Chapter 13: Further Semiconductors Components

1. Conduction in semiconductors

Learning Objectives:

At the end of this topic, you will be able to:

- recall the conduction processes in n-type and p-type semiconductors in terms of electrons and holes
- recall conduction processes at a p-n junction and the reasons for differences in the conducting properties of a p-n diode under forward and reverse bias
- recall the principles of operation of photodiodes and LEDs.

Conductors, semiconductors and insulators

All substances are made up from atoms, in which tiny negatively-charged particles called electrons orbit around the positively charged nucleus. The electrons are bound to the nucleus by the attractive force between their opposite electrical charges.

That force depends on factors like:

- the average distance from electron to nucleus;
- the shielding effect of other electrons between the electron and the nucleus;
- how great the positive charge on the nucleus.

As a result, some electrons are more tightly bound than others. Given enough energy, an electron can escape from an atom, though usually not from the material. It can then wander through the material, responding to the effect of external electric fields and taking part in an electric current.

In some materials, it takes a lot of energy to free an electron from an atom. The currents that flow in these materials are tiny. We call these materials **insulators** (e.g. glass).

In others, some electrons receive enough energy even at room temperature to escape and wander through the material. These allow considerable currents to flow in response to external electric fields and are known as **conductors** (e.g. copper).

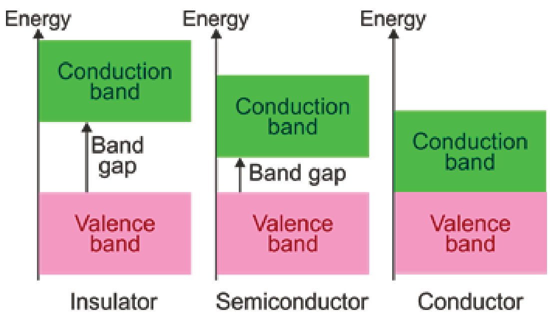
In between these extremes, electrons in a third class of materials require more energy to escape. At room temperatures, they act as poor insulators. When warmed, they begin to conduct a little. These are called **semiconductors** (e.g. silicon).

Energy band diagrams

This gives the same information as above but in the form of a graph having only one axis - energy.

Electrons that have sufficient energy to escape from their ‘home’ atom are found in the *conduction band.* These wander through the material and can take part in an electric current.

Electrons with less energy cannot do this and are found in the *valence band*. They are known as ‘valence electrons’, and their motion serves to bind the atoms together.



In insulators, there is a sizeable gap between the highest energy level in the valence band and the lowest level in the conduction band. This *energy band gap* means that an electron must gain considerable energy before it can take part in conduction. This is unlikely and so insulators do not conduct electricity very well. In other words, they have a high resistance to the flow of electricity.

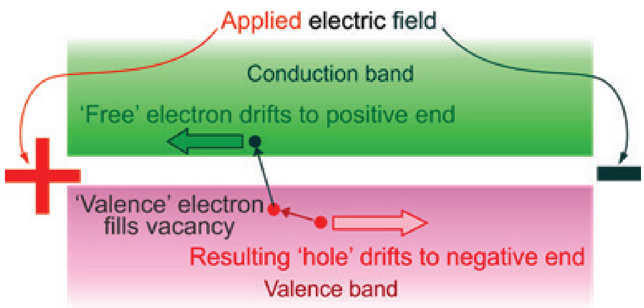
At the other extreme, conductors have no energy band gap at all. At absolute zero temperature, all the electrons remain in the valence band and no conduction takes place. Above that temperature, however, electrons can gain enough energy to jump into the conduction band and take part in an electric current. Conductors have a low resistance.

In between these extremes, semiconducting materials have a small energy band gap. At low temperatures, all the electrons remain in the valence band and no conduction takes place. When the temperature rises high enough, they start to gain enough energy to jump into the conduction band and a current can flow. A significant, though somewhat peculiar, effect is that the gaps (vacant energy levels) they leave behind in the valence band can also allow conduction. This ‘hole conduction’ is described next.

Electrons and holes

The previous section described how ‘free’ electrons (those with enough energy to leap into the conduction band and escape from their ‘home’ atom) can take part in electric currents, attracted by any external electric field applied to the material.

This escape leaves behind a ‘vacancy’ in the valence band. A nearby electron can make a small jump in energy to occupy that vacancy. In the process, it leaves behind a vacancy - and so the process goes on.



Electrons that do not have enough energy to escape from their ‘home’ atom can nevertheless move towards the positive end of the material by jumping into the vacancies left by electrons that do transfer to the conduction band.

