XMUT204 Electronic Design

Note 2a: Diodes and PN Junctions

1. Introduction to Diodes

We will now take our basic knowledge of semiconductor materials from the previous chapter and use it to study the structure and function of different electronic devices. We will first look at the diodes, which can be defined as a two-terminal device with asymmetric (non-linear) conductance.

The circuit symbol for the diode is shown below, where the line/bar in the symbol indicates the cathode lead, i.e. the more negative lead for normal operation.

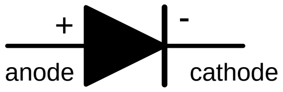


Figure 1: The circuit symbol for a diode.

A diode should have a very low resistance to current in the one direction and a very high resistance to current in the opposite direction, acting as an electrical valve with a current – voltage that can be approximated as shown in the following figure.

In the given figure, three different types of electrical connection can be distinguished:

- *Zero bias*: The diode is not connected to a voltage source so that V=0 and as expected we should not have any current flow (I=0) in this condition.

- *Forward bias:* The diode is connected with its cathode to the more negative side of the voltage source and its anode to the more positive side of the voltage source. In this condition, the diode offers a very low resistance to current flow and high currents flow through the diode when relatively small voltages are applied.

- *Reverse bias:* The diode is now reversed so that its cathode is connected to the more positive side of the voltage source and its anode to the more negative side of the voltage source. In this condition, the diode offers a very high resistance to current flow (infinite resistance in the ideal case) and zero (or very little) current flows for any applied voltage.

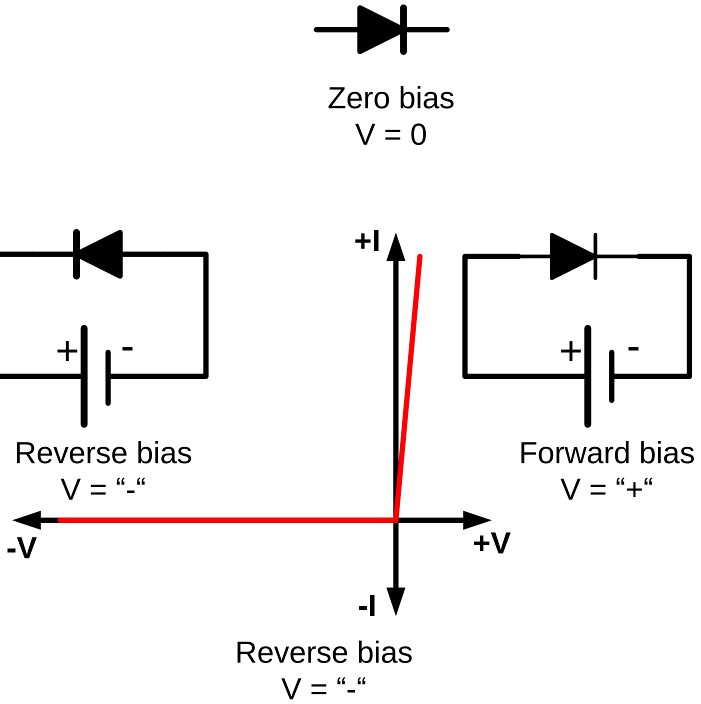


Figure 2: The current-voltage (I-V) characteristics of an ideal diode.

2.1. Early Versions of the Diode

It is clear that the diode as described above will be an essential circuit element when we want to rectify current and such elements has been an essential component of electronic circuits since the start of the electronics industry.

In the early years, this role was filled by vacuum tube diodes, consisting a metal anode and cathode in an evacuated glass enclosure as shown in the following figure. The cathode (negative terminal under forward bias) could be heated by passing a current through it and at sufficiently high temperatures, electrons would be emitted from the cathode and be attracted to the positive anode, allowing a current to flow.

If the bias was reversed by making the cathode positive, the emitted electrons will simply stay on the cathode and the current will be zero. This vacuum diode suffered from high power consumption (heating of cathode required), a fragile glass envelope and a physically large structure and has been replaced by semiconductor diodes in modern electronics.

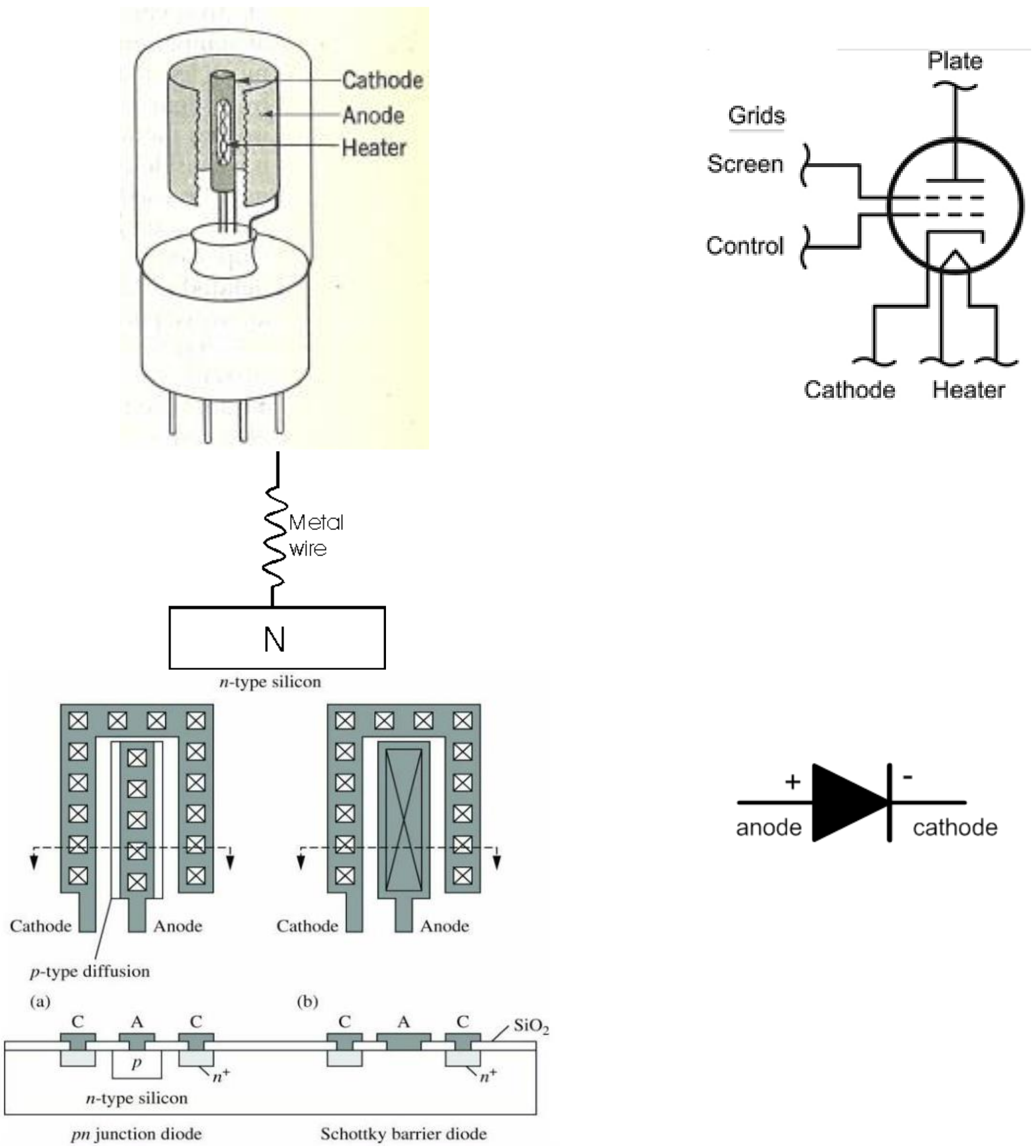


Figure 3: A vacuum tube diode (top), an early point contact diode (cat's whisker) and a contemporary diode (bottom).

