

## 10 A Short Introduction to Electrons in Solids

v1.6 – June 2021

*Prerequisite knowledge required: Newtonian Physics, Electric and Magnetic Fields*

### 10.1 Introduction

It's quite easy to understand the behaviour of electrons in free space: they have a negative charge, so they accelerate towards positive charges and away from negative charges<sup>1</sup>. However, when dealing with just about any practical application of electronics, electrons are not in free space: they are contained within a solid, and this makes a significant difference to how they behave. An understanding of how electrons behave in solids is essential for predicting the behaviour of electronic components including resistors, capacitors and diodes.

This chapter is a short introduction to the sometimes rather curious behaviour of electrons when they are bound within solids.

### 10.2 Electrons and atoms

When discussing electrons in matter, it's useful to start with atoms. Atoms are the fundamental components of matter: all solids are composed of groups of atoms, which themselves are composed of smaller, elementary particles: electrons, protons and neutrons.

A semi-classical<sup>2</sup> model of an atom consists of a central positively-charged nucleus (containing protons and neutrons), with some negatively-charged electrons orbiting around it. You've probably seen images representing this idea before, something like this:

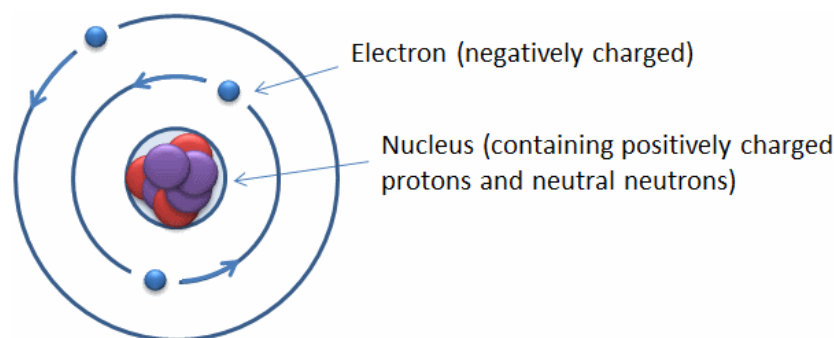


Figure 10.1 A classical picture of an atom

There are a few important things to note about this picture: firstly that there are the same number of protons (positively charged) in the nucleus as electrons (negatively charged) rotating around it, so the whole atom is electrically neutral.

---

<sup>1</sup> I'm neglecting the effect of any magnetic fields here; things get a bit more complicated when there are magnetic fields about as well. See the chapter on Electromagnetism for more details.

<sup>2</sup> In other words one that isn't quite right, since it doesn't fully include the ideas of quantum mechanics, which suggests that electrons aren't really particles with a fixed location in space at all... but that's a whole level of complexity that we don't need at the moment.

Secondly, the electrons are rotating around the nucleus in a similar way to the orbit of planets around the sun (which is why they don't just fall straight towards the nucleus).

Thirdly, and very importantly, there are different orbits that electrons can be in. Some are closer to the nucleus, and these will be experiencing a greater force of attraction to the nucleus, and be orbiting with a higher speed; others are further away, and move more slowly.

At this point we need to consider two rather curious but important results from quantum mechanics:

- Electrons can only be in certain orbits, at certain distances, and with certain energy levels.
- No more than two electrons can be in the same orbit at the same time<sup>3</sup>.

and one slightly less curious result from classical physics:

- Even though the electrons closer to the nucleus are moving faster, they actually have less energy than the electrons further away.

Electrons close to the nucleus do have more kinetic energy (since they are moving faster), but they have less potential energy, and the net result is a reduction of energy in these orbits. Like everything else, if left alone electrons will try and reduce their potential energy, which in this case means they tend to move to lower energy states closer to the nucleus when they can.

### 10.2.1 Sodium lights

Perhaps the most accessible evidence for these quantum effects comes from the spectra of light sources. For example, if you look carefully at the spectrum of the familiar yellow light coming from a sodium-vapour street-lamp, you'll see that most of the light comes out at just one particular frequency.

This is because the photons (elementary particles of light) of different colours have different amounts of energy, given by:

$$E = hf \tag{10.1}$$

where  $E$  is the energy in a photon,  $f$  is the frequency of the photon (corresponding to the colour) and  $h$  is a constant (known as Planck's constant).

At this point it might be useful to introduce a new unit of energy: the electron-Volt (eV). One electron volt is the energy required to move an electron down<sup>4</sup> through a potential difference of one

---

<sup>3</sup> This is a consequence of electrons being *fermions*. A fermion is a particle which cannot be the same quantum state (in other words in the same orbit with the same energy) as another fermion. It is possible to have two electrons in the same orbit, since the quantum state also includes a property called *spin*, and electrons can have either 'up' or 'down' spin, so you can have an 'up' spin electron and a 'down' spin electron in the same orbit, but that's all. (Don't worry if this sounds confusing; at this stage the only important thing to know for now is that electrons can't all drop down to the lowest-energy orbits.)

<sup>4</sup> This isn't a typo: it does require energy to move an electron from a high potential to a lower potential, since electrons have negative charge, and they are attracted to bodies with positive charge; and moving closer to a body with a positive charge means moving to a higher potential. Anything with a negative charge is actively

volt (alternatively, it's the energy released when an electron moves up through a potential difference of one volt).

One electron volt is  $1.6 \times 10^{-19}$  Joules (or around 160 zeptoJoules<sup>5</sup>), since the charge on one electron is  $-1.6 \times 10^{-19}$  Coulombs.

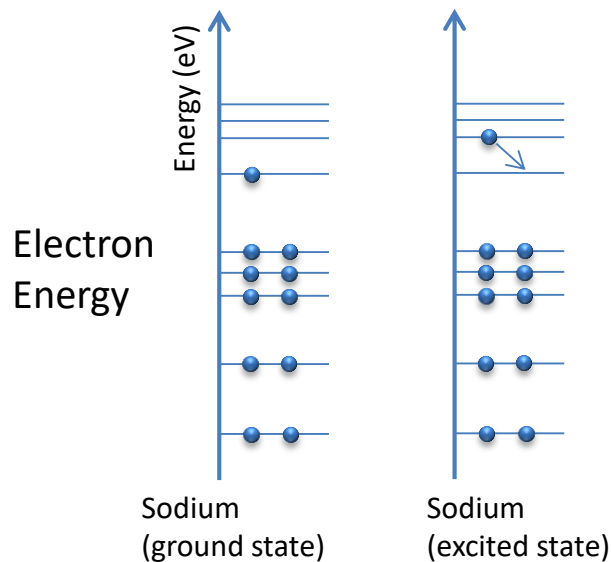


Figure 10.2 Ground and excited sodium atoms

The sodium atom's "ground state" (in other words, when all electrons are at the lowest energy they can have, bearing in mind that only two electrons can share an energy level), is illustrated on the left-hand side of Figure 10.2. However, when energy is provided to the atom (in the case of a sodium light by means of an electrical current), the first excited state it can enter is when the highest energy electron (the one in an energy level by itself; sodium atoms have eleven electrons) moves up to the next highest energy band (see the right-hand side of Figure 10.2).

