thermionic Current

The flow of thermions gives rise to the flow of current known as thermionic current. Mathematically the thermionic equation which gives the current density of electrons is expressed as:

Where:

- T is the absolute temperature,
- k_B is the Boltzmann Constant,
- Φ_W is the work function,
- e is the electron charge
- A is a constant.

Applications of Thermionic Emission

Thermionic emission forms the basic principle on which many of the devices used in the field of electronics and communication operates. Example applications of thermionic emission include vacuum tubes, diode valves, cathode ray tube, electron tubes, electron microscopes, X-ray tubes, thermionic converters, and electrodynamic tethers.

Thermionic Emitter

The metallic structure used to facilitate thermionic emission is called **thermionic emitter**. The emitter is also called cathode. The emitter or cathode is sufficiently heated in vacuum or evacuated space to initiate thermionic emission i.e. emission of electrons from the body of emitter or cathode. The metal or metallic substances used to construct a thermionic emitter should have three main features:

1. It should have a low work function. A low work function helps to emit electrons from the cathode surface in a comparatively lower temperature.
2. It should have a high melting point. The temperature required to emit an electron from the cathode surface is quite high compared to the melting point of normal metals. Some of the common metals have low work function but till they are not suitable for constructing a thermionic emitter. This is because the lower melting point causes vaporization of metals before they emit electrons. For example, copper has low work function but we can not use it as a **thermionic emitter**, because it’s melting point is only 810°C. So at thermionic emission temperature, the copper gets vaporized instead of emitting electrons from its solid surface.
3. It should have high mechanical strength. The absolute vacuum cannot be created in space surrounding the cathode, so there may always be some gaseous molecules present in the space. After a collision with emitted electrons from the cathode, these gaseous molecules produce positive ions in the space. Due to electrostatic farce, these positive ions strick the cathode. It sufficiently high electric field is applied, these bombardments may be significantly high to create damage on the cathode. To avoid the damage of the cathode due to ions collisions, the mechanical strength of the materials used for constructing cathode must be high enough. Considering the above-mentioned properties, we normally use, tungsten, thoriated tungsten, oxide-coated metals for constructing cathode of thermionic emission.

Tungsten

- Work function = 4.52 eV
- Melting point = 3650°K
- Tensile strength = 100000 – 500000 psi @ room temperature
- The Thermionic emission temperature = 2327°C
- Emission efficiency 4 mA/watt

Tungsten was previously used as the materials for the thermionic emitter. It has high work function but still, it was used as the cathode because of its high melting point and the material is mechanically very strong. Due to work function, the operating temperature of tungsten cathode is high and at the same time, the emission efficiency is low since for maintaining the high temperature of the cathode the input energy to the system is high compared to emitted current from the cathode.

Thoriated Tungsten

Sometimes the addition of one metal to others makes the work function of mixture lower. Thoriated Tungsten is a mixture of thorium and tungsten. Thorium has work function 3.4 eV and tungsten has work function 4.52 eV. When a small quantity of thorium is mixed with tungsten to make thoriated tungsten, the work function comes down to 2.63 eV. This causes the operating temperature of thermionic emission at 1700°C when the cathode is made of thoriated tungsten. So, the power input for heating the cathode elements is reduced hence, the emission efficiency is increased accordingly.

Oxide Coated Cathode

Here, the cathode for thermionic emission is made of nickel ribbon coated with barium and strontium oxide. The oxide coating reduces the work function of the system to a quite low value. It is about 1.1 eV. Low work function causes low operating temperature and high emission efficiency of the system. The operating temperature and thermionic emission efficiency of the system are 750°C and 200 mA/watt respectively.

Construction of Cathode for Thermionic Emission

The cathode or thermionic emitter is placed inside a vacuum container. So, only possible way to heat up the cathode is electrical heating. There are two types of electric heating used in thermionic emission, one is direct heating and other is indirect heating.

Directly Heated Cathode

In a directly heated cathode, the cathode is made of in the form of a filament. The filament is normally made of oxide-coated nickel. When the current from the input source passes directly through the filament, it gets hot and emits electrons.

A direct heating method is more efficient as the input current (input energy) directly heats the filament cathode to emit electrons. As the heating is quick, starting time of thermionic emission quick and at the same time, it is an efficient process. As the emitter is directly heated, any of fluctuation in input source will affect the emission. This is the main disadvantage of directly heated cathode thermionic emission.

Indirectly Heated Cathode

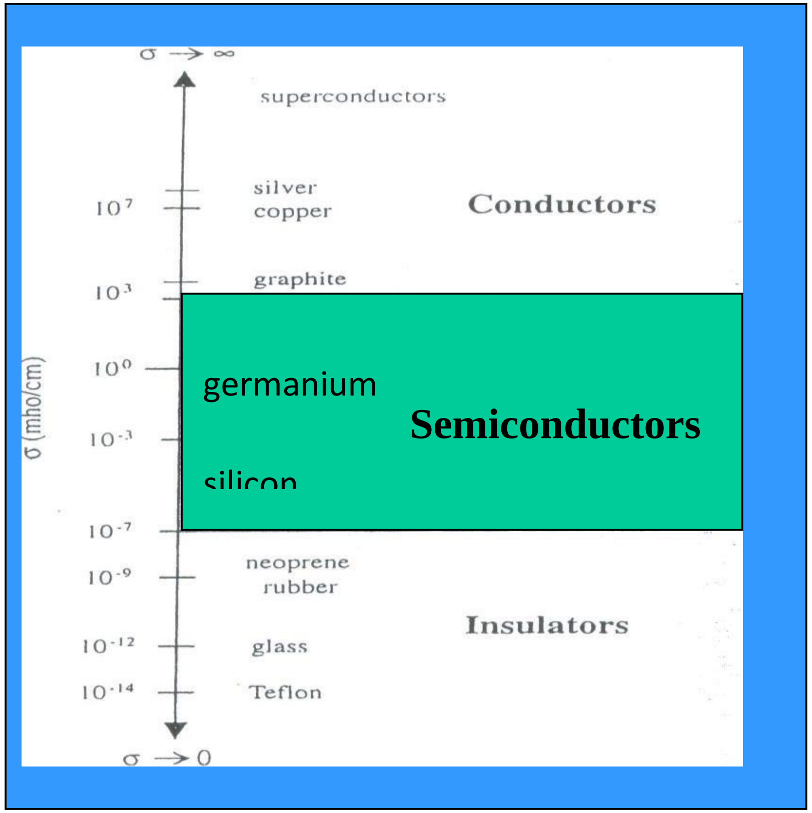
Here, the heating filament and emitting surface are separate and they are insulated to each other. The filament is surrounded by thin oxide coated metal sleeve. The input current possesses through the heating filament and hence it heats up the metallic sleeve from where electrons are emitted. Most modern thermionic emitters are indirectly heated cathode this is because of the following facts.

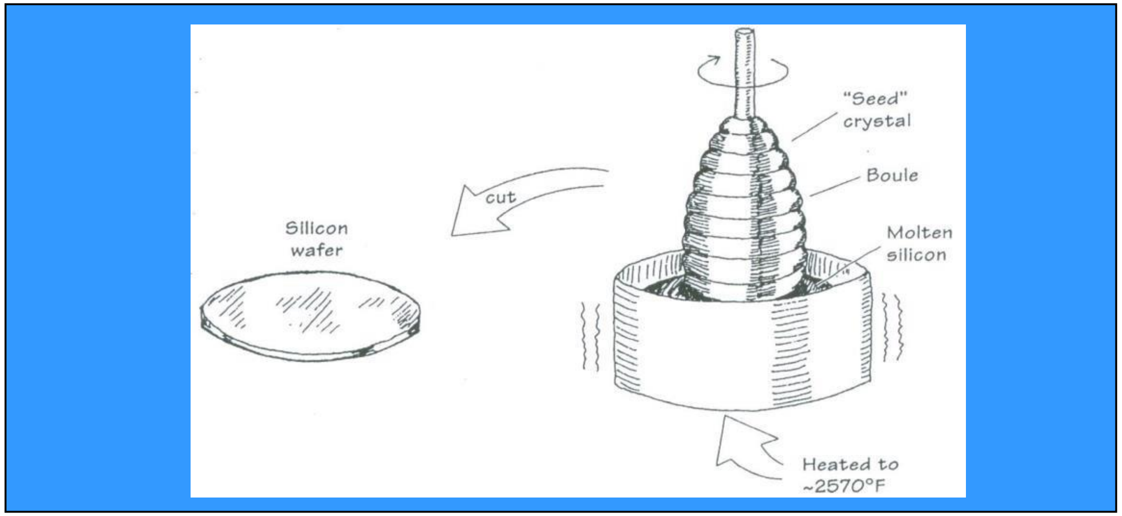
1. The emission potential and heating potential are separate. The emitter can be connected to any required potential irrespective of the heating potential.
2. The fluctuations in input heating potential don’t affect the emission.
3. Alternating current can also be used as a heating current of the system.

Chapter 3: Semiconductor Theory

Classification of Materials

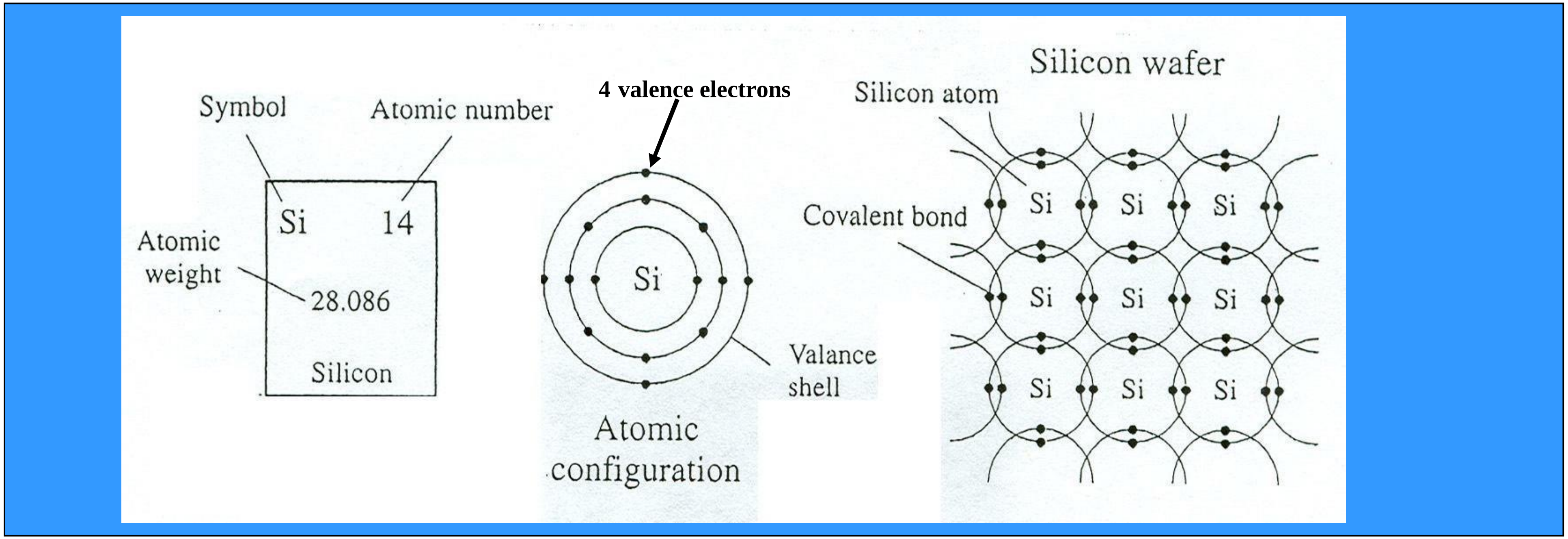
- Materials that permit flow of electrons are called conductors (e.g., gold, silver, copper, etc.).
- Materials that block flow of electrons are called insulators (e.g., rubber, glass, Teflon, mica, etc.).
- Materials whose conductivity falls between those of conductors and insulators are called semiconductors.
- Semiconductors are "part-time" conductors whose conductivity can be controlled.
- Silicon is the most common material used to build semiconductor devices.
- Si is the main ingredient of sand and it is estimated that a cubic mile of seawater contains 15,000 tons of Si.
- Si is spun and grown into a crystalline structure and cut into wafers to make electronic devices.





Atomic Theory

- Atoms in a pure silicon wafer contains four electrons in outer orbit (called valence electrons). — Germanium is another semiconductor material with four valence electrons.
- In the crystalline lattice structure of Si, the valence electrons of every Si atom are locked up in covalent bonds with the valence electrons of four neighboring Si atoms.
  - In pure form, Si wafer does not contain any free charge carriers.
  - An applied voltage across pure Si wafer does not yield electron flow through the wafer.
  - A pure Si wafer is said to act as an insulator.
- In order to make useful semiconductor devices, materials such as phosphorus (P) and boron (B) are added to Si to change Si's conductivity.

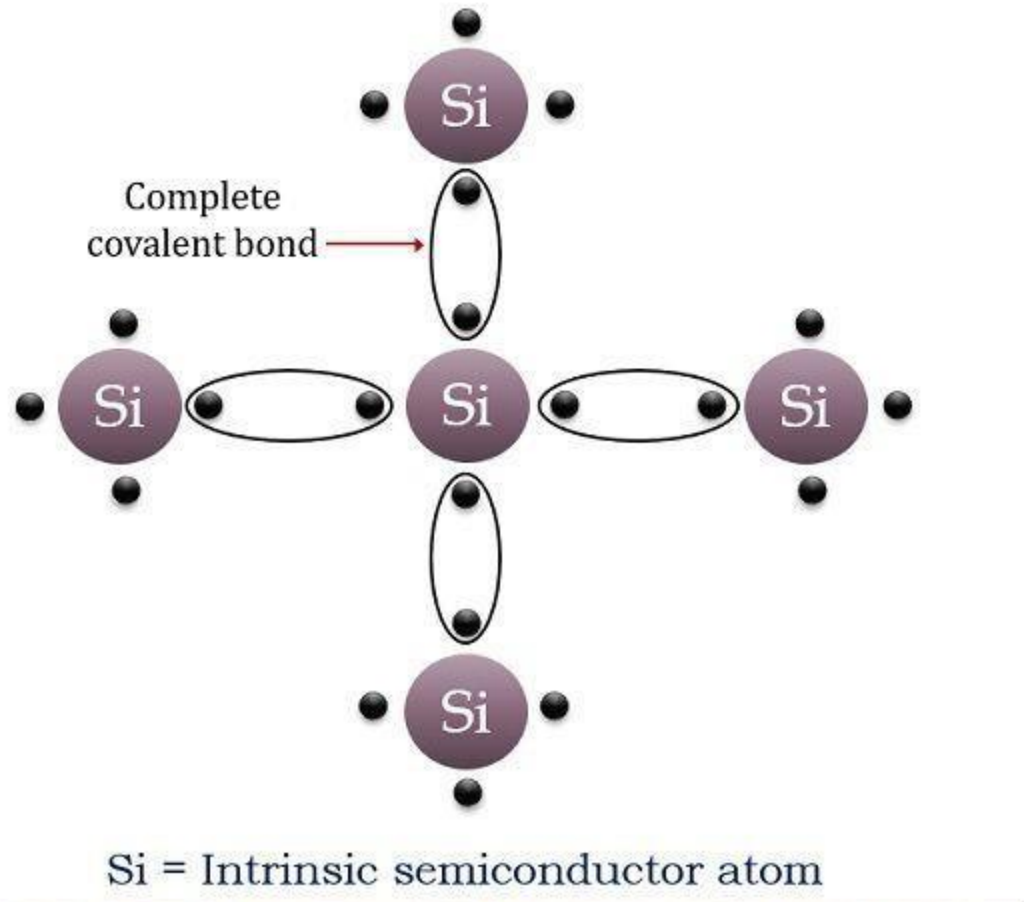


Intrinsic semiconductor

An intrinsic semiconductor is formed from a highly pure semiconductor material thus also known as pure semiconductors. These are basically undoped semiconductors that do not have doped impurity in it.

At room temperature, intrinsic semiconductors exhibit almost negligible conductivity. As no any other type of element is present in its crystalline structure. The group IV elements of the periodic table form an intrinsic semiconductor. However, mainly silicon and germanium are widely used. This is so because in their case only small energy is needed in order to break the covalent bond.

The figure below shows the crystalline structure of silicon:



Crystalline structure of Intrinsic semiconductor

The figure above clearly shows that silicon consists of 4 electrons in the valence shell. Here, 4 covalent bonds are formed between the electrons of the silicon atom. When the temperature of the crystal is increased then the electrons in the covalent bond gain kinetic energy and after breaking the covalent bond it gets free. Thus, the movement of free electrons generates current.

The rise in temperature somewhat increases the number for free electrons for conduction.

Extrinsic semiconductor

Extrinsic Semiconductors are those that are the result of adding an impurity to a pure semiconductor. These are basically termed as an impure form of semiconductors.

The process by which certain amount of impurity is provided to a pure semiconductor is known as **doping**. So, we can say a pure semiconductor is doped to generate an extrinsic semiconductor.

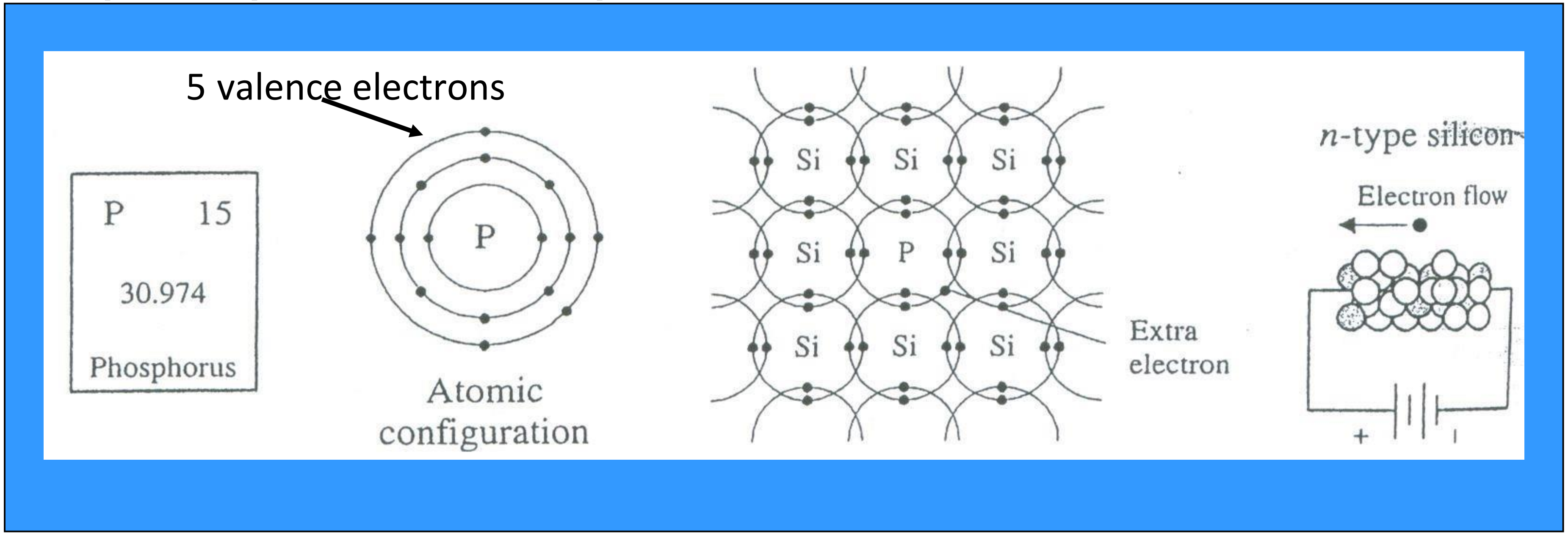
These are highly conductive in nature. However, unlike intrinsic semiconductor, extrinsic semiconductors are of two types p-type and an n-type semiconductor.

It is noteworthy here that the classification of the extrinsic semiconductor depends on the type of element doped to the pure semiconductor.

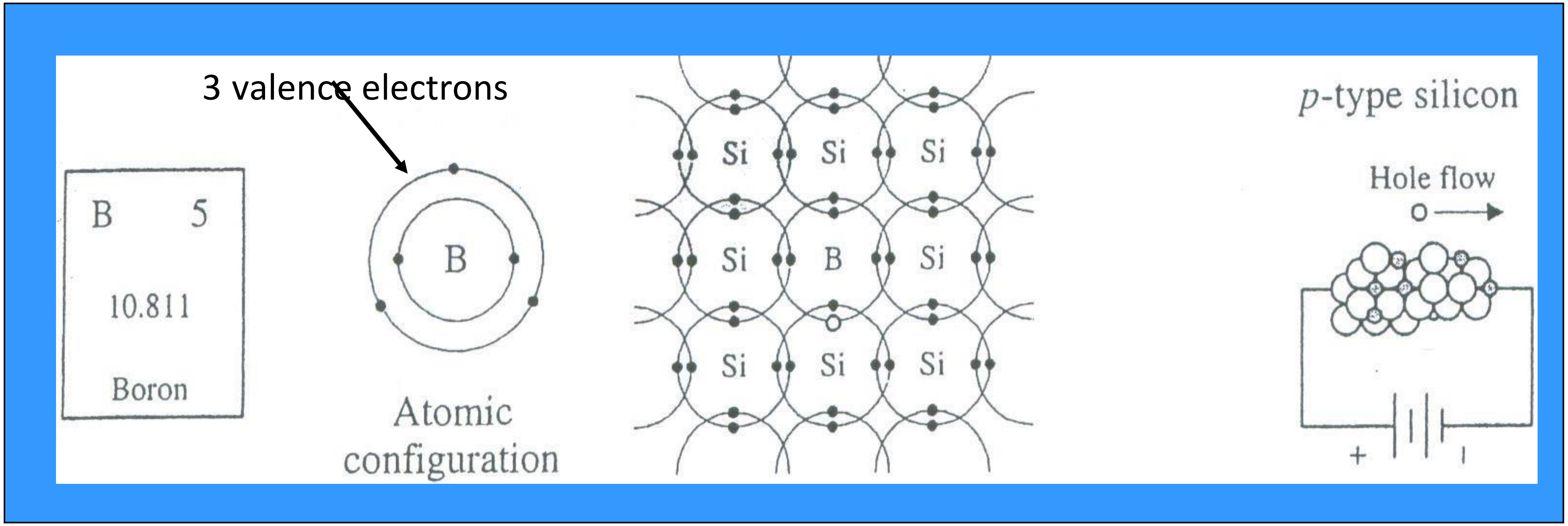
P-type semiconductors are formed by introducing group III elements or trivalent impurity into the pure semiconductor. These are also known as an acceptor impurity, as a trivalent impurity has only 3 electrons in the valence shell.

N-type semiconductors are formed by the addition of group V elements or pentavalent impurity to a pure semiconductor. These are termed as donor impurity, as a pentavalent impurity holds 5 electrons in its valence shell.

- **N-Type Silicon**
- Pentavalent impurities such as phosphorus, arsenic, antimony, and bismuth have 5 valence electrons.
- When phosphorus impurity is added to Si, every phosphorus atom's four valence electrons are locked up in covalent bond with valence electrons of four neighboring Si atoms. However, the 5^{th} valence electron of phosphorus atom does not find a binding electron and thus remains free to float. When a voltage is applied across the silicon-phosphorus mixture, free electrons migrate toward the positive voltage end.
- When phosphorus is added to Si to yield the above effect, we say that Si is doped with phosphorus. The resulting mixture is called N-type silicon (N: negative charge carrier silicon).
- The pentavalent impurities are referred to as donor impurities.



- **P-Type Silicon**
- Trivalent impurities e.g., boron, aluminum, indium, and gallium have 3 valence electrons.
- When boron is added to Si, every boron atom's three valence electrons are locked up in covalent bond with valence electrons of three neighboring Si atoms. However, a vacant spot "hole" is created within the covalent bond between one boron atom and a neighboring Si atom. The holes are considered to be positive charge carriers. When a voltage is applied across the silicon-boron mixture, a hole moves toward the negative voltage end while a neighboring electron fills in its place.
- When boron is added to Si to yield the above effect, we say that Si is doped with boron. The resulting mixture is called P-type silicon (P: positive charge carrier silicon).
- The trivalent impurities are referred to as acceptor impurities.



- The hole of boron atom points towards the negative terminal.
- The electron of neighboring silicon atom points toward positive terminal.
- The electron from neighboring silicon atom falls into the boron atom filling the hole in boron atom and creating a "new" hole in the silicon atom. • It appears as though a hole moves toward the negative terminal!

Chapter 4: Semiconductor Diodes

The PN junction Diode

- **Diode**

•A diode is a 2-lead semiconductor that acts as a one-way gate to electron flow. –
Diode allows current to pass in only one direction.

•A pn-junction diode is formed by joining together n-type and p-type silicon.

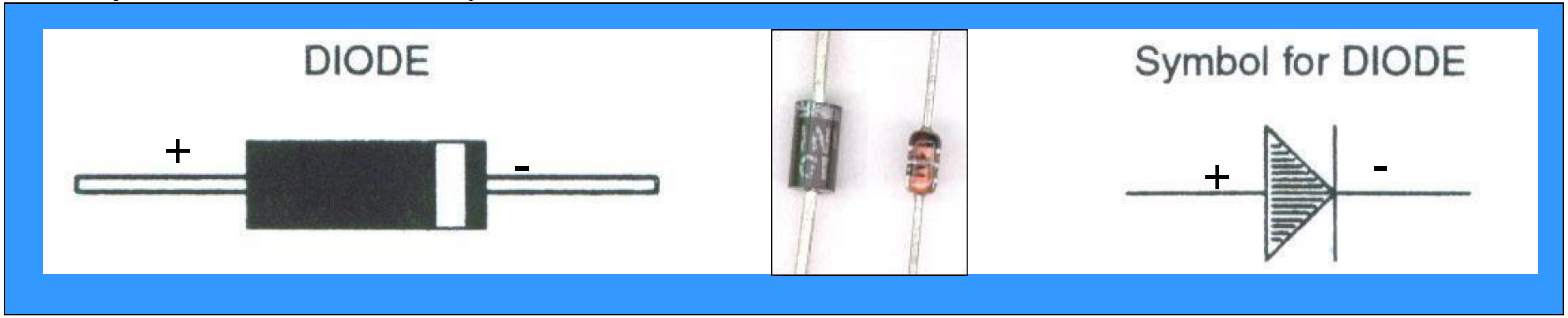
•In practice, as the n-type Si crystal is being grown, the process is abruptly altered to grow p-type Si crystal. Finally, a glass or plastic coating is placed around the joined crystal.

•The p-side is called anode and the n-side is called cathode.

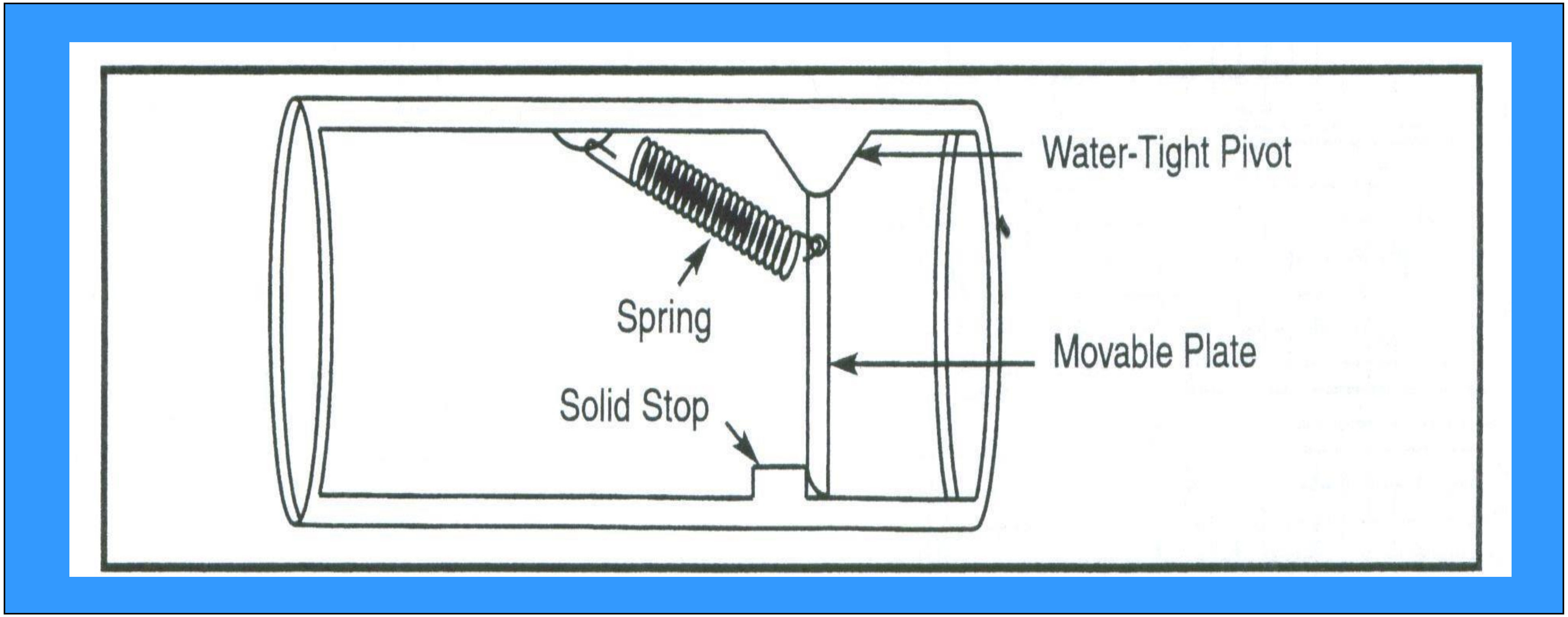
•When the anode and cathode of a pn-junction diode are connected to external voltage such that the potential at anode is higher than the potential at cathode, the diode is said to be forward biased.

–In a forward-biased diode current is allowed to flow through the device.

•When potential at anode is smaller than the potential at cathode, the diode is said to be reverse biased. In a reverse-biased diode current is blocked.



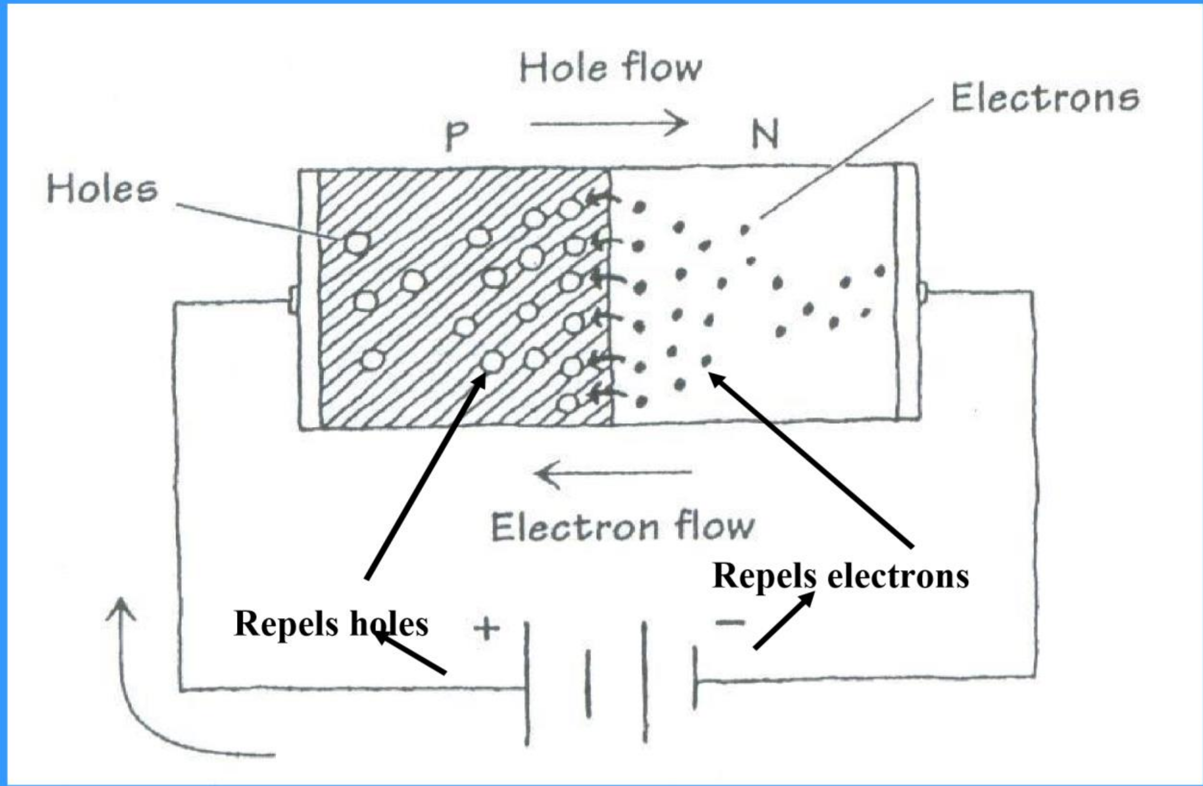
Water Analogy of Diodes



- When water pressure on left overcomes the restoring force of spring, the gate is opened and water is allowed to flow.
- When water pressure is from right to left, the gate is pressed against the solid stop and no water is allowed to flow.
- Spring restoring force is analogous to 0.6V needed to forward bias a Si diode.

- **Diode: How it Works**

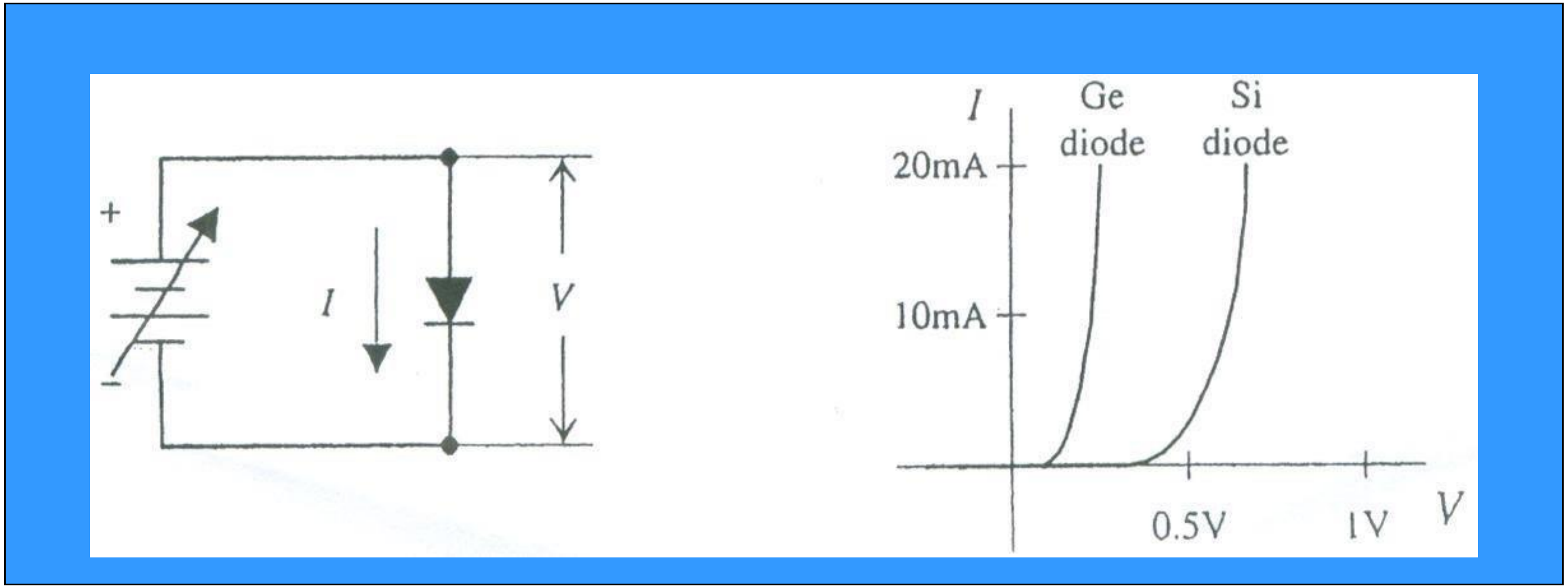
- When a diode is connected to a battery as shown, electrons from the n-side and holes from the p -side are forced toward the center by the electrical field supplied by the battery. The electrons and holes combine causing the current to pass through the diode. When a diode is arranged in this way, it is said to be <u>forward</u> <u>-</u> <u>biased</u> .



Forward -biased (“open door ”)

- A diode’s one-way gate feature does not work all the time.

- Typically for silicon diodes, an applied voltage of 0.6V or greater is needed, otherwise, the diode will not conduct.
- This feature is useful in forming a voltage-sensitive switch.
- I-V characteristics for silicon and germanium diodes is shown below.



