Lecture 3A – Semiconductors

Semiconductor structure. p-*type semiconductor.* n-*type semiconductor. The* p-n *junction. The* p-n *junction characteristic (diode* v-i *characteristic). Diode models. The Hall-effect device. Breakdown diodes. The photodiode. The light emitting diode (LED). The Schottky diode. The varactor diode.*

Semiconductor Structure

The predominant semiconductor material is silicon. Silicon is one of the most abundant elements on Earth, and is always found in compound form in nature (sand is mainly SiO_2). It is purified by chemical means so that the concentration of troublesome impurities is about 1 in a billion. The valence of silicon is 4, like carbon. It is the valence electrons that participate in chemical bonding when the atoms form compounds. Silicon crystallizes in a diamond-like structure, because this minimizes the free energy – each atom has four neighbours, set in a tetrahedral structure.

Silicon is abundant and has a tetrahedral crystal structure

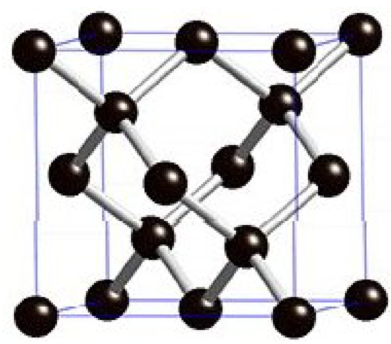


Figure 3A.1

Each of the 4 valence electrons is shared with a neighbour, which is called covalent bonding. It does not involve electric charge transfer between different locations in the lattice.

Pure crystalline silicon will possess the same structure as diamond, but it is nowhere near as hard a substance as diamond. There are two reasons for this. Firstly, the silicon-silicon bond is much weaker than the carbon-carbon bond (silicon is a bigger atom). Secondly, carbon is a significantly smaller-thanaverage atom, and there are vastly more bonds per unit of volume in a diamond than in any other substance.

To simplify things, we can describe the *Si* crystal in a two-dimensional form:

Silicon is held together with covalent bonds

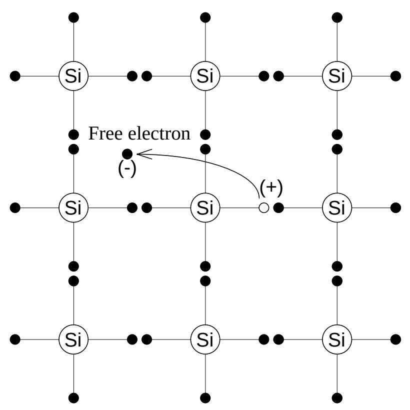


Figure 3A.2

At very low temperatures, pure silicon behaves as an insulator, since a shared electron is bound to its locality and there is no source of extra energy to free itself from its bonds and make itself available for conduction. The extra energy can be obtained from thermal vibrations of the crystal lattice atoms.

Thermal energy can break the covalent bonds, releasing an electron-hole pair

When a valence electron is freed, two charge carriers are created. The first is the electron itself. The second one, called a hole, is the charge located in the area vacated by the electron. That vicinity is left with a *net* positive charge (obviously caused by a silicon nucleus). Any one of the other valence electrons moving nearby can step into the vacated site. This shifts the *net* positive charge – the hole – to a new location. Both the free electron and the hole can therefore move around in the semiconductor crystal.

The conductivity of pure silicon is therefore proportional to the free carrier concentration, and is very small.

p-type Semiconductor

To make devices like diodes and transistors, it is necessary to increase the electron and hole population. This is done by intentionally adding specific impurities in controlled amounts – a process known as doping. Doping defined

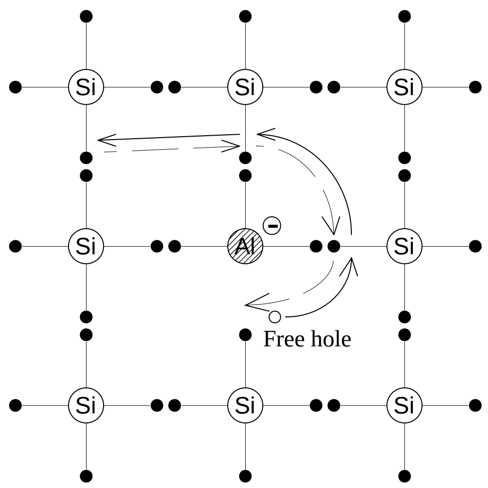


Figure 3A.3

If *Si* is doped with an element with only 3 valence electrons, then at the location of that impurity one of the covalent bonds is missing. The location is electrically neutral. One of the other valence electrons can cross over and complete the missing bond. When this happens a hole is created at the position vacated by that valence electron.

The impurity atom, having accepted an additional electron, is called an acceptor and now has a net negative charge. A semiconductor doped with acceptors is rich in holes, i.e. positive charge carriers, and therefore called *p*-type. *p*-type semiconductor defined

In this case the holes are called the majority carriers, the electrons are called the minority carriers.

n-type Semiconductor

If *Si* is doped with an element with 5 valence electrons, then four of the valence electrons will take part in the covalent bonding with the neighbouring *Si* atoms while the fifth one will be only weakly attached to the impurity atom location. The thermal energy of a semiconductor at room temperature is more than enough to free this electron, making it available for conduction.

n-type semiconductor defined

The impurity atom, having donated an additional electron, is called a donor. The semiconductor in this case is called *n*-type, because it is rich in negative charge carriers. The electrons are the majority carriers and the holes are the minority carriers for this type of semiconductor.

