CHAPTER 12

From semiconductors to solid state devices

Limitations of past technologies and increased research into the structure of the atom resulted in the invention of transistors

Introduction

The development of semiconductors from the first germanium 'crystal set' radio to modern microprocessors containing millions of transistor connections was largely driven by the need for reliable, portable radio transceivers. Parallel to this need was the development of more sophisticated and complex electronic circuits, which were based on an increasing number of valves—unreliable, power-hungry and bulky vacuum tubes. The photograph (right) shows such a device from an old radio navigation unit, the modern equivalent of which could easily fit in the palm of your hand. The increased knowledge of the behaviour of semiconducting elements—germanium and later silicon, along with the process of 'doping', allowed huge improvements in electronics to occur—improvements that today we largely take for granted, but without which our modern society could not function.

An old valve-based radio navigation unit

Valence shell and valence electrons

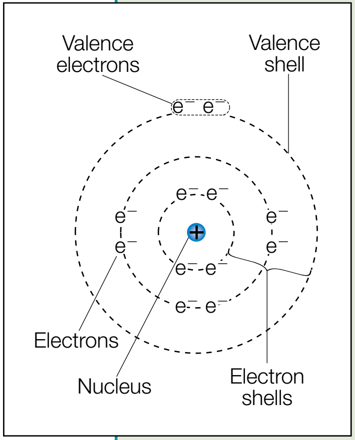
12.1

An atom consists of a small and dense positive nucleus that contains protons and neutrons; around the nucleus are electrons, which are organised into distinct orbits called electron shells. Different atoms have different numbers of electron shells, but each electron shell can only hold a certain number of electrons.

The outermost electron shell of all atoms is given the name **valence shell**, and the electrons contained in the valence shell are called **valence electrons**. There can be a maximum of *eight* valence electrons in a valence shell, and they have higher energy compared to the electrons in the other shells. (The innermost electron shell has the lowest energy level.)

Consider the magnesium atom, which has 12 electrons in total: the organisation of these electrons into electron shells is shown in Figure 12.1. There are three electron shells for the magnesium atom, and the valence shell contains two valence electrons. These two electrons have higher energy than any of the other 10 electrons.

Figure 12.1 Electron shells



12.2 Metals: metallic bonds and the sea electron model

■ *Identify that some electrons in solids are shared between atoms and move freely*

Metallic bonds are interactions through which metal atoms are joined together to form a lattice structure; it is best described by the 'sea of electrons' model.

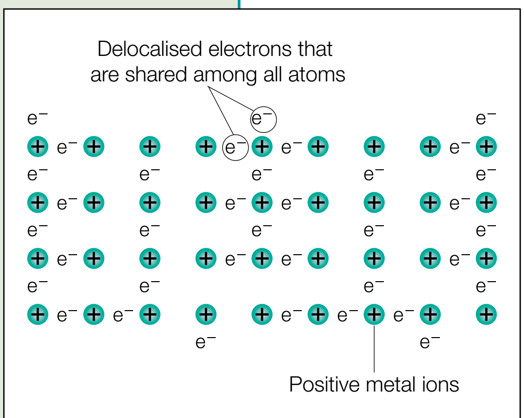


Figure 12.2
Lattice of sodium metal

Definition

According to the **'sea of electrons' model**, the bondings between metal atoms are described as a lattice of positive metal ions surrounded by a sea of delocalised valence electrons (see Fig. 12.2).

Delocalised valence electrons refers to the fact that the valence electrons of all metal atoms are freed and are shared among other metal atoms. These electrons therefore do not belong to any specific atom, they have high energy and are moving freely about.

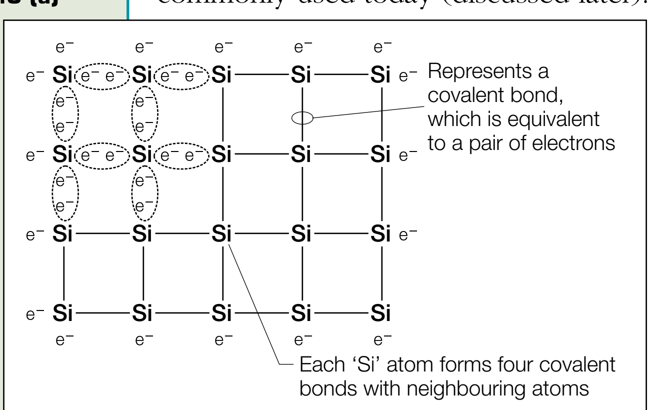
Delocalised electrons are responsible for stablising and holding metal atoms together in a lattice. It is also important to remember it is these delocalised electrons that give metals their unique physical properties, for instance, high thermal and electrical conductance.

In Figure 12.2, we have a lattice of sodium metal. Each sodium atom has one valence electron that is delocalised and shared among all other atoms.

12.3 The structure of semiconductors

You need to have some ideas about what conductors and insulators are. The easiest way to define **semiconductors** is that they have properties in between that of conductors and insulators. Common semiconductors are germanium and silicon, with silicon being more commonly used today (discussed later).

Figure 12.3 (a)



Note: This is a three-dimensional representation of the structure of silicon. Each silicon atom is bonded to 4 other silicon atoms in a tetrahedral fashion. This structure is repeated indefinitely and constitutes the entire crystal lattice. Bonds for silicon atoms on the edges are not drawn for the purpose of clarification.

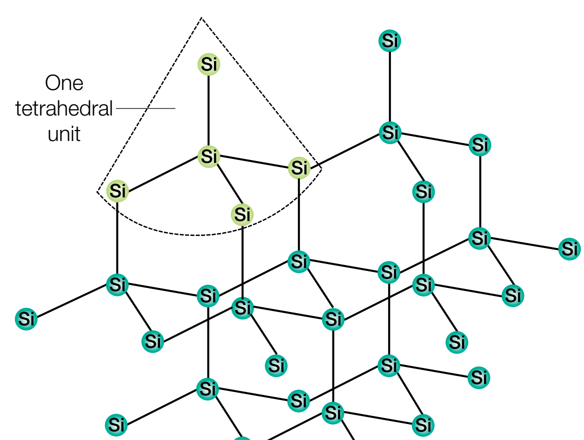


Figure 12.3 (b)

As both germanium and silicon are group IV elements, they have *4* valence electrons. This allows each silicon or germanium atom to make *4* covalent bonds with the neighbouring silicon or germanium atoms to form a macro-covalent lattice structure as represented two-dimensionally in Figure 12.3 (a). The four covalent bonds around each atom form a tetrahedral shape three-dimensionally, as shown in Figure 12.3 (b).

12.4 Band structure and conductivity for metals, semiconductors and insulators

- ***Describe the difference between conductors, insulators and semiconductors in terms of band structures and relative electrical resistance***
- ***Compare qualitatively the relative number of free electrons that can drift from atom to atom in conductors, semiconductors and insulators***

When an atom is by itself, the energy level of its electrons is sharply defined; that is, the energy is the same for all the electrons in the same (corresponding) electron shell, and different from those in the other electron shells.

