Bipolar Junction Transistors

1. Introduction

Transistors were invented by J.Bardeen and W.H.Brattain of bell laboratories in USA in 1947 and that was a point contact transistor.

William shockey first invented junction transistor in 1951 which consists of two back to back pn junctions.

These transistors are also known as bipolar junction transistor (BJI) and are our subject of study in this section.

Thus a jucntion transistor is formed by sandwiching a thin layet of p-type semi-conductor between two layers of n-type semi-conductor and vice versa.

Transistors are of two types

PNP transistors : Here the n-type thin layer is sandwiched between two p-type layes

NPN transistors : Here p-type thin layer is sandwiched between two n-layers

There are three regions in a transistors

emiiter

base

collector

Figure (1) shows digramatic representation of PNP and NPN transistors along with the sysmbols used to represent them

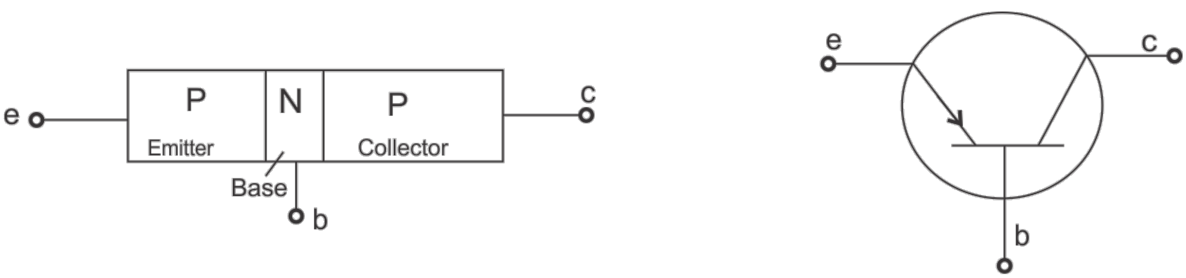


Figure 1(a) PNP transistor and its symbolic representation

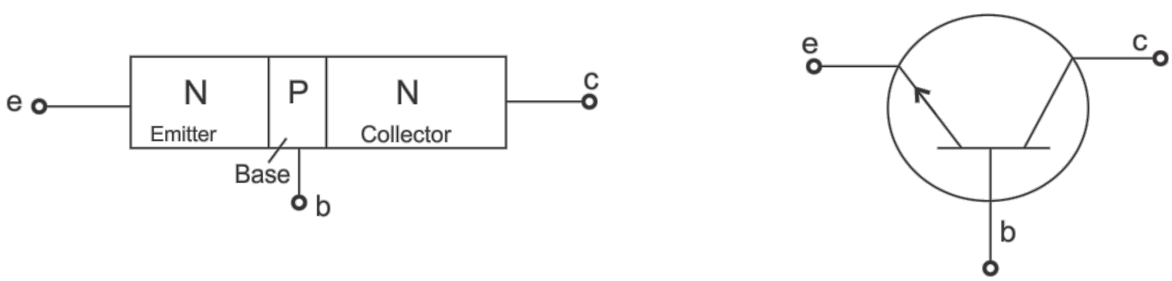


Figure 1(b) NPN transistor and its symbolic representation

All the three regions of the transistors have different thickness and their doping levels are also different.

In symoblic representation of transistors the arrow head shows the direction of conventional current in the transistor.

Brief discription of all the three region of the transistors

Emmitter: Emitter region of transistor is of moderate size and it is heavilly doped .Function of emiiter is to inject electrons or holes depedning on types of transitor into the base

Base: Middle region of transistor is known as base region.This region of transistor is very thin and very lightly doped.Function of the base is to pass most the electron or holes onto the collector

Collector: Collector region is moderately doped and is made physically larger than the emiiter since it is required to dissipate more heat .Collector has a job of collecting electrons or holes from the base.