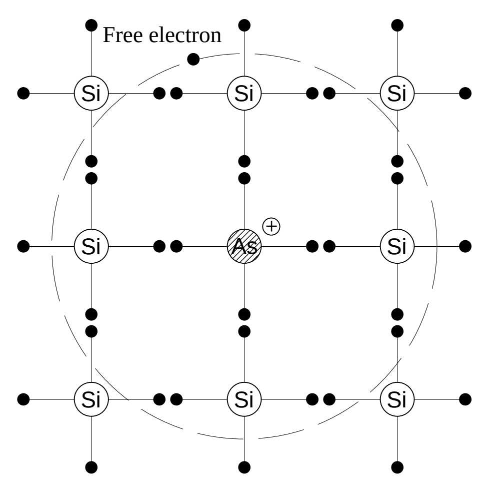


Figure 3A.4

The *p-n* Junction

A *p-n* junction is a location in a semiconductor where the impurity type changes from *p* to *n*, while the lattice structure continues undisturbed. It is the most important region in any semiconductor device.

A *p-n*-junction is formed by doping a pure semiconductor

Initially assume a situation where the *p-n* junction is completely neutral. We can show each lattice site with a neighbouring free charge carrier:

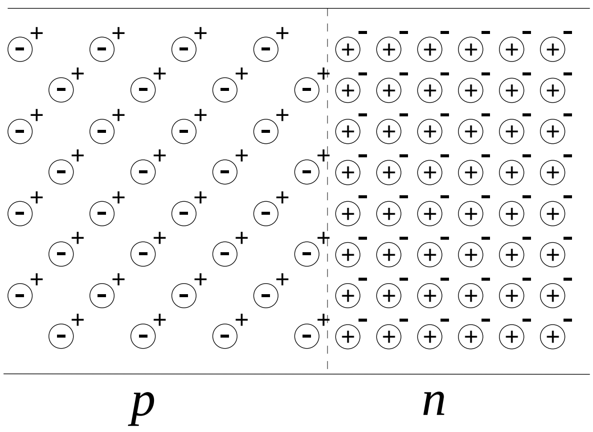


Figure 3A.5

This is a very unnatural state of affairs. Imagine that we have a gas cylinder full of oxygen. We open the valve to release the oxygen into the room. *What happens?* The oxygen and the air in the room mix – a process known as diffusion. Nature wants things to spread out in an even fashion.

A *p-n* junction is subject to diffusion of the majority carriers,

This is what happens with the free, gas-like particles in the *p-n* junction. The holes and electrons will diffuse to try and cover the whole semiconductor in an even fashion. As soon as the holes and electrons move, they "uncover", or leave behind, a charge equal but opposite at that site in the crystal. This uncovered and fixed charge will create an electric field, according to Coulomb's law.

thus creating an electric field in the junction

Where do the holes and electrons go? Since they will try to diffuse across the *p-n* junction, they will find themselves in a region full of opposite charges. If an electron meets up with a hole, then *recombination* takes place. The electron

The process of recombination

"falls into" the hole so that the net charge is zero. They effectively disappear from our diagram.

The resulting electric field that exists at the *p*-*n* junction due to the "uncovered" charge forms a potential barrier. (*Refer back to Lecture 2A*). The field opposes further diffusion of electrons and holes. The field *does* cause drift currents of minority carriers. *Show how this happens*. At equilibrium, the two components of current exactly balance.

A depletion region is formed at the junction

A region exists, on both sides of the junction, in which there is a depletion of mobile carriers, since the field sweeps them away. This region is called the *depletion region*:

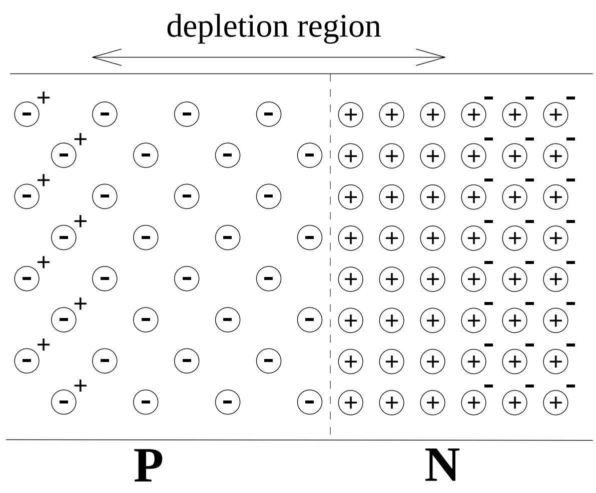


Figure 3A.6

A diode can be a *p*-*n* junction

A *p*-*n* junction forms what is called a diode. Its circuit symbol is:

The diode's circuit symbol

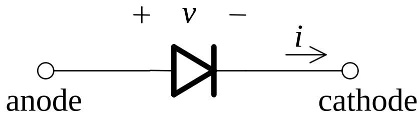


Figure 3A.7

Under open circuit conditions, the inside of the diode will have a potential barrier, as previously discussed:

Conditions inside the diode when open-circuited

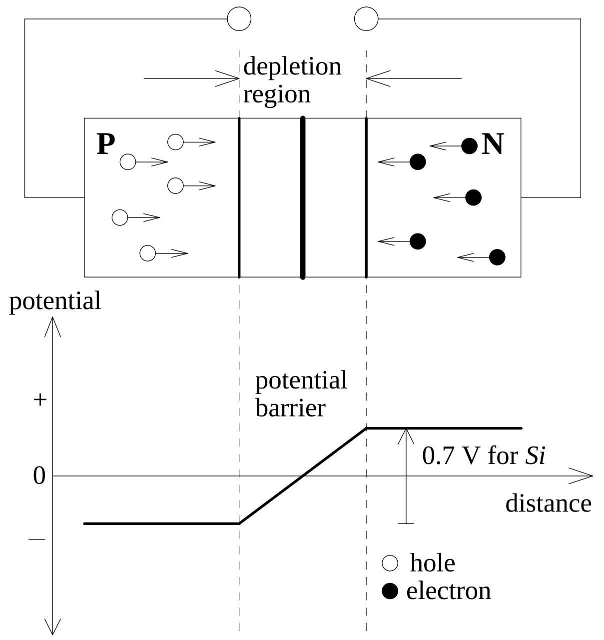


Figure 3A.8

To make an analogy, consider a slope at which we are rolling balls. Each ball will rise up the slope to a level where the gravitational potential energy equals its initial kinetic energy. If a ball has enough energy, it will reach the top of the slope (the "other" side). If a ball does not have enough energy, it will roll back down. You should be able to look at the diode and imagine electrons and holes trying to get over the barrier in this fashion.

Holes and electrons find it difficult to surmount the potential barrier - ball and hill analogy

Is the internal potential (the barrier) available as a voltage source? *No*. For a diode, there are always metal contacts at the diode terminals which form a semiconductor-metal interface. These semiconductor-metal interfaces also create internal potential barriers, which cancel (or balance) the *p-n* junction's potential barrier. The result is an electrically neutral device, as expected.

The internal potential is not a voltage source

Reverse Bias

Conditions inside the diode when reverse biased

When a voltage is applied as shown in Figure 3A.9, the diode is reverse biased. Majority carriers are taken out of the diode. The depletion region widens and the potential barrier increases. The diffusion current is very small. Leakage current in a diode biased in this fashion is due to drift of thermally generated minority carriers.

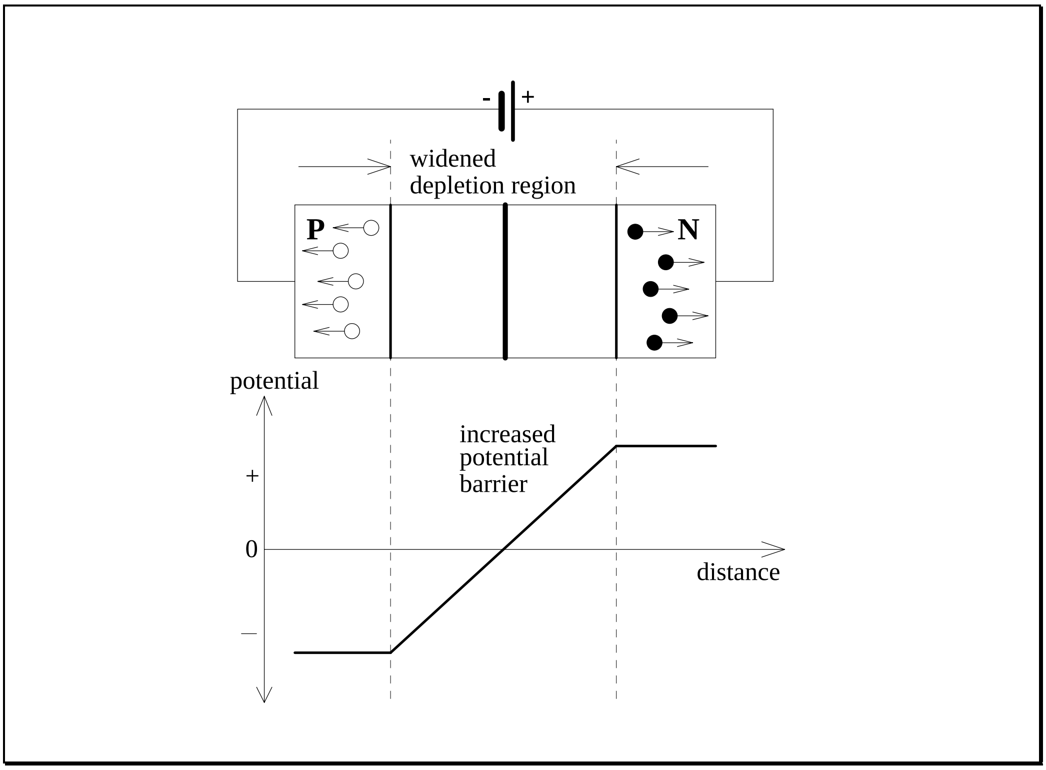


Figure 3A.9

Forward Bias

Majority carriers are supplied to the diode. The depletion layer gets smaller, and more charge carriers are able to overcome the reduced potential barrier. An increase in the diffusion current component results. The drift current remains the same.