However, when atoms are packed together quite closely to form a lattice structure, electrons in a single atom will be constantly interacting with neighbouring electrons of other atoms or even positive nuclei of other atoms. This results in the energy level of the individual electrons changing slightly and therefore being no longer sharply defined. In a broad sense, the energy of all the electrons in the lattice is not sharp and can take all ranges of values. This blurring of the energy level of electrons in a lattice structure results in the formation of electron **energy bands**.

Definition

The **energy band** is the range of energy electrons possess in a lattice.

There are two types of energy band you need to know about:

- *Valence band*: The **valence band** is made up of the energy levels of the valence electrons of individual atoms. It has higher energy than energy bands formed by the electrons found in the inner shells.
- *Conduction band*: When valence electrons gain energy, they might move up to even higher energy shells that were previously empty. These electrons (energy levels) make up the **conduction band**. Once in the conduction band, these electrons are free to move and therefore are able to conduct electricity (hence the name conduction band).

Except for conductors, electrons in the conduction band usually have higher energy than those in the valence band. The energy gap that electrons have to overcome to move from the valence band to the conduction band is referred to as the **forbidden energy gap**.

We are now going to examine the band structures for metals, semiconductors and insulators and analyse how they are related to the conductivity of each. It is easier to start with semiconductors, to illustrate the basic principles mentioned above, then the differences will be noted for metals and insulators.

Semiconductors

■ *Identify absences of electrons in a nearly full band as holes, and recognise that both electrons and holes help to carry current*

Band structure

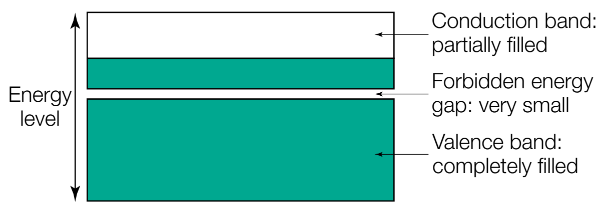
You have learnt that the **4** valence electrons of each semiconductor atom form **4** covalent bonds with the neighbouring atoms, therefore all are locked in position. Also the formation of the covalent bonds means the valence shells now have **8** electrons, consequently the **valence band** is full.

NOTE: One covalent bond equals a pair of (shared) electrons.

By the nature of a semiconductor, the valence electrons only need to gain a small amount of energy if they are going to move into the conduction band; hence, the conduction band is separated from the valence band by a very **small forbidden energy gap**.

At room temperature, a minority of the electrons in the valence band can possess enough thermal energy to overcome the small forbidden energy gap to 'jump' into the conduction band. Hence at room temperature, the **conduction band** of a semiconductor is partially filled. See Figure 12.4.

Figure 12.4 The band structure of a semiconductor



Conductivity

The small number of electrons in the conduction band means at room temperature the conductivity of a semiconductor is *moderate*.

It is also important to remember:

As the temperature of a semiconductor increases, its conductivity increases and its resistance decreases.

NOTE: Conductivity (conductance) is inversely related to resistance; that is

$$\text{conductivity} = \frac{1}{\text{resistance}}.$$

This is because as the temperature increases, the electrons in the valence band gain more thermal energy. This allows more electrons in the valence band to possess enough energy to overcome the forbidden energy gap to move into the conduction band. More electrons in the conduction band effectively translate to a higher conductivity for the semiconductor. Although the increase in the temperature also increases the total number of **undesirable collisions** between the conducting electrons and the lattice (which tends to decrease the conductivity), this is largely overruled by the substantial increase in the number of electrons in the conduction band.

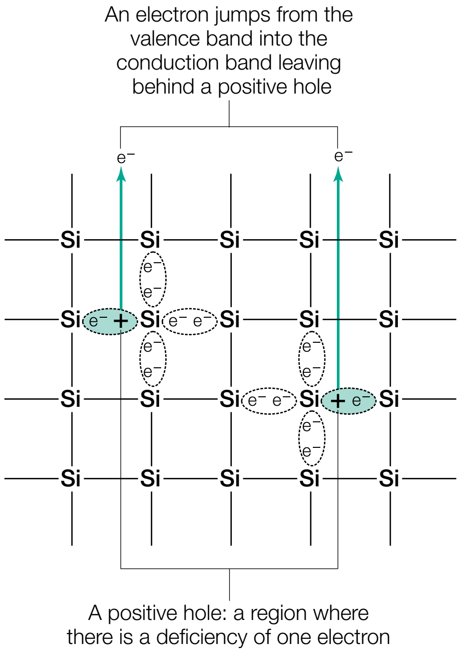


Figure 12.5 (a) Electron hole pairs

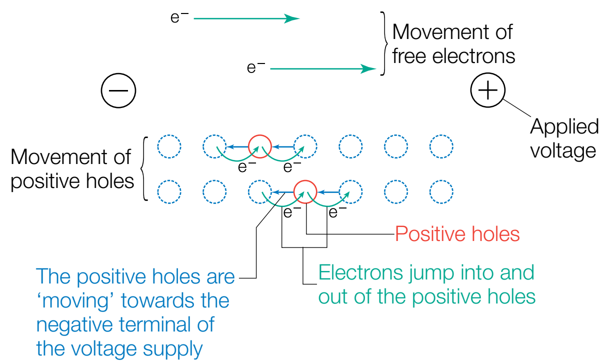


Figure 12.5 (b) Electron hole pair conduction in the presence of a voltage

A characteristic property: electron-hole pair conduction

As one electron 'jumps' across the forbidden energy gap to occupy the conduction band, there will be an electron deficiency in the valence shell, which should have **8** electrons to be full. The missing of one electron from the almost full valence shell constitutes a **positive hole**. See Figure 12.5 (a).

Positive holes are imaginary and do not exist in reality, as they are simply regions in the valence band where there are deficiencies of valence electrons. They are quantum mechanic models that allow electrons to jump in and out without spending a lot of energy. Although positive holes do not migrate, when an electron jumps into a positive hole, there will be another positive hole left at the position where the electron was before. It follows that the movement of electrons creates the movement of positive holes in the opposite direction. Hence when a voltage is applied across a semiconductor, conduction can occur in the conduction band by the movement of free electrons, as well as in the valence band by the movement of positive holes in the opposite direction due to the movement of electrons into and out of these holes. See Figure 12.5 (b). This type of conduction is known as the **electron-hole pair conduction**.

NOTE: A positive hole is only formed in a region where there is one electron deficiency within an electron filled surrounding, that is, in the case of a semiconductor. For this reason, no positive holes can form for metals as all the valence electrons are delocalised.

Conductors (metals)

Band structure

As discussed early in this chapter, the valence electrons for metal atoms are all delocalised and shared. These electrons have gained high energy and are free to move, and hence are able to conduct; these electrons are all in the conduction band. Since all the valence electrons of a conductor (metal) are in the conduction band, thus the valence band of a conductor is said to **merge** with the conduction band. The forbidden energy gap of course is non-existent. This is represented in Figure 12.6.