Bipolar Junction Transistors

2. Action of transistor

Transistor has two junctions

Emiiter base (EB) junction

collector base(CB) junction

To operate the transistor a suitable potential difference must be applied across two of its junction .This is known as biasing of the transistor

The charge carriers move across different region of the transistor whne all three terminals of transistors are properly biased

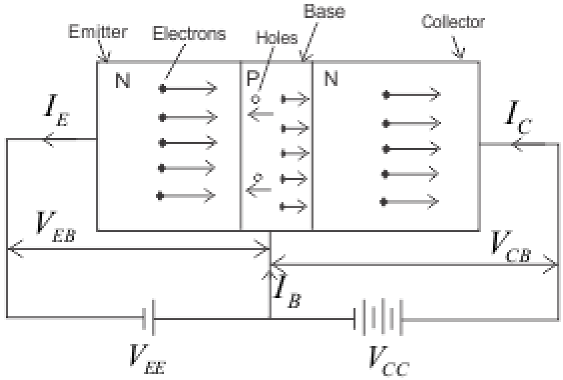
There are four possible ways to bias these functions given belwo in the table

Condition	EB junction	CB junction	Region of operation
1	FB	RB	Active
2	FB	FB	Saturation
3	RB	RB	cut off
4	RB	FB	Inverted

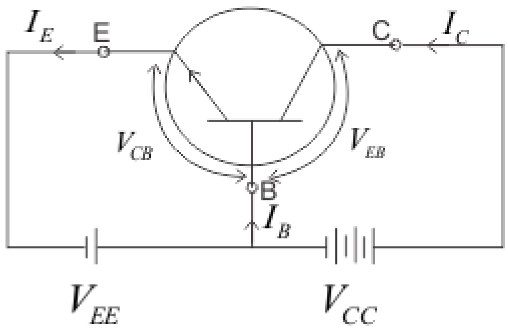
Only one of these conditions interests us at the moment that is condition 1 in which EB junction is FB and CB junction is RB

We will consider the case of npn transistor when its EB junction is FB and CB junction is RB.Case of pnp transistor can be also be understood on the same basis

the figure 2 shows that an npn transistor is properly biased using the two power supplies V_{EE} and V_{CC} and base terminal is common to both these power supplies



(a) Physical picture of electrons and holes in npn transistor



(b) Symbolic representation of an npn transistor

Figure 2 - Bias voltage applied on an npn transistor

Consider figure 2a ,since emmiter base terminal is FB the majority charge carriers there are electrons in this case enters the base region in large numbers

Base region which is very lightly doped is short of majority charge carriers which are holes in case of npn transistors.Since base is a p-type semiconductor ,thus large numbers of electrons entering the base from the emitter combines with the very small number of holes present there in base region

Since the base collector junction is RB so these electrons which are minorty charge carriers in base ,can easily cross the jucntion and enter the collectror region

These electron just entered the collector region moves forwards to reach the collector terminal

These electron go through batteries V_{cc} and V_{EE} and are taken back to emmiter

The electrons going from source V_{EE} to emitter constitute the electric current I_E in the direction oppostie to the direction of the flow of electron .This current is known as *emmitter current*.

Similary electron moving from the collector to the battery constitute the collectror current I_C

Collector current I_C is slightly less then that of emitter current I_E becuase some of the electrons comes out of the base terminal instead of going to the collector

This small fraction of current comming out of base terminal is known as base current I_B.Base current is a small fraction of the emiiter current

Using kirchoof law we can write

$I_E=I_C+I_B$ ---(1)

There is always a reverse leakage current in the transistor whioninorty charge carriers

α and β parameters for transistor are defined as

$\alpha=I_C/I_E$

$\beta =I_C/I_B$ ----(2)

These parameters are helpful while analyzing transistors as a circuit element

We have to keep one connection in mind that all currents entering into transistor are taken to be positive and current flowing out is negative

