# SEMICONDUCTOR DEVICES

## SEMICONDUCTOR MATERIALS

### Introduction

A very simple but somewhat superficial definition of semiconductors is that they are those materials whose resistivity lies between that of a perfect insulator and that of a good conductor. Thus glass (resistivity about $2 \times 10^{13}$ Ω.cm) is an insulator, copper (resistivity about $1.7 \times 10^{-6}$ Ω.cm) is a conductor, whereas germanium (47 Ω.cm at 27°C) and silicon ($3 \times 10^{5}$ Ω.cm at 27°C) are semiconductors.

Silicon and germanium are the best known semiconductor materials because it is from these that the most common devices, transistors and diodes, are made. These days there is a preference for silicon devices. Silicon has a lower leakage current, is less affected by temperature changes than germanium and permits easier mass production of high quality devices. Many other semiconductor materials are also in common use, for example cadmium sulphide, lead sulphide, gallium arsenide and indium antimonide, but these tend to be found mainly in specialist devices such as photo-electric cells and electroluminescent devices. This note will be confined to the properties of silicon and germanium.

### The Crystal Structure

The outer electrons of an atom are called the VALENCE electrons. Silicon and germanium are 'tetravalent', that is their atoms have four electrons in their outer orbits (tetra = 4). The silicon atom, which is simpler than that of germanium, is represented in figure 1.

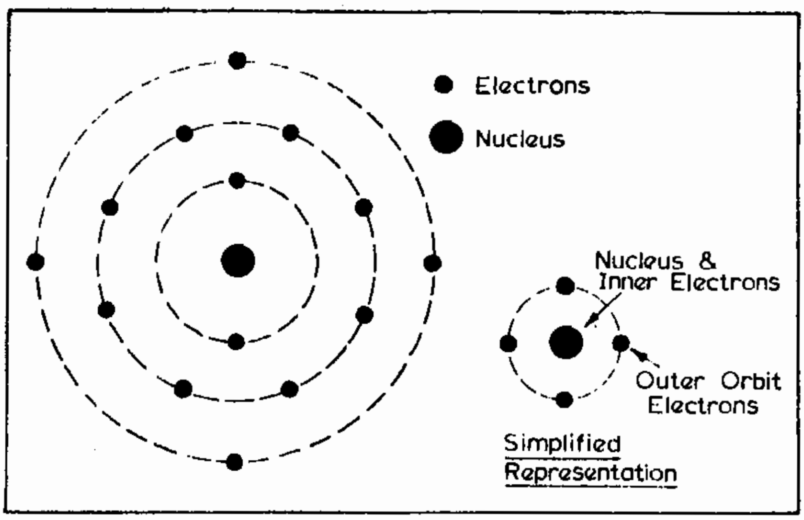


Fig. 1

Silicon Atom

SEMICONDUCTOR DEVICES

It is well known that the nucleus is positively charged and the electrons negatively charged and that the positive charge on the nucleus is equal to the total negative charge on all the electrons. The complete atom is therefore electrically neutral. Since it is the outer or valence electrons which take part in the flow of current, it is convenient to simplify the representation as shown in the diagram.

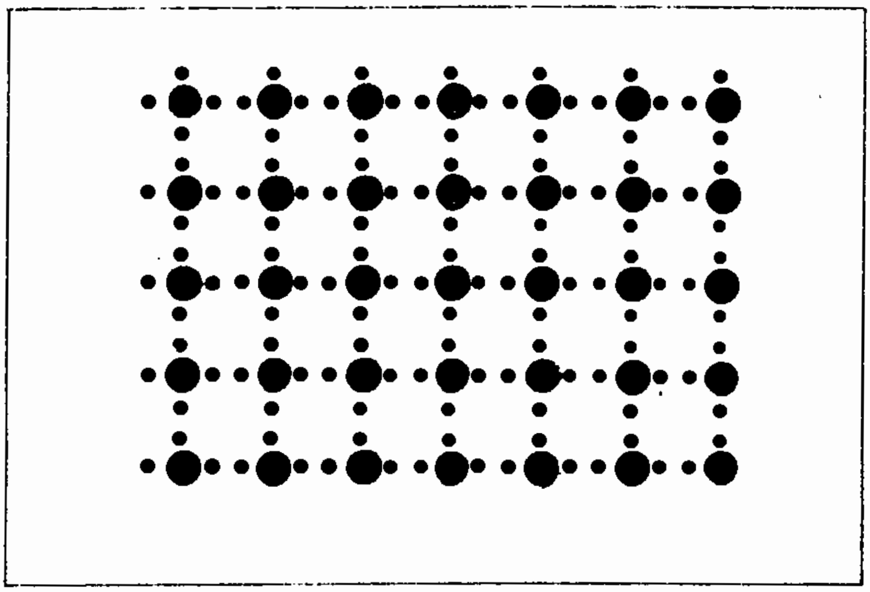

Fig. 2 : Diagramatic representation of Germanium or Silicon Crystal

When two atoms are close together there is often a tendency towards electron sharing, the electrons orbit both nuclei and this results in a bonding between the atoms that would not otherwise exist. The phenomenon is called 'covalent bonding'.

Figure 2 illustrates diagramatically what happens in tetravalent semi-conductor crystals. Each of the four valence electrons of any one atom is shared with four neighbouring nuclei so that each nucleus is in effect orbited by eight electrons, causing strong bonds to exist between neighbouring atoms in the crystal. As it is very difficult for an electron to break away from its bonds to participate in current flow such crystals tend to be rather poor conductors. In practice pure silicon and germanium are perfect insulators only at a temperature of absolute zero. At room temperatures, owing to thermal agitation, the occasional electron is able to break away from the crystal lattice to carry a small current when a battery is connected across the crystal. The resistivity of germanium ($\rho = 47\ \Omega.cm$ at $27^{\circ}C$) is far less than that of silicon ($\rho = 3 \times 10^{5}\ \Omega.cm$ at $27^{\circ}C$).