After a short time, the electron will return to the lower state, and the energy is released in a photon; the energy of the photon being given by the difference between these two energy states.

The energy that is released when an electron moves from a higher orbit to a lower orbit has to go somewhere, and in sodium lights it goes into a photon of light. Since these photons are produced by electrons moving from the same higher orbit to the same lower orbit in different atoms, all the photons have the same amount of energy<sup>6</sup>, and hence all the photons are the same colour.

---

trying to get to higher potential. (A body with negative mass (if any such thing existed) if left free to move would tend to fly upwards, away from the Earth.)

<sup>5</sup> I don't get the chance to use the 'zepto' SI prefix very often. It's comparatively new; it was only adopted in 1991. It means multiply by  $10^{-21}$ .

<sup>6</sup> This is not quite true, since there is some interaction between the atoms, and the atoms in a sodium lamp are a gas moving around, so there are some Doppler effects here as well. However, it's a good approximation, and if you look at the spectrum from a sodium lamp you will indeed see almost all of the light coming out in photons of approximately the same colour.

### 10.3 Metals, semiconductors and insulators

Some elements conduct electricity very well (e.g. copper), some hardly conduct any at all (e.g. sulphur), others do allow some electricity to flow, but they have a significant resistance (e.g. silicon).

As we'll see later, the reason for this difference has to do with what happens to the energy levels in the atoms when they accumulate to form a solid, and the number of electrons each atom has.

#### 10.3.1 The Fermi-Dirac distribution

Since electrons (like everything else) tend to drop from high energy states to lower energy states when given the chance, but you can't have more than two electrons in any given state, electrons tend to fill up the lower energy states first and leave the higher energy states empty (see Figure 10.2 again).

However, this "ground state" (where electrons fill up from the bottom) is only true at a temperature of absolute zero. At all other temperatures there is some random movement of electrons, and this means that some of the electrons at the higher energy levels can randomly acquire a bit more energy and move into one of the vacant states immediately above it, leaving a *hole* (an unfilled energy state where an electron could be) below. The smaller the gap in energy between where the electron is, and the next higher unoccupied state, the more likely it is that the electron will gather enough additional energy to move up.

For example, excited sodium atoms may be in states like these:

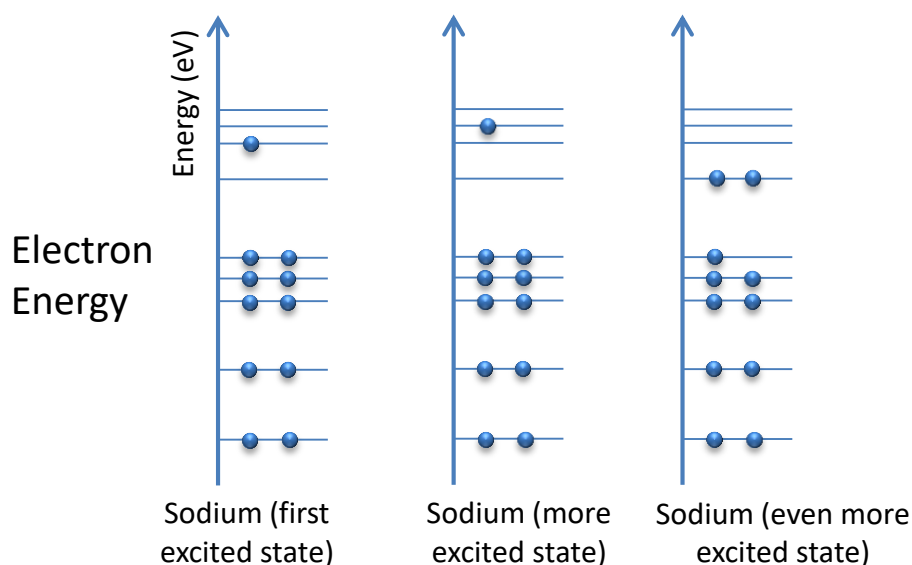


Figure 10.3 Excited states of a sodium atom

There's an equation which relates the probability that a state is occupied to the energy of the state: it's known as the Fermi-Dirac distribution:

$$p(E) = \frac{1}{e^{(E-\eta)/kT} + 1} \quad (10.2)$$

Here  $p(E)$  is the probability that there is an electron in a state with energy  $E$ ,  $T$  is the absolute temperature in Kelvins, and  $k$  is a constant known as the Boltzmann constant.  $\eta$  is a constant known as the *Fermi Level*: it's the energy level where there is a 50% chance that the state is occupied, since:

$$p(\eta) = \frac{1}{e^{(\eta-\eta)/kT} + 1} = \frac{1}{e^0 + 1} = \frac{1}{1+1} = \frac{1}{2} \quad (10.3)$$

Plotting the Fermi-Dirac distribution for different temperatures gives results like that shown in Figure 10.4. (Note I've normalised the plot here to a Fermi level of 1.0; just scale the x-axis to get the plot for any other Fermi level.)

