Module
2

www.learnabout-electronics.org

Semiconductor Diodes

Module 2.0
Diodes

What you'll learn in Module 2.0

Introduction

Diodes are one of the simplest, but most useful of all semiconductor devices. Many types of diode are used for a wide range of applications. Rectifier diodes are a vital component in power supplies where they are used to convert AC mains (line) voltage to DC. Zener diodes are used for voltage stabilisation, preventing unwanted variations in DC supplies within a circuit, and to supply accurate reference voltages for many circuits. Diodes can also be used to prevent disastrous damage to battery powered equipment when batteries are connected in the wrong polarity.

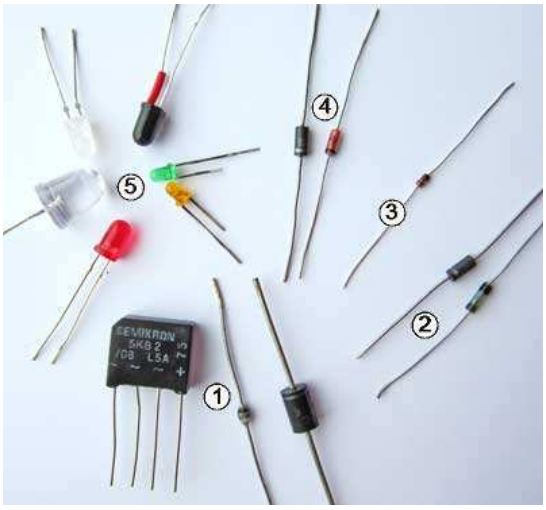


Figure 2.0.1. Diodes

Signal diodes also have many uses in processing signals in electronic equipment; they are used to obtain the audio and video signals from transmitted radio frequency signals (demodulation) and can also be used to shape and modify AC signal waveforms (clipping, limiting and DC restoration). Diodes are also built into many digital integrated circuits to protect them from dangerously large voltage spikes.

LEDs produce light of many colours in a very wide range of equipment from simple indicator lamps to huge and complex video displays. Photo diodes also produce electrical current from light.

Diodes are made from semiconductor materials, mainly silicon, with various compounds (combinations of more than one element) and metals added depending on the function of the diode. Early types of semiconductor diodes were made from Selenium and Germanium, but these diode types have been almost totally replaced by more modern silicon designs.

Fig. 2.0.1 shows a selection of common wire ended diodes as follows:

1. Three power rectifiers, (a Bridge rectifier for use with mains (line) voltages, and two mains voltage rectifier diodes).

2. A point contact diode (with glass encapsulation) and a Schottky diode.

3. A small signal silicon diode.

4. Zener Diodes with glass or black resin encapsulation.

5. A selection of light emitting diodes. Counter-clockwise from red: Yellow and green indicator LEDs, an infra red photodiode, a 5mm warm white LED and a 10mm high luminosity blue LED.

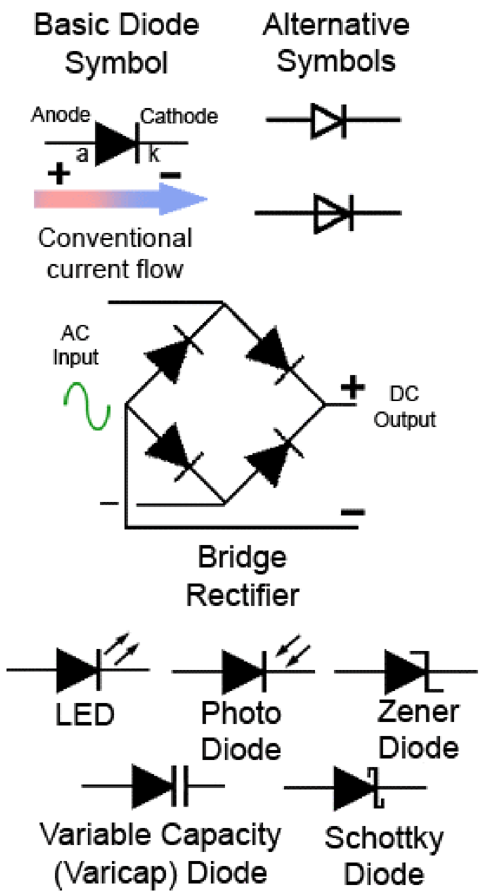


Fig 2.0.2 Diode Circuit Symbols

Diode Circuit Symbols

A diode is a one-way conductor. It has two terminals, the anode or positive terminal and the cathode or negative terminal. Ideally a diode will pass current when its anode is made more positive than its cathode, but prevent current flow when its anode is more negative than its cathode. In the circuit symbols shown in Fig. 2.0.2, the cathode is shown as a bar and the anode as a triangle. On some circuit diagrams the anode of a diode may also be indicated by the letter ‘a’ and the cathode by the letter ‘k’.

Which way does diode current flow?

Notice from Fig. 2.0.2 that conventional current flows from the positive (anode) terminal to the negative (cathode) terminal although the movement of electrons (electron flow) is in the opposite direction, from cathode to anode.

Silicon Diode Construction

Modern silicon diodes are generally produced using one of various versions of the planar process, also used for manufacturing transistors and integrated circuits. The layered construction used in Silicon Planar methods give a number of advantages such as predictable performance and reliability as well as being advantageous to mass production.

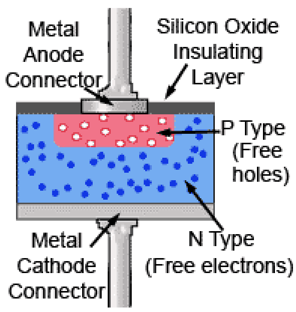


Fig 2.0.3 Silicon Planar Diode

A simplified planar silicon diode is illustrated in Fig. 2.0.3. Using this process for silicon diodes produces two differently doped layers of silicon, which form a ‘PN junction’. Un-doped or ‘intrinsic’ silicon has a lattice structure of atoms, each having four valence electrons, but P type silicon and N type silicon are doped by adding a relatively very small amount of material having either an atomic structure with three valence electrons (e.g. Boron or Aluminium) to make P type, or five valence electrons (e.g. Arsenic or Phosphorus) to make N type silicon. These doped versions of silicon are known as ‘extrinsic’ silicon. The P type silicon now has a shortage of valence electrons in its structure, which can also be considered to be a surplus of ‘holes’ or positive charge carriers, whereas the N type layer is doped with atoms having five electrons in its valence shell and therefore has a surplus of electrons, which are negative charge carriers.

Diode PN Junction

When P and N type silicon are brought together during manufacture, a junction is created where the P type and N type materials meet, and holes close to the junction in the P type silicon are attracted into negatively charged N type material at the other side of the junction. Also, electrons close to the junction in the N type silicon are attracted into the positively charged P type silicon. Therefore along the junction between the P and N type silicon, a small natural potential is set up between the P and N semiconductor material with negatively charged electrons now on the P type side of the junction, and positively charged holes on the N side of the junction. This layer of opposite polarity charge carriers builds up until it is just sufficient to prevent the free movement of any further holes or electrons. Because of this natural electrical potential across the junction, a very thin layer has been formed between the P and N layers at the PN junction that is now depleted of charge carriers and so is called the Depletion Layer. When a diode is connected into a circuit therefore, no current can flow between anode and cathode until the anode is made more positive than the cathode by a forward potential or voltage(V_F) at least sufficient to overcome the natural reverse potential of the junction. This value depends mainly on the materials the P and N layers of the diode are made from and the amount of doping used. Different types of diode have natural reverse potentials ranging from approximately 0.1V to 2 or 3V. Silicon PN junction diodes have a junction potential of about 0.6V to 0.7V

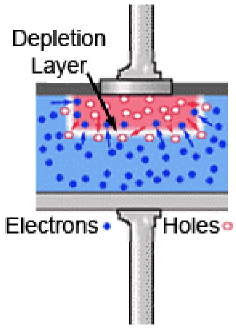


Fig 2.0.4 Diode Depletion Layer

Diode Forward Conduction

Once the voltage applied to the anode is made more positive than the cathode by an amount greater than the depletion layer potential, forward conduction from anode to cathode conventional current commences, as shown in Fig. 2.0.5.

As the voltage applied between anode and cathode increases, forward current increases slowly at first, as charge carriers begin to cross the depletion layer then increasing rapidly in an approximately exponential manner. The resistance of the diode, when ‘turned on’ or conducting in a ‘forward biased’ mode is therefore not zero ohms, but is very low. Because forward conduction increases after the depletion potential is overcome in an approximately following exponential curve, forward resistance (V/I) varies slightly depending on the voltage applied.

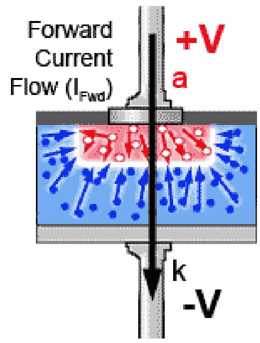


Fig 2.0.5 Diode Forward Conduction

Reverse Biased Diode

When the diode is reverse biased (the anode connected to a negative voltage and the cathode to a positive voltage), as shown in Fig. 2.0.6, positive holes are attracted towards the negative voltage on the anode and away from the junction. Likewise the negative electrons are attracted away from the junction towards the positive voltage applied to the cathode. This action leaves a greater area at the junction without any charge carriers (either positive holes or negative electrons) as the depletion layer widens. Because the junction area is

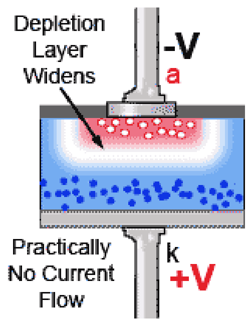


Fig 2.0.6 Diode Reverse Biased

now depleted of charge carriers it acts as an insulator, and as higher voltages are applied in reverse polarity, the depletion layer becomes wider still as more charge carriers away from the junction. The diode will not conduct with a reverse voltage (a reverse bias) applied, apart from a very small ‘Reverse Leakage Current’(I_R), which in silicon diodes is typically less than 25nA. However if the applied voltage reaches a value called the ‘Reverse Breakdown Voltage’ (V_{RRM}) current in the reverse direction increases dramatically to a point where, if the current is not limited in some way, the diode will be destroyed.

Diode I/V Characteristics

The operation of diodes, as described above, can also be described by a special graph called a ‘characteristic curve’. This graphs shows the relationship between the actual currents and voltages associated with the different terminals of the device. An understanding of these graphs helps in understanding how the device operates.

For diodes the characteristic curve is called an I/V characteristic because it shows the relationship between the voltage applied between the anode and cathode, and the resulting current flowing through the diode. A typical I/V characteristic is shown in Fig. 2.0.7.

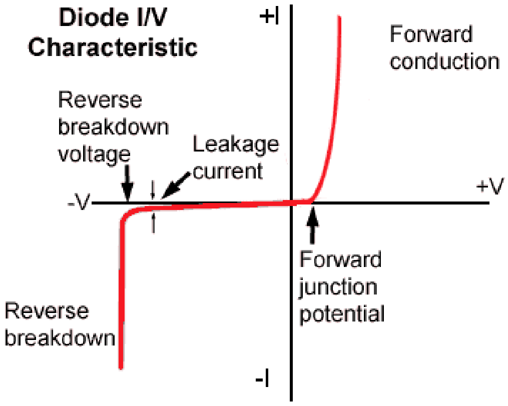


Fig 2.0.7. Typical Diode I/V Characteristic

The axes of the graph show both positive and negative values and so intersect at the centre. The intersection has a value of zero for both current (the Y axis) and voltage (the X axis). The axes +I and +V (top right area of the graph) show the current rising steeply after an initial zero current area. This is the forward conduction of the diode when the anode is positive and cathode negative. Initially no current flows until the applied voltage exceeds the forward junction potential. After this, current rises steeply in an approximately exponential manner.

The -V and -I axes show the reverse biased condition (bottom left area of the graph). Here it can be seen that a very small leakage current increases with the increase in reverse voltage. However once the reverse breakdown voltage is reached, reverse current flow (-I) increases dramatically.

Module 2.1
Silicon Rectifiers

What you'll learn in Module 2.1

After studying this section, you should be able to:

- Describe typical rectifier applications.
- Recognise rectifier polarity markings.
- Describe typical rectifier Parameters.
  - Junction p.d.
  - Average Forward Current.
  - Repetitive Peak Forward Current.
  - Reverse Leakage Current.
  - Repetitive Peak Reverse Voltage.
  - Reverse Recovery Time.
- Describe temperature effects on rectifiers.
  - Thermal runaway.

Silicon Rectifier Diodes

Rectifier diodes, like those shown in Figure. 2.1.1 are typically used in applications such as power supplies using both high voltage and high current, where they rectify the incoming mains (line) voltage and must pass all of the current required by whatever circuit they are supplying, which could be several Amperes or tens of Amperes.

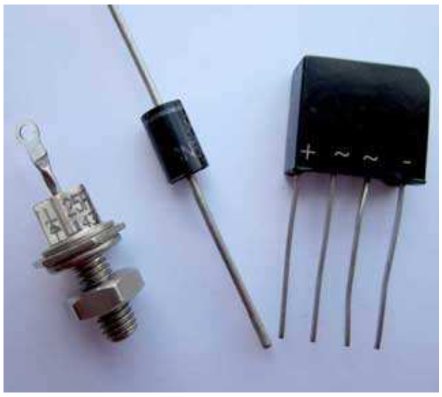


Figure 2.1.1. Silicon Rectifier Diodes

Carrying such currents requires a large junction area so that the forward resistance of the diode is kept as low as possible. Even so, the diode is likely to get quite warm. A black resin case or even a bolt on heat sink helps dissipate the heat.

The resistance of the diode in the reverse direction (when the diode is 'off') must be high, and the insulation offered by the depletion layer between the P and N layers extremely good to avoid the possibility of reverse breakdown, where the insulation of the depletion layer fails and the diode is permanently damaged by the high reverse voltage across the junction.

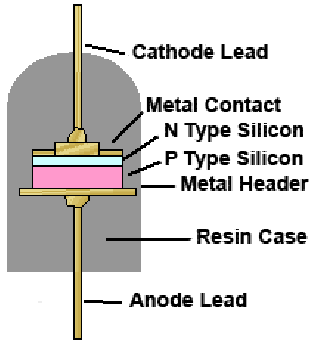


Figure 2.1.2. Silicon Rectifier Construction

Diode Polarity Markings

On the resin case of the diodes, the cathode is usually indicated by a line around one end of the diode casing. Alternative indications do exist however, on some resin encapsulated rectifier diodes a rounded end on the casing indicates the cathode as shown in Fig. 2.1.2. On metal stud rectifier diodes, the polarity of the diode may be shown by diode symbol printed on the case. The stud end of the diode is often the cathode, but his cannot be relied on, as Fig. 2.1.1 shows, it may be the anode! On bridge rectifier dioded the + and - (plus and minus) symbols shown on rectifier case indicate the polarity of the DC output and not the anode or cathode of the device, the AC input terminals are indicated by small sine wave symbols. One corner of the casing on some in line bridge rectifiers is also often chamfered off, but this should not be taken as a reliable guide to polarity, as rectifiers are available that use this indication as either the + or - output terminal.

Silicon rectifier diodes are made in many different forms with widely differing parameters. They vary in current carrying ability from milliamps to tens of amps, some will have reverse breakdown voltages of thousands of volts.

Rectifier Parameters

What the parameters mean.

Depletion layer (Junction) p.d.

The depletion layer or junction p.d. is the potential difference (voltage) that is naturally set up across the depletion layer, by the combination of holes and electrons during the manufacture of the diode. This p.d. must be overcome before a forward biased diode will conduct. For a silicon junction the p.d is about 0.6V.

Reverse leakage current (I_R).

When a PN junction is reverse biased a very small leakage current (I_R) will flow due mainly to thermal activity within the semiconductor material, shaking loose free electrons. It is these free electrons that form a small leakage current. In silicon devices this is only a few nano-Amperes (nA).

Maximum Repetitive Forward Current (I_{FRM}).

This is the maximum current that a forward biased diode may pass without the device being damaged whilst rectifying a repetitive sine wave. I_{FRM} is usually specified with the diode rectifying a sine wave having a maximum duty cycle of 0.5 at a low frequency (e.g. 25 to 60Hz) to represent the conditions occurring when a diode is rectifying a mains (line) voltage.