- **Caution: How a Diode does not work**

- When a pn-diode is connected to a battery as shown, holes in the nside are forced to the left while electrons in the p-side are forced to the right. This results in an empty zone around the pn-junction that is free of charge carries creating a ***depletion region***. This depletion region acts as an insulator preventing current from flowing through the diode. When a diode is arranged in this way, it is said to be reversebiased.
- **Diode Applications —Half Wave Rectifier**

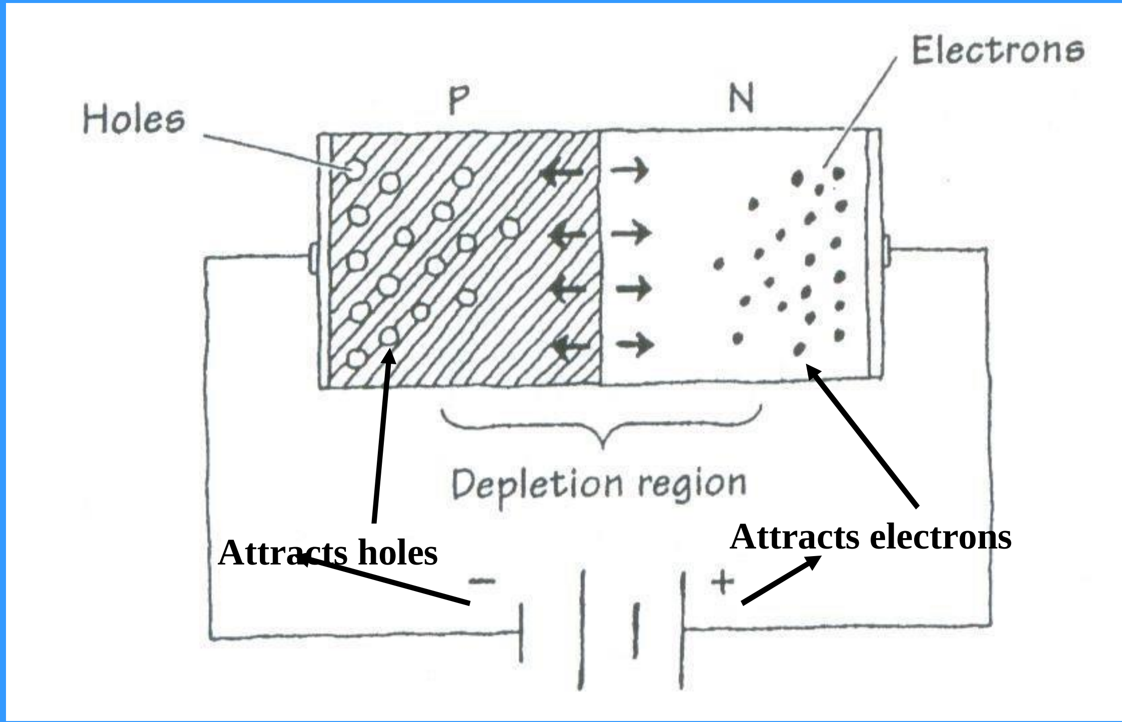
•Diode converts ac input voltage to a pulsed dc output voltage.

•Whenever the ac input becomes negative at diode's anode, the diode blocks current flow.

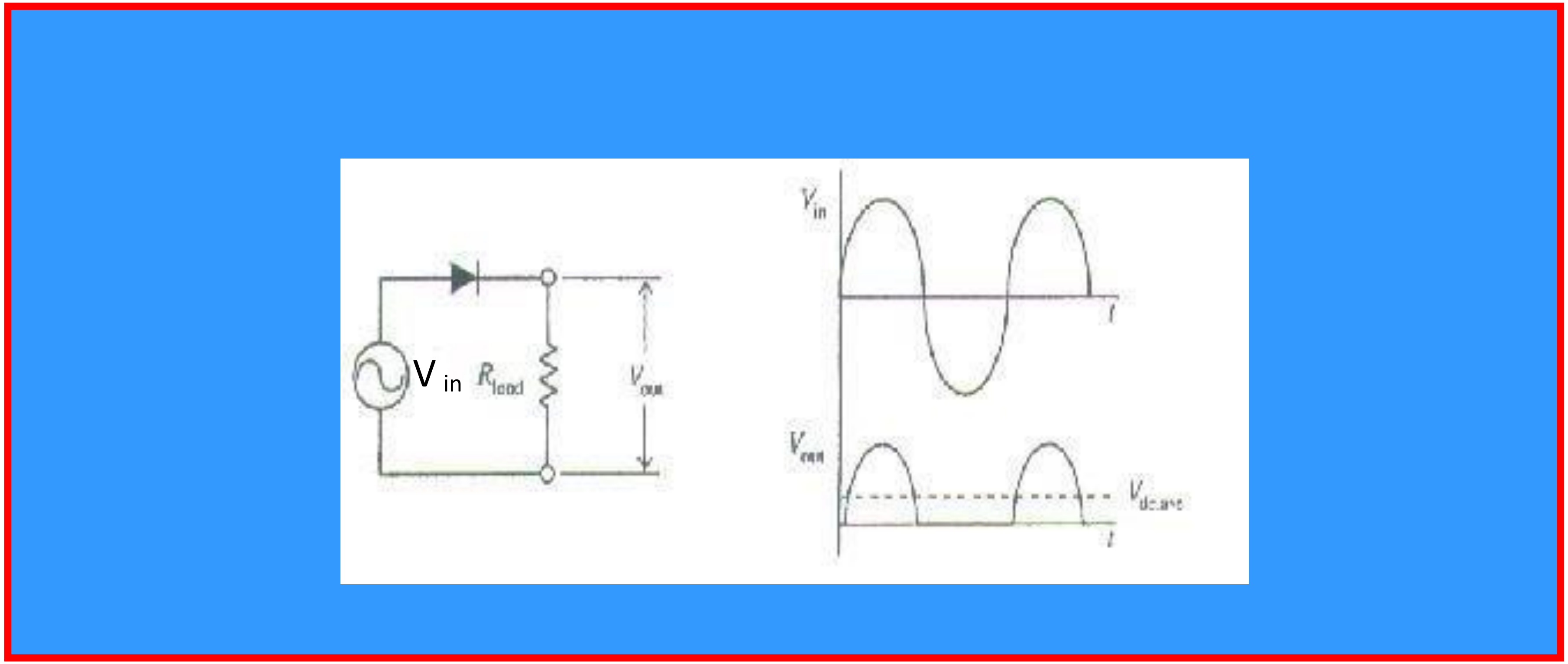
–□ o/p voltage become zero.

•Diode introduces a 0.6V drop so o/p peak is 0.6V smaller than the i/p peak.

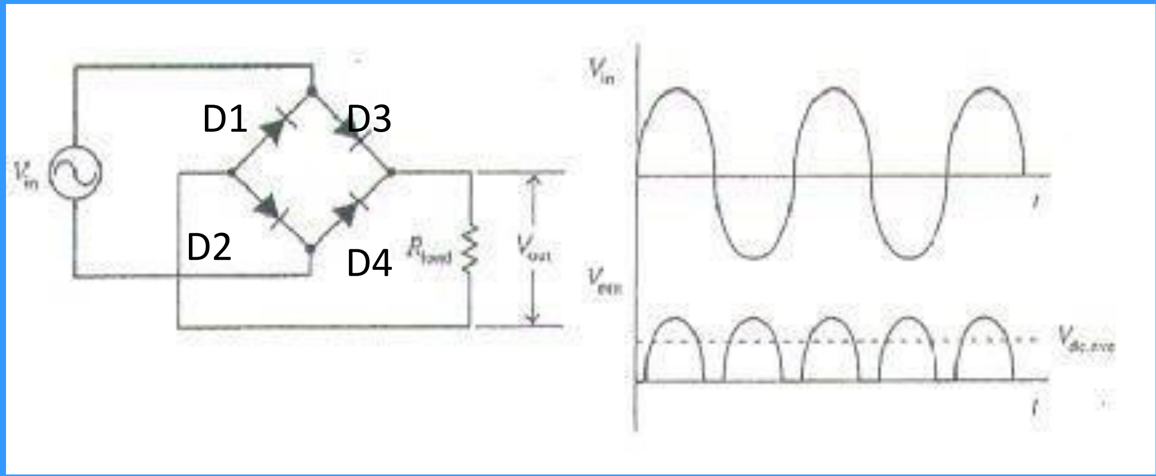
•The o/p frequency is same as the i/p frequency.



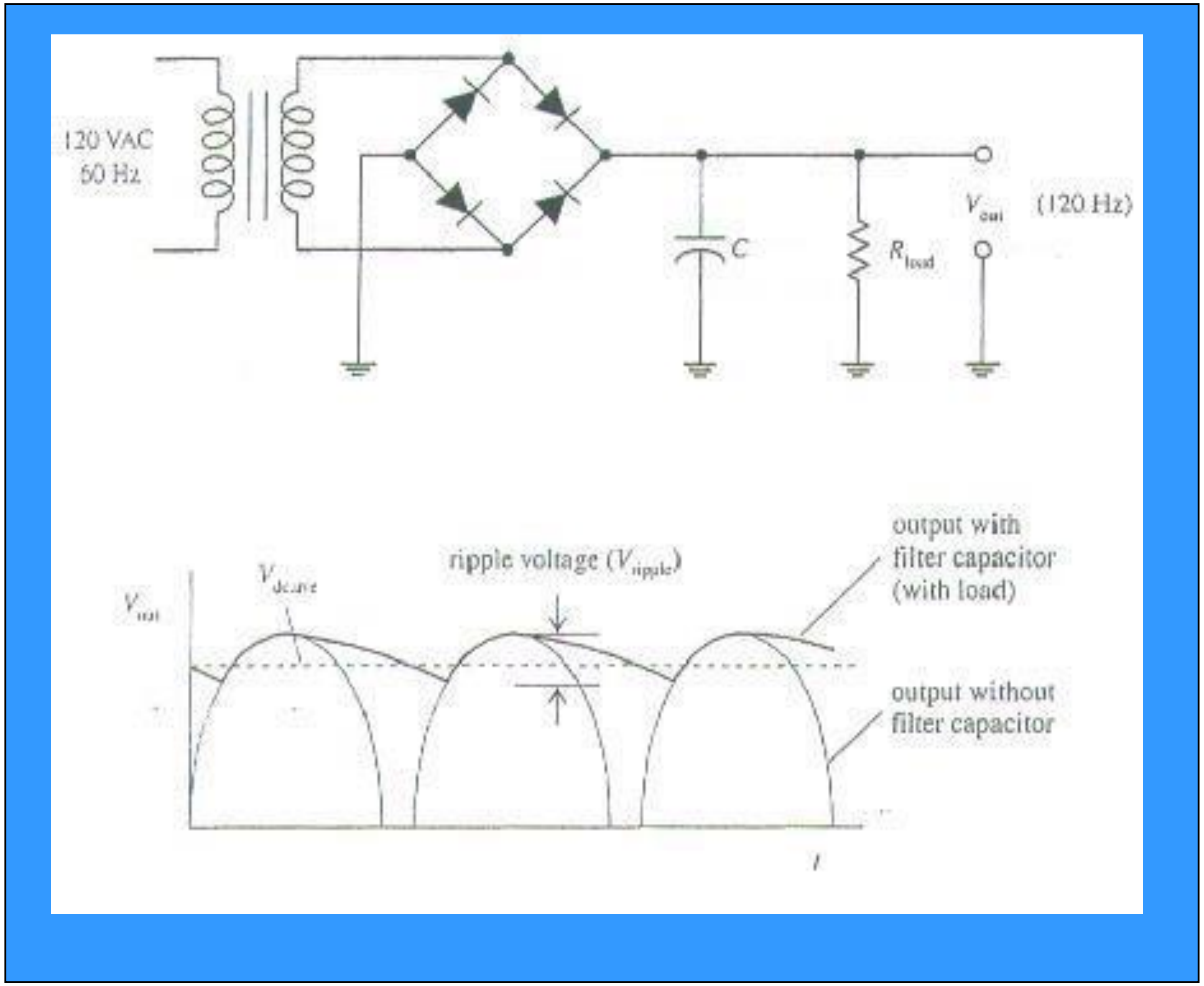
Reverse-biased ("closed door")



- **Diode Applications —Full Wave Rectifier**
- A full-wave rectifier does not block negative swings in the i/p voltage, rather it transforms them into positive swings at the o/p.
- To gain an understanding of device operation, follow current flow through pairs of diodes in the bridge circuit.
- It is easily seen that one pair (D3-R_{out}-D2) allows current flow during the +ve half cycle of V_{in} while the other pair (D4-R_{out}-D1) allows current flow during the -ve half cycle of V_{in}.
  - o/p voltage peak is 1.2V below the i/p voltage peak.
  - The o/p frequency is twice the i/p frequency.



- **Diode Applications —AC2DC Power Supply**



•An AC2DC power supply is built using a transformer and a full-wave rectifier.

•Transformer is used to step down the voltage i/p.

•Rectifier converts AC to pulsed DC.

•A filter capacitor is used to smooth out the pulses. •Capacitor must be large enough to store sufficient charge so as to provide a steady current supply to the load:

$R_{Load}C$ □□1/ f f is rectified signal's frequency (120Hz).

4b

Other diodes

Introduction

Diodes are two-terminal electronic devices / components that functions as a one-way switch i.e., they allow current to flow only in one direction. These diodes are manufactured using semiconductor materials like Silicon, Germanium and Gallium Arsenide.

The two terminals of the diode are known as Anode and Cathode. Based on the potential difference between these two terminals, the operation of diode can be classified in two ways:

- If anode has higher potential than cathode, then the diode is said to be in Forward Bias and it allows current to flow.
- If cathode has higher potential than anode, then the diode is said to be in Reverse Bias and it doesn't allow current to flow.

Different types of diodes have different voltage requirements. For Silicon Diodes, the forward voltage is 0.7V and for Germanium diodes, it is 0.3V. Usually, in Silicon Diodes, the dark band on one end of the diode indicates the Cathode terminal and the other terminal is anode.

One of the main application of Diodes is Rectification i.e., to convert AC to DC. Since diodes allow current to flow only in one direction and blocks current flow in the other direction, diodes are used in reverse polarity protector and transient protector applications.

There are many Different Types of diodes and some of them are listed below.

Different Types of Diodes

Let us now briefly see about few common types of diodes.

1. Small Signal Diode

It is a small device with disproportional characteristics whose applications are mainly involved at high frequency and very low current applications such as radios and televisions etc. To protect the diode from contamination it is enveloped with a glass so it is also named as Glass Passivated Diode. One of the popular diodes of this type is the 1N4148.

Appearance wise, signal diodes are very small when compared with power diodes. To indicate the cathode terminal, one edge is marked with black or red color. For applications at high frequencies, the performance of the small signal diode is very effective.

With respect to the other functionalities, the signal diodes usually have a small current carrying capability and power dissipation. Usually, these are in the range of 150mA and 500mW respectively.

The Small Signal Diode can be made of either Silicon or Germanium type semiconductor material, but the characteristics of the diode varies depending up on the doping material.

Small Signal Diodes are used in general purpose diode applications, high speed switching, parametric amplifiers and many other applications.

Some important characteristics of Small Signal Diode are:

- Peak Reverse Voltage (V_{PR}) – It is the maximum reverse voltage that can be applied to the diode before it breaks down.
- Reverse Current (I_R) – The current (very small value) that flows when it is reverse biased.
- Maximum Forward Voltage at Peak Forward Current (V_F at I_F)
- Reverse Recovery Time – The time required for reverse current to fall down from forward current to I_R.

2. Large Signal Diode

These diodes have large PN junction layer. Thus, they are usually used in rectification i.e., converting AC to DC. The large PN Junction also increases the forward current carrying capacity and reverse blocking voltage of the diode. The large signal diodes are not suitable for high frequency applications.

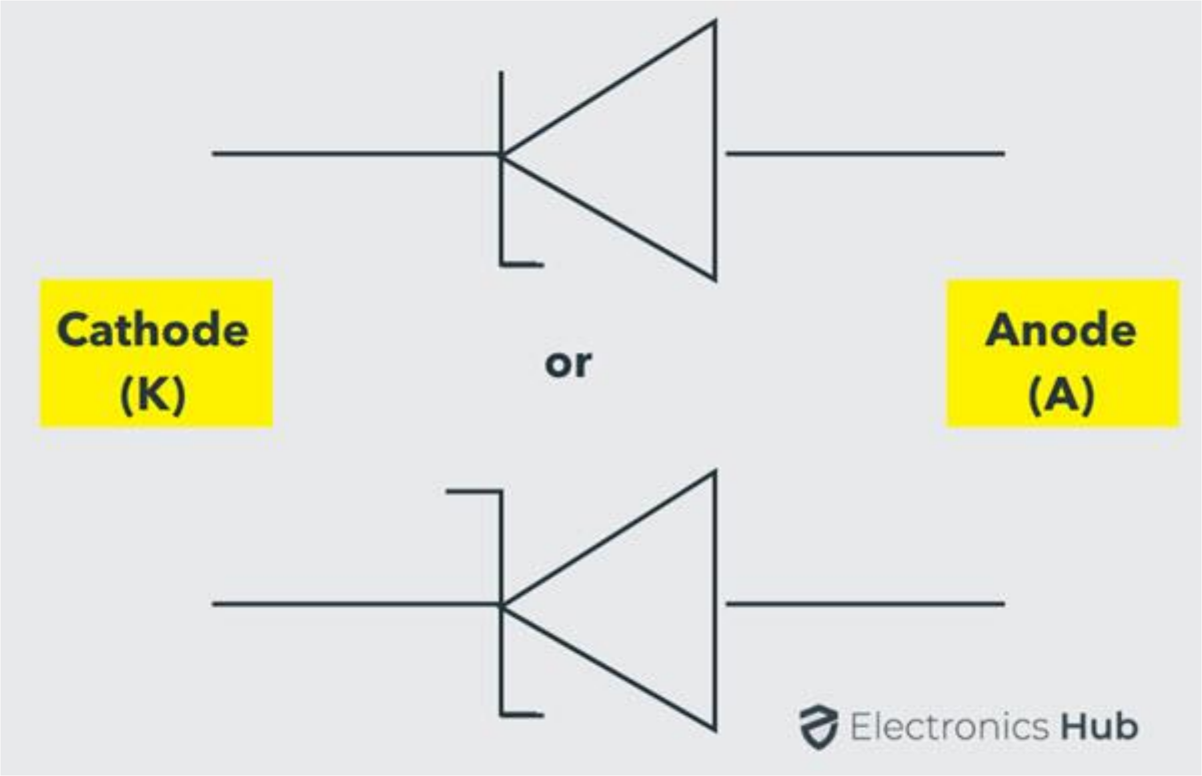
The main applications of these diodes are in Power Supplies (rectifiers, converter, inverters, battery charging devices, etc.). In these diodes, the value of forward resistance is few Ohms and the value of reverse blocking resistance is in Mega Ohms.

Since it has high current and voltage performance, these can be used in electrical devices which are used to suppress high peak voltages.

3. Zener Diode

It is a passive element which works under the principle of ‘Zener Breakdown’. First produced by Clarence Zener in 1934, it is similar to normal diode in forward bias condition i.e., it allows current to flow.

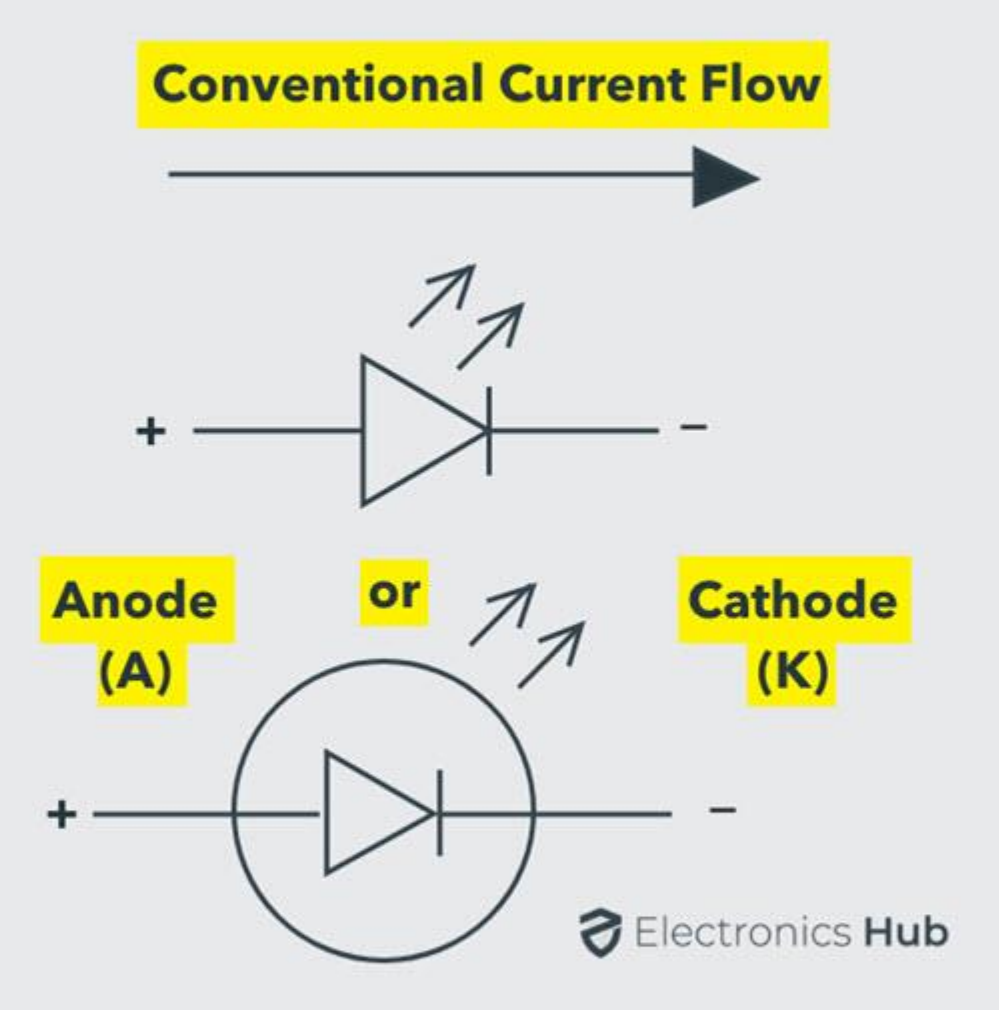
But in reverse bias condition, the diode conducts only when the applied voltage reaches the breakdown voltage, known as Zener Breakdown. It is designed to prevent the other semiconductor devices from momentary voltage pulses. It acts as voltage regulator.



4. Light Emitting Diode (LED)

These diodes convert the electrical energy in to light energy. First production started in 1968. It undergoes electroluminescence process in which holes and electrons are recombined to produce energy in the form of light in forward bias condition.

In the early days, LEDs are very costly and used only in special application. But over the years, the cost of the LEDs has comedown significantly. This and the fact they are extremely power efficient, makes LEDs as the main source of lighting in homes, offices, streets (for street lighting as well as traffic lights), automobiles, mobile phones.



5. Constant Current Diodes

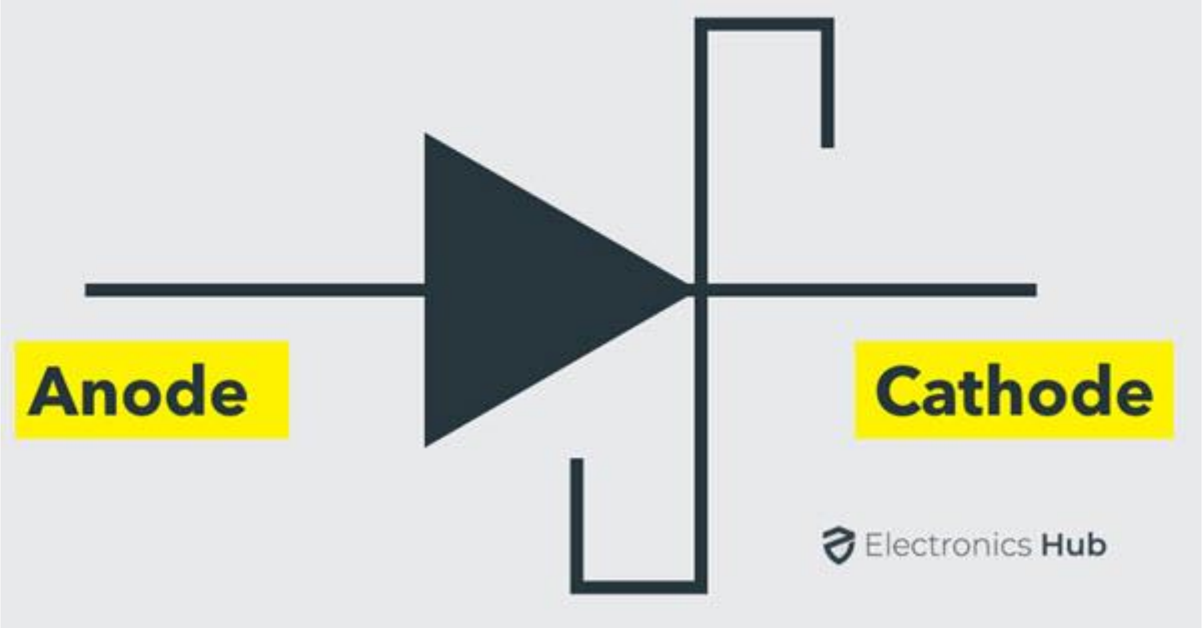
It is also known as Current-Regulating Diode or Current-Limiting Diode or Diode-Connected Transistor. The function of the diode is to regulate the voltage at a particular current.

It functions as a two terminal current limiter. In this, JFET acts as current limiter to achieve high output impedance. The constant current diode symbol is shown below.

6. Schottky Diode

In this type of diode, the junction is formed by contacting the semiconductor material with metal. Due to this, the forward voltage drop is decreased to a minimum. The semiconductor material is N-type silicon, which acts as an anode and metals such as Chromium, Platinum, Tungsten etc. acts as cathode.

Due to the metal junction, these diodes have high current conducting capability and hence the switching time is reduced. So, Schottky Diode has greater use in switching applications. Mainly because of the metal – semiconductor junction, the voltage drop is low, which in turn increases the diode performance and reduces power loss. So, these are used in high frequency rectifier applications. The symbol of Schottky diode is as shown below.

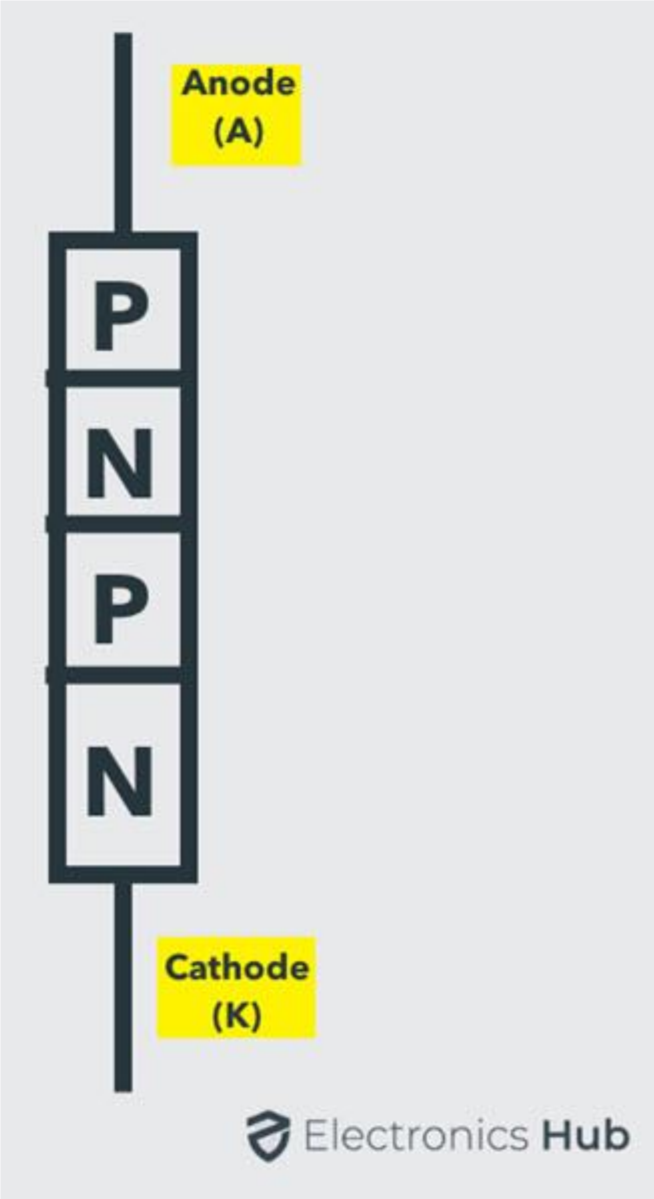


Applications	Purpose
Voltage clamping	To prevent transistor from saturation this is possible due to the higher current density of schottky diode
Reverse current and discharge protection	It prevents the reverse currents and helps the batteries to discharge
Power supply	Acts as rectifier

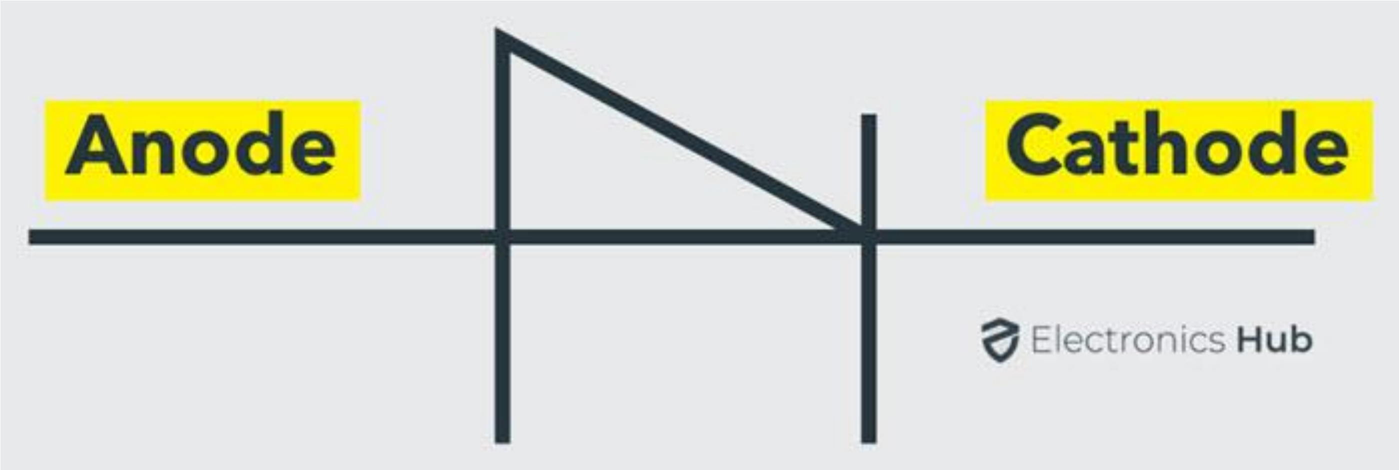
7. Shockley Diode

It was one of the first semiconductor devices to be invented. Shockley Diode has four layers. It is also called as PNPN diode. It is equal to a thyristor without a gate terminal, which means the gate terminal is disconnected. As there is no trigger input, the only way the diode can conduct is by providing forward voltage.

It stays ON once it turned “ON” and stays OFF once it turned “OFF”. The diode has two operating states conducting and non-conducting. In non-conducting state the diode conducts with less voltage.



The symbol of the Shockley diode is as follows:



Shockley Diode Applications

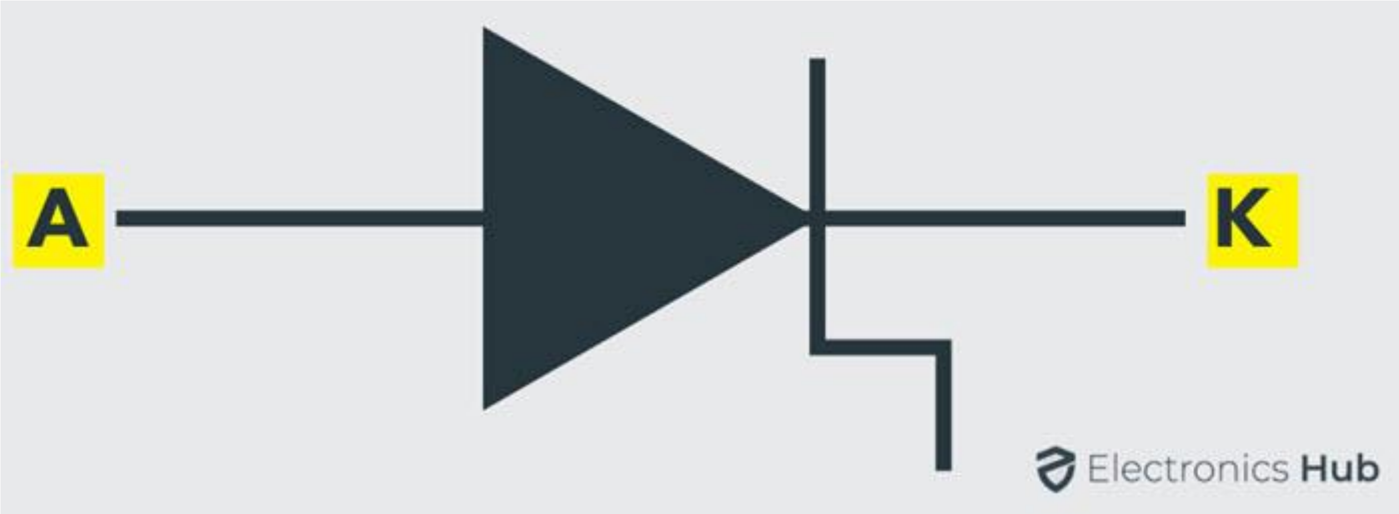
- Trigger switches for SCR.
- Acts as relaxation oscillator.

8. Step Recovery Diodes

It is also called as snap-off diode or charge-storage diode. These are the special type of diodes which stores the charge from positive pulse and uses in the negative pulse of the sinusoidal signals. The rise time of the current pulse is equal to the snap time. Due to this phenomenon, it has speed recovery pulses.

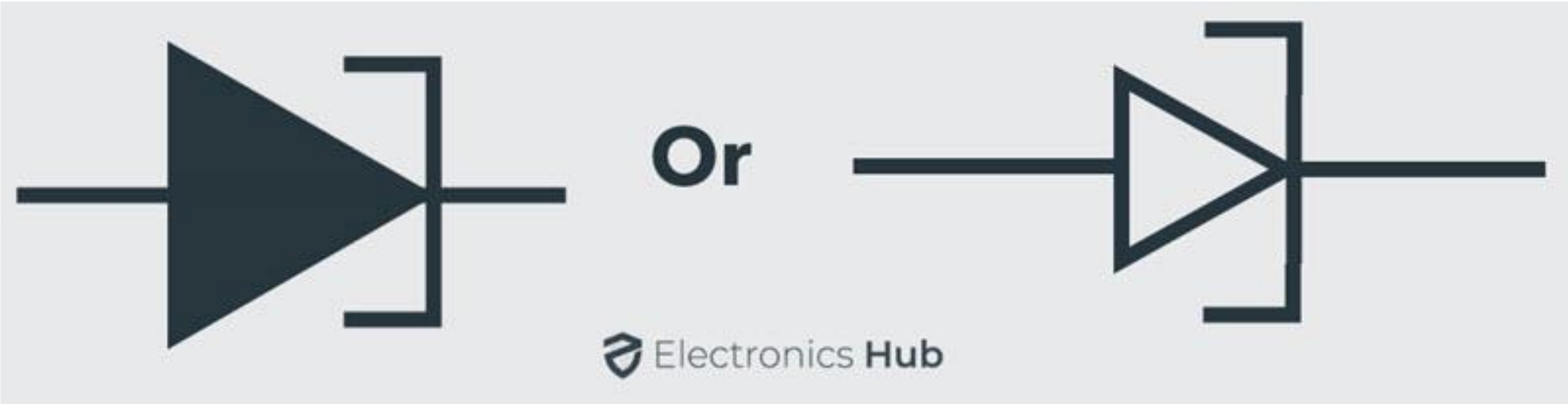
The applications of these diodes are in higher order multipliers and in pulse shaper circuits. The cut-off frequency of these diodes is very high which are nearly at Giga hertz order.

As multiplier, this diode has the cut-off frequency range of 200 to 300 GHz. In the operations which are performing at 10 GHz range, these diodes play a vital role. The efficiency is high for lower order multipliers. The symbol for this diode is as shown below.



9. Tunnel Diode

It is used as high-speed switch, with switching speed in the order of few nano-seconds. Due to tunneling effect it has very fast operation in microwave frequency region. It is a two-terminal device in which concentration of dopants is too high.
The transient response is being limited by junction capacitance plus stray wiring capacitance. Mostly used in microwave oscillators and amplifiers. It acts as most negative conductance device. Tunnel diodes can be tuned both mechanically and electrically. The symbol of tunnel diode is as shown below.



Tunnel Diode Applications

- Oscillatory circuits.

- Microwave circuits.
- Resistant to nuclear radiation.

10. Varactor Diode

These are also known as Varicap diodes. It acts like the variable capacitor. Operations are performed mainly at reverse bias state only. These diodes are very famous due to its capability of changing the capacitance ranges within the circuit in the presence of constant voltage flow.

They can be able to vary capacitance up to high values. In varactor diode, we can decrease or increase the depletion layer by changing the reverse bias voltage. These diodes have many applications as voltage-controlled oscillator for cell phones, satellite pre-filters etc. The symbol of varactor diode is given below.

Varactor Diode Applications

- Voltage-controlled capacitors
- Voltage-controlled oscillators
- Parametric amplifiers
- Frequency multipliers
- FM transmitters and Phase locked loops in radio, television sets and cellular phone

11. Laser Diode

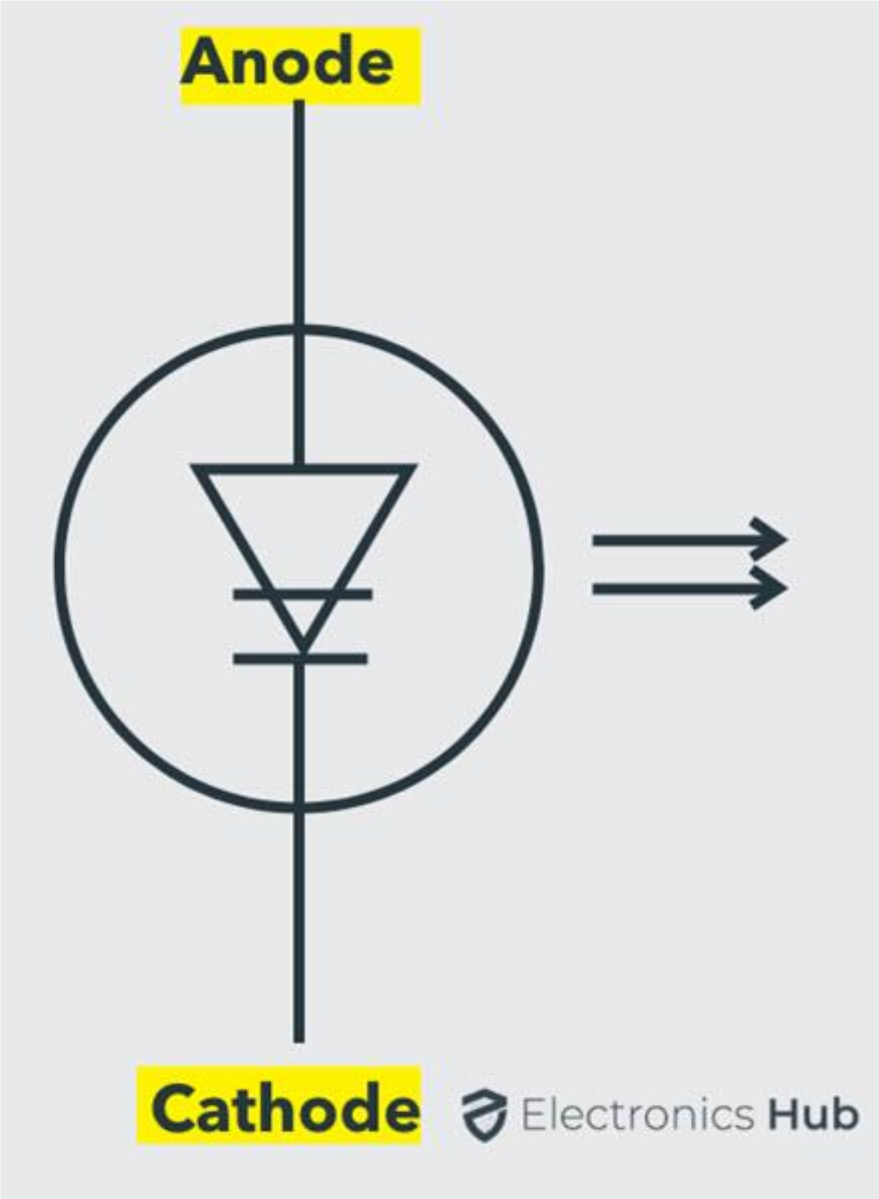
Similar to LED in which active region is formed by p-n junction. Electrically laser diode is P-I-N diode in which the active region is in intrinsic region. Used in fiber optic communications, barcode readers, laser pointers, CD/DVD/Blu-ray reading and recording, Laser printing.

