THE BRITISH BROADCASTING CORPORATION ENGINEERING TRAINING DEPARTMENT

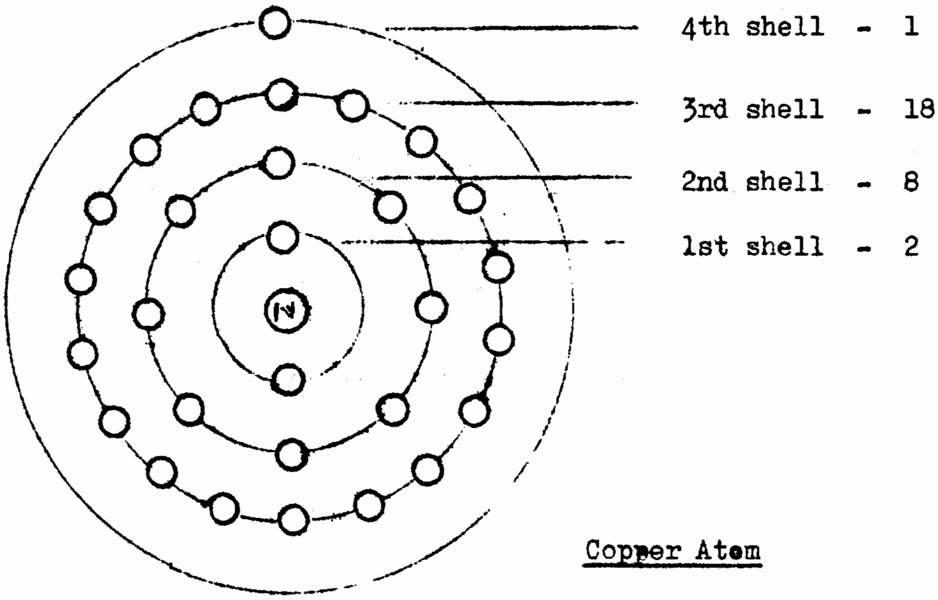
INFORMATION SHEET

SEMICONDUCTORS

All matter is made up of atoms. An atom consists of a nucleus and its orbitting electrons. Each electron has a negative charge while each nucleus is positively charged. The positive charge on the nucleus is equal to the total negative charge on all its orbitting electrons.

The electrons are grouped into several orbits or shells. The shell nearest to the nucleus has a maximum of two electrons in it. The next shell has a maximum of eight and the next has a maximum of 18.

As an example, copper has 29 electrons. These are in four shells with two in the first, eight in the second, 18 in the third and only one in the fourth.



Copper Atom

The electrons in the outer shell are called valence electrons. It is mainly these electrons which affect the conductivity of an element.

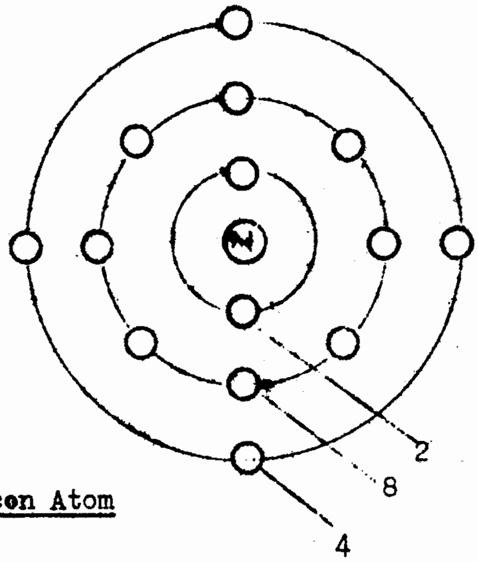
Materials can be split into three groups, conductors, insulators and semiconductors.

If a valence electron has enough energy it can break away from its orbit and move into the orbit of another atom. This energy can be given to an electron by the action of heat or light. In a conductor the valence electrons can be easily moved from one atom to another. This movement takes place in a random manner unless an e.m.f. is applied across the material. Under the influence of the e.m.f. the electrons move towards the positive terminal and are replaced by electrons entering at the negative terminal. This constitutes a flow of current.