Average Forward Current (I_{FAV}).

This is the average rectified forward current or output current (I_{FAV}) of the diode, typically this would be the forward current when rectifying a 50Hz or 60Hz sine wave, averaged beteween the period when a (half wave) rectifier diode is conducting, and the period of the wave when the diode is reverse biased. Notice that this average value will be considerably less than the repetitive value quoted for I_{FRM}. This (and other parameters) are also largely dependant on the junction temperature of the diode. The relationship between the various parameters and junction temperature is usually specified as a series of footnotes in manufacturers data sheets.

Repetitive Peak Reverse Voltage (V_{RRM})

The maximum peak voltage that may be repetitively applied to a diode when it is reverse biased (anode - cathode +) without damage to the device. This is an important parameter and refers normally to mains (Line) operation. E.g. a diode used as a half wave rectifier for rectifying the 230V AC mains voltage will conduct during the positive half cycle of the mains waveform and turn off during the negative half cycle. In a power supply circuit the cathode of the rectifier diode will usually be connected to a large electrolytic reservoir capacitor, which will maintain the cathode voltage of the rectifier at a voltage close to the peak voltages of the mains waveform. Remember that the 230V AC wave refers to the RMS value of the wave, so the peak value will be about 230V x 1.414 = approximately +325V. During the negative half cycle of the mains waveform the anode of the diode will fall to a maximum negative value of about -325V. Therefore there will be repetitive periods (50 or 60 times per second when the reverse voltage across the diode will be 325V x 2 = 650V. For this task therefore it would be necessary to use a rectifier diode with a V_{RRM} parameter of at least 650V, and to ensure reliability there must be a safety margin for such an important component, so it would be wiser to select a diode with a V_{RRM} of 800 or 1000V.

Maximum Working Peak Reverse Voltage (V_{RWM})

This is the maximum allowable reverse voltage. The reverse voltage across the diode at any time, whether the reverse voltage is an isolated transient spike or a repetitive reverse voltage.

Maximum DC Reverse Voltage (V_R)

This parameter sets the allowable limit for reverse voltage and is usually the same value as V_{RRM} and V_{RWM}. Theoretically these maximum parameters could each be different, but as any voltage (instantaneous, repetitive or constant) that is greater by no more than about 5% than any of these parameters could potentially destroy the diode, it is always advisable to be cautious when fitting diodes and build in a reasonable margin to allow for unexpected spikes in voltage. One common safety measure to protect power supply rectifiers from externally generated spikes is to connect a small capacitance, high voltage capacitor, typically a disc ceramic type across each of the four diodes in a bridge rectifier as shown in Fig. 2.1.3.

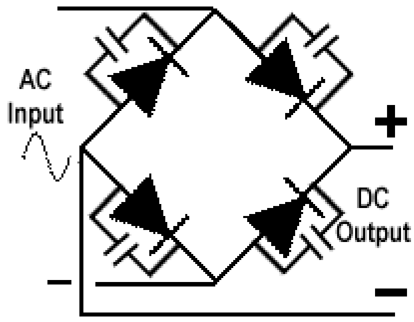


Fig 2.1.3 Spike Suppression

Reverse Recovery Time (t_{rr})

The time required for the current to fall to a specified low level of reverse current when switching from a specified forward current (diode turned on) to a specified reverse current (diode turned off, typically <10% of the value of the 'on' current). Typical t_{rr} times for rectifier diodes, though not as fast as small signal diodes, and depends somewhat on the voltages and currents involved, can be found to be in the tens of nanoseconds (ns) e.g. 30ns for a BYV28 3.5A I_{AF} 50V rectifier and <60ns for a BYV44 dual 30A I_{AF} 500V rectifier.

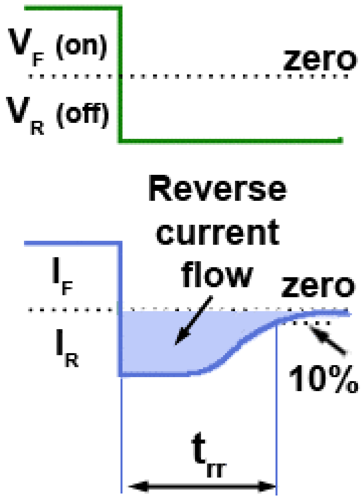


Fig 2.1.4 Reverse Recovery Time (t_{rr})

When a rectifier diode is used in a high speed switching operation such as in a switched mode power supply The reverse current should ideally fall to zero instantly. However when the diode is conducting (before switch off) there will be a large concentration of minority carriers on either side of the junction; these will be holes that have just crossed to the N type layer and electrons that have just crossed to the P type layer, and before they have been neutralised by joining with majority carriers.

If a reverse voltage (V_R) is now suddenly applied, as shown in Fig. 2.1.4, the diode should be turned off, but instead of the current through the diode falling instantly to zero, a reverse current (I_R) is set up as these minority carriers are attracted back across the junction (holes back into the P layer and electrons back into the N layer). This reverse current will continue to flow, until all these charge carriers return to their natural side of the junction.

Maximum Temperature

Each of these parameters can be affected by other factors, such as the ambient temperature in which the diode is operating, or the junction temperature of the device itself. Any semiconductor generates heat, especially those used in power supplies. Therefore it is essential that the design of such circuits takes into account the effects of temperature. One of the greatest problems is the prevention of Thermal Runaway where a diode (or any other semiconductor) increases its temperature, leading to an increase in current through the device, which leads to a further increase in temperature and so on until the device is destroyed. To prevent this problem each of the diode parameters references temperature, for example the reverse leakage current of a silicon PN diode is usually quoted at an ambient temperature of 25°C but is likely to approximately double for each 10°C above that figure. Also an increase in temperature will cause a decrease in the forward junction potential of about 2 to 3 mV for every 1°C of temperature increase. Temperature has an even greater effect on Schottky rectifiers.

Module 2.2
Schottky Diodes

What you'll learn in Module 2.2

After studying this section, you should be able to:

- Understand construction methods used in Schottky Diodes.
- Recognise advantages & disadvantages of Schottky Diodes.
- Describe typical applications for Schottky Diodes.

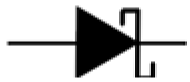


Figure 2.2.1. Schottky Diode Circuit Symbol

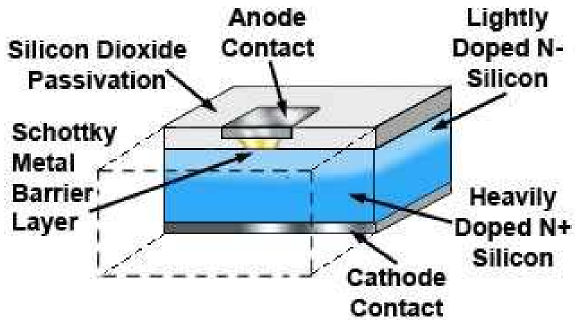


Figure 2.2.2. Small Signal Schottky Diode

The Schottky Diode

Schottky diodes, also called Hot Carrier Diodes or Schottky Barrier Diodes use a metal/semiconductor junction instead of a P semiconductor/N semiconductor junction, a basic principle that dates back to the earliest 'Cats Whisker' diodes at the end of the 19th century. Although germanium diodes using the cats whisker or point contact principle illustrated in Fig. 2.2.3 fell into disuse by the late 20th century, a Metal/semiconductor junction is still used in Schottky diodes manufactured using silicon planar technology in place of the cats whisker, and can be manufactured with more reliable characteristics in both discrete component and integrated circuit form to provide the advantages of these early diodes in many modern circuits.

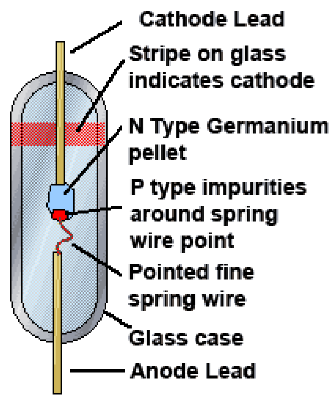


Figure 2.2.3. Germanium Point Contact Diode

Low Junction Potential

The metal to silicon junction used in Schottky diodes provides several advantages (and some disadvantages) compared with a PN silicon diode. The P type region of the PN diode is replaced by a metal anode, usually gold, silver, platinum, tungsten, molybdenum or chromium. When the diode is formed during manufacture a small junction potential occurs between the metal anode and the N type silicon. Typically this will be about 0.15V to 0.3V depending on the metal used, and the difference between the energy levels of the electrons in the metal and the adjoining silicon, all of these metals produce a junction potential called the Schottky Barrier. Because this potential barrier is smaller than the 0.6V junction potential of a PN silicon junction, this makes Schottky diodes such as the BAT49 and the 1N5711 from ST Microelectronics very suitable for small signal radio frequency applications in circuits such as the RF mixer, modulator and demodulator stages in many radio communication systems, as well as high speed switching in digital logic circuits.

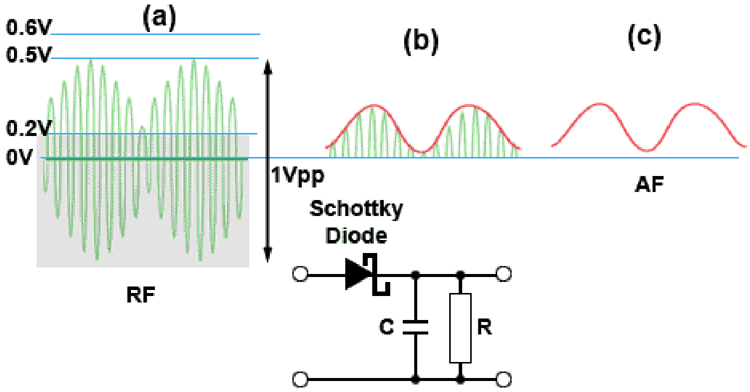


Figure 2.2.4. AM Demodulation Using a Schottky Diode

Basic AM Demodulation

Fig. 2.2.4 illustrates the advantage of using Schottky diodes for demodulating small amplitude AM waves. Amplitude Modulated signals are used in both broadcasts and communictaions as they can be transmitted over much longer distances using relatively low power transmitters than would be possible using VHF or UHF signals. When an AM signal is recieved its amplitude at the receiver may only be a few millivolts or even microvolts. This signal is greatly amplified by the receiver but may still be quite small (e.g. 1Vpp as shown in Fig. 2.2.4) when it is applied to the demodulator to recover the modulating signal. It would not therefore have sufficient amplitude (0.5V) to overcome the junction voltage of a silicon PN diode (0.6V), so no signal would be demodulated. Using a Schottky diode with a junction potential of only 0.2V however allows the demodulator to produce usable information from weaker signals than would be possible using a silicon PN diode.

The demodulation process involves applying the amplitude modulated signal to the Schottky diode, which only conducts when the positive half cycles of the RF are greater than 0.2V. (Fig. 2.2.4a) This produces an assymetrical RF signal that is applied to the capacitor C, which charges to nearly the peak value of each half cycle of the RF to produce a signal, (Fig. 2.2.4b) following the envelope shape of the RF signal, this is now an audio frequency waveform (shown in red)(Fig. 2.2.4c) that varies with the same shape as the audio signal originally used to modulate the RF. This demodulated audio signal is now amplified and used to drive the radio loudspeaker.

High Speed Switching

A typical metal/N type Schottky junction works because when the junction is forward biased, the depth of the barrier decreases, allowing majority charge carriers (electrons) from the silicon to flood into the metal anode, where they are at a higher energy level than the electrons in the metal. Here they rapidly lose some of their energy and add to the free electrons in the metal, creating an electron flow from cathode to anode. When a reverse voltage is applied the junction however, the Schottky barrier level increases and the great majority of the electrons in the metal layer do not have a high enough energy level to re-cross the junction into the silicon, so only a very small leakage current flows, although the leakage current is greater than in a comparable PN diode.

Because, in a Schottky diode there is no exchanging and re-exchanging of holes and electrons across the junction, as happens in the PN diode, the switching speed is much faster. Schottky diodes therefore have a minimal Reverse Recovery Time (t_{rr}). Any delay in switching, which can be as low as 100 pico-seconds is mainly due to the capacitance of the junction, which especially in small signal switching types of Schottky diodes, as illustrated in Fig.2.2.2, is very small due to the small area of the junction. The junction capacitance is therefore typically less than 10pF, allowing some specialist types of Schottky diodes to operate at low voltages at frequencies in the GigaHertz and TeraHertz ranges.

Schottky Power Rectifiers

In Schottky power rectifiers similar to that illustrated in Fig.2.2.5, this low junction potential is less important but does have the advantage that when the diode is conducting there is less power dissipated at the Schottky junction than in a comparable PN diode, so less heat is generated at the junction.

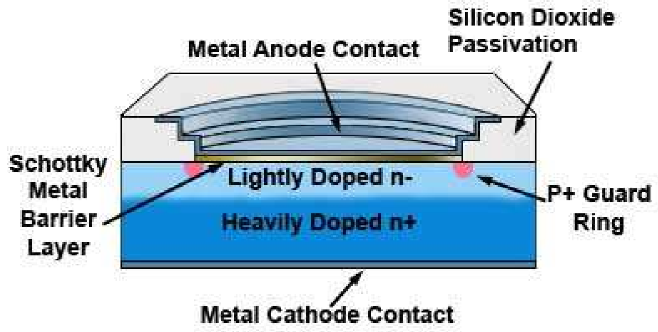


Figure 2.2.5. Schottky Rectifier Diode

High Speed Switching Rectifiers

The main advantage in using Schottky diodes in power supplies is in its very fast switching speed. Many modern circuits use Switched Mode Power Supplies, which operate using square waves at high frequencies that need to be rectified at the power supply output. The fast switching speed of Schottky diodes such as the BYV44 from NXP or the BYV28 from Vishay are ideal for this purpose. However, the Schottky rectifier diode also has its drawbacks.

Schottky Reverse Current Limitations

Rectifier diodes are generally designed to handle large currents and large reverse voltages but the Schottky design is not as capable at either of these requirements as comparable PN junction diodes. Forward current generates heat at the diode junction and although the low junction potential of the Schottky design may generate less heat, the low junction potential of the Schottky depends on a very thin (the thinner the junction the lower the potential) metal layer at the junction. A thinner layer also means that the reverse leakage current of the diode will be greater. This can be seen from a comparison of typical PN and Schottky characteristic curves (not to scale) shown in Fig. 2.2.6. Also, although the Schottky junction may be considered to generate less heat per Watt than the PN junction, in order to keep its reverse leakage current within acceptable limits, the maximum junction temperature must be kept below typically 125°C to 175°C (depending on type) compared with 200°C or more for a PN diode.

Over Voltage Protection

If the reverse leakage current is not carefully controlled and the diode also protected against sudden spikes in voltage, it is possible that the current may become large enough (even momentarily) to take the reverse current into the reverse breakdown region and destroy the diode. To prevent this, it is common in Schottky rectifiers to include a Guard Ring round the junction area, this consists of a ring of heavily doped P+ type silicon embedded into the N- type cathode region, in effect forming a reverse biased PN junction within the Schottky diode structure, as can be seen in Fig. 2.2.5. Because the guard ring is heavily doped it behaves rather like a Zener diode with pronounced avalanche characteristics, i.e. it will suddenly conduct heavily in its reverse current mode at a precise reverse voltage. This point is designed to be at a lower voltage than the breakdown voltage of the Schottky junction, therefore the Schottky diode is protected as the current taken by the PN junction will be sufficient to prevent the reverse voltage rising above safe limits.