P- and N-Type Semiconductors

In order to obtain the controlled flow of current required in transistors, diodes and other devices, minute traces of certain impurity elements are added to the INTRINSIC (i.e. pure) semiconductor crystal. For instance, if impurity atoms of a pentavalent element such as antimony or arsenic are added, figure 3, only four of the impurity's five outer electrons can 'fit' into the lattice and the surplus fifth electron becomes free to act as a current carrier. There will of course be a free electron for each atom of impurity added. A crystal treated in this way is called N-TYPE semiconductor material because it contains free electrons which are negative charge carriers. However, it should be noted that the overall charge on the crystal remains zero because each of the individual atoms present is electrically neutral. The overall effect is a large increase in the conductivity. The resistivity is reduced to only a few Ω.cm by the addition of about one impurity atom to every $10^8$ atoms of pure semiconductor. The pentavalent impurity atoms are called DONOR atoms since each donates one free electron to the lattice.

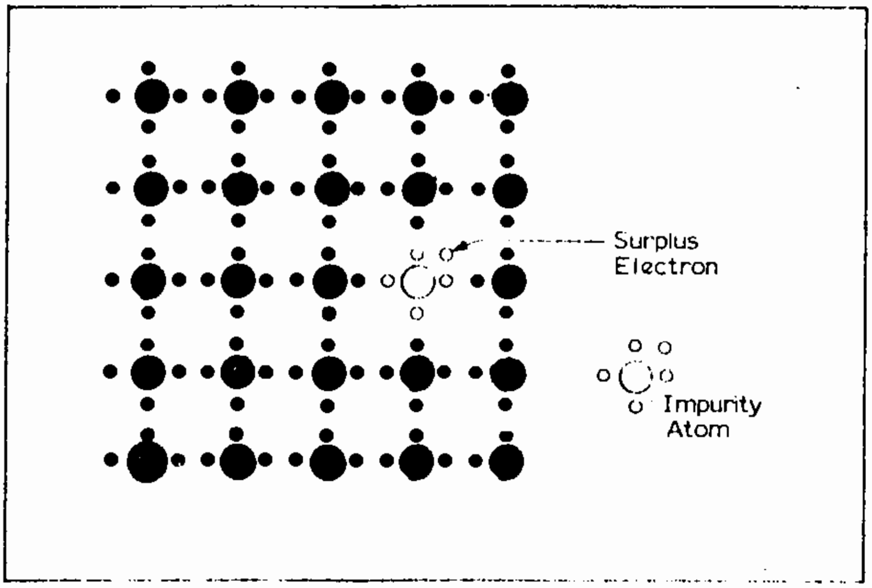


Fig. 3 : N-Type Semiconductor Material

If on the other hand pure semiconductor crystal is contaminated or DOPED with atoms of a trivalent element such as aluminium or indium, figure 4, a deficiency of electrons is introduced into the lattice. For each impurity atom there is a deficiency of one electron. These deficiencies are called HOLES. Because a hole exerts an attractive force on neighbouring electrons in the lattice, it constitutes a virtual positive charge and may be considered as a positively charged particle. A hole need not stay near the impurity atom which introduced it. An electron from a neighbouring atom can move it and cancel it so that the hole moves to the neighbouring atom. A semiconductor crystal which contains mainly holes is called P-TYPE (p for positively charged particles). Whilst the presence of holes increases the conductivity, p-type material is electrically neutral. The impurity atoms are called ACCEPTORS because each atom can accept one electron from the lattice.

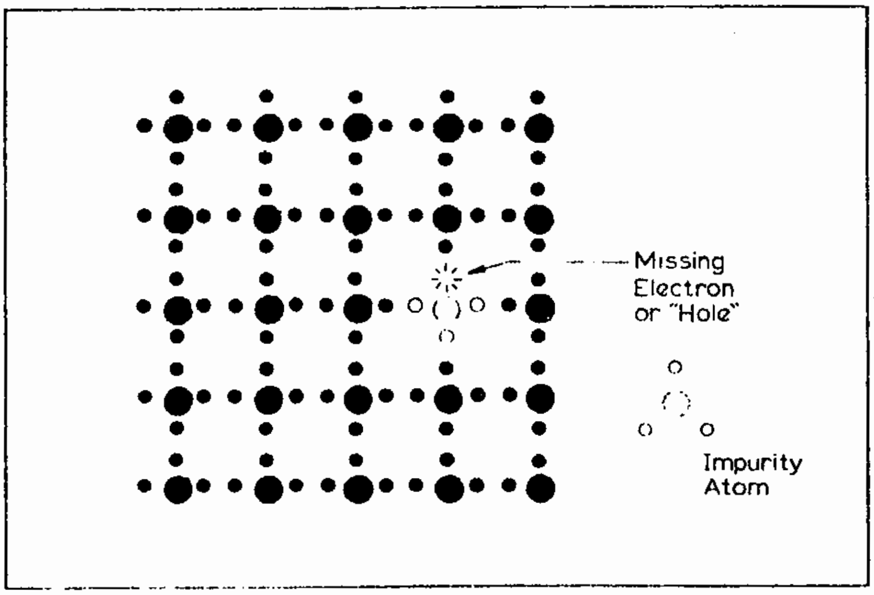


Fig. 4 : P-Type Semiconductor Material

In future illustrations we shall indicate only the free electrons and holes, as shown in figure 5. The donor and acceptor atoms are omitted for the sake of clarity. In practice, at room temperatures, semiconductor materials contain a few holes and electrons generated by thermal agitation. In fact thermal generations and recombinations are taking place continuously. Thus p-type material will at all times contain a few free electrons and n-type material will contain a few holes. However most of the carriers in the n- (or p-) type material are the electrons (or holes) introduced deliberately, these are called MAJORITY CARRIERS. The thermally generated holes (or electrons) are called MINORITY CARRIERS.