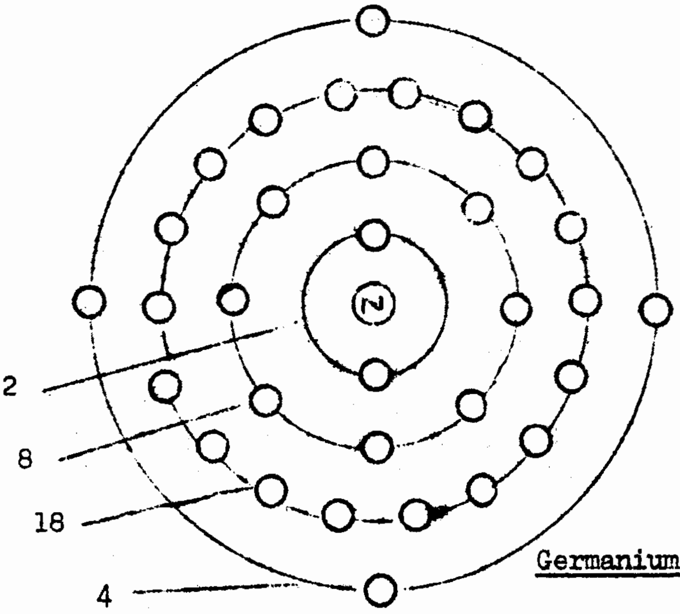
In a good insulator the valence electrons are firmly bound to their parent nuclei. Thus there are no free electrons to form a current if a potential is applied across the material.

Semiconductors are materials which are neither good insulators nor good conductors. Silicon and germanium are semiconductors. They both have four valence electrons.

Silicon has fourteen electrons in three orbits and germanium has 32 in four orbits.

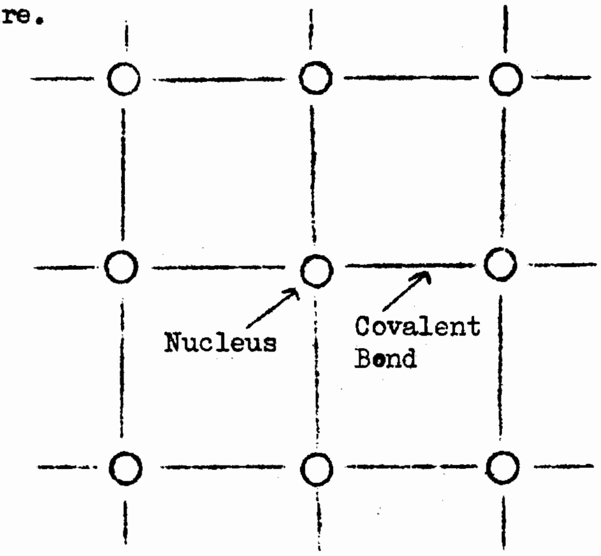


Silicon Atom



Germanium Atom

Adjacent atoms of semiconductor material link together by sharing valence electrons. These links are called covalent bonds. Each atom of silicon (or germanium) links with four other atoms. This forms a three dimensional crystal lattice structure. For convenience we shall use a two dimensional diagram to show the lattic structure.



Simplified diagram of the crystal lattice structure of silicon or germanium

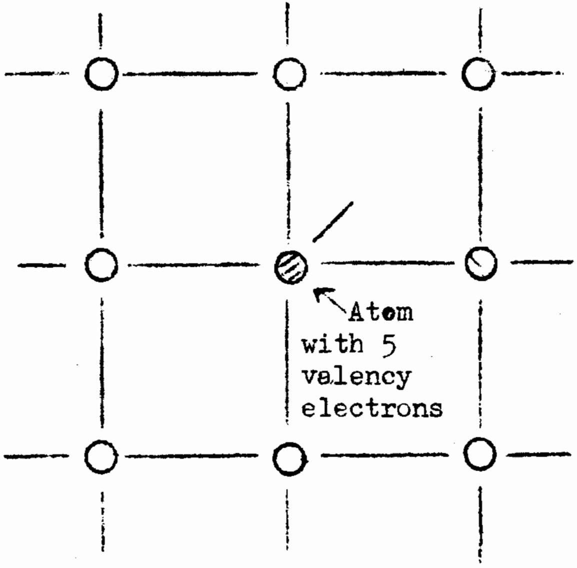
At very low temperatures the valence electrons have little energy and cannot escape from the covalent bonds. Thus a pure semiconductor will behave as a good insulator at very low temperatures.

