Online Radio & Electronics Course

Reading 24

Ron Bertrand VK2DQ
http://www.radioelectronicschool.com

SEMICONDUCTORS Part 1

The objective of this reading and the following, is to provide a basic coverage of the most generally employed solid state devices, such as diodes of various types, bipolar junction transistors (BJTs), FETs MOSFETs, SCRs, and ICs, as well as a few of the characteristic circuits in which they are used.

Information that is required for examination purposes will be brought to the attention of the reader. Pay attention to highlighting.

SOLID STATE DEVICES

Solid state devices are much smaller physically, more rugged and lighter than vacuum types, but cannot withstand heat as well, changing their characteristics as they warm. However, many newer special solid state devices will operate at much higher frequencies than any vacuum devices can.

SEMICONDUCTORS

The outer orbit, or valence electrons of some atoms, such as metals (conductors), can be detached with relative ease at almost any temperature, and may be called free electrons. This is why metals are good conductors. These valence electrons are able to move outward from a normal outer orbit level into a conduction level or band, from which they can be easily dislodged. Such materials make good electrical conductors. Other substances such as glass, rubber, and plastics, have no free valence electrons in their outer conduction bands at room temperatures, and are therefore good insulators.

A few materials have a limited number of electrons in the conduction level (outer orbit or valence band) at room temperatures and are called semiconductors. Applying energy in the form of photons (small packets of light or heat energy) to the valence electrons moves some of them up into the conduction band, and the semiconductors then become better electrical conductors. Energy of some form is required to raise semiconductor electrons to a conduction level. Conversely, if an electron drops to the valence level from a higher conduction level, it will radiate energy in some high frequency form such as heat, light, infrared, ultraviolet, and, if the fall is great enough, X-rays.

Two semiconductors, germanium and silicon, have four outer ring electrons. Crystals of these can be laboratory grown, so this makes mass production of them easy.

NATURE PREFERS EIGHT

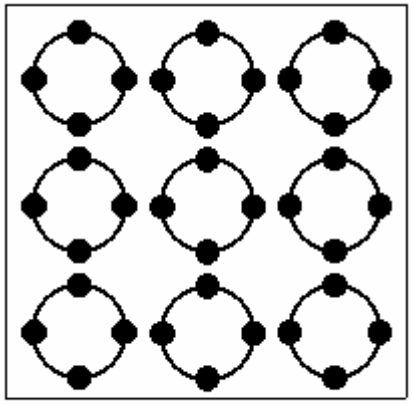
It just so happens in nature that atoms are more stable and 'like' to have eight valence (outer) electrons. In the case of germanium and silicon which only have four valence electrons, a special arrangement occurs. Germanium or silicon in pure form create what is called a crystal lattice structure. The four valence electrons of each atom in a crystal lattice

structure share themselves with the adjacent orbits of all other electrons on the crystal lattice structure. In this way by borrowing electrons from neighbouring atoms a type of shared valency of eight exists. Consequently, the valence electron arrangement is very stable and it is difficult to make these electrons participate in current flow. The great breakthrough in physics (electronics) was the discovery of how the characteristics of pure germanium or silicon could be changed dramatically by adding impurities (other atoms) to the crystal. Adding impurities disturbs or upsets the crystal lattice structure.

INTRINSIC SEMICONDUCTOR LATTICE

A perfectly formed intrinsic semiconductor crystal lattice is illustrated in figure 1. Such a crystal acts more like an insulator than a conductor at room temperature. I have drawn the crystal in only two dimensions (not being much of an artist). Shown is the valence band (outer orbit of each atom) and the four valance electrons. The valence electrons are of course not stationary, but orbiting around the atom as if on the surface of a sphere. The sharing of valence electrons is called covalent bonding. This arrangement is very stable electrically. Electrons are locked into the crystal lattice, and at normal temperatures the crystal is an insulator.

Figure 1.