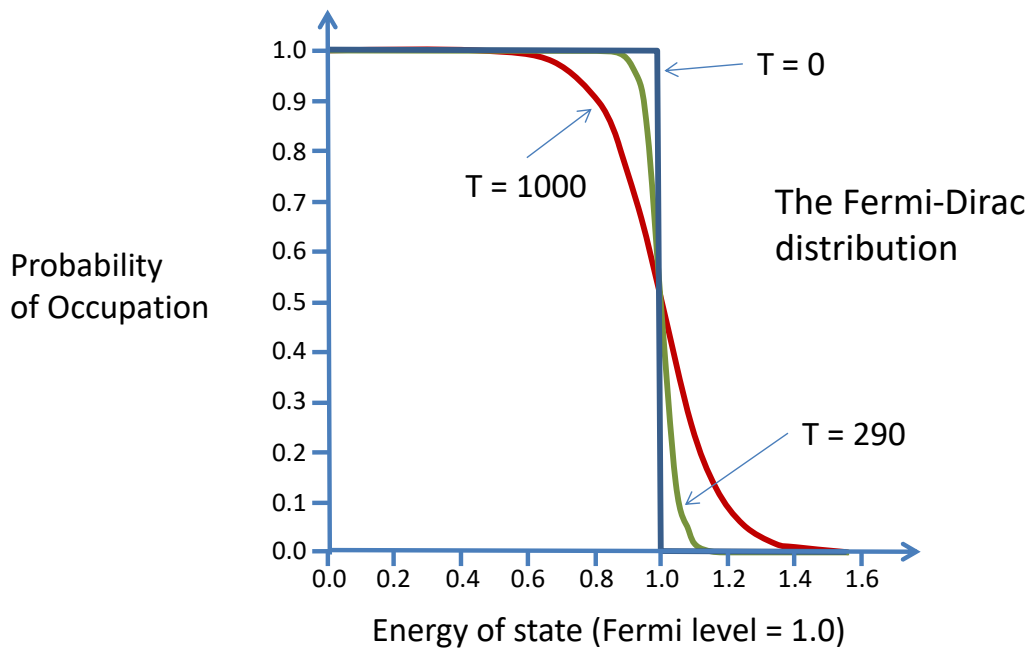


Figure 10.4 The Fermi-Dirac distribution

As you see, at very low temperatures it's unlikely that there are many electrons above the Fermi level: all the atoms will be in their *ground states* with their electrons at their lowest possible energy levels. At a temperature of 290 Kelvin (approximately room temperature) there is a significant probability that any state less than 10% above the Fermi level will be occupied, and that any state less than 10% below the Fermi level will be vacant.

As the temperature rises, there is a higher probability that electrons with considerably more energy than the Fermi level will be found (and that states with considerably less energy than the Fermi level will be unoccupied).

### 10.3.2 Two atoms

If you have two atoms with electrons in them and bring them together into one tiny two-atom block of matter, then the electron orbits will overlap and combine, and by the rules above the resultant system can't have more than two electron states with the same energy levels. So what happens to the lowest possible electron states in the two atoms?

What happens is that when the atoms are brought close together, the permitted energy levels of the two atoms split, and become slightly different: instead of two single atoms both with two electrons in the lowest energy state, there is a two-atom system with four electrons in two states with slightly different energies. The closer the nuclei get, the greater the spreading of the energy states.

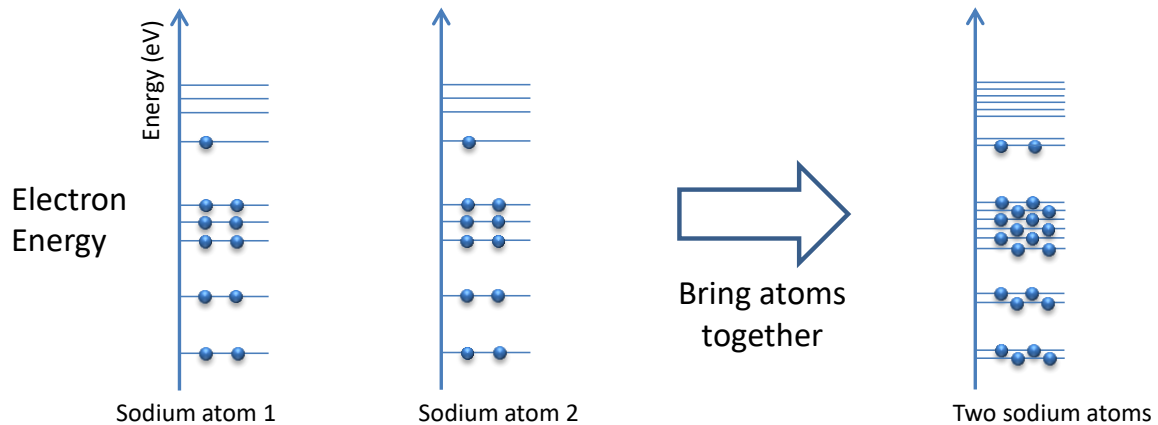


Figure 10.5 Energy states in two sodium atoms (not to scale, energy bands are closer than shown in two-atom system)

What's important (and a bit mind-bending) is that these new states cannot be associated with just one of the two atoms. You have to think of the electrons as moving around a two nuclei system, although some of the energy states are not symmetric: electrons in these states will spend more time on one side of the pair than the other side.

You might also notice that while all the energy bands in the two sodium atom system are full (they have two electrons in them), there is an empty energy band just higher in energy than the highest occupied state. This means that at any temperature above absolute zero, there are likely to be a significant number of electrons in the next highest state. (This is an important observation: it explains why sodium is a good conductor.)

### 10.3.3 Lots of atoms

The same principle applies to a solid with billions of atoms in it: the energy levels split so each of the original energy levels divides up into an (almost) continuous band with billions of closely packed states with slightly different energy levels. Whether a material is a conductor, a semi-conductor or an insulator is a function of how far these bands are apart, and where the Fermi level is in relationship to them.

The electrons in these energy states move around between several atoms in the solid<sup>7</sup>. You could perhaps think about the concept illustrated in Figure 10.6 below: at the highest occupied energy

<sup>7</sup> This is another approximation which is sufficiently inaccurate to warrant a footnote. The different electron states are actually different probability distributions of where the electron is likely to be found in the solid. The electrons with very low energy levels tend to be tightly bound to individual atoms, so that they are always found close to those atoms; electrons with higher energy levels they tend to be more spread out and the probability distributions cover several atoms.

But the idea that electrons are actually small particles in orbit around several adjacent nuclei as suggested in the figure is a huge over-simplification. More accurate representations will have to wait for future modules after you've covered probability theory and some quantum mechanics.

levels the different electron states are associated with different places in the solid, but not with just one atom. To move around the solid, electrons have to move from one possible energy state to another centred around a different place.

But please read footnote 7.