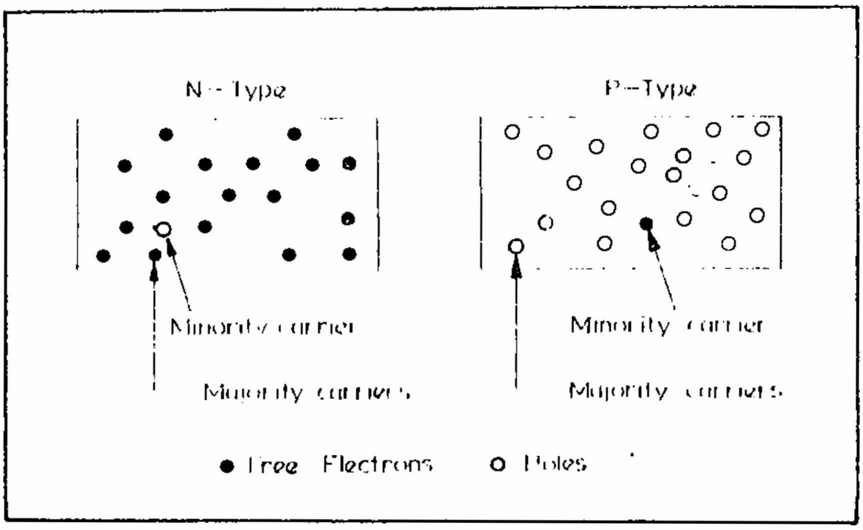


Fig. 5 : Diagramatic representation of P and N Materials

Figure 6 illustrates what happens when a battery is connected to n- and p-type semiconductors. In the n-type, figure 6(a), electrons are attracted to the positive terminal of the battery (because unlike charges attract). The loss of electrons from the left-hand end of the crystal would cause the material to become positively charged. However, this positive charge is cancelled by electrons flowing into the right-hand end from the negative terminal of the battery. This action repeats continuously giving rise to a flow of electrons (i.e. current) through the crystal but the net number of electrons within the crystal at any given instant remains constant.

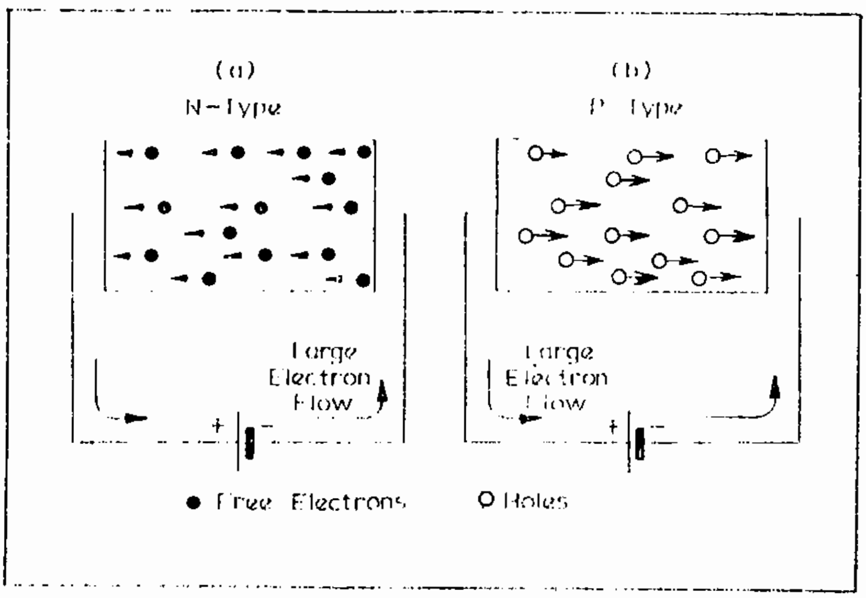


Fig. 6 : Current flow in P- and N- Type Semiconductors

SEMICONDUCTOR DEVICES

In the p-type, Fig. 6(b), holes reaching the right-hand side are neutralised by a flow of electrons from the negative terminal of the battery. This gain of electrons would cause the crystal to become negatively charged. However, equilibrium is maintained by the necessary number of electrons being ejected from the left-hand end to the positive terminal of the battery. A continuous flow of electrons from the battery results but the number of holes within the crystal at any instant remains constant.

THE SEMICONDUCTOR DIODE

Introduction

Basically the semiconductor diode consists of a piece of n-type and a piece of p-type semiconductor joined together (Fig. 7). The two electrodes are the anode (p-type) and the cathode (n-type).

Naturally, as soon as the junction is formed there will be a drift of carriers across it. Holes move into the n-type and electrons move into the p-type. At first sight it might be expected that all the electrons in the n-type would move into the p-type and that all the holes would move into the n-type resulting in a total disappearance of all carriers. However, once a few electrons have entered the p-type they make it negatively charged (both regions were initially neutral). Similarly the gain of holes by the n-type makes it positively charged (Fig. 8). Since like charges repel, the negative charge opposes the further flow of electrons and the positive charge opposes the further flow of holes. The situation then is just as if there were a small battery across the junction. For simplicity, in Fig. 8, the region near the junction is shown devoid of carriers. This carrier free area is called the BARRIER or DEPLETION region.