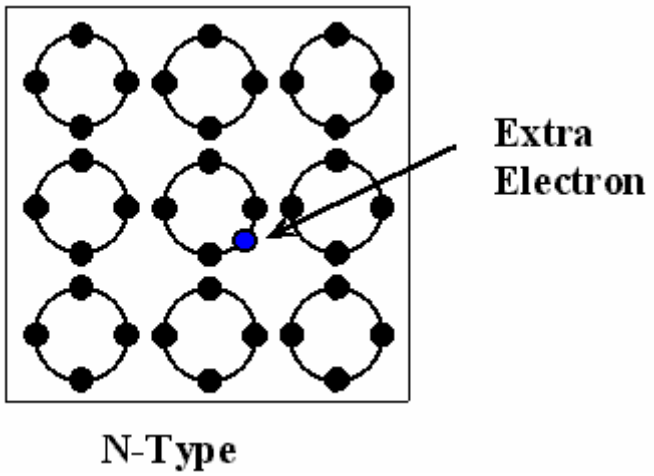


Semiconductor

DOPING (N)

Doping is the process of deliberately adding impurities to the crystal during manufacture. With germanium crystal, about one in a million atoms of an impurity such as arsenic with its five outer ring electrons, is added. The resulting crystal is imperfect. Arsenic is penta-valent (five outer electrons) and cannot fit into the crystal lattice structure. What happens is four of the arsenic electrons participate in the sharing (covalent bonding) and one is left out! The crystal lattice shown in figure 2 has one atom in a million with an excess outer ring electron not being tightly held.

Figure 2.



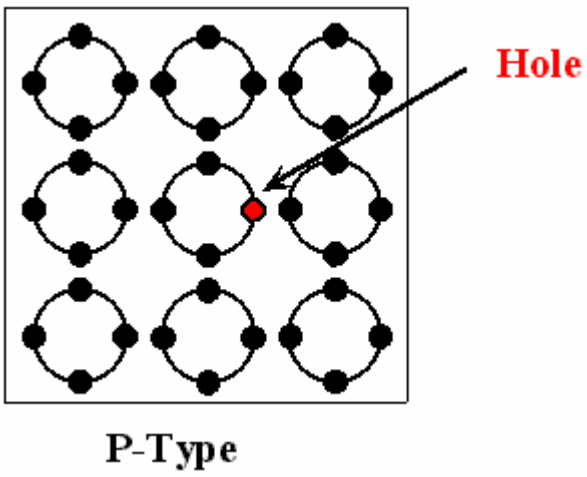
It should be understood that ‘N' type semiconductor does not have extra electrons electrically. N-type semiconductor does not have a negative charge. What is ‘extra’ in N-type material is electrons which do not fit into the crystal lattice structure. These extra electrons are not locked into the crystal lattice structure so they are much easier to move.

When an electrostatic field (by application of an emf) is developed across such arsenic-doped germanium, a current will flow. The N-type semiconductor is about 1,000 times better as a conductor than the intrinsic semiconductor. Doped germanium with such relatively free electrons is known as N germanium, and is a reasonably good conductor. (To form N silicon, phosphorus can be used as the dopant).

So, merely by adding a small amount of impure pentavalent atoms to a pure crystal, we convert it into a conductor by disrupting the harmony of the crystal lattice structure.

DOPING (P)

Figure 3.



When germanium is doped with gallium, which has three valence electrons (tri-valent), the crystal lattice is again disrupted. This time there is an area, or hole, in the crystal lattice structure, a region that apparently lacks an electron. While the hole may not actually be positive, at least it is an area in which electrons might be repelled to by a negative charge. This positive appearing semiconductor material is called P germanium. When an electrostatic field is impressed across a P-type semiconductor, the hole areas act as stepping stones for electron travel through the material. It can be said that hole current flows in a direction opposite to the electron flow. Note that both N germanium and P germanium have zero electric charge because both have an equal number of electrons and protons in all of their atoms. (One dopant used to produce P silicon is boron).