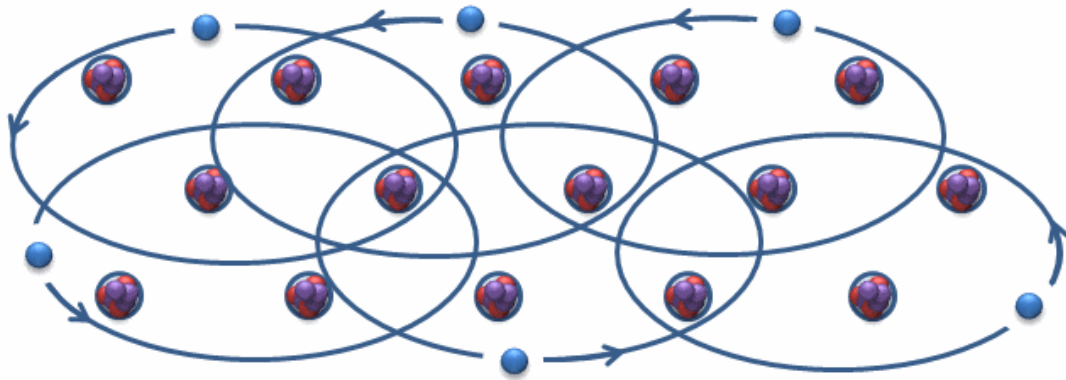


Figure 10.6 Horribly over-simplified conceptual diagram of electron orbits in a solid

Back to the distribution of the possible energy states: billions of individual atoms, and hence billions of possible energy states, when combined together into a solid, form energy bands which can look like one of the three possibilities shown in Figure 10.7.

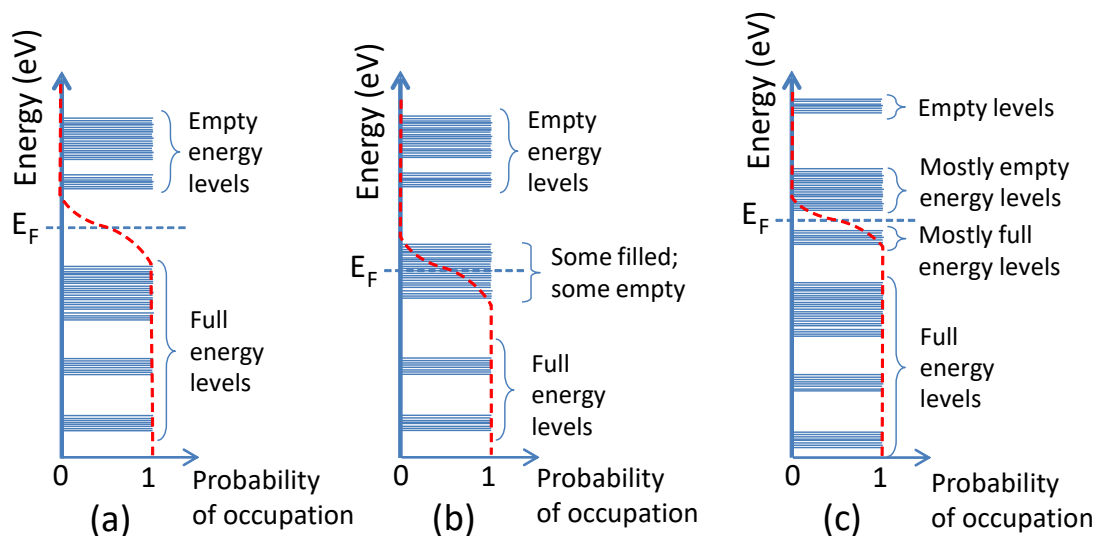


Figure 10.7 Energy levels in three types of solid

The first type of solid has no possible energy states anywhere near the Fermi level. Therefore, the probability of occupancy of the possible states below the Fermi level is approximately one, and the probability of occupancy of the possible states above the Fermi level is approximately zero. This means that no electron can change state: electrons below the Fermi level cannot move because there are no empty states of approximately equal energy for them to move in, and there aren't any electrons above the Fermi level at all. So the electrons cannot move through the material: this is an insulator.

The second solid (b) has a continuous series of possible energy levels all around the Fermi level. At any temperature other than zero degrees, there are going to be a lot of electrons in the energy states just above the Fermi level, and a lot of available states with no electrons in them<sup>8</sup> in the energy states just below. This makes it easy for electrons to move around from one state to another in this band of energy levels, and in moving from one state to another they can move around the solid. This means current can flow: this is a conductor.

The third solid has a smaller gap just around the Fermi level, so at low temperatures it will have full energy states below the Fermi level and empty states above, and hence not conduct electricity. However at room temperature there will be a few electrons in the higher states and a few empty states in the lower levels, so putting an electric field across this solid will result in some electrons moving, but not that many. These solids are called *semiconductors*.

As the temperature rises, more electrons appear in the states above the Fermi level and find it easy to move to other states just above the Fermi level. As a result these materials become significantly more conductive: in other words the resistance of these solids is a strong function of temperature (which is why no-one makes resistors out of semiconductors).

#### 10.4 Conductors, resistance and the Drude model

A solid becomes a conductor if there are a large number of possible electron states around the Fermi level, so that when an external electric field is applied, electrons can easily move into other states, thus allowing the electrons to move around the solid. The result is a flow of charge, otherwise known as a current. How much current flows?

From experience we know that the current is not infinite, and Ohm's law suggests that the current is proportional to the voltage applied. One explanation for this effect was proposed by Paul Drude, who suggested a model to explain the finite resistance of conductors.

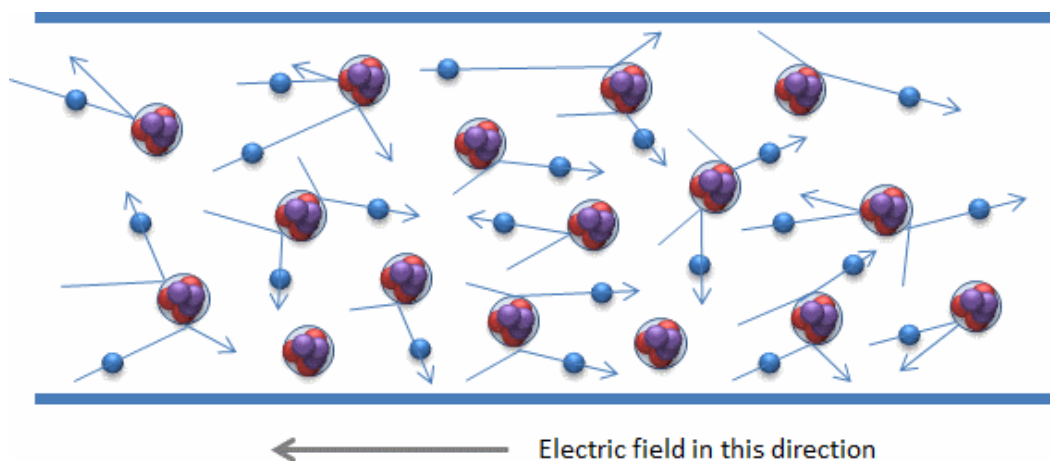


Figure 10.8 Electrons moving in Drude's model of a solid

Drude's model is based on classical (Newtonian) physics, so it's relatively easy to visualise. The theory is that electrons in a conductor accelerate when they experience an electric field (as expected), but once they start moving they are likely to collide with a nucleus at some point, at

<sup>8</sup> These possible energy levels below the Fermi level that are unoccupied are often called *holes*.



which point they will bounce off in a random direction (nuclei are several thousand times heavier than electrons), as shown in Figure 10.8.