The early development of radio was also dependant on the development of the so called point contact diodes. These diodes consisted of a piece of semiconductor (typically PbS) into which a metal wire (called a cat's whisker) was pressed. The semiconductor-metal junction produced a crude rectifying junction

Today, the most common method of forming a semiconductor diode is by the junction between an n-type and a p-type semiconductor and we will look at this junction in more detail. Such a diode structure is not only useful for current rectification, but also forms the basis for light emitting diodes (LEDs), semiconductor lasers, photodetectors and solar cells.

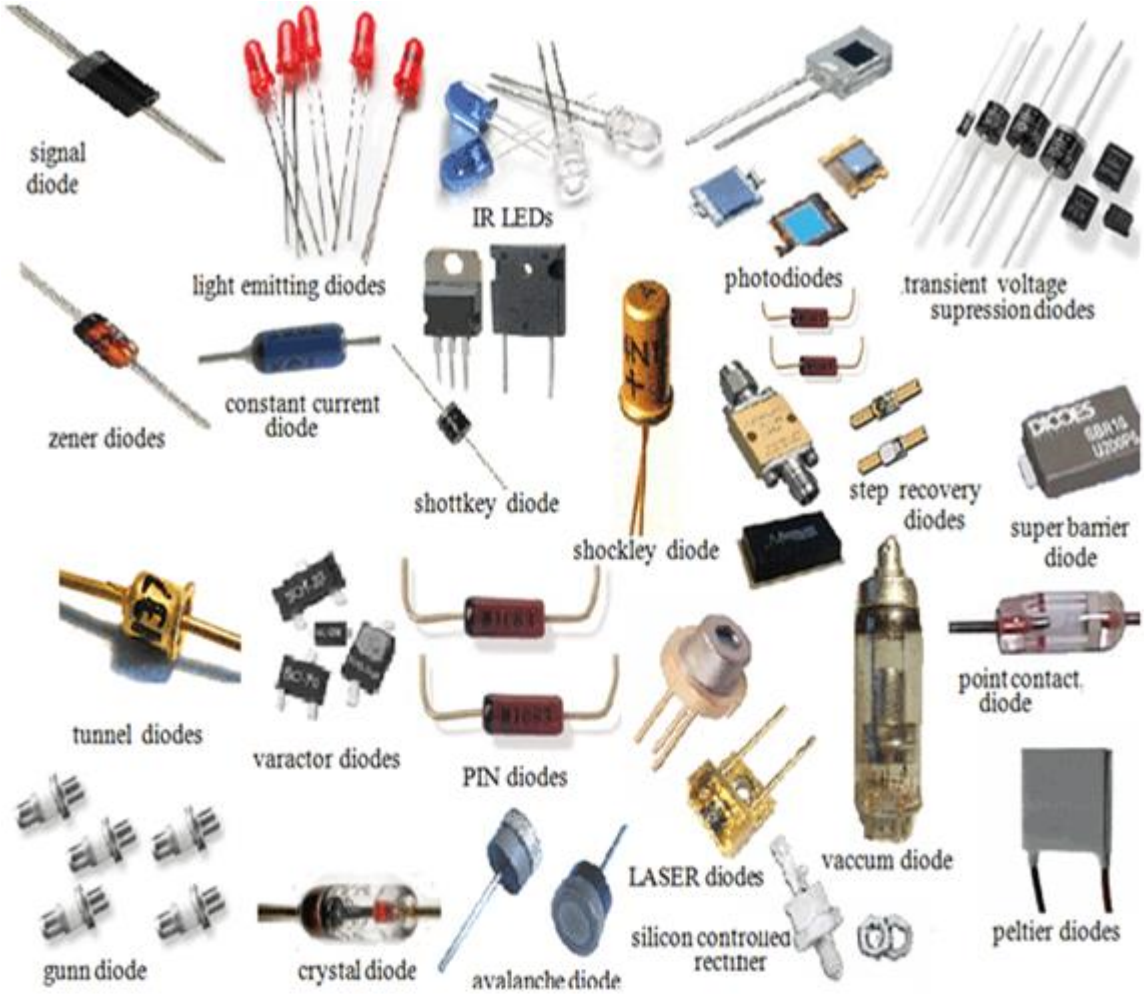


Figure 4: Various packages of diodes

The diagram above illustrates some of types of diode packages which are available in the market and the following diagram shows the symbols of common diodes.

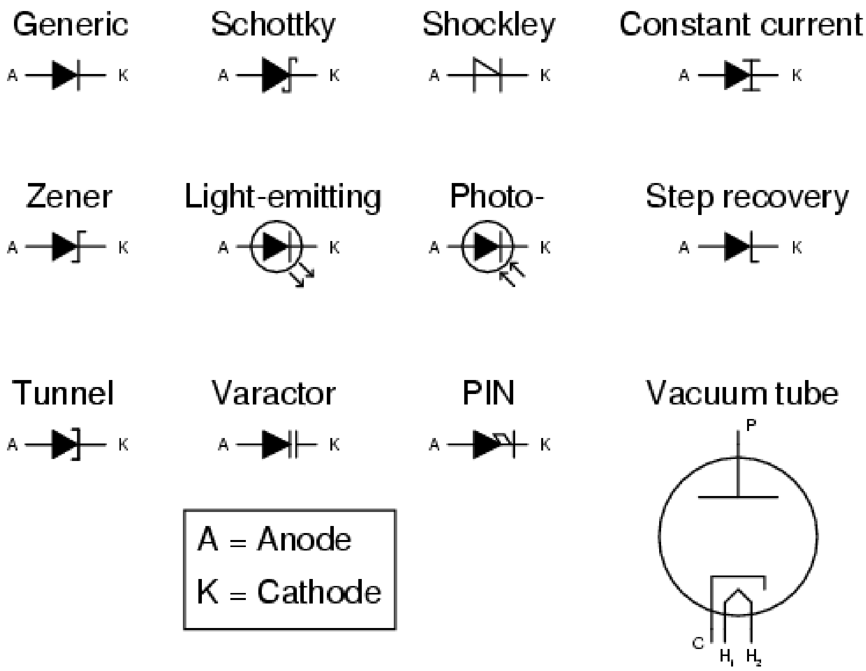


Figure 5: Symbols of common diodes

The following diagram shows some of the extended symbols of diodes.

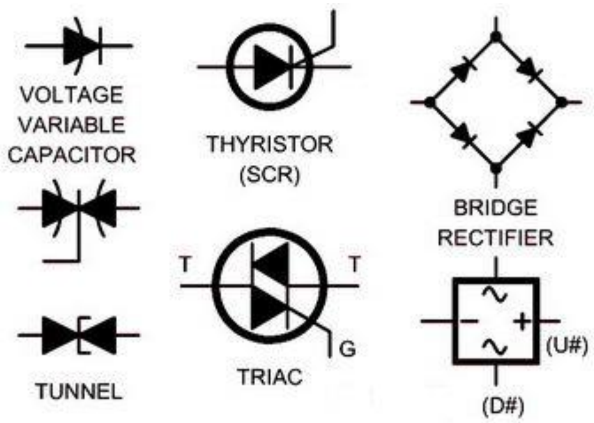


Figure 6: Extended symbols of diodes

In the next section, we will study the structure of a p-n junction and see how this structure enables non-linear electrical conduction behaviour.

2.2. The PN Junction – Zero Bias Behaviour

Assume a slab of n-type semiconductor that is brought in contact with a slab of p-type material, forming a *metallurgical junction* as shown in the following figure. This contact is normally achieved by the processes of diffusion or ion implantation during the fabrication process.