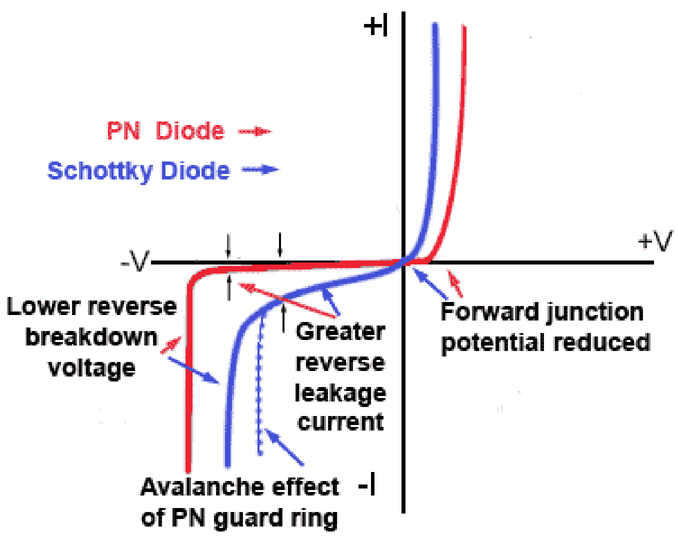


Figure 2.2.6. Schottky & PN Characteristics Compared

For any circuit design it is important to carefully consider the advantages and disadvantages of both Schottky and PN Junction diodes to ensure that the chosen components will perform both efficiently and reliably. There is no simple answer to which of these types of diode is most suited to a particular purpose. It is a matter of selecting a diode whose individual parameters match the required purpose. Schottky rectifier diodes may be preferable for switching speed and efficiency, and PN diodes better for higher current and voltage designs. But the final choice depends on the characteristics of the individual components.

Figure 2.2.7. Surface Mount Schottky Rectifier in a DO-214 (5.3 x 3.6mm) Package

Module 2.3
Small Signal Diodes

What you'll learn in Module 2.3

After studying this section, you should be able to:

Describe typical construction methods used in Small Signal Diodes.

Describe typical applications for Small Signal Diodes.

- Waveform Clipping.
- Waveform clamping or DC restoration.
- Input protection.

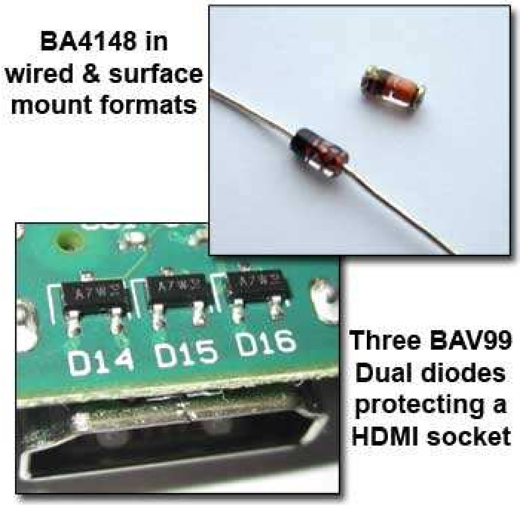


Figure 2.3.1. Small Signal Diodes

Small Signal Diodes

Many electronic systems use small signals, these can be categorised as AC (continuously varying) signals or pulses whose amplitude is limited to a few volts, or in many cases a few milli or even micro-volts. Such uses include radio, audio and video signals as well as digital signals that may be found in the home, in industrial equipment, automotive, aeronautic and musical systems.

Small Signal PN Diode Construction

There are very many small types of signal diodes available in both wire ended and surface mount (SMT) format. They differ from rectifier diodes in several ways, they generally have smaller junction areas giving the junction less capacitance making them more useful at higher frequencies. High speed small signal diodes, often called switching diodes are designed to be used in circuits with high frequency or fast pulses and have a very fast 'Reverse Recovery Time', typically a few nanoseconds or less. They are physically smaller than dedicated power rectifiers and have lower maximum reverse voltage parameters. Some designs are for general purpose use and others for more specialised purposes. The basic construction for a silicon planar diode is to infuse a layer of doped silicon (e.g. N type) with a region of P type silicon for example, as shown in Fig. 2.3.2. In this simple example, a layer of N type silicon has a layer of P type silicon infused into its upper surface and a PN junction is formed. However the curvature of the junction makes the parameters of the diode more difficult to predict accurately and one way to produce an improved small signal diode is shown in Fig. 2.3.3.

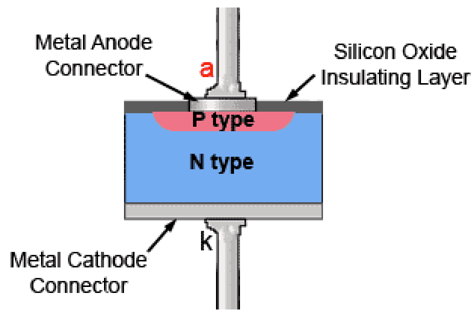


Figure 2.3.2. Simple Planar Diode

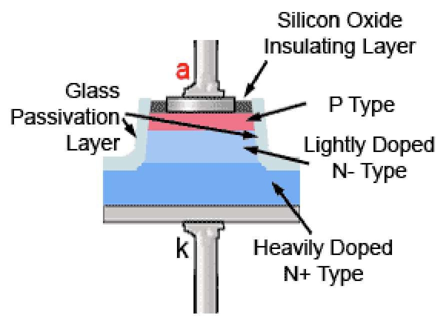


Figure 2.3.3. Simple Mesa Diode

After the P type layer is created, and while the diode is still part of a silicon wafer containing many identical components, a section of the PN block is etched away to form what is known as a Mesa (an isolated high topped plateau) Diode. This device now has a smaller junction area, giving a lower junction capacitance and a flatter junction giving more predictable characteristics. Additionally the N type layer is divided into two layers, a N- (lightly doped) and a N+ (heavily doped) layer, which gives a better distribution of electric charge across the diode, helping to reduce the reverse potential across the junction and to a give better reverse voltage capability. The sides of the 'mesa' are also coated with an insulating layer of either silicon oxide or glass to 'passivate' the junction area protecting it from deterioration due to oxidisation. Such modifications produce a more reliable diode with more accurately controlled characteristics.

Wave Shaping With Diodes

Clipping

Signal processing may also require the process of clipping, which is a process much the same as rectification, that is altering the shape of a wave by removing part or all the positive or negative peak of a wave. Clipping action on a sine wave is shown in Fig. 2.3.4.

In Fig.2.3.4(a) a simple circuit consisting of a resistor and diode is used to clip the positive half cycles of a sine wave. Assuming an ideal diode (with zero junction potential) is used, the output wave will be half the amplitude of the input wave and completely negative going.

In Fig. 2.3.4(b) the diode is simply reversed and now (again assuming an ideal diode) the negative half cycles are removed, leaving just the positive half cycles of the sine wave.

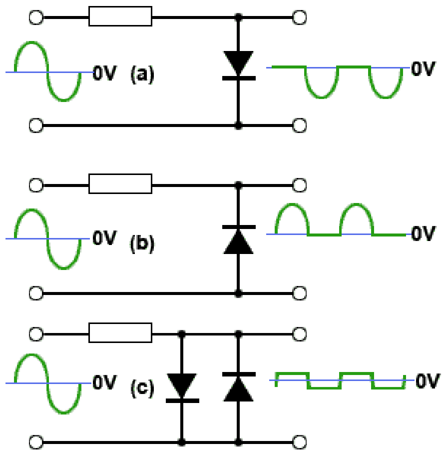


Figure 2.3.4. Clipping Action

In Fig. 2.3.4 (c) two diodes are used, and this time it is assumed that the diodes are not theoretical ideal diodes, but small signal silicon types. These will have a forward junction potential of about 0.7V and so do not clip the sine wave at 0V, but at +0.7V and -0.7V leaving a square(ish) wave of 1.4Vpp.

Clipping With Real Diodes

In practical silicon diodes the anode must be more positive than the cathode by a certain amount. A Silicon junction diode must have an anode voltage which is about 0.7v more positive than its cathode before current will flow. If the waveforms with which the diode is used are fairly large, then the small voltages mentioned above are insignificant. In Fig. 2.3.5 the same clipping circuit as that shown in Fig. 2.3.4(a) is used, (actually with a 1K resistor and a BA4148 diode). This circuit should remove the positive half cycle of the waveform, and with a sine wave input of 20Vpp it does quite a good job. The output wave was carefully measured and found to be 10.7Vpp. Assuming that the waveform has a large amplitude, the output wave will be close enough to half the amplitude of the input wave.

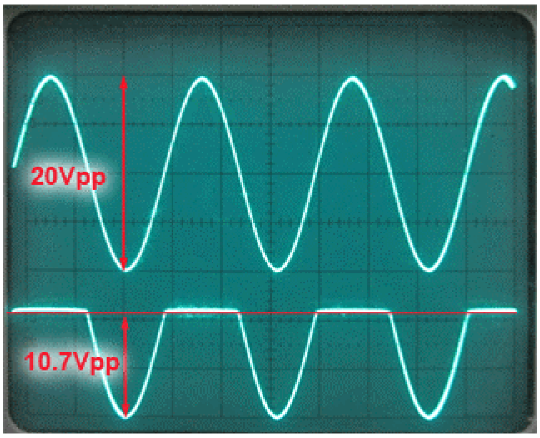


Figure 2.3.5. Clipping a 20Vpp Sine Wave

In Fig. 2.3.6 the same circuit is used but with an input wave of only 2vpp. This time the output wave is not too different from the input wave above it on the CRO screen. Its a bit squashed on the positive half cycles and also rather rounded instead of being flat. The reason for this is firstly because the junction potential of the diode is nearly half the peak to peak value of the wave so hardly any clipping of the positive half cycles takes place.

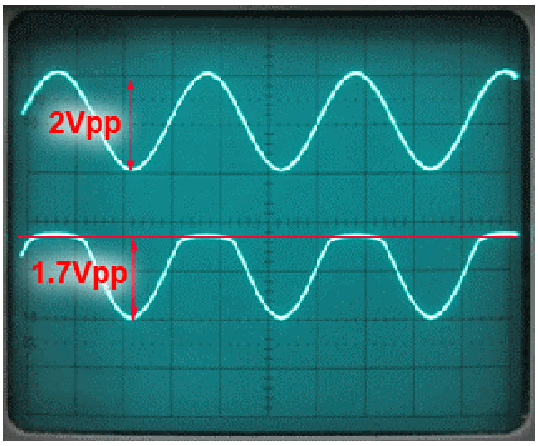


Figure 2.3.6. Clipping a 2Vpp Sine Wave

Additionally the 'clipped' half cycle is rounded because the junction potential is only 'about' 0.7V. In fact the diode begins to conduct when the forward bias is about 5.5V and is just about fully conducting at about 0.7V. Remember that the diode characteristic curve shown in Fig. 2.0.7 (on the Diodes page 2.0) showed how the current increases in an exponential curve in the forward biased condition. This also means that the forward resistance of the diode is not constant but varies (slightly) depending on the slope of the curve and will have slightly different values at different forward current values.

Diode Clamping or DC Restoration

Diodes can be used to clamp some point on a waveform to a particular DC level. Sometimes also called 'DC Restoration' a simple circuit consisting of a diode and a capacitor can be used to set the DC level of an AC wave to any required level. This can be useful in many electronic systems, for example a digital circuit where the upper or lower limits of square waves or pulses are required to have specific voltage values, or video circuits where the brightness of any picture element depends on a particular voltage.

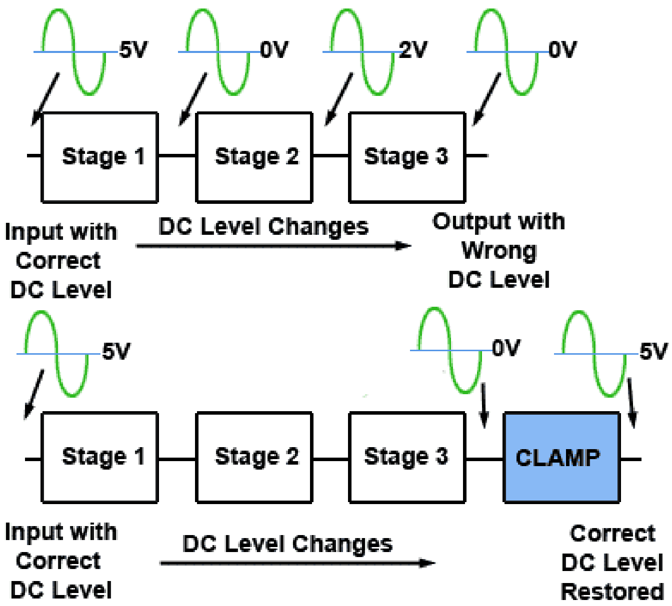


Figure 2.3.7. Clamping or DC Restoration

Fig. 2.3.7 illustrates a typical situation where clamping may be needed. When an AC signal passes through a number of stages (individual circuits such as a series of amplifiers) within a system, it is quite possible that the original DC level of a waveform will be altered. A DC restorer or clamp circuit in the final stage can then be used to restore the original DC level or set a new DC level as required. A simple clamping circuit can do this by clamping one point (e.g. the upper or lower tips of the waveform) to an appropriate voltage.

A simple DC restorer or clamp circuit is shown in Fig. 2.3.8 and consists of a capacitor and a diode. The resistor R_L represents the load on the circuit but also forms a time constant with the capacitor (C). The operation of the circuit is as follows:

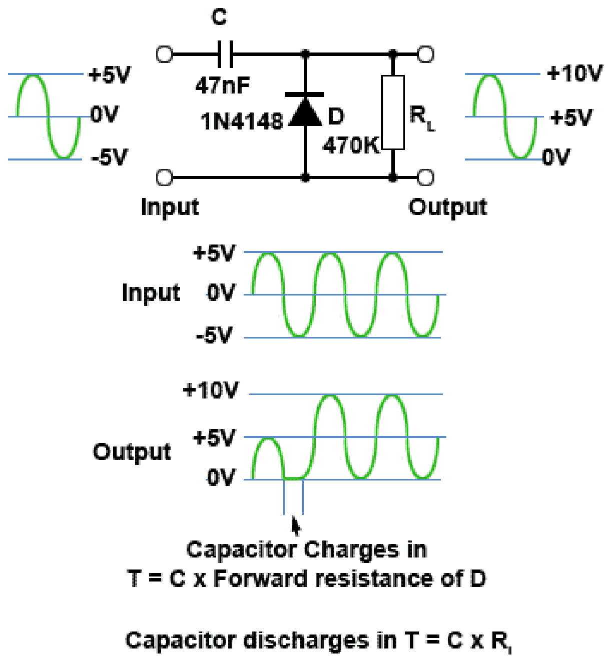


Figure 2.3.8. How Clamping Works

How Clamping Works

Initially suppose that the capacitor C is fully discharged and has 0V on both plates.

As the input waveform goes positive the changes in voltage occurring at the left hand plate are passed through to the right hand plate. As this plate also goes positive, the diode D1 will be reverse biased and will not conduct, and so far plays no part in the action.

However the first time the input wave goes negative at the end of the positive half cycle of the wave, so does the left hand plate of C and the right hand plate tries to 'follow' it. The anode of the diode cannot go negative however as the diode immediately begins to conduct, holding the potential on the right hand plate of C at 0v.

The left hand plate of C continues to go negative, due to the input wave, and C begins to charge at a rate that depends on the time constant T = C x the forward resistance of the diode, which will be very low, so giving a very short time constant. When the input wave reaches its most negative value, C is charged to a value that is the same as negative peak value of the wave (e.g. -5V).

Once the value of the input wave begins to rise back towards zero however, the voltage on the left hand plate of C begins to rise, and the potential and the right hand plate follows it. This causes the right hand plate of C to immediately go positive, so the diode stops conducting, and by the time the input wave has returned to 0V, the voltage on the right hand plate of C (and therefore on the circuit output) has a positive value equal to half the peak to peak value of the wave. In order for this to happen the time constant formed by C and R_L must be considerably longer than half of the periodic time of the input wave, so that very little of the charge on C leaks away during the time the diode is turned off.

Provided that during following cycles, only a little of the DC voltage on the output is allowed to leak away, each time the input wave goes negative and returns to zero, the capacitor will charge sufficiently through the diode to hold the DC level of the output wave constant. Therefore the action described above sets the DC level of the output such that the negative peaks of the wave are always at 0V, i.e. the negative tips of the waveform are CLAMPED to zero volts.