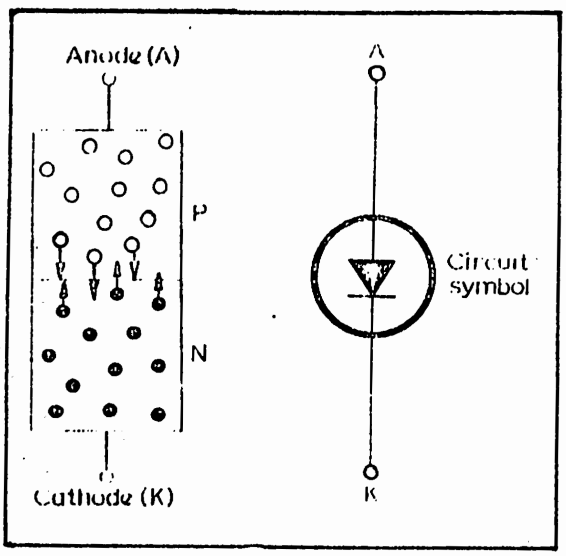


Fig. 7 : P-N Junction Diode

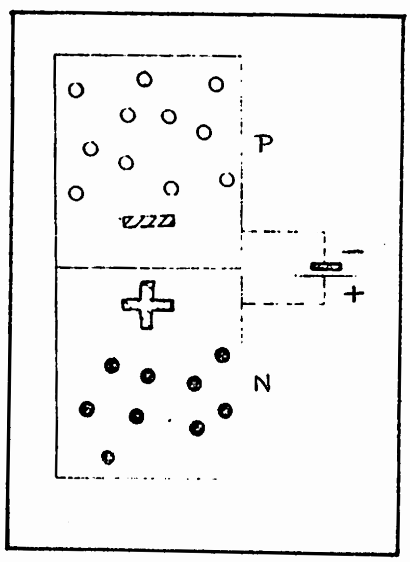


Fig. 8 : Virtual Battery across a P-N Junction

## Reverse Bias

If a real battery is now connected the same way round as the virtual battery, Fig. 9, the barrier region is broadened and no majorities can cross the junction. Only the occasional minority carriers generated thermally cross the junction and a very small flow of current, called the LEAKAGE CURRENT results. The diode is now said to be REVERSE BIASED. The leakage current in germanium diodes is far greater than that in silicon diodes.

## Forward Bias

If on the other hand the real battery which is larger than the virtual battery is connected in opposition to it, Fig. 10, a considerable flow of current results. Holes attracted by the negative terminal of the battery drift to the right towards the junction; on reaching it they are cancelled by electrons from the n-type; this loss of electrons from the n-type is compensated for by more electrons being simultaneously injected by the negative terminal of the battery. As holes from the p-type are cancelled at the junction, electrons are simultaneously released from its left-hand to the positive terminal of the battery. The small leakage current due to minorities adds to the main majority flow. A large current flows and the diode is now said to be FORWARD BIASED.

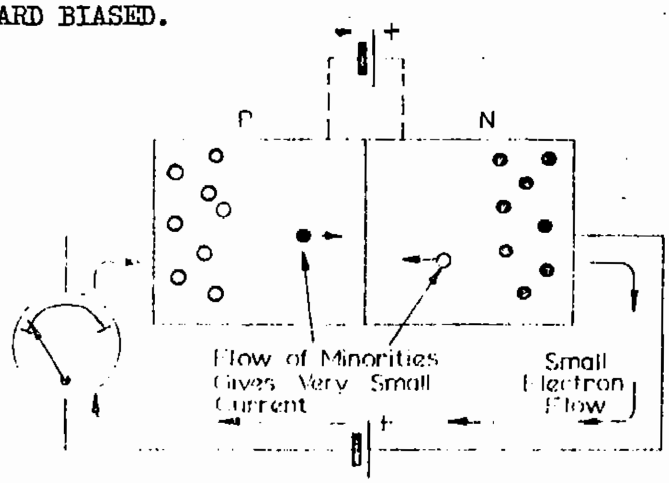


Fig. 9 : Reverse Bias

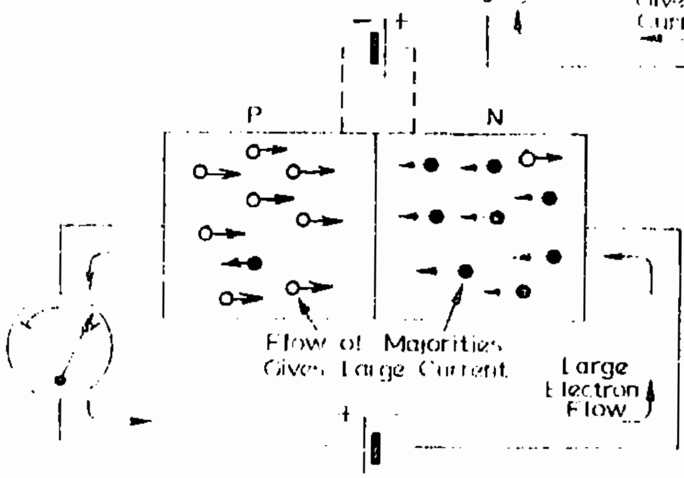


Fig. 10 : Forward Bias

SEMICONDUCTOR DEVICES

The Diode Characteristic

Figure 11 illustrates the shape of the characteristic. The virtual battery across the junction is small and so only a fraction of a volt need be applied in the forward direction to start a current. This 'turn-on' voltage is of the order of 200 mV for germanium diodes and 600 mV for silicon diodes. Once this voltage has been reached an extremely small change in voltage results in large changes in current. This situation is illustrated by the fact that the curves become very steep.

The forward current could quickly become large enough to damage the device and hence, in the test circuit, a limiting resistor R is added.