Using this model, we can analyse the net current due to an electric field. We need to make a few assumptions:

- The average time between collisions is  $\tau$
- After a collision, electrons bounce off atoms in an entirely random direction with a constant velocity independent of how fast the electron was going before the collision. This implies that the average speed of an electron in the direction of the current just after a collision is zero (it is equally likely to be positive and negative).

From these assumptions we can deduce that:

- In-between the collisions, the electrons experience a force given by  $Eq$  where  $E$  is the external electric field and  $q$  is the charge on an electron, so they will reach (on average) a final speed<sup>9</sup> in the direction of the current of:

$$v = \frac{qE}{m_e} \tau \quad (10.4)$$

- This is the final speed, but the electrons start from a speed of zero, so over a long time, the average speed that the electrons are travelling at is actually  $v / 2$ .
- If the density of mobile electrons in the solid is  $n$  electrons/m<sup>3</sup>, and the cross-sectional area of the conductor is  $A$ , then the average number of electrons which pass any given point along the conductor in unit time is:

$$N = n \times A \times v = nA \frac{qE}{2m_e} \tau \quad (10.5)$$

- The number of electrons passing a point multiplied by the charge per electron will be the amount of charge passing a point in unit time, but this is just the current:

$$I = qN = nA \frac{q^2 E}{2m_e} \tau \quad (10.6)$$

It's not far from here to derive Ohm's law. Just two more steps: firstly you have to consider that for the current flowing through a resistor, for every electron that leaves the resistor by flowing into the wire at one end, another electron arrives into the resistor from the wire at the other end, so the total number of electrons in the resistor is constant. Secondly, the electric field strength is related to the potential by:

---

<sup>9</sup> This comes from the well-known formula  $v = u + at$  where  $u$  is the initial velocity,  $v$  is the final velocity,  $a$  the acceleration and  $t$  the time, combined with Newton's second law of motion which states that the force (in this case  $qE$ ) is equal to the mass times the acceleration, so acceleration is force divided by mass.

$$E = \frac{dV}{dx} \quad (10.7)$$

In a homogeneous material, we could reasonably assume that  $dV/dx$  (the gradient of the curve of voltage against distance) is constant, and therefore that:

$$E = \frac{V}{L} \quad (10.8)$$

where  $V$  is the potential difference across the resistor, and  $L$  is the length of the resistor. Combining equations (10.6) and (10.8) gives:

$$I = nA \frac{q^2}{2m_e L} \tau V \quad (10.9)$$

which is in agreement with Ohm's law: both show that the current is proportional to the potential difference across the resistor. This suggests that the resistance of a resistor should be given by:

$$R = \frac{2}{nA\tau} \frac{m_e L}{q^2} \quad (10.10)$$

and this predicts that the resistance should increase linearly with length of the resistor, and be inversely proportional to the cross-sectional area of the resistor. Both predictions can be tested experimentally, and found to be true (at least at low frequencies and with wires of a reasonable cross-section).

## 10.5 Semiconductors and doping

First another bit of terminology: the group of energy states just below the Fermi level in a semiconductor is known as the *valence band*. The group of energy states just above the Fermi level is known as the *conduction band*.

A pure semiconductor (for example silicon) has a resistance that decreases rapidly as the temperature rises; this is because the number of electrons in the conduction band increases significantly when the temperature rises (see Figure 10.4).

More consistent behaviour can be introduced using a technique called *doping*. The idea is to introduce a few atoms with either one-more or one-fewer electron per atom into the solid. As noted before, you can't really consider the electrons in the higher energy bands as being bound to any one atom; their orbits spread out around the solid crystal. Introducing a small number of atoms with slightly different nuclei doesn't disturb the structure of the solid crystal or the energy of the possible quantum states, it just changes the number of electrons that are trying to occupy them.

While most semi-conductors are in group IV of the periodic table (e.g. silicon), the elements used for doping are typically from group III (with one fewer electron per atom, e.g. boron) or group V (with one more, e.g. phosphorus).

In terms of the Fermi level, what happens is that it just moves up or down slightly, since there are either slightly more or slightly fewer electrons trying to fit into the same number of states. (Very

heavy doping in which so many additional electrons are introduced that the Fermi level moves into the conduction band results in a solid which behaves much more like a conductor, and that's not so interesting; if you want a conductor it's easier to use a metal.)

A piece of *n-type* semiconductor has a few more electrons than the intrinsic (pure) semiconductor would; a piece of *p-type* semiconductor has fewer.