If the temperature rises some of the valence electrons will have energy to break the bonds. If there is an e.m.f. across the material these electrons form a current. This current is called a thermal current. The number of electrons released from their covalent bonds increases with temperature thus the thermal current also rises.

Less energy is required to release a valence electrons from germanium than from silicon.

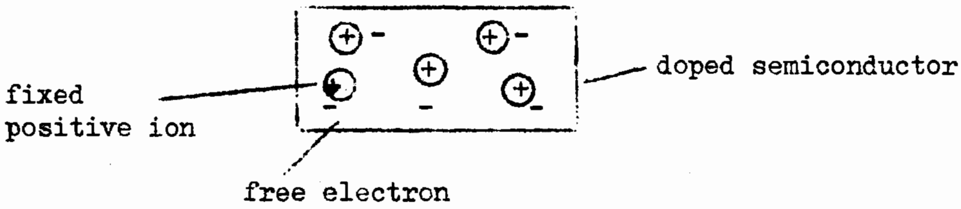
Doping

If a small amount of an element with 5 valence electrons is introduced into a semiconductor lattice structure an irregularity is produced. This process is called <u>doping</u>. The atoms of this element fit into the crystal lattice structure and set up covalent bonds with four adjacent atoms. This leaves one electron which is not linked.



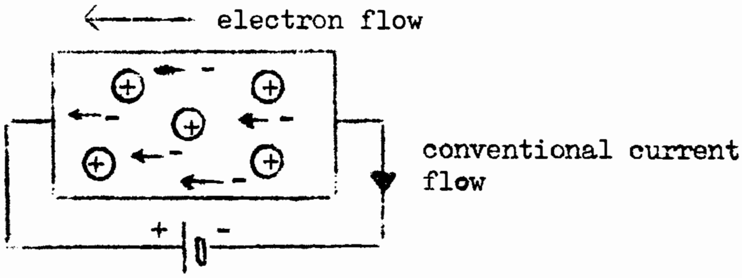
The unlinked electron is easily detached from its parent atom. This atom will then be one electron short and will have a positive charge. It is then called a positive <u>ion</u>. There will be as many positive ions as there are free electrons.

We shall show the free electrons and ions as follows:-



The positive ions are locked into the structure and cannot move. When an e.m.f. is applied across the material conduction takes place by the movement of the free electrons.

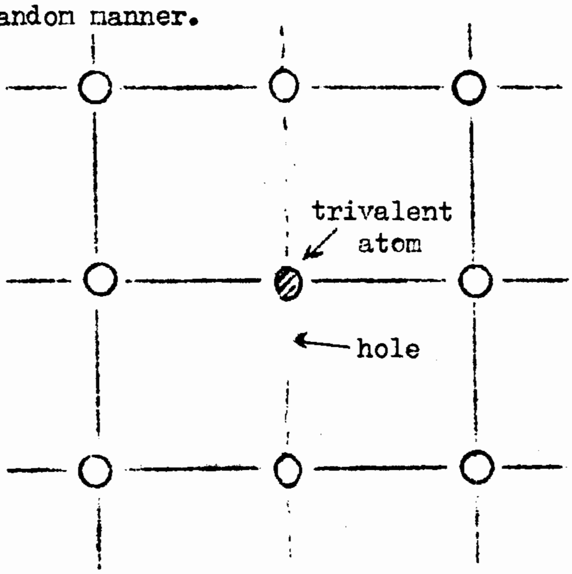
The electron flow is towards the positive terminal of the supply. This is the opposite direction to 'conventional current' flow.



An element having 5 valence electrons is called a pentavalent element. Semiconductor material doped with a pentavalent element is called N type. This is because the current flow is due to the Negatively charged electrons.