HOLES

I am going to talk about holes for a bit, as it seems to be a stumbling block for many. I have drawn the 'hole' as a red circle in figure 3. The hole is a missing electron in the crystal lattice structure, which destroys the crystals insulating properties. It is very easy with N type material to visualise that electron flow can take place. With a hole it is a little harder, and I find some textbooks a little confusing on this issue. A hole is a hole in the crystal lattice structure. Because the lattice is not complete in the location where a hole is, electrons can move into the hole, and in doing so they create a hole from where they came from.

SOME FOOD FOR THOUGHT

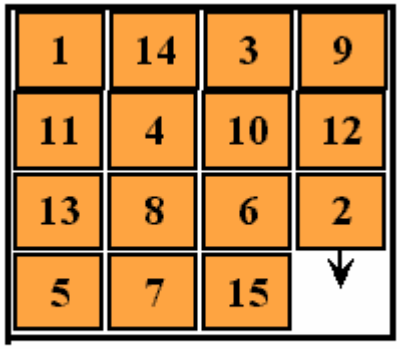
- A hole is a missing electron in the crystal lattice structure.
- A hole is not a positive charge.
- A hole like any hole, can be filled.
- A hole can be filled with an electron.
- When an electron does fill a hole – then the filled hole disappears, BUT where the electron came from there is now a hole.

If an electron falls into a hole, then where that electron came from will be a hole.
A hole can be thought of as positive for behavioural description purposes.
An electric current is an ordered movement of electrons.
Holes allow electrons to move in the crystal by giving them somewhere to go ie. filling a hole.
When an electron moves out of the covalent bond to fill a hole it leaves a hole from whence it came.

I think most of us have seen the toy shown in figure 4. A flat panel of plastic squares with pictures or number on them, and the objective is to move the squares around into some order, either to get the numbers in order or to make a picture. Would you be able to slide

the squares around if the game was made without a missing square? No, of course not. By leaving a square out (leaving a hole in the puzzle) it makes it possible to slide the squares around by moving them into a hole. In moving a square into a hole you create a hole, making it possible to slide other squares into that hole. Think of the squares as electrons and their ability to move is made possible by the presence of the hole (missing square).

Figure 4.



Puzzle Game

Slide 2 (or 15) can be moved into the hole, however in doing so they would leave a hole. Now, if you were to put this puzzle on auto pilot and sit back and watch it, what would you notice about the way the hole moves? It moves in the opposite direction to the slides. So if the numbered squares are electrons the hole is behaving like a positive charge in that it moves in the opposite direction to electrons. Some references do not explain this very well and go on to talk about "hole current" moving from positive to negative, I believe confusing the reader further. Electrons moving forms the only current. Though I will be talking shortly about holes moving, holes do really move when you fill them. However, all of the **real** moving is done by electrons falling into holes in P-type semiconductor.

P-type semiconductor material is a conductor because of the presence of holes in the crystal lattice structure. Doped silicon has considerably more resistance than germanium, but it is useful in higher voltage applications, does not change its resistance as much when heated, and can withstand greater temperatures without its crystalline structure being destroyed.

So what?

Well it does not seem like we have achieved much. We have taken a perfectly good semi-insulator (semiconductor) and turned it into a conductor by adding penta-valent impurities (N-type) or tri-valent impurities (P-type). The magic starts when we combine the two, that is, we make one side of the crystal N-type and the other P-type. In reality they are not made separately and then stuck together, crystals are grown and doped on different sides of the same crystal in the laboratory.

SOLID STATE DIODES

Before we start, the term solid state is only used because the alternative devices before them were the electron tube devices.

Stick'em together.

Lets take a piece of N-type and P-type semiconductors and join then together The area in which the N and P substances join is called the junction.

Some of the relatively free electrons in the N-type material at the junction fall into some of the holes in the P-type material. So, right at the junction, there are no free electrons (in the crystal lattice structure) and no holes, as free electrons from the N material have filled some of the holes in the P material. This creates a region at the junction, which has neither free electrons nor holes. It is a region which can be thought of as depleted of free electrons and holes, and is called the depletion zone.