Thus for an NPN transistor

I_E -> negative

I_C ,I_B-> positive

Bipolar Junction Transistors

3. Transistor circuit configuration and transistor characterstics

There are three terminals in the transistor i.e emmitter,base and collector

In circuits input and output connections are made such that one of the terminals of the transistor is common to both the input and output

Keeping this in mind ,transistor can be connected in any of the three following configuration

Common emmitter configuration(CE)

Common base configuration(CB)

common collector configuration(CC)

In all these configurations voltage connection are made such thata EB jucntion is always foward baised and CB junction is always reverse biased so that transistor works in the active region

Out of all the three transistor configuration CE is most widely used so our discussion would be restricted to this configuration only

Bipolar Junction Transistors

3(a). Common Emitter transistor configuration

In CE configuration emitter is made common to both input and output or the input is between the base and emitter and output is between emitter and collector.

Figure below shows an transistor circuit using npn transistor in CE mode

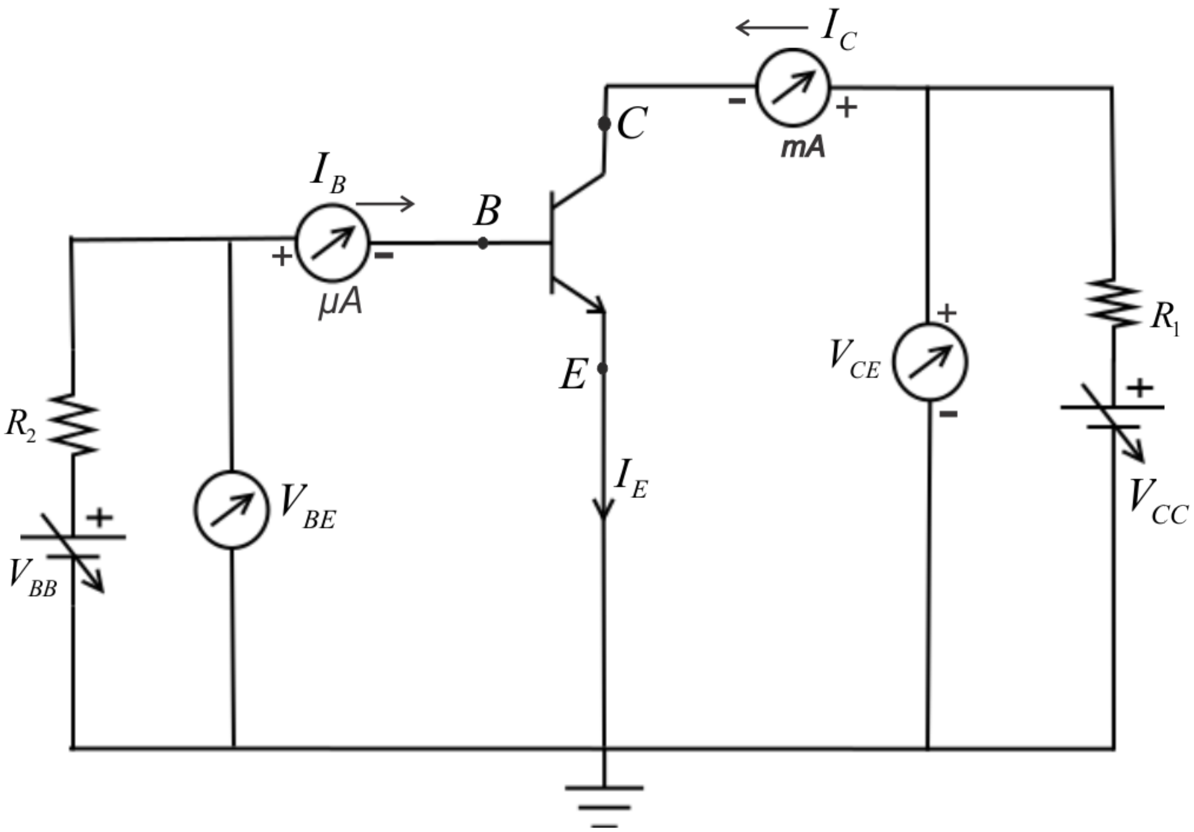


Figure 3- Circuit arrangement for studying input and output characteristics of npn transistor in CE configuration