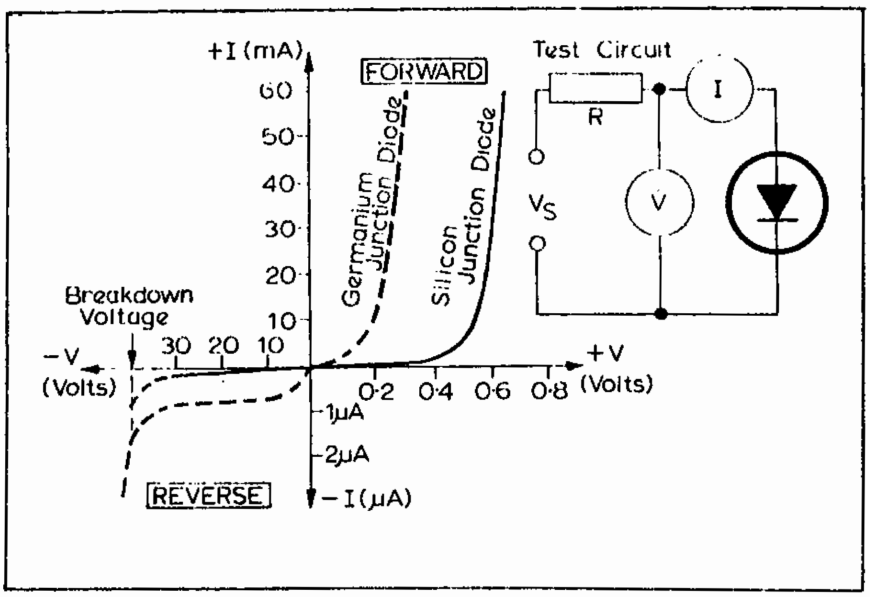


Fig. 11 : Diode Characteristic

In the reverse direction, however, so little current flows that it is necessary to inflate the 'I' scale in order to see the current at all. The fact that the leakage current for silicon diodes is much less than for germanium diodes is clearly illustrated by the reverse characteristic. There comes a point, however, when 'breakdown' occurs and a considerable current flows. At low voltages breakdown is caused by electrons breaking away from their covalent bonds (the ZENER effect). At higher voltages it is caused by the minority electrons gaining sufficient velocity to dislodge other electrons from their atoms. These new electrons are then themselves accelerated and, in turn, can produce more electrons. There is therefore a sudden build-up of current and this is known as the AVALANCHE effect.

These breakdown effects are exploited in VOLTAGE REGULATOR DIODES (sometimes called zener diodes). These diodes make use of the fact that once the breakdown voltage has been reached the characteristic is almost parallel to the 'I' axis and thus they can be used in voltage regulated power supplies. The voltage at which breakdown occurs is determined by impurity levels and other physical factors.

## Diode Applications

Like the thermionic diode, the semiconductor device is commonly used for rectification, modulation and demodulation. It is also used a great deal in logic circuits. In the rectifier circuit, Fig. 12(a), the diode passes current only during positive half cycles of the input so that the load voltage is a half-wave rectified signal. This can be smoothed to obtain a steady direct voltage. Other possibilities are full-wave rectifier and voltage doubler circuits.

The function of the demodulator, Fig. 12(b), is to abstract the wanted audio frequency signal content from the unwanted radio frequency carrier. The diode permits only the positive going half-cycles of the carrier to reach the load. The average value of the load voltage is the wanted audio frequency signal, the unwanted radio frequency being bypassed through the capacitor whose reactance is chosen to be low at the radio frequency.

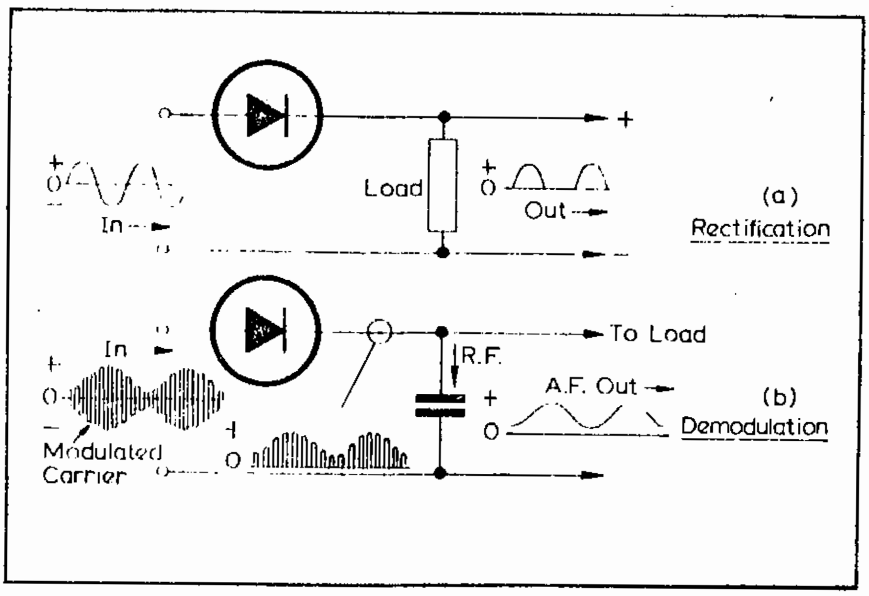


Fig. 12 : Diode Applications

## SEMICONDUCTOR DEVICES

## THE TRANSISTOR

### Introduction

The conventional transistor is a two-junction, three-layer semiconductor device capable of current, voltage and power amplification. Clearly there are two possible configurations: n-p-n and p-n-p, Fig. 13. There are three electrodes: the BASE (b), EMITTER (e) and COLLECTOR (c). Note that the arrow in the symbol for the p-n-p device points towards the base whereas in the n-p-n transistor it points away from it. The arrow indicates the direction in which 'conventional' current would normally flow in the device, electron flow is in the opposite direction to the arrows.

### Transistor Action

Figures 14 and 15 illustrate current amplification in an n-p-n silicon transistor. The base-emitter junction forms a diode which is forward biased by the 600 mV supply (remember that for forward bias the p-type is connected to the positive terminal of the supply. On the other hand the collector is 5.4 volts (i.e. 6-0.6) positive with respect to the base (taking the emitter as datum = 0 volts and ignoring the meter resistances). Hence the base-collector junction is reverse biased.