As was shown in previous chapter, the n-type material will contain a high concentration of free electrons, while the p-type material will contain a high concentration of mobile holes as the majority carriers. A concentration gradient will thus exist for electrons from the n-type (high concentration of electrons) to the p-type material (low concentration) and similarly a concentration gradient will exist for holes from the p-type to the n-type material.

Both these mobile charge carriers will thus start to flow (diffuse) along these concentration gradients, with electrons flowing into the p-type material and holes flowing into the n-type material. This movement of carriers is then called *diffusion current*.

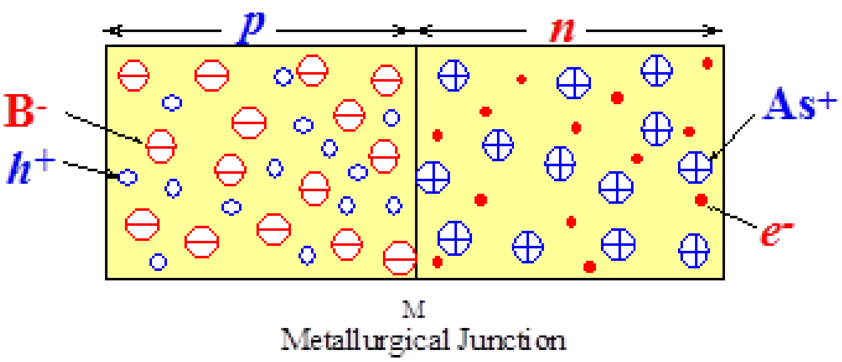


Figure 7: A piece of n-type semiconductor is placed in contact with a piece of p-type semiconductor, forming a metallurgical junction. This picture presents a snapshot immediately after contact, before any flow of charge has taken place.

Holes diffusing and entering the n-side will recombine with electrons on that side near the junction, while electrons diffusing into the p-side will recombine with holes near the junction, with the recombination process leading to the removal of these free carriers from the material.

The effect of this electron-hole recombination is the formation of a region around the metallurgical junction that becomes depleted of mobile charge carriers – the so-called *depletion layer*. This region can also be called *the space charge layer*, as the depletion of carriers in this region will expose the stationary ions that form part of the crystal lattice – negative ions (B^-) on the p-side and positive ions (As^+) on the n-side. A space charge layer thus forms around the metallurgical junction with fixed positive charge on the n-side and fixed negative charge on the p-side.

Note that the space charge layer does not need to be symmetrically around the junction. The width of the space charge layers on each side of the junction is determined by the doping density on that side; a low doping density will lead to a wider space charge layer and a high doping density will lead to narrow space charge region. Of course, the charge on the one side of the junction must equal the charge on the other side, so that for charge neutrality to hold we must have

$$N_a W_p = N_d W_n$$

Where:

N_a = acceptor (p-type) doping level.

W_p = depletion width on the p-side.

N_d = donor (n-type) doping level.

W_n = depletion width on the n-side.

Due to the space charge layer, we will then have an internal electric field across the junction; the direction of this field will be from the positive ions on the n-side to the negative ions on the p-side (remember the direction of the electric field is defined as the direction of a positive charge carrier).

The direction of this field will then oppose the movement of charge carriers due to diffusion. An electron which enters this internal electric field will then experience a force ($F = qE$) and drift from the p-side to the n-side under the influence of the field, while a hole that enters this field will drift from the n-side to the p-side under influence of the field.

This movement of charge carriers under the influence of the internal electric field is termed the *drift current*. In each case, the direction of drift is then opposite to the direction of the diffusion current due to the concentration gradient. The result is that an equilibrium situation is reached where the diffusion current for each particle is exactly balanced by the drift current.

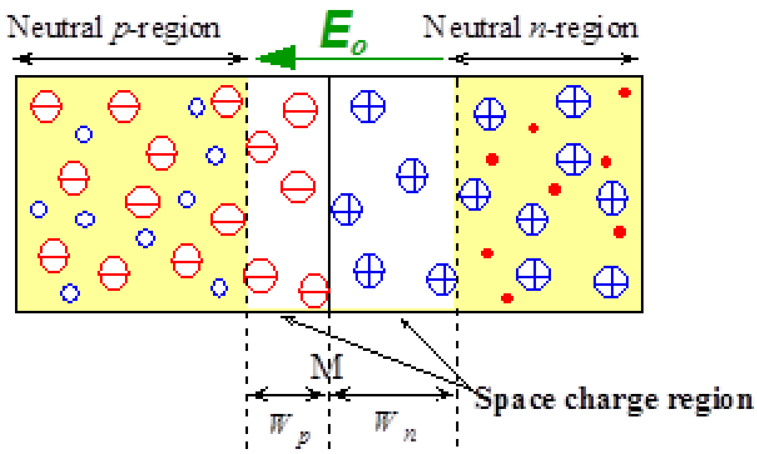


Figure 8: The formation of a depletion region or a space charge region around the metallurgical junction.

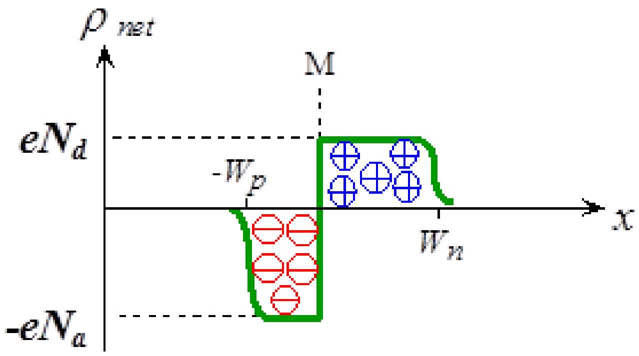


Figure 9: The charge distribution in the space charge layer

It is interesting to note that the electric field will not be uniform as it would have been in the case of a parallel plate capacitor, but starts at zero at the one edge of the space charge layer, reaches a peak at the metallurgical junction and then declines to zero at the opposite edge of the space charge layer. The magnitude of the electric field over the depletion region will then look approximately as below in part (a) of the following figure.

The relationship between the electrical potential (V) and the electric field (E) is given by:

$$E = -\frac{dV}{dx}$$

and if we integrate the field, we obtain the electrical potential (the voltage) as illustrated in part (b) in the following figure. If we now calculate the variation in potential energy across the junction for a hole from the expression $PE(x) = eV$, we get the curve as shown in part (c) in the following figure. As expected, we get inverted curves for electrons and holes – an electron in n-type material will be at a relatively low potential energy, moving it to the p-type material raises the potential energy.

Conversely, a hole in p-type material will be at low potential energy and by moving it into the n-type material we will have to raise its potential energy. The potential energy curve for electrons in part (d) in the following figure is then simply the reverse of that for holes.

The change in electrical potential across the junction can thus be seen as a built-in potential V_0 across the junction, going from p-type to n-type semiconductor. An expression for this built-in voltage is given by:

$$V_0 = \frac{kT}{e}\ln\left(\frac{N_a N_d}{n_i^2}\right) = V_T \ln\left(\frac{N_a N_d}{n_i^2}\right)$$

Where:

k = Boltzman constant (= 1.38 x 10^{-23}).

T = temperature in Kelvin (i.e. 298° K).

e = electron charge (=1.6 x 10^{-19}).

N_a = acceptor doping concentration.

N_d = donor doping concentration.

n_i = intrinsic carrier concentration.

The term V_T is called the thermal voltage and has a value of about 25 mV at 298°K.

Note that this is *the internal voltage across the junction with no applied voltage*. It is not the voltage that will be measured at the electrodes attached to the semiconductors to form the complete diode. The bandgap E_g of the material from which the p-n junction is fabricated will be reflected in the intrinsic concentration n_i.