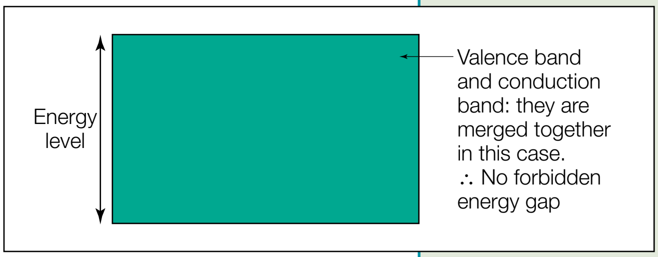


Figure 12.6 The band structure of a conductor

Conductivity

Since the valence band of a conductor is merged with the conduction band, this means that there are as many electrons found in the conduction band as there are in the valence band. Since a single metal lattice will contain billions of metal atoms thus billions of valence electrons, the population of electrons in the conduction band is very high. Consequently a conductor has a very good electric conductivity, or low electric resistance.

It is important to remember:

As the temperature of a conductor increases, its conductivity decreases and its resistance increases.

This is because as the temperature of a conductor increases, the lattice will possess more thermal energy, which causes it to vibrate more vigorously. These vibrations lead to more collisions between the conducting electrons and lattice and therefore impede the motion of these electrons; this results in a decrease in conductance or an increase in resistance of the conductor.

NOTE: No extra electrons can be recruited into the conduction band for a metal at a higher temperature, as the population of electrons in the valence band is already a maximum.

A word about drift velocity

Even when there is no potential difference applied across a conductor, its electrons are still moving at very high speed in all directions. This speed can reach 10^5 to 10^6 m s^{-1}; however, due to their randomness, no net current will flow. See the solid lines in Figure 12.7.

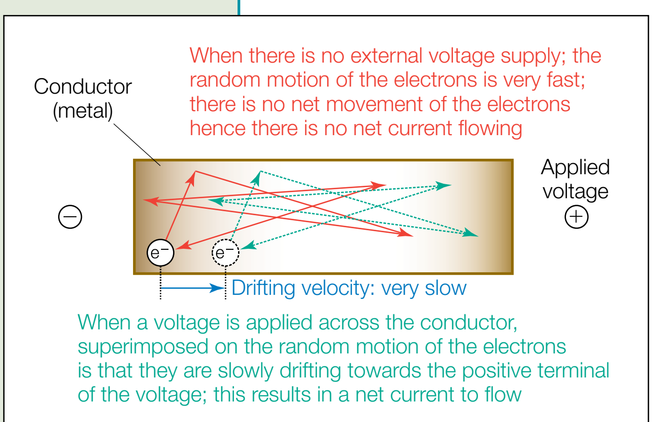


Figure 12.7 Drift velocity of electrons

When a voltage is applied across a conductor, superimposed on top of the random motions of the electrons is that all the electrons start to 'drift' slowly in one uniform direction. This results in a net current to flow. See the dashed lines in Figure 12.7. How fast the electrons will drift in response to the applied voltage is known as the **drift velocity**. Such a velocity is generally very slow, usually only a few centimetres per second.

The drift velocity can be quantitatively described by the equation: $\boldsymbol{v} = \dfrac{\mathbf{I}}{\mathbf{neA}}$; where $\boldsymbol{v}$ is the drift velocity (m s^{-1}), **I** is the current flowing through the conductor (A), **n** is the electron density of the conductor (how many electrons in a given volume), **e** is the charge of electron which has a value of 1.602×10^{-19} C, **A** is the cross sectional area of the conductor (m^2).

NOTE: You are not required to perform any calculation using the equation $\boldsymbol{v} = \dfrac{\mathbf{I}}{\mathbf{neA}}$.
However, you do need to know what *factors* determine the size of the drift velocity.

Insulators

Band structure

For an insulator, all the valence electrons are used to form covalent bonds to hold its atoms together within the lattice (similar to a semiconductor). Thus the valence band is full, and the valence electrons are locked in position and can not move. The valence band is separated from the conduction band via a **very large forbidden energy gap**. At room temperature, virtually no electrons in the valence band can gain enough energy to jump across the large forbidden energy gap to occupy the conduction band. Therefore the conduction band is virtually empty. See Figure 12.8.

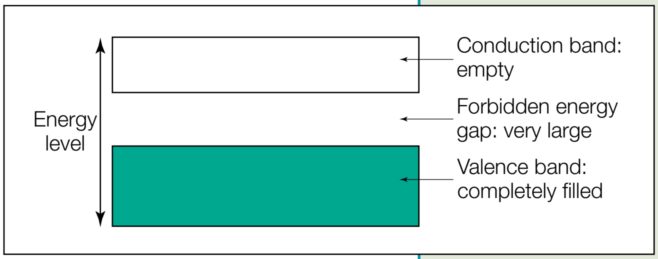


Figure 12.8 Band structure of an insulator

Conductivity

Since the conduction band of an insulator is empty, its conductivity at room temperature is almost zero, and its resistance is infinite.

It is important to note if enough energy is applied to an insulator, that is, by applying a very high voltage or heating it to a very high temperature, the electrons in the valence band will eventually gain enough energy to overcome the forbidden energy gap to occupy the conduction band. In these cases, the insulation property will break down and the insulator will start to conduct; however, during such processes its structure might have already been damaged.

Insulators

12.5 A closer look: intrinsic and extrinsic semiconductors

- ***Describe how 'doping' a semiconductor can change its electrical properties***
- ***Identify differences in p- and n-type semiconductors in terms of the relative number of negative charge carriers and positive holes***

Semiconductors can be broadly classified into two categories:

1. **Intrinsic semiconductors**: Pure semiconductors. These semiconductors conduct electricity by electron-hole pair conduction, and have moderate conductivity at room temperature, as discussed in the previous section.
2. **Extrinsic semiconductors**: Semiconductors that have other types of impurities added. The impurities added to semiconductors can dramatically increase their conductivity by modifying their electrical properties. This will be discussed in detail in the next section.

Extrinsic semiconductors

Extrinsic semiconductors can be further classified into two types, depending on the types of substances (impurities) added (see Fig. 12.9a and b for summary).

1. p-type semiconductors
2. n-type semiconductors

The process of adding substances (impurities) to semiconductors in order to change their electrical properties is known as **doping**. It is important to remember that the amount of impurities added is *very small*, about 0.001%.

p-type semiconductors

A p-type semiconductor is created when a pure semiconductor such as silicon is doped with any group III elements, for instance, boron. All group III atoms have