This develops an area at the junction that is actually slightly negative on the P side (because electrons have filled holes) of the junction, and slightly positive on the N side (because electrons have left to go and fill holes). This produces a barrier to any further electron flow of about 0.2 V with germanium and 0.6 V with silicon diodes.

We have created a semiconductor diode (PN junction). From now on I am going to be drawing electrons (blue) and holes (red). In figure 5 the depletion zone is shown at the junction, drawn in green. The whole (no pun intended) diagram is exaggerated greatly as the depletion zone is extremely narrow, and there are many more holes and electrons in a real PN junction.

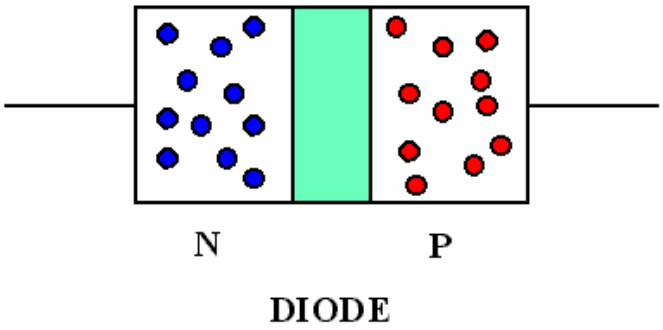
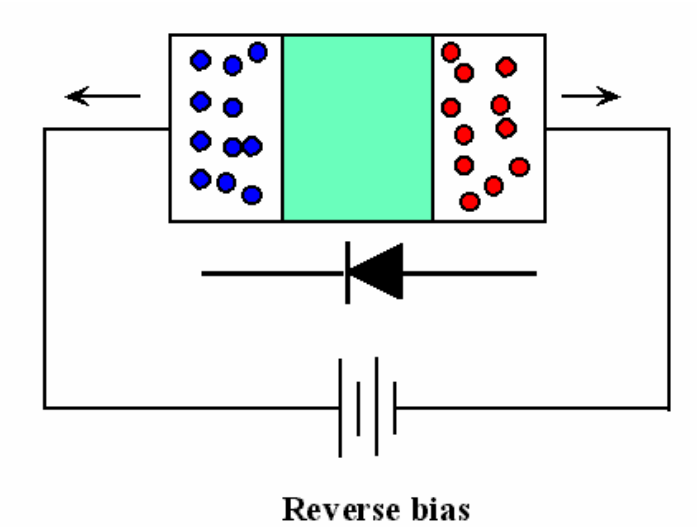


Figure 5.

P and N junctions aren't made separately and then stuck together as I inferred above. Different parts of the one crystal is dope to create the junction. This way millions of junctions can be created on the one "chip".

APPLYING AN EMF TO THE PN JUNCTION

Figure 6.



Reverse Bias:

In figure 6 we have applied an external electrical pressure. Electrons are attracted toward the positive terminal of the battery (the long stroke on the battery symbol is positive) and holes (thinking of them as positive charges) are attracted to the negative terminal.

No current flows through the PN junction, which from now on, we will call a diode. The depletion zone is widened as electrons and holes are moving away from the junction. The diode is said to be reverse biased. I have drawn the schematic symbol of a diode, so you can see that for reverse bias, positive is connected to the cathode and negative to the anode. The left hand side of the diode shown is the cathode.

Also, recall when we discussed power supplies without really describing how a diode worked. I suggested you look at the diode schematic symbol as an arrow. In figure 6 the arrow is pointing to the left. I asked you to remember that conduction was only possible in the opposite direction to the arrow. Electrons can only flow in the direction cathode-to-anode, against the arrow.

We have said that no current flows through the diode when it is reverse biased. To be truthful, there is an extremely low *leakage* current which for most practical purposes can be considered to be zero. Also, while we are being truthful, if you increase the reverse bias voltage high enough you will blow the crapper out of the diode and conduction will take place (PIV).