In summary, at equilibrium in the unbiased p-n junction no current will flow. Any diffusion current of majority carriers (holes from p-side to n-side or electrons from n-side to p-side) will be balanced by a drift current of minority carriers (holes from the n-side to the p-side and electrons from the p-side to the n-side). In the absence of an applied bias voltage, the net flow of charge in any one direction across the semiconductor junction will thus be zero.

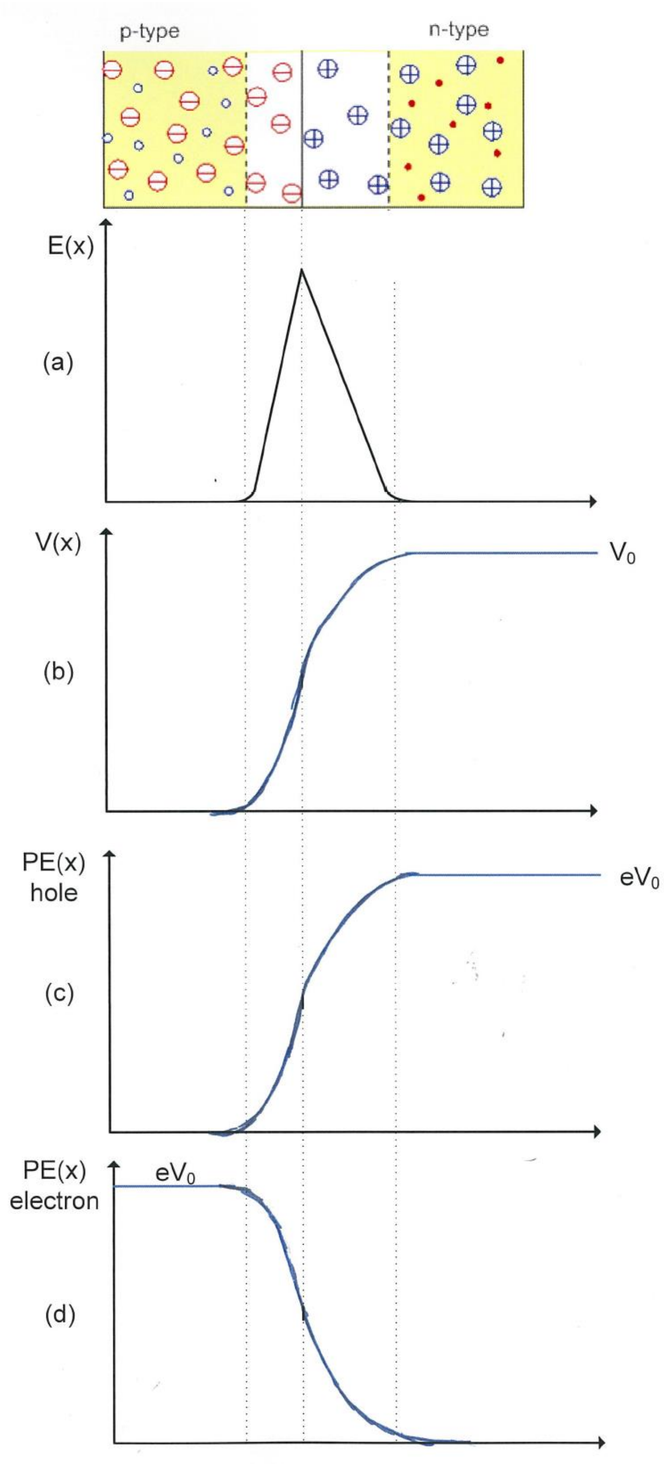


Figure 10: The variation of electric field, electric potential and potential energy for both electrons and holes across the metallurgical junction.

Example for Tutorial 1 – PN Junction of Diode

1. When the n-type semiconductor is joined with the p-type semiconductor, a PN junction is formed. Describe briefly how the potential energies are formed in the PN junction.
[2.5 marks]

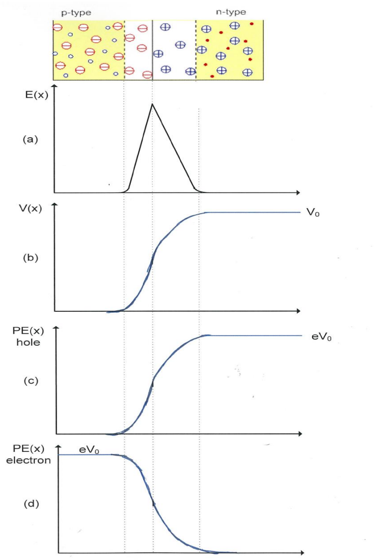
Answer:

For the given PN junction:

Due to the space charge layer, we will then have an internal electric field across the junction; the direction of this field will be from the positive ions on the n-side to the negative ions on the p-side.

The direction of this field will then oppose the movement of charge carriers due to diffusion.

An electron which enters this internal electric field will then experience a force ($F = qE$) and drift from the p-side to the n-side under the influence of the field, while a hole that enters this field will drift from the n-side to the p-side under influence of the field.



As a result, the potential energies are formed due to electrons concentration in the n-side and due to holes concentration in the p-side respectively (see the PE diagrams for hole and electron concentrations on the right-hand side).

2. Given that intrinsic carrier density of Si is 1 x 10^{10} cm^{-3}, determine built-in potential in the PN junction (e.g. p-doping level is 1.5 x 10^{15} cm^{-3} and n-doping level is 7.5 x 10^{16} cm^{-3}).
[2.5 marks]

Answer:

The built-in potential in the diode junction is found from the following equation:

$$V_0 = V_T \text{Ln}\left[\frac{N_A N_B}{(n_i)^2}\right]$$

Putting in the values into the equation:

$$V_0 = 0.025 \times \text{Ln}\left[\frac{(7.5 \times 10^{16})(1.5 \times 10^{15})}{(1 \times 10^{10})^2}\right]$$

$$= 0.025 \times \text{Ln}\,(11.25 \times 10^{11}) = 0.694\ \text{V}$$

3. The Forward Biased PN Junction

In the previous section, we have considered the p-n junction with no applied bias, i.e. no external voltage applied to the junction. We will now first look at how the junction and in particular the flow of current across the junction changes when we have a forward bias as in the following figure, i.e. a positive voltage applied to the p-side of the junction and a more negative voltage applied to the n-side.

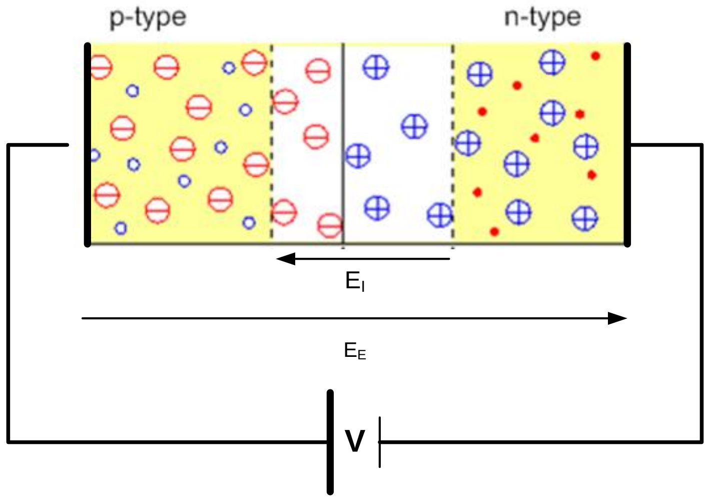


Figure 11: Forward biasing a p-n junction

The total structure over which the voltage V is applied consists of three regions:

(1) the p-type bulk material.