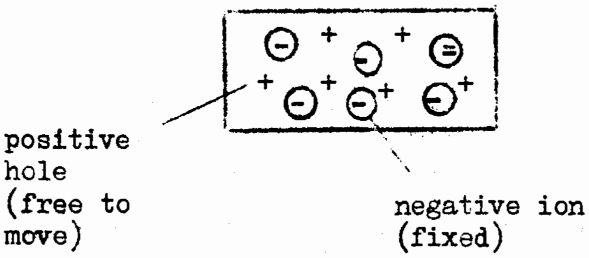
The semiconductor material may also be doped with an element having three valence electrons. A doping element of this type is called a trivalent element.

Each atom of the trivalent element forms bonds with three atoms of the semiconductor leaving a hole in the lattice structure where the fourth bond is missing. This weakness in the lattice structure encourages an electron from a nearby atom to leave its own atom and orbit the trivalent atom. The hole has therefore moved from the trivalent atom to a semiconductor atom. Another electron can now move in to fill this hole and so the hole moves about the material in a random manner.

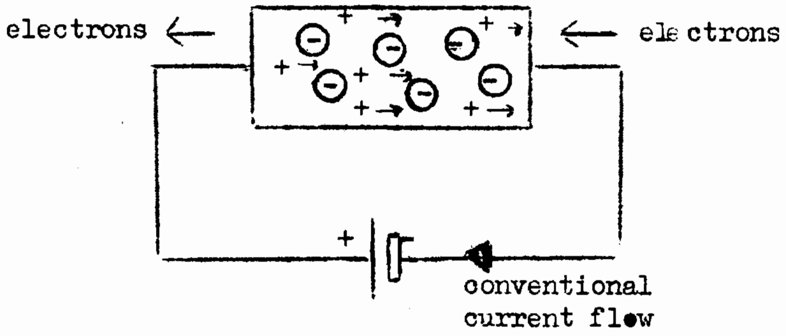


When it receives a fourth valence electron the atom of the trivalent element will have a negative charge. It is then a negative ion. The negative ions are fixed in the lattice structure and are unable to move.

We shall show the holes and negative ions in the following way:-



If an e.m.f. is applied across a piece of semiconductor doped with a trivalent element a current flows. The random movement of holes about the material is directed by the e.m.f, to the negative terminal where they are neutralised by electrons from the supply. At the positive terminal electrons leave the material producing holes which move across to the negative terminal.



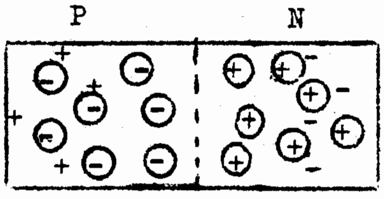
It is convenient to consider the current as being made up of positively charged holes. Semiconductor material doped with a trivalent element is then called P type.

The doping is usually about one atom of the doping element in 10^7 atoms of pure semiconductor.

Junction Diode

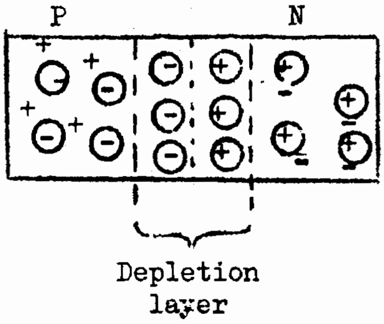
A piece of semiconductor can be doped such that one half is N type and the other is P type. The region where the two meet is called the P-N junction. The holes and electrons in the junction drift across and meet each other. As a hole is an atom which has lost an electron, these holes and electrons combine and neutralise each other.

The combinations in the junction leave a deficit of holes and electrons. There is therefore a surplus of positively charged ions along the N side of the junction and negatively charged ions along the P side.

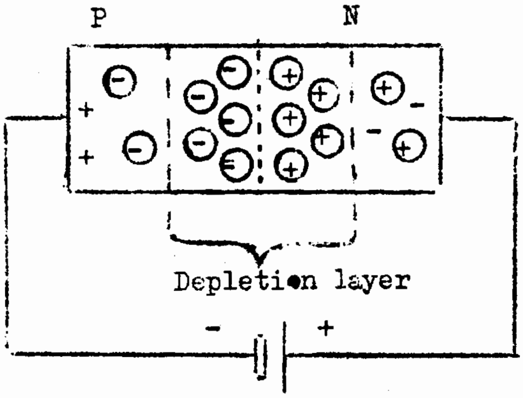


A P-N Junction

The region in the junction where there are no free holes or electrons is called the depletion layer. The ions in this layer produce a potential barrier which stops any more holes or electrons from entering.