Conditions inside the diode when forward biased

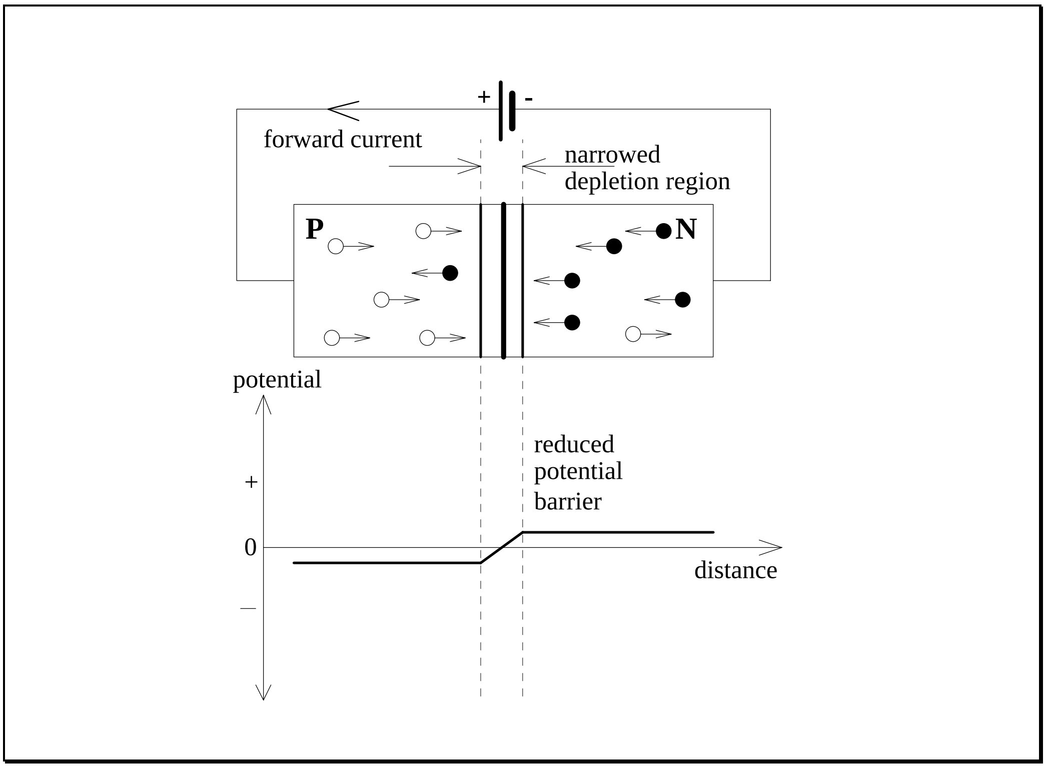


Figure 3A.10

Junction Capacitance

The depletion region of a diode effectively forms a capacitance (there are two conducting regions separated by a high permittivity region). *Show how the junction capacitance is dependent upon the depletion layer width.*

The diode's depletion region is a small capacitor

The *p-n* Junction Characteristic (Diode *v-i* Characteristic)

The diode's terminal electrical characteristics can be obtained using the following circuit:

Obtaining a diode's terminal characteristics

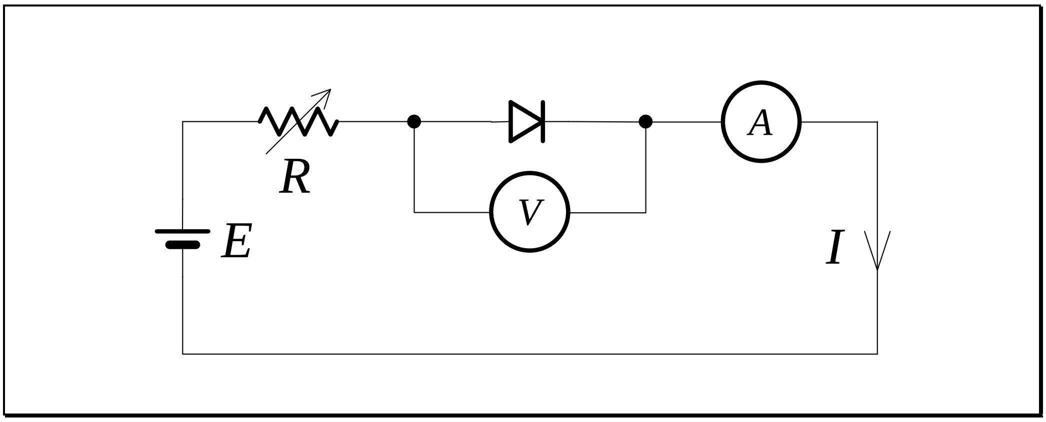


Figure 3A.11

With the battery as shown, we can vary R and measure V and I to obtain the forward-bias characteristic. We could also use a curve tracer to obtain the characteristic. We can reverse the polarity of E to obtain the reverse-bias characteristic. The total characteristic looks like:

A typical characteristic for a silicon diode

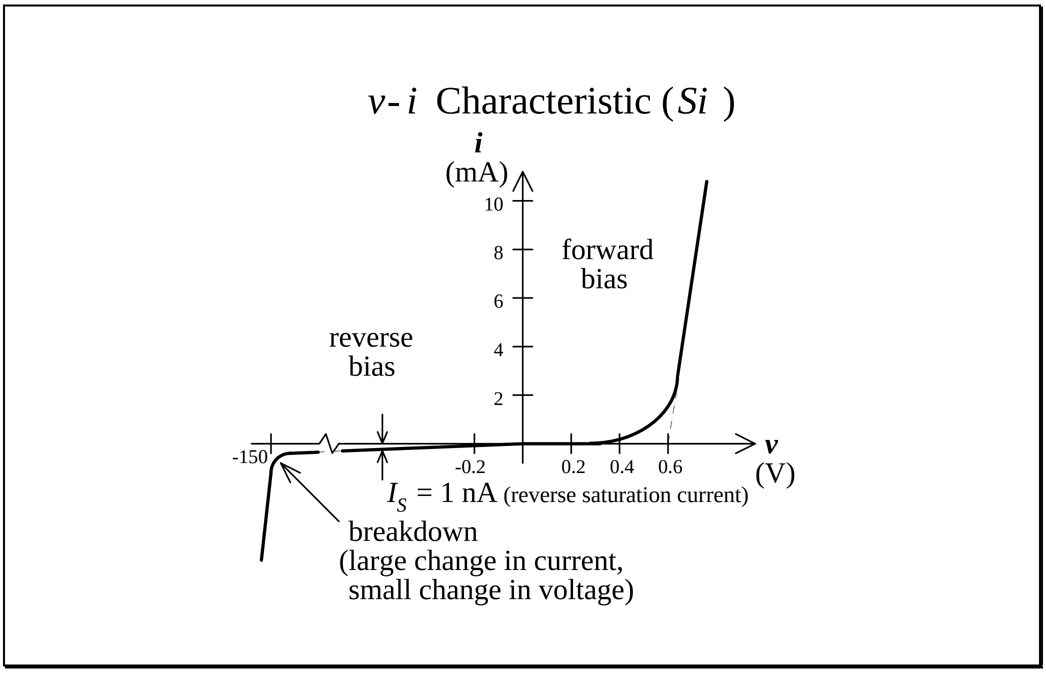


Figure 3A.12

The characteristic can be divided up into two main regions.