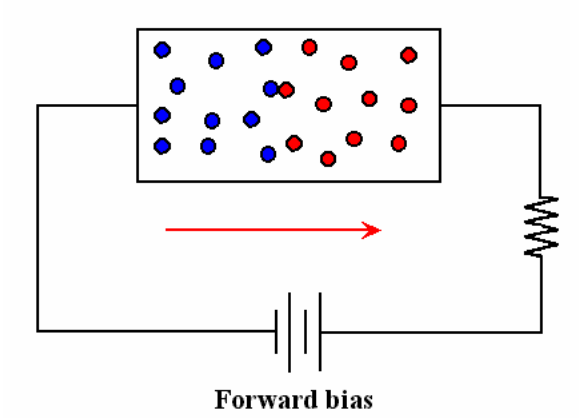
Just as a matter of interest, the reason why you do get a reverse leakage current in a diode is that some of the N material will have just a few holes in it and some of the P material will have some electrons in it. These are due to extremely small amounts of contaminants. So there is like a minor ‘ghost’ diode the opposite way around to the ‘real’ diode. These contaminants are called minority current carriers. The real holes and electrons are called the majority current carriers. If you like, you have a majority diode one way and a minority diode the other way. When the majority diode is reversed biased the minority diode is forward biased and this accounts for the small leakage current.

The reverse leakage current of a PN junction increases with temperature. Reverse biased PN junctions can be used to measure temperature, amplifying, and measuring the reverse leakage current.

Forward Bias:

Now let's reverse the battery. I know you know that the diode is going to conduct but let's look at exactly what goes on. We are going to think of the holes as positive charges when we really know they aren't - we covered that issue with the 'puzzle' example earlier. However, do let me know if any part of this reading is not clear enough.

Figure 7.

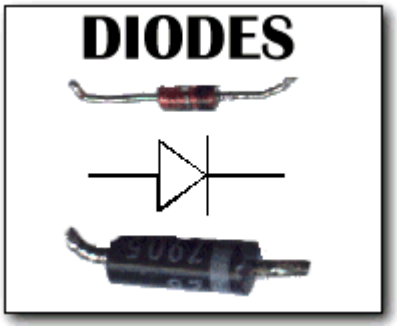


We have now applied a negative potential to the N material (cathode) and positive to the P material (anode). **Provided this potential is greater than the barrier potential** (0.2V germanium, 0.6V silicon) the depletion zone will be flooded with electrons and the diode will conduct as shown in figure 7. The resistor is added to limit the forward current.

You may well ask, "why don't all the electrons move across the junction and fill up all the holes". We all know that current flow is electrons. In figure 7 the right hand side is P-type (holes), and electrons leave the P-type anode and flow to the positive terminal of the battery. Every electron that leaves the anode, creates a hole. The holes move toward the junction to be filled by more electrons.

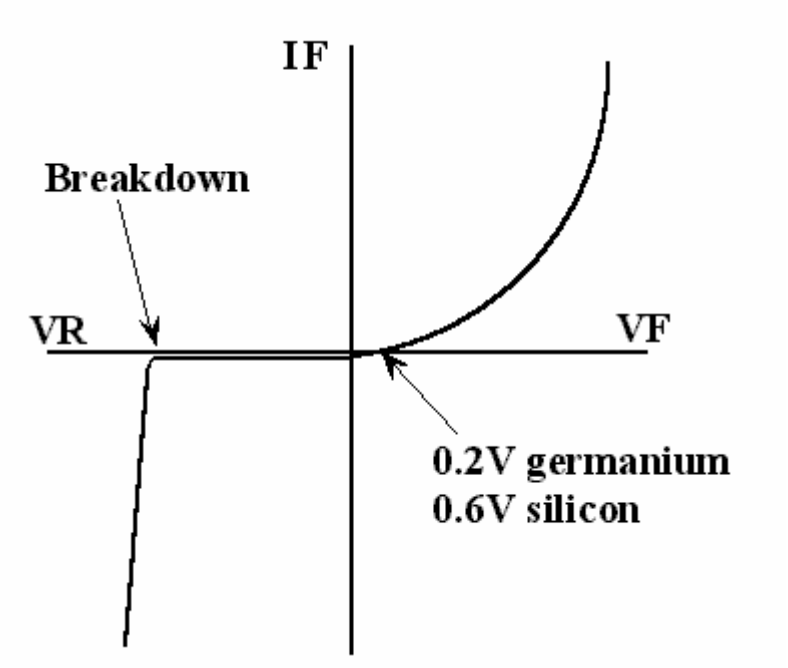
Such a diode could be used in a rectifier circuit. All diodes have a maximum current rating as well as a peak inverse voltage rating. A diode for example may be rated at 1Amp 400V PIV.