For this reason the circuit is called a 'Clamping Circuit'. The two time constants created by the diode in Fig. 2.3.8 switching on (short time constant) and off (long time constant) fulfil the necessary criteria to produce the shift in level of the output waveform that can be seen in the Spice analysis shown in Fig. 2.3.9

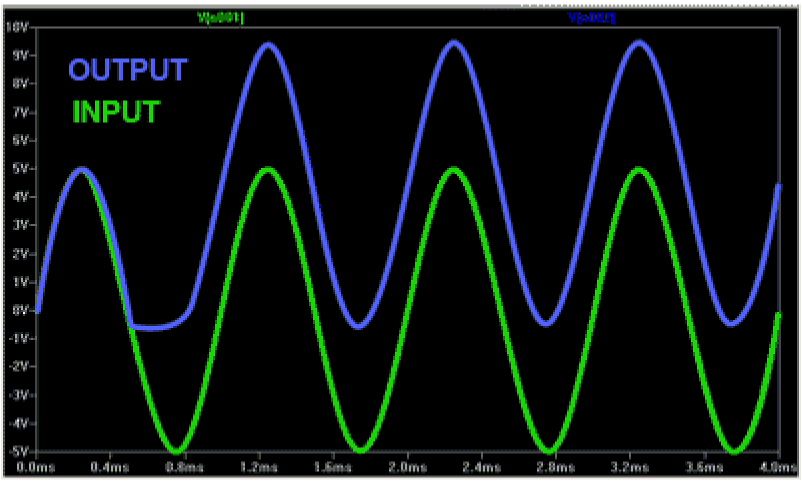


Figure 2.3.9. Spice Circuit Analysis of Fig 2.3.8

It also may be that a waveform needs to be clamped to some other value than the negative tips described above. To clamp the positive tips of the waveform to 0V so that the whole waveform is negative only requires that the diode is connected the opposite way round, with its cathode connected to 0V.

If a clamping level, and therefore a DC level of the wave of some other value is required, this can be easily arranged as shown in Fig. 2.3.10 by connecting the anode of the diode to some potential.

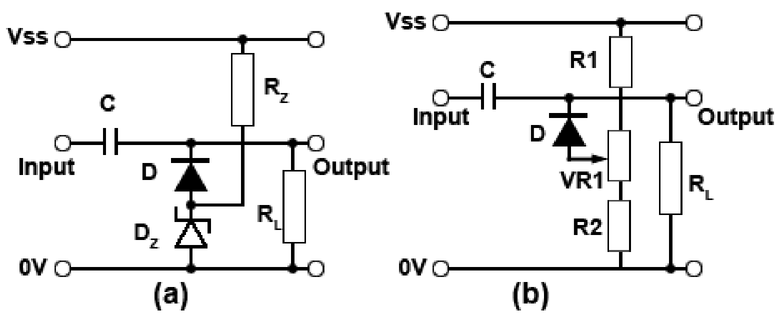


Figure 2.3.10. Changing Clamping Levels

Fig. 2.3.10(a) shows how a zener diode may be used to obtain an accurate clamping level and Fig. 2.3.10(b) shows how a variable clamping level may be obtained from a potential divider network.

Diodes for High Frequency Applications

Diodes used for high frequency applications need to have as low a junction capacitance as possible and a very short Reverse Recovery Time (t_{rr}) to enable them to switch from conducting to non-conducting mode and back again as quickly as possible. Small signal diodes can generally be divided into fast and ultra fast diodes. Fast types, such as the popular 1N4148 have a t_{rr} parameter of 4 to 8ns, making it suitable for use on VHF signals having a frequency of up to about 100MHz. Ultra fast diodes, which are usually Schottky designs can have t_{rr} figures measured in pico seconds, making them suitable for use up to UHF and microwave frequencies. Both fast and ultra fast diodes are physically very small and have very small junction areas. Schottky diodes in particular also have a low junction potential of around 0.2V (similar to the older germanium diodes) compared with around 0.6V for silicon PN junction diodes, this makes Schottky signal diodes particularly suitable for RF applications such as demodulation, where the audio or video signal is retrieved from the radio frequency carrier wave.

Protection Diodes

Diodes are used in many circuits to protect other semiconductor devices against excessive voltages and sudden voltage spikes present on external power supplies, or back emf spikes caused when inductive loads such as motors and relays are switched off. Fig. 2.3.11 shows a typical application where three BAV99 Dual diodes are being used to protect a HDMI input against over voltage or spikes.

In the circuit shown in Fig.2.3.11 each of three digital input lines from a HDMI socket is being prevented from attaining a voltage higher than +5V as pin 2 of each BAV99 is connected to the +5V supply, so that if a voltage higher that +5V occurs on any of the relevant HDMI input pins (13,15 or 16) connected to pin 3 of each BAV99, the upper diode will conduct and clamp the input line at 5V (plus the forward junction potential which, according to the data sheet for the BAV99 can vary between 0.715V and 1.75V, depending on the forward current). Similarly any voltage on the input line that falls below zero volts will be clamped by the diode between pins 1 and 3.

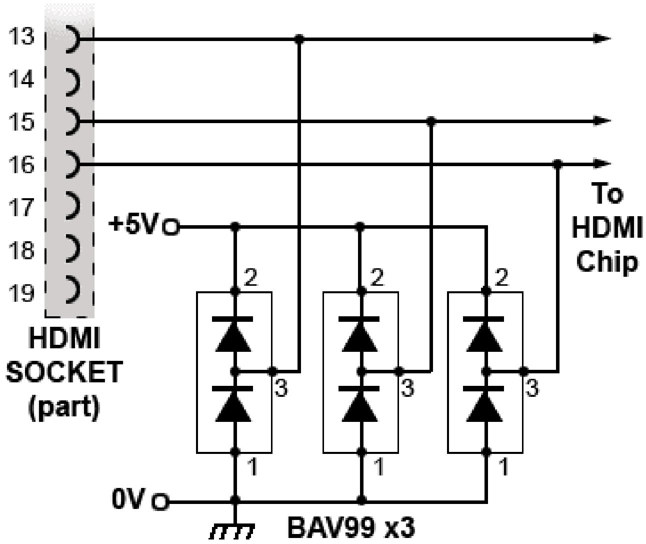


Figure 2.3.11. Protection Diodes

Other methods and applications for using protection diodes are discussed in relevant modules on Learnabout-electronics, click any of the links below to find out more.

- Back emf Protection.
- Protection diode for voltage regulators.
- Anti-static protection for CMOS gates.
- Driving motors with a 555 timer.

Module 2.4
Zener Diodes

What you'll learn in Module 2.4

After studying this section, you should be able to:

Describe typical Zener Diode construction.

Describe the Zener effect.

Describe the Avalanche effect in Zener Diodes.

Describe typical Applications for Zener Diodes.

- Shunt voltage regulation.
- Series voltage regulation.

Calculate appropriate values for current limiting resistors for Zener Diodes.

Zener Diode Construction

Zener diodes are a modified form of PN silicon diode used extensively for voltage regulation. The P type and N type silicon used is doped more heavily than a standard PN diode. As shown in Fig. 2.4.1, this results in a relatively thin junction layer, and consequently a reverse breakdown voltage that can be much lower than in a conventional diode. The actual breakdown voltage is controlled during manufacture by adjusting the amount of doping used. Breakdown voltages can be selected in this way to occur at precise preset values anywhere between about 3V and 300V. Zener diodes can also withstand higher reverse current flow than comparable PN diodes, and are available with various power ratings, typically from 500mW to 50W.

When Zener diodes are biased in their forward direction, with the anode voltage higher than the cathode, they behave in the same way as a normal silicon diode. When they are reverse biased they exhibit a very high resistance, and consequently a low value of reverse leakage current. However when a reverse bias reaches the value of the diode´s reverse breakdown voltage (the Zener voltage) a rapid drop in resistance and increase in current occurs. To prevent this current increasing to a value that would exceed the diode´s power rating and destroy it, the Zener diode uses a resistor connected in series with the diode to limit the reverse current to a safe value.

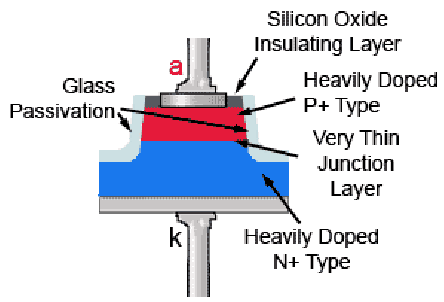


Figure 2.4.1. Zener Diode Construction

Operating the diode in this condition means that, due to the very steep slope of the diode´s reverse characteristic, any slight change in voltage across the diode will cause a large change in current through the diode. This effect is very useful in voltage regulator circuits, as explained in our Power Supplies Modules 2.1(Shunt Voltage Regulators) and 2.2 (Series Voltage Regulators). Zener diodes are also useful for providing an accurate reference voltage for purposes such as waveform clamping. This rapid increase in reverse current in Zener diode operation is due to one or both of two effects:

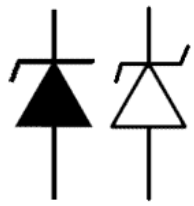


Figure 2.4.2. Alternative Zener Diode Symbols

1.The Zener Effect

2.The Avalanche effect.

1. The Zener Effect

Because of the heavily doped P and N materials at either side of the junction, which are therefore good conductors, and the very thin depletion layer, the electric field strength across the depletion layer is very strong, and it becomes relatively easy, even at low voltages, for holes and electrons to cross the depletion layer and combine to create a reverse current. This effect mostly happens in Zener diodes with a low reverse breakdown voltage, typically 5 to 6V or less and leads to a gradual, rather than a sudden increase in reverse current.

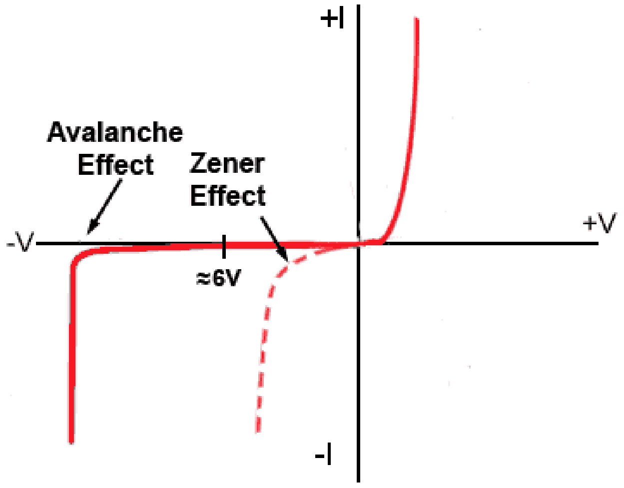


Figure 2.4.3. Zener & Avalanche Effects

2. The Avalanche Effect

In Zener diodes with wider depletion layers and therefore higher breakdown voltages, the increase in current at the breakdown voltage is much more sudden, giving an abrupt reduction in the reverse resistance of the diode and a nearly vertical region to the diode's reverse current characteristic. This effect happens mainly in diodes with a higher reverse breakdown voltage (above about 5V) and less heavily doped P and N regions. Below the reverse breakdown voltage, although only a small reverse leakage current is flowing, some current does flow and therefore electrons and holes are entering the depletion layer. As the reverse voltage approaches the reverse breakdown voltage the electrons and holes entering the depletion layer come under the effect of a strong electric field and are rapidly accelerated. In this accelerated state they begin to collide with other atoms and knock electrons from their atomic bonds in a process called 'impact ionisation', so creating more electron/hole pairs that are also greatly accelerated by the electric field. These secondary current carriers in turn ionise other atoms, creating a very rapid increase in reverse current through the diode. This process is called 'Avalanche Breakdown'

Practical Zener Diodes

Practical zener diodes may use either the zener effect or the avalanche effect and in some diodes both effects can also occur at the same time, but it is common practice to call all these diodes zener diodes. Both the zener and avalanche effects are also dependant to some degree on the junction temperature of the diode. However, whist the current in a purely zener diode has a negative temperature coefficient i.e. the current reduces with an increase in temperature, the opposite effect occurs in a diode using the avalanche effect. Therefore it is possible to manufacture zener diodes that use both effects and so these temperature effects tend to cancel each other out, producing diodes that have very minimal current variation due to temperature.

Zener diodes are widely used in power supply circuits for both voltage regulation and over voltage protection, how they are used is discussed in much more detail in our Power Supplies Module 2.1.

Module 2.5
Light Emitting Diodes

What you'll learn in Module 2.5

After studying this section, you should be able to:

Describe typical construction methods for LEDs.

Understand the operation of LEDs.

- Coloured LEDs
- Methods for producing White LEDs.
- Current limiting for LEDs.
- Multiple LED arrays.

Describe methods for testing LEDs.

Figure 2.5.1. LEDs

Light Emitting Diodes (LEDs)

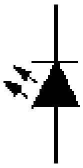


Fig. 2.5.1 shows a range of LEDs illustrating some of the wide range of styles and sizes of LEDs available. Colours range across the visible light spectrum from deep red to ultra violet as well as shades of white. Additionally infra-red LEDs are used in many sensors and remote control applications.

From left to right the LEDs in Fig. 2.5.1 are 5mm warm white, 10mm ultra high brightness blue, standard 5mm red & green, miniature yellow & green, tri-colour (red/green/blue), infra-red opto-coupler, infra-red transmitter/receiver, and an infra-red opto-isolator. Below is a 230V 8W 230 lumens warm white LED light bulb and a 7-segment display.

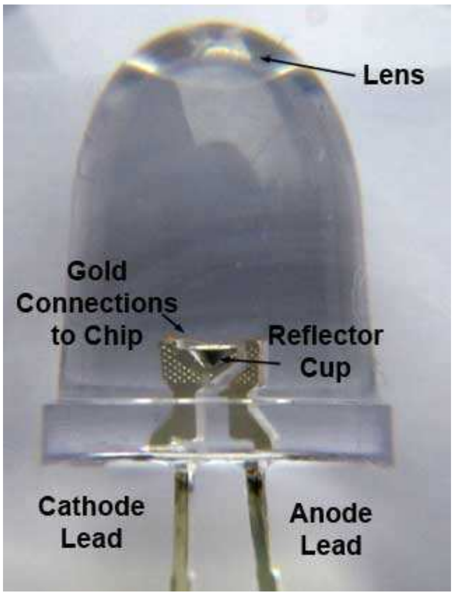


Figure 2.5.2. LED Construction (Side View)

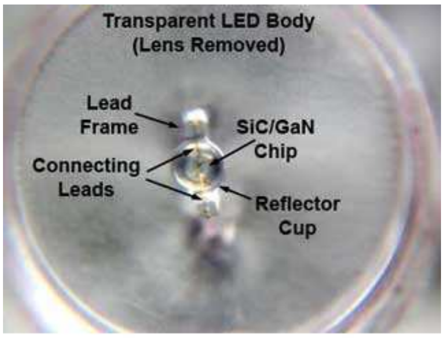


Figure 2.5.3. LED Construction (Top View)

How LEDs Work

In semiconductor diodes whenever an electron recombines with a hole, energy is released for a brief moment in the form of a photon. Ordinary silicon diodes are not suitable for light emission as in the silicon PN junction, the photons produced are mostly converted to heat within the silicon, and only a very small amount of light can escape the diode structure. This light also has a wavelength limited to the infrared region. For several decades, light emitting diodes have used materials such as gallium arsenide(GaAs), gallium arsenide phosphide(GaAsP), or gallium phosphide(GaP), which make PN junctions more efficient at producing light. These compound materials also have carefully controlled amounts of indium(In) or aluminium(Al) added and can be doped with other elements such as magnesium(Mg). This enables the production of the more common LED colours of red, orange, yellow and green. Blue LEDs are now also made possible by using silicon carbide (SiC) and gallium nitride (GaN). The colour and brightness of an LED depends on the combination of materials used and the energy gaps of the P and N materials on either side of the junction.

The energy gap (the amount of energy needed to move an electron from the valence band of an atom into its conductance band) of the semiconductor material on ether side of the PN junction is different in different semiconductor materials, and as current flows through the LED, electrons in the higher energy band recombine across the junction layer with holes in the lower energy band. In doing so the electrons lose some energy and it is this energy that is emitted by the LED as light.