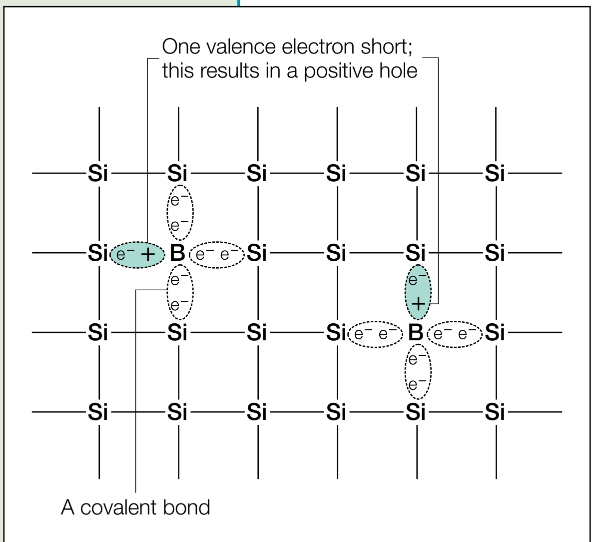


Figure 12.9 (a) A p-type semiconductor

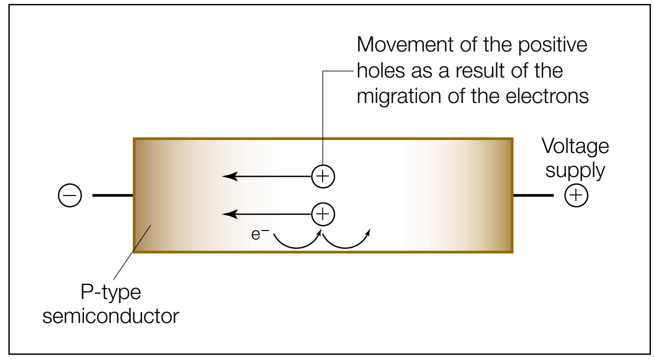


Figure 12.9 (b) Conduction by a p-type semiconductor

3 valence electrons, therefore can only form *3* covalent bonds with neighbouring atoms. This means in the regions of the silicon lattice where the silicon atoms are replaced by the boron atoms, there will be one electron short for the formation of the fourth covalent bond with the neighbouring silicon atoms. These electron deficiencies within an almost filled valence electron band not surprisingly constitute **positive holes** (see Figure 12.9a). Therefore:

p-type semiconductors contain positive holes.

NOTE: Remember 'p' for **p**ositive holes.

NOTE: p-type semiconductors are not positive, as the number of electrons is still equal to the number of protons within the lattice.

When a voltage is applied across a p-type semiconductor, conduction occurs as electrons can easily migrate from the negative to the positive terminal by jumping in to and out of the positive holes, which effectively results in the migration of these holes in the opposite direction (see Fig. 12.9b). Therefore we say the conduction in a p-type semiconductor is carried out by the positive holes. The presence of positive holes allows p-type semiconductors to have much higher conductivity compared to intrinsic semiconductors.

n-type semiconductors

An n-type semiconductor is created when a pure semiconductor such as silicon is doped with any group V elements, for instance, phosphorous. All group V atoms have *5* valence electrons, therefore are capable of forming *5* covalent bonds with neighbouring atoms. Therefore after the phosphorous atoms have formed *4* covalent bonds with the neighbouring silicon atoms in the lattice, each atom will have one spare electron that is not required for bonding (see Fig. 12.10a). These electrons are free to move and have sufficiently high energy to occupy the conduction band. Therefore:

n-type semiconductors contain free electrons.

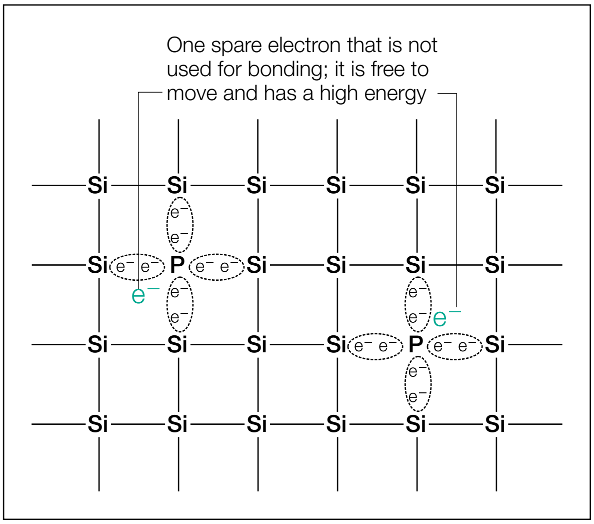


Figure 12.10 (a) An n-type semiconductor

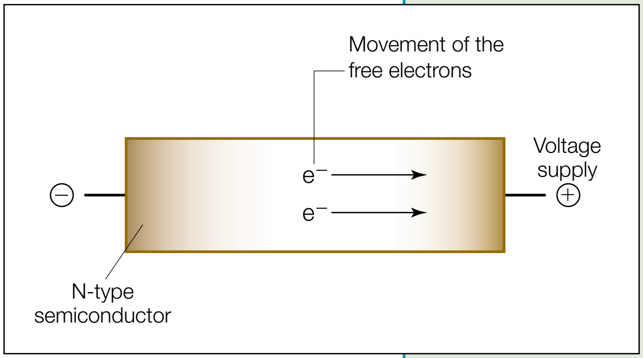


Figure 12.10 (b) Conduction by an n-type semiconductor

NOTE: Remember 'n' for negative electrons.

NOTE: By the same logic, n-type semiconductors are not negative.

When a voltage is applied across an n-type semiconductor, the free electrons can quite readily conduct electricity, as shown in Figure 12.10 (b). Therefore we say the conduction in an n-type semiconductor is carried out by the free electrons. Again the presence of free electrons increases the conductivity of n-type semiconductors quite dramatically.

When a p-type semiconductor is joined with an n-type semiconductor

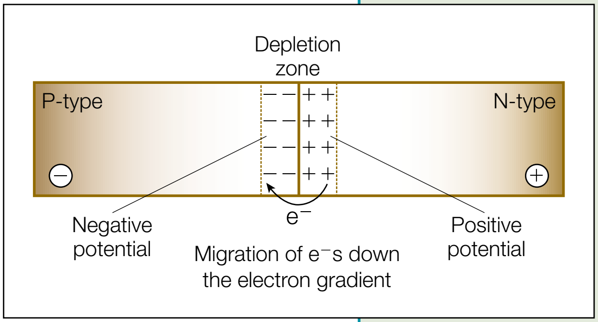
As p-type semiconductors have positive holes and therefore lack electrons and n-type semiconductors have excessive free electrons, when a p-type semiconductor and an n-type semiconductor are joined together, electrons from the n-type semiconductor will migrate into the p-type semiconductor at the junction to fill up positive holes in this area. This is shown in Figure 12.11.

As a result, the p-type semiconductor will now possess more electrons than its protons, consequently displaying a negative charge, whereas the n-type will display a positive charge due to the loss of electrons. This creates a potential difference (electric field) across the junction where electron diffusions had occurred, which is also known as the **depletion zone**. At equilibrium, this potential difference also opposes further diffusion of electrons from the n-type to the p-type semiconductor.

So remember:

The p-type semiconductor becomes negative and the n-type semiconductor becomes positive when they are joined together.