Laser Diode Types:

- Double Heterostructure Laser: Free electrons and holes available simultaneously in the region.
- Quantum Well Lasers: lasers having more than one quantum well are called multi quantum well lasers.
- Quantum Cascade Lasers: These are heterojunction lasers which enables laser action at relatively long wavelengths.

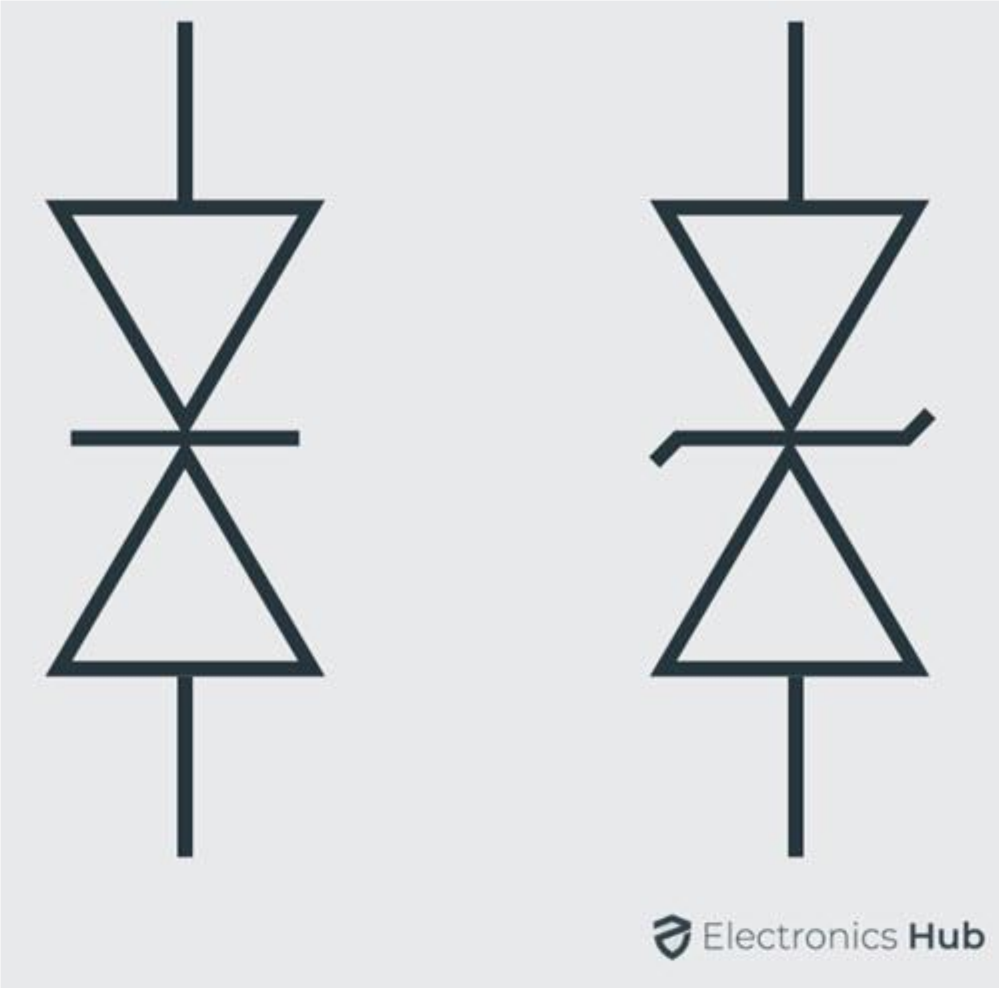
- Separate Confinement Heterostructure Lasers: To compensate the thin layer problem in quantum lasers we go for separate confinement heterostructure lasers.
- Distributed Bragg Reflector Lasers: It can be edge emitting lasers or VCSELS.

The symbol of the Laser Diode is as shown:



12. Transient Voltage Suppression Diode

In semiconductor devices, transients will occur due to the sudden change in the state voltage. They will damage the device's output response. To overcome this problem, Voltage Suppression Diodes are used. The operation of voltage suppression diode is similar to Zener diode operation.
The operation of these diodes is normal as p-n junction diodes but at the time of transient voltage its operation changes. In normal condition, the impedance of the diode is high. When any transient voltage occurs in the circuit, the diode enters in to the avalanche breakdown region in which a low impedance is provided.
It is very spontaneously because the avalanche breakdown duration ranges in Pico seconds. Transient voltage suppression diode will clamp the voltage to the fixed levels, mostly its clamping voltage is in minimum range.
These are having applications in the telecommunication fields, medical, microprocessors and signal processing. It responds to over voltages faster than Varistors or gas discharge tubes.
The symbol for Transient voltage suppression diode is as shown below.



The diode is characterized by:

- Leakage current
- Maximum reverse stand-off voltage
- Breakdown voltage
- Clamping voltage
- Parasitic capacitance
- Parasitic inductance

- Amount of energy it can absorb

13. Gold Doped Diodes

In these diodes, Gold is used as a dopant. These diodes are faster than other diodes. In these diodes, the leakage current in reverse bias condition is also less. Even at the higher voltage drop it allows the diode to operate in signal frequencies. In these diodes, Gold helps for the faster recombination of minority carriers.

14. Super Barrier Diodes

It is a rectifier diode having low forward voltage drop as Schottky diode with surge handling capability and low reverse leakage current as P – N junction diode. It was designed for high power, fast switching and low-loss applications. Super barrier rectifiers are the next generation rectifiers with low forward voltage than Schottky diode.

15. Peltier Diode

In this type of diode, it generates heat at the two-material junction of a semiconductor, which flows from one terminal to another terminal. This flow is done in only single direction which is same as the direction of current flow.

This heat is produced due to electric charge produced by the recombination of minority charge carriers. This is mainly used in cooling and heating applications. This type of diodes used as sensor and heat engine for thermo electric cooling.

16. Crystal Diode

This is also known as Cat's whisker, which is a type of point contact diode. Its operation depends on the pressure of contact between semiconductor crystal and the point.

In this, a metal wire is present, which is pressed against the semiconductor crystal. In this, the semiconductor crystal acts as cathode and metal wire acts as anode. These diodes are obsolete in nature. Mainly used in microwave receivers and detectors.

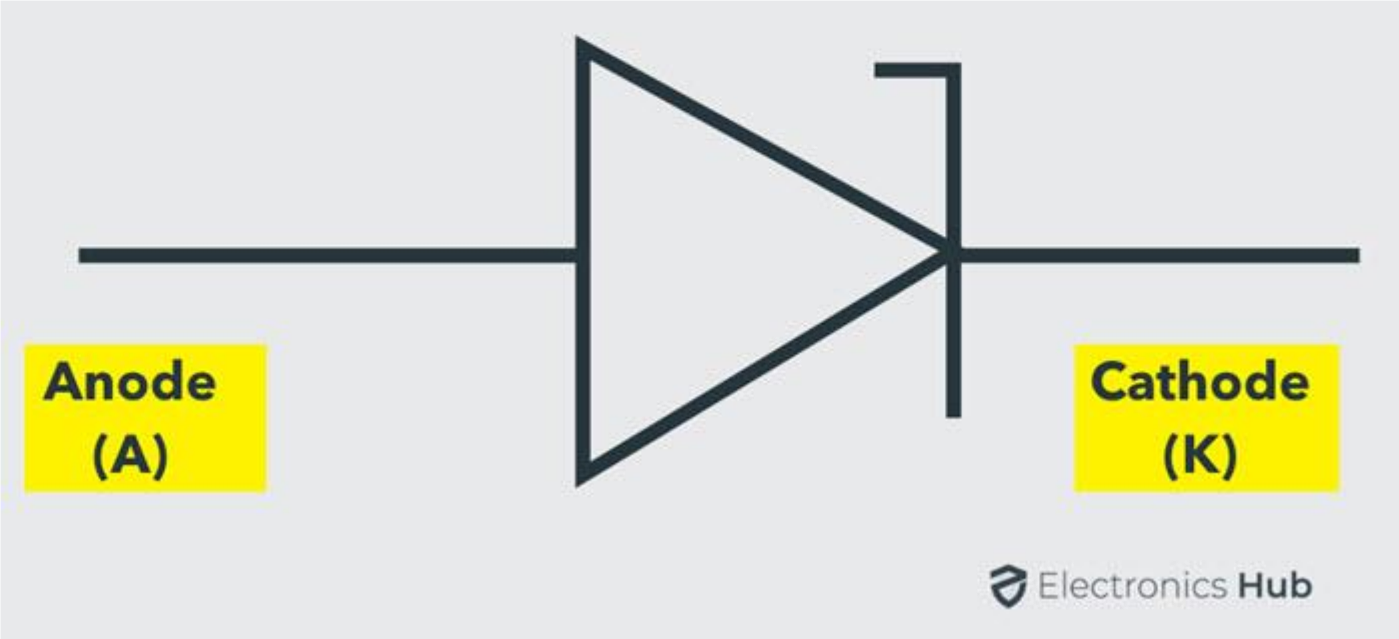
Crystal Diode Applications

- Crystal diode rectifier
- Crystal diode detector
- Crystal radio receiver

17. Avalanche Diode

This is passive element works under principle of Avalanche Breakdown. It works in reverse bias condition. It results in a large current due to the ionization produced by P – N junction during reverse bias condition.

These diodes are specially designed to undergo breakdown at specific reverse voltage to prevent the damage. The symbol of the avalanche diode is as shown below:

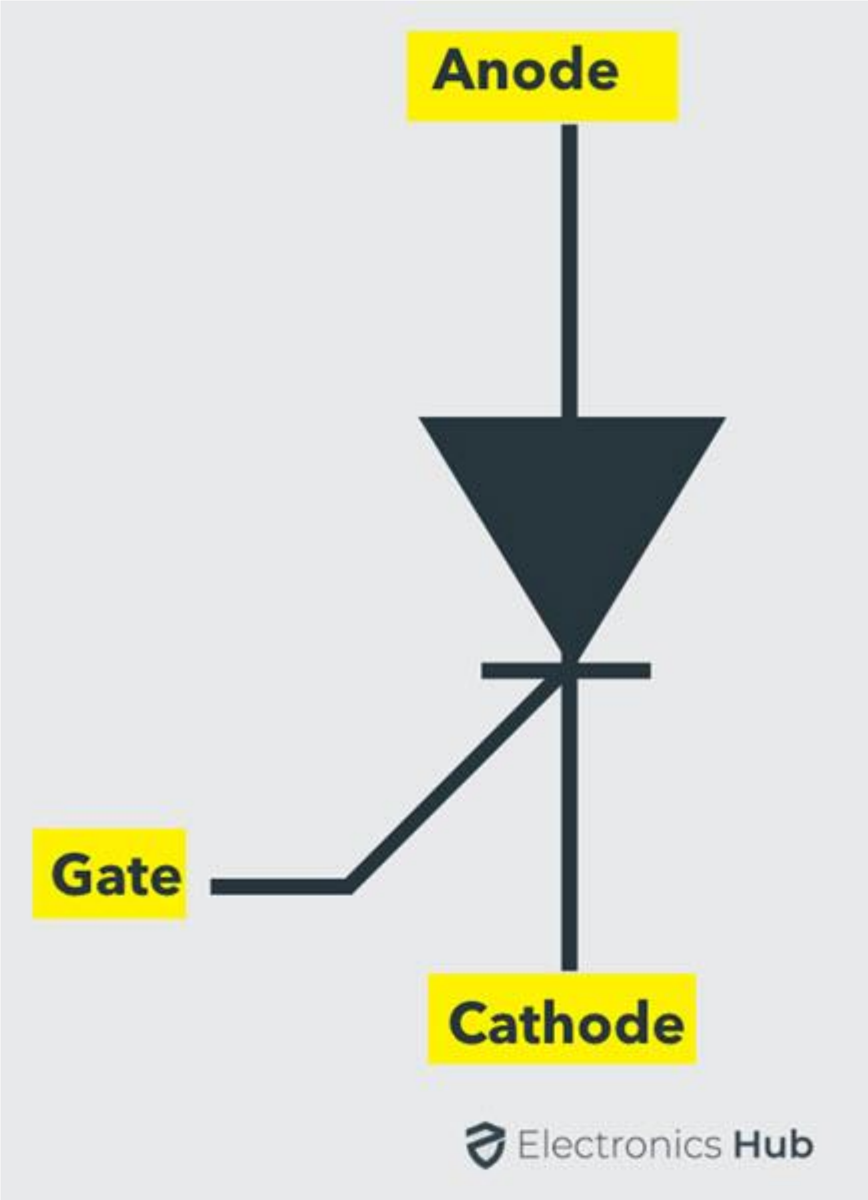


Avalanche Diode Uses

- RF Noise Generation: It acts as source of RF for antenna analyzer bridges and also as white noise generators.
- Used in radio equipment and also in hardware random number generators.
- Microwave Frequency Generation: In this the diode acts as negative resistance device.
- Single Photon Avalanche Detector: These are high gain photon detectors used in light level applications.

18. Silicon Controlled Rectifier

It consists of three terminals they are anode, cathode and a gate. It is nearly equal to the Shockley diode. As its name indicates it is mainly used for the control purpose when small voltages are applied in the circuit. The symbol of the Silicon Controlled Rectifier is as shown below:



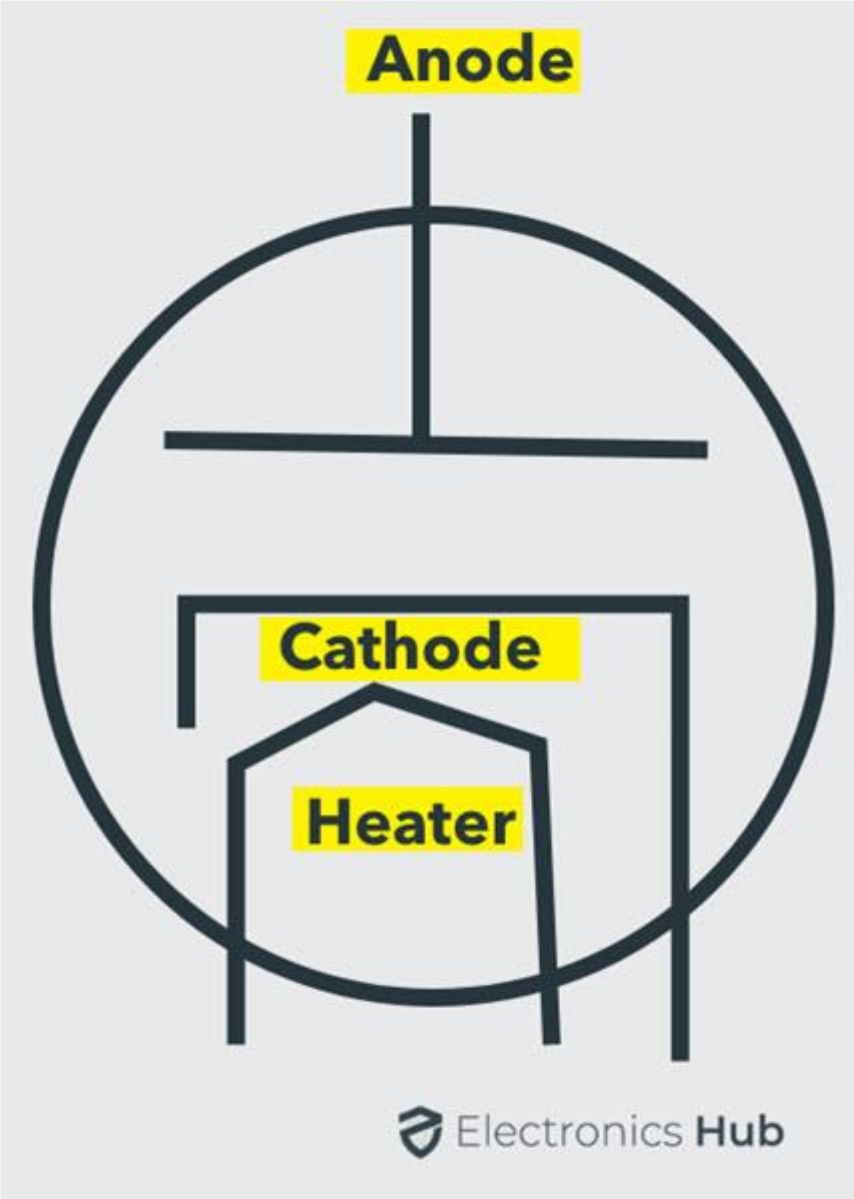
Modes of Operation:

1. Forward blocking mode (off state): In this J1 and J3 forward biased and J2 is reverse biased. It offers high resistance below breakover voltage and hence it is said to be off state.
2. Forward conduction mode (on state): By increasing the voltage at anode and cathode or by applying positive pulse at the gate we can turn ON. To turn off the only way is to decrease the current flowing through it.
3. Reverse blocking mode (off state): SCR blocking the reverse voltage is named as asymmetrical SCR. Mostly used in current source inverters.

19. Vacuum Diodes

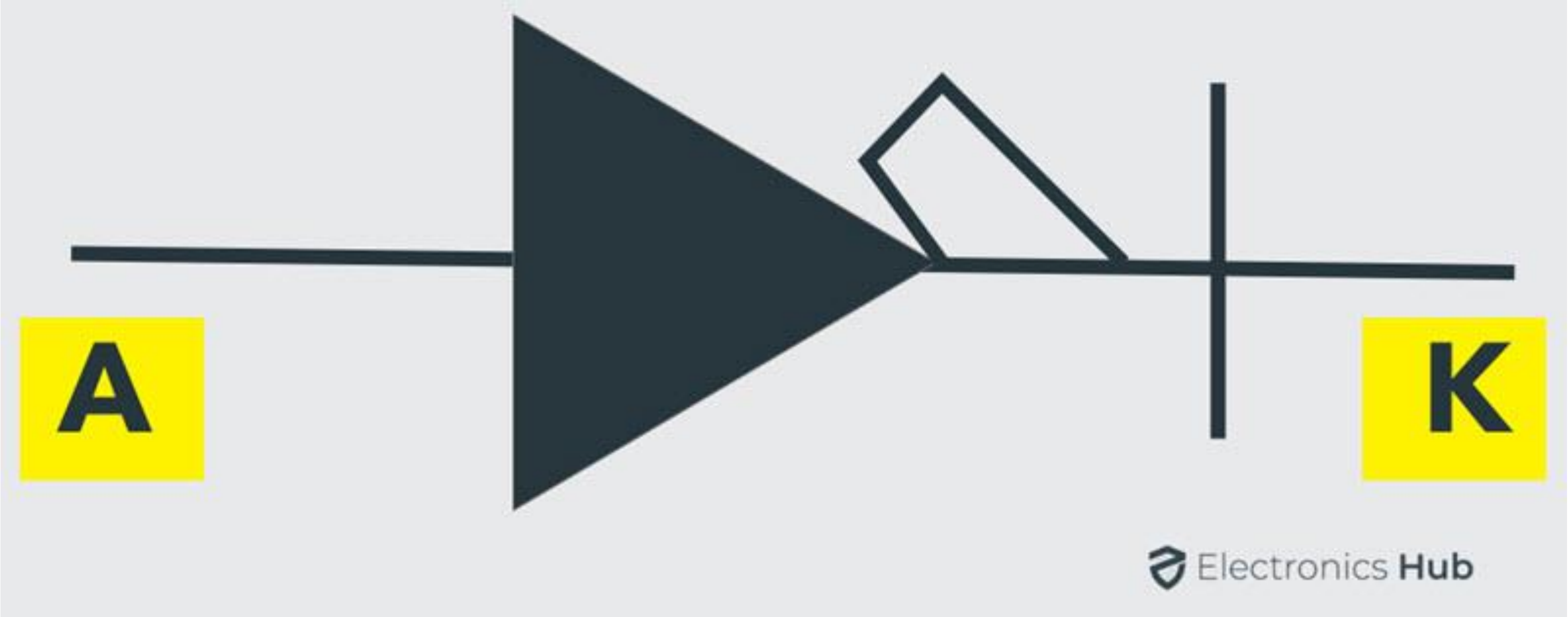
Vacuum diodes consist of two electrodes which will acts as an anode and the cathode. Cathode is made up of Tungsten, which emits the electrons in the direction of anode. Always electron flow will be from cathode to anode only. So, it acts like a switch.

If the cathode is coated with oxide material, then the electrons emission capability is high. Anode is a bit long in size and in some cases their surface is rough to reduce the temperatures developing in the diode. The diode will conduct only in one case that is when the anode is positive with respect to cathode terminal. The symbol is as shown in figure:



20. PIN Diode

The improved version of the normal P-N junction diode gives the PIN diode. In PIN diode doping is not necessary. The intrinsic material i.e., the material which has no charge carriers, is inserted between the P and N regions, which increase the area of depletion layer.
When we apply forward bias voltage, the holes and electrons will be pushed into the intrinsic layer. At some point due to this high injection level, the electric field will conduct through the intrinsic material also. This field makes the carriers to flow from two regions. The symbol of PIN diode is as shown below:

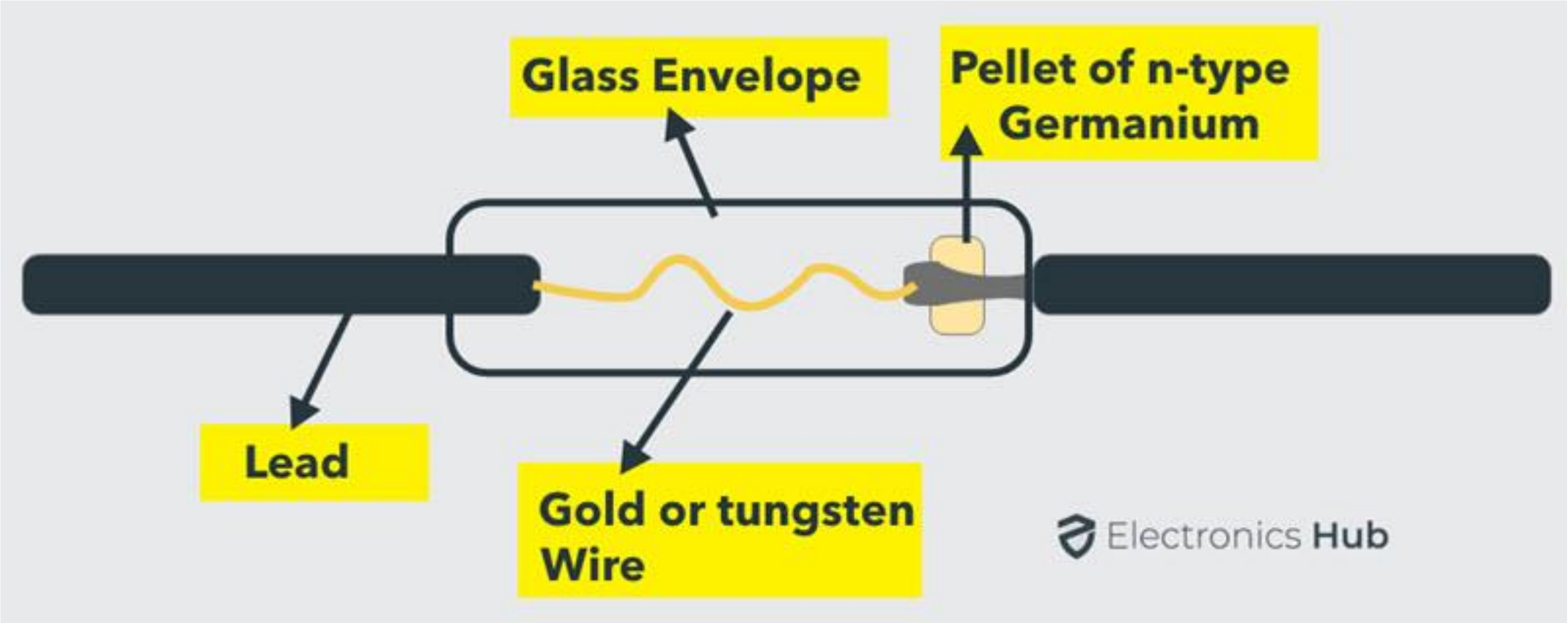


PIN Diode Applications:

- RF Switches: PIN diode is used for both signal and component selection. For example, PIN diodes acts as range-switch inductors in low phase noise oscillators.
- Attenuators: it is used as bridge and shunt resistance in bridge-T attenuator.
- Photo Detectors: it detects x-ray and gamma ray photons.

21. Point Contact Devices

A gold or tungsten wire is used to act as the point contact to produce a PN junction region by passing a high electric current through it. A small region of PN junction is produced around the edge of the wire which is connected to the metal plate which is as shown in the figure.

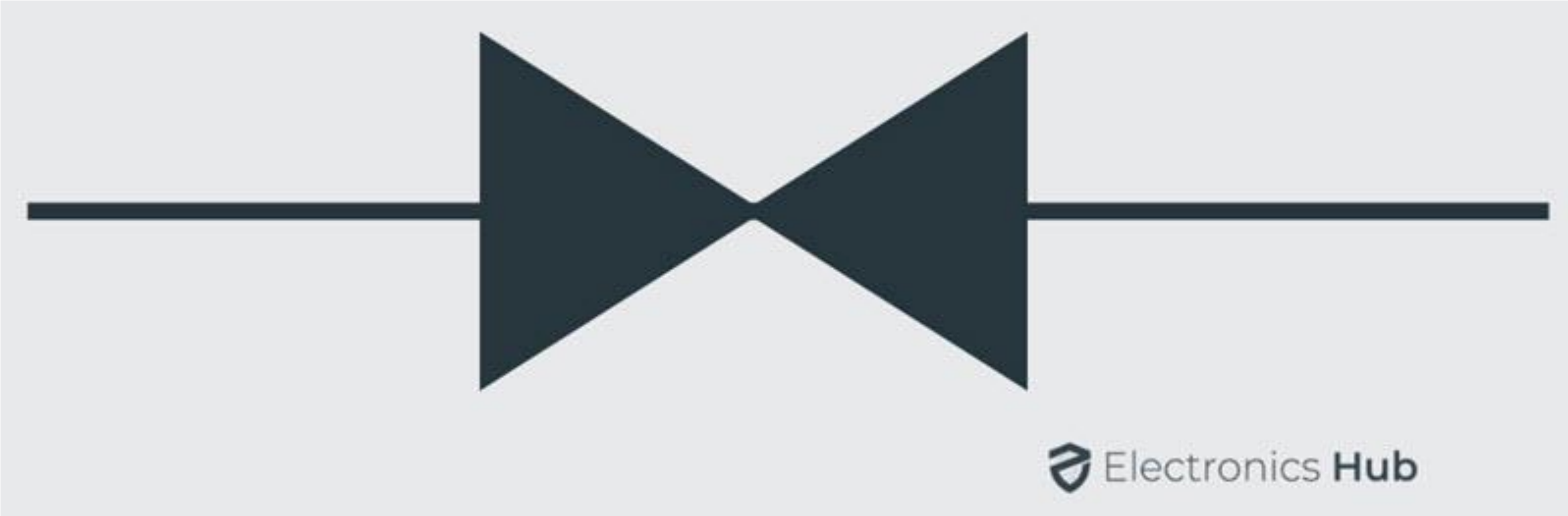


In forward direction, its operation is quite similar but in reverse bias condition the wire acts like an insulator. Since this insulator is between the plates, the diode acts as a capacitor. In general, the capacitor blocks the DC currents but the AC currents can flow in the circuit at high frequencies. So, these are used to detect the high frequency signals.

22. Gunn Diode

Gunn diode is fabricated with n-type semiconductor material only. The depletion region of two N-type materials is very thin. When voltage increases in the circuit, the current also increases. After certain level of voltage, the current will exponentially decrease, thus this exhibits the negative differential resistance.

It has two electrodes with Gallium Arsenide and Indium Phosphide. Due to this, it has negative differential resistance. It is also termed as transferred electron device. It produces micro wave RF signals so it is mainly used in Microwave RF devices. It can also use as an amplifier. The symbol of Gunn diode is shown below:



Chapter 5: Bipolar Junction Transistors

The Bipolar Junction Transistor is a semiconductor device which can be used for switching or amplification
In the diode tutorials we saw that simple diodes are made up from two pieces of semiconductor material to form a simple pn-junction and we also learnt about their properties and characteristics.
If we now join together two individual signal diodes back-to-back, this will give us two PN-junctions connected together in series which would share a common *Positve*, (P) or *Negative*, (N) terminal. The fusion of these two diodes produces a three layer, two junction, three terminal device forming the basis of a **Bipolar Junction Transistor**, or **BJT** for short.
Transistors are three terminal active devices made from different semiconductor materials that can act as either an insulator or a conductor by the application of a small signal voltage. The transistor's ability to change between these two states enables it to have two basic functions: "switching" (digital electronics) or "amplification" (analogue electronics). Then bipolar transistors have the ability to operate within three different regions:

- Active Region – the transistor operates as an amplifier and $Ic = \beta * Ib$
- Saturation – the transistor is "Fully-ON" operating as a switch and Ic = I(saturation)
- Cut-off – the transistor is "Fully-OFF" operating as a switch and Ic = 0

A Typical
Bipolar Transistor

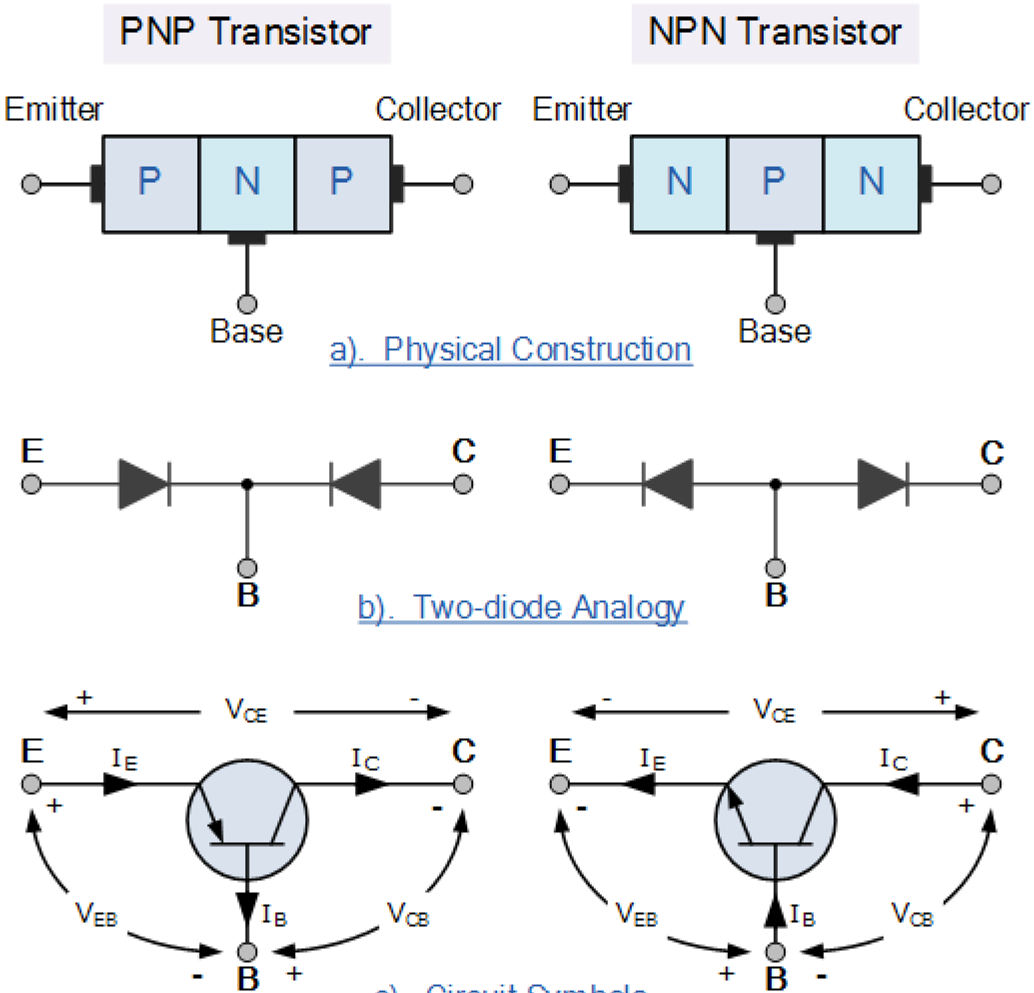
The word Transistor is a combination of the two words Transfer Varistor which describes their mode of operation way back in their early days of electronics development. There are two basic types of bipolar transistor construction, PNP and NPN, which basically describes the physical arrangement of the P-type and N-type semiconductor materials from which they are made.

The **Bipolar Transistor** basic construction consists of two PN-junctions producing three connecting terminals with each terminal being given a name to identify it from the other two. These three terminals are known and labelled as the Emitter (E), the Base (B) and the Collector (C) respectively.

Bipolar Transistors are current regulating devices that control the amount of current flowing through them from the Emitter to the Collector terminals in proportion to the amount of biasing voltage applied to their base terminal, thus acting like a current-controlled switch. As a small current flowing into the base terminal controls a much larger collector current forming the basis of transistor action.

The principle of operation of the two transistor types PNP and NPN, is exactly the same the only difference being in their biasing and the polarity of the power supply for each type.

Bipolar Transistor Construction



a). Physical Construction
b). Two-diode Analogy
c). Circuit Symbols

The construction and circuit symbols for both the PNP and NPN bipolar transistor are given above with the arrow in the circuit symbol always showing the direction of "conventional current flow" between the base terminal and its emitter terminal. The direction of the arrow always points from the positive P-type region to the negative N-type region for both transistor types, exactly the same as for the standard diode symbol.

Bipolar Transistor Configurations

As the **Bipolar Transistor** is a three terminal device, there are basically three possible ways to connect it within an electronic circuit with one terminal being common to both the input and output signals. Each method of connection responding differently to its input signal within a circuit as the static characteristics of the transistor vary with each circuit arrangement.

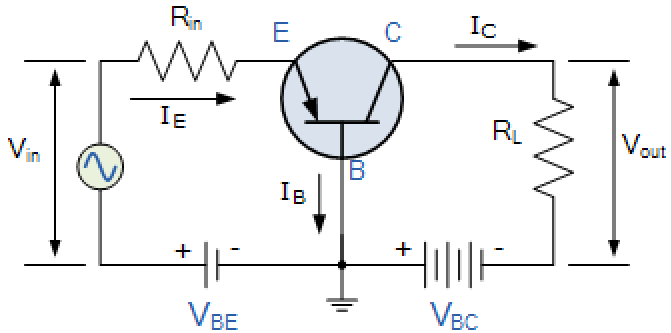
- Common Base Configuration – has Voltage Gain but no Current Gain.
- Common Emitter Configuration – has both Current and Voltage Gain.
- Common Collector Configuration – has Current Gain but no Voltage Gain.

The Common Base (CB) Configuration

As its name suggests, in the **Common Base** or grounded base configuration, the BASE connection is common to both the input signal AND the output signal. The input signal is applied between the transistors base and the emitter terminals, while the corresponding output signal is taken from between the base and the collector terminals as shown. The base terminal is grounded or can be connected to some fixed reference voltage point.

The input current flowing into the emitter is quite large as its the sum of both the base current and collector current respectively therefore, the collector current output is less than the emitter current input resulting in a current gain for this type of circuit of "1" (unity) or less, in other words the common base configuration "attenuates" the input signal.

The Common Base Transistor Circuit



This type of amplifier configuration is a non-inverting voltage amplifier circuit, in that the signal voltages Vin and Vout are "in-phase". This type of transistor arrangement is not very common due to its unusually high voltage gain characteristics. Its input characteristics represent that of a forward biased diode while the output characteristics represent that of an illuminated photo-diode.

Also this type of bipolar transistor configuration has a high ratio of output to input resistance or more importantly "load" resistance (RL) to "input" resistance (Rin) giving it a value of "Resistance Gain". Then the voltage gain (Av) for a common base configuration is therefore given as:

Common Base Voltage Gain

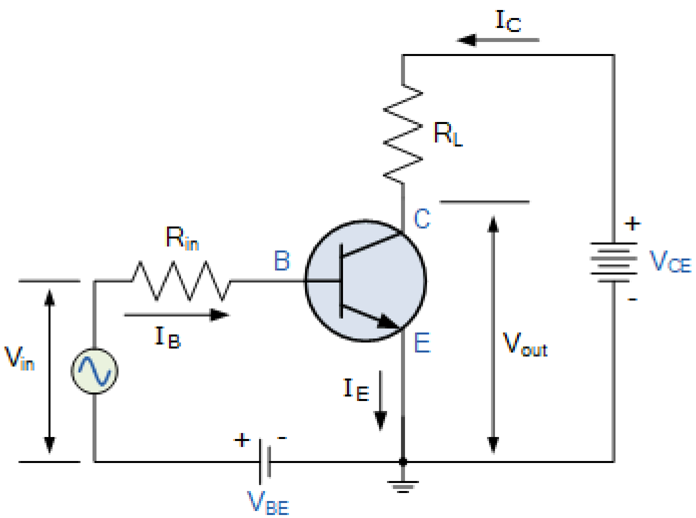
$$A_V = \frac{Vout}{Vin} = \frac{I_C \times R_L}{I_E \times R_{IN}}$$

Where: Ic/Ie is the current gain, alpha (α) and RL/Rin is the resistance gain.
The common base circuit is generally only used in single stage amplifier circuits such as microphone pre-amplifier or radio frequency (R*f*) amplifiers due to its very good high frequency response.