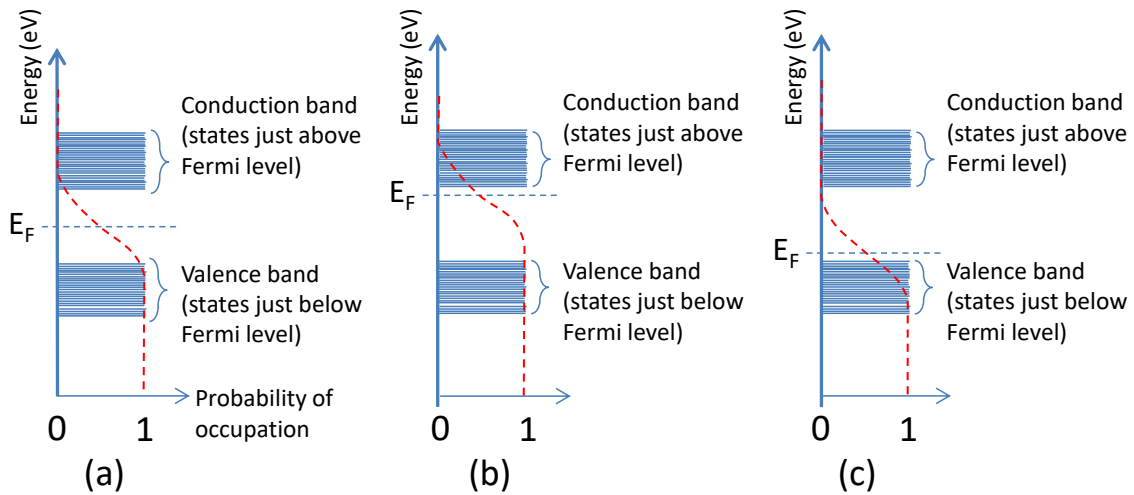


Figure 10.9 Intrinsic, n-type doped and p-type doped semiconductors

First consider a piece of n-type semiconductor, see Figure 10.9(b). The Fermi level rises, and the Fermi-Dirac distribution shows that the lower valence band is full<sup>10</sup>: all possible states in the valence band are occupied by electrons, which means that none of them can move. However, there are an increased number of electrons in the higher conduction band, and lots of unoccupied states of very similar energy: which means these electrons can move quite easily, and can conduct current.

On the other hand, if group III elements are used to effectively remove electrons from the bands, the semiconductor becomes p-type, and the Fermi level falls. With fewer electrons around, the higher energy occupied states in the conduction band empty out, but there are still a few unoccupied states left in the valence band. Some of the electrons in the valence band can now move, but they move more slowly, as they often have to wait for another electron to vacate any state that they want to move into. This is why p-type semiconductors tend to have higher resistance than n-type semiconductors.

### 10.5.1 p-n junctions

Where this really gets interesting is when you take a piece of semiconductor, and dope it so that part of it becomes n-type (more electrons) and another part p-type (fewer electrons).

If you imagine the junction between these two parts of the semiconductor you might think it looks something like this:

<sup>10</sup> Not quite true... there are always a few vacancies except at absolute zero, but for the purposes of this argument we can assume that for all practical purposes the valence band is full.

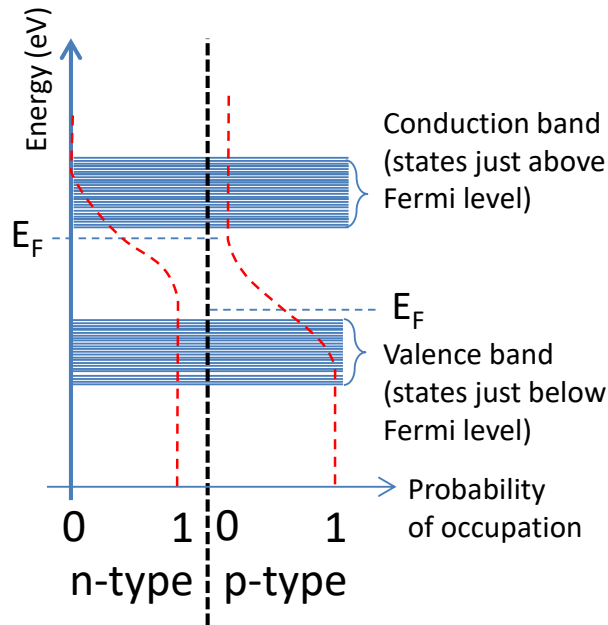


Figure 10.10 What doesn't happen at a p-n junction

However what actually happens around the junction is that some of the electrons in the conduction band of the n-type semiconductor near the boundary drop down into the vacant states in the p-type semiconductor. (They will tend to do this, as it's a chance to move down to a lower energy state, and everything likes to do that.)

These additional electrons start to give the part of the p-type region close to the junction a negative charge, and the fact that electrons have left the n-type region leave it with a net positive charge.

The result is an electric field across the junction, which means that the electric potential of the n-type region is now higher than that of the p-type region<sup>11</sup>, this is known as the *built-in potential*. At some point there is no longer an energy advantage for any more electrons to move from the n-type region to the p-type region, since they have to move against the electric field produced by all the electrons that have already left. At this point the Fermi levels in two regions are the same<sup>12</sup>, and the energy band diagram looks like this:

<sup>11</sup> If you're wondering, yes this does mean that any such device just lying on a lab bench would have a potential difference of approximately the gap between the valence and conduction bands between its terminals. This voltage is however difficult to measure: just putting a voltmeter across the device would not give any reading since no current would be flowing.

<sup>12</sup> They have to be the same, otherwise there would be an energy advantage in electrons to move from one side to the other, and at equilibrium there can be no net motion of electrons since there is no external circuit to flow around in this case.

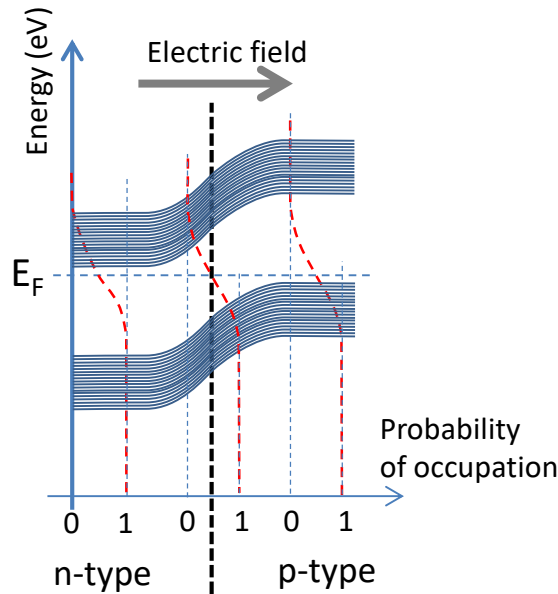


Figure 10.11 Energy states at a p-n junction

Note that at the junction of the p-type and n-type material, the Fermi level is directly between the conduction band and the valence band, which means that there are almost no free electrons in the conduction band, and almost no unoccupied states in the valence band, just like in an undoped semiconductor. This region is known as the *depletion region* since it is depleted of most of the electrons that are easily able to move.