Forward Bias

There is not much increase in current until the internal barrier voltage is overcome (approximately 0.6 V in silicon). Then large conduction results. The current is due to drift of the majority carriers and diffusion of the minority carriers. There is a "forward" capacitance associated with the diode. The thin depletion region gives rise to a "junction" capacitance, and the excess concentration of minority carriers on each side of the depletion region (caused by diffusion) gives rise to a "diffusion" capacitance. The diode's forward capacitance is then given by:

When forward biased, the diode conducts

$$C_{fd} = C_j + C_D \tag{3A.1}$$

where: $C_j =$ junction capacitance $(\approx \mu\text{F})$
$C_D =$ diffusion capacitance $(\approx \text{pF})$

Reverse Bias

A small leakage current exists due to minority carriers. Before breakdown, the depletion region is very large so there is a small capacitance:

When reverse biased, the diode does not conduct

$$C_{rd} = C_j \tag{3A.2}$$

where: $C_j =$ junction capacitance $(\approx \text{pF})$

Breakdown

If enough reverse bias is applied, the diode will "break down" and start conducting. It is not a destructive process unless the device cannot dissipate the heat produced in the breakdown process. Breakdown is actually exploited in certain types of diodes (e.g. the Zener diode) because of the near vertical characteristic in this region.

Breakdown occurs eventually for a large enough reverse bias

Diode Models

Why we model the diode

The curve describing the diode's terminal characteristics is non-linear. How can we use this curve to do circuit analysis? We only know how to analyze linear circuits. There is therefore a need for a linear circuit model of the diode.

The concept of modelling

When we model something, we transform it into something else – usually something simpler – which is more amenable to analysis and design using mathematical equations. Modelling mostly involves assumptions and simplifications, and the only requirement of a model is for it to "work" reasonably well. By "work" we mean that it agrees with experimental results to some degree of accuracy.

Models are sometimes only valid under certain operating conditions, as we shall see when modelling the diode.

The Ideal Diode Model

The diode as an ideal (controlled) switch

As a first approximation, we can model the diode as an ideal switch:

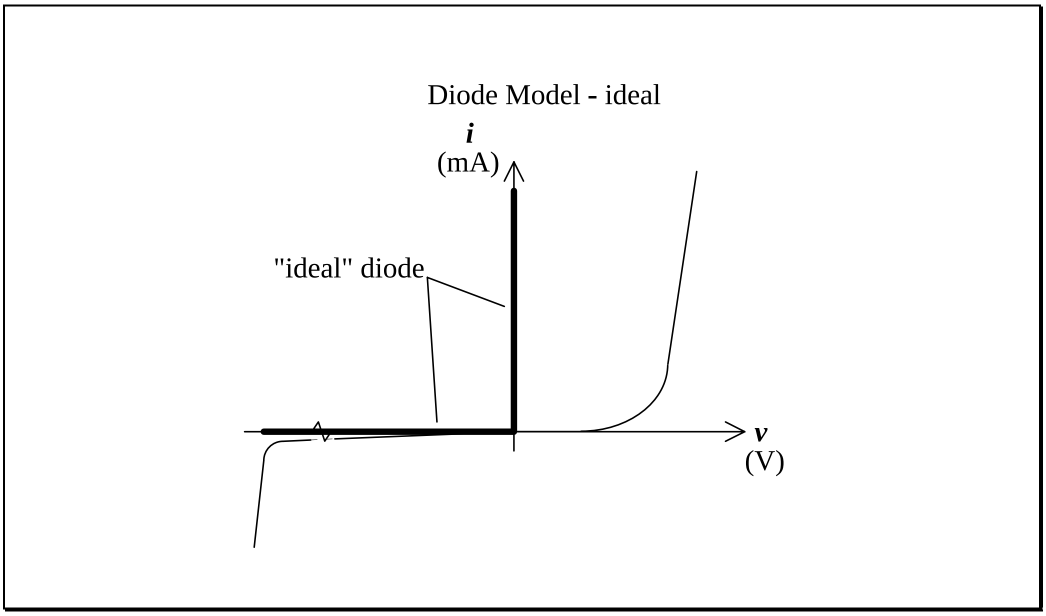


Figure 3A.13

The characteristic in this case is approximated by two straight lines – the vertical representing the "on" state of the diode, and the horizontal representing the "off" state. To determine which of these states the diode is in, we have to determine the conditions imposed upon the diode by an external circuit. This model of the diode is used sometimes where a quick "feel" for a diode circuit is needed. The above model can be represented symbolically as:

The ideal diode model

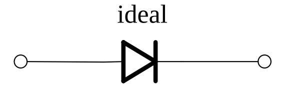


Figure 3A.14

Example

(i) Find the current, I, in the circuit shown below, using the ideal diode model.

(ii) If the battery is reversed, what does the current become?