Figure 12.11 Joining a p-type and n-type semiconductor



Implementations of semiconductors: solid state devices

SECONDARY SOURCE INVESTIGATION

PFAs

H1, H3

PHYSICS SKILLS

H12.3A, B, D
H13.1A, B, C
H12.4F
H14.1G, H

- ***Gather, process and present secondary information to discuss how shortcomings in available communication technology lead to an increased knowledge of the properties of materials with particular reference to the invention of the transistor***

The need for the transistor

During World War II, the use of thermionic devices (vacuum tubes; see section 12.6, page 222) became important in communications and in other developing technologies such as radar. Reliable communication between pilots of aircraft and between pilots and control towers, as well as between field commands and troops in the field was needed. Rugged, light and reliable transceivers that could be powered by batteries rather than mains power for portability were needed. Radar became an increasingly important tool for defence as it was able to detect approaching bombers long before they could be seen or heard by lookouts. With electronic circuits becoming more complex, more vacuum tubes, or valves, were required, increasing the unreliability, size and power requirements of the devices being built.

The need to replace vacuum tubes, with all their inherent problems, saw the attention of researchers turn to solid state semiconductors. Germanium had already been identified as the first such material able to be obtained with sufficient purity as to be useful. However, it took over a decade of research and experimentation until the transistor was invented by John Bardeen and William Shockley in 1947, two years after the war had ended.

WWW→

USEFUL WEBSITE

For a comprehensive timeline leading to the invention of the transistor:
http://www.pbs.org/transistor/index.html

Definition

Solid state devices are electronic devices made from semiconductors.

Diodes

In the above section, we saw what happens when a p-type semiconductor is joined with an n-type semiconductor. This type of arrangement also has practical uses. Such a device is known as a solid state **diode**.

*A **diode** is an electronic device which only allows electric current to flow in one direction.* When the p-type part of a diode is connected to the positive terminal of a power source and the n-type part to the negative terminal, a current will flow in the circuit unimpeded (see Fig. 12.12a). In this case, we say the diode is **forward biased**. However, if the connection is reversed, that is, connecting the p-type to the negative terminal and the n-type to the positive terminal, then no current can flow through the diode, in which case we say the diode is **reverse biased**.

NOTE: How diodes are able to carry this function is basically due to the movement of electrons and positive holes within the semiconductors. The principle is not very complicated; however, because this knowledge is not explicitly required by the syllabus, it will not be discussed in this book.

Diodes are very important in electronics; for instance they form the basis of current rectifiers, which are devices that are able to convert AC to DC.

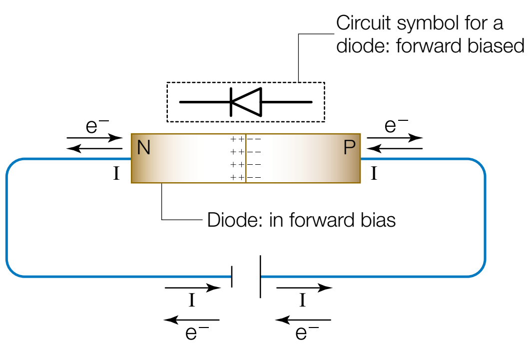


Figure 12.12 (a)

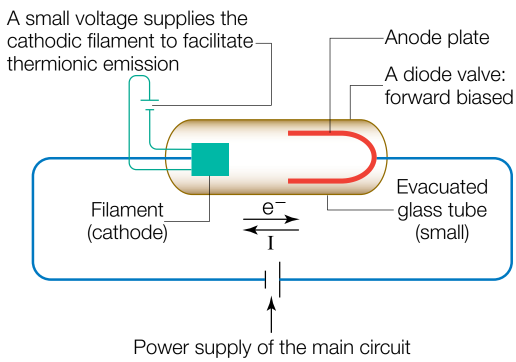


Figure 12.12 (b)

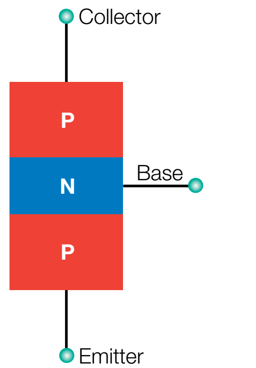
Diodes are not the only examples of solid state devices. By combining p-type and n-type semiconductors in different ways, many other useful solid state devices are created. **Transistors** will be discussed here; others will be discussed in a later section.

Diodes

Transistors

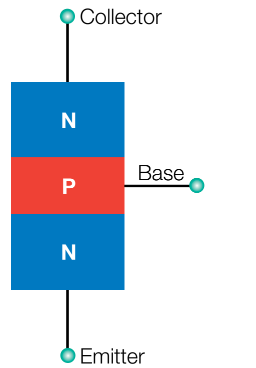
Structure of transistors: Transistors are another type of solid state electronic device. A transistor can be created by either sandwiching a thin layer of p-type semiconductor in between two pieces of n-type semiconductor or a thin layer of n-type in between two p-type semiconductors. The first type of arrangement is more common. A transistor has three connecting leads: collector, base and emitter. These arrangements are illustrated in Figures 12.13 (a) and (b).

Functions of transistors: Transistors have extensive applications in electronics. The three leads of a transistor allow it to be connected across two circuits. One circuit goes through the emitter and the base, while the other one goes through the emitter and the collector (see Fig. 12.14). The flow of charge carriers (electrons) through the emitter and the base alters the electric property of the middle piece semiconductor; this will affect the conductivity of the transistor between the emitter and collector, thereby affecting the flow of charges in that circuit. (Students do not need to know the details of how that happens.) It follows that a



Conduction is by positive holes

Figure 12.13 (a) A PNP transistor

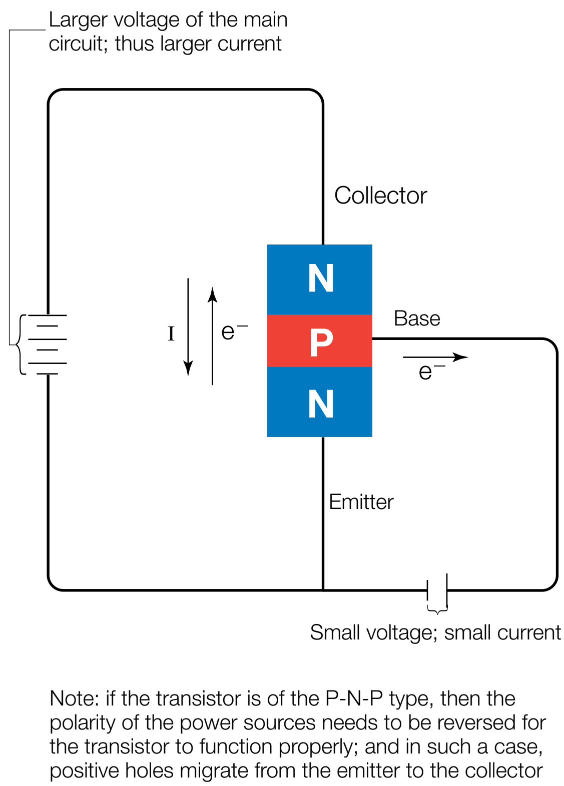


Conduction is by free electrons

Figure 12.13 (b) An NPN transistor

Transistors

Figure 12.14
A single NPN transistor in the circuit



small current flowing through the base can modulate the flow of—usually—a larger current in the main circuit that goes through the collector. This allows the small current to make a larger copy of itself in the main circuit; hence, transistors are often used in electronics as **amplifiers**. A small current flowing through the base may also facilitate or completely stop the current flow in the main circuit (through the collector); hence, transistors can also act as electronic **switches**.