The Common Emitter (CE) Configuration

In the **Common Emitter** or grounded emitter configuration, the input signal is applied between the base and the emitter, while the output is taken from between the collector and the emitter as shown. This type of configuration is the most commonly used circuit for transistor based amplifiers and which represents the "normal" method of bipolar transistor connection.
The common emitter amplifier configuration produces the highest current and power gain of all the three bipolar transistor configurations. This is mainly because the input impedance is LOW as it is connected to a forward biased PN-junction, while the output impedance is HIGH as it is taken from a reverse biased PN-junction.

The Common Emitter Amplifier Circuit



In this type of configuration, the current flowing out of the transistor must be equal to the currents flowing into the transistor as the emitter current is given as Ie = Ic + Ib.

As the load resistance (R_L) is connected in series with the collector, the current gain of the common emitter transistor configuration is quite large as it is the ratio of Ic/Ib. A transistors current gain is given the Greek symbol of Beta, (β).
As the emitter current for a common emitter configuration is defined as Ie = Ic + Ib, the ratio of Ic/Ie is called Alpha, given the Greek symbol of α. Note: that the value of Alpha will always be less than unity.
Since the electrical relationship between these three currents, Ib, Ic and Ie is determined by the physical construction of the transistor itself, any small change in the base current (Ib), will result in a much larger change in the collector current (Ic).
Then, small changes in current flowing in the base will thus control the current in the emitter-collector circuit. Typically, Beta has a value between 20 and 200 for most general purpose transistors. So if a transistor has a Beta value of say 100, then one electron will flow from the base terminal for every 100 electrons flowing between the emitter-collector terminal.
By combining the expressions for both Alpha, α and Beta, β the mathematical relationship between these parameters and therefore the current gain of the transistor can be given as:

$$\text{Alpha}, (\alpha) = \frac{I_C}{I_E} \quad \text{and} \quad \text{Beta}, (\beta) = \frac{I_C}{I_B}$$

$$\therefore I_C = \alpha . I_E = \beta . I_B$$

$$\text{as:} \quad \alpha = \frac{\beta}{\beta + 1} \qquad \beta = \frac{\alpha}{1 - \alpha}$$

$$I_E = I_C + I_B$$

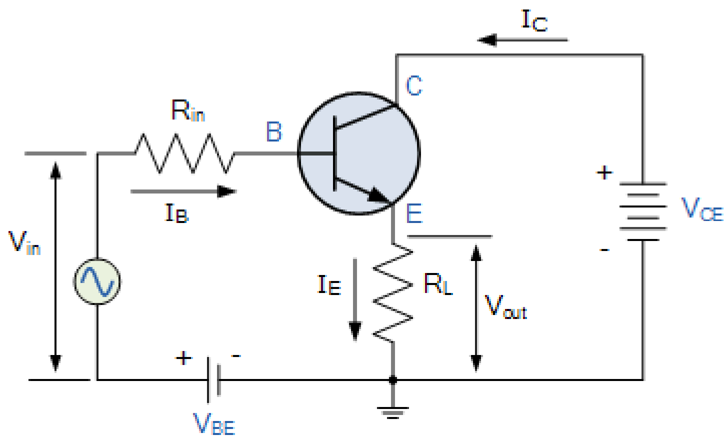
Where: “Ic” is the current flowing into the collector terminal, “Ib” is the current flowing into the base terminal and “Ie” is the current flowing out of the emitter terminal.
Then to summarise a little. This type of bipolar transistor configuration has a greater input impedance, current and power gain than that of the common base configuration but its voltage gain is much lower. The common emitter configuration is an inverting amplifier circuit. This means that the resulting output signal has a 180^{o} phase-shift with regards to the input voltage signal.

The Common Collector (CC) Configuration

In the **Common Collector** or grounded collector configuration, the collector is connected to ground through the supply, thus the collector terminal is common to both the input and the output. The input signal is connected directly to the base terminal, while the output signal is taken from across the emitter load resistor as shown. This type of configuration is commonly known as a **Voltage Follower** or **Emitter Follower** circuit.

The common collector, or emitter follower configuration is very useful for impedance matching applications because of its very high input impedance, in the region of hundreds of thousands of Ohms while having a relatively low output impedance.

The Common Collector Transistor Circuit



The common emitter configuration has a current gain approximately equal to the β value of the transistor itself. However in the common collector configuration, the load resistance is connected in series with the emitter terminal so its current is equal to that of the emitter current. As the emitter current is the combination of the collector AND the base current combined, the load resistance in this type of transistor configuration also has both the collector current and the input current of the base flowing through it. Then the current gain of the circuit is given as:

The Common Collector Current Gain

$$I_E = I_C + I_B$$

$$A_i = \frac{I_E}{I_B} = \frac{I_C + I_B}{I_B}$$

$$A_i = \frac{I_C}{I_B} + 1$$

$$A_i = \beta + 1$$

This type of bipolar transistor configuration is a non-inverting circuit in that the signal voltages of Vin and Vout are "in-phase". The common collector configuration has a voltage gain of about "1" (unity gain). Thus it can considered as a voltage-buffer since the voltage gain is unity. The load resistance of the common collector transistor receives both the base and collector currents giving a large current gain (as with the common emitter configuration) therefore, providing good current amplification with very little voltage gain.

Having looked at the three different types of bipolar transistor configurations, we can now summarise the various relationships between the transistors individual DC currents flowing through each leg and its DC current gains given above in the following table.

Relationship between DC Currents and Gains

$I_E = I_B + I_C$	$\alpha = \frac{I_C}{I_E} = \frac{\beta}{1+\beta}$
$I_C = I_E - I_B$	
$I_B = I_E - I_C$	$\beta = \frac{I_C}{I_B} = \frac{\alpha}{1-\alpha}$

$$I_B = \frac{I_C}{\beta} = \frac{I_E}{1+\beta} = I_E(1-\alpha)$$

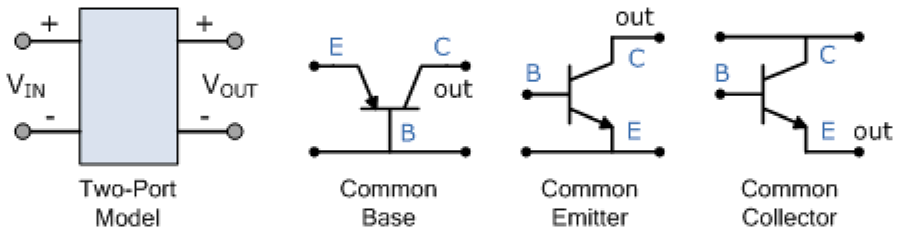
$$I_C = \beta . I_B = \alpha . I_E \qquad I_E = \frac{I_C}{\alpha} = I_B(1+\beta)$$

Note that although we have looked at *NPN Bipolar Transistor* configurations here, PNP transistors are just as valid to use in each configuration as the calculations will all be the same, as for the non-inverting of the amplified signal. The only difference will be in the voltage polarities and current directions.

Bipolar Transistor Summary

Then to summarise, the behaviour of the bipolar transistor in each one of the above circuit configurations is very different and produces different circuit characteristics with regards to input impedance, output impedance and gain whether this is voltage gain, current gain or power gain and this is summarised in the table below.

Bipolar Transistor Configurations



with the generalised characteristics of the different transistor configurations given in the following table:

Characteristic	Common Base	Common Emitter	Common Collector
Input Impedance	Low	Medium	High
Output Impedance	Very High	High	Low
Phase Shift	0°	180°	0°

Voltage Gain	High	Medium	Low
Current Gain	Low	Medium	High
Power Gain	Low	Very High	Medium

NPN Transistor

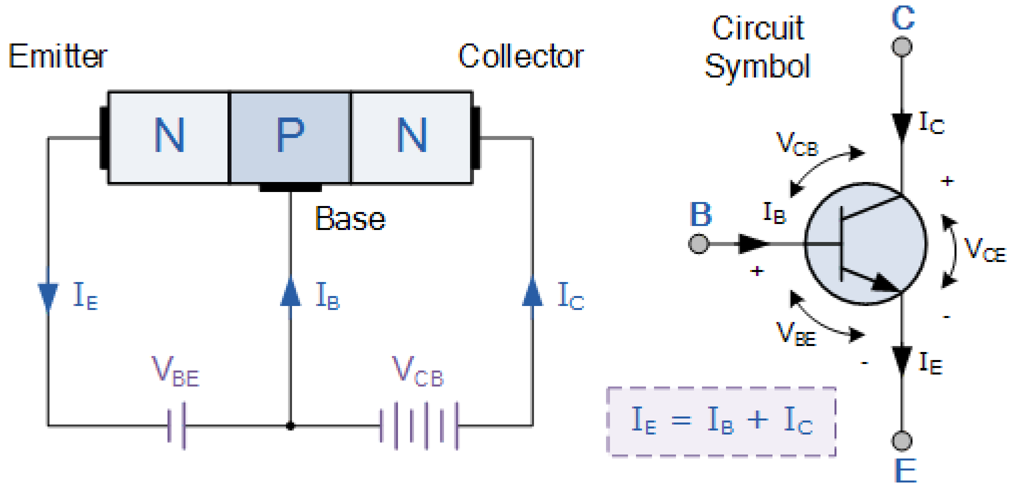
NPN Transistors are three-terminal, three-layer devices that can function as either amplifiers or electronic switches

In the previous tutorial we saw that the standard **Bipolar Transistor** or BJT, comes in two basic forms. An **NPN** (**N**egative-**P**ositive-**N**egative) type and a **PNP** (Positive-**N**egative-**P**ositive) type.

The most commonly used transistor configuration is the **NPN Transistor**. We also learnt that the junctions of the bipolar transistor can be biased in one of three different ways – **Common Base**, **Common Emitter** and **Common Collector**.

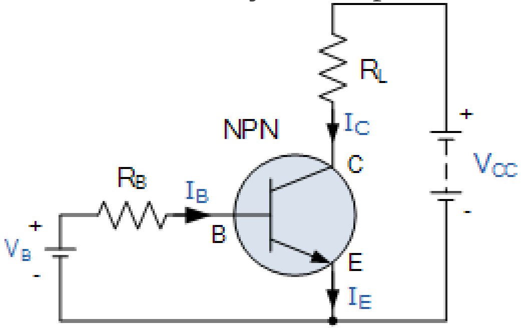
In this tutorial about bipolar transistors we will look more closely at the “Common Emitter” configuration using the **Bipolar NPN Transistor** with an example of the construction of a NPN transistor along with the transistors current flow characteristics is given below.

A Bipolar NPN Transistor Configuration



(Note: Arrow defines the emitter and conventional current flow, “out” for a Bipolar NPN Transistor.)

The construction and terminal voltages for a bipolar NPN transistor are shown above. The voltage between the Base and Emitter (V_{BE}), is positive at the Base and negative at the Emitter because for an NPN transistor, the Base terminal is always positive with respect to the Emitter. Also the Collector supply voltage is positive with respect to the Emitter (V_{CE}). So for a bipolar NPN transistor to conduct the Collector is always more positive with respect to both the Base and the Emitter.



NPN Transistor Connection

Then the voltage sources are connected to an NPN transistor as shown. The Collector is connected to the supply voltage V_{CC} via the load resistor, RL which also acts to limit the maximum current flowing through the device. The Base supply voltage V_B is connected to the Base resistor R_B, which again is used to limit the maximum Base current.

So in a NPN Transistor it is the movement of negative current carriers (electrons) through the Base region that constitutes transistor action, since these mobile electrons provide the link between the Collector and Emitter circuits. This link between the input and output circuits is the main feature of transistor action because the transistors amplifying properties come from the consequent control which the Base exerts upon the Collector to Emitter current.

Then we can see that the transistor is a current operated device (Beta model) and that a large current (Ic) flows freely through the device between the collector and the emitter terminals when the transistor is switched “fully-ON”. However, this only happens when a small biasing current (Ib) is flowing into the base terminal of the transistor at the same time thus allowing the Base to act as a sort of current control input. The current in a bipolar NPN transistor is the ratio of these two currents (Ic/Ib), called the *DC Current Gain* of the device and is given the symbol of hfe or nowadays Beta, (β).

The value of β can be large up to 200 for standard transistors, and it is this large ratio between Ic and Ib that makes the bipolar NPN transistor a useful amplifying device when used in its active region as Ib provides the input and Ic provides the output. Note that Beta has no units as it is a ratio.

Also, the current gain of the transistor from the Collector terminal to the Emitter terminal, Ic/Ie, is called Alpha, (α), and is a function of the transistor itself (electrons diffusing across the junction). As the emitter current Ie is the sum of a very small base current plus a very large collector current, the value of alpha (α), is very close to unity, and for a typical low-power signal transistor this value ranges from about 0.950 to 0.999

α and β Relationship in a NPN Transistor

$$\text{DC Current Gain} = \frac{\text{Output Current}}{\text{Input Current}} = \frac{I_C}{I_B}$$

$$I_E = I_B + I_C \; ..... \; (\text{KCL}) \quad \text{and} \quad \frac{I_C}{I_E} = \alpha$$

$$\text{Thus:} \quad I_B = I_E - I_C$$

$$I_B = I_E - \alpha I_E$$

$$I_B = I_E(1 - \alpha)$$

$$\therefore \quad \beta = \frac{I_C}{I_B} = \frac{I_C}{I_E(1 - \alpha)} = \frac{\alpha}{1 - \alpha}$$

By combining the two parameters α and β we can produce two mathematical expressions that gives the relationship between the different currents flowing in the transistor.

$$\alpha = \frac{\beta}{\beta + 1} \quad \text{or} \quad \alpha = \beta(1 - \alpha)$$

$$\beta = \frac{\alpha}{1 - \alpha} \quad \text{or} \quad \beta = \alpha(1 + \beta)$$

$$\text{If} \quad \alpha = 0.99 \qquad \beta = \frac{0.99}{0.01} = 99$$

The values of Beta vary from about 20 for high current power transistors to well over 1000 for high frequency low power type bipolar transistors. The value of Beta for most standard NPN transistors can be found in the manufactures data sheets but generally range between 50 – 200.

The equation above for Beta can also be re-arranged to make Ic as the subject, and with a zero base current (Ib = 0) the resultant collector current Ic will also be zero, (β*0). Also when the base current is high the corresponding collector current will also be high resulting in the base current controlling the collector current. One of the most important properties of the **Bipolar Junction Transistor** is that a small base current can control a much larger collector current. Consider the following example.

NPN Transistor Example No1

A bipolar NPN transistor has a DC current gain, (Beta) value of 200. Calculate the base current Ib required to switch a resistive load of 4mA.

$$I_B = \frac{I_C}{\beta} = \frac{4 \times 10^{-3}}{200} = 20\mu A$$

Therefore, β = 200, Ic = 4mA and Ib = 20µA.

One other point to remember about **Bipolar NPN Transistors**. The collector voltage, (Vc) must be greater and positive with respect to the emitter voltage, (Ve) to allow current to flow through the transistor between the collector-emitter junctions. Also, there is a voltage drop between the Base and the Emitter terminal of about 0.7V (one diode volt drop) for silicon devices as the input characteristics of an NPN Transistor are of a forward biased diode.

Then the base voltage, (Vbe) of a NPN transistor must be greater than this 0.7V otherwise the transistor will not conduct with the base current given as.

$$I_B = \frac{V_B - V_{BE}}{R_B}$$

Where: Ib is the base current, Vb is the base bias voltage, Vbe is the base-emitter volt drop (0.7v) and Rb is the base input resistor. Increasing Ib, Vbe slowly increases to 0.7V but Ic rises exponentially.

NPN Transistor Example No2

An NPN Transistor has a DC base bias voltage, Vb of 10v and an input base resistor, Rb of 100kΩ. What will be the value of the base current into the transistor.

$$I_B = \frac{V_B - V_{BE}}{R_B} = \frac{10 - 0.7}{100k\Omega} = 93\mu A$$

Therefore, Ib = 93µA.

The Common Emitter Configuration.

As well as being used as a semiconductor switch to turn load currents "ON" or "OFF" by controlling the Base signal to the transistor in either its saturation or cut-off regions, **Bipolar NPN Transistors** can also be used in its active region to produce a circuit which will amplify any small AC signal applied to its Base terminal with the Emitter grounded.

If a suitable DC "biasing" voltage is firstly applied to the transistors Base terminal thus allowing it to always operate within its linear active region, an inverting amplifier circuit called a single stage common emitter amplifier is produced.

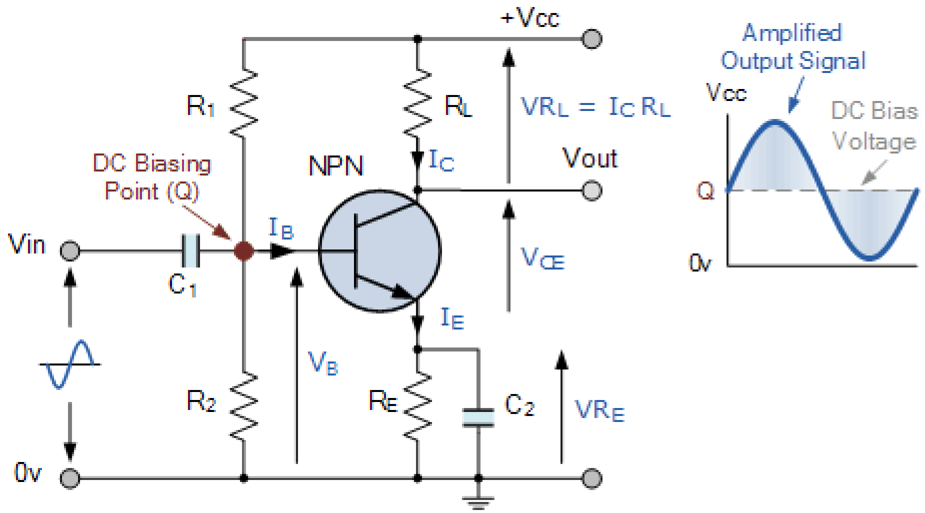
One such *Common Emitter Amplifier* configuration of an NPN transistor is called a Class A Amplifier. A "Class A Amplifier" operation is one where the transistors Base terminal is biased in such a way as to forward bias the Base-emitter junction.

The result is that the transistor is always operating halfway between its cut-off and saturation regions, thereby allowing the transistor amplifier to accurately reproduce the positive and negative halves of any AC input signal superimposed upon this DC biasing voltage.

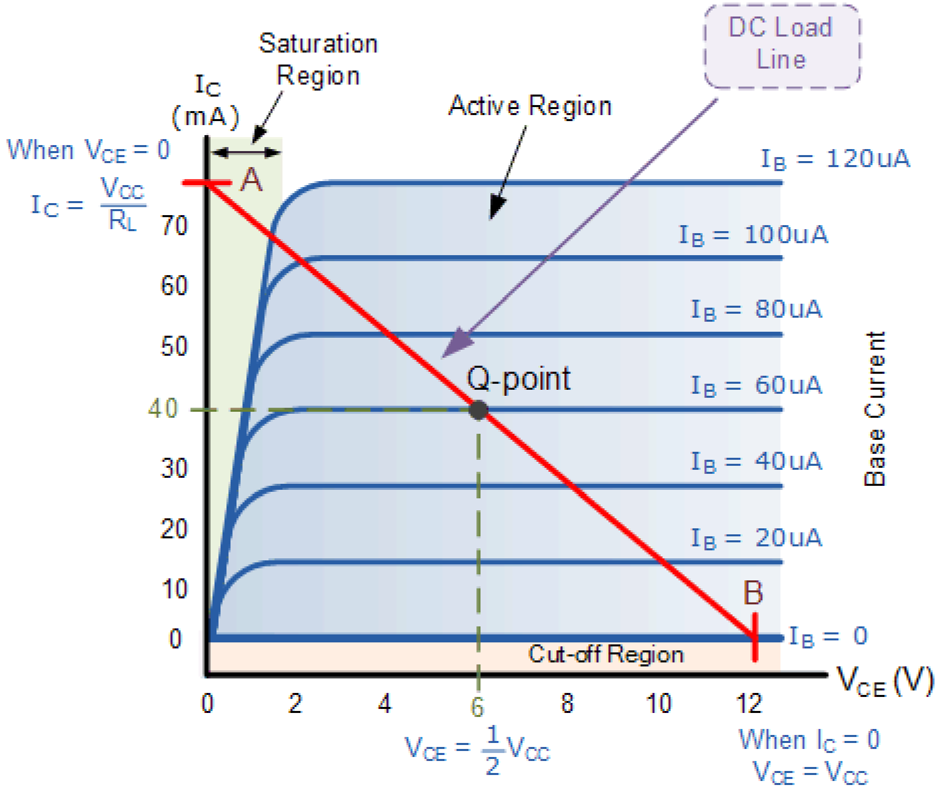
Without this "Bias Voltage" only one half of the input waveform would be amplified. This common emitter amplifier configuration using an NPN transistor has many applications but is commonly used in audio circuits such as pre-amplifier and power amplifier stages.

With reference to the common emitter configuration shown below, a family of curves known as the **Output Characteristics Curves**, relates the output collector current, (Ic) to the collector voltage, (Vce) when different values of Base current, (Ib). Output characteristics curves are applied to the transistor for transistors with the same β value.
A DC "Load Line" can also be drawn onto the output characteristics curves to show all the possible operating points when different values of base current are applied. It is necessary to set the initial value of Vce correctly to allow the output voltage to vary both up and down when amplifying AC input signals and this is called setting the operating point or Quiescent Point, **Q-point** for short and this is shown below.

Single Stage Common Emitter Amplifier Circuit



Output Characteristics Curves of a Typical Bipolar Transistor



The most important factor to notice is the effect of Vce upon the collector current Ic when Vce is greater than about 1.0 volts. We can see that Ic is largely unaffected by changes in Vce above this value and instead it is almost entirely controlled by the base current, Ib. When this happens we can say then that the output circuit represents that of a “Constant Current Source”.

It can also be seen from the common emitter circuit above that the emitter current Ie is the sum of the collector current, Ic and the base current, Ib, added together so we can also say that Ie = Ic + Ib for the common emitter (CE) configuration.

By using the output characteristics curves in our example above and also Ohm´s Law, the current flowing through the load resistor, (R_L), is equal to the collector current, Ic entering the transistor which in turn corresponds to the supply voltage, (Vcc) minus the voltage drop between the collector and the emitter terminals, (Vce) and is given as:

$$\text{Collector Current, } I_C = \frac{V_{CC} - V_{CE}}{R_L}$$

Also, a straight line representing the **Dynamic Load Line** of the transistor can be drawn directly onto the graph of curves above from the point of “Saturation” (A) when Vce = 0 to the point of “Cut-off” (B) when Ic = 0 thus giving us the “Operating” or **Q-point** of the transistor. These two points are joined together by a straight line and any position along this straight line represents the “Active Region” of the transistor. The actual position of the load line on the characteristics curves can be calculated as follows:

$$\text{When: } \left(V_{CE} = 0\right) \quad I_C = \frac{V_{CC} - 0}{R_L}, \quad I_C = \frac{V_{CC}}{R_L}$$

$$\text{When: } \left(I_C = 0\right) \quad 0 = \frac{V_{CC} - V_{CE}}{R_L}, \quad V_{CC} = V_{CE}$$

Then, the collector or output characteristics curves for **Common Emitter NPN Transistors** can be used to predict the Collector current, Ic, when given Vce and the Base current, Ib. A Load Line can also be constructed onto the curves to determine a suitable Operating or **Q-point** which can be set by adjustment of the base current. The slope of this load line is equal to the reciprocal of the load resistance which is given as: $-1/R_L$

Then we can define a **NPN Transistor** as being normally “OFF” but a small input current and a small positive voltage at its Base (B) relative to its Emitter (E) will turn it “ON” allowing a much large Collector-Emitter current to flow. NPN transistors conduct when Vc is much greater than Ve.

PNP Transistor

The **PNP Transistor** is the exact opposite to the **NPN Transistor** device we looked at in the previous tutorial.

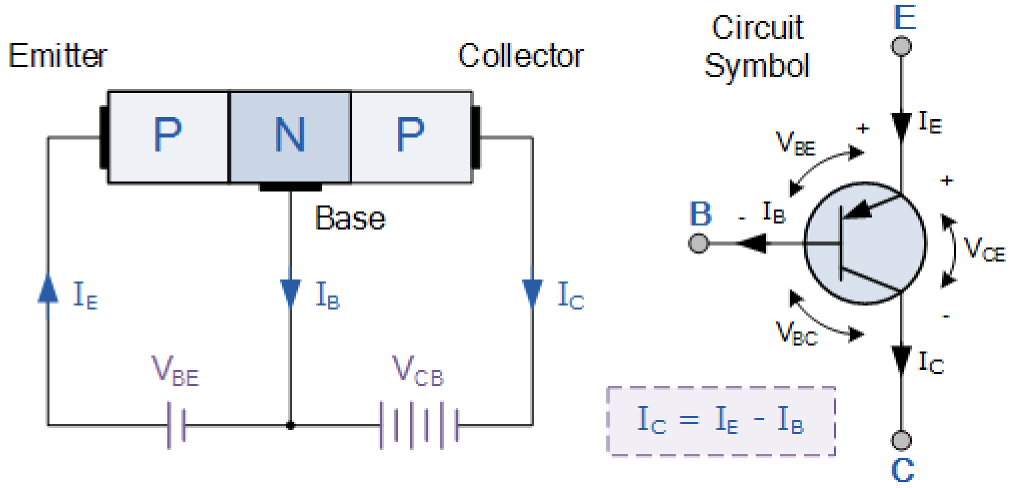
Basically, in this type of transistor construction the two diodes are reversed with respect to the NPN type giving a **P**ositive-**N**egative-**P**ositive type of configuration, with the arrow which also defines the Emitter terminal this time pointing inwards in the transistor symbol.

Also, all the polarities for a *PNP transistor* are reversed which means that it “sinks” current into its Base as opposed to the NPN transistor which “sources” current through its Base. The main difference between the two types of transistors is that holes are the more important carriers for PNP transistors, whereas electrons are the important carriers for NPN transistors.

Then, PNP transistors use a small base current and a negative base voltage to control a much larger emitter-collector current. In other words for a PNP transistor, the Emitter is more positive with respect to the Base and also with respect to the Collector.

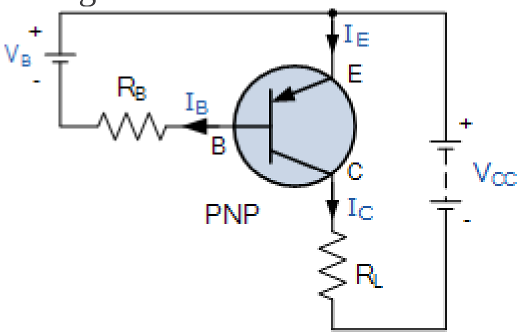
The construction of a “PNP transistor” consists of two P-type semiconductor materials either side of an N-type material as shown below.

A PNP Transistor Configuration



(Note: Arrow defines the emitter and conventional current flow, "in" for a PNP transistor.)

The construction and terminal voltages for an NPN transistor are shown above. The **PNP Transistor** has very similar characteristics to their NPN bipolar cousins, except that the polarities (or biasing) of the current and voltage directions are reversed for any one of the possible three configurations looked at in the first tutorial, Common Base, Common Emitter and Common Collector.



PNP Transistor Connection

The voltage between the Base and Emitter (V_{BE}), is now negative at the Base and positive at the Emitter because for a PNP transistor, the Base terminal is always biased negative with respect to the Emitter.

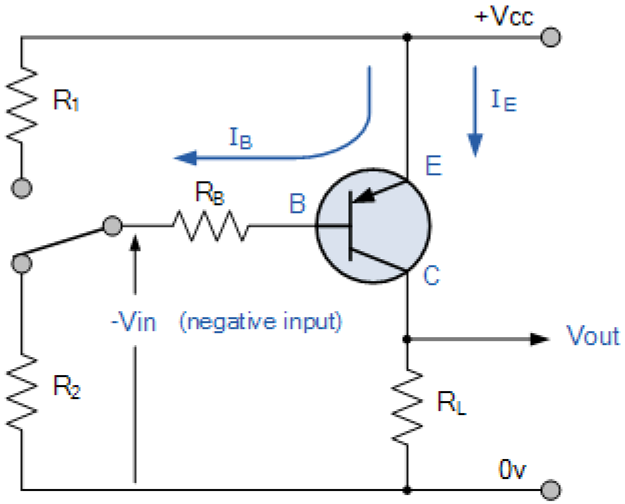
Also the Emitter supply voltage is positive with respect to the Collector (V_{CE}). So for a PNP transistor to conduct the Emitter is always more positive with respect to both the Base and the Collector.
The voltage sources are connected to a PNP transistor are as shown. This time the Emitter is connected to the supply voltage V_{CC} with the load resistor, RL which limits the maximum current flowing through the device connected to the Collector terminal. The Base voltage V_B which is biased negative with respect to the Emitter and is connected to the Base resistor R_B, which again is used to limit the maximum Base current.
To cause the Base current to flow in a PNP transistor the Base needs to be more negative than the Emitter (current must leave the base) by approx 0.7 volts for a silicon device or 0.3 volts for a germanium device with the formulas used to calculate the Base resistor, Base current or Collector current are the same as those used for an equivalent NPN transistor and is given as.

$$I_C = I_E - I_B$$

$$I_C = \beta . I_B \qquad I_B = \frac{I_C}{\beta}$$

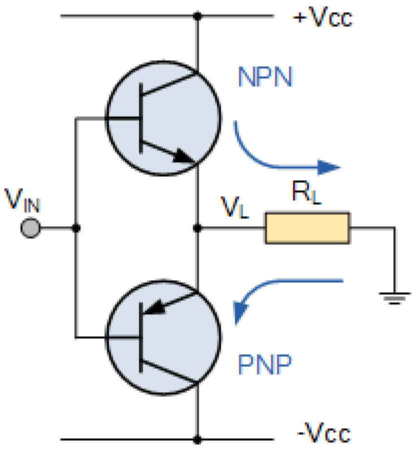
We can see that the fundamental differences between a NPN Transistor and a PNP Transistor is the proper biasing of the transistors junctions as the current directions and voltage polarities are always opposite to each other. So for the circuit above: Ic = Ie – Ib as current must leave the Base.
Generally, the PNP transistor can replace NPN transistors in most electronic circuits, the only difference is the polarities of the voltages, and the directions of the current flow. PNP transistors can also be used as switching devices and an example of a PNP transistor switch is shown below.

A PNP Transistor Circuit



The **Output Characteristics Curves** for a PNP transistor look very similar to those for an equivalent NPN transistor except that they are rotated by 180° to take account of the reverse polarity voltages and currents, (that is for a PNP transistor, electron current flows out of the base and collector towards the battery). The same dynamic load line can be drawn onto the I-V curves to find the PNP transistors operating points.

Transistor Matching



Complementary Transistors

You may think what is the point of having a **PNP Transistor**, when there are plenty of NPN Transistors available that can be used as an amplifier or solid-state switch?. Well, having two different types of transistors "PNP" and "NPN", can be a great advantage when designing power amplifier circuits such as the Class B Amplifier.
Class-B amplifiers uses "Complementary" or "Matched Pair" (that is one PNP and one NPN connected together) transistors in its output stage or in reversible H-Bridge motor control circuits where we want to control the flow of current evenly through the motor in both directions at different times for forward and reverse motion.
A pair of corresponding NPN and PNP transistors with near identical characteristics to each other are called **Complementary Transistors** for example, a TIP3055 (NPN transistor) and the TIP2955 (PNP transistor) are good examples of complementary or matched pair silicon power transistors. They both have a DC current gain, Beta, (Ic/Ib) matched to within 10% and high Collector current of about 15A making them ideal for general motor control or robotic applications.
Also, class B amplifiers use complementary NPN and PNP in their power output stage design. The NPN transistor conducts for only the positive half of the signal while the PNP transistor conducts for negative half of the signal.
This allows the amplifier to drive the required power through the load loudspeaker in both directions at the stated nominal impedance and power resulting in an output current which is likely to be in the order of several amps shared evenly between the two complementary transistors.

Identifying the PNP Transistor

We saw in the first tutorial of this transistors section, that transistors are basically made up of two Diodes connected together back-to-back. We can use this analogy to determine whether a transistor is of the PNP type or NPN type by testing its Resistance between the three different leads, Emitter, Base and Collector. By testing each pair of transistor leads in both directions with a multimeter will result in six tests in total with the expected resistance values in Ohm's given below.

- 1. Emitter-Base Terminals – The Emitter to Base should act like a normal diode and conduct one way only.
- 2. Collector-Base Terminals – The Collector-Base junction should act like a normal diode and conduct one way only.
- 3. Emitter-Collector Terminals – The Emitter-Collector should not conduct in either direction.

Terminal Resistance Values for PNP and NPN Transistors

Between Transistor Terminals		PNP	NPN
Collector	Emitter	R_{HIGH}	R_{HIGH}
Collector	Base	R_{LOW}	R_{HIGH}

Emitter	Collector	R_{HIGH}	R_{HIGH}
Emitter	Base	R_{LOW}	R_{HIGH}
Base	Collector	R_{HIGH}	R_{LOW}
Base	Emitter	R_{HIGH}	R_{LOW}

Then we can define a **PNP Transistor** as being normally "OFF" but a small output current and negative voltage at its Base (B) relative to its Emitter (E) will turn it "ON" allowing a much large Emitter-Collector current to flow. PNP transistors conduct when Ve is much greater than Vc.

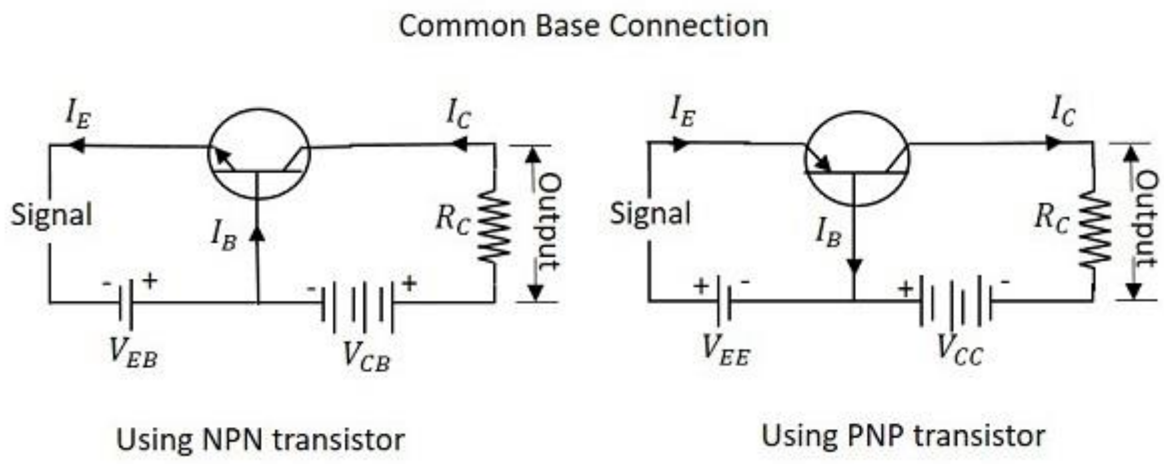
In other words, a **Bipolar PNP Transistor** will ONLY conduct if both the Base and Collector terminals are negative with respect to the Emitter

BJT CONFIGURATIONS

The three types of configurations are **Common Base, Common Emitter** and **Common Collector** configurations. In every configuration, the emitter junction is forward biased and the collector junction is reverse biased.

Common Base CBCB Configuration

The name itself implies that the Base terminal is taken as common terminal for both input and output of the transistor. The common base connection for both NPN and PNP transistors is as shown in the following figure.



For the sake of understanding, let us consider NPN transistor in CB configuration. When the emitter voltage is applied, as it is forward biased, the electrons from the negative terminal repel the emitter electrons and current flows through the emitter and base to the collector to contribute collector current. The collector voltage $\mathbf{V_{CB}}$ is kept constant throughout this.

In the CB configuration, the input current is the emitter current $\mathbf{I_E}$ and the output current is the collector current $\mathbf{I_C}$.

Current Amplification Factor αα

The ratio of change in collector current $ΔI_C$$ $ΔIC$ to the change in emitter current $ΔI_E$$ $ΔIE$ when collector voltage $\mathbf{V_{CB}}$ is kept constant, is called as **Current amplification factor**. It is denoted by α.

$$\alpha = \frac{\Delta I_C}{\Delta I_E} \text{at constant} V_{CB} \quad \alpha=\Delta IC\Delta IE \text{atconstant} VCB$$

Expression for Collector current

With the idea above, let us try to draw some expression for collector current. Along with the emitter current flowing, there is some amount of base current IB which flows through the base terminal due to electron hole recombination. As collector-base junction is reverse biased, there is another current which is flown due to minority charge carriers. This is the leakage current which can be understood as $\mathbf{I_{leakage}}$. This is due to minority charge carriers and hence very small.

The emitter current that reaches the collector terminal is

$$\alpha I_E \quad \alpha IE$$

Total collector current

$$I_C = \alpha I_E + I_{leakage} \quad IC=\alpha IE+Ileakage$$

If the emitter-base voltage V_{EB} = 0, even then, there flows a small leakage current, which can be termed as I_{CBO} collector−basecurrentwithoutputopencollector−basecurrentwithoutputopen.