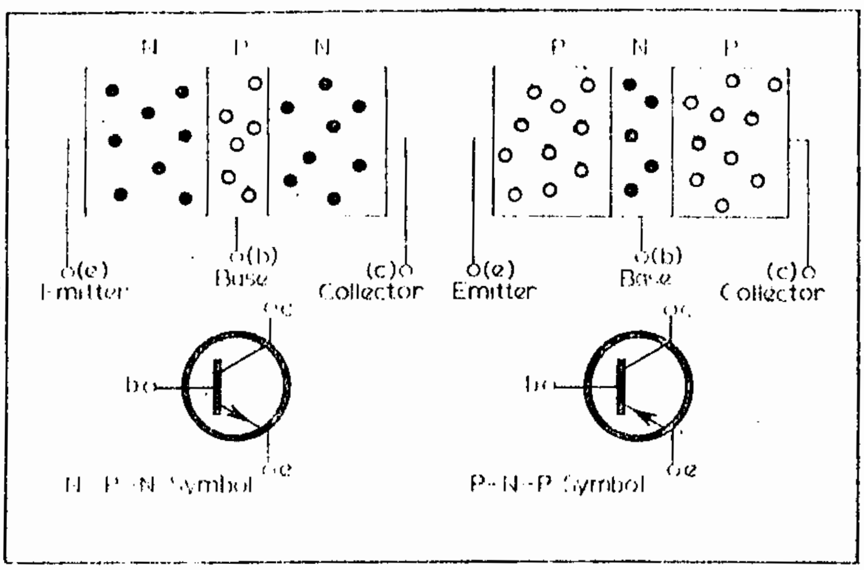


Fig. 13 : N-P-N and P-N-P Transistors

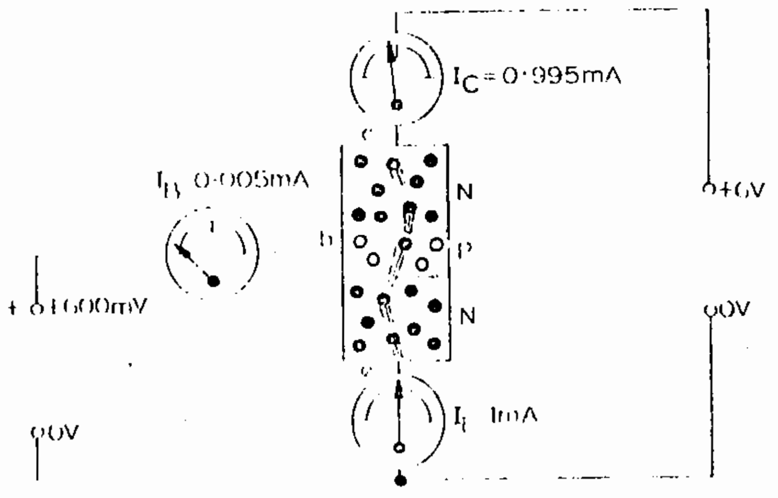


Fig. 14
N-P-N Transistor
Action (i)

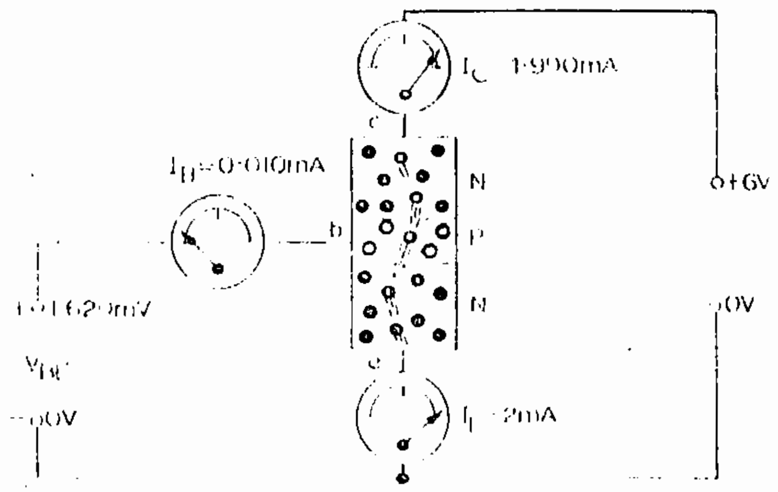


Fig. 15
N-P-N Transistor
Action (ii)

A large number of electrons enters the emitter n-region under the influence of the forward biased base-emitter junction. The base region is designed to be very thin and the attractive force due to the positive collector potential is very strong, and so most of the electrons from the emitter pass straight through the base into the collector. Of course the occasional electron meets a hole in the base and combines with it and this gives rise to a small base current. For a typical silicon transistor with an emitter current ($I_E$) of 1 mA, the collector current ($I_C$) might be 0.995 mA and the base current ($I_B$) only 0.005 mA (1 - 0.995).

Now suppose the base voltage is increased sufficiently to cause the emitter current to be doubled to 2 mA, Fig. 15. Referring back to the diode characteristics in Fig. 11, it will be remembered that a large increase in forward current can be achieved by just a few millivolts change in forward voltage. Thus, in a typical case, an increase in base voltage from about 600 mV to 620 mV would produce the change in emitter current from 1 mA to 2 mA. The collector current doubles to 1.990 mA and the base current to 0.010 mA since now there are twice as many electrons (per second) entering the emitter, twice as many passing through to the collector and twice as many electron-hole combinations in the base.

## Current Amplification

Comparing Figures 14 and 15 it can be seen that a change of 0.005 mA in $I_B$ caused a change of 0.995 mA in $I_C$, a current gain or amplification of :-

$$\frac{0.995 \text{ (change in } I_C)}{0.005 \text{ (change in } I_B)} \simeq 200$$

In practice approximately the same figure is obtained by taking the ratio of static values of $I_B$ and $I_C$. The ratio is defined by the British Standards Institution as the STATIC VALUE OF THE SHORT-CIRCUIT FORWARD CURRENT TRANSFER RATIO (symbol $h_{FE}$). This is the ratio between the continuous output current, in our case $I_C$, and the continuous input current, in our case $I_B$, the output voltage being held constant. Thus when $I_B$ is 0.010 mA and $I_C$ is 1.990 mA the ratio is $\frac{1.990}{0.010} \simeq 200$.