Now consider putting an external electric field across this junction, to try and make a current flow through the junction. First, try this by raising the n-type material to a higher potential than the p-type material by applying an external voltage; the circuit would look something like this:

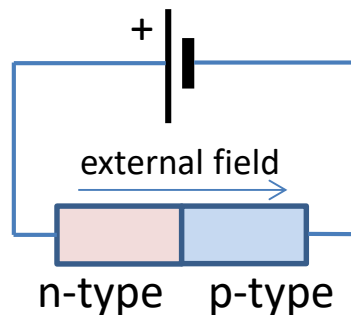


Figure 10.12 A p-n junction with an external reverse bias voltage applied

The block of semiconductor has three regions: two conduct electricity comparatively well, but the third, the depletion region, has a much higher resistance since it has no electrons able to move. Since this is the part of the circuit with the highest resistance, it's the part where most of the applied potential difference is found; there is only a small potential drop across the rest of the n-type and p-type regions. The result is an additional electric field across the depletion region between the two types of semiconductors.

The electrons in the conduction band of the n-type material are now being pulled away from the junction (electrons are negatively charged, and so experience a force in the direction opposite to the electric field). These electrons are comparatively free to move, so many of them do, resulting in the depletion region becoming wider.

The electrons in the valence band of the p-type semiconductor are being pulled towards the junction, and since these electrons are at a higher potential than the bottom of the conduction band in the n-type region on the other side of the depletion region, these electrons have an incentive to move across the depletion region. However this is difficult: the depletion region is now wider than before; and there are very few free states in the valence band in the depletion region to move into. Those empty states here that do exist close to the boundary have lower energy, so it's less likely that electrons will be able to find enough energy to jump up and make it into the conduction band of the n-type material.

As soon as they do get across however, and successfully arrive in the conduction band of the n-type region, they are free to move easily across this region and out into the rest of the circuit. But the problems of getting across the depletion region result in a very small current<sup>13</sup> flowing in the circuit.

The energy-level diagram for this case is illustrated in Figure 10.13. Note that the Fermi levels on the two sides of the boundary are now different: the difference in the potentials of the n-type and p-type material away from the boundary is a direct result of attaching the battery.

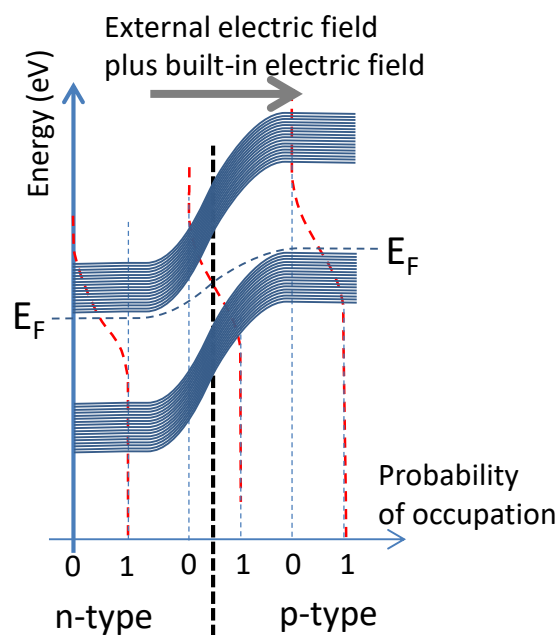


Figure 10.13 Energy bands in a reverse-biased p-n junction

<sup>13</sup> This small current is known as the *saturation current*. It's perhaps interesting to note that the size of this current is not a strong function of the external potential applied: it's just the probability (from the Fermi-Dirac distribution) that there is an electron in the conduction band in the depletion region, and that's a function of the energy gap between the valence and conduction bands, not the rate at which they are sloping.

Next, consider what happens if you raise the potential of the p-type material to a higher potential than the n-type material by applying an external voltage:

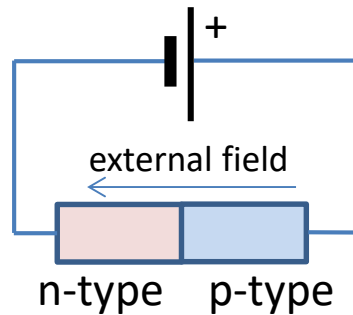


Figure 10.14 A p-n junction with an external forward bias voltage applied

Now, the electrons in the conduction band are being pushed towards the depletion region, and the electrons free to move in the p-type semiconductor are being drawn away from the depletion region and out of the semiconductor.

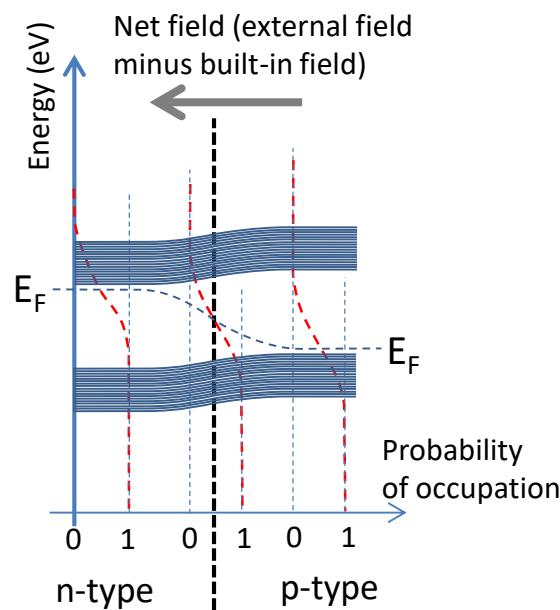


Figure 10.15 Energy bands in a forward-biased p-n junction

The result is a current quite happily flowing: the electrons at the top of the valence band in the p-type semiconductor are being drawn away from the boundary by the applied electric field, this leave more unfilled gaps in the valence band of the p-type material near the boundary. These *holes* are filled by electrons dropping down from the conduction band, and these electrons are easily replaced by electrons from the n-type material, which are experiencing a force towards the boundary.

What we've built is a component which can readily conduct current in one direction, but not the other. They're called *diodes*. They turn out to be very useful.

## 10.6 Insulators and dielectric properties

While insulators don't conduct electricity, it's not true to say their properties have no interest to electronic engineering. One particular property of insulators is extremely interesting: their dielectric constant.

To understand what this is, and why it is so important, it's necessary to go back and think about a solid containing a lot of atoms, all of which have electrons tightly bound to them. The electrons cannot move around the whole solid, but they can move slightly to one side of their atom.

When an external electric field is applied, the negatively charged electrons are pulled in one direction, and the positively charged nuclei in another. The result is a small net dipole as the centre of the electron orbits (where the negative charge is on average) is now offset from the nucleus (where the positive charge is).