(2) the depletion or space charge layer.

(3) the bulk n-type region.

The resistances of these three regions will differ greatly; the two bulk regions will have reasonable conductivities while the depletion region will have a very low conductivity (high resistance) as there are no free charge carriers in this region.

Nearly the entire voltage drop V will then appear over the space charge layer, with very little voltage drop over the two bulk regions. The sign of V is opposite to V_0 and the result of this external voltage drop V across the depletion region is that the built-in potential V_0 will be reduced by V, so that the potential barrier to diffusion of carriers across the junction is now (V_0–V). At the same time, it will reduce the electric field $E(x)$ which opposes diffusion.

The reduction of these factors has the effect of a dramatic (exponential!) increase in holes that diffuse from the p-side to enter the n-side as minority carriers. Similarly, excess electrons can now diffuse from the n-side to the p-side where they will be injected as minority carriers.

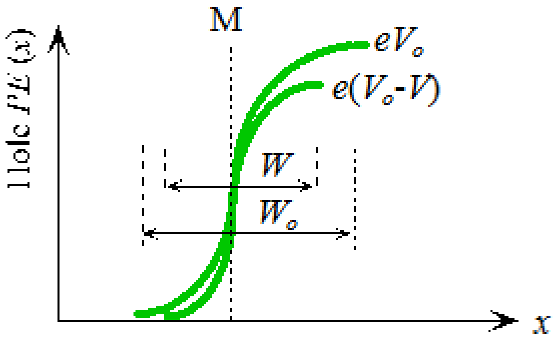


Figure 12: Forward biased p-n junction

The excess holes injected across the depletion region into the n-side region will drift towards the negative terminal, but will eventually recombine with electrons on the n-side as there are many electrons available on this side. Electrons lost by recombination can now be readily replenished by the negative terminal of the battery connected to this side, while the current due to holes will be sustained by more holes diffusing from the p-side. These holes will be replenished by the positive terminal of the battery.

Similar to the above description for holes, electrons will be injected into the p-type region as minority carriers and will diffuse towards the positive terminal. They will recombine with the holes in this region, but the lost holes will be replenished by the positive terminal of the battery while lost electrons will be replaced by more electrons diffusing across the junction and which is itself replenished by electrons from the negative battery terminal.

4. The Reverse Biased PN Junction

When a reverse bias voltage V_r is applied to a p-n junction, the voltage drop will again be mainly across the depletion layer. The negative battery terminal will now attract holes in the p-side to move away from the depletion region, while the positive terminal will attract electrons from the n-side. The effect is a widening of the space charge layer on both sides of the junction.

The movement of electrons in the n-side towards the positive terminal of the battery and the movement of holes in the p-side towards the negative terminal cannot be sustained because there is not a constant supply of these carriers available.

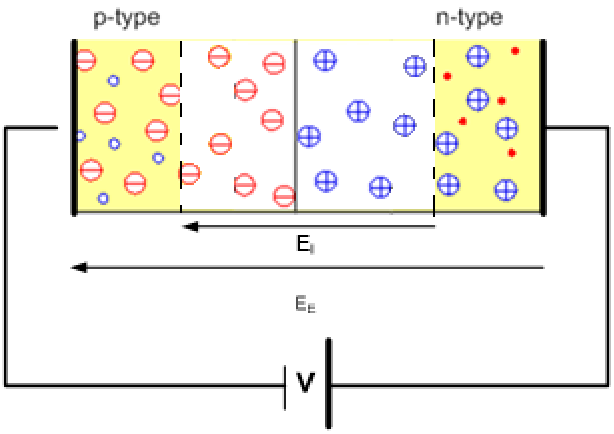


Figure 13: A p-n junction in reverse bias

The applied reverse bias voltage will also increase the built-in potential barrier as shown in the following figure as well as the electric field across the junction. This will assist in sweeping any minority carrier holes close to the space charge layer on the n-side across the junction to the p-side.

Similarly, electrons which are minority carriers in the p-side close to the space charge layer will be drifted across to the n-side by the field. The flow of these holes results in a very small current that can be maintained at near constant values under reverse bias conditions. This current that exists under reverse bias conditions is called the reverse saturation current and is typically re-presented by I_S. This current is typically in the nano ampere range for silicon diodes.

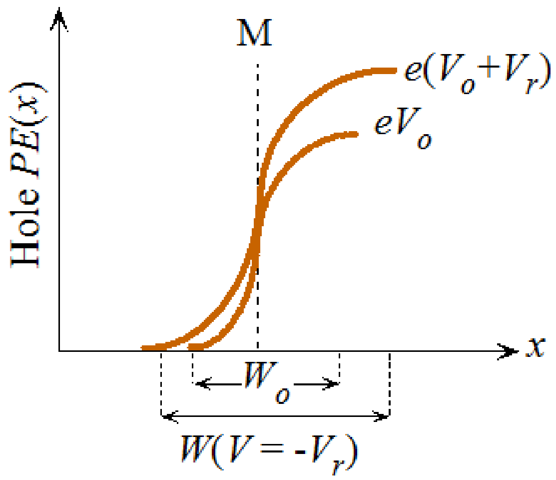


Figure 14: Potential energy barrier across a reverse biased p-n junction.

5. Reverse Bias – Avalanche and Zener Breakdown

The reverse bias voltage across a p-n junction cannot be increased without limit, as eventually the junction will break down with a sudden increase in current observed. This will occur due to the avalanche effect at high values of reverse bias (~tens of volts). This is due to the fact that the velocity of the minority carriers swept out by the increasing field will also increase. At some point, their velocity and associated kinetic energy will become so high that they will start to release additional carriers through collisions.

Once this start to happen, a sudden dramatic increase in the number of carriers will be observed as an avalanche process will be started that will produce a dramatic increase in current. This large current may increase the temperature of the device which may lead to further increase current and potential destruction of the device unless it is operated with an external resistor to limit the current to within power dissipation specifications.

A related but slightly different breakdown process occurs when the p-n junction is very heavily doped in order to enhance breakdown (Zener breakdown) and produces a diode that is specifically used under reverse bias breakdown conditions.

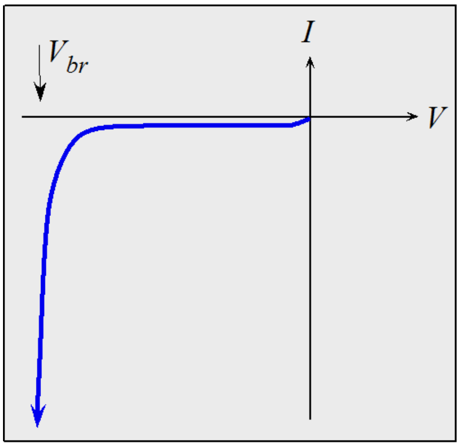


Figure 15: Reverse bias breakdown

6. The Current-Voltage Curve of a Diode

The complete current-voltage (I-V) curve of a diode thus looks as shown as follows. Note the change of scale in the different regions of the graph. It can be shown that this diode current I_D in both the forward and reverse bias regions (not breakdown) of the diode depends on the voltage drop across the diode (V_D) according to the Ebers-Moll equation:

$$I_D = I_S\left[e^{\left(\frac{V_D}{nV_T}\right)} - 1\right] = I_S e^{\left(\frac{V_D}{nV_T}\right)} - I_S$$

Where:

V_D = voltage drop across diode

I_s = reverse bias saturation current

n = ideality factor of the diode (i.e. its value is between 1 to 2)

V_T = thermal voltage factor

The parameter I_s is the reverse bias saturation current and has typical values of 10^{-15} to 10^{-13} A for Si p-n junctions. The exact value will depend on the doping concentrations and cross-sectional area of the junction. The parameter n is called the ideality factor of the diode and takes into account electron-hole recombination in the space charge region and can normally assumed to be $n \approx 1$ - 2. The factor V_T is the thermal voltage as encountered before.