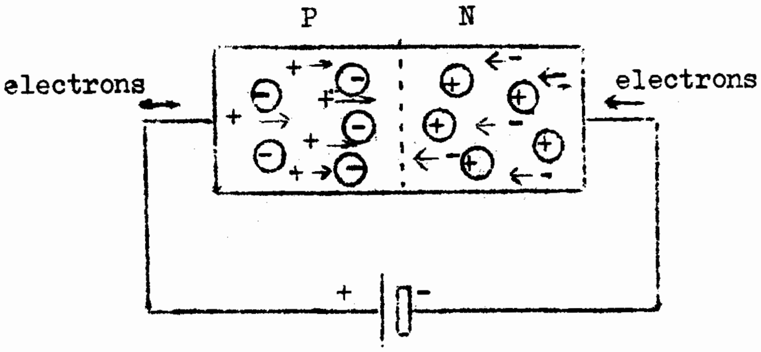


Let us now apply an e.m.f. across the junction such that the negative terminal is connected to the P type and the positive is connected to the N type. This will attract more holes and electrons away from the junction increasing the potential barrier and widening the depletion layer.



As the holes and electrons cannot get into the junction no current can flow. The junction is then said to be reverse biased.

If we now reverse the e.m.f. we will forward bias the junction. In this condition the electrons lost in combination with holes in the junction are replaced from the negative terminal of the supply. Holes are formed at the positive terminal by electrons leaving and replace the holes which are lost in combination at the junction. We shall thus have a current of holes and electrons entering the junction and combining. A current is therefore drawn from the supply and the device is therefore a conductor.



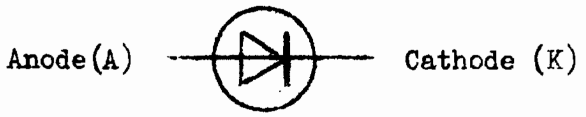
So far in our explanation of the P-N junction we have ignored the thermal current. This current will flow when an e.m.f. is applied whether the semiconductor material is doped or not. As this current is very small its effects will only be of importance when the junction is reverse biased.

At room temperatures the thermal current in germanium is about 4µA. It doubles for approximately every 9°C rise in temperature. For silicon the thermal current is about a thousand times less than for germanium.

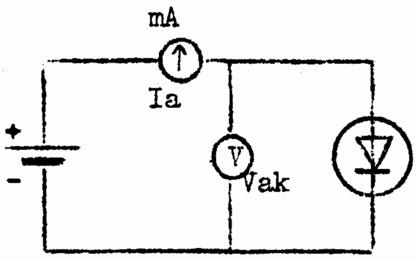
When dealing with the P-N junction we must consider two currents. The 'wanted' current is made up of the holes and electrons released by the doping element. The thermal current made up of electrons released by thermal energy is an 'unwanted' current.

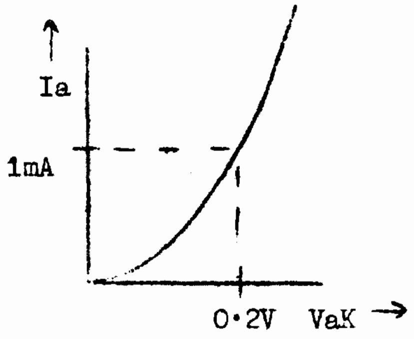
The P-N junction which we have been considering is a junction diode. It conducts when the P layer is connected to the positive supply terminal. The P layer is therefore the anode and the N layer is the cathode.