The collector current therefore can be expressed as

$$I_C = \alpha I_E + I_{CBO}$$

$$I_E = I_C + I_B$$

$$I_C = \alpha (I_C + I_B) + I_{CBO}$$

$$I_C (1-\alpha) = \alpha I_B + I_{CBO}$$

$$I_C = \left(\frac{\alpha}{1-\alpha}\right) I_B + \left(\frac{I_{CBO}}{1-\alpha}\right)$$

$$I_C = \left(\frac{\alpha}{1-\alpha}\right) I_B + \left(\frac{1}{1-\alpha}\right) I_{CBO}$$

Hence the above derived is the expression for collector current. The value of collector current depends on base current and leakage current along with the current amplification factor of that transistor in use.

Characteristics of CB configuration

- This configuration provides voltage gain but no current gain.
- Being $\mathbf{V_{CB}}$ constant, with a small increase in the Emitter-base voltage V_{EB}, Emitter current $\mathbf{I_E}$ gets increased.
- Emitter Current $\mathbf{I_E}$ is independent of Collector voltage $\mathbf{V_{CB}}$.
- Collector Voltage $\mathbf{V_{CB}}$ can affect the collector current $\mathbf{I_C}$ only at low voltages, when V_{EB} is kept constant.
- The input resistance ri is the ratio of change in emitter-base voltage ΔV_{EB} to the change in emitter current ΔI_E at constant collector base voltage $\mathbf{V_{CB}}$.

$$\eta = \frac{\Delta V_{EB}}{\Delta I_E} \text{ at constant } V_{CB}$$

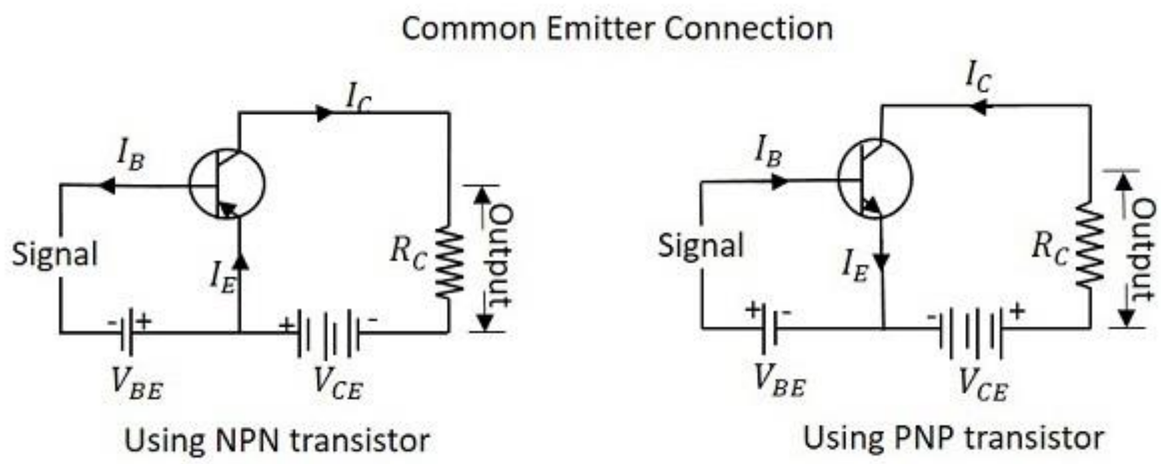
- As the input resistance is of very low value, a small value of V_{EB} is enough to produce a large current flow of emitter current $\mathbf{I_E}$.
- The output resistance r_o is the ratio of change in the collector base voltage ΔV_{CB} to the change in collector current ΔI_C at constant emitter current $\mathbf{I_E}$.

$$r_o = \frac{\Delta V_{CB}}{\Delta I_C} \text{ at constant } I_E$$

- As the output resistance is of very high value, a large change in $\mathbf{V_{CB}}$ produces a very little change in collector current $\mathbf{I_C}$.
- This Configuration provides good stability against increase in temperature.
- The CB configuration is used for high frequency applications.

Common Emitter (CE) Configuration

The name itself implies that the **Emitter** terminal is taken as common terminal for both input and output of the transistor. The common emitter connection for both NPN and PNP transistors is as shown in the following figure.



Just as in CB configuration, the emitter junction is forward biased and the collector junction is reverse biased. The flow of electrons is controlled in the same manner. The input current is the base current $\mathbf{I_B}$ and the output current is the collector current $\mathbf{I_C}$ here.

Base Current Amplification factor ββ

The ratio of change in collector current $\$\Delta I_C\$\$\Delta IC\$$ to the change in base current $\$\Delta I_B\$\$\Delta IB\$$ is known as **Base Current Amplification Factor**. It is denoted by β

$$\beta=\Delta I_C\Delta I_B\beta=\Delta IC\Delta IB$$

Relation between β and α

Let us try to derive the relation between base current amplification factor and emitter current amplification factor.

$$\beta=\Delta I_C\Delta I_B\beta=\Delta IC\Delta IB$$

$$\alpha=\Delta I_C\Delta I_E\alpha=\Delta IC\Delta IE$$

$$I_E=I_B+I_CIE=IB+IC$$

$$\Delta I_E=\Delta I_B+\Delta I_C\Delta IE=\Delta IB+\Delta IC$$

$$\Delta I_B=\Delta I_E-\Delta I_C\Delta IB=\Delta IE-\Delta IC$$

We can write

$$\beta=\Delta I_C\Delta I_E-\Delta I_C\beta=\Delta IC\Delta IE-\Delta IC$$

Dividing by \$\$

$$\beta=\Delta I_C\Delta I_E\Delta I_E\Delta I_E-\Delta I_C\Delta I_E\beta=\Delta IC\Delta IE\Delta IE\Delta IE-\Delta IC\Delta IE$$

$$\alpha=\Delta I_C\Delta I_E\alpha=\Delta IC\Delta IE$$

We have

$$\alpha=\Delta I_C\Delta I_E\alpha=\Delta IC\Delta IE$$

Therefore,

$$\beta=\alpha1-\alpha\beta=\alpha1-\alpha$$

From the above equation, it is evident that, as α approaches 1, β reaches infinity.
Hence, **the current gain in Common Emitter connection is very high**. This is the reason this circuit connection is mostly used in all transistor applications.

Expression for Collector Current

In the Common Emitter configuration, $\mathbf{I_B}$ is the input current and $\mathbf{I_C}$ is the output current.
We know

$$I_E=I_B+I_CIE=IB+IC$$

And

$$I_C=\alpha I_E+I_{CBO}IC=\alpha IE+ICBO$$

$$=\alpha(I_B+I_C)+I_{CBO}=\alpha(IB+IC)+ICBO$$

$$I_C(1-\alpha)=\alpha I_B+I_{CBO}IC(1-\alpha)=\alpha IB+ICBO$$

$$I_C=\alpha1-\alpha I_B+11-\alpha I_{CBO}IC=\alpha1-\alpha IB+11-\alpha ICBO$$

If base circuit is open, i.e. if $\mathbf{I_B} = 0$,
The collector emitter current with base open is I_{CEO}

$$I_{CEO}=11-\alpha I_{CBO}ICEO=11-\alpha ICBO$$

Substituting the value of this in the previous equation, we get

$$I_C=\alpha1-\alpha I_B+I_{CEO}IC=\alpha1-\alpha IB+ICEO$$

$$I_C=\beta I_B+I_{CEO}IC=\beta IB+ICEO$$

Hence the equation for collector current is obtained.

Knee Voltage

In CE configuration, by keeping the base current $\mathbf{I_B}$ constant, if $\mathbf{V_{CE}}$ is varied, $\mathbf{I_C}$ increases nearly to 1v of $\mathbf{V_{CE}}$ and stays constant thereafter. This value of $\mathbf{V_{CE}}$ up to which collector current $\mathbf{I_C}$ changes with $\mathbf{V_{CE}}$ is called the **Knee Voltage**. The transistors while operating in CE configuration, they are operated above this knee voltage.

Characteristics of CE Configuration

- This configuration provides good current gain and voltage gain.
- Keeping $\mathbf{V_{CE}}$ constant, with a small increase in $\mathbf{V_{BE}}$ the base current $\mathbf{I_B}$ increases rapidly than in CB configurations.

- For any value of V_{CE} above knee voltage, I_C is approximately equal to βI_B.
- The input resistance r_i is the ratio of change in base emitter voltage \$$\Delta V_{BE}$\$\$ΔVBE\$ to the change in base current \$$\Delta I_B$\$\$ΔIB\$ at constant collector emitter voltage V_{CE}.

$$r_i = \Delta V_{BE} \Delta I_B \text{ at constant } V_{CE}$$ ri=ΔVBEΔIBatconstantVCE

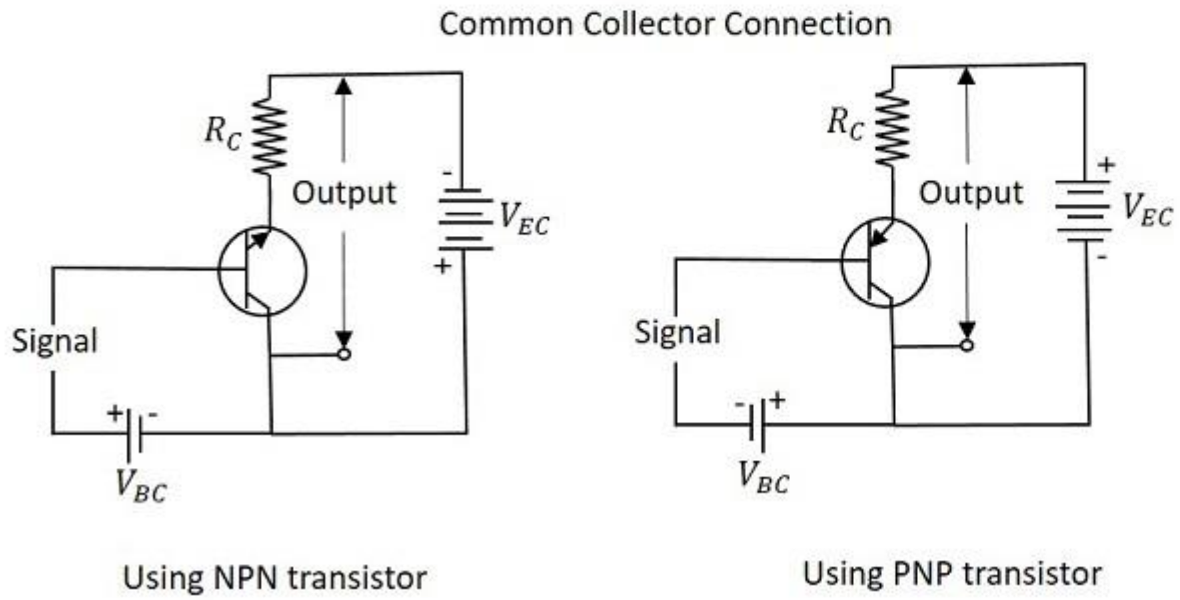
- As the input resistance is of very low value, a small value of V_{BE} is enough to produce a large current flow of base current I_B.
- The output resistance r_o is the ratio of change in collector emitter voltage \$$\Delta V_{CE}$\$\$ΔVCE\$ to the change in collector current \$$\Delta I_C$\$\$ΔIC\$ at constant I_B.

$$r_o = \Delta V_{CE} \Delta I_C \text{ at constant } I_B$$ ro=ΔVCEΔICatconstantIB

- As the output resistance of CE circuit is less than that of CB circuit.
- This configuration is usually used for bias stabilization methods and audio frequency applications.

Common Collector CCCC Configuration

The name itself implies that the **Collector** terminal is taken as common terminal for both input and output of the transistor. The common collector connection for both NPN and PNP transistors is as shown in the following figure.



Just as in CB and CE configurations, the emitter junction is forward biased and the collector junction is reverse biased. The flow of electrons is controlled in the same manner. The input current is the base current I_B and the output current is the emitter current I_E here.

Current Amplification Factor γ

The ratio of change in emitter current ΔI_E to the change in base current ΔI_B is known as **Current Amplification factor** in common collector CC configuration. It is denoted by $\boldsymbol{\gamma}$.

$$\gamma = \frac{\Delta I_E}{\Delta I_B}$$

- The current gain in CC configuration is same as in CE configuration.
- The voltage gain in CC configuration is always less than 1.

Relation between γ and α

Let us try to draw some relation between γ and α

$$\gamma = \frac{\Delta I_E}{\Delta I_B}$$

$$\alpha = \frac{\Delta I_C}{\Delta I_E}$$

$$I_E = I_B + I_C$$

$$\Delta I_E = \Delta I_B + \Delta I_C$$

$$\Delta I_B = \Delta I_E - \Delta I_C$$

Substituting the value of I_B, we get

$$\gamma = \frac{\Delta I_E}{\Delta I_E - \Delta I_C}$$

Dividing by ΔI_E

$$\gamma = \frac{\frac{\Delta I_E}{\Delta I_E}}{\frac{\Delta I_E}{\Delta I_E} - \frac{\Delta I_C}{\Delta I_E}}$$

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

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

Expression for collector current

We know

$$I_C = \alpha I_E + I_{CBO}$$

$$I_E = I_B + I_C = I_B + (\alpha I_E + I_{CBO})$$

$$I_E(1-\alpha) = I_B + I_{CBO}$$

$$I_E = \frac{I_B}{1-\alpha} + \frac{I_{CBO}}{1-\alpha}$$

$$I_C \cong I_E = (\beta+1) I_B + (\beta+1) I_{CBO}$$

The above is the expression for collector current.

Characteristics of CC Configuration

- This configuration provides current gain but no voltage gain.
- In CC configuration, the input resistance is high and the output resistance is low.
- The voltage gain provided by this circuit is less than 1.
- The sum of collector current and base current equals emitter current.
- The input and output signals are in phase.
- This configuration works as non-inverting amplifier output.
- This circuit is mostly used for impedance matching. That means, to drive a low impedance load from a high impedance source.

Chapter 6: Field Effect Transistors

A Field Effect Transistor (FET) is a three-terminal semiconductor device. Its operation is based on a controlled input voltage. By appearance JFET and bipolar transistors are very similar. However, BJT is a current controlled device and JFET is controlled by input voltage. Most commonly two types of FETs are available.

- Junction Field Effect Transistor (JFET)
- Metal Oxide Semiconductor FET (IGFET)

Junction Field Effect Transistor

The functioning of Junction Field Effect Transistor depends upon the flow of majority carriers (electrons or holes) only. Basically, JFETs consist of an **N** type or **P** type silicon bar containing PN junctions at the sides. Following are some important points to remember about FET −

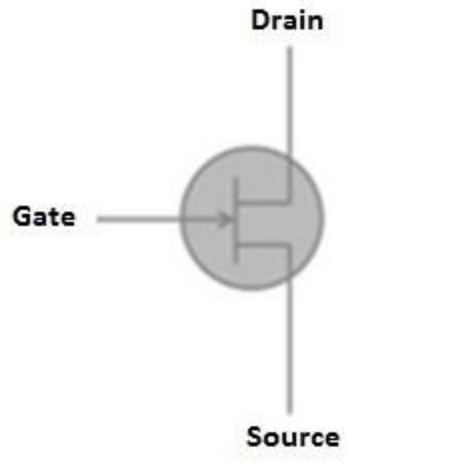
- **Gate** − By using diffusion or alloying technique, both sides of N type bar are heavily doped to create PN junction. These doped regions are called gate (G).
- **Source** − It is the entry point for majority carriers through which they enter into the semiconductor bar.
- **Drain** − It is the exit point for majority carriers through which they leave the semiconductor bar.
- **Channel** − It is the area of N type material through which majority carriers pass from the source to drain.

There are two types of JFETs commonly used in the field semiconductor devices: **N-Channel JFET** and **P-Channel JFET**.

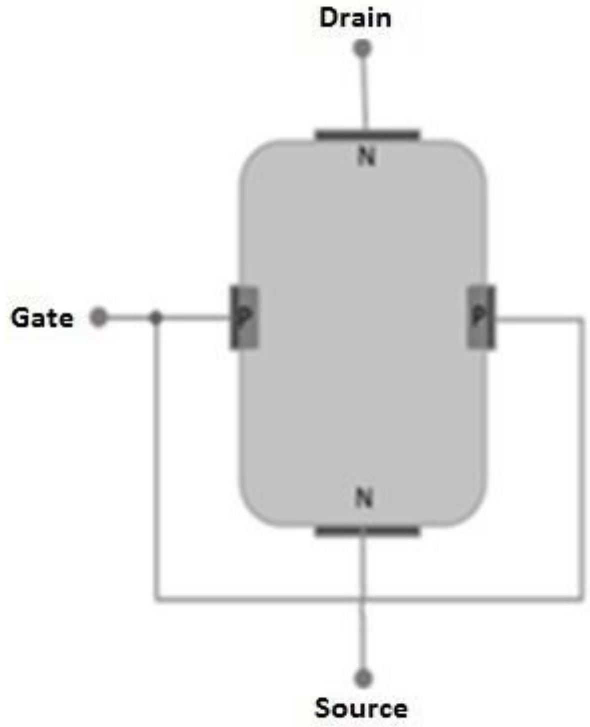
N-Channel JFET

It has a thin layer of N type material formed on P type substrate. Following figure shows the crystal structure and schematic symbol of an N-channel JFET. Then the gate is formed on top of the N channel with P type material. At the end of the channel and the gate, lead wires are attached and the substrate has no connection. When a DC voltage source is connected to the source and the drain leads of a JFET, maximum current will flow through the channel. The same amount of current will flow from the source and the drain terminals. The amount of channel current flow will be determined by the value of V_{DD} and the internal resistance of the channel.

A typical value of source-drain resistance of a JFET is quite a few hundred ohms. It is clear that even when the gate is open full current conduction will take place in the channel. Essentially, the amount of bias voltage applied at ID, controls the flow of current carriers passing through the channel of a JFET. With a small change in gate voltage, JFET can be controlled anywhere between full conduction and cutoff state.



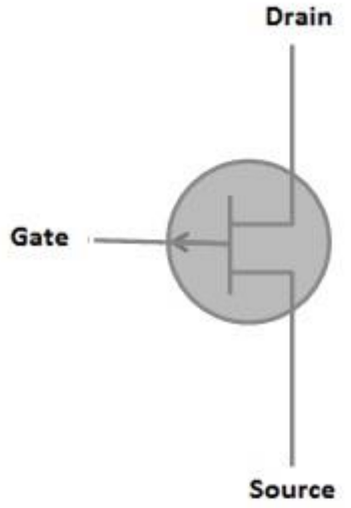
Symbol of N-Channel JFET



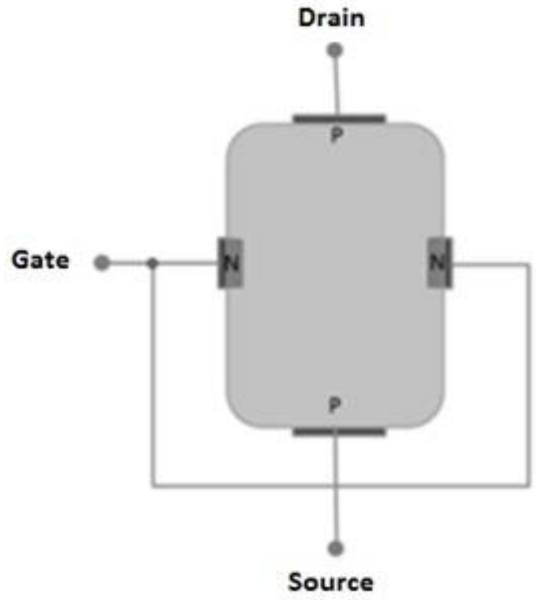
N-Channel JFET

P-Channel JFETs

It has a thin layer of P type material formed on N type substrate. The following figure shows the crystal structure and schematic symbol of an N-channel JFET. The gate is formed on top of the P channel with N type material. At the end of the channel and the gate, lead wires are attached. Rest of the construction details are similar to that of N- channel JFET.



Symbol of P-Channel JFET

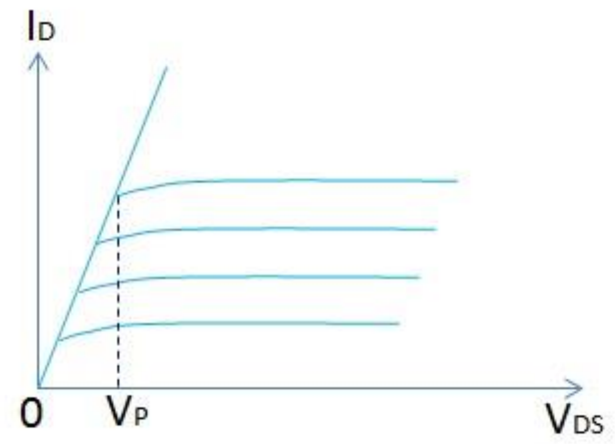


P-Channel JFET

Normally for general operation, the gate terminal is made positive with respect to the source terminal. The size of the P-N junction depletion layer depends upon fluctuations in the values of reverse biased gate voltage. With a small change in gate voltage, JFET can be controlled anywhere between full conduction and cutoff state.

Output Characteristics of JFET

The output characteristics of JFET are drawn between drain current (I_D) and drain source voltage (V_{DS}) at constant gate source voltage (V_{GS}) as shown in the following figure.



Initially, the drain current (I_D) rises rapidly with drain source voltage (V_{DS}) however suddenly becomes constant at a voltage known as pinch-off voltage (V_P). Above pinch-off voltage, the channel width becomes so narrow that it allows very small drain current to pass through it. Therefore, drain current (I_D) remains constant above pinch-off voltage.

Parameters of JFET

The main parameters of JFET are −

- AC drain resistance (Rd)
- Transconductance
- Amplification factor

AC drain resistance (R_d) − It is the ratio of change in the drain source voltage (ΔV_{DS}) to the change in drain current (ΔI_D) at constant gate-source voltage. It can be expressed as,

$R_d = (\Delta V_{DS})/(\Delta I_D)$ at Constant V_{GS}

Transconductance (g_{fs}) − It is the ratio of change in drain current (ΔI_D) to the change in gate source voltage (ΔV_{GS}) at constant drain-source voltage. It can be expressed as,

$g_{fs} = (\Delta I_D)/(\Delta V_{GS})$ at constant V_{DS}

Amplification Factor (u) − It is the ratio of change in drain-source voltage (ΔV_{DS}) to the change in gate source voltage (ΔV_{GS}) constant drain current (ΔI_D). It can be expressed as,

$u = (\Delta V_{DS})/(\Delta V_{GS})$ at constant I_D

The MOSFET

MOSFET's operate the same as JFET's but have a gate terminal that is electrically isolated from the conductive channel.

As well as the Junction Field Effect Transistor (JFET), there is another type of Field Effect Transistor available whose Gate input is electrically insulated from the main current carrying channel and is therefore called an **Insulated Gate Field Effect Transistor**.

The most common type of insulated gate FET which is used in many different types of electronic circuits is called the **Metal Oxide Semiconductor Field Effect Transistor** or **MOSFET** for short.

The **IGFET** or **MOSFET** is a voltage controlled field effect transistor that differs from a JFET in that it has a "Metal Oxide" Gate electrode which is electrically insulated from the main semiconductor n-channel or p-channel by a very thin layer of insulating material usually silicon dioxide, commonly known as glass.

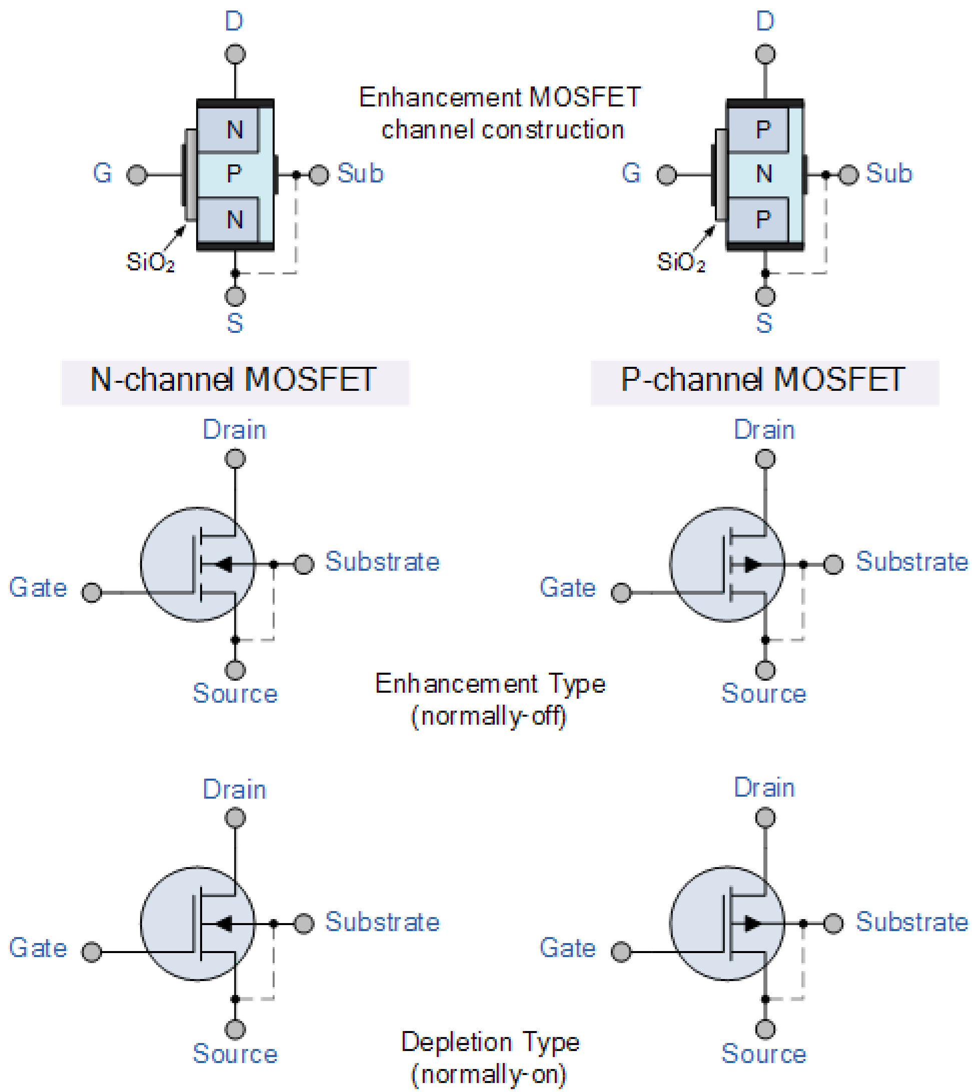
This ultra thin insulated metal gate electrode can be thought of as one plate of a capacitor. The isolation of the controlling Gate makes the input resistance of the **MOSFET** extremely high way up in the Mega-ohms (MΩ) region thereby making it almost infinite.

As the Gate terminal is electrically isolated from the main current carrying channel between the drain and source, "NO current flows into the gate" and just like the JFET, the MOSFET also acts like a voltage controlled resistor where the current flowing through the main channel between the Drain and Source is proportional to the input voltage. Also like the JFET, the MOSFETs very high input resistance can easily accumulate large amounts of static charge resulting in the **MOSFET** becoming easily damaged unless carefully handled or protected.

Like the previous JFET tutorial, MOSFETs are three terminal devices with a Gate, Drain and Source and both P-channel (PMOS) and N-channel (NMOS) MOSFETs are available. The main difference this time is that MOSFETs are available in two basic forms:

- Depletion Type – the transistor requires the Gate-Source voltage, (V_{GS}) to switch the device "OFF". The depletion mode MOSFET is equivalent to a "Normally Closed" switch.
- Enhancement Type – the transistor requires a Gate-Source voltage, (V_{GS}) to switch the device "ON". The enhancement mode MOSFET is equivalent to a "Normally Open" switch.

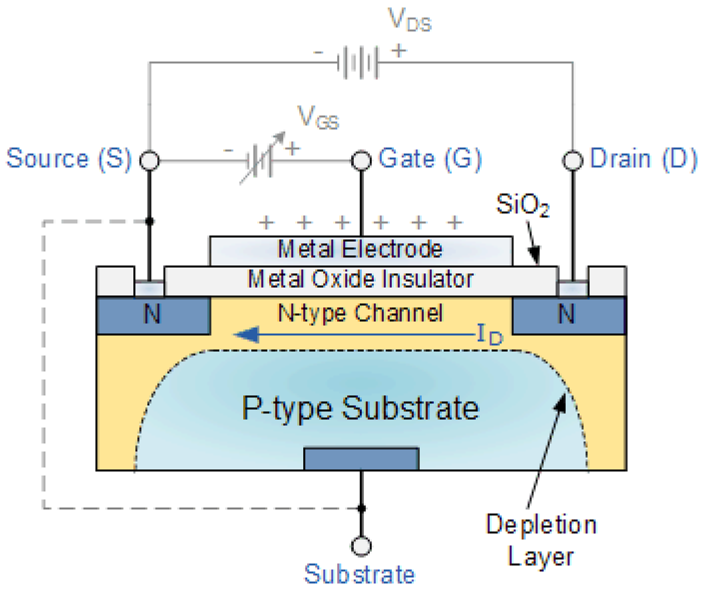
The symbols and basic construction for both configurations of MOSFETs are shown below.



The four MOSFET symbols above show an additional terminal called the Substrate and is not normally used as either an input or an output connection but instead it is used for grounding the substrate. It connects to the main semiconductive channel through a diode junction to the body or metal tab of the MOSFET.

Usually in discrete type MOSFETs, this substrate lead is connected internally to the source terminal. When this is the case, as in enhancement types it is omitted from the symbol for clarification.
The line in the MOSFET symbol between the drain (D) and source (S) connections represents the transistors semiconductive channel. If this channel line is a solid unbroken line then it represents a “Depletion” (normally-ON) type MOSFET as drain current can flow with zero gate biasing potential.
If the channel line is shown as a dotted or broken line, then it represents an “Enhancement” (normally-OFF) type MOSFET as zero drain current flows with zero gate potential. The direction of the arrow pointing to this channel line indicates whether the conductive channel is a P-type or an N-type semiconductor device.

Basic MOSFET Structure and Symbol



The construction of the Metal Oxide Semiconductor FET is very different to that of the Junction FET. Both the Depletion and Enhancement type MOSFETs use an electrical field produced by a gate voltage to alter the flow of charge carriers, electrons for n-channel or holes for P-channel, through the semiconductive drain-source channel. The gate electrode is placed on top of a very thin insulating layer and there are a pair of small n-type regions just under the drain and source electrodes.
We saw in the previous tutorial, that the gate of a junction field effect transistor, JFET must be biased in such a way as to reverse-bias the pn-junction. With a insulated gate MOSFET device no such limitations apply so it is possible to bias the gate of a MOSFET in either polarity, positive (+ve) or negative (-ve).

This makes the MOSFET device especially valuable as electronic switches or to make logic gates because with no bias they are normally non-conducting and this high gate input resistance means that very little or no control current is needed as MOSFETs are voltage controlled devices. Both the p-channel and the n-channel MOSFETs are available in two basic forms, the **Enhancement** type and the **Depletion** type.

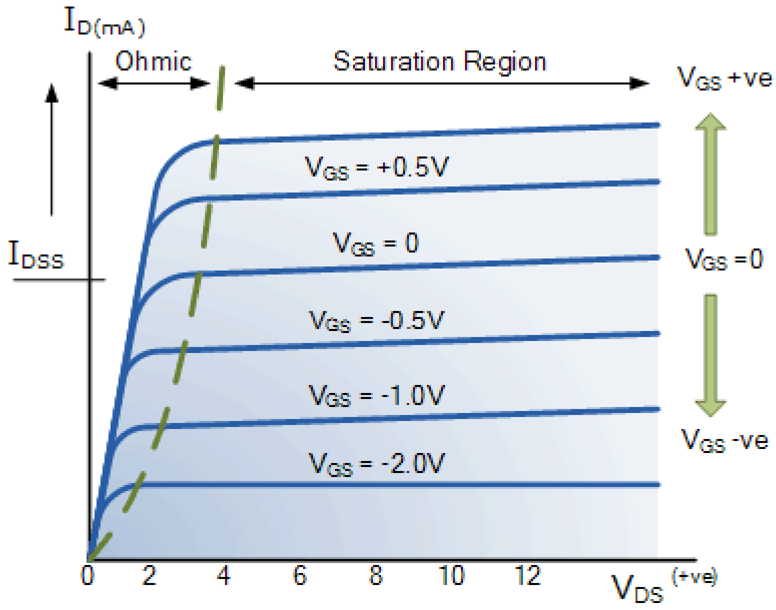
Depletion-mode MOSFET

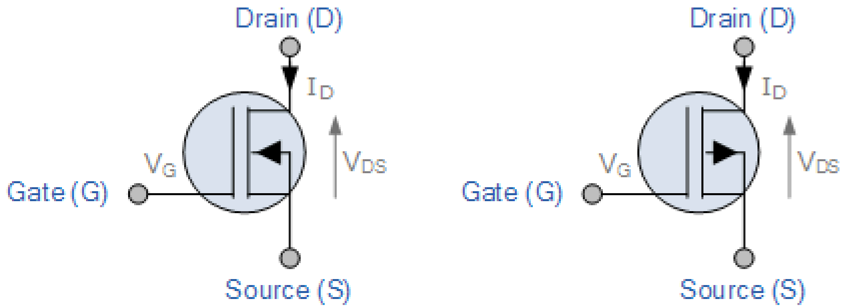
The **Depletion-mode MOSFET**, which is less common than the enhancement mode types is normally switched "ON" (conducting) without the application of a gate bias voltage. That is the channel conducts when $V_{GS} = 0$ making it a "normally-closed" device. The circuit symbol shown above for a depletion MOS transistor uses a solid channel line to signify a normally closed conductive channel.

For the n-channel depletion MOS transistor, a negative gate-source voltage, $-V_{GS}$ will deplete (hence its name) the conductive channel of its free electrons switching the transistor "OFF". Likewise for a p-channel depletion MOS transistor a positive gate-source voltage, $+V_{GS}$ will deplete the channel of its free holes turning it "OFF".

In other words, for an n-channel depletion mode MOSFET: $+V_{GS}$ means more electrons and more current. While a $-V_{GS}$ means less electrons and less current. The opposite is also true for the p-channel types. Then the depletion mode MOSFET is equivalent to a "normally-closed" switch.

Depletion-mode N-Channel MOSFET and circuit Symbols





N-channel P-channel

The depletion-mode MOSFET is constructed in a similar way to their JFET transistor counterparts were the drain-source channel is inherently conductive with the electrons and holes already present within the n-type or p-type channel. This doping of the channel produces a conducting path of low resistance between the Drain and Source with zero Gate bias.

Enhancement-mode MOSFET

The more common **Enhancement-mode MOSFET** or eMOSFET, is the reverse of the depletion-mode type. Here the conducting channel is lightly doped or even undoped making it non-conductive. This results in the device being normally "OFF" (non-conducting) when the gate bias voltage, V_{GS} is equal to zero. The circuit symbol shown above for an enhancement MOS transistor uses a broken channel line to signify a normally open non-conducting channel.

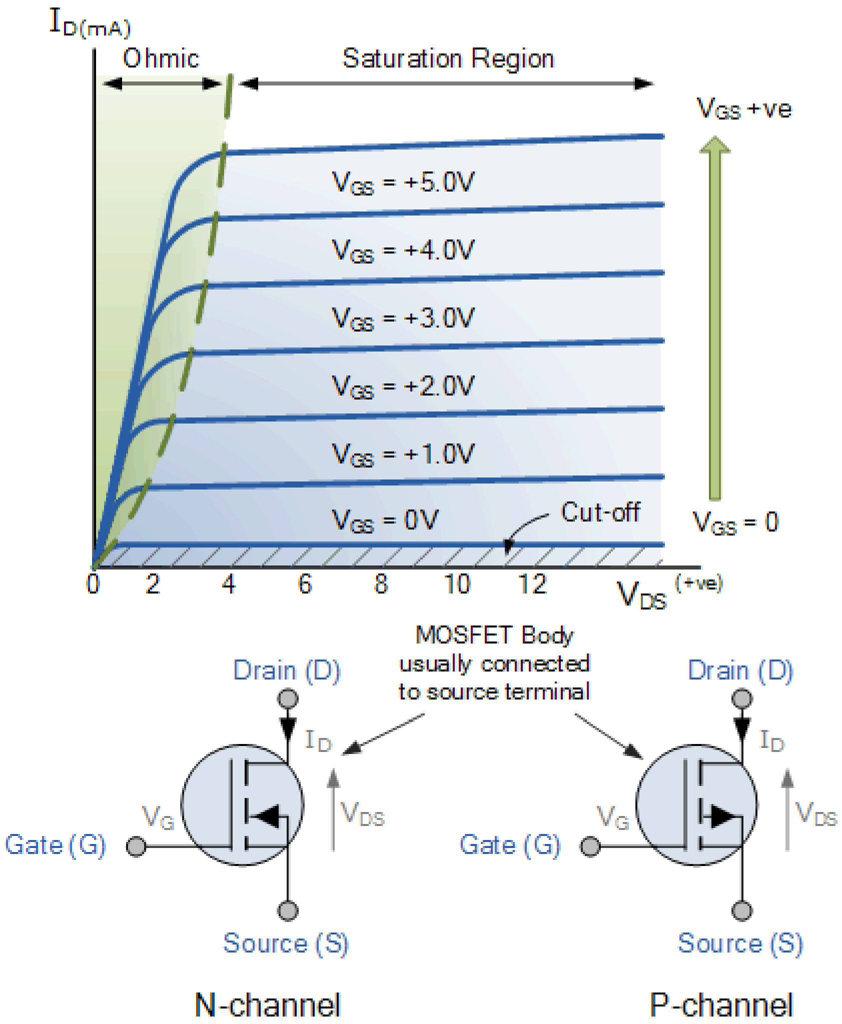
For the n-channel enhancement MOS transistor a drain current will only flow when a gate voltage (V_{GS}) is applied to the gate terminal greater than the threshold voltage (V_{TH}) level in which conductance takes place making it a transconductance device.

The application of a positive (+ve) gate voltage to a n-type eMOSFET attracts more electrons towards the oxide layer around the gate thereby increasing or enhancing (hence its name) the thickness of the channel allowing more current to flow. This is why this kind of transistor is called an enhancement mode device as the application of a gate voltage enhances the channel.

Increasing this positive gate voltage will cause the channel resistance to decrease further causing an increase in the drain current, I_D through the channel. In other words, for an n-channel enhancement mode MOSFET: $+V_{GS}$ turns the transistor "ON", while a zero or $-V_{GS}$ turns the transistor "OFF". Thus the enhancement-mode MOSFET is equivalent to a "normally-open" switch.