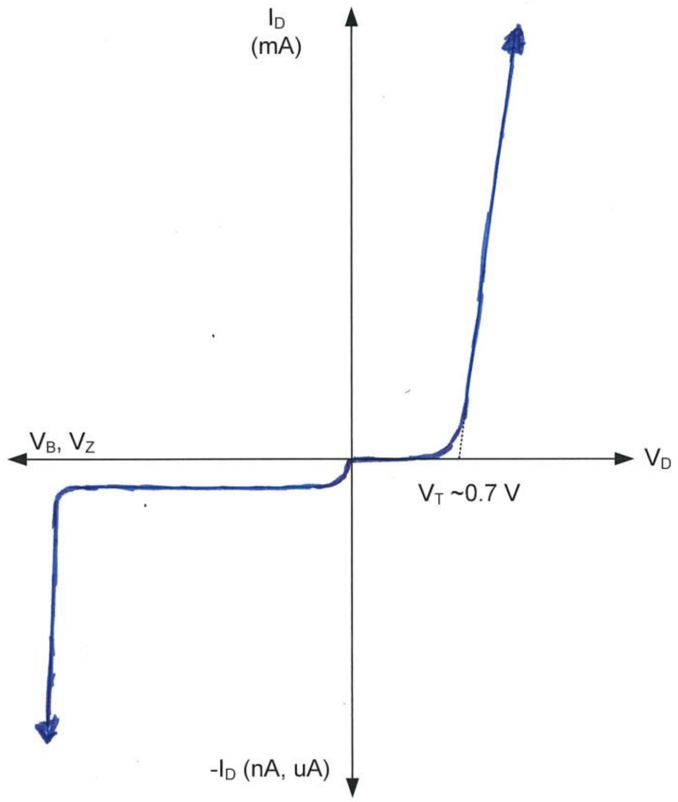


Figure 16: The complete I-V characteristics of a p-n junction. Note the change in current scales.

From this curve, the behaviour can be summarised as follows:

In the *forward bias region*, as the applied voltage is increased from zero, the current observed will initially be zero (at V_D = 0) but then increases exponentially. The current will remain very small, until a certain threshold voltage (V_B) is reached. From this graph, we clearly see the exponential growth in current with voltage.

This growth in current is now so fast that the voltage drop across the diode remains approximately at V_B while the current tremendously increases. The diode resistance has now become close to zero and it is essential at this point that we have an external current limiting resistor in the circuit which will limit the current flowing through the resistor.

The exact value of the breakthrough voltage V_B will be determined by the properties of the semiconductor from which the junction is made. For a silicon p-n junction, this breakthrough voltage is approximately 0.65 – 0.7 V, while for germanium with a much smaller bandgap this value is only ~ 0.3 V.

In the reverse bias region, we will have a sharp increase in current as the voltage is decreased from zero, but the current will then settle at the value of the saturation current which will be typically be in the order of 10^{-6} to 10^{-15} A depending on the type of semiconductor and the physical properties of the junction.

At very large values of reverse bias, we will get a reverse bias breakdown and the current will again dramatically increase with only a small change in the voltage drop across the diode.

In practice, we can obtain the data for V-I graph. We need to set up a diode circuit experiment in the lab as shown in the figure below. Then, we measure the voltage across and current that flows in the diode using voltmeter and ammeter respectively.

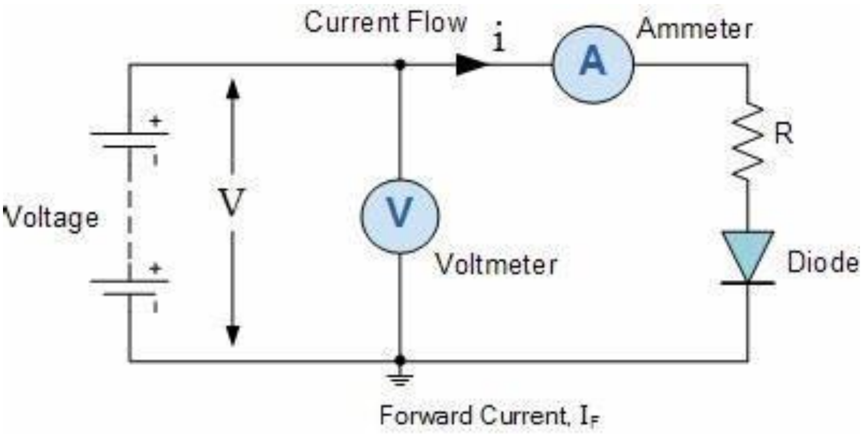


Figure 17: The diode circuit set up for the V-I characteristics experiment

From the circuit given above, we gather the measurements data from experiment in the lab with the condition of the experiment as shown in the table below.

Parameter	Value	Unit
k	1.38×10^{-23}	
T	300	K
e	1.60×10^{-19}	C
N_a	1.00×10^{22}	m^{-3}
N_d	1.00×10^{22}	m^{-3}
n_i	1.00×10^{16}	m^{-3}

V_0	0.71	V
I_s	1.00×10^{-14}	A
V_T	0.026	V

Table 1: Conditions and variables of the experiment

The following table lists all of the voltage and current measurements from the diode circuit.

V_d **(V)**	I_d **(mA)**
0.05	5.84E-11
0.1	4.58E-10
0.15	3.19E-09
0.2	2.19E-08
0.25	1.50E-07
0.3	1.03E-06
0.35	7.02E-06
0.4	4.80E-05
0.45	3.29E-04
0.475	8.59E-04
0.5	2.25E-03
0.525	5.88E-03
0.55	1.54E-02
0.575	4.02E-02
0.6	1.05E-01
0.625	2.75E-01
0.65	7.20E-01
0.675	1.88E+00
0.7	4.93E+00
0.725	1.29E+01
0.75	3.37E+01
0.775	8.82E+01
0.8	2.31E+02

Table 2: Results of measurement from the experiment

Then, we draw and simulate the I-V curve of a diode as shown in the figure below.

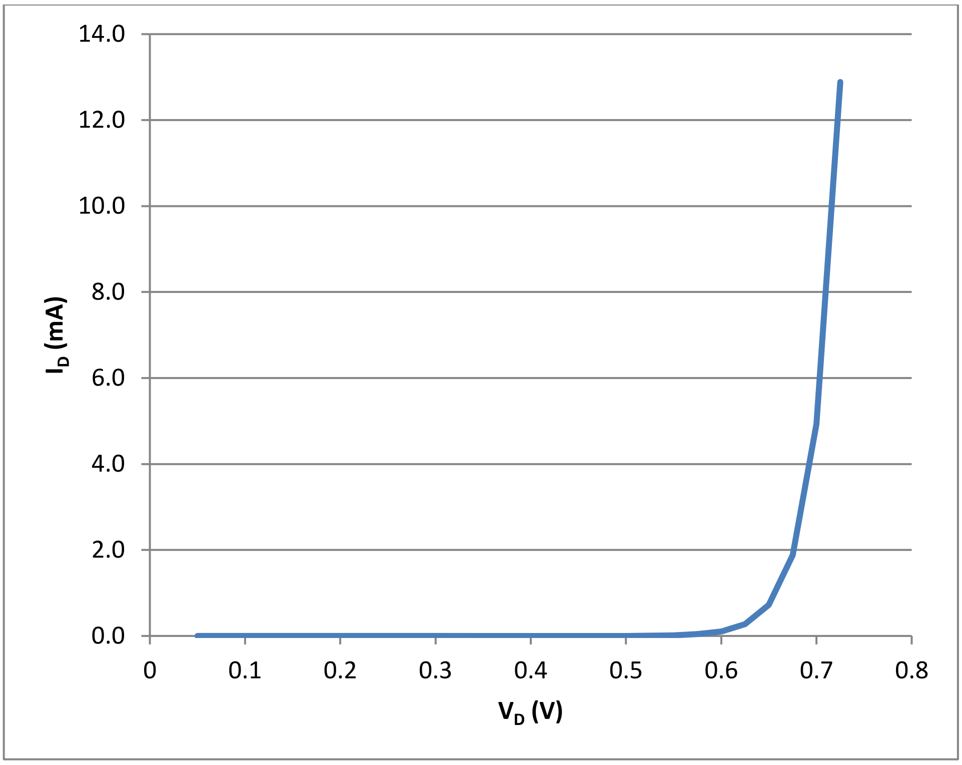


Figure 18: The V-I graph with data obtained from the experiment

Example for Tutorial 2 – Thermal Voltage of Diode

At a temperature of 27°C (i.e. common temperature for components in an enclosed operating system), determine the thermal voltage V_T of a semiconductor device.