12.6 Solid state devices versus thermionic devices

■ ***Describe differences between solid state and thermionic devices and discuss why solid state devices replaced thermionic devices***

A word about thermionic devices

The easiest way to define a **thermionic device** is that it consists of a vacuum tube in which is embedded two or more electrodes, very much resembling the structure of a cathode ray tube, although not as big. Emission of electrons is from the cathode with the aid of the **thermionic effect** (described in Chapter 10), that is, by heating.

A **diode valve** is an example of a thermionic device; it is the thermionic counterpart of a **solid state diode** made from semiconductors. It consists of a small vacuum tube with *two* embedded electrodes, an anode and a thermionic cathode. Its structure and the way it is connected in a circuit to give a forward bias are shown in Figure 12.12 (b). (Try to compare this with Figure 12.12a.) Diode valves were used in electronics before the invention of solid state diodes to restrict the flow of current to one direction.

Diode valves

A **triode valve** is another example of thermionic device. It has a similar structure to a diode valve, however as its name suggests it has a third electrode inserted in between the anode and cathode. It is the thermionic counterpart of a **transistor**. Not surprisingly, the function of a triode valve is such that a small current passing through the third electrode (which makes it more or less negative) can be used to control and modify a large current in the main circuit that is going through the anode and cathode. Although today they have all been replaced by transistors, before the invention of transistors, many electronic devices relied on triode valves. For instance, the world's first computer was built based on triode valves.

It was undeniable that the invention of thermionic devices revolutionised the whole electronics industry and made lots of electronic devices possible many decades ago. However, due to the problems associated with their use, such as their large size, fragility and inefficiency, scientists were constantly seeking suitable replacements. Eventually the thermionic devices were all replaced by the newer solid state devices, which could carry out similar functions but were far superior in many other aspects.

Some advantages of solid state devices over thermionic devices

- *Miniature size*: The relatively large size of thermionic devices limits their uses. Solid state devices such as diodes and transistors are considerably smaller (in millimetres). Further reduction in size can be achieved in a microchip (discussed later in this chapter) which may contain millions of transistors within the size of a fingernail. The trend of miniaturisation of electronic devices, such as mobile phones and lap-top computers, means the tiny solid state devices are much preferred.
- *Durable and long lasting*: Solid state devices are quite tough and can withstand a reasonable amount of physical impact. (Dropping a transistor might not necessarily break it.) On the other hand, thermionic devices are made from glass bulbs, which make them extremely fragile, so they need to be handled with care. Also, solid state devices generally have a longer life span than thermionic devices, which must be replaced after certain number of uses.
- *More rapid operational speed*: Solid state devices operate at a much faster rate than thermionic devices; this makes them particularly valuable in the production of fast operating microchips and microprocessors. In addition, solid state devices will function as soon as they are switched on, whereas thermionic devices require warming up.
- *More energy efficient*: Thermionic devices require very high voltages for their operation, whereas solid state devices can function at voltages less than 1 V. In addition, a large amount of heat is dissipated during the operation of thermionic devices so that there is a considerable amount of energy wasted. Solid state devices on the other hand only dissipate a small amount of energy during their operation.
- *Cheap to produce*: Solid state devices are much cheaper to make than the thermionic devices, so they are more economical when large quantities are needed.

12.7 Why silicon not germanium?

- ***Identify that the use of germanium in early transistors is related to lack of ability to produce other materials of suitable purity***

Early solid state devices were mostly made from the semiconductor **germanium**. This was because the semiconductors used to make solid state devices must be extremely pure. At this early time, there was only the technology for extracting and purifying germanium, but no technology was available to prepare silicon with sufficiently high purity for the production of solid state devices.

Today, technology allows the manufacture of extremely pure silicon. Therefore, almost all solid state devices are made from **silicon**, as silicon has many superior properties compared to germanium. These include:

- *More economical*: Silicon can be extracted from sand (SiO_2), which is very abundant. The abundance of silicon allows it to be obtained at a much lower price and consequently reduces the price of solid state devices.
- *Functions well under high temperature*: Heat is produced while electronic devices are operating, which elevates the temperature of such devices. Under these high temperatures, silicon will still maintain its semiconductivity, while germanium tends to become a better conductor, losing its semiconductor properties.
- *The ability to form an oxide layer*: Silicon is able to form an impervious silicon dioxide layer when it is treated by heat in the presence of high oxygen content. This is an essential property for the production of microchips.

More solid state devices: solar (photovoltaic) cells

SECONDARY SOURCE INVESTIGATION

PHYSICS SKILLS

H12.3A, B, D
H12.4F
H13.1A, B, C, E
H14.1E, F, G, H

PFA H3

'Assesses the impact of particular advances in physics on the development of technologies'

- ***Identify data sources, gather, process and present information to summarise the effect of light on semiconductors in solar cells***
- ***Summarise the effect of light on semiconductors in solar cells***

What is the advance in physics in this case?

Utilising the photoelectric effect to transform light energy into electrical energy, photovoltaic cells, or solar cells, use the photoelectric effect within semiconductor materials that have been arranged in such a way as to generate an electromotive force, or EMF. This EMF can in turn be used to do useful work in an external circuit.

Which technologies arose from this advance in physics?

Solar cell technology, or photovoltaics, is one of the most exciting developments in the area of alternative energy sources. Its application is already widespread in remote communities and in smaller, mobile applications such as boats where connection to the mains power grid is not possible or practical. Satellites and space stations such as the International Space Station use solar panels as a source of electricity for their long-term missions.

Millions of dollars are being spent worldwide on further developing this technology. The University of New South Wales Centre for Photovoltaics is one of several Australian research facilities attempting to improve the efficiency of the conversion of sunlight to electricity and to lower the cost of solar cells. Application of solar cells as a viable alternative to base-load coal-fired power stations is hampered by the inability to store the huge amounts of electricity required for the periods when the sun is not shining.

Assessment of the impact of this advance in physics on the development of new technology

WWW→

USEFUL WEBSITES

How solar cells work:
http://www.howstuffworks.com/solar-cell.htm

With the use of the photoelectric effect (see Chapter 11) photovoltaic cells or solar cells are yet another example of solid state devices; they are capable of converting sunlight energy into electricity. A solar cell consists of a joined p-type and n-type semiconductor, sandwiched in between two metal contacts that are responsible for conducting electricity into and out of such a device. A schematic drawing of a small section of a panel of solar cell is shown in Figure 12.15 (a).