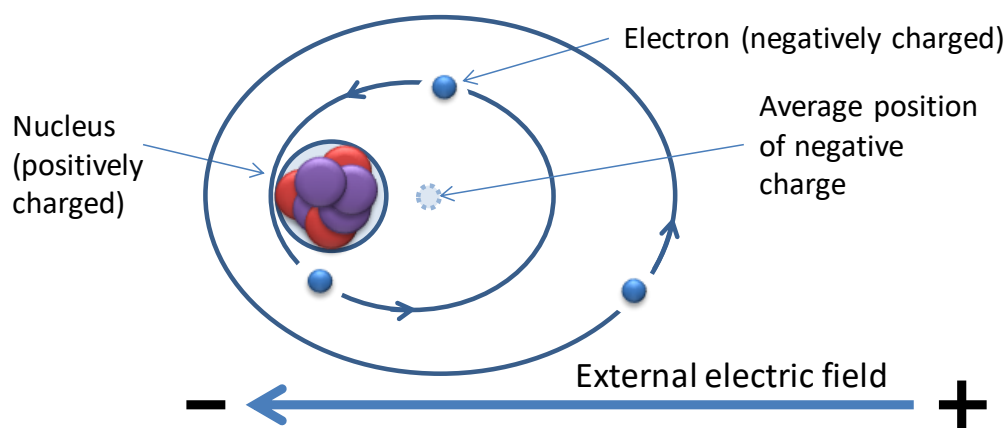


Figure 10.16 Simplified diagram of dipole production in a dielectric material

The total electric field inside the material is then the sum of the external electric field and the electric field due to all of these small dipoles. Since the negative charge in the atoms is pulled towards the positive charges outside, and vice versa, the net effect of the dipoles is to reduce the electric field inside the material.

This is treated mathematically by considering two different fields: the E-field (which is the amount of electric field actually there) and the D-field (which is known as the *displacement field* and can be thought of as proportional to the electric field that would be there if the insulating material was not there at all). The difference between them is known as the *polarisation field* and it's the effect of all of the small dipoles in the material.

We can write:

$$E = \frac{1}{\epsilon_0}(D - P) \quad (10.11)$$

(Why the factor of  $\epsilon_0$ ? It's just convention; this is how the fields are always defined.)

A relative permittivity  $\epsilon_r$  is commonly used in the case where the polarisation is directly proportional to the applied field (which is usually a good approximation), and is defined as:



$$\epsilon_r = 1 + \frac{P}{\epsilon_0 E} \quad (10.12)$$

so that we can write:

$$E = \frac{1}{\epsilon_0 \epsilon_r} D \quad (10.13)$$

### 11.1.1 Dielectrics and capacitors

One of the most important uses of dielectric in electronics is in capacitors. The induced dipoles in the dielectric occur in the opposite direction to the applied field, so the polarisation field they produce acts to reduce the electric field inside the material, which reduces the potential difference between one side of the material and the other for a given amount of charge on the plates (remember for parallel charged plates the potential difference  $V = E * d$  where  $d$  is the distance between the plates).

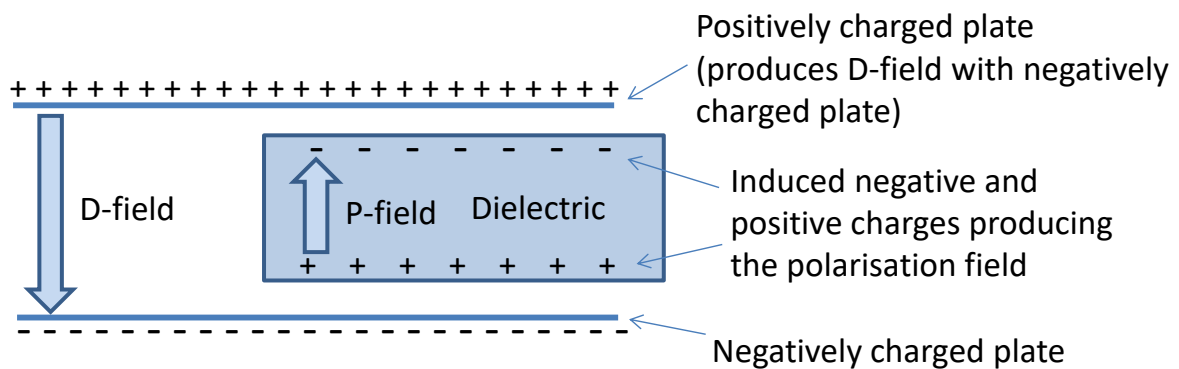


Figure 10.17 Induced charges and polarisation field in a capacitor

So, if you put a dielectric between the plates of a capacitor, with the same amount of charge on the capacitor plates the electric field between the plates will go down, which means a smaller amount of work has to be done to move a charge from one plate to the other, and hence a smaller electric potential difference (voltage) must exist between the plates.

Since for a capacitor:

$$Q = CV \quad (10.14)$$

if the voltage across the plates  $V$  reduces for the same amount of charge  $Q$  on the plates, the capacitance  $C$  must get bigger. This is why just about all capacitors have some sort of dielectric between their plates: it makes the capacitance larger for the same physical size of capacitor.

## 10.7 Summary: the most important things to know

- Electrons in solids occur in discrete energy states. There are some energies for which there are no possible states, so electrons cannot exist in the solid with these energies.
- The Fermi level is the energy where the probability of a state being occupied is 50%.

- All states with energies below the Fermi level are occupied at absolute zero, no states with energies above the Fermi level are occupied at absolute zero. At higher temperatures, the probability of occupancy of states depends on how far they are from the Fermi level.
- Insulators do not conduct electricity because the Fermi level is in a large gap in possible energy states, so there are no electrons in higher energy states that can move, and no spare states for electrons to move into in the lower energy states.
  - Insulators act as dielectrics, and create a polarisation field which reduces the electric field inside the material.
- Conductors can conduct electricity because there are a lot of possible states around the Fermi level so lots of electrons with empty states they can easily move into.
  - With some simplifying assumptions the Drude model can predict Ohm's law
- In semiconductors the Fermi level is in a small gap between mostly full states (the valence band) and mostly empty states (the conduction band).
- Doping is the process of adding a small amount of another material to a semiconductor to move the location of the Fermi level towards one of the bands
  - p-doping involves moving the Fermi level down towards the valence band
  - n-doping involves moving the Fermi level up towards the conduction band
- A p-n junction forms a diode which can pass current in only one direction.