[5 marks]

$$V_T = \frac{kT_K}{q}$$

Where:

k = Boltzman coefficient (1.38×10^{-23}).

e = Electron charge (1.60×10^{-19}).

T_k = Temperature in Kelvin.

Answer

Converting the temperature in Celsius to Kelvin, we obtain: [2.5]

$$T = 273 + °\text{C} = 273 + 27 = 300°\text{K}$$

Thus, the thermal voltage of the semiconductor device is: [2.5]

$$V_T = \frac{kT_K}{q} = \frac{(1.38 \times 10^{-23}\text{J/K})(300°\text{K})}{1.6 \times 10^{-19}\text{C}} = 25.875\text{ mV} \cong 26\text{ mV}$$

As a result, the thermal voltage of the device is found to be approximately 26 mV.

7. Temperature Effects

An expression for the built-in potential across a junction was earlier given as:

$$V_0 = \frac{kT}{e}\ln\left(\frac{N_a N_d}{n_i^2}\right) = V_T \ln\left(\frac{N_a N_d}{n_i^2}\right)$$

Where:

N_a = acceptor doping concentration.

N_d = donor doping concentration.

n_i = intrinsic carrier concentration

V_T = thermal voltage factor (i.e. 25 mV at room temperature)

In this equation, the biggest variable will be the value of the intrinsic carrier concentration n_i – this is not only determined by the bandgap of the semiconductor material, but is also an exponential function of the temperature.

Increasing the temperature will thus lower V_0 result in the diode turning on at a lower voltage. The thermal voltage V_T is also a function of temperature, but as the dependence is linear the effect will be much smaller.

The reverse bias saturation current is also dependant on n_i, so that I_s increases very rapidly with increasing temperature. A rough rule of thumb is that the reverse bias diode current will double for every 10°C rise in temperature.

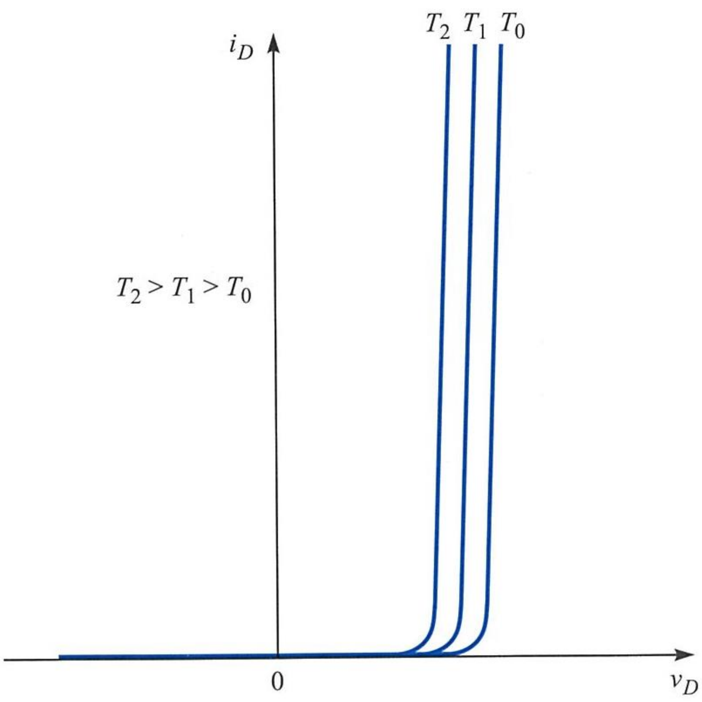
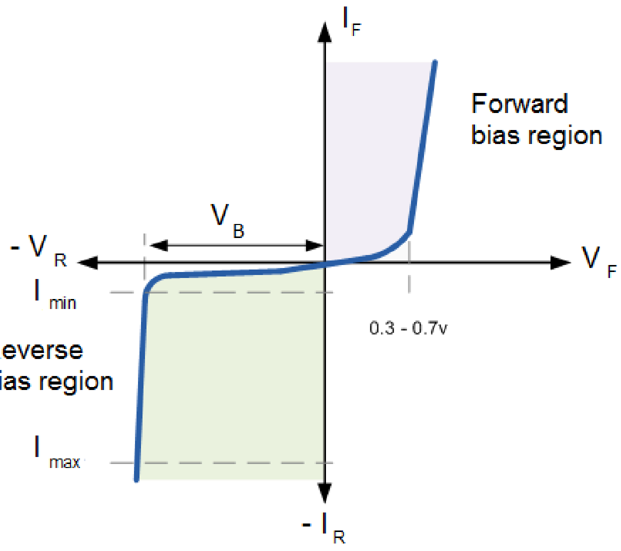


Figure 17: Temperature dependence of the forward bias diode curve.

Example for Tutorial 3 – V-I Graph of Diode

Given a V-I curve graph, briefly explain how temperature influences diode characteristics. [5 marks]