A small section of this panel of solar cell is enlarged as shown in Figure 12.15 (b) to clearly illustrate the principle of a solar cell: when the sunlight reaches the p-n junction, electrons are freed from the semiconductor at the junction as a result of the **photoelectric effect**. These electrons were once in the valence band but now have gained high enough energy so that they occupy the conduction band. Since these electrons are free to move, they can be easily accelerated by the electric field existing naturally at the p-n junction towards the n-type semiconductor, that is, against the direction of the electric field. Recall this electric field is created as a result of the migration of free electrons down the electron gradient from the n-type semiconductor into the p-type semiconductor when they are joined together; and since the n-type carries a positive potential and p-type carries a negative potential, the direction of the electric field at the junction is from the n-type to the p-type. After being accelerated through the n-type semiconductor, the electrons are collected by the front metal grids to enter the external circuit to do work. This is the electricity. These electrons are then returned by the back metal

Solar panel

Section of solar cell

Figure 12.15 (a) and **(b)** A photovoltaic (solar) cell

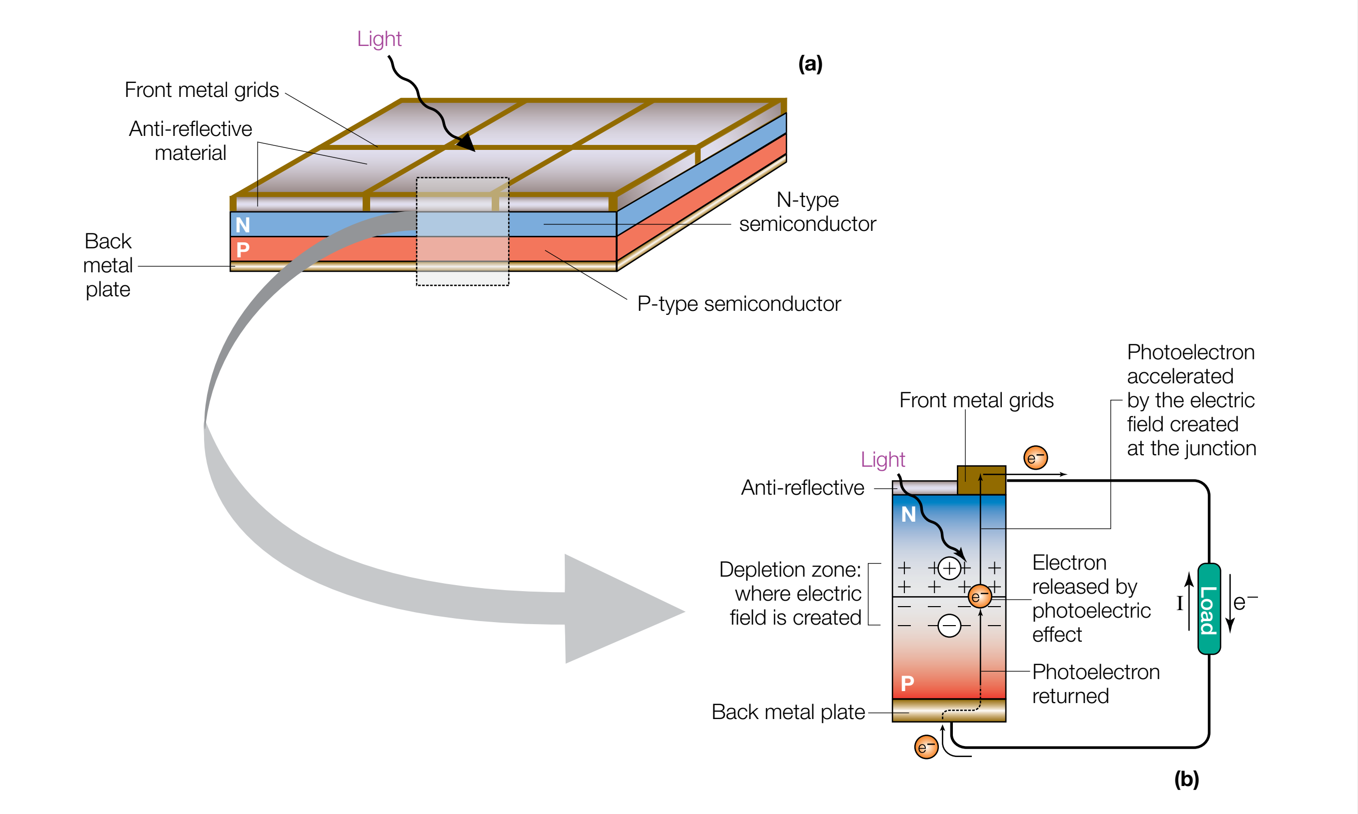


plate to fill up the position of electron deficiency created by the initial photoelectric effect, and the whole process starts again. Also, it is important to note, since the electrons flow from the n-type through the external circuit to the p-type, it follows that the conventional current flows in the opposite direction.

NOTE: Recall that conventional current flows in the opposite direction to the direction of the electron current.

Integrated circuits and microchips: extensions of transistors

SECONDARY SOURCE INVESTIGATION

PFAs

H4

PHYSICS SKILLS

H12.3A, B, D
H12.4F
H13.1A, B, C, E
H14.1E, F, G, H
H14.3D

- ***Identify data sources, gather, process, analyse information and use available evidence to assess the impact of the invention of transistors on society with particular reference to their use in microchips and microprocessors***

Integrated circuits

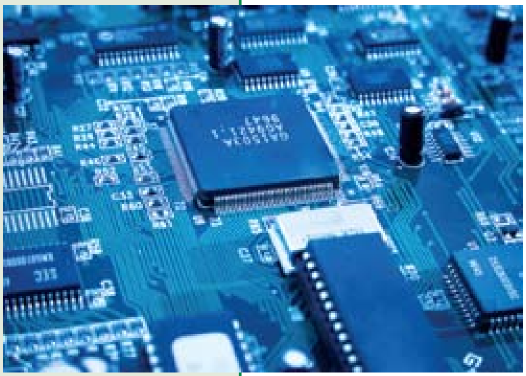
The next application involves a more complicated use of transistors: this is to put transistors together with other electronic devices in a way called **integrated circuits**. Imagine you have some semiconductor devices like transistors and diodes and other devices like capacitors or resistors; normally they can be connected into a circuit via copper wires and this circuit is supposed to carry out a certain operation. In contrast, in integrated circuits, these devices can be built along with their interconnections on a single chip of semiconducting material. Integrated circuits are also called **microchips**.

NOTE: In an integrated circuit, the transistors and diodes do not have the same appearance as those described in the previous section. They do not have a particular morphology in the integrated circuit; the only thing in common is that they carry out the same function.

Definition

An **integrated circuit** is an assembly of electronic devices and their connections, fabricated in a single unit (a single chip), which is designed to carry out specific tasks as they would if they were made individually and connected by wires.

Integrated circuits



ANALOGY: Integrated circuits are almost like normal circuits shrunk into a single unit.

The production of integrated circuits is very complicated and is not required by the HSC course. The basic concept is that integrated circuits are made by processes of lithographic definition, deposition, and etching on common substrates, in most cases silicon. The processes are such that the required electronic devices, for example, transistors and diodes, are made and their interconnections are also established. A typical integrated circuit or microchip is very small, usually 1.5 cm^2, but contains numerous transistors and diodes and their interconnections to carry out a very complex operation.