Figure 8.



The line on the end of the diode is used to indicate the cathode end.

Figure 9 – Diode Operating Characteristics



The graph in figure 9 shows the operating characteristics of a diode. VF is the forward voltage and VR is the reverse voltage. Conduction in the forward direction does not start until VF exceeds the barrier potential shown. If the reverse voltage becomes too high the diode will break down and conduct in the reverse direction.

This (break down) is called *zener effect* if the emf value at breakdown is less than about 5 V, and *avalanche* if it is more than about 5 V. This is not a normal operating condition for most semiconductor diodes and may cause lattice damage, ruining the diode.

ZENER DIODES

Figure 10 – Zener.

The reverse voltage breakdown effect, however, is used in special zener diodes. These diodes are deliberately operated under enough voltage to cause them to conduct in the reverse direction. A resistor must be connected in series with a zener diode to prevent the junction from being destroyed. The circuit across which the diode is connected will not increase in voltage over the zener breakdown voltage. For this reason, zener diodes are used as shunt (parallel) **voltage-regulating** devices. Zener diodes are used to provide a regulated DC voltage of low power. So if some part of a 12 volt DC circuit required 5V DC at low power, then a zener diode with an appropriate series resistor could be used to provide a regulated 5 volt DC output in spite of variations in the 12 volt DC input voltage. A good way to remember the symbol of a zener diode is to note the shape of the cathode line on the diode as representing the forward and reverse current characteristics of a diode shown earlier in figure 9.

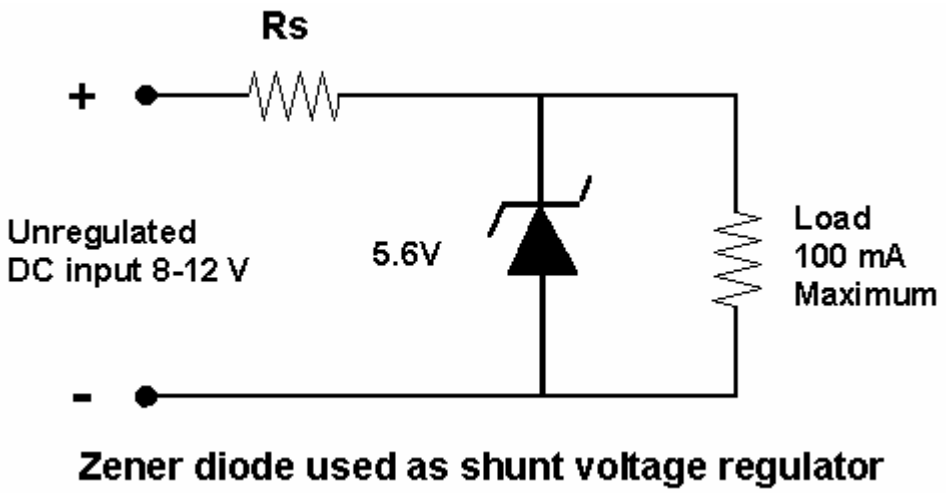
THE CIRCUIT OF A ZENER VOLTAGE REGULATOR

The circuit diagram of a zener used as a voltage regulator is shown in figure 11. The unregulated DC to the circuit varies from 8 to 12 volts. R_S is a current limiting resistance to prevent the zener from being destroyed. Remember, a zener regulator is operated with reverse bias and to the point where it breaks down and conducts in the reverse direction.