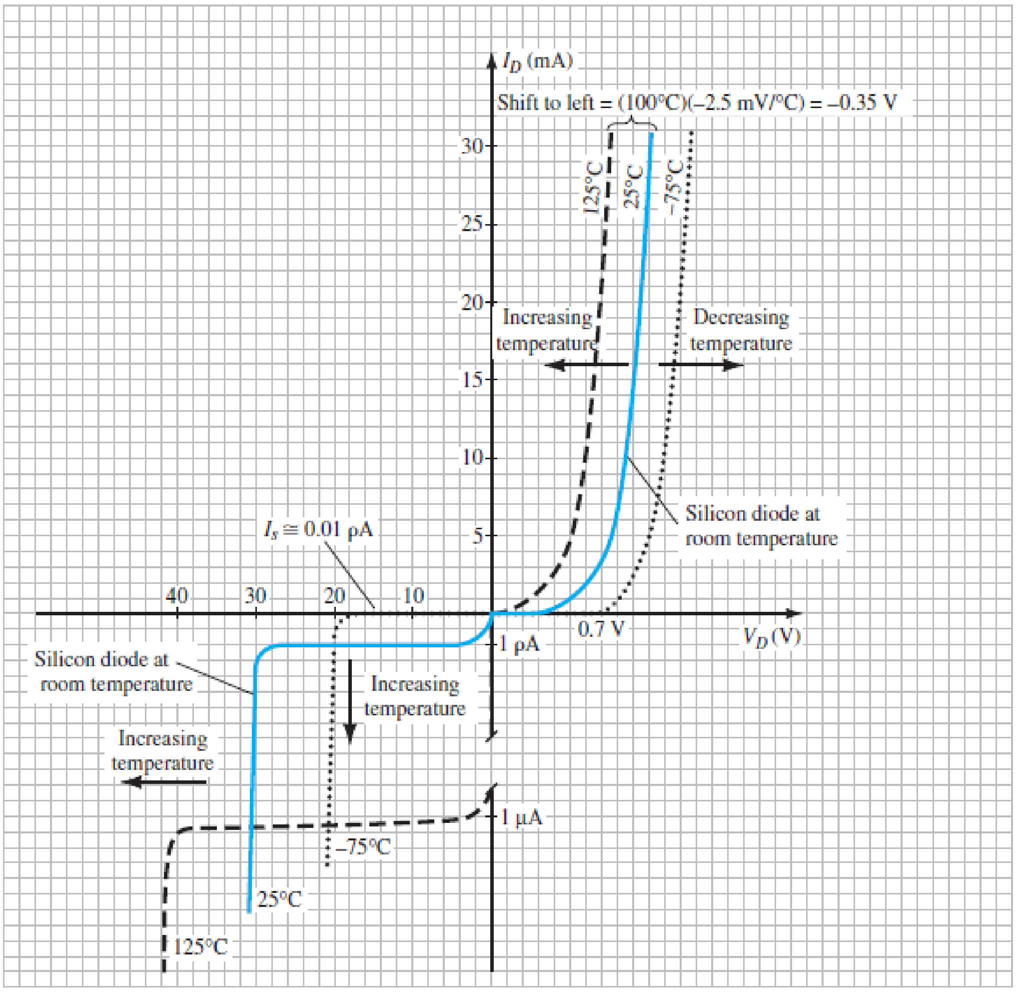


Answer

Temperature could influence the V-I curve: Increasing the temperature could shift the curve towards Y-axis in the plot. [2.5]

This would make the forward bias voltage of the diode less than before and therefore the reverse bias voltage would increase slightly.

Decreasing the temperature would provide the opposite effect towards the temperature characteristics of the diode. [2.5]



8. Applications of Diode

There are many applications of diode in electronics. Mainly, its applications are due to its characteristic i.e. allowing electrical current to flow in one direction and opposing on the other direction.

Restrict the flow of current in one direction e.g. snubber or free wheel circuit applications for allowing electrical energy to commutate safely in the circuit.

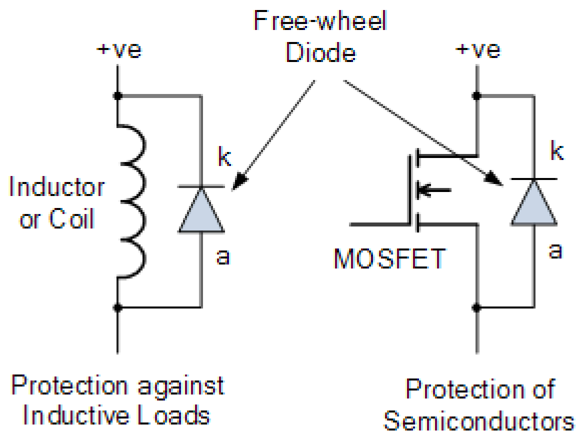


Figure 18: Snubber/free wheel circuit applications of diode

Rectifiers – which is convert AC to DC signal or energy (e.g. electronic chargers), and diode is also used for voltage limiters and regulators - providing safe and stable operating environment for the circuits.

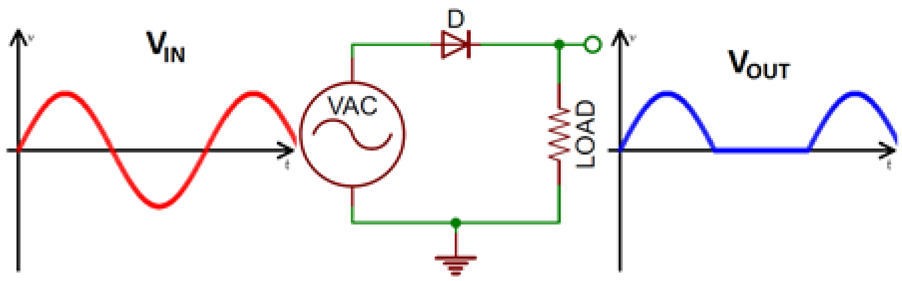


Figure 19: Signal or energy rectifier applications of diode

AM radio detector which is used for detecting the AM wave broadcasting radio signal and convert this into useful sound and information.

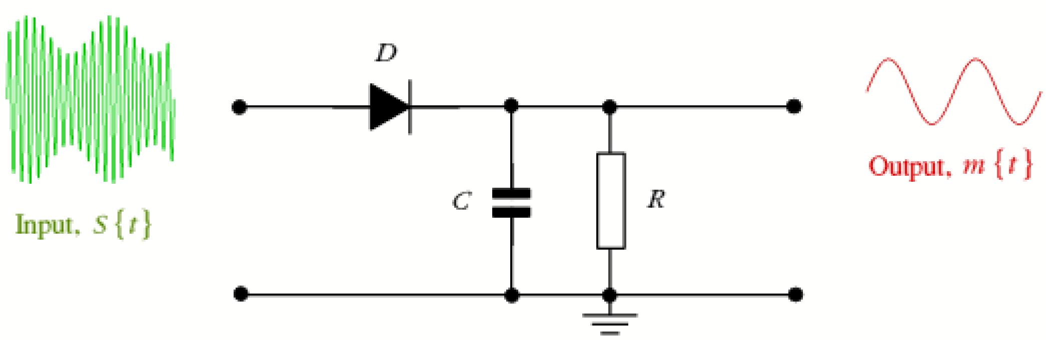


Figure 20: AM radio receiver application of diode

Light emitting diodes (LED) which are for indicator and lighting applications. Then, we have photodiodes and solar cells for sensing of light and generation of electrical energy. The mode of these applications depends on the operating regions of the devices.

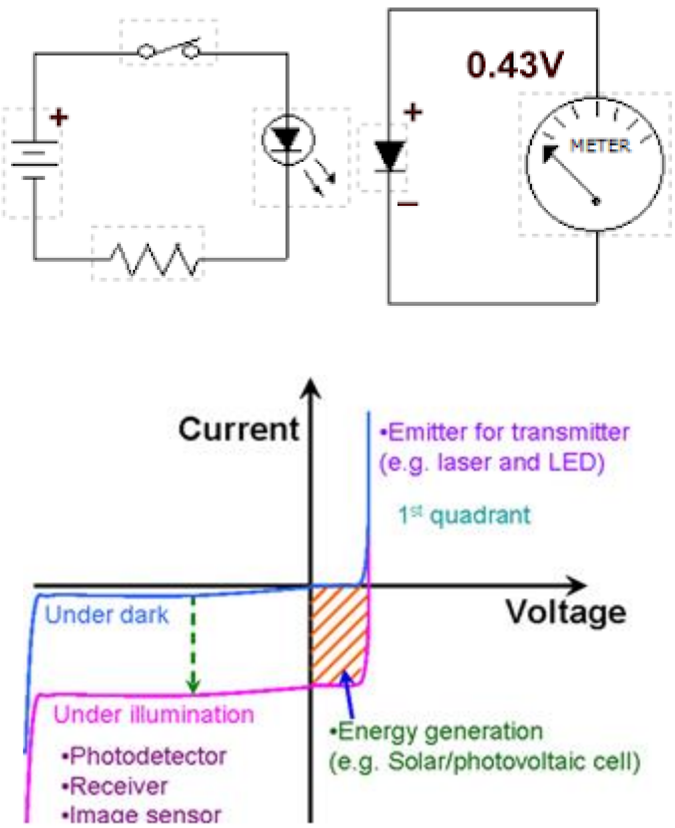


Figure 21: Photo electric applications of diode