In published data the ratio is always quoted for a specified value of $I_C$, 1 mA for example.

## Voltage and Power Amplification

In order to obtain a voltage output a load resistor ($R_L$) must be added. The base-collector junction is reverse biased and therefore its resistance is of the order tens of thousands of ohms (in practice it varies with $I_C$). A load resistor of the order hundreds of ohms can therefore be added in series with the collector without changing the values of $I_C$ very much. In Fig. 16 then, a 1 kΩ load has been added. The voltage developed across this is the output ($V_{OUT}$). We have seen that in a typical case when the base-emitter voltage ($V_{BE}$) is changed by 20 mV, $I_C$ changes by 0.995 mA and so, assuming that the presence of $R_L$ does not significantly alter $I_C$, $V_{OUT}$ changes by : 0.995 mA x 1 kΩ = 1 volt. The change in $V_{OUT}$ is therefore about 50 times ($= \frac{1}{0.020}$) the change in $V_{BE}$; a voltage gain of 50. In Fig. 16 the initial conditions are indicated above the final conditions.

The change in input power is (620 x 0.010 - 600 x 0.005) µW = 3 µW. The corresponding change in output power is approx. (4 x 2 - 5 x 1) = 3 mW. The power gain is about 1000.

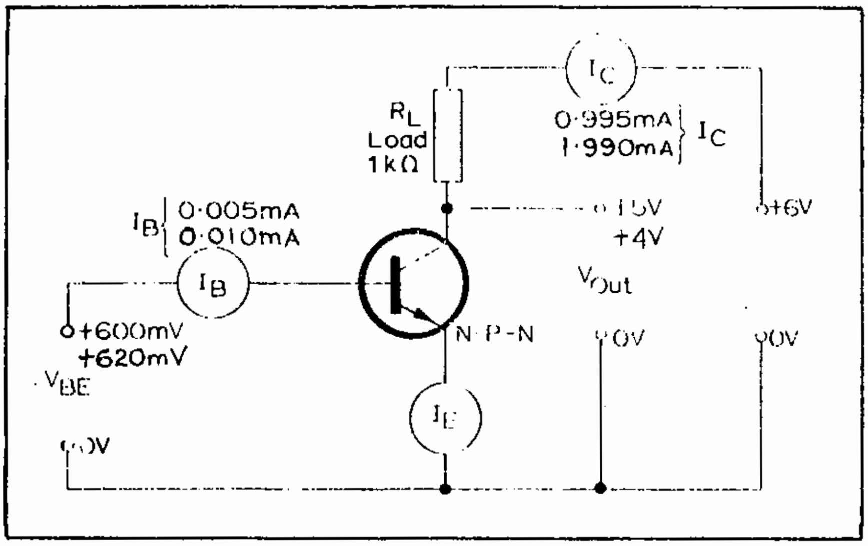


Fig. 16 : Voltage Amplification

## A.C. Amplification

Figure 17 is the basic circuit for an a.c. amplifier. The a.c. input signal of 20 mV peak to peak is applied, through d.c. blocking capacitor C1, to the base. The input signal is superimposed on the 610 mV d.c. base bias - for a germanium transistor this is much lower, see Figure 11. Assuming that the same current transfer ratio appears as in Figure 16, the a.c. output waveform has a peak-to-peak value of 1 volt. In practice the a.c. gain, or to be precise the SMALL-SIGNAL SHORT-CIRCUIT FORWARD CURRENT TRANSFER RATIO ($h_{fe}$), is slightly larger than the static ratio. It should be noted that the output signal is in antiphase with the input.

The principle of operation of p-n-p transistors is similar to that for n-p-n devices and can be described in terms of hole flow from emitter to collector (rather than electron flow as in the case of n-p-n). It is also important that the polarity of the base-emitter and collector-emitter supplies are reversed.