Input section of the circuit contains base emitter base junction .Thus two input variables are

I_B ,the base current

V_{BE} ,the voltage in the input section

Similarly output section of the circuit contains CE junction .Thus two output varibel are

I_C ,the collector current

V_{CE} ,the collector emitter voltage

Transistor characterstics are the curves that relate transistor currents and voltage .There are two types of characterstics curves

Input characterstics :- relates input current with the input voltage for a given output voltage

Output characterstics :- relates output current with output voltage for a given input current

We now know what are transistor characterstics and now we shall discuss the characterstics curves for a npn transistor in CE configuration (connection known in fig 3)

i) Input characterstics:-

To study input characterstics of transistors in CE configuration a curve is plotted against base current I_B and base emitter voltage V_{BE} keeping collector emitter voltage V_{CE} fixed

Input characterstics relate I_B and V_{BE} for different values of V_{CE}

Graph below shows the input characterstics of npn transisator which are similar to those of a forward biased pn junction FIG

Change in output voltage V_{CE} does not result in lrage deviation of the curves

To keep transistor in active state ,it is required to keep V_{CE} large enough to make CB junction reverse biased

Dynamic resistance of the transistor at a given V_{CE} is $r_i=\Delta V_{BE}/\Delta I_B$ for V_{CE}=constant ---(3) input resistnace r_i is defined as the ratio of change in base emitter voltage (ΔV_{BE}) to the resulting change in base current (ΔI_B) at constant CE voltage(V_{CE})

Value of r_i can vary from few hundered to few thousand ohms

ii) Output characterstics:-

These curves relate output current I_C to the voltage between collector and emmitter for various value sof the input current I_B

Output characterstics are shown below in figure 5 FIG

In active region ,I_C increase slowly as V_{CE} incraeses

From output chracterstics we can determine dynamic output resistnace whihc is defined as the ratio of the change in CE voltgae (ΔV_{CE}) to the change in collector current (ΔI_C) at a constant base current I_B Thus,

$r_o=\Delta V_{CE}/\Delta I_C$ for I_B=constant ---(4)

current amplification factor (B) is defined as the ratio of change in collector current to the change in base current at constant CE voltage => $B_{ac}=\Delta I_C/\Delta I_B$ at V_{CE}=constant ---(5) B_{ac} is also known as ac current gain

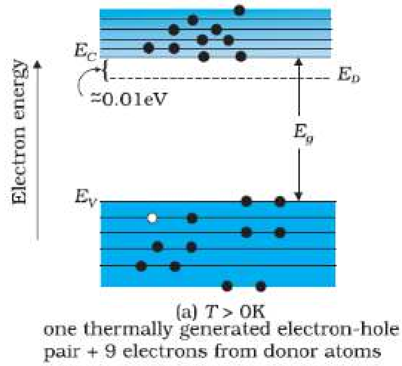
Simple ratio of I_C and I_B is known as DC current gain.Thus $B_{dc}=I_C/I_B$ ---(6) From fig (5) it is clear that for constant I_B ,current I_C increases with V_{CE}. This indicates that B_{dc} incraese with V_{CE}

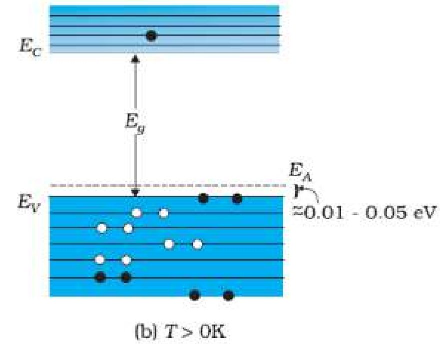
When V_{CE} drops below V_{BE},the CB junction becomes forward baised and I_C decraese rapidly with V_{CE} is