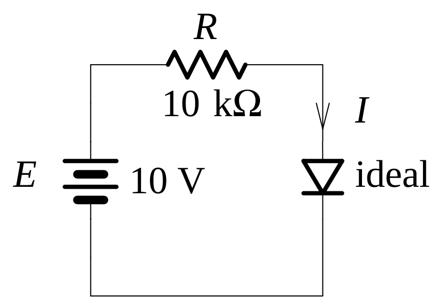
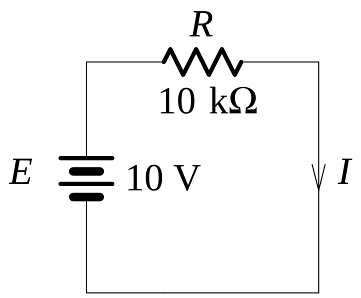


Figure 3A.15

(i) Firstly, we must determine whether the diode is forward biased or reverse biased. In this circuit, the positive side of the battery is connected (via the resistor) to the anode. Therefore, the anode is positive with respect to the cathode, and the diode is *forward biased*. In order to use the ideal diode model, the diode is simply replaced by the ideal diode model (forward bias model), and the simplified circuit is analysed accordingly.

The *equivalent circuit* is shown below, where the diode has now been replaced by a short circuit.



$$I = \frac{E}{R} = \frac{10}{10\text{ k}} = 1\text{ mA}$$

Figure 3A.16

Ohm's Law may be used to determine the current, *I*, as shown:

(ii) If the battery is reversed, the diode becomes *reverse biased*. In this case, the diode is replaced by the ideal diode model for reverse bias. Since the reverse biased ideal diode model is simply an *open circuit*, there is no current, i.e. $I = 0$.

The Constant Voltage Drop Model

A better model is to approximate the forward bias region with a vertical line that passes through some voltage called e_{fd}:

A model that takes into account the forward voltage drop

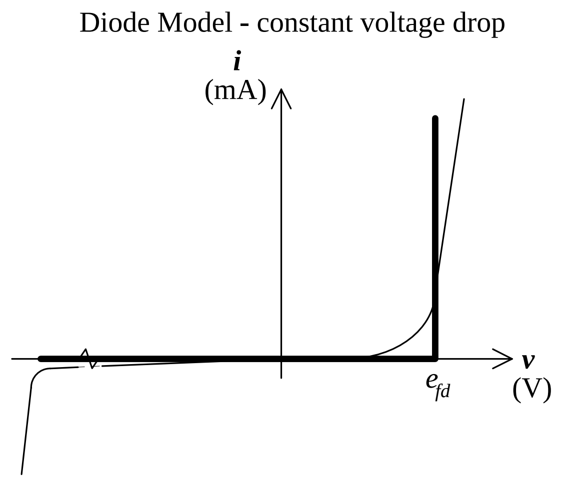


Figure 3A.17

This "constant voltage drop" model is better because it more closely approximates the characteristic in the forward bias region. The "voltage drop" is a model for the barrier voltage in the *p*-*n* junction. The model of the diode in this case is:

The constant voltage drop diode model

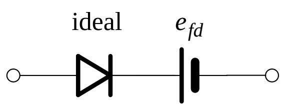


Figure 3A.18

Example

(i) Find the current, I, in the circuit shown below, using the constant voltage drop model of the diode (assume $e_{fd} = 0.7\text{ V}$).

(ii) If the battery is reversed, what does the current become?

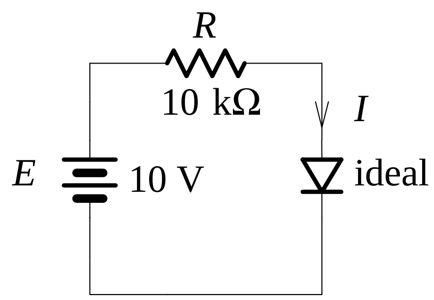
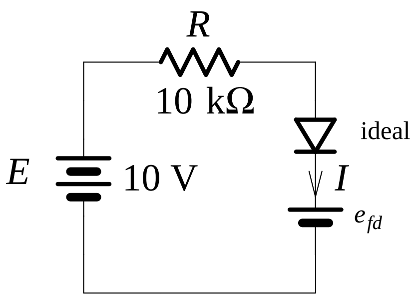


Figure 3A.19

(i) Analysis proceeds in exactly the same manner as the previous example, except that the constant voltage drop diode model is used instead. The diode is again forward biased, and so the equivalent circuit is shown below, along with the calculation for I.



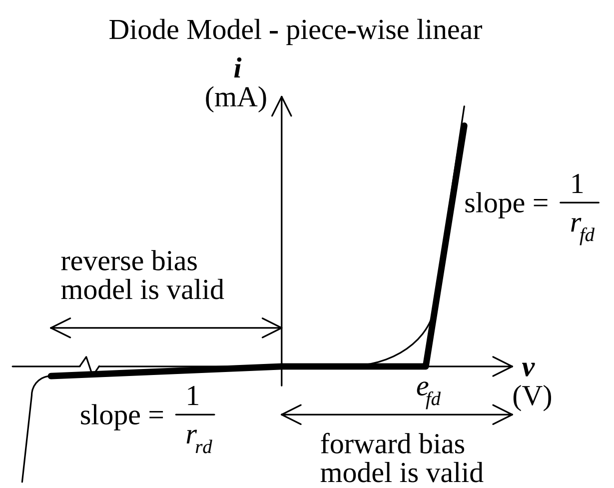
$$I = \frac{E - e_{fd}}{R} = \frac{10 - 0.7}{10\text{k}} = 0.93\text{ mA}$$

Figure 3A.20

(ii) If the battery is reversed, the diode becomes *reverse biased*, resulting in no current, i.e. $I = 0$.