The vacancies created by this process, known as ‘holes’, migrate towards the negative end of the material, as if they were positively-charged. Holes behave like positively-charged particles. There are two components to the electric current: the current of ‘free’ electrons wandering towards the positive end of the material; and the current of ‘holes’ drifting towards the negative end of the material.

A free electron in the conduction band may lose energy and drop back into the valence band, recombining with and annihilating a hole. Its energy may be converted to thermal vibrations (warming up the crystal) or it may create a pulse of light (the principle behind the **LED**).

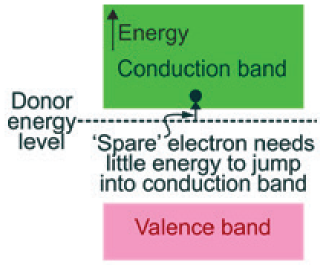
The reverse process takes place in a **photodiode**, where a pulse of light is absorbed and creates an electron-hole pair in the process. When the device is biased appropriately, an appreciable current can be created in this way.

The behaviour described above is known as *intrinsic* behaviour. One hole is created when one electron jumps to the conduction band. The number of holes is equal to the number of free electrons.

The effect of impurities

The addition of *impurity* atoms (i.e. not atoms of the semiconducting material) changes the electrical properties radically. Useful impurities are either *donors* or *acceptors*.

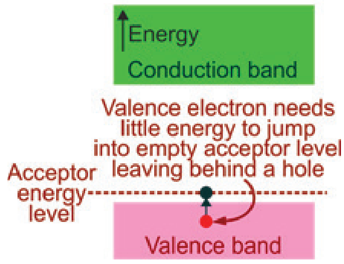
A donor atom offers an extra electron to the semiconductor crystal. This electron takes no part in bonding the atoms together and can be easily liberated from the atom. It can then take part in an electrical current. In this case, a free electron is created without a corresponding hole.



The material contains more free electrons than holes.

Semiconductors 'doped' with donor atoms are known as *'n-type'* extrinsic semiconductors.

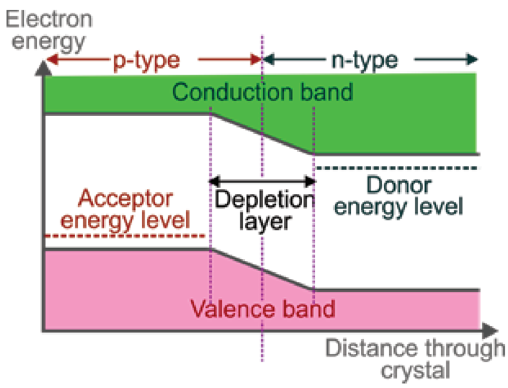
An acceptor atom has one fewer electron able to take part in bonding than the semiconductor. As a result, an electron from a neighbouring semiconductor atom jumps into the vacancy (hole) created. No free electron was created and so now there are more holes than free electrons.



Semiconductors doped with acceptor atoms are known as 'p-type' extrinsic semiconductors.

The p-n junction

It is possible to diffuse different impurities into different sections of a single semiconductor crystal. In this way, the crystal can be part p-type and part n-type semiconductor. Where they meet, a p-n junction is formed.



The excess holes in the p-type region diffuse into the n-type region. The excess free electrons in the n-type region diffuse into the p-type region. They all quickly recombine.

The results:

- A depletion region is created which has neither free electrons nor holes.
- The energy bands in the p-type region are raised in energy compared to those in the n-type region.
- A 'built-in voltage' (around 0.7 V for silicon). This is better called a 'built-in potential' as it cannot be measured directly on a voltmeter and cannot drive a current.

This energy 'step' causes the p-n junction to act as a rectifier. It makes it difficult for free electrons in the n-type region to flow into the p-type region unless they gain some energy. Equally, it is difficult for the holes to migrate from the p-type region into the n-type region.

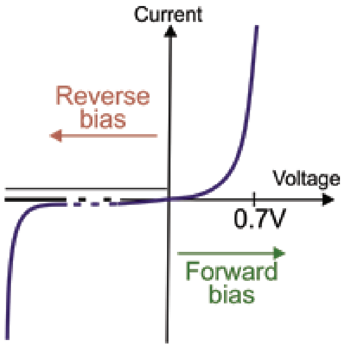
However, if the p-type region is connected to a positive voltage supply, free electrons in the n-type region are attracted to the p-type region (i.e. given potential energy). The holes, effectively positively charged, are pushed into the n-type region. When the free electrons reach the end of the p-type region, they flow out into the external circuit.