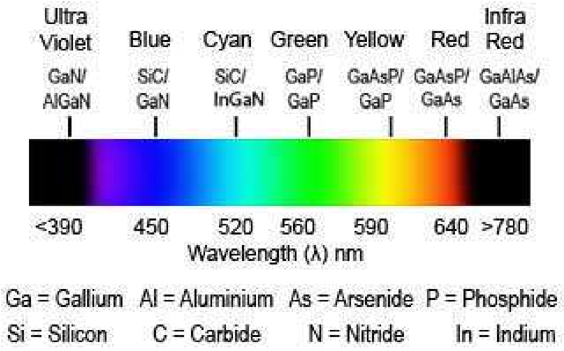


Figure 2.5.4. LED Colours

The more energy the electrons lose in this process, the higher the frequency (and the shorter the wavelength) of the light produced. Fig. 2.5.4 illustrates the different combinations of semiconductor materials used to produce light of different colours.

Generally the light resulting from each electron/hole recombination is very directional and short lived, but the millions of recombinations happening when the LED is forward biased produce light continuously.

Because the light is emitted in narrow beams at many angles, to make this scattered light more useful, the LED chip is mounted at the focus point of a reflector cup, which focuses the light emitted by the chip into a cone shaped beam.

The clear plastic body of the LED also contains a lens to better focus the light into a beam. Some LEDs use a coloured plastic body, generally red, yellow or green, but the body colour is only to identify the colour of the LED in its unlit state and makes little or no difference to the colour emitted by the LED chip. The range of different colours available from LEDs is illustrated in Fig. 2.5.4

Maximising Light Output

A problem with traditional LEDs however, is that the amount of light leaving the LED chip might only be around 20% of the actual photons generated within the chip. The reason for this is that whilst a beam of light (a photon) approaching the chip wall at an angle perpendicular to the surface, passes easily from the chip into the surrounding medium (e.g. the transparent plastic of the LED body), light approaching the chip surface at other angles is deflected by refraction as it meets interface between the chip and the surrounding material. This is due to a change in the speed of light between the different materials. When a beam of light arrives at the interface of of the chip and the surrounding plastic at an angle greater than the ‘critical angle’ for the two materials concerned, it is reflected back into the chip, where the photon energy is dissipated as heat.

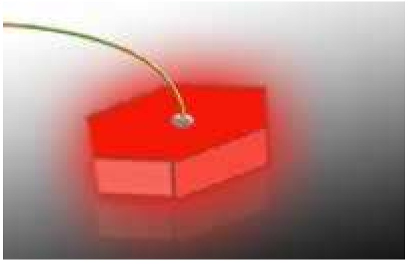


Figure 2.5.5. Hexagonal LED Chip

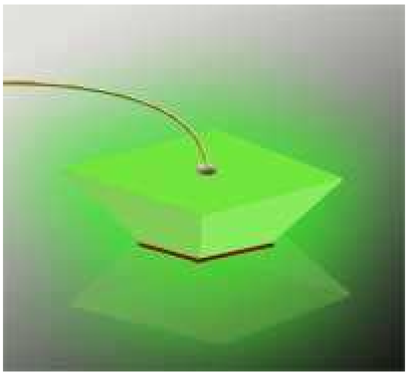


Figure 2.5.6. Pyramid LED Chip

To overcome this problem and increase the light output from the chip as well as reducing the heat produced during operation, a number of manufacturers are producing LED chips that do not have the regular rectangular shape, by cutting the individual chips into polygons instead of rectangles as shown in Fig. 2.5.5.

Another approach is to cut the sides of a rectangular chip at an angle, forming a partial pyramid as shown in Fig. 2.5.6. These techniques increase the chances of internally reflected photons arriving at another chip surface at an angle that allows them to pass through the surface, so increasing light output.

By replacing sides that are at right angles to each other with sides at various angles, the chance of a photon emitted from the LED junction at any random angle arriving at the boundary of the chip at an angle that allows it to leave the chip rather than being reflected internally is increased, as illustrated in Fig. 2.5.7. In this way light extraction is increased and internal heat generation is decreased, allowing for the manufacture of more efficient LEDs.

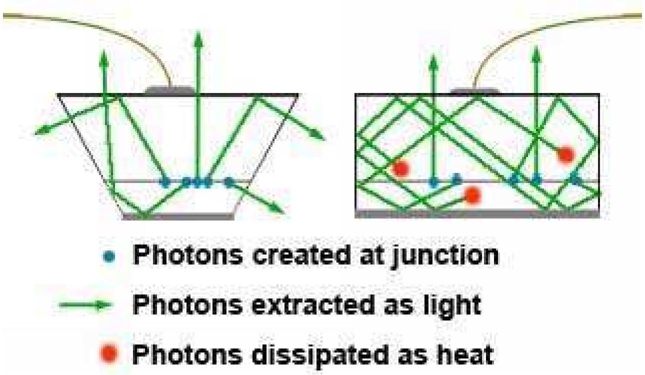


Figure 2.5.7. Internal Reflection

White LEDs

Three main types of white LEDs are available, the first is actually a blue LED in which the emitted light activates a yellow phosphor coating to give the effect of white light, however the white light produced by this blue/yellow combination, unlike sunlight does not have an even energy spread over the whole light spectrum so is not particularly suitable for accurate colour matching applications.

A second type of white LED actually comprises three LEDs (red green and blue) in a single package. This type gives a more even light spectrum and some versions allow for each of the three individual colours to be independently varied. This is an important feature because one problem with LEDs is that their output can vary with temperature, and red, green and blue LEDs do not all vary by the same amount for a given temperature change. These LEDs therefore require more complex (temperature sensitive) control circuits if a pure white is to be maintained, however this system also gives the opportunity to create variable multi-colour lighting.

A third approach is to use an ultra-violet LED to stimulate a mixture of phosphors designed to give a white light covering the full visible spectrum without the ‘gaps’ left by the blue/yellow system and, because only one (invisible) LED is involved, overcoming the temperature stability issues of the three LED system.

LED Colour

LEDs currently cover a wide range of colours, light outputs and power requirements and are generally many times more reliable and use much less power than competitive incandescent or fluorescent alternatives.

To make a LED glow it is necessary to forward bias the diode sufficiently to pass an appropriate amount of current, generally the shorter the wavelength of the light produced, the higher the voltage required for forward bias, and typical forward bias voltages vary from 1.5 to 1.7V for infrared LEDs to 3.3V or more for blue and ultraviolet. The commoner red yellow and green versions require around 2V forward bias and white LEDs about 3.6V.Fig. 2.5.8 shows typical characteristics for various colours of LED. Notice the considerable difference between the forward voltage (V_F) for blue types and red to green types. Blue LEDs also typically have a greater reverse leakage current (I_{REV}) than other LEDs but the safe limit for most LEDs is considered to be around -5V, a VERY LOW value compared with silicon diodes, which may have reverse breakdown voltages measured in tens or hundreds of volts. Therefore LEDs are more easily damaged by relatively small values of either excess forward current, or reverse voltage compared to ordinary silicon diodes, so to take advantage of the LED´s excellent reliability, care must be taken to ensure that any LED operates within its safe operating area.

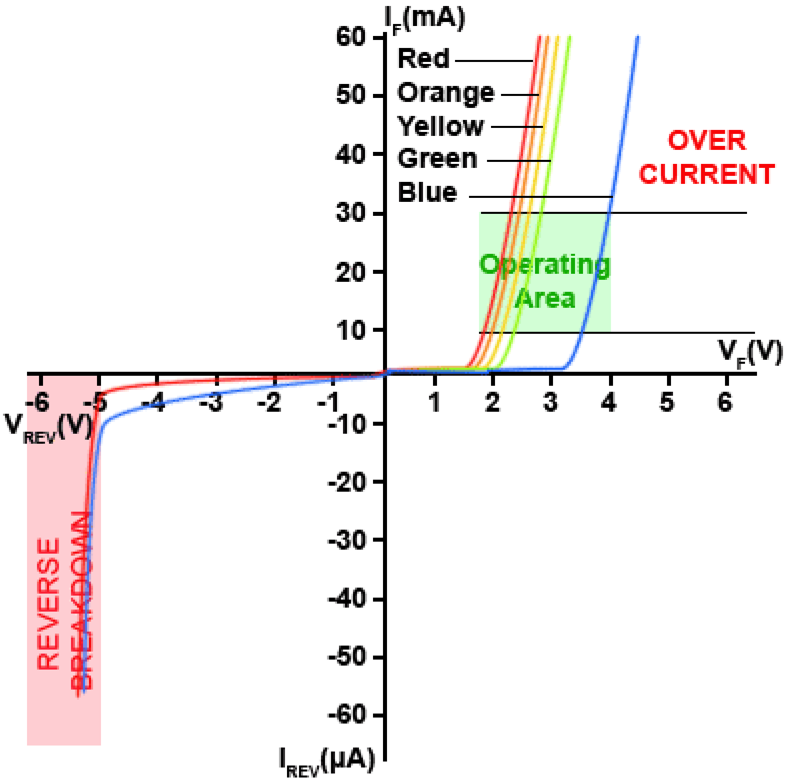


Figure 2.5.8. LED Characteristics Compared

Connecting LEDs

To ensure that the forward current through the LED is correct for the type of LED used, it is best to consult the relevant data sheet. If the appropriate data sheet is not available good and comprehensive data on LED's is available on the web sites of specialist manufacturers such as Kingbright.

Having found an appropriate figure for forward current, a current limiting resistor (R_{LIM}) such as that shown in Fig. 2.5.9 can easily be calculated by subtracting the appropriate forward voltage for the LED (V_F) from the supply voltage (V_S) to give the required voltage across the resistor (V_R), and then dividing V_R by the required forward current (I_F). It will be unlikely that the result of your calculation will be a preferred value of resistor, in this case choose the next higher preferred value. The wattage required for R_{LIM} should be the next higher power rating available greater than the power calculated by R_{LIM} x I_F.

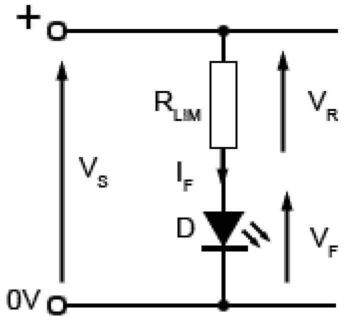


Figure 2.5.9. LED Current Limiting Resistor

$$R_{LIM} = \frac{V_S - V_F}{I_F}$$

LED Arrays

LEDs are often used in multiple arrays as shown in Fig. 2.5.10 and a typical method of connection is to connect a number of LEDs in series, fed via a single limiting resistor from a higher voltage supply(V_S) than would be required for a single LED. The current through each series connected LED is identical to that required for a single LED but the the voltage across the four LEDs in Fig. 2.5.10 is four times that required for a single LED (i.e. V_{F1} + V_{F2} + V_{F3} + V_{F4}).

A number of identical series groups may be connected in parallel as shown in Fig.2.5.10. There are several advantages to this method of connection:

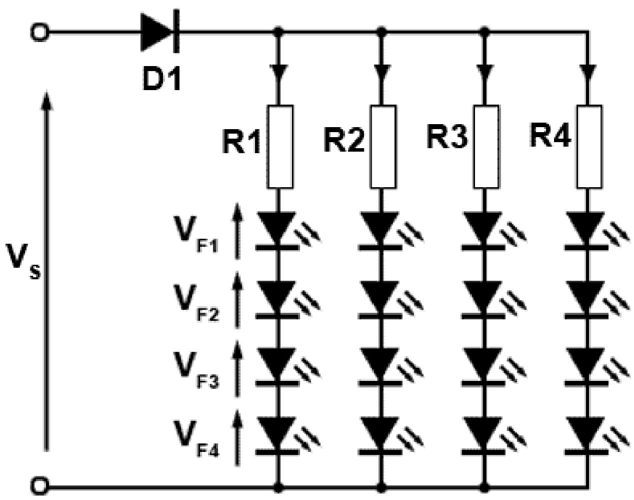


Figure 2.5.10. Series-Parallel LED Array

1. The supply voltage does not need to be as high as if many LEDs were connected in series, making this method more suitable for battery supplies.

2. A smaller number of limiting resistors are needed, only one per series group rather than one per LED, compared to an all parallel connection.

3. The effect of a faulty LED on the total light output is reduced. If one LED develops a short circuit only that LED will fail to light, however the remaining LEDs in the four diode series group will experience a 25% increase in current. With more LEDs in the group this effect would be reduced. If any one of the diodes develops an open circuit, then only the four LEDs in the affected series group will fail to light, the rest of the LEDs will work normally.

D1 might typically be included in a battery-powered circuit to prevent damage from any accidental reverse polarity battery connection.

Testing LEDs

Correct polarity must be observed when connecting LEDs and unfortunately the anode and cathode connections of LEDs are not always readily apparent. Fig. 2.5.11 illustrates the two ways that polarity is indicated on standard 5mm LEDs but these depend on the connecting leads not being shortened (as they often are) or being able to see the slight ‘flat’ near the cathode lead, which is not always easy. With larger clear LEDs it may be possible to determine the cathode of the device by looking at the internal structure of the LED. In this case the larger of the two internal lead structures is the cathode (See Fig 2.5.2).

As LEDs are available in so many shapes and sizes it is often not possible to determine which is the anode and which the cathode connections visibly.

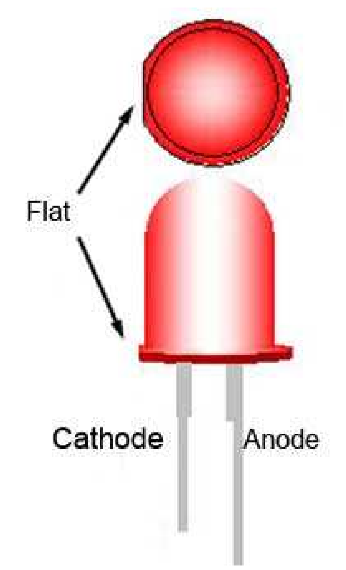


Figure 2.5.11. LED Polarity

To overcome this problem there are a number of testing devices on the market varying in price from less than $10 to around $160 but simple (and cheaper) tests can be carried out with a basic multi meter switched to the **diode range** that will not only reveal the polarity of the diode but also show whether the LED is faulty or not. In Fig. 2.5.12 a typical problem is shown where a miniature green LED that has been mounted in a wired holder, needs testing.

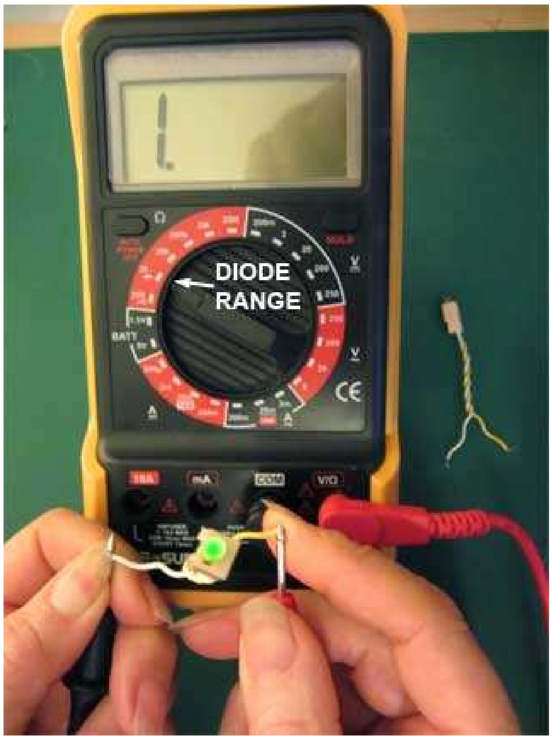


Figure 2.5.12. Testing a LED

Simply connecting the multi meter leads is usually sufficient to make the LED glow (often weakly) when the red lead is connected to the anode and black to cathode. Reversing the leads will not make the diode glow, so revealing in this case, that the yellow lead on the LED holder, is connected to the anode. Notice however that the meter is still showing 1 on its display indicating that the LED is open circuit when it clearly is not because its working!

Some LEDs will fail to light in this test, no matter which way round the meter leads are connected, suggesting an open circuit diode but at the same time will give an infinity (1 on the display) reading in one direction, indicating a very high resistance, and a reading of perhaps several hundred or perhaps slightly more than 1KΩ in the other direction, indicating a good LED. The results depend on both the characteristics of the diode and on the meter used, as well as on the condition of the meter battery. If all this seems confusing, just assume that when testing an LED out of circuit:

- If **either** testing an out of circuit LED for a glow, **or** for different resistance readings as described above indicate a good LED, then it very probably is a good LED.
- If **both** tests suggest a faulty LED then the LED is very probably faulty.