For some, the term 'breakdown' is confusing. Break down does not mean the zener is destroyed or busted! Break down means it is forced to conduct in the opposite direction (from anode to cathode). Normally a diode operated beyond break down is destroyed. However zener diode are designed to operate in the breakdown region with a small breakdown current. A zener would be destroyed just like any other diode except for the current limiting resistance R_S. Under these condition the voltage across the zener is constant. In figure 11 the voltage across the zener will be 5.6 volts (it is a 5.6 volt zener – you buy them with a voltage rating). Irrespective of input voltage fluctuations, the voltage across the zener will be a very constant 5.6 volts.

R_S and the zener form a series circuit. The sum of the voltage across the zener (5.6V) and across R_S is equal to the input voltage. Suppose the unregulated input voltage was at a maximum (12V), then the voltage across R_S would be 12-5.6 = 6.4 volts.

Figure 11.



Zener diode used as shunt voltage regulator

As an important rule-of-thumb, the reverse zener current is about 1/10th of the maximum current drawn by the load – the load draws 100mA so one could expect the reverse current of the zener to be 10mA. The zener and the load form a parallel circuit – so the sum of the branch currents is equal to the current through R_S. The current through R_S must then be 100 + 10 = 110mA. So we know the maximum current through R_S is 110mA. What is the maximum voltage across R_S? The maximum input voltage is 12V so the maximum voltage across R_S will be 12-5.6 = 6.4 volts. We can now calculate the resistance of R_S from Ohms Law:

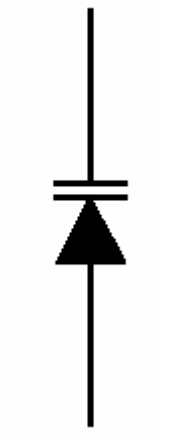
R_S = E(across R_S) / I(through R_S) = 6.4 volts / 110 mA = 58 Ω.

Just as important is the power rating of R_S:

Power (of R_S) = E(across R_S) x I(through R_S) = 6.4 volts x 110 mA = 0.704 Watts.

That's all there is to designing a simple voltage regulator using a zener. Zener diodes come in a range of voltages up to about 18 volts.

Figure 12.



VARACTOR OR VARICAP

A reverse biased diode (or PN junction) does not conduct. If you refer back to when we applied a reverse bias to a PN junction, we saw the width of the depletion zone increased. If we increase the amount of reverse bias further (without reaching breakdown), the width of the depletion zone would widen even further. If we reduce the reverse bias, the depletion zone would narrow. Now, if we continually varied the amount of reverse bias, the width of the depletion zone would also continuously vary. Each time the depletion zone changes in size there must be some movement of electrons in the circuit. However, electrons never move across the junction. The junction, under reverse bias, is an insulator.

Does this movement of current on each side on an insulator remind you of capacitance? It should, because **a reversed biased diode will act just like a small capacitor whose capacitance can be changed by the amount of reverse bias**. If you like to think of it another way, the depletion zone is the dielectric, which can be made to change in thickness or width by the amount of reverse bias.

So a reverse biased diode can be used to create a voltage variable capacitor, the symbol of which is shown in figure 12. There are purpose made diodes for use as varactors, though in practice almost any diode can be used for this effect. This ability of a diode to behave as a variable capacitance is extremely useful. The frequency of a tuned circuit or a quartz crystal can be made to vary by using a varactor diode. By using a variable resistor to adjust the reverse bias, the capacitance of the diode can be made to vary, which in turn will affect the frequency of the tuned circuit.