When holes reach the end of the n-type region, they combine with electrons in the external circuit. In other words, the p-n junction passes a current. This condition, where the p-type region is positive with respect to the n-type region, is known as **forward bias**.

When the external voltage is applied in the opposite direction, i.e. p-type region being negative with respect to n-type region, free electrons tend to remain in the (relatively positively-charged) n-type region while the holes (effectively positively-charged) remain in the p-type region.
The p-n junction no longer passes a current.
This condition is called **reverse bias**.

This broadly explains the current / voltage characteristic of the p-n junction, shown below.

- When forward biased, very little current flows until the applied voltage reaches 0.7V. This is enough to counteract the 'built-in potential' and give free electrons and holes sufficient energy to flow across the p-n junction.
- When reverse biased, very little current flows. When the applied voltage is big enough, it can trigger the creation of huge numbers of free electrons and holes. This can be the result of electrons gaining sufficient energy to jump into the conduction band, or can be the result of an electron avalanche, where highly energetic electrons collide with and liberate other electrons.



2. The MOSFET

Learning Objectives:

At the end of this topic, you will be able to:

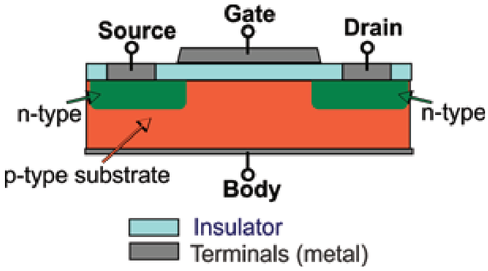
- explain the properties of an n-channel enhancement mode MOSFET in terms of the effects of bias voltage on the conducting channel (pinching).

MOSFETs are available in two basic forms: depletion type and enhancement type. This course focuses on the enhancement type.

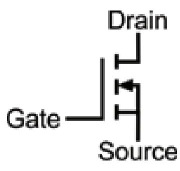
The first picture (not a literal plan of a device) aims to illustrate the layout of the n-channel enhancement mode MOSFET.

It consists of:

- a p-type crystal, called the **substrate**;
- two identical n-type regions, one of which becomes the **source** and the other the **drain**;
- an insulating 'skin' separating these from the control terminal, the **gate**;
- a **body** connection, which can be used to ground the substrate to prevent its voltage from floating (it is often connected to the source terminal and to ground. In some diagrams, it is omitted to simplify the treatment).



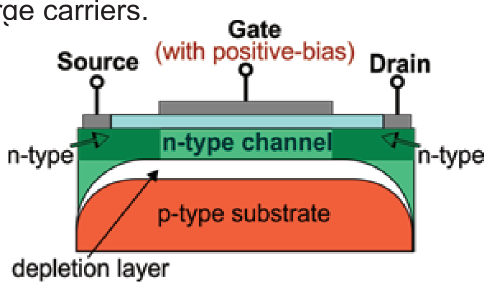
A MOSFET can be called a voltage-controlled device because the output current (from drain to source) is controlled by the voltage applied to the gate terminal. This controls the resistance of the conducting channel between the drain and the source, and hence controls the output current.



The channel region is only lightly 'doped' with p-type impurity atoms, and there are very few free electrons or holes in it. As a result, when no voltage is applied to the gate, the channel resistance is high and negligible current flows. The channel is 'pinched' off. The circuit symbol reflects this by having a broken line between drain and source terminals.

When a positive voltage is applied to the gate, electrons are attracted towards it and holes migrate away from it. Initially, the electrons moving near to the gate terminal recombine with holes in the p-type semiconductor and create a depletion layer devoid of free charge carriers.

When the applied voltage is increased, eventually enough electrons are attracted to the region just below the gate to form a conducting channel. This is known as an inversion layer as this region of the crystal, initially p-type is now n-type, containing more free electrons than holes.



The voltage needed to do this is called the **threshold voltage**. Once the applied gate voltage exceeds the threshold voltage (about 3 V), the width of the channel and the number of free electrons it contains increases, and so its resistance falls, leading to an increase in output current (again, the second picture aims only to **illustrate** the effect.).

The n-channel enhancement mode MOSFET behaves like a ‘normally-open’ switch:

- When the gate voltage is zero (or negative) the device is ‘off’.
- When a sufficient positive gate voltage is applied, it turns ‘on’.

They make excellent switches, having:

- low ‘on’ resistance
- extremely high ‘off’ resistance
- infinitely high input resistance thanks to the layer of insulator between the gate terminal and the substrate.