Module 2.6
Laser Diodes

What you'll learn in Module 2.6

After studying this section, you should be able to:

Describe LASER light.

Describe LASER diode operation at the atomic level.

- Junction p.d.
- Describe precautions needed for LASER operation.
- Recognise Safety labelling used on LASER equipment.

Laser Light

White light is made up of all the colours of the visible light spectrum, which is a very wide band of many different frequencies. Ordinary LEDs give a light output often consisting of one colour, but even that light contains electromagnetic waves covering quite a wide band of frequencies.

Any light, such as white light, that contains multiple frequencies or wavelengths, is difficult to focus to a very fine point. This is because the lens system focussing the light has a fixed focal length, but the focal length required to focus various wavelengths (colours) of light is different. Therefore each colour will focus at different points, causing what is called 'chromatic aberration'. This can be seen, even in good quality lens systems as coloured fringing around images seen through the lens.

The light from a laser contains only a single frequency and therefore can be focussed by even a simple lens system to an extremely small point. There is no chromatic aberration since only one wavelength exists; also all of the energy from the light source is concentrated into a very small spot of light. Typically the diameter of the focussed light beam from a laser such as that found in a CD player would be about 1.6µm (less than two thousandths of a millimetre). This means that if the laser diode output were only 0.5mW the focussed (infra red) power of the beam (allowing for losses in the focussing lens) would be about 12kW/cm^2.

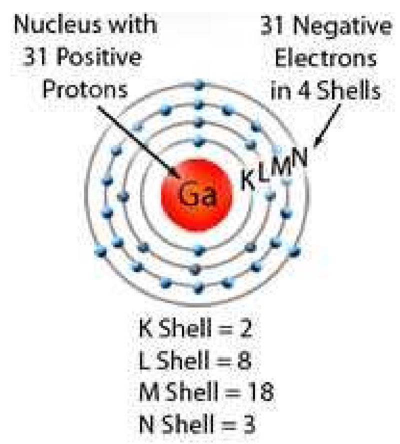


Fig. 2.6.1. An Atom of Gallium

Low power Lasers are used in an increasing number of familiar applications including CD and DVD players and recorders, bar code readers, security systems, optical communications and surgical instruments. LASER is an acronym (a bunch of initials made into a word) for:

Light **A**mplification by the **S**timulated **E**mission of **R**adiation

The name gives a description of how the device works, so long as a few basic principles are understood.

The Basics of an Atom

Everything in the universe is made up from only about 100 different atoms, either in a pure form (an element), or in an unlimited number of combinations (compounds and mixtures). Atoms are constantly in motion. They continuously vibrate, move and rotate. Even solid materials are actually in motion. This motion is called excitation. Atoms can be in different states of excitation. In other words, they can have different levels of energy. If a lot of energy is applied to an atom via heat, light, or electricity, it can leave what is called the ground-state energy level and go to an excited level. The level of excitation depends on the amount of energy applied to the atom.

Fig. 2.6.1 is a simplified diagram of an atom. It consists of a nucleus (containing protons and neutrons) surrounded by an electron cloud, although in this simplified electronics model the neutrons are ignored as they have no electric charge, and it´s also helpful to think of the electrons orbiting the nucleus in several fixed orbits. Although other models of the atom do not depict discrete orbits for the electrons, it can be useful to think of these orbits as the different energy levels of the atom. In other words, if some heat is applied to an atom, some of the electrons in the lower energy orbits would jump to higher energy orbits further away from the nucleus. This is a highly simplified view of things, but contains the basic idea of how atoms work in terms of lasers.

As an atom absorbs energy and some of its electrons jump to a higher-energy orbit, it eventually wants to return to the ground state. When it does, each electron releases its energy as a photon − a particle of light. Atoms can be seen releasing energy as photons all the time. For example, when metal glows red hot, the red light seen is the atoms of the hot metal releasing red photons. When looking at a picture on a TV screen, what you are seeing is phosphor atoms that coat the screen, or the surface of white LEDs being excited as electron of high energy atoms releasing photons, and directly or indirectly producing different colours of light. Many devices produce light in this way − fluorescent lights, neon signs, LED street lighting and even traditional incandescent light bulbs, all emit light through the action of electrons changing orbits and releasing photons.

Laser Diode Construction

There are several variations of construction used for laser diodes, each aimed at achieving the maximum efficiency for converting electric current into laser light.

Fig. 2.6.2 shows a simplified construction for a laser diode, which in this case is similar to a light emitting diode (LED) in that it uses gallium arsenide, doped with elements such as aluminium, silicon or selenium to produce P type and N type semiconductor materials. However a laser diode has an additional active layer of un-doped (intrinsic) gallium arsenide only a few nanometers thick, sandwiched between the P and N layers, effectively creating a PIN (P type/Intrinsic/N type) diode. It is in this layer that the laser light is produced.

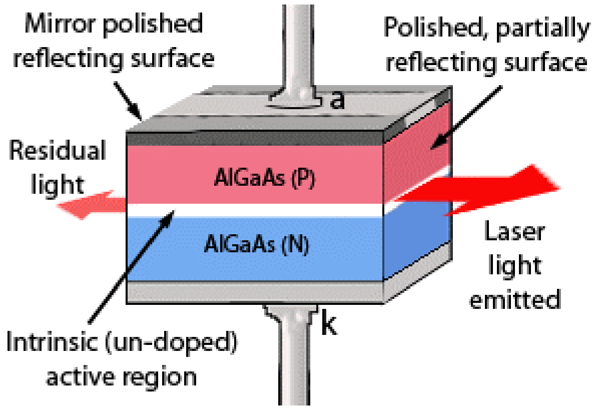


Fig. 2.6.2. Laser Diode Construction

Laser Diode Action

The laser diode passes a large amount of forward current from P to N. Much greater than that used in a LED as the Laser diode will only produce laser light when operated at above about 80% of its maximum current.

Laser Pumping

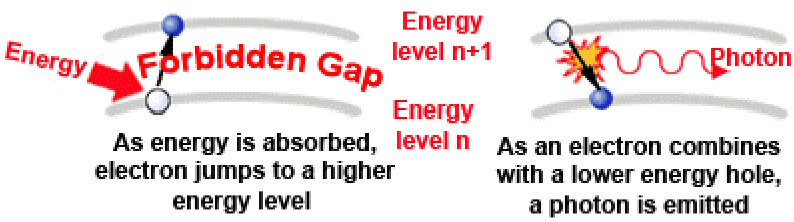


Fig. 2.6.3. The Absorption of Energy

Under these conditions the atoms are in a highly energised (pumped) state, and as charge carriers (electrons and holes) enter the active layer at the PN junction. Electrons are at an energy level higher than that of the holes, and as the electrons and holes re-combine energy is lost in the form of photons. The photons produced, all oscillate at a particularly precise frequency as they ‘bounce’ up and down between the light reflective walls of the active layer. Some photons collide with other atoms and so create additional energised electrons that produce even more photons. This process is called ‘pumping’ and increases the number of highly energised electrons until there are more electrons in the excited state than in the un-energised ‘ground’ state. At this point, known as population inversion, a constant stream of coherent or laser light is produced, as the extra excited photons more than make up for any losses due to photons being re-absorbed within the semiconductor material. Because the photons oscillate at a single precise frequency they produce laser light that has only a single wavelength.

The Resonant Cavity

As the photons increase in number, the light increases in power. Although some light escapes in different directions or is absorbed within the semiconductor material, some of the photons run in a direction parallel to the laser’s axis, as shown in Fig. 2.6.4 these bounce back and forth off the ends of the laser material. The end surfaces are very accurately cut and polished to create parallel reflecting mirrors. The distance between these reflecting surfaces is an exact multiple of one wavelength, so that as the light waves (the photons) reflect from each end of the cavity, they stay in phase. The amplitude of the reflected wave adds to the amplitude of other waves within the cavity, so the waves keep adding as they bounce backwards and forwards between the mirrors. In this way, the active layer forms a ‘Resonant Cavity’ that aids the amplification of the light. As the photons pass through the crystal laser material, they also stimulate emission in other atoms. As a result, amplified, monochromatic, single-phase light leaves the resonant cavity of the laser through the partially reflecting mirror.

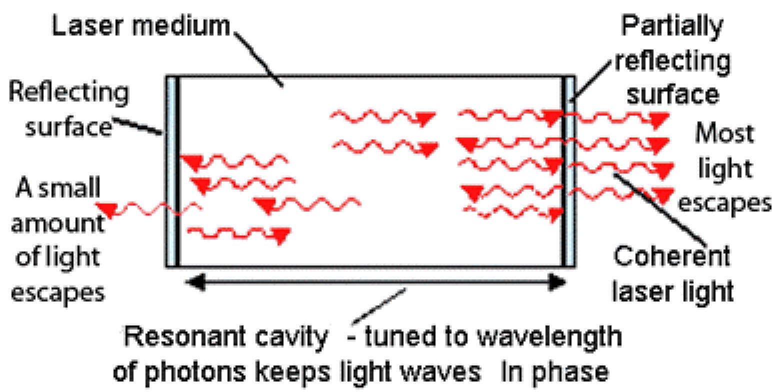


Fig. 2.6.4. The Resonant Cavity

Controlling the Laser Diode.

A laser diode is basically a LED producing laser light; to do this the laser diode is operated at a much higher current, typically about 10 times greater than a normal LED. Fig. 2.6.5 compares a graph of the light output of a normal LED and that of a laser diode. In a LED the light output increases steadily as the diode current is increased. In a laser diode however, laser light is not produced until the current level reaches the threshold level, when stimulated emission starts to occur. The threshold current is normally more than 80% of the maximum current the device will pass before being destroyed! For this reason the current through the laser diode must be carefully regulated. Another problem is that the emission of photons is very dependent on temperature, the diode is already being operated close to its limit and so gets hot, therefore changing the amount of light emitted (photons) and the diode current. By the time the laser diode is working efficiently it is operating on the brink of disaster! If the current reduces and falls below the threshold current, stimulated emission ceases; just a little too much current and the diode is destroyed.

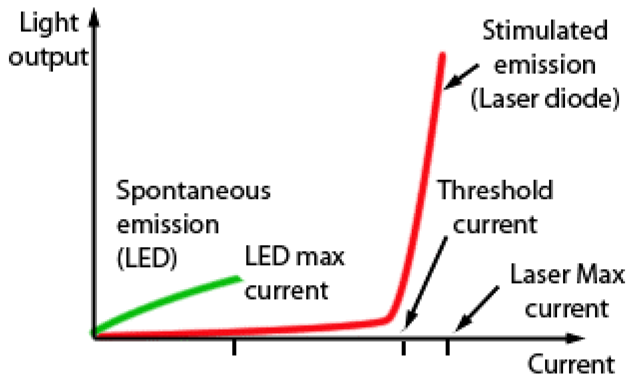


Fig. 2.6.5. Comparison between a LED and a Laser diode

As the active layer is filled with oscillating photons, some (typically about 60%) of the light escapes in a narrow, flat beam from the edge of the diode chip. As shown in Fig 2.6.6 some residual light also escapes at the opposite edge and is used to activate a photo diode, which converts the light back into electric current. This current is used as a feedback to the automatic diode driver circuit, to measure the activity in the laser diode and so make sure by controlling the current through the laser diode, that the current and light output remain at a constant and safe level.

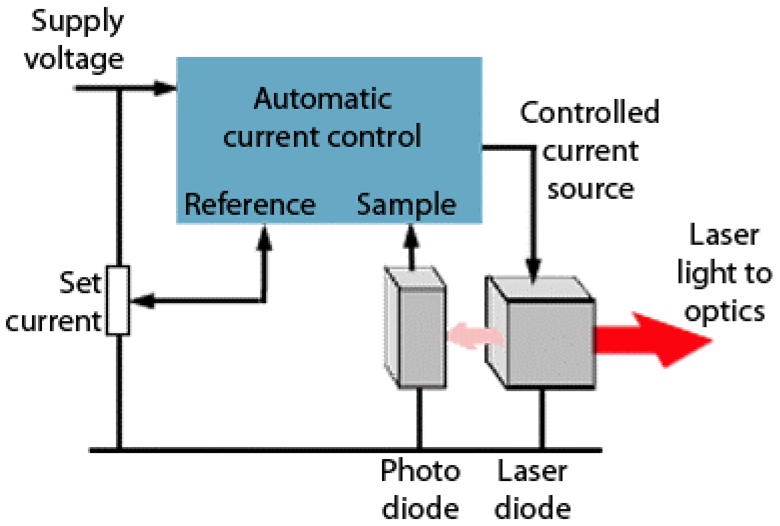


Fig. 2.6.6. Controlling the Laser Diode

Laser Module

For these reasons, laser diodes are rarely used on their own; they are normally supplied as a laser diode module, which contains:

- The diode itself.
- A photodiode light sensor.
- A current regulation circuit.
- A collimating lens.

Fig. 2.6.7 Typical Laser Module

The laser diode module is a self-regulating circuit that senses its own light output and automatically regulates the supply current and temperature to keep the diode operating in the critical conditions where laser light is produced.

Optical Correction

The beam of light produced by the laser diode still needs some modification to change it from an elliptical, spreading beam produced as the laser light leaves the thin active layer of the diode, into a circular parallel beam. This process is carried out by an optical device called a collimating lens (Fig. 2.6.8). This may be a simple spherical lens or an aspherical type, which can convert an elliptical beam into a circular one.

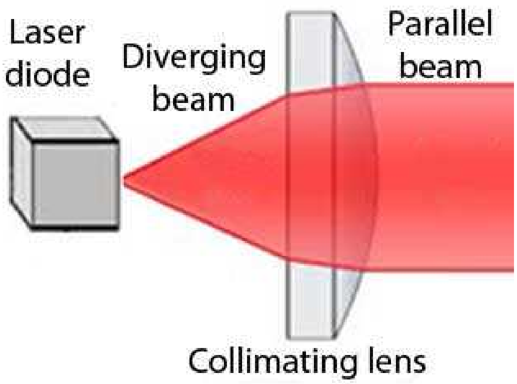


Fig. 2.6.8 Collimating Lens

Laser diode modules are available complete with optics and electronics. A typical module would feature facilities such as built-in power stabilisation, built-in slow start and heat sink. Low power modules (Class 2) are used in laser pointers, bar-code readers, sighting, levelling and positioning equipment, as well as a wide range of educational and laboratory uses. They produce a continuous wave, rather than a pulsed wave output, at various wavelengths between 500nm and 900nm, and so have different colours of laser light (green at 532nm and red at 650nm) as well as infra-red and near infra-red. They are simple to use, requiring typically a 3V to 5V DC supply to operate.

Laser Classifications

Lasers are classified into four broad areas (plus sub areas) depending on the potential for causing biological damage. When you see a laser, it should be labelled with appropriate class designation, briefly described below:

• **Class 1** − Safe under conditions of normal use. Output is restricted to less than 0.39mW at 600nm (less for shorter wavelengths).

• **Class 1M** − Lasers producing divergent beams, safe for all conditions of use, except when passed through magnifying optics such as microscopes and telescopes.

• **Class 2** − Low power visible lasers that emit above Class 1 levels but at a radiant power not above 1mW. The concept is that the human aversion reaction to bright light (the blink reflex) will protect a person. This class is used for laser pointers.

• **Class 2M** − Similar to Class 2 but can allow more power as beams with this classification must produce wide or diverging beams. Light passing through a viewer's pupil must not be greater than that allowed in class 2.

• **Class 3R** − Moderate power lasers below 5mW, considered safe where viewing is restricted, giving low risk of injury.

Fig. 2.6.9 Typical Laser Warning Stickers

• **Class 3B** − High-power pulsed lasers up to 500mW: Hazardous to view (except as reflected light from a matt surface) significant controls such as protective eyewear and safety interlocks are required of Class 3B laser facilities.

• **Class 4** − High power lasers greater than 500mW. They can burn the skin, and cause potentially devastating and permanent eye damage as a result of direct or diffuse beam viewing. They may also ignite combustible materials, and thus may represent a fire risk. This classification includes many industrial, scientific, medical and military lasers.

Equipment using lasers of any of the above classifications will be found to carry a warning label similar to those in Fig. 2.6.9 outlining the hazards and the classification of the laser used.