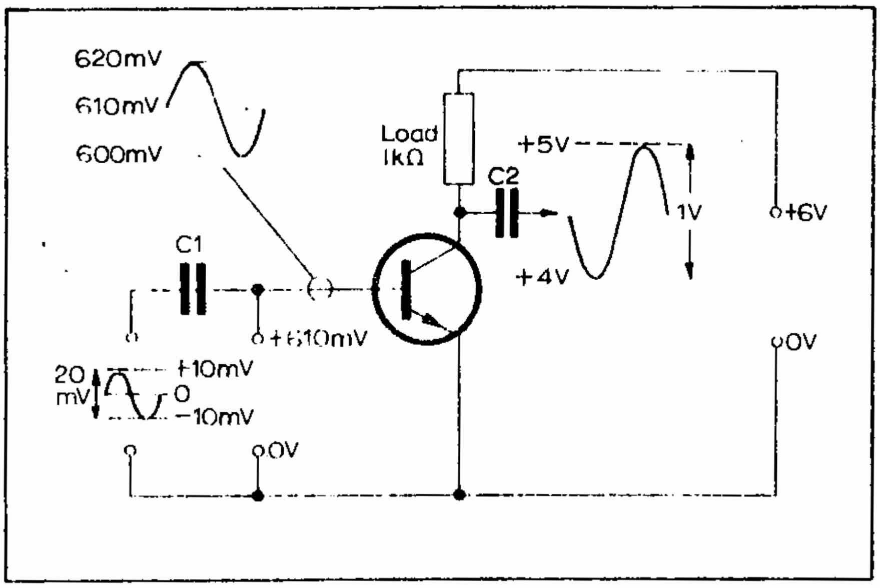


Fig. 17 : A.C. Amplification

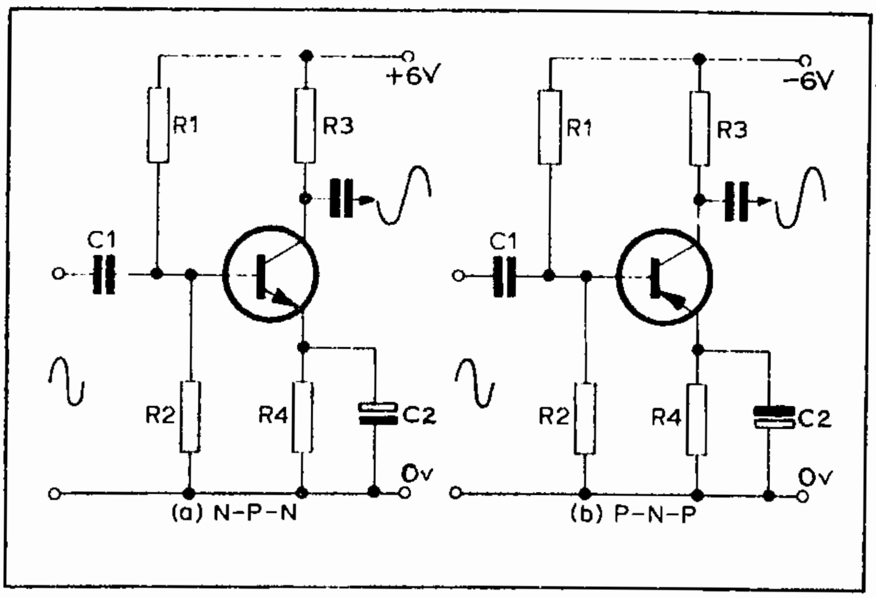


Fig. 18 : Practical Amplifier Circuits

## Practical Amplifier Circuits

Figure 18 illustrates typical practical amplifier circuits for n-p-n and p-n-p transistors. The standing base current is provided by potential divider resistors R1 and R2, so that the circuit may be powered by a single battery. Resistor R4 is included to minimise the effects of temperature changes on collector current. Due to minority carriers there is always a small temperature dependent leakage current flowing between emitter and collector even when the base is open circuited. The leakage current causes a small voltage drop across R4 which raises the emitter potential so reducing the emitter-base bias. This reduction in bias reduces $I_C$ and largely compensates for the effects caused by the increase in temperature. R4 also provides a degree of negative feedback for wanted a.c. signals and, if such feedback is not required, bypass capacitor C2 is included in the circuit. Thermal effects are more of a problem in germanium than in silicon devices. With silicon devices the inclusion of R4 is often unnecessary. At higher collector currents, if precautions are not taken, the heat generated can cause many minority carriers to be liberated. These carriers add to the collector current generating still more heat. If this cumulative build-up, called 'thermal runaway', is allowed to continue it can result in destruction of the device. Therefore it is sometimes necessary to mount the transistor a carefully designed heatsink.

## Circuit Configurations

The amplifier circuits so far considered have been of the 'common emitter' type, so called because, as far as a.c. is concerned, the emitter is common to both input and output terminals. Whilst the common or 'grounded' emitter circuit is by far the most frequently used, there are applications for which the 'common base' and 'common collector' configurations are useful. Figure 19 illustrates the three configurations. An important consideration when choosing a circuit is the input and output impedance requirements. Some important features of the three circuits are illustrated in the table on page 16.

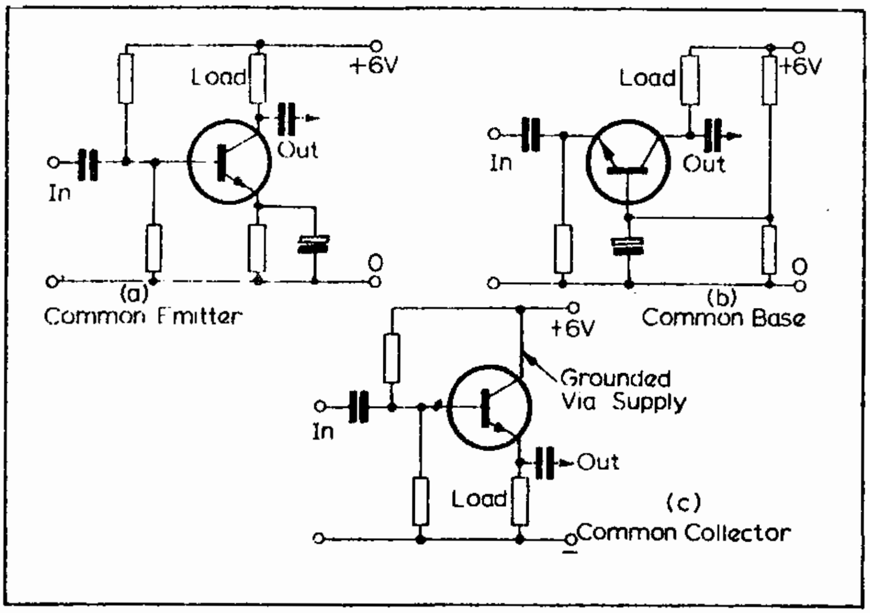


Fig. 19 : Basic Circuit Configurations

SEMICONDUCTOR DEVICES

| | Common Base | Common Emitter | Common Collector |
|---|---|---|---|
| Current gain | about 1 | High | High |
| Voltage gain | High | High | about 1 |
| Input impedance | Low | Medium | High |
| Output impedance | High | Medium | Low |
| Power gain | Medium | High | Medium |
| $180^{o}$ phase inversion (output w.r.t. input) | No | Yes | No |
| Static s/c forward current transfer ratio symbol | $h_{FB}$ | $h_{FE}$ | $h_{FC}$ |
| Small signal s/c forward current transfer ratio symbol | $h_{fb}$ | $h_{fe}$ | $h_{fc}$ |

Diagrams and text extracted from the Mullard Mini Book "Semioonductor Devices", now out of print.

J. Fraser/MAG
Retyped 16th September 1977