The reverse is true for the p-channel enhancement MOS transistor. When $V_{GS} = 0$ the device is "OFF" and the channel is open. The application of a negative (-ve) gate voltage to the p-type eMOSFET enhances the channels conductivity turning it "ON". Then for an p-channel enhancement mode MOSFET: $+V_{GS}$ turns the transistor "OFF", while $-V_{GS}$ turns the transistor "ON".

Enhancement-mode N-Channel MOSFET and Circuit Symbols



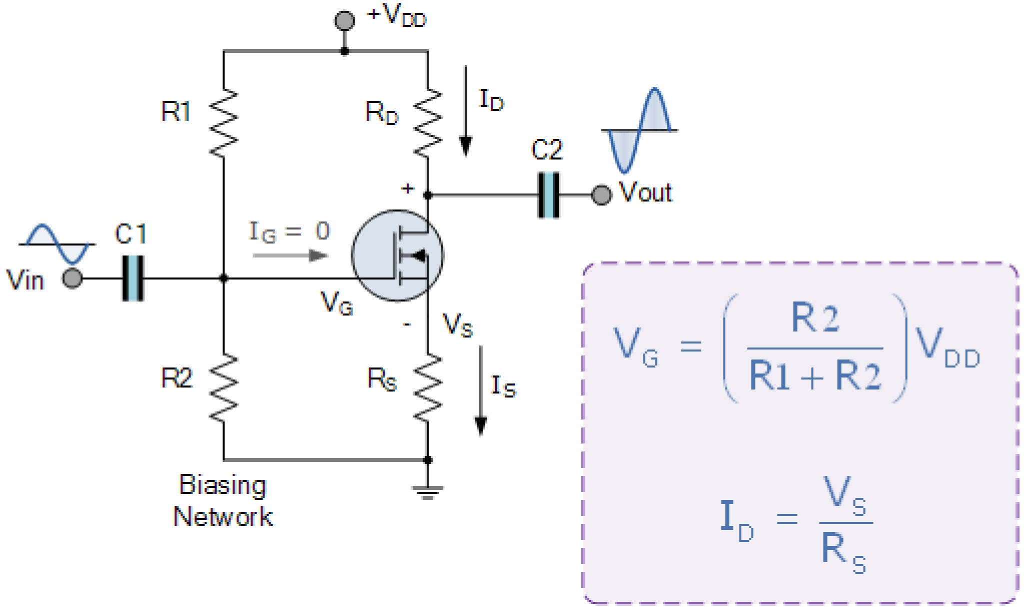
Enhancement-mode MOSFETs make excellent electronics switches due to their low "ON" resistance and extremely high "OFF" resistance as well as their infinitely high input resistance due to their isolated gate. Enhancement-mode MOSFETs are used in integrated circuits to produce CMOS type *Logic Gates* and power switching circuits in the form of as PMOS (P-channel) and NMOS (N-channel) gates. CMOS actually stands for *Complementary MOS* meaning that the logic device has both PMOS and NMOS within its design.

The MOSFET Amplifier

Just like the previous Junction Field Effect transistor, MOSFETs can be used to make single stage class "A" amplifier circuits with the enhancement mode n-channel MOSFET common source amplifier being the most popular circuit. Depletion mode MOSFET amplifiers are very similar to the JFET amplifiers, except that the MOSFET has a much higher input impedance.

This high input impedance is controlled by the gate biasing resistive network formed by R1 and R2. Also, the output signal for the enhancement mode common source MOSFET amplifier is inverted because when V_G is low the transistor is switched "OFF" and V_D (Vout) is high. When V_G is high the transistor is switched "ON" and V_D (Vout) is low as shown.

Enhancement-mode N-Channel MOSFET Amplifier



The DC biasing of this common source (CS) MOSFET amplifier circuit is virtually identical to the JFET amplifier. The MOSFET circuit is biased in class A mode by the voltage divider network formed by resistors R1 and R2. The AC input resistance is given as $R_{IN} = R_G = 1M\Omega$.

Metal Oxide Semiconductor Field Effect Transistors are three terminal active devices made from different semiconductor materials that can act as either an insulator or a conductor by the application of a small signal voltage.

The MOSFETs ability to change between these two states enables it to have two basic functions: "switching" (digital electronics) or "amplification" (analogue electronics). Then MOSFETs have the ability to operate within three different regions:

- 1. Cut-off Region – with $V_{GS} < V_{threshold}$ the gate-source voltage is much lower than the transistors threshold voltage so the MOSFET transistor is switched "fully-OFF" thus, $I_D = 0$, with the transistor acting like an open switch regardless of the value of V_{DS}.
- 2. Linear (Ohmic) Region – with $V_{GS} > V_{threshold}$ and $V_{DS} < V_{GS}$ the transistor is in its constant resistance region behaving as a voltage-controlled resistance whose resistive value is determined by the gate voltage, V_{GS} level.
- 3. Saturation Region – with $V_{GS} > V_{threshold}$ and $V_{DS} > V_{GS}$ the transistor is in its constant current region and is therefore "fully-ON". The Drain current I_D = Maximum with the transistor acting as a closed switch.

MOSFET Tutorial Summary

The Metal Oxide Semiconductor Field Effect Transistor, or **MOSFET** for short, has an extremely high input gate resistance with the current flowing through the channel between the source and drain being controlled by the gate voltage. Because of this high input impedance and gain, MOSFETs can be easily damaged by static electricity if not carefully protected or handled.

MOSFET's are ideal for use as electronic switches or as common-source amplifiers as their power consumption is very small. Typical applications for metal oxide semiconductor field effect transistors are in Microprocessors, Memories, Calculators and Logic CMOS Gates etc.

Also, notice that a dotted or broken line within the symbol indicates a normally "OFF" enhancement type showing that "NO" current can flow through the channel when zero gate-source voltage V_{GS} is applied.

A continuous unbroken line within the symbol indicates a normally "ON" Depletion type showing that current "CAN" flow through the channel with zero gate voltage. For p-channel types the symbols are exactly the same for both types except that the arrow points outwards.

This can be summarised in the following switching table.

MOSFET type	V_{GS} = +ve	V_{GS} = 0	V_{GS} = -ve
N-Channel Depletion	ON	ON	OFF
N-Channel Enhancement	ON	OFF	OFF
P-Channel Depletion	OFF	ON	ON
P-Channel Enhancement	OFF	OFF	ON

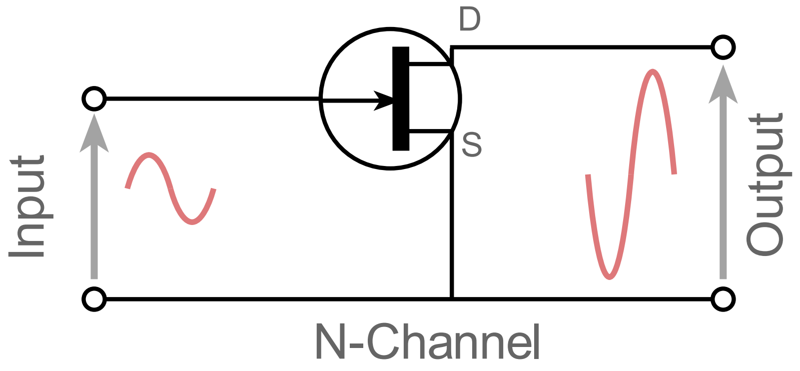
So for n-type enhancement type MOSFETs, a positive gate voltage turns "ON" the transistor and with zero gate voltage, the transistor will be "OFF". For a p-channel enhancement type MOSFET, a negative gate voltage will turn "ON" the transistor and with zero gate voltage, the transistor will be “OFF”. The voltage point at which the MOSFET starts to pass current through the channel is determined by the threshold voltage V_{TH} of the device.

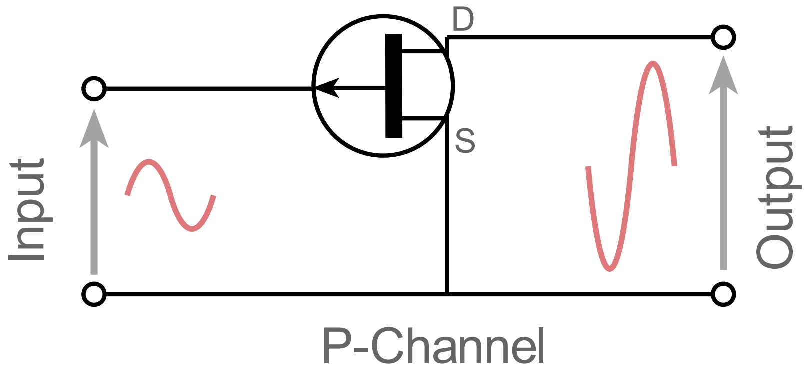
FET circuit configurations are the common source, common gate, and common drain formats. Each have their own characteristics of voltage and current gain as well as input and output impedance.
The choice of the FET circuit configuration or topology is one of the key design parameters on which the overall circuit design is based.

FET configuration basics

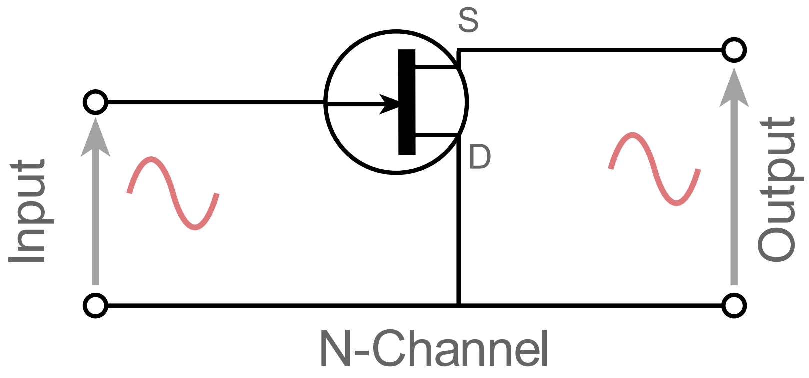
The terminology used for denoting the three basic FET configurations indicates the FET electrode that is common to both input and output circuits. This gives rise to the three terms: common gate, common drain and common source.
The three different FET circuit configurations are:

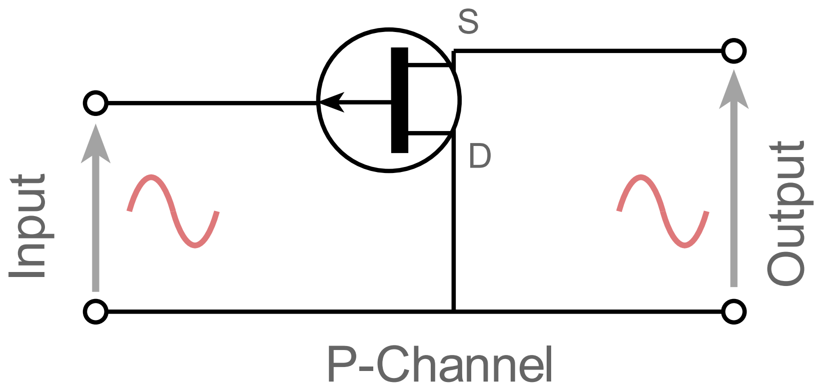
- ***Common source:*** This FET configuration is probably the most widely used. The common source circuit provides a medium input and output impedance levels. Both current and voltage gain can be described as medium, but the output is the inverse of the input, i.e. 180° phase change. This provides a good overall performance and as such it is often thought of as the most widely used configuration.



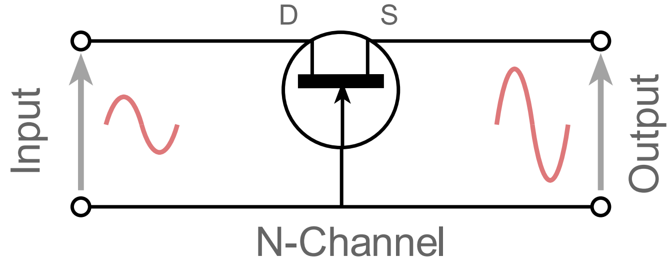


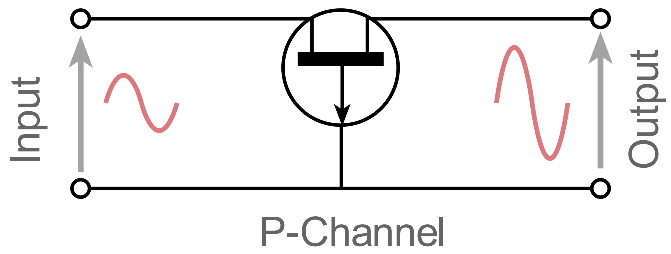
- ***Common drain:*** This FET configuration is also known as the source follower. The reason for this is that the source voltage follows that of the gate. Offering a high input impedance and a low output impedance it is widely used as a buffer. The voltage gain is unity, although current gain is high. The input and output signals are in phase.





- ***Common gate:*** This transistor configuration provides a low input impedance while offering a high output impedance. Although the voltage is high, the current gain is low and the overall power gain is also low when compared to the other FET circuit configurations available. The other salient feature of this configuration is that the input and output are in phase.





FET circuit configuration summary table

The table below gives a summary of the major properties of the different FET circuit configurations.

FET CONFIGURATION SUMMARY TABLE

FET CONFIGURATION	COMMON GATE	COMMON DRAIN (SOURCE FOLLOWER)	COMMON SOURCE
Voltage gain	High	Low	Medium
Current gain	Low	High	Medium
Power gain	Low	Medium	High
Input resistance	Low	High	Medium
Output resistance	High	Low	Medium
Input / output phase relationship	0°	0°	180°

As can be seen the different configurations or topologies have different characteristics. The common source is the most widely used FET circuit configuration and it equates to the common emitter transistor amplifier. The FET common drain or source follower is used as a buffer amplifier and it equates to the transistor common emitter amplifier.

Chapter 7: Power Supply Units

Power supplies are an important element in many items of electronics equipment. While some are battery driven, others need mains power supplies, and the power supply electronics circuitry and design is of paramount importance to the successful operation of the whole equipment.

Power supply electronics circuits can be split into a number of sections or building blocks. Each is important to the operation of the power supply as a whole, but each section of the power supply electronics is required to perform its function satisfactorily for the successful operation of the whole unit.

WARNING!: *Many power supplies will contain mains or line voltages which can be hazardous. Extreme care must be taken when dealing with these circuits, as electric shocks could be fatal. Only qualified personnel should deal with the internal circuitry of power supply electronics circuits.*

Types of electronics power supply

There are three main types of power supply that can be used. Each ash their own advantages and disadvantages and as a result each is used under slightly different circumstances.

The three major types of electronics power supply are:

- ***Rectified and smoothed power supply:*** These electronics power supplies are the simplest types, and are generally used for non-critical applications where performance is not a major issue. This type of power supply was widely used in thermionic valve or vacuum tube equipment as it was not so easy to regulate supplies, and often the requirements were not so critical.
- ***Linear regulated power supply:*** This form of electronics power supply is able to provide a very high level of performance. However the fact that it uses a series regulator element means that it can be comparatively inefficient, dissipating a significant proportion of the input power as heat. Nevertheless these power supplies can offer very high levels of regulation with low values of ripple, etc. . . . . *Read more about* ***Linear Power Supplies.***
- ***Switch mode power supply:*** In this form of power supply, electronics circuits use switching technology to regulate the output. Although spikes are present on the output, they offer very high levels of efficiency and in view of this they can be contained in much smaller packages than their linear equivalents. . . . . *Read more about* ***Switch Mode Power Supplies, SMPS.***

The different types of power supply are each used for different types of application according to their advantages. As such they are all widely used, but in different areas of electronics. Each type of building block and power supply is covered in greater detail on other pages on this website. Links to these pages can be found on the left hand side of the page below the main menu in the "Related Articles" section.

Major power supply electronics blocks

A power supply can be split into a number of elements, each providing a function within the overall power supply. Naturally these areas can be rather arbitrary, and may vary slightly dependent upon the actual power supply design, but they can be sued as a rough overall guide.

- ***Power input filtering:*** It some instances it is necessary to ensure that spikes from the power line do not enter the power supply, and that noise that might be generated by the power supply does not enter the power lines. To achieve this circuitry to remove noise and limit the effects of incoming spikes is placed at the input to the power supply. In many cases any filtering at this point is quite minimal, although for specialist supplies more complicated circuits may be used.
- ***Input transformer:*** If a power supply using mains / line supplies of 110 or 240 volts AC is used, then the input usually has a transformer to transform the incoming line voltage to required level for the power supply design.

- ***Rectifier:*** It is necessary to change the incoming AC waveform to a DC waveform. This is achieved using an AC rectifier circuit. Two types of rectifier circuit may be

  used - full wave and half wave rectifiers. These effectively block the part of the waveform in one sense and allow through the part of the waveform in the other sense.
  The rectifying action of a diode
- ***Rectifier smoothing:*** The output from the AC rectifier circuit consists of a waveform varying from zero volts to 1.414 times the RMS input voltage (less any losses introduced by the rectifier). In order that this can be used by electronics circuits, it needs to be smoothed. This is achieved using a capacitor. It will charge up over part of the cycle and then as the voltage falls it will supply the current to the circuit, charging up again as the voltage rises.

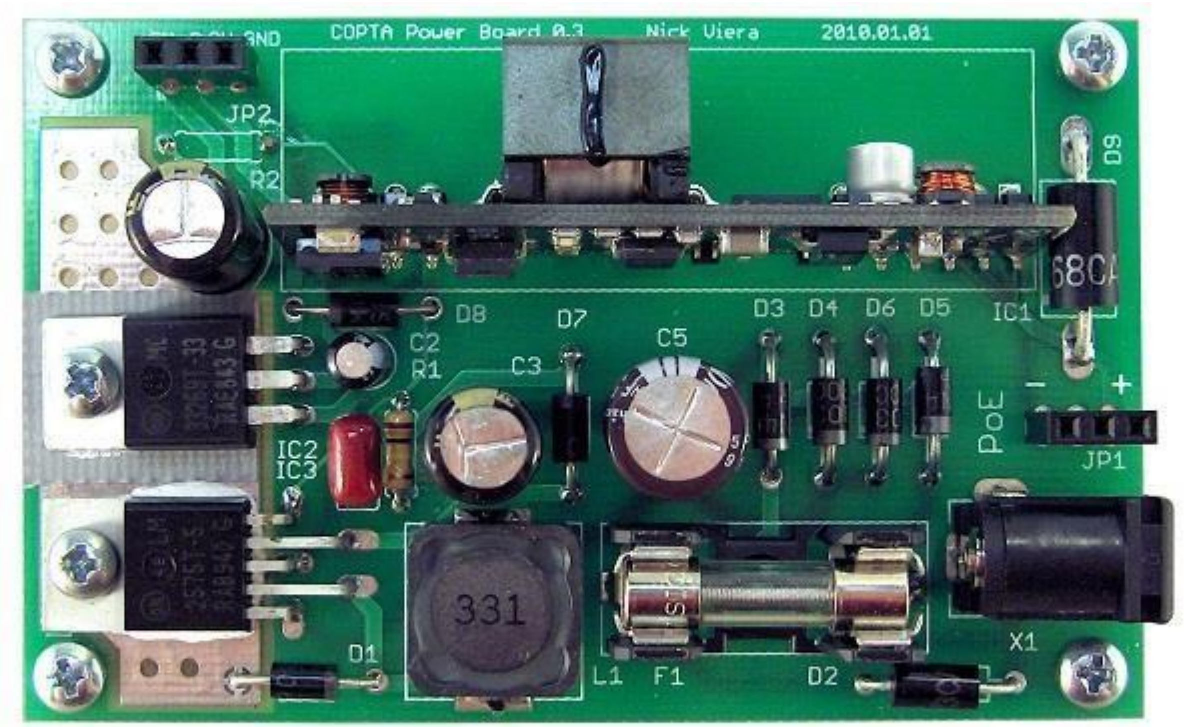
. . . . . *Read more about* ***[Capacitor Smoothing Circuits.]***

- ***Regulation:*** Even after the rectified voltage has been smoothed, there may still be significant levels of residual hum. Also the voltage will vary as different levels of current are drawn. To provide a stable voltage output from the power supply with little residual hum and noise a voltage regulator circuit is required. Regulators are able to provide a stable voltage at a set or variable level dependent upon the requirement. This may use either linear or switching mode techniques to bring the output voltage to the required level.
- ***Over voltage protection:*** In the event of the failure of the regulator it is possible under some circumstances that the output voltage from the power supply could rise to a level that could damage the circuitry being powered. To prevent this occurrence over-voltage protection circuitry can be used. This circuit element detects the level of the output voltage and if it starts to rise above its acceptable limits it will trip, removing the supply from the regulator and usually clamping the output from the regulator to zero volts, thereby protecting the remaining circuitry from damage.

Not all of these power supply electronics building blocks are used in every power supply. Most will have a transformer, smoothing and a regulator, but the other elements may or may not be included dependent upon the specification.
B

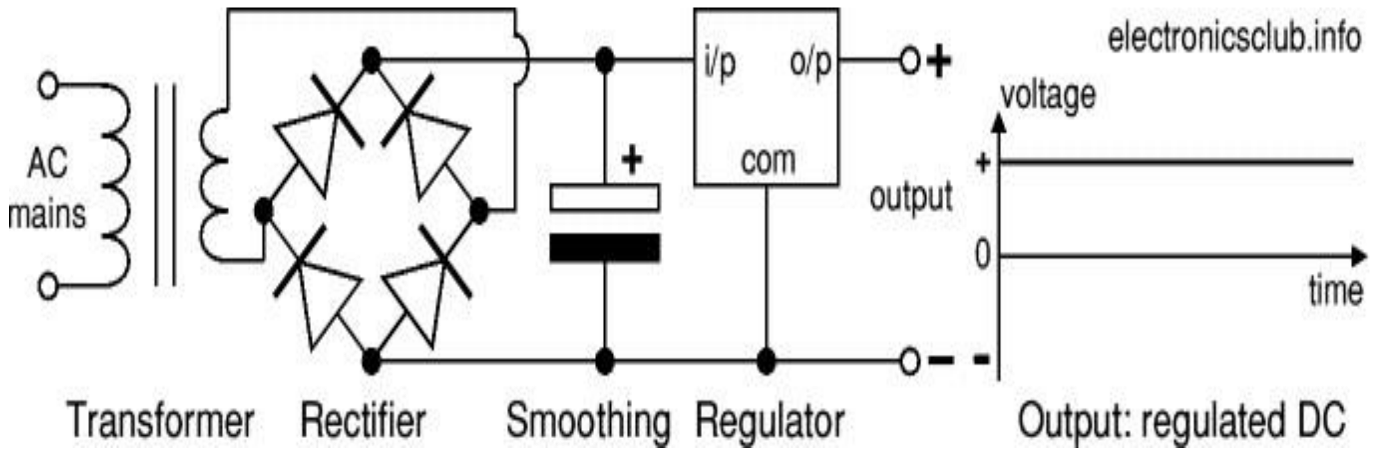
Introduction to power supplies

Power supplies are used in computers, television, laptops, Cathode ray Oscilloscope, broadcast radios, among other many electronic circuits. Special power supplies have outputs that supply power at different voltages to sections that require different power inputs. For example, some inputs may need a 12v DC while some others may need a 30v DC. In order to provide the required dc voltages, the incoming 230v AC supply has to be converted into pure DC for the usage.



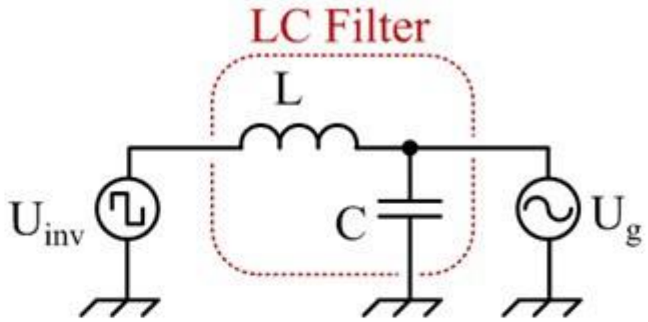
✞ **Major electronics blocks in a power supply circuit.**

A power supply can be split into a number of elements, each providing a function within the overall power supply. Naturally these areas can be rather arbitrary, and may vary slightly dependent upon the actual power supply design, but they can be sued as a rough overall guide.



- **Power input filtering**

Input filters consist of capacitors and inductors. They help protect the power supply from spikes from the power line. The filters also ensure that noise generated by the power supply does not enter the power lines; the filters prevent electromagnetic interference, generated by the switching **source** from reaching the **power** line and affecting other equipment. To achieve this circuitry to remove noise and limit the effects of incoming spikes is placed at the input to the power supply. In many cases any filtering at this point is quite minimal, although for specialist supplies more complicated circuits may be used.



- **Input transformer**

If a power supply using mains / line supplies of 110- or 240-volts AC is used, then the input usually has a transformer to transform the incoming line voltage to required level for the power supply design.

- **Rectifier**

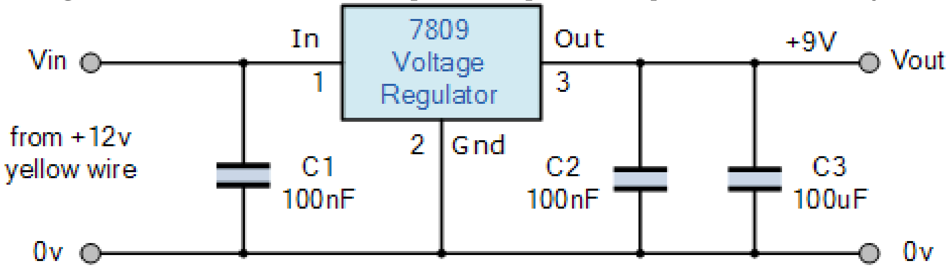
It is necessary to change the incoming AC waveform to a DC waveform. This is achieved using an AC rectifier circuit. Two types of rectifier circuit may be used - full wave and half wave rectifiers. These effectively block the part of the waveform in one sense and allow through the part of the waveform in the other sense. The rectifying action of a diode.

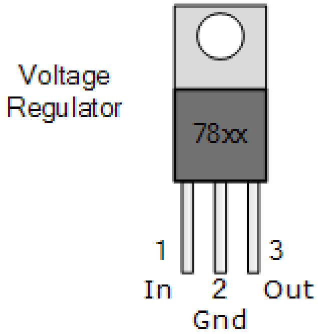
- **Rectifier smoothing**

The output from the AC rectifier circuit consists of a waveform varying from zero volts to 1.414 times the RMS input voltage (less any losses introduced by the rectifier). In order that this can be used by electronics circuits, it needs to be smoothed. This is achieved using a capacitor. It will charge up over part of the cycle and then as the voltage falls it will supply the current to the circuit, charging up again as the voltage rises.

- **Regulation**

Even after the rectified voltage has been smoothed, there may still be significant levels of residual hum. Also, the voltage will vary as different levels of current are drawn. To provide a stable voltage output from the power supply with little residual hum and noise a voltage regulator circuit is required. Regulators are able to provide a stable voltage at a set or variable level dependent upon the requirement. This may use either linear or switching mode techniques to bring the output voltage to the required level.

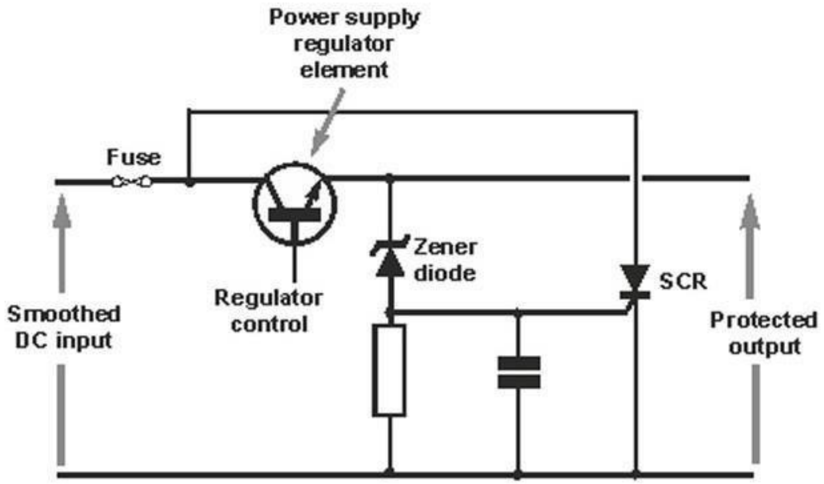




Type	Min Input Voltage	Output Voltage
7805	7V	+5V
7806	8V	+6V
7808	10V	+8V
7809	11V	+9V
7812	15V	+12V
7815	18V	+15V
7818	22V	+18V
7824	30V	+24V

- **Over voltage protection**

In the event of the failure of the regulator it is possible under some circumstances that the output voltage from the power supply could rise to a level that could damage the circuitry being powered. To prevent this occurrence over-voltage protection circuitry can be used. This circuit element detects the level of the output voltage and if it starts to rise above its acceptable limits it will trip, removing the supply from the regulator and usually clamping the output from the regulator to zero volts, thereby protecting the remaining circuitry from damage.

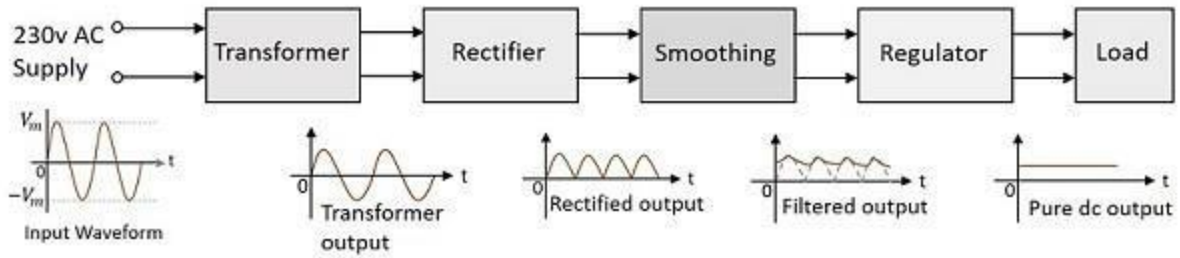


Parts of a Power supply A typical Power supply unit consists of the following.

- **Transformer** − An input transformer for the stepping down of the 230v AC power supply.
- **Rectifier** − A Rectifier circuit to convert the AC components present in the signal to DC components.
- **Smoothing** − A filtering circuit to smoothen the variations present in the rectified output.
- **Regulator** − A voltage regulator circuit in order to control the voltage to a desired output level. • **Load** − The load which uses the pure dc output from the regulated output.

Block Diagram of a Power Supply Unit

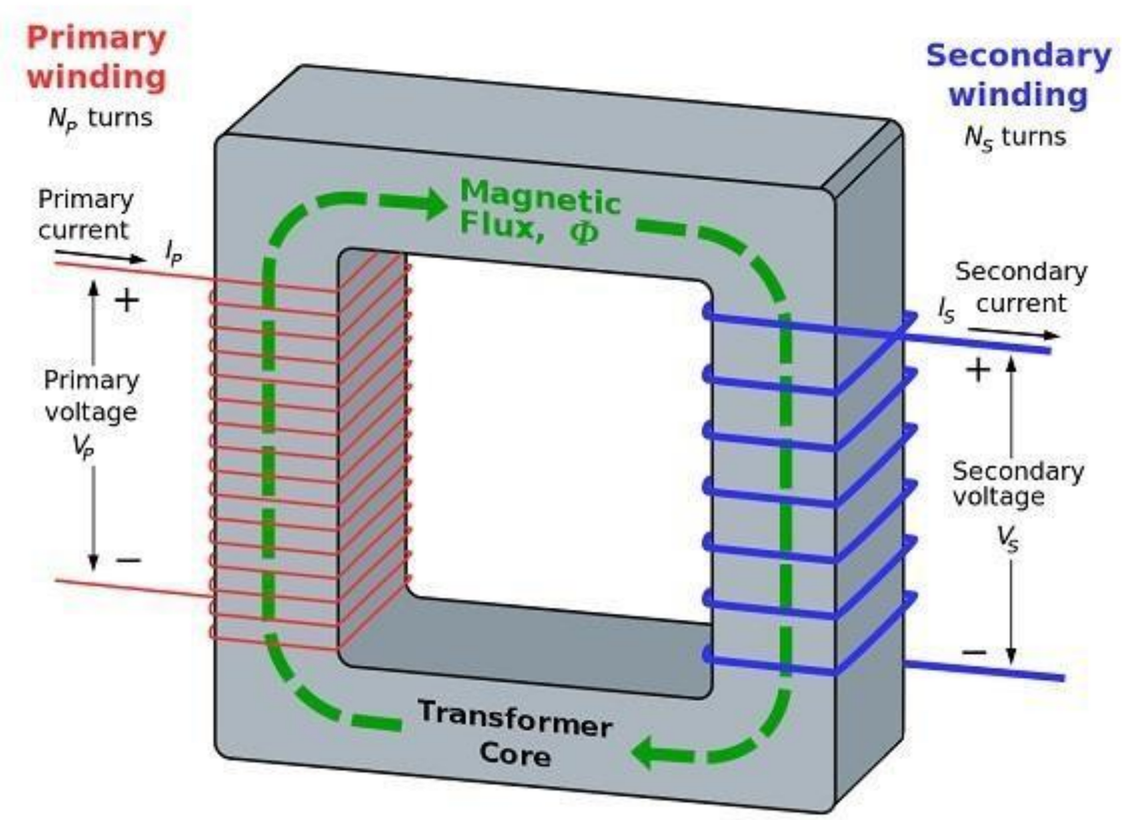
The block diagram of a Regulated Power supply unit is as shown below.



From the diagram above, it is evident that the transformer is present at the initial stage. Though we had already gone through the concept regarding transformers in BASIC ELECTRONICS tutorial, let us have a glance over it.

- **Transformer**

A transformer has a **primary coil** to which **input** is given and a **secondary coil** from which the **output** is collected. Both of these coils are wound on a core material. Usually an insulator forms the **Core** of the transformer.
The following figure shows a practical transformer.

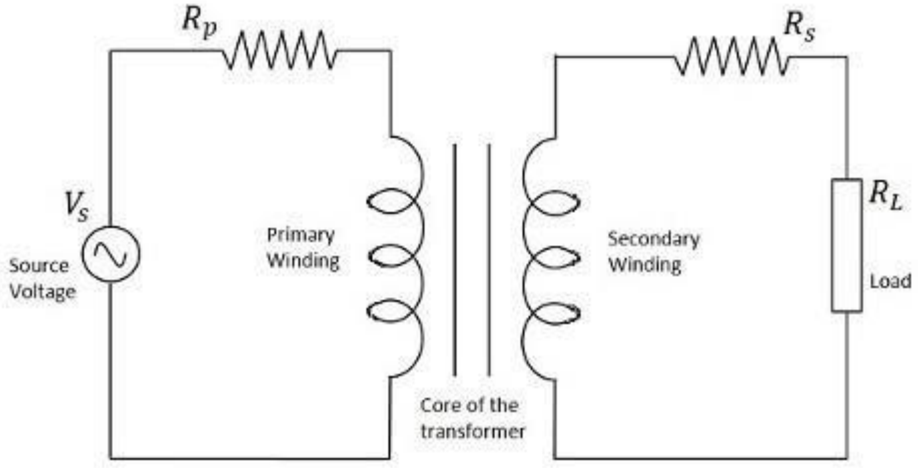


From the above figure, it is evident that a few notations are common. They are as follows − • NpNp = Number of turns in the primary winding

- NsNs = Number of turns in the secondary winding
- IpIp = Current flowing in the primary of the transformer
- IsIs = Current flowing in the secondary of the transformer
- VpVp = Voltage across the primary of the transformer
- VsVs = Voltage across the secondary of the transformer
- ϕϕ = Magnetic flux present around the core of the transformer

- ***Transformer in a Circuit***

The following figure shows how a transformer is represented in a circuit. The primary winding, the secondary winding and the core of the transformer are also represented in the following figure.

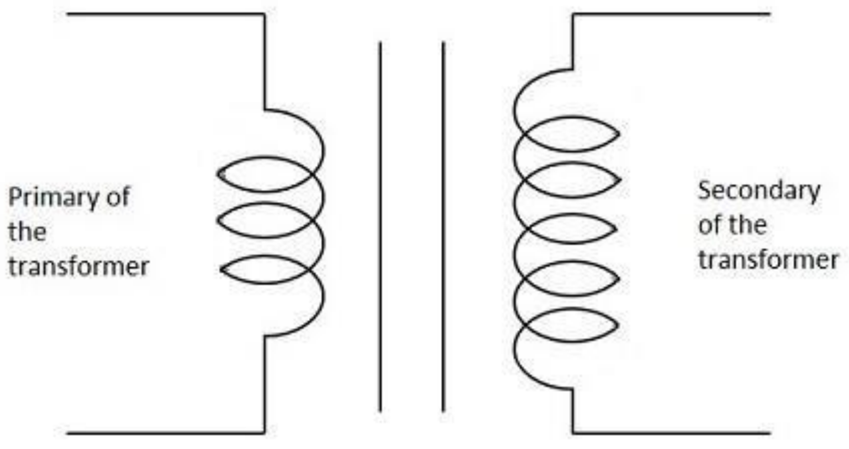


Hence, when a transformer is connected in a circuit, the input supply is given to the primary coil so that it produces varying magnetic flux with this power supply and that flux is induced into the secondary coil of the transformer, which produces the varying EMF of the varying flux. As the flux should be varying, for the transfer of EMF from primary to secondary, a transformer always works on alternating current AC.

Depending upon the number of turns in the secondary winding, a transformer can be classified either as a **Step-up** or a **Step-down** transformer. ➢ ***Step-Up Transformer***

When the secondary winding has a greater number of turns than the primary winding, then the transformer is said to be a **Step-up** transformer. Here the induced EMF is greater than the input signal.

The figure below shows the symbol of a step-up transformer.