The above list is an abridged version of the laser specifications contained in the IEC 60825-1 standard and should not be relied on as a comprehensive guide. The full standard, together with other relevant safety information can be purchased from the International Electrotechnical Commission Webstore

Module 2.7
Photodiodes

What you'll learn in Module 2.7

After studying this section, you should be able to:

Describe Different methods of photodiode operation:

- Photovoltaic.
- Photoconductive.

Describe the basic construction of photodiodes.

Describe the operation of different types of photodiode:

- Photoconductive diodes.
- PIN Photodiodes.
- Avalanche Photodiodes.

Describe typical limitations in photodiode operation.

- Dark current.
- Noise.

Describe the reason for the choice of common materials used in photodiode construction.

- Silicon.
- Germanium.
- Gallium Arsenide.
- Indium Gallium Arsenide.

Photodiode Basics

Photodiodes basically perform the opposite effect to LEDs and laser diodes. Instead of using electric current to cause electrons and holes to combine to create photons, photodiodes absorb light energy (photons) to generate electron/hole pairs, so creating an electric current flow.

Fig. 2.7.1 Photovoltaic diode array (Solar Panels)

Photodiode Families

Two basic methods for generating electricity from light, using photodiodes are photovoltaic and photoconductive operation. Both methods use light sensitive semiconductor diodes, the chief difference is that photovoltaic devices, mainly used in solar panels (Fig. 2.7.1) do not use any bias voltage applied to the diode, but in photoconductive operation (Fig. 2.7.2) the photodiode has a reverse bias voltage applied from some external source.

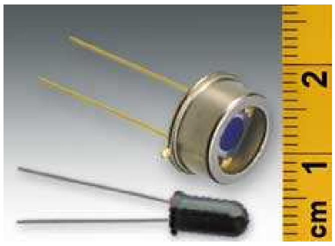


Fig. 2.7.2 Photoconductive diode

Photodiode Applications

Photoconductive diodes are used in electronic systems such as fibre optic communications (this text has been delivered to you using photodiodes). Cameras use photodiodes to measure light, and to control the shutter, focus and flash. Medical uses include X-ray detection and pulse measurement. Photoconductive diodes are the sensor of choice for many industrial systems where light needs to be measured, from bar code scanners and position sensors to smoke detectors and surveying instruments. In applications involving high frequency changes in light levels, such as fibre optic communications it is important to keep the junction capacitance of the diode to a minimum, as quite a small capacitance would remove the higher frequencies and seriously reduce the efficiency of the photodiode receiver. Photoconductive diodes are therefore manufactured in small physical sizes, which generate very small amounts of electric current. Photovoltaic diodes by contrast are manufactured as very large size solar panels to maximise the efficiency of light collection. Solar panels necessarily have a much larger junction capacitance than photoconductive devices, but their efficiency is not reduced as they are designed to produce (much greater) electric current at DC (0Hz).

Photodiode Construction

The typical construction of a photodiode is illustrated in Fig. 2.7.3. This example uses a construction technique called ion implantation where the surface of a layer of N type is bombarded with P type silicon ions to produce a P type layer about 1 µm thick. During the formation of the diode, electrons from the N type layer are attracted into the P type material and holes from the P type are attracted into the N type layer, resulting in the removal of free charge carriers close to the PN junction, so creating a depletion layer (shown in white in Fig. 2.7.3).

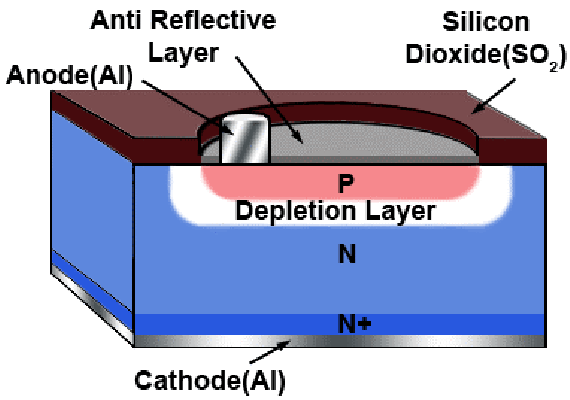


Fig. 2.7.3 Photodiode Construction

The (light facing) top of the diode is protected by a layer of Silicon Dioxide (SO_2) in which there is a window for light to shine on the semiconductor. This window is coated with a thin anti-reflective layer of Silicon Nitride (SiN) to allow maximum absorption of light and an anode connection of aluminium (Al) is provided to the P type layer. Beneath the N type layer is a more heavily doped N+ layer to provide a low resistance connection to the cathode.

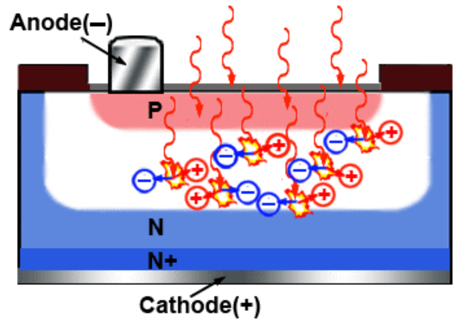


Fig. 2.7.4 Photons Create Electron/Hole Pairs

Photodiode Operation

For a diode operating in Photoconductive Mode, it is usual to use reverse bias by applying a DC voltage to make the cathode more positive than the anode. This has the effect of widening the depletion layer as shown in Figs. 2.7.4 and 2.7.5.

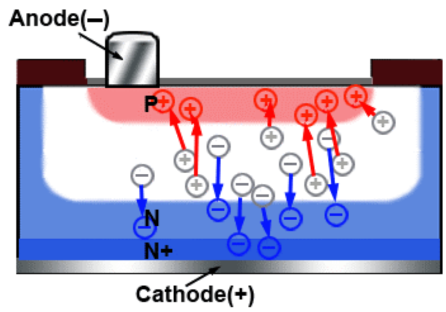


Fig. 2.7.5 Holes and Electrons are Attracted by Reverse Bias

Because the P and N layers with the depletion layer between them effectively form a capacitor, widening the depletion layer reduces the capacitance of the PN junction and increases the maximum frequency at which the diode can operate; a desirable property, especially in photodiodes that operate as digital information receivers.

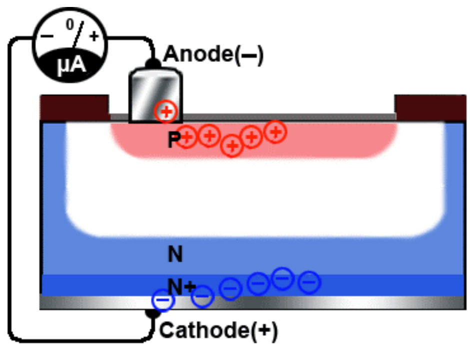


Fig. 2.7.6 Holes and Electrons Form a Photoelectric Current

When the surface of the photodiode is illuminated, as shown in Fig. 2.7.4 photons are absorbed within the diode and, mainly in the depletion layer, energise negative electrons in the valence layer of atoms, to jump to the higher energy level in the atom’s conduction band.

This leaves positively charged holes in the valence band, so producing ‘electron/hole pairs’ in the depletion layer. Some electron hole pairs are also produced in the P and N layers, but apart from those produced in the diffusion region N layers, most will be re-absorbed within the P and N materials as heat. The electrons in the depletion layer are then swept towards the positive potential on the cathode, and the holes swept towards the negative potential on the anode, so creating a photocurrent, as shown in Fig. 2.7.6.

Although Figs. 2.7.4 to 2.7.6 show different stages in the conversion of light energy into electric current, it should be realised that these steps all take place simultaneously and as a continuous process as long as the receiving surface of the photodiode is illuminated. An alternative way of illustrating photodiode action is to use an energy diagram, as shown in Fig. 2.7.7. This plots the energy levels of the valence and conduction bands of the (silicon) atom on the vertical axis of the diagram, against the distance between the anode and cathode of the photodiode on the horizontal axis.

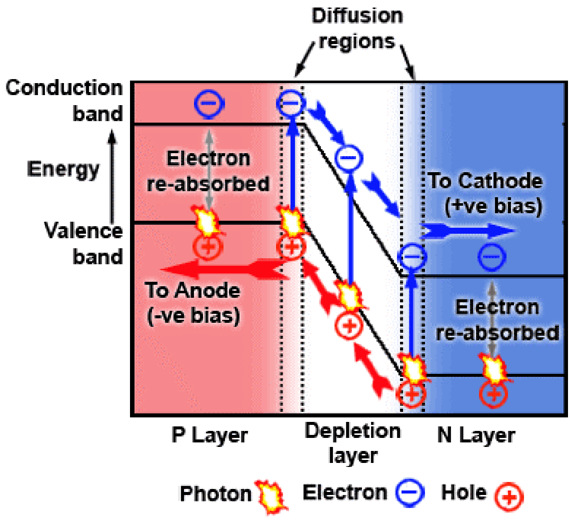


Fig.2.7.7 Energy Band Diagram of Photodiode Action

In Fig.2.7.7 photons striking atoms within the depletion layer and the diffusion regions of the P and N layers are illustrated as little flashes of energy, each of which excites an electron to jump (vertical blue arrows) to the higher energy level conduction band. Note that electron/hole pairs created within the body of the P and N layers are re-absorbed as heat. Once the electrons and holes are separated, the reverse bias applied to the anode and cathode of the diode takes over, sweeping electrons towards the (positive) cathode and holes towards the (negative) anode, (large blue and red arrows).

PIN Photodiode

This photodiode uses a layer of intrinsic (un doped or sometimes lightly doped N-) semiconductor between the P and N layers, see Fig. 2.7.8. This has the effect of reducing the capacitance of the PN junction and therefore improving the maximum switching speed, especially suited for fibre optic communications. The comparatively deep intrinsic layer also provides a larger volume for photon to electron/hole conversion.

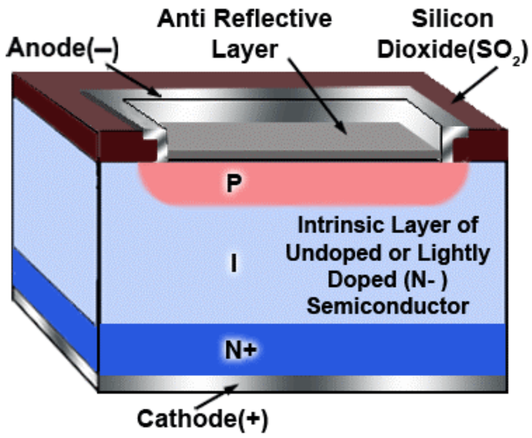


Fig. 2.7.8 PIN Photodiode

PIN photodiodes are used in the Photoconductive Mode with a reverse bias applied, the relationship between the amount of light received and electric current produced is practically linear and they are also relatively stable over their normal temperature range.

Dark Current and Noise

The current produced by the photodiode process is extremely small, in the region of micro amps to a few milliamps, and although the relationship between the amount of light shining on the photodiode and the current produced is quite linear, under very low light conditions the photocurrent produced is masked by the normal reverse leakage current due to thermal activity within the atomic structure of the diode. This current is referred to as 'dark current' as it is still present when the diode is not illuminated.

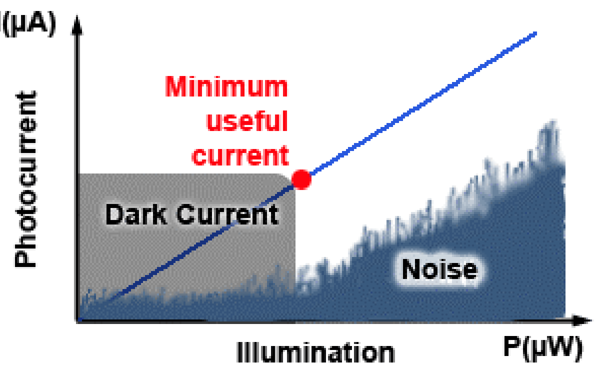


Fig. 2.7.9 Dark Current and Noise

The small value of photoelectric current produced by the photodiode, and the presence of the thermally produced dark current, results in the photodiode's useful range being significantly limited at low light levels.

Due to the extremely low signal levels obtainable from photodiodes, thermally produced noise is also a problem, especially where photodiodes may be used to detect low light levels. The 'Minimum Useful Current' for light detection is the photocurrent that is equal to the dark current plus the thermally produced noise generated by the diode as shown in Fig. 2.7.9.

Generally the signal from the photodiode will be amplified in some way, but it is useful to create the largest signal amplitude possible before it is amplified by an external circuit, as any electronic amplifier will also introduce some noise. The answer to this is to use the photodiode itself to produce a useful degree of amplification; this is the purpose of the Avalanche Photodiode.

The Avalanche Photodiode

The purpose of the avalanche photodiode is to provide an initial amplification of photocurrent within the diode itself. It does so by operating with a much larger reverse bias than other photodiodes. This can mean that the diode is operating close to the reverse breakdown area of its characteristics.

Fig. 2.7.8 shows one typical structure of an avalanche photodiode. Notice that the P+ anode is made negative and the N+ cathode layer is positive to provide the reverse bias.

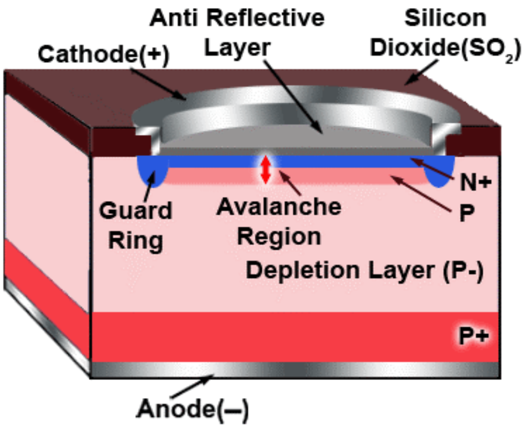


Fig. 2.7.8 The Avalanche Photodiode

Using such a high voltage bias provides a wide depletion layer, which forms a large collection area where photons create electron/hole pairs. This high voltage across the depletion layer also creates a strong electric force field that accelerates the electrons towards the positive potential on the cathode (and the holes towards the anode).

Amplification by Impact Ionisation

Notice particularly the doping used for the various layers of the photodiode. The N + layer immediately beneath the anti-reflective layer is heavily doped. Beneath this is a normally doped P layer, forming the PN junction of the diode; the main body of the diode is a lightly doped P- layer with a heavily doped P+ layer next to the anode connection.

The level of doping in a semiconductor affects its resistance, the more heavily doped layers having the lowest resistance. For a particular value of current flow through the diode layers, which are effectively a series of different value resistances, causes different voltage values across the different layers. This creates an uneven electric force field across the diode as illustrated in Fig. 2.7.11.

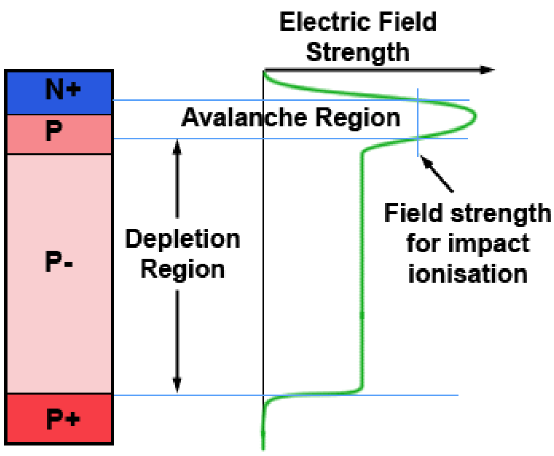


Fig 2.7.11 Impact Ionisation

The greater the electric field strength, the more acceleration is given to the electrons in the semiconductor. At the bottom of the diagram (Fig. 2.7.9) the heavily doped P+ layer next to the diode anode has a low resistance to aid efficient coupling to the metal anode connector. In the depletion region the resistance of the P- semiconductor is higher, providing sufficient field strength to accelerate the electron/hole pairs created by the photons. It is necessary, because of the depth of this area to move the charge carriers (electrons and holes) as quickly as possible to give the photodiode a fast response to changes in light level.

As electrons are attracted into the avalanche region around the more heavily doped P N+ junction, the higher resistance of these layers creates a higher voltage and so a higher field strength, which accelerates the electrons even further. When these highly accelerated electrons impact valence electrons in the atoms of the semiconductor material, they cause these previously bound valence electrons to jump into the conduction band, creating extra charge carriers. These new charge carriers (electrons) now also have sufficient energy to dislodge more electrons by impact and so on, creating an avalanche of extra electrons, which of course creates extra current.