The BS symbol for a semiconductor diode is:-



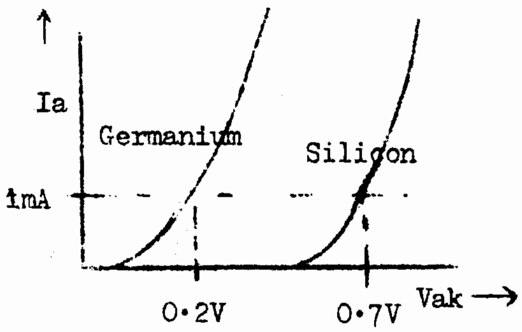
Let us now look at the characteristics of a germanium diode. If we plot the anode current against the anode-cathode voltage for the forward biased conditions we shall get:-





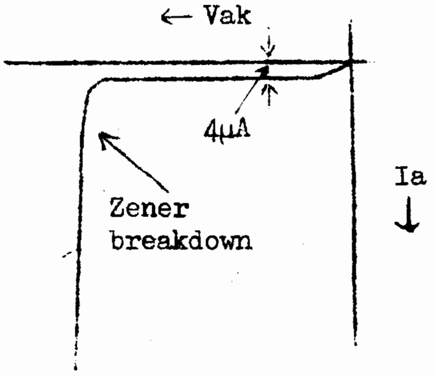
The curve is approximately exponential in shape.

If we now plot the forward bias characteristics of a silicon diode we shall find that the curve is similar to that of a germanium diode but requires a greater voltage for the same current.



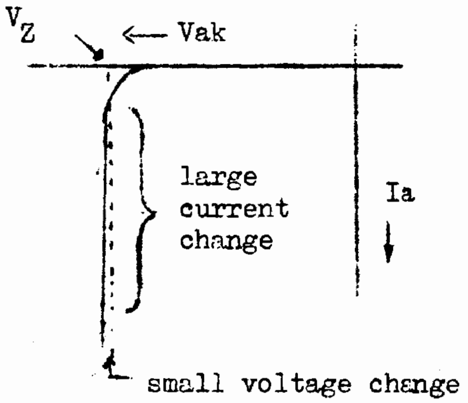
When a germanium diode is reverse biased a small thermal current of a few µA flows. If the reverse bias is increased too far the lattice structure becomes strained and valence electrons can break away from their bonds. This is called Zener breakdown.

When Zener breakdown takes place the junction will be destroyed unless the current is limited by external resistance.



When a diode is used as a rectifier it is important that Zener breakdown never occurs. Rectifier diodes can be manufactured to withstand reverse voltages of up to about 1000V without breaking down.

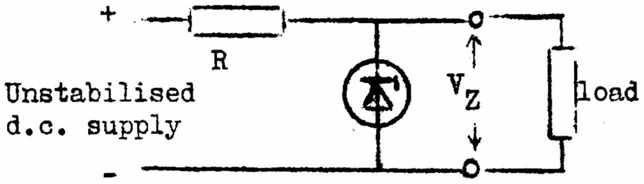
Zener breakdown can be put to use in voltage stabilisers. When in the Zener breakdown condition the voltage across a silicon diode is almost constant.



The reverse voltage at which Zener breakdown occurs can be controlled during manufacture. This voltage is called the Zener Voltage (V_z).

Diodes used for this purpose are called Zener Diodes. The BS symbol for a Zener diode is:-

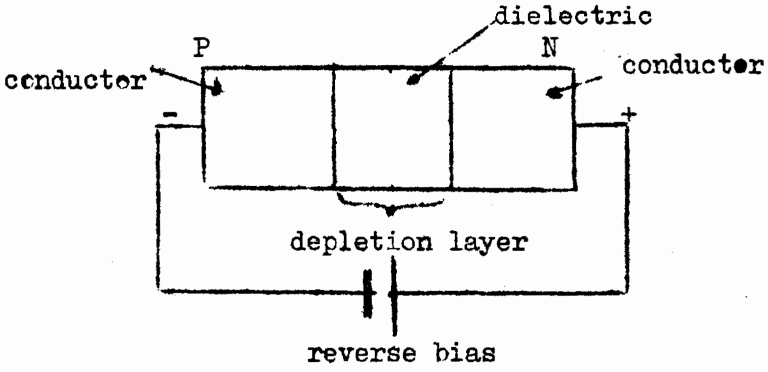
An example of a Zener diode voltage stabiliser is:-



V_z is the stabilised d.c. output voltage.

Zener diodes can be made with Zener voltages from about 3 to 300V.