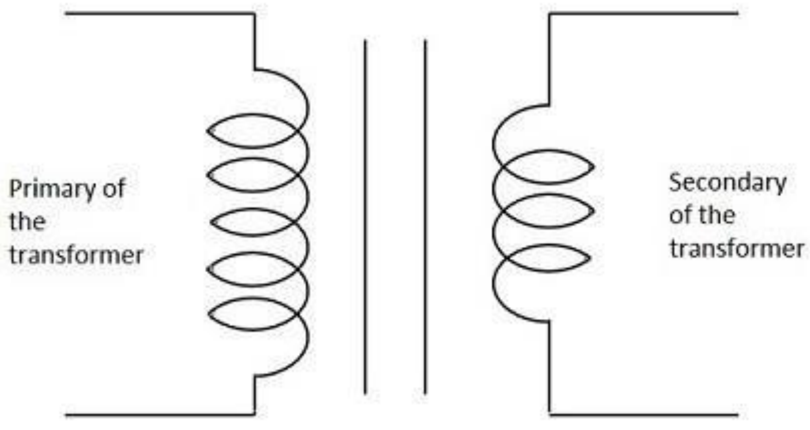


Step-up transformer

- ***Step-Down Transformer***

When the secondary winding has lesser number of turns than the primary winding, then the transformer is said to be a **Step-down** transformer. Here the induced EMF is lesser than the input signal.

The figure below shows the symbol of a step-down transformer.



Step-down transformer

In our Power supply circuits, we use the **Step-down transformer**, as we need to lessen the AC power to DC. The output of this Step-down transformer will be less in power and this will be given as the input to the next section, called **rectifier**. We will discuss about rectifiers in the next chapter.

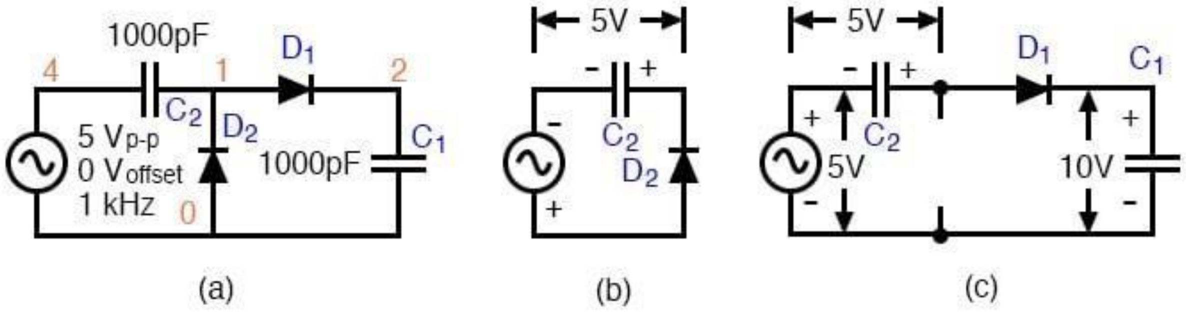
❖ Voltage multipliers

A *voltage multiplier* is a specialized rectifier circuit producing an output which is theoretically an integer times the AC peak input, for example, 2, 3, or 4 times the AC peak input. Thus, it is possible to get 200 VDC from a 100 Vpeak AC source using a doubler, 400 VDC from a quadrupler. Any load in a practical circuit will lower these voltages. We'll first go over several types of voltage multipliers—voltage doubler (half- and full-wave), voltage tripler, and voltage quadrupler—then make some general notes about voltage multiplier safety and finish up with the Cockcroft-Walton multiplier.

- **Voltage Doubler**

A voltage doubler application is a DC power supply capable of using either a 240 VAC or 120 VAC source. The supply uses a switch selected full-wave bridge to produce about 300 VDC from a 240 VAC source. The 120 V position of the switch rewires the bridge as a doubler producing about 300 VDC from the 120 VAC. In both cases, 300 VDC is produced. This is the input to a switching regulator producing lower voltages for powering, say, a personal computer. ➢ ***Half-Wave Voltage Doubler***

The half-wave voltage doubler in Figure below (a) is composed of two circuits: a clamper at (b) and peak detector (half-wave rectifier) in Figure prior, which is shown in modified form in Figure below (c). C2 has been added to a peak detector (half-wave rectifier).

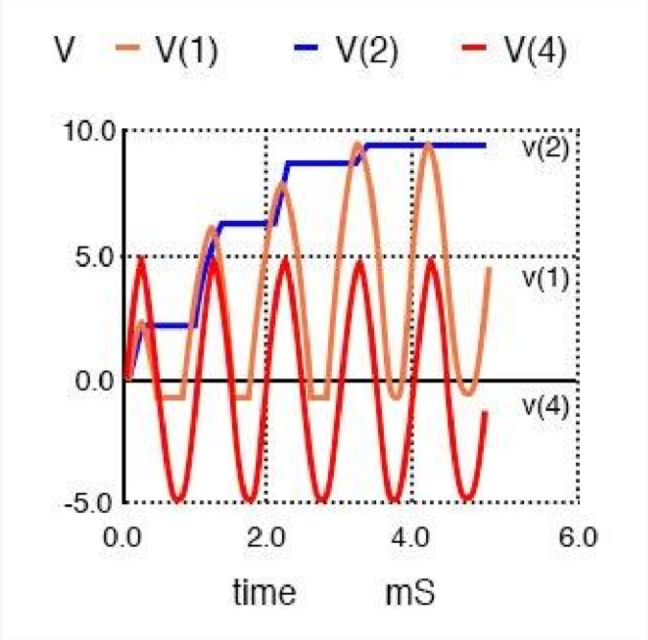


Half-wave voltage doubler (a) is composed of (b) a clamper and (c) a half-wave rectifier.

Half-wave Voltage Doubler Operation Circuit Analysis

Referring to Figure(b) above , C2 charges to 5 V (4.3 V considering the diode drop) on the negative half cycle of AC input. The right end is grounded by the conducting D2. The left end is charged at the negative peak of the AC input. This is the operation of the . clamper

During the positive half cycle, the half-wave rectifier comes into play at Figure(c) above . Diode D2 is out of the circuit since it is reverse biased. C2 is now in series with the voltage source. Note the polarities of the generator and C2, series aiding. Thus, rectifier D1 sees a total of 10 V at the peak of the sinewave, 5 V from generator and 5 V from C2. D1 conducts waveform v(1) (figure), charging below C1 to the peak of the sine wave riding on 5 V DC (figure below v(2)). Waveform v(2) is the output of the doubler, which stabilizes at 10 V (8.6 V with diode drops) after a few cycles of sine wave input.



*SPICE 03255.eps C1 2 0 1000p D1 1 2 diode C2 4 1 1000p D2 0 1 diode V1 4 0 SIN(0 5 1k) .model diode d .tran 0.01m 5m .end

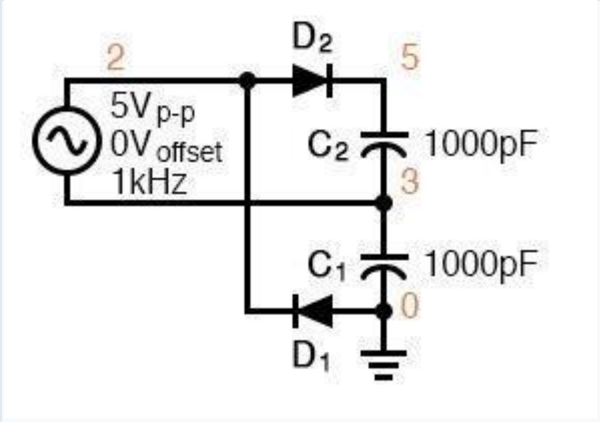
Voltage doubler: v(4) input. v(1) clamper stage. v(2) half-wave rectifier stage, which is the doubler output.

➢ ***Full-Wave Voltage Doubler***

The *full-wave voltage doubler* is composed of a pair of series stacked half-wave rectifiers. (Figure below) The corresponding netlist is in Figure below.

Full-Wave Voltage Doubler Operation Analysis

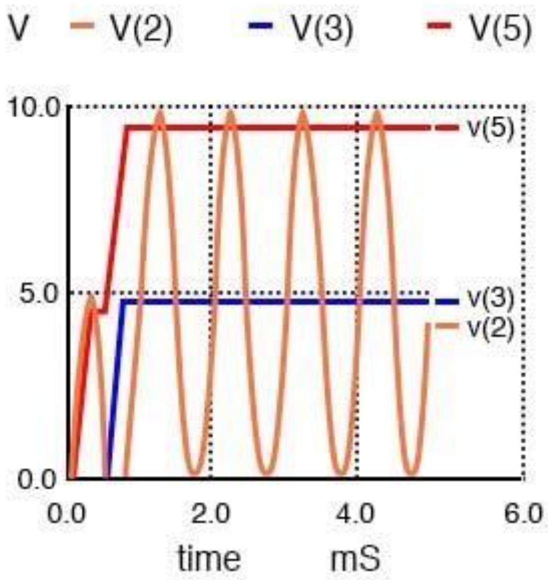
The bottom rectifier charges C1 on the negative half cycle of input. The top rectifier charges C2 on the positive halfcycle. Each capacitor takes on a charge of 5 V (4.3 V considering diode drop). The output at node 5 is the series total of C1 + C2 or 10 V (8.6 V with diode drops).



*SPICE 03273.eps *R1 3 0 100k *R2 5 3 100k D1 0 2 diode D2 2 5 diode C1 3 0 1000p C2 5 3 1000p V1 2 3 SIN(0 5 1k) .model diode d .tran 0.01m 5m .end

Full-wave voltage doubler consists of two half-wave rectifiers operating on alternating polarities.

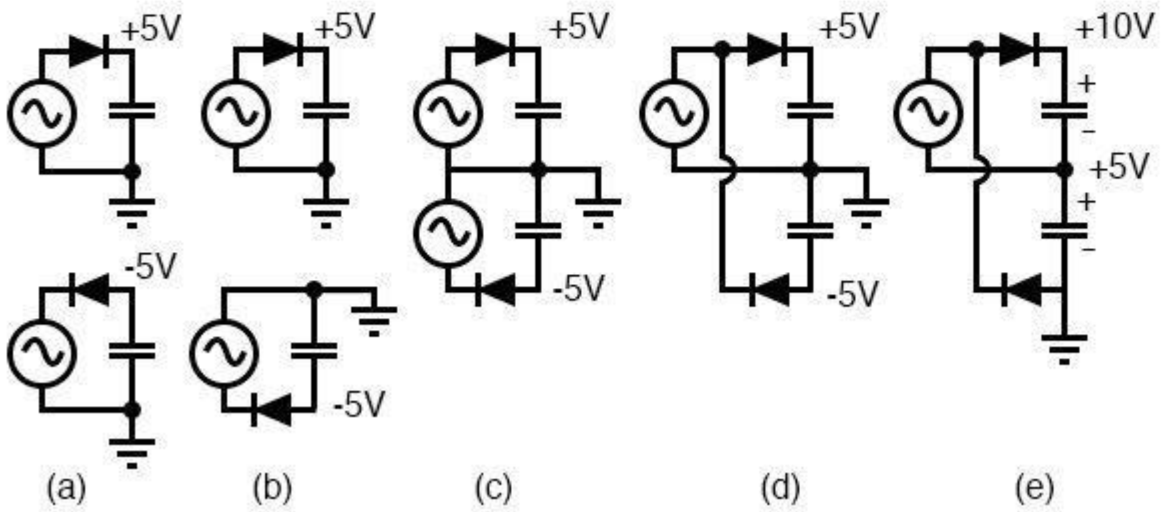
Note that the output v(5) Figure below reaches full value within one cycle of the input v(2) excursion.



Full-wave voltage doubler: v(2) input, v(3)voltage at mid point, v(5) voltage at output

• **Deriving Full-wave Doublers from Half-wave Rectifiers**

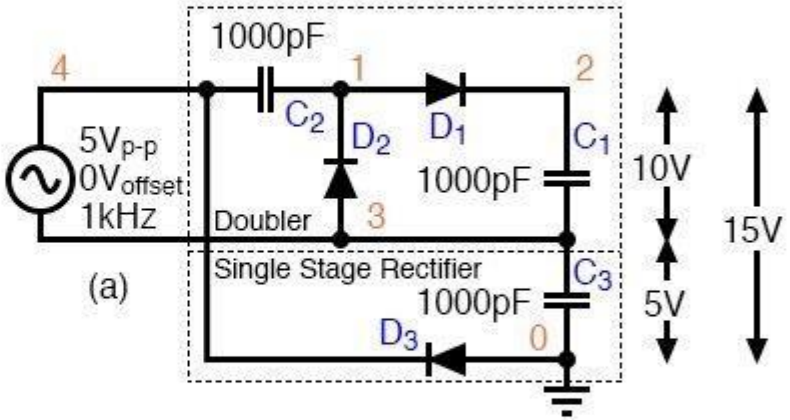
Figure below illustrates the derivation of the full-wave doubler from a pair of opposite polarity half-wave rectifiers (a). The negative rectifier of the pair is redrawn for clarity (b). Both are combined at (c) sharing the same ground. At (d) the negative rectifier is re-wired to share one voltage source with the positive rectifier. This yields a ±5 V (4.3 V with diode drop) power supply; though, 10 V is measurable between the two outputs. The ground reference point is moved so that +10 V is available with respect to ground.



Full-wave doubler: (a) Pair of doublers, (b) redrawn, (c) sharing the ground, (d) share the same voltage source. (e) move the ground point.

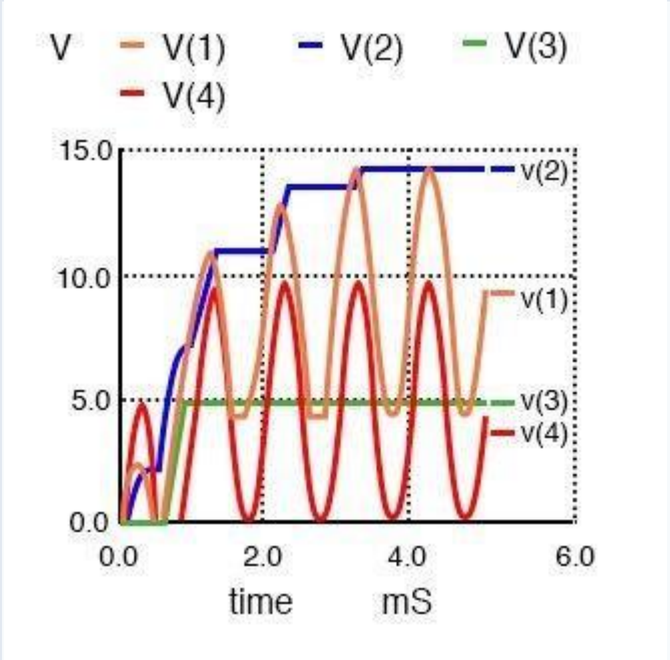
• **Voltage Tripler**

A *voltage tripler* (Figure below) is built from a combination of a doubler and a half wave rectifier (C3, D3). The half-wave rectifier produces 5 V (4.3 V) at node 3. The doubler provides another 10 V (8.4 V) between nodes 2 and 3. for a total of 15 V (12.9 V) at the output node 2 with respect to ground. The netlist is in Figure below.



Voltage tripler composed of doubler stacked atop a single stage rectifier.

Note that V(3) in Figure below rises to 5 V (4.3 V) on the first negative half cycle. Input v(4) is shifted upward by 5 V (4.3 V) due to 5 V from the half-wave rectifier. And 5 V more at v(1) due to the clamper (C2, D2). D1 charges C1 (waveform v(2)) to the peak value of v(1).



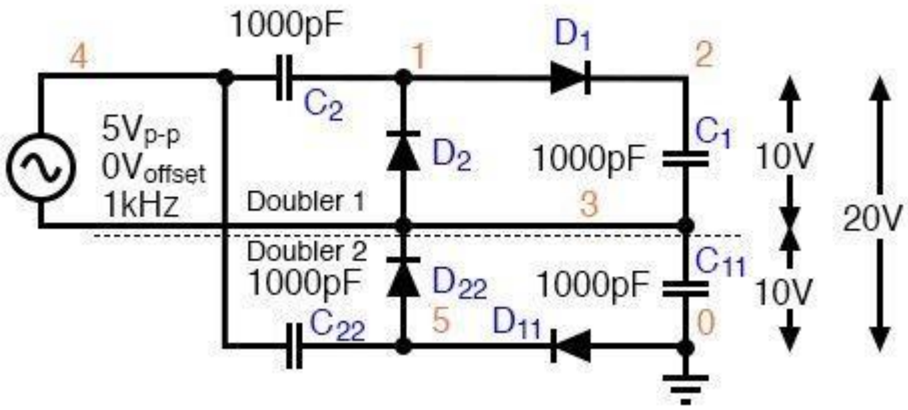
*SPICE 03283.eps C3 3 0 1000p D3 0 4 diode C1 2 3 1000p D1 1 2 diode C2 4 1 1000p D2 3 1 diode V1 4 3 SIN(0 5 1k) .model diode d .tran 0.01m 5m .end

Voltage tripler: v(3) half-wave rectifier, v(4) input+ 5 V, v(1) clamper, v(2) final output.

- **Voltage Quadrupler**

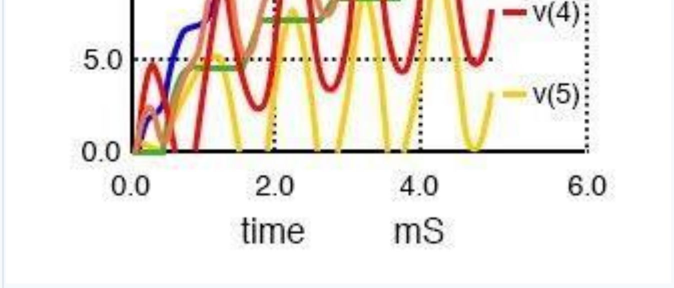
A *voltage quadrupler* is a stacked combination of two doublers shown in Figure below. Each doubler provides 10 V (8.6 V) for a series total at node 2 with respect to ground of 20 V (17.2 V)

The netlist is in Figure below.



Voltage quadrupler, composed of two doublers stacked in series, with output at node 2.

The waveforms of the quadrupler are shown in Figure below. Two DC outputs are available: v(3), the doubler output, and v(2) the quadrupler output. Some of the intermediate voltages at clampers illustrate that the input sinewave (not shown), which swings by 5 V, is successively clamped at higher levels: at v(5), v(4) and v(1). Strictly v(4) is not a clamper output. It is simply the AC voltage source in series with the v(3) the doubler output. None the less, v(1) is a clamped version of v(4)



*SPICE 03441.eps *SPICE 03286.eps C22 4 5 1000p C11 3 0 1000p D11 0 5 diode D22 5 3 diode C1 2 3 1000p D1 1 2 diode C2 4 1 1000p D2 3 1 diode V1 4 3 SIN(0 5 1k) .model diode d .tran 0.01m 5m .end

Voltage quadrupler: DC voltage available at v(3) and v(2). Intermediate waveforms: Clampers: v(5), v(4), v(1).

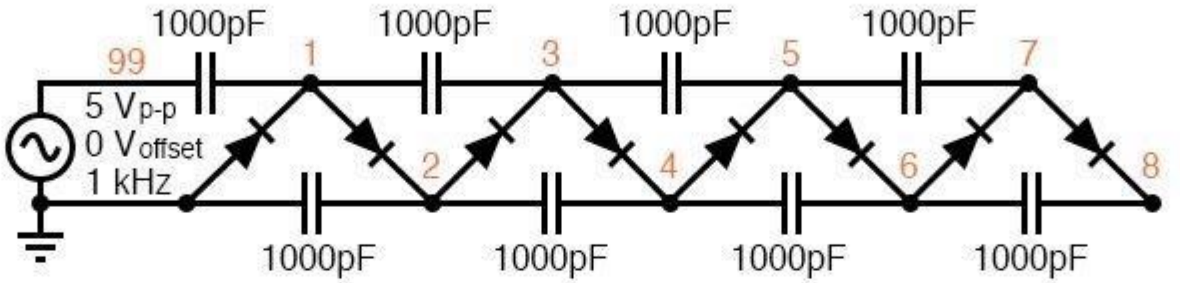
- **Notes on Voltage Multipliers and Line Driven Power Supplies**

Some notes on voltage multipliers are in order at this point. The circuit parameters used in the examples (V= 5 V 1 kHz, C=1000 pf) do not provide much current, microamps. Furthermore, load resistors have been omitted. Loading reduces the voltages from those shown. If the circuits are to be driven by a kHz source at low voltage, as in the examples, the capacitors are usually 0.1 to 1.0 µF so that milliamps of current are available at the output. If the multipliers are driven from 50/60 Hz, the capacitor are a few hundred to a few thousand microfarads to provide hundreds of milliamps of output current. If driven from line voltage, pay attention to the polarity and voltage ratings of the capacitors.

Finally, any direct line driven power supply (no transformer) is dangerous to the experimenter and line operated test equipment. Commercial direct driven supplies are safe because the hazardous circuitry is in an enclosure to protect the user. When breadboarding these circuits with electrolytic capacitors of any voltage, the capacitors will explode if the polarity is reversed. Such circuits should be powered up behind a safety shield.

- **Cockcroft-Walton Multiplier**

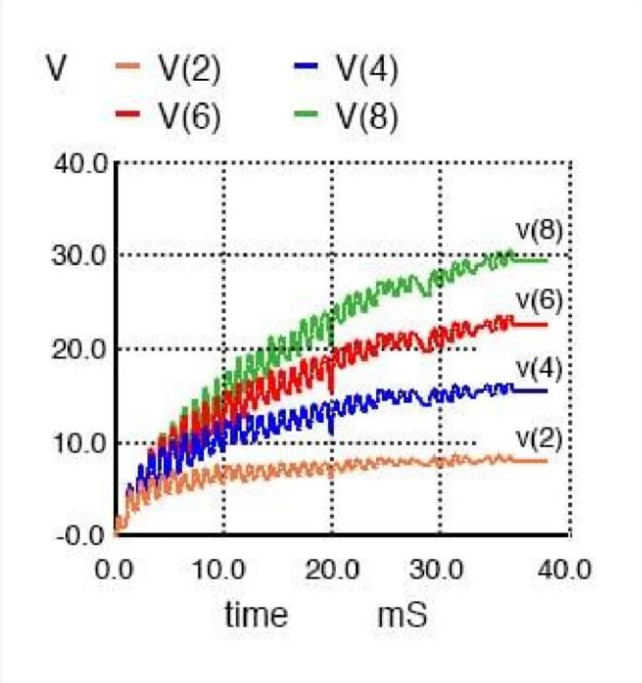
A voltage multiplier of cascaded half-wave doublers of arbitrary length is known as a *Cockcroft-Walton* multiplier as shown in Figure below. This multiplier is used when a high voltage at low current is required. The advantage over a conventional supply is that an expensive high voltage transformer is not required– at least not as high as the output.



Cockcroft-Walton x8 voltage multiplier; output at v(8).

The pair of diodes and capacitors to the left of nodes 1 and 2 in Figure above constitute a half-wave doubler. Rotating the diodes by 45º counterclockwise, and the bottom capacitor by 90º makes it look like Figure prior (a). Four of the doubler sections are cascaded to the right for a theoretical x8 multiplication factor. Node 1 has a clamper waveform (not shown), a sinewave shifted up by 1x (5 V). The other odd numbered nodes are sinewaves clamped to successively higher voltages. Node 2, the output of the first doubler, is a 2x DC voltage v(2) in Figure below. Successive even numbered nodes charge to successively higher

voltages: v(4), v(6), v(8)




```
D1 7 8 diode C1 8 6 1000p D2 6 7 diode C2 5 7 1000p D3 5 6 diode C3 4 6 1000p D4 4 5 diode C4 3 5 1000p D5 3 4 diode C5 2 4 1000p D6 2 3 diode D7 1 2 diode C6 1 3 1000p C7 2 0 1000p C8 99 1 1000p D8 0 1 diode V1 99 0 SIN(0 5 1k) .model diode d .tran 0.01m 50m .end
```

Cockcroft-Walton (x8) waveforms. Output is v(8).

Without diode drops, each doubler yields 2Vin or 10 V, considering two diode drops (10-1.4)=8.6 V is realistic. For a total of 4 doublers one expects 4·8.6=34.4 V out of 40 V.

Consulting Figure above, v(2) is about right; however, v(8) is <30 V instead of the anticipated 34.4 V. The bane of the Cockcroft-Walton multiplier is that each additional stage adds less than the previous stage. Thus, a practical limit to the number of stages exist. It is possible to overcome this limitation with a modification to the basic circuit. [ABR] Also note the time scale of 40 msec compared with 5 ms for previous circuits. It required 40 msec for the voltages to rise to a terminal value for this circuit. The netlist in Figure above has a ".tran 0.010m 50m" command to extend the simulation time to 50 msec; though, only 40 msec is plotted.

The Cockcroft-Walton multiplier serves as a more efficient high voltage source for photomultiplier tubes requiring up to 2000 V. [ABR] Moreover, the tube has numerous *dynodes*, terminals requiring connection to the lower voltage "even numbered" nodes. The series string of multiplier taps replaces a heat generating resistive voltage divider of previous designs.

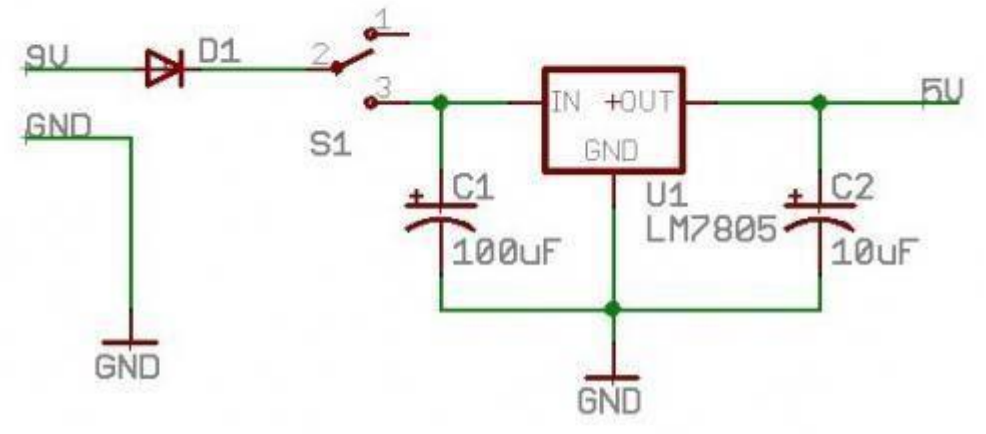
An AC line operated Cockcroft-Walton multiplier provides high voltage to "ion generators" for neutralizing electrostatic charge and for air purifiers.

- **/oltage Multiplier Review:**
  - A voltage multiplier produces a DC multiple (2,3,4, etc) of the AC peak input voltage.
  - The most basic multiplier is a half-wave doubler.
  - The full-wave double is a superior circuit as a doubler.
  - A tripler is a half-wave doubler and a conventional rectifier stage (peak detector).
  - A quadrupler is a pair of half-wave doublers
  - A long string of half-wave doublers is known as a Cockcroft-Walton multiplier.

Power supply methods of protection

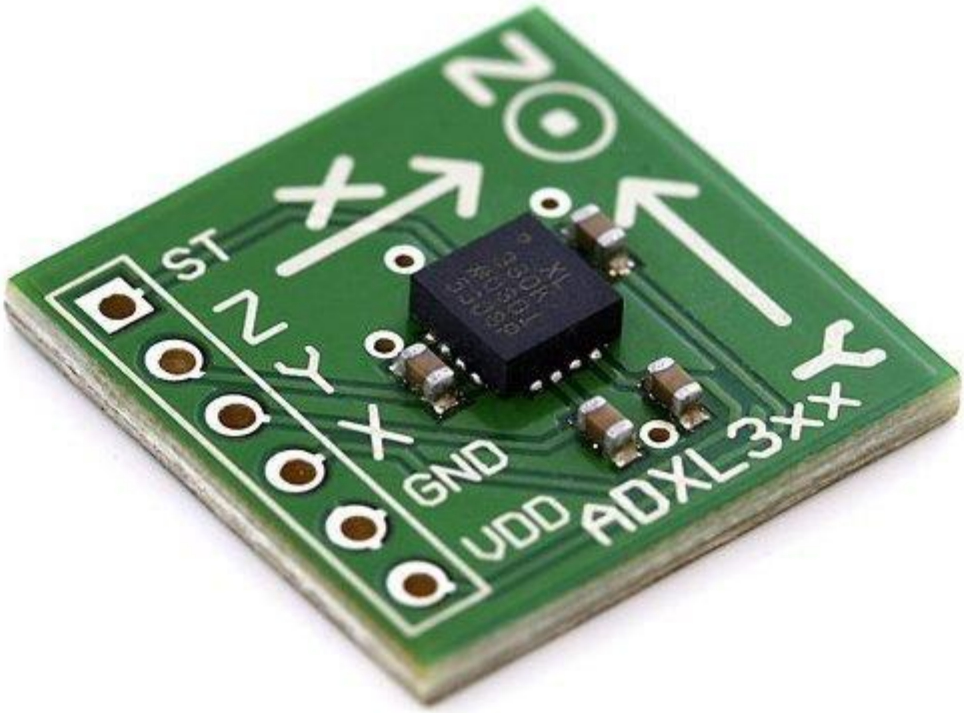
- **protection from reverse polarity: inline protection**

Users of electronic products often make mistakes while connecting loads to DC power supply such as connecting the negative power input to the positive terminal thereby damaging the circuit board. For that reason, most power supplies in a circuit board are protected from reverse polarity using a diode. This small diode is designed into the product to protect against reverse polarization. If someone hooks power up backwards, the diode fails to forward bias and the board simply doesn't turn on, protecting it from damage.



D1 is an 'inline protection diode'

Cheap diodes have a theoretical 0.7V drop. So, if you hook 5V up to the board, you'll get 5-0.7=4.3V delivered to the board. In practice, the forward drop of the diode is actually a bit lower (0.5V) and there are specialty diodes available that have even lower forward drop (germanium?). This all works great if your incoming power is 2-3 volts higher than your output, but if you're running a 5V board from a 5V source, the diode will drop the voltage to your system down significantly.



Checkout the Eagle DFM tutorial for more information about labeling your board.

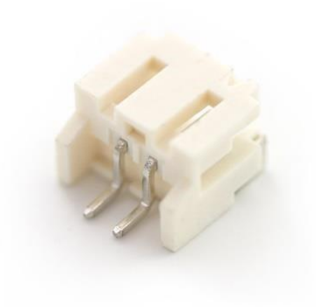
- **Voltage Regulator**

The nice thing about many voltage regulators is that they have short circuit and reverse polarization protection built in!

Note: Reversing the voltage on electrolytic or tantalum caps is a bad thing. A 16V rated tantalum may "pop" (explode with great force) if you apply 10V the wrong way. Electrolytic caps won't explode as violently, but may expand or puff out a bit.

- **Polarized battery connector**

Using a polarized connector like the SMD JST 2-pin connector instead of a bare 0.1" spaced connector.



Polarized power connectors
This connector makes it more difficult for people to willy-nilly attach a battery pack or power supply. This can initially be frustrating, but it forces the person to terminate their battery or power supply correctly. Once terminated, the user can quickly plug in the battery and not have to think about it. This works but may not provide infallible protection (the user may terminate a 12V wall wart to an intended 3.7V connector). We use this option all the time for our LiPo powered projects.

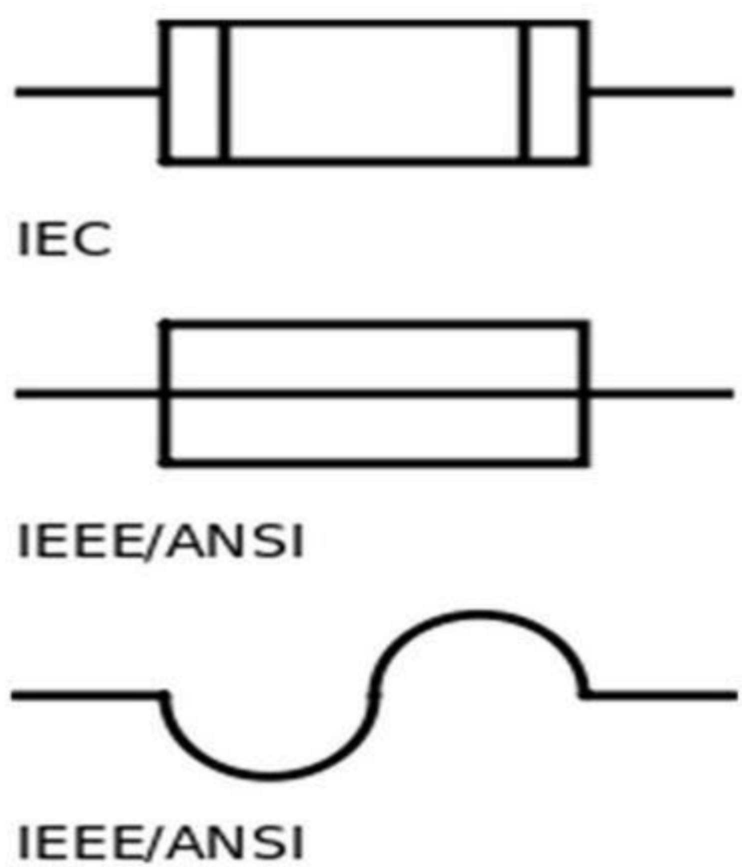
- **Resettable Fuse PTC (Positive temperature coefficient)**

Using a resettable fuse is $0.30, slightly more space than the diode and is cheaper than a full voltage regulator solution.

The PTC will cut off power if there is a current draw of more than 250mA. Removing the problem or short will allow the PTC to cool off and current will flow again. More info in this tutorial. Think about this carefully as the extra peripherals may add up to 250mA causing the PTC to trip incorrectly. We love using PTCs to help protect the electronics from short-circuit failures, however many voltage regulators already have this feature built-in.

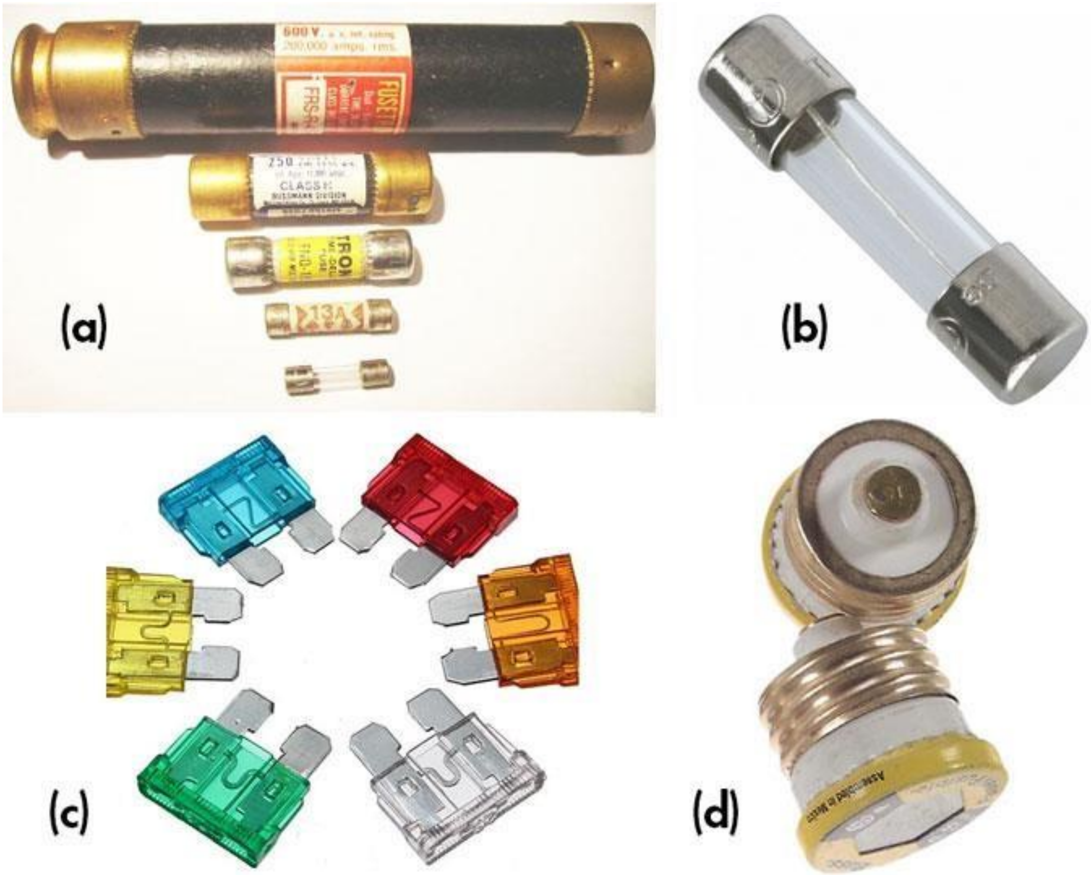
- **Overcurrent protection and Fuses**

Whether a result of a short circuit external to the supply or within it, overcurrent is a major concern. It can initiate a cascade of additional failures, put users at risk, and even start a fire. The oldest solution is a fuse (also called a fusible link) *(Fig. 1)* with apparently simple operation: when the current flow exceeds the fuse's current threshold, the current causes the special wire within the fuse to overheat (I^2R heating), melt, and open, thus cutting the current to zero.



1. Many schematic-diagram symbols exist for the fuse; these are a few of them. (Source: Slideplayer.com) Once the fuse blows open, the current flow is completely cut off, and can be restored only by replacing the fuse itself, which is either a benefit or a negative, depending on the application. The more-complex circuit breaker is an alternative to the fuse which doesn't need replacement after activation. Some breakers are thermally activated, some are magnetically activated; either way, like the fuse, the breaker is a current-triggered device.

Although the fuse is "ancient," it's inexpensive, reliable, easy to design in, and effective. Basic fuses are available with ratings under 1 A to hundreds of amps *(Fig. 2)*. While fuses do have a voltage rating, that's primarily for contact rating and physical spacing, as the fuse itself is triggered only by the current through it and not the voltage.