By this method, called Impact Ionisation, the original very small current created by photons has effectively been amplified. The amount of amplification depends on the accelerating voltage, which may range between about 20V to several hundred volts. Additional factors affecting the amplification are the thickness of the avalanche region and number of electrons taking part in the impact ionisation process.

Because the number of impacts is random, the amount of amplification over any short period of time will vary, and so can only be quoted as an average value. Also because of the random nature of the photon impact, the output current will tend to be noisy due to the rapid fluctuations in amplification.

Avalanche photodiodes do not have as good a linear relationship between the light received and current produced as the other photoconductive diodes already described, but that is not necessarily a serious drawback in their main application, which is as a receiver of digital information in fibre optic communications and other high speed switching applications.

Photodiode Materials

Photodiodes use various semiconductor materials in their construction, chiefly to allow manufacturers to make a range of photodiodes that respond to different parts of the visible spectrum, as well as ultra violet and infra red wavelengths. Fig 2.7.10 shows the approximate wavelengths covered by some common semiconductor materials used for photodiodes.

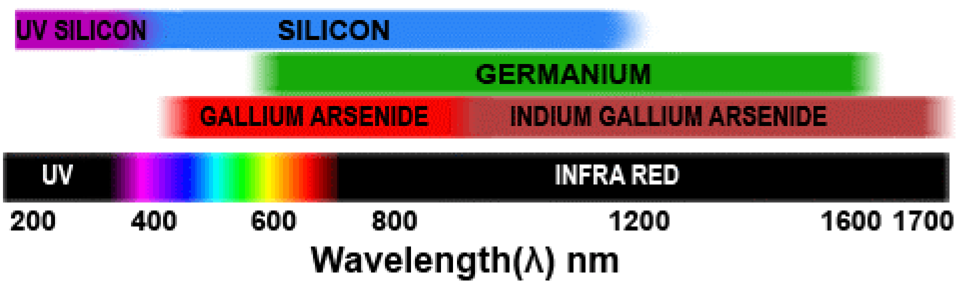


Fig. 2.7.10 Approximate Wavelength Ranges of Common Photodiode Materials

Silicon Photodiodes

Silicon (Si) photodiodes are popular for optical data receivers as they can be manufactured with a low value of junction capacitance, making them suitable for receiving digital data with frequencies up to several GigaHertz. They also generate relatively low amounts of dark current noise. However they also have a worse photon absorption rate than some other materials, which reduces their sensitivity.

Although silicon can be used over a wide range of wavelengths, from ultra violet (using specially designed UV versions) to the infra red wavelengths, silicon photodiodes are most useful in the 800 to 900nm range, as shown in Fig.2.7.13.

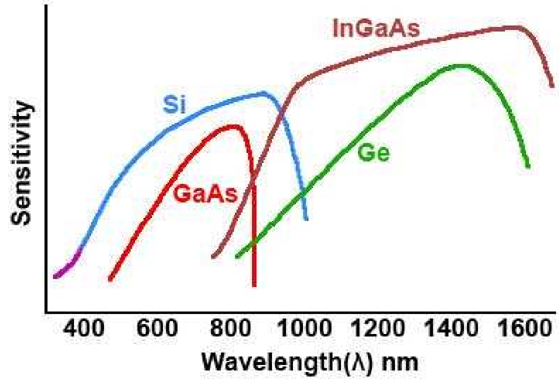


Fig.2.7.13 Relative Sensitivity of Photodiode Semiconductors

Germanium Photodiodes

Although Germanium (Ge) has been superseded in many diode applications it is useful in photodiodes as it provides light sensitivity at wavelengths longer than 900nm where Silicon is less sensitive, and Germanium is less expensive than Indium Gallium Arsenide (InGaAs), which makes it useful in photodiodes with large detection areas (up to around 1cm diameter). However Germanium photodiodes generally have higher levels of dark current and create comparatively more noise than either Silicon or Indium Gallium Arsenide, the noise level also increases at higher temperatures.

Indium Gallium Arsenide Photodiodes

Photodiodes using Indium Gallium Arsenide provide extra sensitivity in low light conditions especially at wavelengths in the infra red regions compared with either Silicon or Germanium. They generate less than half the noise and are more stable over a wide temperature range than Germanium.

Module 2.8
Testing Diodes

What you'll learn in Module 2.8

After studying this section, you should be able to:

Describe methods for testing diodes using digital or analogue multi-meters

Recognise typical faults in diodes.

- Open Circuit.
- Short Circuit.
- Leaky.

Multimeter Diode Testing

Diodes can be tested using a multi meter. It is normally the resistance of the diode in both forward and reverse directions that is tested. There are however a number of points to remember when testing diodes.

Fig. 2.8.1 Digital Meter

With Digital Meters

Most digital multi-meters are suitable for diode testing, and in many cases will have a special 'diode test' range usually marked with a diode symbol. This range should always be used when testing diodes or any other semiconductor device. The reason for this is that the meter tests the diode by applying a voltage across the diode junction. The normal voltages used by the meter on other resistance ranges may not be high enough to overcome the diode's forward junction potential and so will not make the diode conduct, even in the forward direction. This would give an indication that the diode was open circuit (very high resistance). If the diode range is used, the test voltage applied by the meter will be high enough in most cases to overcome the forward junction potential and the diode will conduct. Therefore in the forward direction (meter positive lead to the diode anode, and the negative lead to the cathode) the diode's resistance can be measured.

The actual value of resistance will depend on the slope of the forward characteristic of the diode at the voltage applied by the meter, and so will vary from device to device and from meter to meter, so a precise value cannot be given. When measuring a good silicon diode (not connected to any circuit), a reading in the forward direction of about 500Ω to 1kΩ could be expected, similar or slightly less with germanium diodes. With the meter leads reversed, an out of range (infinity) or open circuit reading (usually indicated by a display something like '**1.**' on a digital meter, as shown in Fig. 2.8.1) should be expected.

If the diode is already in a circuit, the resistances measured, **always with the circuit switched off**, will be affected by any parallel paths. Therefore readings will be lower than those indicated above. However very low or zero ohm readings may indicate a short circuit diode (the most common fault with diodes) making it worthwhile, if no other obvious reason for the very low reading can be seen, to remove at least one end of the diode from the circuit and re-check the diode's forward and reverse resistance.

With Analogue Meters

If an analogue meter is used for testing it must be remembered that because zero on the resistance and voltage scales are reversed, due to the inner workings of the meter, the polarity of the probes when using analogue meters for resistance measurement, is also reversed compared to digital meters. Therefore when measuring resistance of a diode with an analogue meter on any range the BLACK lead is positive and the RED lead is negative. This means that the black lead should be connected to the anode and red to cathode to measure the FORWARD resistance of the diode. Some analogue meters have a specific diode testing range, but most analogue meters will be quite suitable for diode testing. The most suitable analogue range will normally be indicated in the user instructions, but as with digital meters the actual voltage used on the testing range should be checked to to understand its effect on the expected forward and reverse resistances.

Fig. 2.8.2 Analogue Meter

It is also quite usual for the forward resistance measurement across some LEDs, especially those such as blue LEDs that have a higher forward junction potential to appear to very high (infinity) during testing if the meter voltage on the diode range is low, even when the LED is OK. However a meter with a test voltage of around 3V should produce some glow from the LED. Some multi-meters are also available, which instead of displaying the resistance of the diode on the diode test range, display the junction potential (in volts). It is therefore essential to make sure you know what conditions the meter uses before testing any semiconductors.

Making the tests

The diagram below shows how to connect a digital meter to test the diode. There are a number of things to remember:

- Make sure you are using the diode range.
- Using a digital meter, connect the black lead to the cathode and red to the anode (forward bias - around 1kΩ).
- Reverse the meter connections (reverse bias - infinity reading).

REMEMBER - If you are using an analogue meter to measure **resistance** the polarities of the test leads are reversed.

SOME METERS, when measuring diode resistance, give a reading indicating the junction potential (in volts) instead of the diode's resistance (in Ohms) CHECK YOUR METER INSTRUCTIONS so that you are sure what the meter reading indicates.

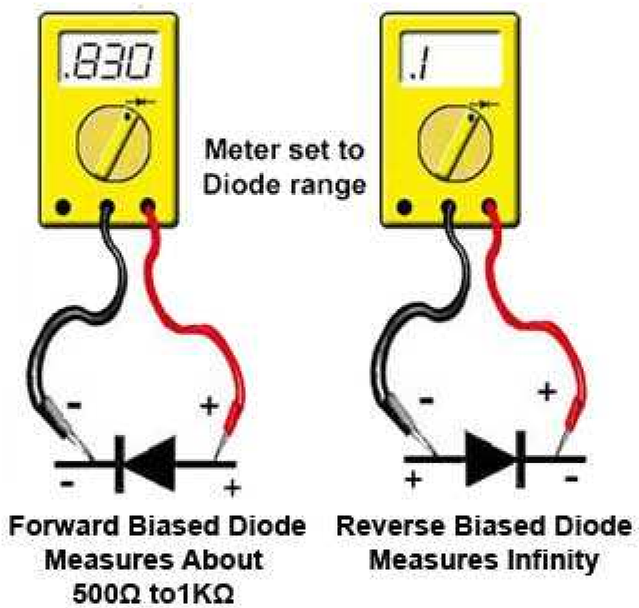


Fig. 2.8.3 Connecting a Digital Meter for Testing a Diode

Identifying Diode Connections

The cathode connection of a diode is marked in various ways. In the case of a bridge rectifier package, the AC input terminals and the DC output terminals are usually marked with a sine wave symbol and plus/minus signs respectively, as shown.

Bridge rectifiers can be tested as ordinary diodes as long as each diode is tested separately. The package pins should be compared with the diagram of the internal layout of the four diodes as shown in Fig. 2.8.4 so you can test each diode´s forward and reverse resistance. Single diodes are generally marked with a band to indicate the cathode, but with stud type rectifiers there is generally a diode symbol printed on the case.

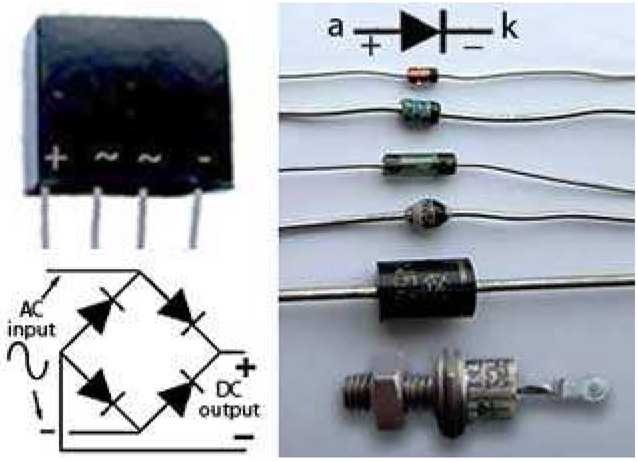


Fig. 2.8.4 Diode Polarity Markings.

Fault Indications

Short Circuit

Diodes can be damaged by high voltages, especially diodes working in high voltage or high power applications such as power supplies, and as a result will usually go short circuit 0Ω when measured in either direction. When a diode in a power supply goes short circuit, large currents can flow and obvious damage occurs such as "cooked" diodes and / or blown fuses. Short circuit diodes that are not obviously damaged show 0Ω or very low resistance in both forward and reverse directions.

Open Circuit

Occasionally, diodes (especially small signal diodes) may go open circuit, and read very high resistance or infinity (shown as 1 on digital meters) in both forward and reverse directions.

Leaky

Sometimes a signal diode may become "leaky". While its forward resistance may be normal, its reverse resistance may be lower than the expected infinity. Often this fault could only be measured with the diode removed from the circuit it is working in because of the parallel resistances of other components connected across the diode.

Testing Zener Diodes

All Zener diodes have a defined voltage, and if the voltage measured across them under working conditions, is higher than that printed in the circuit manual (or on the diode if you can see the markings), then the diode is faulty, (probably open circuit) and must be changed. Zener diodes exhibit similar short and open circuit faults to other diodes, but in addition may become ‘noisy’. The normally very stable voltage across them suffers from very rapid fluctuations similar constant to the ‘background noise’ hiss on a poor audio signal. As Zener diodes are often used to stabilise power supply lines, this rapid fluctuation of voltage can give rise to strange faults, depending on what is being supplied by the power supply in question. The moral is - If a circuit is behaving strangely, and noise on the power supply is suspected, check any Zener diode stabilising that line by substituting it with a known good diode.

Testing LEDs

LED testing is described in Diodes Module 2.5

Module 2.9
Diodes Quiz

Questions on what you have learned from Semiconductors Module 1

Use the information given in Diodes Module 2.0 to 2.8 to answer the questions in this quiz. You can check your answers by visiting our Diodes Quiz on line at:

http://www.learnabout-electronics.org/Semiconductors/diodes_29.php

1 Which of the symbols in Fig. 2.9.1 correctly represents a bridge rectifier?

a)

b)

c)

d)

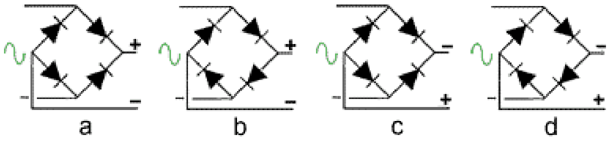


Fig. 2.9.1

2 Which of the following impurities could be used to convert intrinsic silicon to extrinsic P type silicon?

a) Aluminium.

b) Germanium.

c) Arsenic.

d) Zinc.

3 Which of the following values represents a typical junction potential of a silicon diode?

a) 0.2V

b) 0.6V

c) 1.6V

d) 3.3V

4 When selecting a silicon rectifier diode to rectify 230VAC which of the following V_{RRM} values should the diode have for safe and reliable operation?

a) 230V

b) 325V

c) 650V

d) 1000V

5 Compared to a Silicon PN Rectifier diode, which of the following characteristics would be typical of a Schottky Rectifier diode?

a) A lower forward junction potential and less reverse leakage current.

b) A greater reverse leakage current and a higher reverse breakdown voltage.

c) A lower forward junction potential and a greater reverse leakage current.

d) A higher forward junction potential and a lower reverse breakdown voltage.

6 Which of the circuit symbols illustrated in Fig. 2.9.2 represents a Schottky diode?

a)

b)

c)

d)

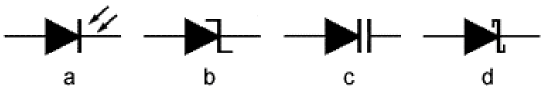


Fig. 2.9.2

7 Which of the output waveforms shown is correct for the circuit diagram shown in Fig.2.9.3?

a)

b)

c)

d)

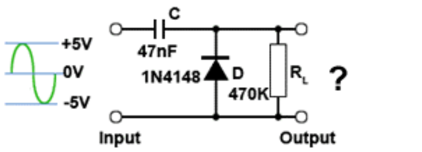
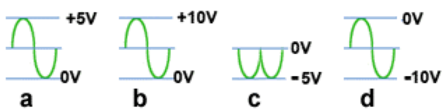


Fig. 2.9.3



8 Which of the following is a common application for Zener diodes.

a) Demodulation.

b) AC rectification.

c) Voltage regulation.

d) Input protection.

9 Refer to Fig. 2.9.4, which shows some typical ratings for an LED. Which of the given resistor values would be the most suitable for a current limiting resistor, if the LED is to be driven from a 12Vdc supply?

a) 220Ω 0.5W

b) 470Ω 0.25W

c) 680Ω 0.25W

d) 820Ω 0.25W

Item	Value	Unit
DC Forward Current I_F	1.5	mA
DC Reverse Current I_R	30 (V_R = 5V)	μA
DC Forward Voltage V_F	2.2 (2.8 Max)	V
DC Reverse Voltage V_R	5	V
Power Dissipation P_D	45	mW
Operating Temperature T_{OPR}	-40 to +95	°C

Fig. 2.9.4

10 Which class of laser diode is suitable for use in hand held laser pointers?

a) Class 1

b) Class 2

c) Class 3R

d) Class 4