To highlight the relationship between transistors and integrated circuits, there are basically two types of transistors used in integrated circuits: **the bipolar transistors** and the **Metal-Oxide-Semiconductor-Field-Effect Transistors** (**MOSFET**). The bipolar transistors are basically those shown in the previous section, Figure 12.13 (although in an integrated circuit they will not have any particular morphology), and they are **current-controlled devices**. They are usually used in situations where high-gain amplifiers (to increase current) are needed, this includes applications such as radios and other **analog applications**. MOSFET, on the other hand, is a **voltage-controlled device**. These devices dissipate far less heat compared to the bipolar transistors, due to the small amount of current used. As a consequence, they are more suitable for integration and therefore can perform complicated functions. They are more often used in **digital circuits** to switch electric signals on and off, therefore operating as switches. Ons and offs generate 1s and 0s, which form the basic computer languages. Hence MOSFET are commonly found in digital devices, of course, the main one being computers.

Microprocessors

Definition

A **microprocessor** is a type of microchip that contains enough complicated electronic devices and their connections to perform arithmetic, logic and control operations.

In a sense, a microprocessor is a more complicated form of a microchip. A microprocessor can perform and execute complicated actions. The rise of microprocessors is due to the ability to integrate a large number of electronic devices into a single chip.

NOTE: The more devices integrated onto a single chip, the more powerful the chip.

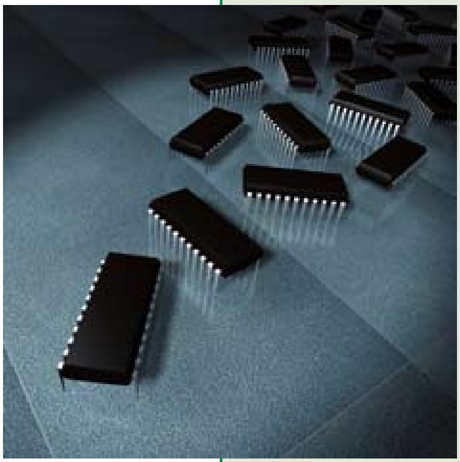
In early days, only a few thousand devices could be integrated onto one chip. Gradually, the number has increased to thousands of thousands, millions and millions of millions, and growing all the time. A typical example of a microprocessor is the central processing unit (CPU) of a computer.

Features and advantages

The number of devices that are able to be integrated onto a single tiny chip has increased from a few to billions. As mentioned before, the more devices that can be integrated onto a single chip, the more powerful is its function. Hence billions of integrated devices will result in the integrated circuits' ability to execute an extremely complicated action, which includes logic and calculations, such as in the case of a CPU and electronic robots.

As the number of semiconductor devices on a single chip increases, the devices are placed closer to each other and so the transmission of signals between the devices becomes more efficient due to the smaller operating distance between the devices. The power dissipation during operation is also reduced due to both the small size of the devices and small distance of separation between them; the small distance results in less resistance to electric signals, and therefore less heat loss. Low heat dissipation means the devices become more power efficient, and the requirement for heat removal is also reduced.

Microprocessor chips



One should note that the manufacturing costs of microchips are proportional to the area of the chip. Hence, being able to integrate more devices onto a single chip, which results in increase in complexity, does *not* actually raise the total cost of production. Effectively, this actually leads to a reduction in cost per unit transistor. Massive production of integrated circuits due to high demand further lowers the cost in the manufacturing of integrated circuits.

Evaluation of the impact of microchips and microprocessors

The invention and development of integrated circuits form the foundation for modern microelectronics and have promoted the development of the so-called information society. The applications of integrated circuits (microchips) are extensive. They are found in many forms of electronic devices, such as medical diagnosis applications, biotechnology, telecommunications, and so forth.

The further development of microprocessors enables computers to be made. As the number of devices integrated in one chip increases, computers are not just becoming more powerful but also smaller. Computers are essential in our day-to-day lives and are used in business and industries and by scientific researchers. The fundamental necessity of computers is self-evident.

The development of microprocessors also leads to the invention of intelligent terminals and robots, which are employed in many areas to carry out work that could replace human labour. This could prove beneficial in areas where heavy labour is involved or dangerous situations are anticipated.

In summary, one can see that from basic semiconductors, diodes and transistors are made. These devices are integrated to make microchips and from there, more and more integration has led to the development of microprocessors. You can easily see the significance of the development of solid state devices such as transistors on society, especially in relation to microchips and microprocessors.

Modelling the behaviour of semiconductors

FIRST-HAND INVESTIGATION

PFAs

H2

PHYSICS SKILLS

H12.3A
H13.1A, B, C, E
H14.1F, G
H14.3A

TR

Risk assessment matrix

- ***Perform an investigation to model the behaviour of semiconductors, including the creation of a hole or positive charge on the atom that has lost the electron and the movement of electrons and holes in opposite directions when an electric field is applied across the semiconductor***

Model the creation of positive holes and the movement of positive holes and electrons in the presence of an external electric field

This concept has already been illustrated in this chapter. To model this, one can use bottle caps and marbles.

Another way of modelling this is to use a Chinese checkers board or a draughts board. Leave a space in one of the holes or squares.

As the playing pieces are moved into the hole, and the pieces behind are moved too, it is observed that pieces (representing electrons) move in one direction, but the space (representing the positive hole) moves in the opposite direction.

Alternative method

An alternative way to show the movement of electrons and holes in the opposite direction is to develop a series of PowerPoint slides with the electron moving one place at a time and then to play the slideshow with each slide showing for one or two seconds.

CHAPTER REVISION QUESTIONS

1. Using the concept of energy bands, explain why a metal is a better conductor of electricity than a semiconductor, which is in turn a better conductor than an insulator.
2. (a) Identify the factors that influence the size of the drift velocity of electrons.
   (b) Two conducting wires are made from the same metal. If one wire has a diameter that is three times bigger than the other wire, compare the drifting velocity of the electrons going through these conductors when equal size currents flow through these conductors.
   (c) Justify why drifting velocity is much slower compared to the actual velocity of the electrons.
3. Compare the change in resistance when a metal conductor and a semiconductor are cooled.
4. Define ‘intrinsic semiconductors’.
5. Besides heating a semiconductor, describe two other ways of decreasing its resistance.
6. The electrical property of a semiconductor can be modified by a process called doping.
   (a) Define the term ‘doping’.
   (b) How are n-type semiconductors produced; what electric property do n-type semiconductors have?
   (c) How are p-type semiconductors produced; what electric property do p-type semiconductors have?
   (d) Describe what will happen when an n-type semiconductor is joined with a p-type semiconductor. In your answer include the charge carried by each type of semiconductor.
   (e) Name the electronic device that is produced by the arrangement described in (d); what is the function of this electronic device?
7. Draw a table to compare and contrast five aspects of solid state devices versus thermionic devices.
8. Give two reasons why most integrated circuits are made from silicon rather than germanium.
9. Quote: A solar cell involves the creation of an electric field, the photoelectric effect and movement of electrons.
   (a) Describe how the electric field is created in a solar cell.
   (b) Describe the occurrence of the photoelectric effect.
   (c) The resulting acceleration of the electrons creates electricity. Briefly explain how this takes place.
10. Evaluate the uses of semiconductors in the development of electronics and computers.

SR

Answers to chapter revision questions