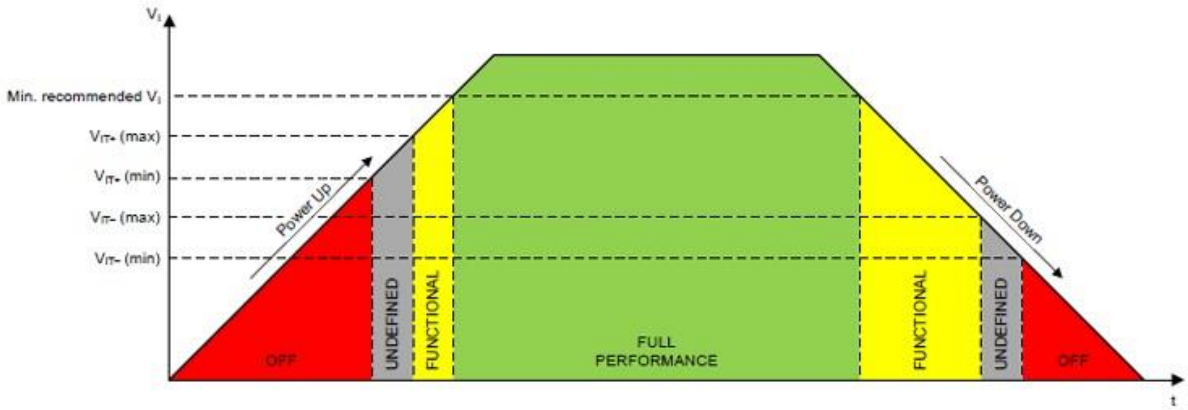
2. Fuses come in a wide range of form factors and current/voltage ratings (not to scale): Larger fuses (50 A and higher) are often housed in cylinders called cartridges (a); low-current “3AG” fuses for up to about 250 V ac (b); blade-type 15- and 20-A fuses commonly used for car circuits (12 V dc) (c); and oldfashioned “S” and “T”-type screw-in fuses rated to 20 and 30 A used in 120-V ac power lines) (d). (Image sources: Sunstore/UK; Source: Electrical Wholesaler/Ireland; RONA Langdon Hardware Ltd/Canada; and reviseOmatic.org)

For some devices, the fuse isn’t a good choice (think of a smartphone’s limited-energy internal power circuits), while it’s the best choice in others, and often used in conjunction with other protection techniques. The fuse is frequently added to help a product meet regulatory safety requirements, due to the directness of its functionality.

Note that despite their simple principle, they’re offered in many variations and subtleties, such as how long it takes for it to react and open the circuit (which is a function of both the current and elapsed time). Fuse datasheets have many charts showing performance under various conditions, and specialty fuses are available for unique situations.

- **Undervoltage Lockout (UVLO)**

UVLO ensures that a power-supply (or dc-dc) converter doesn’t attempt to operate when its own input voltage is too low *(Fig. 3)*. This is done for two reasons. First, circuitry within the supply or converter may malfunction or act in an indeterminate way if the input dc voltage is too low, and some higher-power components may actually be damaged. Second, it prevents the supply/converter from drawing on primary power if it can’t produce valid output power.



3. A power supply doesn't "instantly" come up to full output, but instead has turn-on and turn-off transitional ranges and time. UVLO ensures the supply doesn't attempt to provide a full output when its input voltage is below the minimum needed for proper operation. (Source: Texas Instruments)
To implement UVLO, a small, low-power comparison circuit within the supply/converter compares the input voltage to a preset threshold and puts the unit into quiescent mode until the threshold is crossed. To ensure that the UVLO doesn't "chatter" around the threshold, a small amount of hysteresis is added.

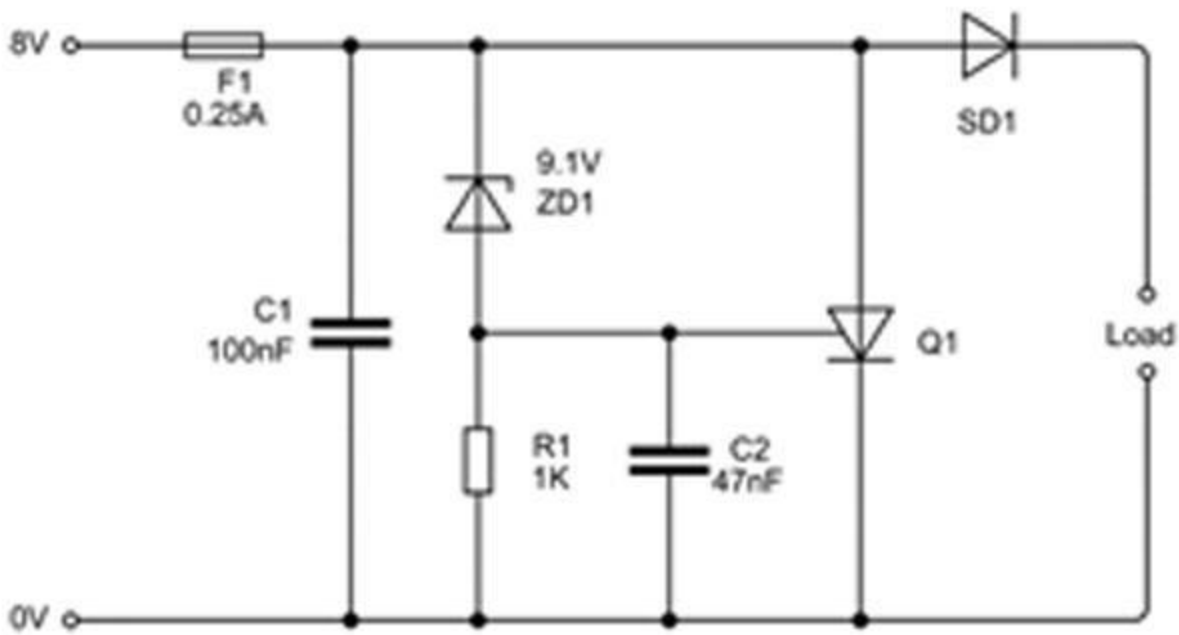
- **Overvoltage Protection (OVP)**

Although a supply or power converter is designed to normally produce a fixed dc-output voltage, an internal failure in the supply may cause this voltage to rise, and possibly damage the load to which the supply is connected. OVP is a function that monitors the supply/converter output versus an internal reference and short-circuits that output if the voltage rises above the threshold. The OVP must do several things:

- Obviously, prevent any excessive voltage from appearing at the protected components.
- Not interfere with normal operation, but instead be "invisible" to the power supply.
- Distinguish between normal transient voltage fluctuations and excessive overvoltage.
- Be fast, and respond before the load is damaged when a genuine overvoltage situation does occur.
- Not have false positives (false trips), which are a nuisance, and not fail to respond to real overvoltage conditions.

➢ ***The Crowbar***

One widely used OVP function is the "crowbar," supposedly so named because it has the same effect as placing a metal crowbar across the output and thus shorting the output voltage. There are two kinds of crowbars: one where the crowbar, once tripped, will only be reset if the power is turned off; and one where it will reset itself once the fault is cleared. The second one is useful when the condition that tripped the crowbar is due to some sort of transient rather than a hard failure in the supply. While most supplies now come with a built-in crowbar, many vendors offer a small, separate crowbar circuit that can be added to an existing supply if needed.
The crowbar is a normally high-impedance circuit across the supply output (or input of the load to be protected) *(Fig. 4)*. It transforms into a low-impedance circuit when an overvoltage situation occurs and triggers it, and it stays in low-impedance mode until the current decreases below the "holding current." Subsequently, it returns to the high-impedance, normal-operation state. The crowbar must be able to handle the current flowing through it during the time the supply is in overvoltage state.



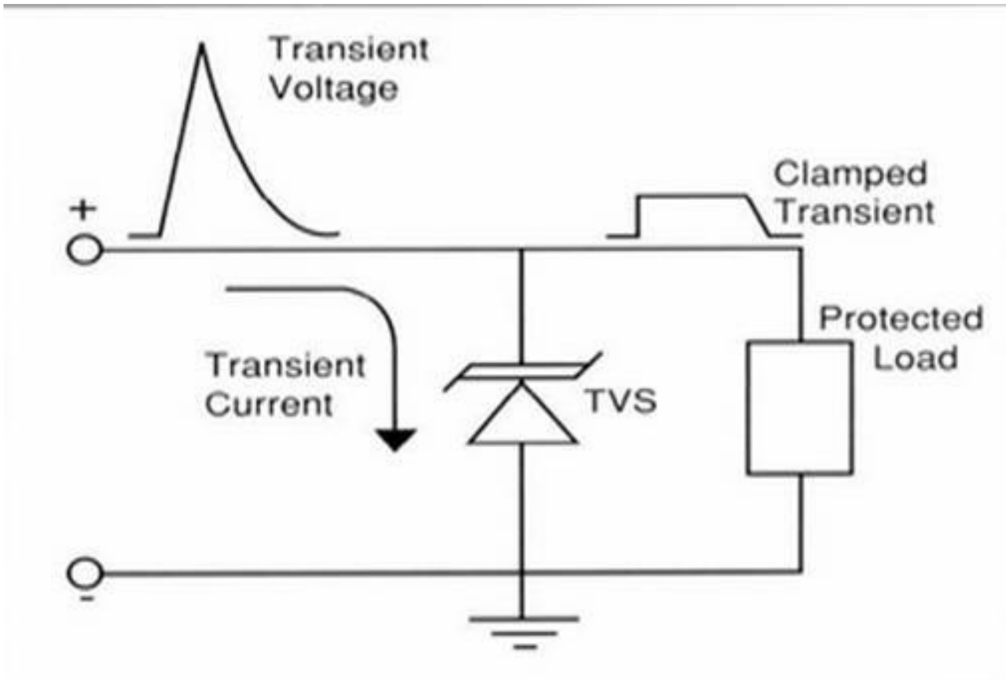
4. This crowbar circuit operates from an 8-V supply. The Zener diode sets overprotection at 9.1 V at that voltage; the diode starts to conduct, causing a trigger signal to switch on the thyristor Q1 (note that the fuse is for protection against excessive current).

Other common crowbars are based on thyristor surge protectors (TSPs). These are silicon-based PNPN devices with a breakdown voltage that can be set precisely by their manufacturer. TSPs are offered in many package types and can dissipate various levels of surges.

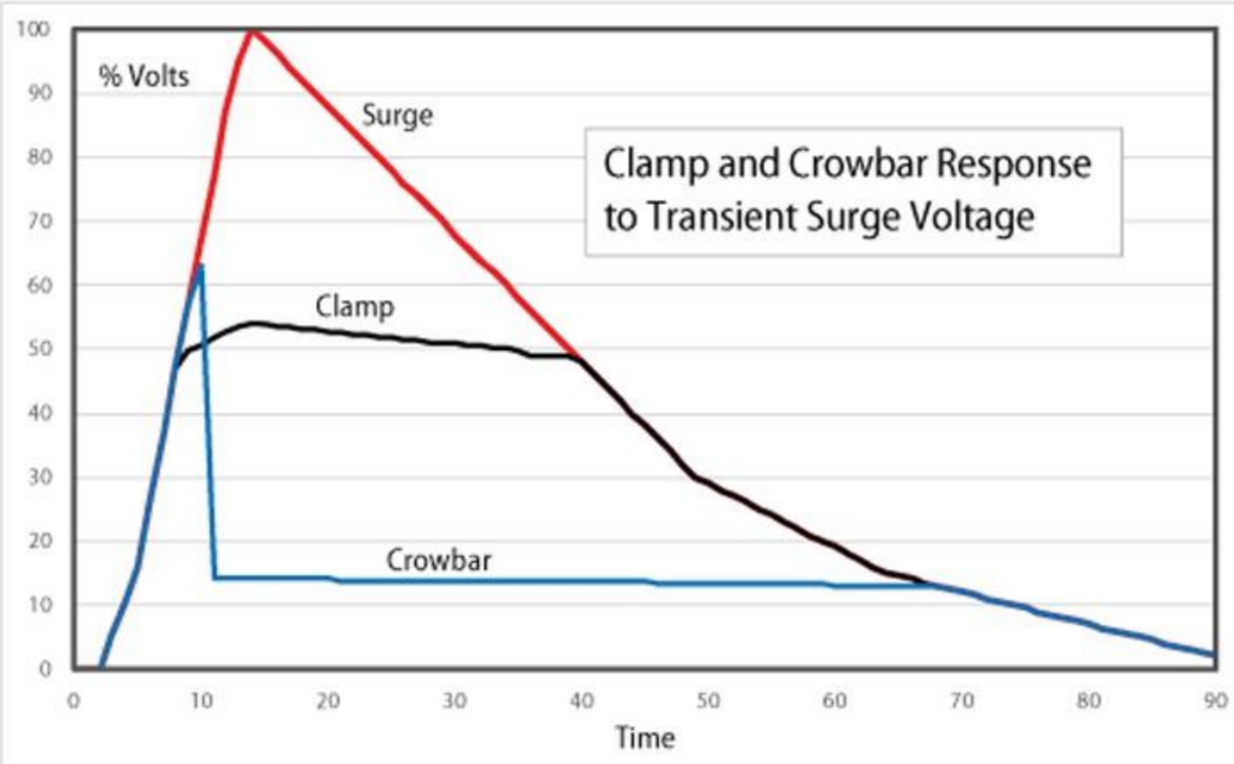
There's also the gas discharge tube (GDT), which is a miniature spark gap usually housed in a ceramic enclosure and PCB-compatible. When triggered by a high voltage, the spark gap conducts and all current flow is diverted. Spark gaps can be manufactured so that they protect from modest voltages (around 100 V) to thousands of volts. When the overvoltage situation clears, the TSP or GDT go back to normal, high-impedance mode.

➢ ***The Clamp***

A complement to the crowbar is the clamp, which prevents the voltage from exceeding a preset level. Clamps are often referred to as transient voltage suppressors (TVSs), since they may be protecting against a startup transient or inductive transient rather than an actual failure *(Fig. 5)*. For most clamps, the clamp function releases when the overvoltage condition clears.



5. The TVS, which is simple to apply, is placed between the voltage source and the load without any interfering components that might affect its performance or impede the current path.
A clamp conducts just enough current to maintain the voltage across it at a safe, desired value when the transient is above the clamp's conduction voltage. It must be rated for the power it will have to dissipate for a specific time, usually a relatively short transient event. The TVS clamp—a silicon bipolar junction device similar to a basic rectifier diode but designed to survive reverse breakdown-voltage situations—is available with breakdown voltages from 4 to 500 V, and in various power ratings to provide different surge-protection capabilities. A TVS is a bipolar junction device.
Compared to a clamp, the crowbar's low holding voltage lets it carry higher fault current without dissipating much power, so that it can handle higher currents and do so for longer periods *(Fig. 6)*.
It's also easier to configure the circuit so that the crowbar also causes a fuse to blow (and thus stop current flow completely), if that's desired.



6. The basic response of a crowbar and a clamp to a short-lived surge shows how the crowbar goes to a near short-circuit while the clamp limits to voltage increase. (Source: Bourns)

A clamp can also be built using a metal-oxide varistor (MOV), a bidirectional semiconductor voltage-transient suppressor device. It conducts (i.e., switches) at a voltage related to the size and number of special grains between its leads. MOV breakdown voltages range from about 14 V to over a 1,000 V, with the larger ones intended to handle several kilovolt-amps (kVA), such as from a lightning surge.

MOVs are low cost, fast acting, easy to use, and offered in many voltage ratings, and their own failure mode is to short circuit (which is preferred in most failsafe designs). However, they can only dissipate small amounts of power, so they're suited only for short-term and transient OVP situations

In general, crowbars are better for long-term faults, while clamps are best suited for transient events rather than outright supply failures. Many commercial power supplies incorporate both a crowbar and a clamp. If the concern is outright failure and associated high-current flow, which would soon overwhelm the dissipation rating of the crowbar or clamp, the design should also include a fuse or circuit breaker. The fuse/breaker will eventually blow from the overcurrent related to the excess voltage and thus provide multi-factor protection.

- **Thermal Protection**

Finally, there's the issue of thermal-overload protection. By its nature, any power supply generates heat because it's less than 100% efficient, and even an efficient supply generates a potentially troublesome amount. For example, a 100-W supply that's 90% efficient still dissipates 10 W, which is very capable of warming up a small, sealed enclosure. For this reason, the supply must be designed with sufficient active cooling (e.g., via a fan) or passive cooling (achieved by convection air flow and conductive cooling paths).

But what happens when the fan fails, the air-flow path is blocked, or another heat source is introduced into the enclosure? The supply may exceed its temperature rating, which shortens its life and may even cause immediate malfunction. The solution is a sensor within the supply (as a discrete device or incorporated within an IC) that senses the ambient temperature and puts the supply into a quiescent mode if it exceeds a preset limit. Some implementations allow the supply to resume operation if the temperature drops, while others do not.

Power-supply protection is, not surprisingly, a nuanced topic. There are issues of current, voltage, and power handling, dissipation by the protection circuit or components, and fault duration, as well as protection component placement, cost, and footprint. But protection is also good engineering practice and often mandated by regulatory standards. Again, it's like insurance: It comes in many forms and covers many types of bad events. You hope you don't need it, but there's a chance you will for a variety of possible reasons.

C

Switch Mode Power Supply

Linear voltage IC regulators have been the basis of power supply designs for many years as they are very good at supplying a continuous fixed voltage output.

Linear voltage regulators are generally much more efficient and easier to use than equivalent voltage regulator circuits made from discrete components such a zener diode and a resistor, or transistors and even op-amps.

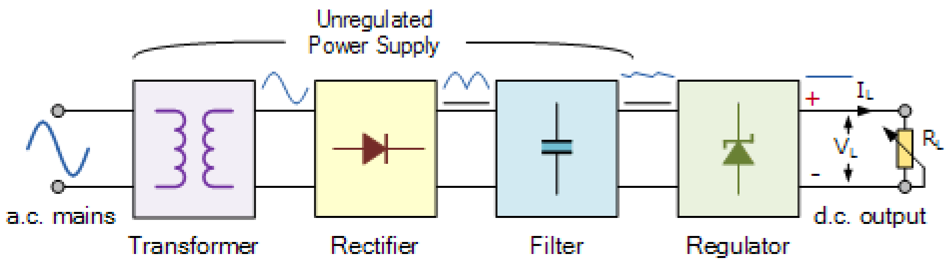
The most popular linear and fixed output voltage regulator types are by far the 78… positive output voltage series, and the 79… negative output voltage series. These two types of complementary voltage regulators produce a precise and stable voltage output ranging from about 5 volts up to about 24 volts for use in many electronic circuits.

There is a wide range of these three-terminal fixed voltage regulators available each with its own built-in voltage regulation and current limiting circuits. This allows us to create a whole host of different power supply rails and outputs, either single or dual supply, suitable for most electronic circuits and applications.

There are even variable voltage linear regulators available as well providing an output voltage which is continually variable from just above zero to a few volts below its maximum voltage output.

Most DC power supplies comprise of a large and heavy step-down mains transformer, diode rectification, either full-wave or half-wave, and a filter circuit to remove any ripple content from the rectified DC to produce a suitably smooth DC output voltage. Also, some form of voltage regulator or stabiliser circuit, either linear or switching can be used to ensure the correct regulation of the power supplies output voltage under varying load conditions. Then a typical DC power supply would look something like this:

Typical DC Power Supply

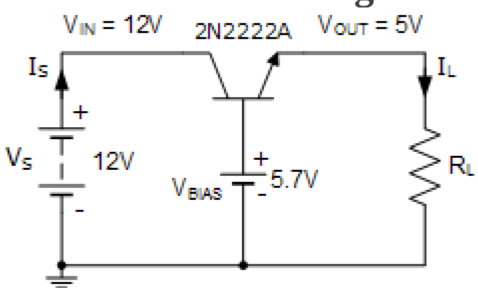


These typical power supply designs contain a large mains transformer (which also provides isolation between the input and output) and a series regulator circuit. The regulator circuit could consist of a single zener diode or a three-terminal linear series regulator to produce the required output voltage. The advantage of a linear regulator is that the power supply circuit only needs an input capacitor, output capacitor and some feedback resistors to set the output voltage.

Linear voltage regulators produce a regulated DC output by placing a continuously conducting transistor in series between the input and the output operating it in its linear region (hence the name) of its current-voltage (i-v) characteristics.

Thus the transistor acts more like a variable resistance which continually adjusts itself to whatever value is needed to maintain the correct output voltage. Consider this simple series pass transistor regulator circuit below:

Series Transistor Regulator Circuit



Here this simple emitter-follower regulator circuit consists of a single NPN transistor and a DC biasing voltage to set the required output voltage. As an emitter follower circuit has unity voltage gain, applying a suitable biasing voltage to the transistors base, a stabilised output is obtained from the emitter terminal.

Since a transistor provides current gain, the output load current will be much higher than the base current and higher still if a Darlington transistor arrangement is used.

Also, providing that the input voltage is sufficiently high enough to get the desired output voltage, the output voltage is controlled by the transistors base voltage and in this example is given as 5.7 volts to produce a 5 volt output to the load as approximately 0.7 volts is dropped across the transistor between the base and emitter terminals. Then depending upon the value of the base voltage, any value of emitter output voltage can be obtained.

While this simple series regulator circuit will work, the downside to this is that the series transistor is continually biased in its linear region dissipating power in the form of heat. Since all of the load current must pass through the series transistor, this results in a poor efficiency, wasted V*I power and continuous heat generation around the transistor.

Also, one of the disadvantages that series voltage regulators have is that their maximum continuous output current rating is limited to just a few amperes or so, so are generally used in applications where low power outputs are required.

When higher output voltage or current power demands are required, the normal practice is to use a switching regulator commonly known as a *switch-mode power supply* to convert the mains voltage into whatever higher power output is required.

Switch Mode Power Supplies, or **SMPS**, are becoming common place and have replaced in most cases the traditional linear AC-to-DC power supplies as a way to cut power consumption, reduce heat dissipation, as well as size and weight.

Switch-mode power supplies can now be found in most PC's, power amplifiers, TV's, dc motor drives, etc., and just about anything that requires a highly efficient supply as switch-mode power supplies are increasingly becoming a much more mature technology.

By definition, a switch mode power supply (SMPS) is a type of power supply that uses semiconductor switching techniques, rather than standard linear methods to provide the required output voltage. The basic switching converter consists of a power switching stage and a control circuit.

The power switching stage performs the power conversion from the circuits input voltage, V_{IN} to its output voltage, V_{OUT} which includes output filtering.

The major advantage of the switch mode power supply is its higher efficiency, compared to standard linear regulators, and this is achieved by internally switching a transistor (or power MOSFET) between its "ON" state (saturated) and its "OFF" state (cut-off), both of which produces lower power dissipation.

This means that when the switching transistor is fully "ON" and conducting current, the voltage drop across it is at its minimal value, and when the transistor is fully "OFF" there is no current flow through it. So the transistor is acting like an ideal ON/OFF switch.

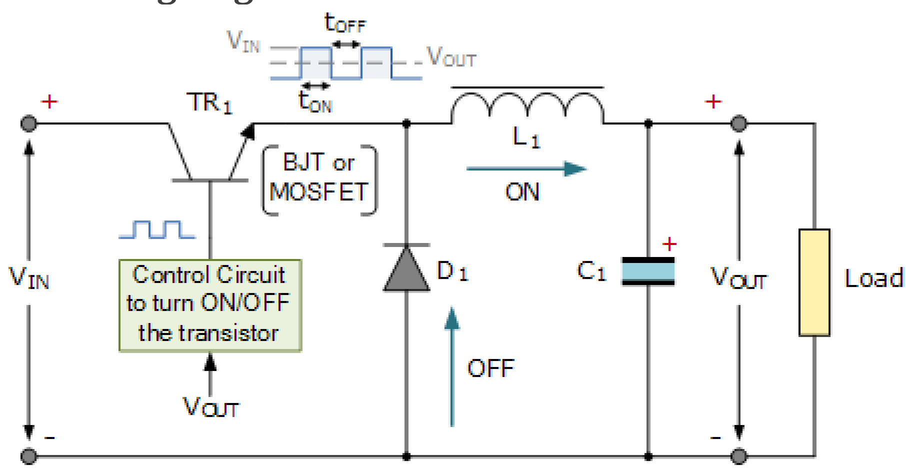
Unlike linear regulators which only offer step-down voltage regulation, a switch mode power supply can provide step-down, step-up and negation of the input voltage using one or more of the three basic switch mode circuit topologies: *Buck*, *Boost* and *Buck-Boost*. These names refer to how the transistor switch, inductor, and smoothing capacitor are connected together within the basic SMPS circuit.

Buck Switch Mode Power Supply

The **Buck switching regulator** is a type of switch mode power supply circuit that is designed to efficiently reduce DC voltage from a higher voltage to a lower one, that is it subtracts or "Bucks" the supply voltage, thereby reducing the voltage available at the output terminals without changing the polarity. In other words, the buck switching regulator is a step-down regulator circuit, so for example a buck converter can convert say, +12 volts to +5 volts.

The buck switching regulator is a DC-to-DC converter and one of the simplest and most popular type of switching regulator. When used within a switch mode power supply configuration, the buck switching regulator uses a series transistor or power MOSFET (ideally an insulated gate bipolar transistor, or IGBT) as its main switching device as shown below.

The Buck Switching Regulator



We can see that the basic circuit configuration for a buck converter is a series transistor switch, TR_1 with an associated drive circuit that keeps the output voltage as close to the desired level as possible, a diode, D_1, an inductor, L_1 and a smoothing capacitor, C_1. The buck converter has two operating modes, depending on if the switching transistor TR_1 is turned "ON" or "OFF".

When the transistor is biased "ON" (switch closed), diode D_1 becomes reverse biased and the input voltage, V_{IN} causes a current to flow through the inductor to the connected load at the output, charging up the capacitor, C_1.

As a changing current flows through the inductor coil, it produces a back-emf which opposes the flow of current, according to Faraday's law, until it reaches a steady state creating a magnetic field around the inductor, L_1. This situation continues indefinitely as long as TR_1 is closed.

When transistor TR_1 is turned "OFF" (switch open) by the controlling circuitry, the input voltage is instantly disconnected from the emitter circuit causing the magnetic field around the inductor to collapse inducing a reverse voltage across the inductor. This reverse voltage causes the diode to become forward biased, so the stored energy in the inductors magnetic field forces current to continue to flow through the load in the same direction, and return back through diode.

Then the inductor, L_1 returns its stored energy back to the load acting like a source and supplying current until all the inductor's energy is returned to the circuit or until the transistor switch closes again, whichever comes first. At the same time the capacitor also discharges supplying current to the load. The combination of the inductor and capacitor forms an LC filter smoothing out any ripple created by the switching action of the transistor.

Therefore, when the transistor solid state switch is closed, current is supplied from the supply, and when the transistor switch is open, current is supplied by the inductor. Note that the current flowing through the inductor is always in the same direction, either directly from the supply or via the diode but obviously at different times within the switching cycle.

As the transistor switch is being continuously closed and opened, the average output voltage value will therefore be related to the duty cycle, D which is defined as the conduction time of the transistor switch during one full switching cycle.

If V_{IN} is the supply voltage, and the "ON" and "OFF" times for the transistor switch are defined as: t_{ON} and t_{OFF}, then the output voltage V_{OUT} is given as:

Buck Converter Duty Cycle

$$V_{OUT} = \frac{t_{ON}}{(t_{ON} + t_{OFF})} \times V_{IN}$$

The buck converters duty cycle can also be defined as:

$$D = \frac{t_{ON}}{(t_{ON} + t_{OFF})} = \frac{t_{ON}}{\text{Total Time}} = \frac{t_{ON}}{T}$$

$$\therefore D \approx \frac{V_{OUT}}{V_{IN}} \quad \text{or} \quad V_{OUT} = D V_{IN}$$

So the larger the duty cycle, the higher the average DC output voltage from the switch mode power supply. From this we can also see that the output voltage will always be lower than the input voltage since the duty cycle, D can never reach one (unity) resulting in a step-down voltage regulator.

Voltage regulation is obtained by varying the duty cycle and with high switching speeds, up to 200kHz, smaller components can be used thereby greatly reducing a switch mode power supply's size and weight.

Another advantage of the buck converter is that the inductor-capacitor (LC) arrangement provides very good filtering of the inductor current. Ideally the buck converter should be operated in a continuous switching mode so that the inductor current never falls to zero. With ideal components, that is zero voltage drop and switching losses in the "ON" state, the ideal buck converter could have efficiencies as high as 100%.

As well as the step-down buck switching regulator for the basic design of a switch mode power supply, there is another operation of the fundamental switching regulator that acts as a step-up voltage regulator called the Boost Converter.

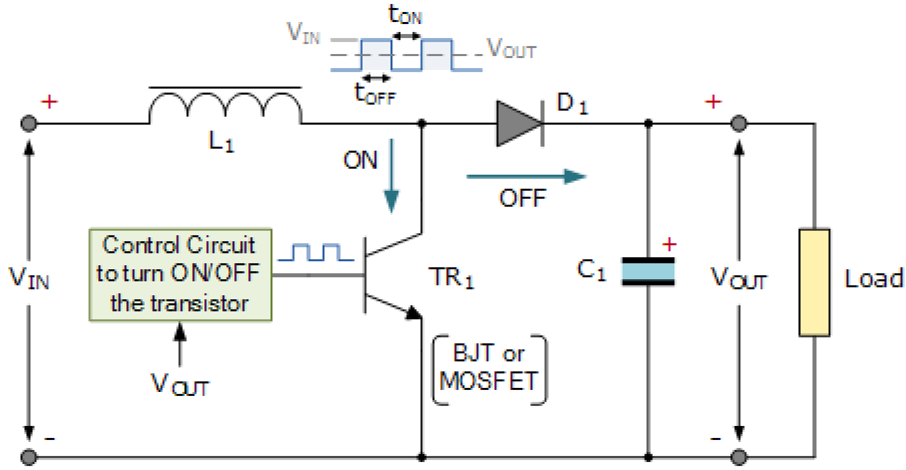
Boost Switch Mode Power Supply

The **Boost switching regulator** is another type of switch mode power supply circuit. It has the same types of components as the previous buck converter, but this time in different positions. The boost converter is designed to increase a DC voltage from a lower voltage to a higher one, that is it adds too or "Boosts" the supply voltage, thereby increasing the available voltage at the output terminals without changing the polarity. In other words, the boost switching regulator is a step-up regulator circuit, so for example a boost converter can convert say, +5 volts to +12 volts.

We saw previously that the buck switching regulator uses a series switching transistor within its basic design. The difference with the design of the *boost switching regulator* is that it uses a parallel connected switching transistor to control the output voltage from the switch mode power supply.

As the transistor switch is effectively connected in parallel with the output, electrical energy only passes through the inductor to the load when the transistor is biased "OFF" (switch open) as shown.

The Boost Switching Regulator



In the *Boost Converter* circuit, when the transistor switch is fully-on, electrical energy from the supply, V_{IN} passes through the inductor and transistor switch and back to the supply. As a result, none of it passes to the output as the saturated transistor switch effectively creates a short-circuit to the output.

This increases the current flowing through the inductor as it has a shorter inner path to travel back to the supply. Meanwhile, diode D_1 becomes reverse biased as its anode is connected to ground via the transistor switch with the voltage level on the output remaining fairly constant as the capacitor starts to discharge through the load.

When the transistor is switched fully-off, the input supply is now connected to the output via the series connected inductor and diode. As the inductor field decreases the

induced energy stored in the inductor is pushed to the output by V_{IN}, through the now forward biased diode.
The result of all this is that the induced voltage across the inductor L_1 reverses and adds to the voltage of the input supply increasing the total output voltage as it now becomes, $V_{IN} + V_L$.
Current from the smoothing capacitor, C_1 which was used to supply the load when the transistor switch was closed, is now returned to the capacitor by the input supply via the diode. Then the current supplied to the capacitor is the diode current, which will always be “ON” or “OFF” as the diode is continually switched between its forward and reverse status by the switching action of transistor. Then the smoothing capacitor must be sufficiently large enough to produce a smooth steady output.
As the induced voltage across the inductor L_1 is negative, it adds to the source voltage, V_{IN} forcing the inductor current into the load. The boost converters steady state output voltage is given by:

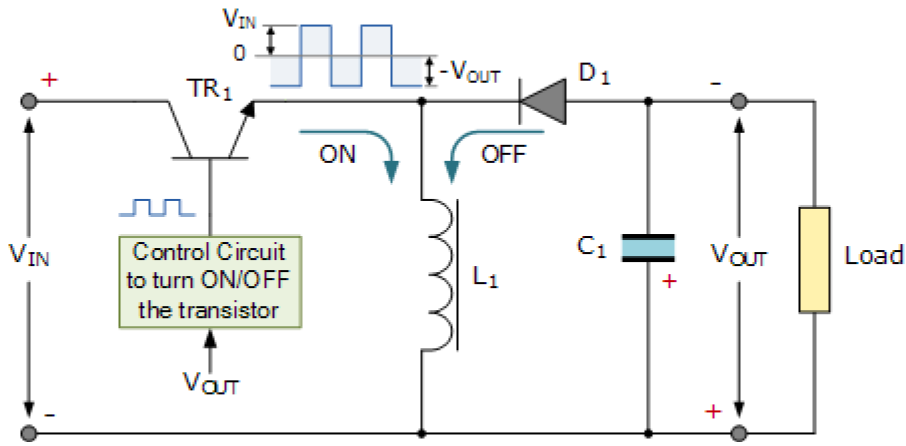
$$V_{OUT} = V_{IN}\frac{1}{(1-\text{duty cycle})} = V_{IN}\left(\frac{1}{1-D}\right)$$

As with the previous buck converter, the output voltage from the boost converter depends upon the input voltage and duty cycle. Therefore, by controlling the duty cycle, output regulation is achieved. Not also that this equation is independent of the value of the inductor, the load current, and the output capacitor.
We have seen above that the basic operation of a non-isolated switch mode power supply circuit can use either a buck converter or boost converter configuration depending upon whether we require a step-down (buck) or step-up (boost) output voltage. While buck converters may be the more common SMPS switching configuration, boost converters are commonly used in capacitive circuit applications such as battery chargers, photo-flashes, strobe flashes, etc, because the capacitor supplies all of the load current while the switch is closed.
But we can also combine these two basic switching topologies into a single non-isolating switching regulator circuit called unsurprisingly, a *Buck-Boost Converter*.

Buck-Boost Switching Regulator

The **Buck-Boost switching regulator** is a combination of the buck converter and the boost converter that produces an inverted (negative) output voltage which can be greater or less than the input voltage based on the duty cycle. The buck-boost converter is a variation of the boost converter circuit in which the inverting converter only delivers the energy stored by the inductor, L_1, into the load. The basic buck-boost switch mode power supply circuit is given below.

The Buck-Boost Switching Regulator



When the transistor switch, TR_1, is switched fully-on (closed), the voltage across the inductor is equal to the supply voltage so the inductor stores energy from the input supply. No current is delivered to the connected load at the output because diode, D_1, is reverse biased. When the transistor switch is fully-off (open), the diode becomes forward biased and the energy previously stored in the inductor is transferred to the load.

In other words, when the switch is "ON", energy is delivered into the inductor by the DC supply (via the switch), and none to the output, and when the switch is "OFF", the voltage across the inductor reverses as the inductor now becomes a source of energy so the energy stored previously in the inductor is switched to the output (through the diode), and none comes directly from the input DC source. So the voltage dropped across the load when the switching transistor is "OFF" is equal to the inductor voltage.

The result is that the magnitude of the inverted output voltage can be greater or smaller (or equal to) the magnitude of the input voltage based on the duty cycle. For example, a positive-to-negative buck-boost converter can convert 5 volts to 12 volts (step-up) or 12 volts to 5 volts (step-down).

The buck-boost switching regulators steady state output voltage, V_{OUT} is given as:

$$V_{OUT} = V_{IN}\left(\frac{D}{1-D}\right)$$

Then the buck-boost regulator gets its name from producing an output voltage that can be higher (like a boost power stage) or lower (like a buck power stage) in magnitude than the input voltage. However, the output voltage is opposite in polarity from the input voltage.

Switch Mode Power Supply Summary

The modern switch mode power supply, or SMPS, uses solid-state switches to convert an unregulated DC input voltage to a regulated and smooth DC output voltage at

different voltage levels. The input supply can be a true DC voltage from a battery or solar panel, or a rectified DC voltage from an AC supply using a diode bridge along with some additional capacitive filtering.
In many power control applications, the power transistor, MOSFET or IGFET, is operated in its switching mode were it is repeatedly turned “ON” and “OFF” at high speed. The main advantage of this is that the power efficiency of the regulator can be quite high because the transistor is either fully-on and conducting (saturated) or full-off (cut-off).
There are several types of DC-to-DC converter (as opposed to a DC-to-AC converter which is an inverter) configurations available, with the three basic switching power supply topologies looked at here being the *Buck*, *Boost*, and the *Buck-Boost* switching regulators. All three of these topologies are non-isolated, that is their input and output voltages share a common ground line.
Each switching regulator design has its own unique properties with regards to the steady-state duty cycles, relationship between the input and output current, and the output voltage ripple produced by the solid-state switch action. Another important property of these switch mode power supply topologies is the frequency response of the switching action to the output voltage.
Regulation of the output voltage is achieved by the percentage control of the time that the switching transistor is in the “ON” state compared to the total ON/OFF time. This ratio is called the duty cycle and by varying the duty cycle, (D the magnitude of the output voltage, V_{OUT} can be controlled.
The use of a single inductor and diode as well as fast switching solid-state switches capable of operating at switching frequencies in the kilohertz range, within the switch mode power supply design, allows for the size and weight of the power supply to be greatly reduced.
This is because there would be no large and heavy step-down (or step-up) voltage mains transformers within their design. However, if electrical isolation is required between the input and output terminals, a transformer must be included before the converter.
The two most popular non-isolated switching configurations are the buck (subtractive) and the boost (additive) converters.
The buck converter is a type of switch-mode power supply that is designed to convert electrical energy from one voltage to a lower one. The buck converter operates with a series connected switching transistor. As the duty cycle, $D < 1$, the output voltage of the buck is always smaller than the input voltage, V_{IN}.
The boost converter is a type of switch-mode power supply that is designed to convert electrical energy from one voltage to a higher one. The boost converter operates with a parallel connected switching transistor which results in a direct current path between V_{IN} and V_{OUT} via the inductor, L_1 and diode, D_1. This means there is no protection against short-circuits on the output.

By varying the duty cycle, (D) of a boost converter, the output voltage can be controlled and with $D < 1$, the DC output from the boost converter is greater than input voltage V_{IN} as a consequence of the inductors self-induced voltage.
Also, the output smoothing capacitors in **Switch-mode Power Supplies** is assumed to be very large, which results in a constant output voltage from the switch mode supply during the transistors switching action.