The Piece-Wise Linear Model

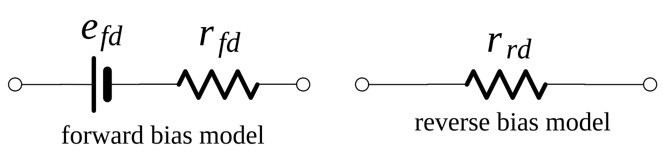
An even better approximation to the diode characteristic is called a "piece-wise" linear model. It is made up of pieces, where each piece is a straight line:



A model that approximates the characteristic by using straight lines

Figure 3A.21

For each section, we use a different diode model (one for the forward bias region and one for the reverse bias region):



The piece-wise linear diode model

Figure 3A.22

Typical values for the resistances are $r_{fd} = 5\,\Omega$ and $r_{rd} > 10^9\,\Omega$.

Notice how we have done away with the ideal diode part of the model. This is because there is a separate equivalent circuit for the forward bias and reverse bias regions, so an ideal diode is not necessary (we apply one equivalent circuit or the other).

You should verify, by using KVL, that these models actually give rise to the straight line characteristics shown in Figure 3A.21.

Example

(i) Find the current, *I*, in the circuit shown below, using the piece-wise linear model of the diode (assume $e_{fd} = 0.7\text{ V}$, $r_{fd} = 5\,\Omega$ and $r_{rd} = \infty$).

(ii) If the battery is reversed, what does the current become?

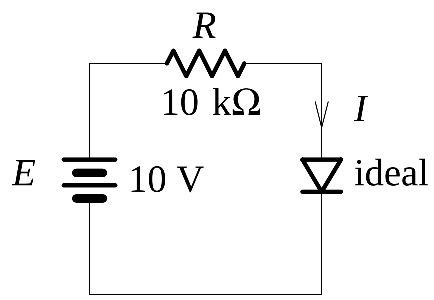
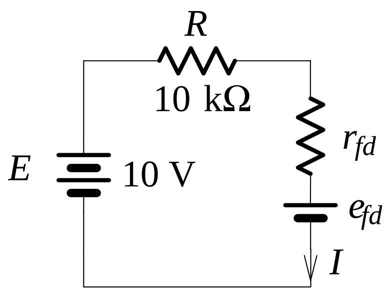


Figure 3A.23

(iii) Analysis proceeds in exactly the same manner as the previous example, except that the piece-wise linear diode model is used instead. The diode is again forward biased, and so the equivalent circuit is shown below, along with the calculation for *I*.



$$I = \frac{E - e_{fd}}{R + r_{fd}} = \frac{10 - 0.7}{10\text{k} + 5} = 0.9295\text{ mA}$$

Figure 3A.24

(iv) If the battery is reversed, the diode becomes *reverse biased*, and the diode is replaced by the piece-wise linear model. Since r_{rd} is infinite, it acts as an open circuit, resulting in no current, i.e. $I = 0$.

The Small Signal Model

Suppose we know the diode voltage and current exactly. Would we still have a need for a linear diode model? *Yes*. Suppose the diode has a DC voltage and current. We may want to examine the behaviour of a circuit when we apply a signal (a small AC voltage) to it. In this case we are interested in small excursions of the voltage and current about some "DC operating point" of the diode. The best model in this instance is the following (the forward bias region is used as an example, but the method applies anywhere):

A model that approximates the characteristic by a tangent at a DC operating point

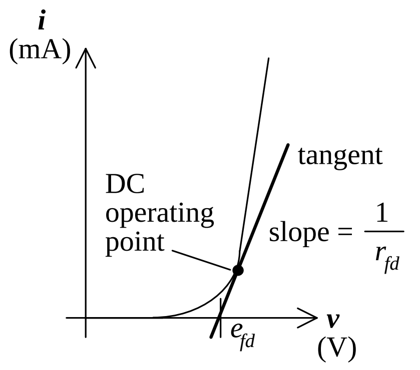


Figure 3A.25

We approximate the curved characteristic by the *tangent* that passes through the operating point. It is only valid for small variations in voltage or current. This is called the *small signal approximation*. A straight line is a good approximation to a curve if we don't venture too far.

A first look at the small signal approximation

The model we get in this case is exactly the same as in Figure 3A.22 except the values of e_{fd} and r_{fd} are different for each DC operating point.

Finally, to complete all our models, we can add a capacitance in parallel to model the forward and reverse capacitance described previously. We will not in general include the capacitance because it only becomes important at very high frequencies.

The capacitance of the diode is added last, but only used at high frequencies

The piece-wise linear model for a diode that includes capacitance

For example, the piece-wise linear models become:

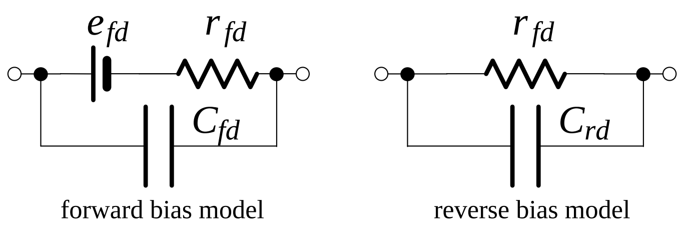


Figure 3A.26

The Hall-effect Device

A Hall-effect device uses semi-conductors and the Lorentz Force Law

Suppose we have a doped semiconductor that has a current passing through it. Now imagine subjecting the semiconductor to a perpendicular magnetic field. *Show that the Lorentz Force Law says that the charge carriers will experience a force downwards regardless of their type (hole or electron)*:

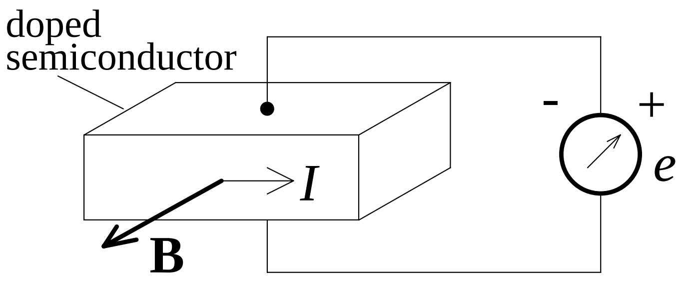


Figure 3A.27

and can be used to measure magnetic fields

An electric field is set up across the semiconductor. *Show the direction of this field for both types of charge carrier.* Its direction (and hence the potential difference across the semiconductor) is dependent upon the charge carrier. We can use this to measure the strength of magnetic fields or to determine the type of semiconductor (*p* or *n*).