Another effect which occurs with reverse biased diodes is the variable capacitance effect. We can consider the depletion layer as being a dielectric and the layers of material on each side as conductors. We can therefore represent the junction as:-

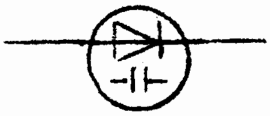


This is effectively a capacitor. The capacitance depends on the width of the depletion layer and therefore on the reverse bias voltage.

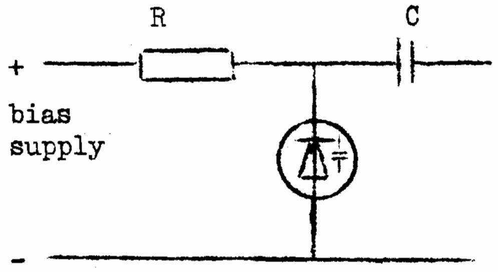
For diodes working at high frequencies a low junction capacitance is required.

The variable capacitance effect can be put to use in certain applications where a voltage variable capacitor is required. An example of this is in automatic frequency control (AFC) circuits. Special high capacitance diodes are used for this purpose, called Varactor Diodes.

The BS symbol for a varactor diode is:-



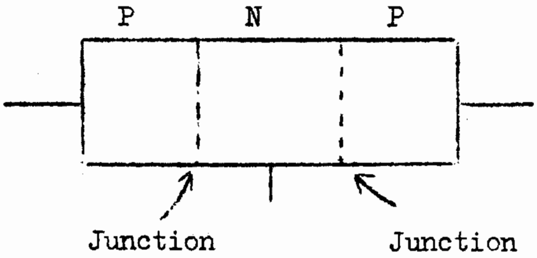
A typical circuit for a varactor diode would be:-



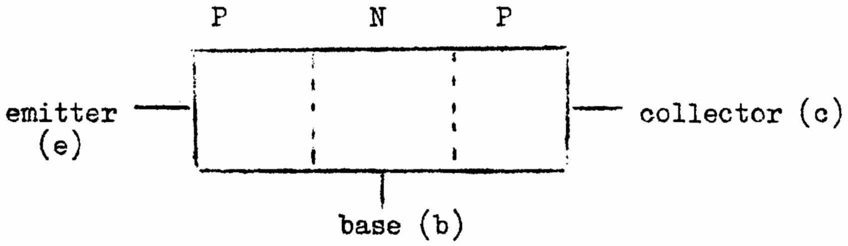
R stops the bias supply loading the a.c. signal and C blocks the d.c. from the rest of the circuit.

The Junction Transistor

Up to now we have only considered devices which have one junction. Let us now look at a transistor which has two junctions.

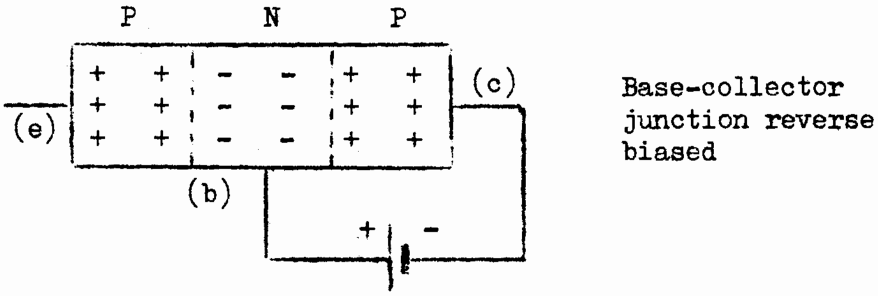


There are three connections to a transistor. These are called the Emitter, Base and Collector.



This type of transistor is called a P-N-P transistor.

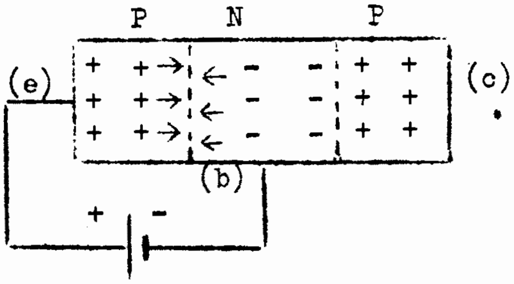
Consider the following circuit in which the base-collector junction is reverse biased.



Only the thermal current flows because the potential barrier stops holes and electrons from entering the base-collector junction.