When both junctions are FB ,transistors works in saturation region in this region I_C no longer depends on I_B

Inactive region collector region is B_{dc} times greater than I_B.This shows that small input current I_B produces large output current I_C

- **Intrinsic Semiconductor:**
The pure semiconductors in which the electrical conductivity is totally governed by the electrons excited from the valence band to the conduction band and in which no impurity atoms are added to increase their conductivity are called intrinsic semiconductors and their conductivity is called intrinsic conductivity. Electrical conduction in pure semiconductors occurs by means of electron-hole pairs. In an intrinsic semiconductor,
$$n_e = n_h = n_i$$
where n_e = the free electron density in conduction band, n_h = the hole density in valence band, and n_i = the intrinsic carrier concentration.
- **Extrinsic Semiconductors:**
A Semiconductor doped with suitable impurity atoms so as to increase its conductivity is called an extrinsic semiconductor.
- **Types of Extrinsic Semiconductors:**
Extrinsic semiconductors are of two types
i) n-type semiconductors
ii) p-type semiconductors
- **n-type semiconductors:**
The pentavalent impurity atoms are called donors because they donate electrons to the host crystal and the semiconductor doped with donors is called n-type semiconductor. In n-type semiconductors, electrons are the majority charge carriers and holes are the minority charge carriers. Thus,
$$n_e >> n_h$$
- **p-type semiconductors:**
The trivalent impurity atoms are called acceptors because they create holes which can accept electrons from the nearby bonds. A semiconductor doped with acceptor type impurities is called a p-type semiconductor. In p-type semiconductor, holes are the majority carriers and electrons are the minority charge carriers Thus,
$$n_h >> n_e$$





Energy Band Diagram of n-type semiconductor   Energy band Diagram of p-type semiconductor

- **Holes:**
  The vacancy or absence of electron in the bond of a covalently bonded crystal is called a hole. A hole serves as a positive charge carrier.
- **Mobility:**
  a) The drift velocity acquired by a charge carrier in a unit electric field is called its electrical mobility and is denoted by μ .

  $$\mu = \frac{V_d}{E}$$

  b) The mobility of an electron in the conduction band is greater than that of the hole (or electron) in the valence band.
- **Electrical conductivity of a Semiconductor:**
  a) If a potential difference V is applied across a conductor of length L and area of cross-section A, then the total current I through it is given by,

  $$I = eA(n_e v_e + n_h v_h)$$

  where n_e and n_h are the electron and hole densities, and v_e and v_h are their drift velocities, respectively.

  b) If μ_h are the electron and hole mobilities, then the conductivity of the semiconductor will be,

  $$\rho = e(n_e \mu_e + n_h \mu_h)$$

  and the resistivity will be,

  $$\rho = \frac{1}{e(n_e \mu_e + n_h \mu_h)}$$

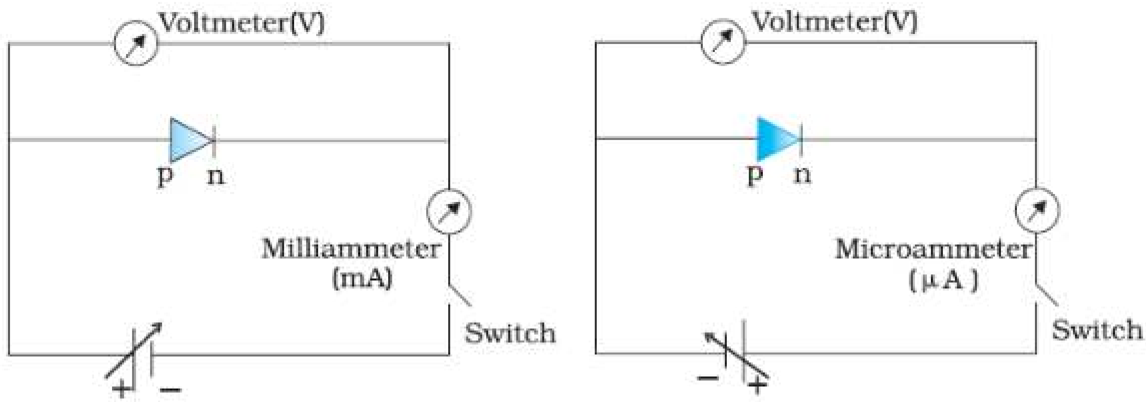
  c) The conductivity of an intrinsic semiconductor increases exponentially with temperature as,