Breakdown Diodes

Some diodes are designed to operate in the breakdown region. It is usually a sharper transition than the forward bias characteristic, and the breakdown voltage is higher than the forward conduction voltage. There are two main types of breakdown.

Some diodes are designed to operate in the breakdown region

Zener Breakdown

The electric field in the depletion layer of a *p*-*n* junction becomes so large that it rips covalent bonds apart, generating holes and electrons. The electrons will be accelerated into the *n*-type material and the holes into the *p*-type material. This constitutes a reverse current. Once the breakdown starts, large numbers of carriers can be produced with negligible increase in the junction voltage.

Zener breakdown is caused by a large internal electric field

Avalanche Breakdown

If the minority carriers are swept across the depletion region of a *p*-*n* junction too fast, they can break the covalent bonds of atoms that they hit. New electron-hole pairs are generated, which may acquire sufficient energy to repeat the process. An avalanche starts.

Avalanche breakdown is caused by electrons with a large kinetic energy

The Photodiode

In a photodiode, the *p*-*n* junction is very close to the surface of the crystal. The Ohmic contact with the surface material is so thin, it is transparent to light. Incident light (photons) can generate electron-hole pairs in the depletion layer (a process called photoionisation).

A photodiode is controlled by light

The Light Emitting Diode (LED)

When a light-emitting diode is forward biased, electrons are able to recombine with holes within the device, releasing energy in the form of light (photons). The color of the light corresponds to the energy of the photons emitted, which is determined by the “energy gap” of the semiconductor. LEDs present many advantages over incandescent and compact fluorescent light sources including lower energy consumption, longer lifetime, improved robustness, smaller size,

An LED emits photons when forward biased faster switching, and greater durability and reliability. At the moment LEDs powerful enough for room lighting are relatively expensive and require more precise current and heat management than compact fluorescent lamp sources of comparable output.

LEDs are used in diverse applications. The compact size of LEDs has allowed new text and video displays and sensors to be developed, while their high switching rates are useful in advanced communications technology. Infrared LEDs are also used in the remote control units of many commercial products including televisions, DVD players, and other domestic appliances

The Schottky Diode

A Schottky diode is a metal-semiconductor junction

A Schottky diode is the result of a metal-semiconductor junction. The Schottky diode is a much faster device than the general purpose silicon diode. There are three main reasons for this: 1) the junction used is a metal-semiconductor junction, which has less capacitance than a *p*-*n* junction, 2) often the semiconductor used is gallium arsenide (GaAs) because electron mobility is much higher, and 3) the device size is made extremely small. The result is a device that finds applications in high speed switching and decoupling operations.

The Varactor Diode

This device is also known as a variable capacitance diode. It has a relatively large capacitance, brought about by a large junction area and narrow depletion region. The applied reverse voltage changes the length of the depletion region, which changes the capacitance. Thus, the device can be used in applications that rely on a voltage controlled capacitance. Applications include electronic tuning circuits used in communication circuits, and electronic filters.

Summary

- Semiconductors are crystals which are insulators at low temperatures. Thermal energy in a pure (intrinsic) semiconductor can create an electron-hole pair, allowing conduction.

- Specific impurities are added to semiconductors in controlled amounts (a process called doping) to increase the number of charge carriers. A semiconductor doped with acceptors is rich in holes and is therefore called *p*-type. A semiconductor doped with donors is rich in electrons and is therefore called *n*-type.

- A *p*-*n* junction is formed in a semiconductor where the impurity type changes from *p* to *n*, while the lattice structure continues undisturbed. The *p*-*n* junction creates a depletion region which forms a potential barrier. The potential barrier can be increased or decreased with the application of an external voltage. The external voltage can therefore be used to change the conductivity of the semiconductor.

- A *p*-*n* junction forms a circuit element known as a diode. A diode's characteristic is broken down into two regions – the forward bias region and the reverse bias region. In the reverse bias region, the diode is effectively an open circuit. In the forward bias region, and once the internal potential barrier is overcome, the diode can conduct.

- There are numerous circuit models for the diode. In general, the choice of diode model to be used is based on three main issues: the available diode data, the accuracy required, and the relative complexity of analysis involved. The most commonly used model is the constant voltage drop model.

References

Sedra, A. and Smith, K.: *Microelectronic Circuits*, Saunders College Publishing, New York, 1991.

Bar-Lev, Adir: *Semiconductors and Electronic Devices*, Prentice-Hall International, Inc., Englewood Cliffs, New Jersey, 1984.

Neudeck, G. and Pierret, R. (Eds.): *Modular Series on Solid State Devices, Volume II - The PN Junction Diode*, Addison-Wesley Publishing Company, Inc., USA, 1983.

Millman, J. and Halkias, C.: *Integrated Electronics: Analog and Digital Circuits and Systems*, McGraw-Hill Kogakusha, Ltd., Tokyo, 1972.