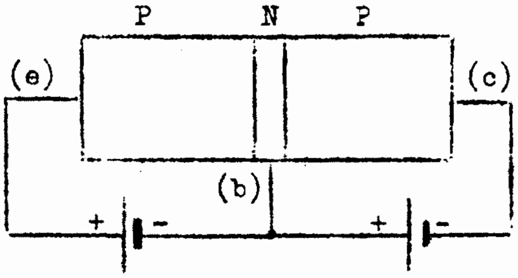
In this diagram only the free electrons and holes have been shown. The fixed ions have been omitted.

If a forward bias is applied across the emitter-base junction a current flows between emitter and base.

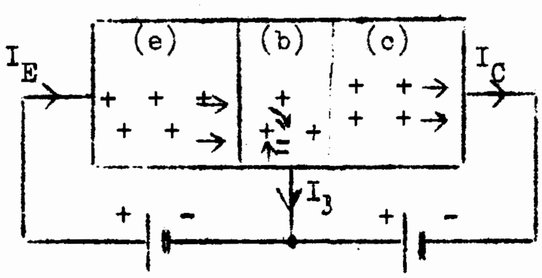


Emitter-base junction forward biased

Let us now apply the reverse bias to the base-collector junction and the forward bias to the base-emitter junction and make the base (N) region very thin.



Holes are attracted from the emitter into the base by the forward bias. If the N region is very lightly doped compared with the P regions, there will be some combination of the holes with the electrons in the N region, but most of the holes will drift towards the base-collector junction. When they reach this junction they come under the influence of the collector voltage and move towards the collector terminal where electrons enter from the battery supply to combine with the holes. Thus there is a flow of holes from emitter to collector although a small proportion never reach the collector owing to combination in the base region. Therefore a collector current flows almost equal to the emitter current and there is a small base current because electrons must enter the base to make up for those lost in combination.



The number of holes entering the base depends upon the forward bias applied to the emitter-base junction. The ratio of the number of holes reaching the collector to the number of holes combining with electrons in the base is a fixed ratio for a given type of transistor.

Expressed in terms of the external currents I_E, I_C and I_B (in the last diagram) we can say:

$\frac{I_C}{I_B}$ is a constant and may be as high as several hundred.

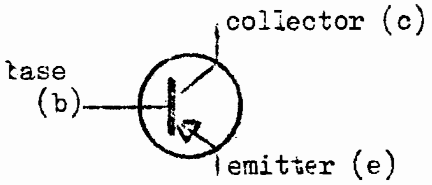
Also: $I_E = I_C + I_B$.

and $\frac{I_C}{I_E}$ is slightly less than unity.

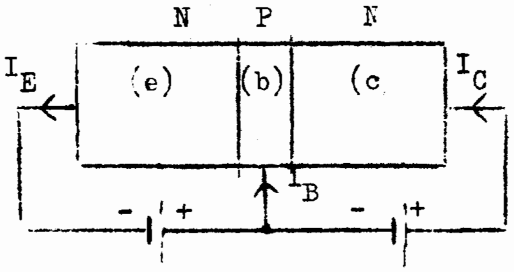
The collector current I_C is hardly affected by the magnitude of the collector-base voltage V_{CB} (provided that it exceeds a certain minimum value). I_C is therefore determined by V_{BE}.

Instead of controlling I_C by variation of V_{BE} it is possible to control the base current, I_B, by an external circuit. This will control V_{BE} and therefore I_C. Further, if I_B is varied by means of an alternating signal, I_C will also vary and will be a magnified version of the base current change.

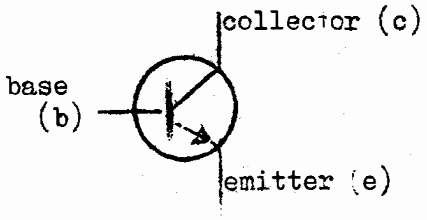
The BS symbol for the P-N-P transistor is:-



The simple description of the operation of the NPN transistor is similar to that for the PNP except that the current carriers from the emitter will be electrons instead of holes. The d.c. conditions will be as follows:-



The BS symbol for the NPN transistor is:-



M.R.H. Scott/MP
23rd January 1970