  $$\sigma = \sigma_0 \exp\left(-\frac{E_g}{2k_B T}\right)$$

- **Forward Biasing of a pn-junction:**
  If the positive terminal of a battery is connected to the p-side and the negative terminal to the n-side, then the pn-junction is said to be forward biased. Both electrons and holes

move towards the junction. A current, called forward current, flows across the junction. Thus a pn-junction offers a low resistance when it is forward biased.

- **Reverse Biasing of a pn-junction:**
  If the positive terminal of a battery is connected to the n-side and negative terminal to the p-side, then pn-junction is said to be reverse biased. The majority charge carriers move away from the junction. The potential barrier offers high resistance during the reverse bias. However, due to the minority charge carriers a small current, called reverse or leakage current flows in the opposite direction. Thus junction diode has almost a unidirectional flow of current.



(a) Forward Bias: pn Junction (b) Reverse Bias: pn Junction

- **Action of a transistor:**
  When the emitter-base junction of an npn-transistor is forward biased, the electrons are pushed towards the base. As the base region is very thin and lightly doped, most of the electrons cross over to the reverse biased collector. Since few electrons and holes always recombine in the base region, so the collector current I_c is always slightly less then emitter current I_E.

$$I_E = I_C + I_B$$

  Where I_B is the base current.

- **Three Configurations of a Transistor:**
  A transistor can be used in one of the following three configurations:
  a) Common-base (CB) circuit.
  b) Common-emitter (CE) circuit.
  c) Common-collector (CC) circuit.

- **Current Gains of a Transistor:**
  Usually low current gains are defined:
  a) **Common base current amplification factor or ac current gain α:**

  It is the ratio of the small change in the collector current to the small change in the emitter current when the collector-base voltage is kept constant.

$$\alpha = \left[\frac{\delta I_C}{\delta I_E}\right]_{V_{CB=\text{constant}}}$$

b) **Common emitter current amplification factor or ac current gain β:**

It is the ratio of the small change in the collector current to the small change in the base current when the collector emitter voltage is kept constant.

$$\beta = \left[\frac{\delta I_C}{\delta I_B}\right]_{V_{CE=\text{constant}}}$$

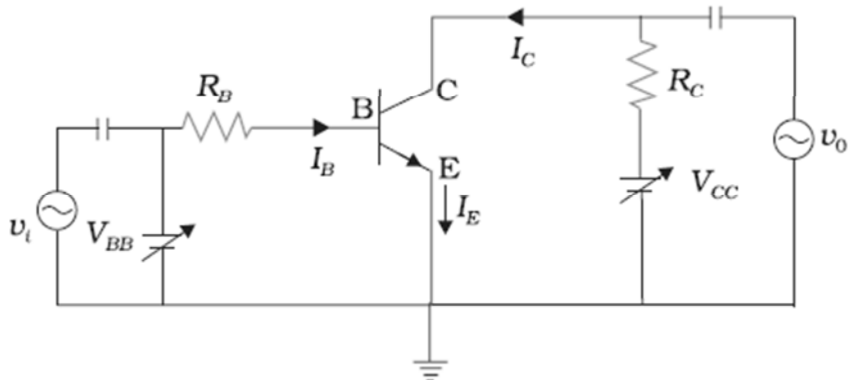
- **Relations between α and β:**

The current gains α and β are related as,

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

- **Transistor as an amplifier:**

An amplifier is a circuit which is used for increasing the voltage, current or power of alternating form. A transistor can be used as an amplifier.



- **AC Current Gain:**

AC current gain is defined as,

$$\beta_{ac} \text{ or } A_i = \left[\frac{\delta I_C}{\delta I_B}\right]_{V_{CE=\text{ constant}}}$$

- **DC Current Gain:**

DC current gain is defined as,

$$\beta_{dc} = \left[\frac{I_C}{I_B}\right]_{V_{CE=\text{ constant}}}$$

- **Voltage Gain of an Amplifier:**

It is defined as,

$$A_v = \frac{V_0}{V_i} = \frac{\text{A small change in output voltage}}{\text{A small change in input voltage}}$$

$$A_V = \frac{\delta V_{CE}}{\delta V_{BE}}$$

$$A_V = \beta_{ac}.\frac{R_{out}}{R_{in}} = A_i.A_r$$

i.e., Voltage gain = Current gain X Resistance gain

- **Power Gain of an Amplifier:**
  It is defined as,

$$A_p = \frac{\text{Output power}}{\text{Input power}} = \text{ current x voltage gain}$$

or

$$\mathrm{A}_p = A_i.A_v = \beta_{ac}^2.\frac{R_{out}}{R_{in}}$$

- **Logic Gate:**
  A logic gate is a digital circuit that has one or more inputs but only one output. It follows a logical relationship between input and output voltage.
- **Truth Table:**
  This table shows all possible input combination and the corresponding output for a logic gate.
- **Boolean Expression:**
  It is a shorthand method of describing the function of a logic gate in the form of an equation or an expression. It also relates all possible combination of the inputs of a logic gate to the corresponding outputs.
- **Positive and Negative Logic:**
  If in a system, the higher voltage level represents 1 and the lower voltage level represent 0, the system is called a positive logic. If the higher voltage represents 0 and the lower voltage level represents 1, then the system is called a negative logic.
- **OR Gate:**
  An OR gate can have any number of inputs but only one output. It gives higher output (1) if either input A or B or both are high (1), otherwise the output is low (0).

  A + B = Y

  Which is read as 'A or B equals Y'.
- **AND gate:**
  An AND gate can have any number of inputs but only one output. It gives a high output (1) if inputs A and B are both high (1), or else the output is low (0). It is described by the Boolean expression.

  A . B = Y

  Which is read as 'A and B equals Y'.
- **NOT Gate:**
  A NOT gate is the simplest gate, with one input and one output. It gives as high output (1) if the input A is low (0), and vice versa.
  Whatever the input is, the NOT gate inverts it. It is described by the Boolean expression:

$$\bar{\mathrm{A}} = Y$$

  Which is read as 'not A equal Y'.

- **NAND (NOT+AND) gate:**
  It is obtained by connecting the output of an AND gate to the input of a NOT gate. Its output is high if both inputs A and B are not high. If is described by the Boolean expression.
  $$\overline{A.B} = Y \text{ or } \overline{AB} = Y$$
  Which is read as 'A and B negated equals Y'.
- **NOR (NOT+OR) Gate:**
  It is obtained by connecting the output of an OR gate to the input of a NOT gate. Its output is high if neither input A nor input B is high. It is described by the Boolean expression.
  $$\overline{A+B} = Y$$
  Which is read as 'A and B negated equals Y'.
- **XOR or Exclusive OR gate.** The XOR gate gives a high output if either input A or B is high but not when both A and B are high or low. It can be obtained by using a combination of two NOT gates, two AND gates and one OR gate. It is described by Boolean expression:
  $$Y = \overline{AB} + \overline{AB}$$
  The XOR gate is also known as difference gate because its output is high when the inputs are different.
- **Integrated Circuits:**
  The concept of fabricating an entire circuit (consisting of many passive components like R and C and active devices like diode and transistor) on a small single block (or chip) of a semiconductor has revolutionized the electronics technology. Such a circuit is known as Integrated Circuit (IC).