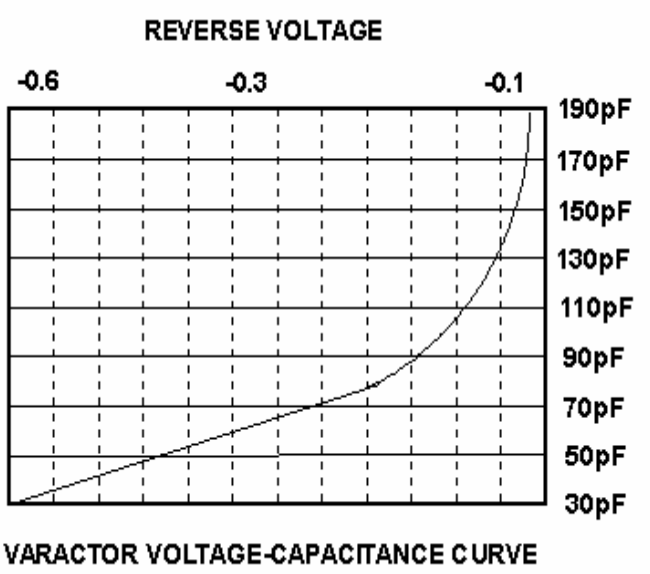


Figure 13.

Though we have not discussed single sideband (SSB) reception yet, many readers will know that with an SSB receiver, say a CB radio, you have to finely tune the radio to the received stations with a control most often called a clarifier. The clarifier is usually a variable resistor in combination with a varactor to make small changes to the receiver's frequency.

LIGHT EMITTING DIODES

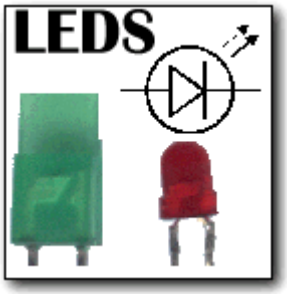
In any forward biased diode, free electrons cross the junction and fall into holes. When electrons recombine with holes they radiate energy. In the rectifier diode, the energy is given off as heat. In the light emitting diode (LED), this energy radiates as light.

Figure 14.

It takes energy from the source to move an electron from the valence level to the conduction level. When an electron drops back to the valence level it will emit energy in the form of photons (light).

An electron moving across the PN junction moves to a hole area. This can allow a nearby conduction electron to fall to its valence level, radiating energy. In common diodes and transistors made from germanium, silicon, or gallium arsenide, this **electromagnetic radiation is usually at a heat frequency, which is lower than light frequencies**. With gallium arsenide phosphide the radiation occurs at red light frequencies. Gallium phosphides produce still higher frequency (yellow through green) radiations. Gallium nitride radiates blue light.

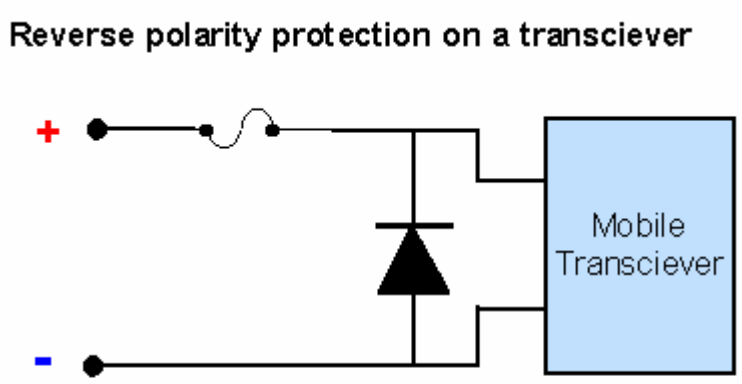
Figure 15.



There are several photodiode and photosensitive devices. The photodiodes convert photons to electric emf. There are also a number of other specialist diodes. For examination purposes, we have more than covered enough material here. Also, the depth of the material is more than adequate. You will *not* be asked to describe the operation of a diode at the electron/hole level, although you should remember the terms used thus far and what they mean.

A SIMPLE APPLICATION FOR A DIODE

Figure 16.



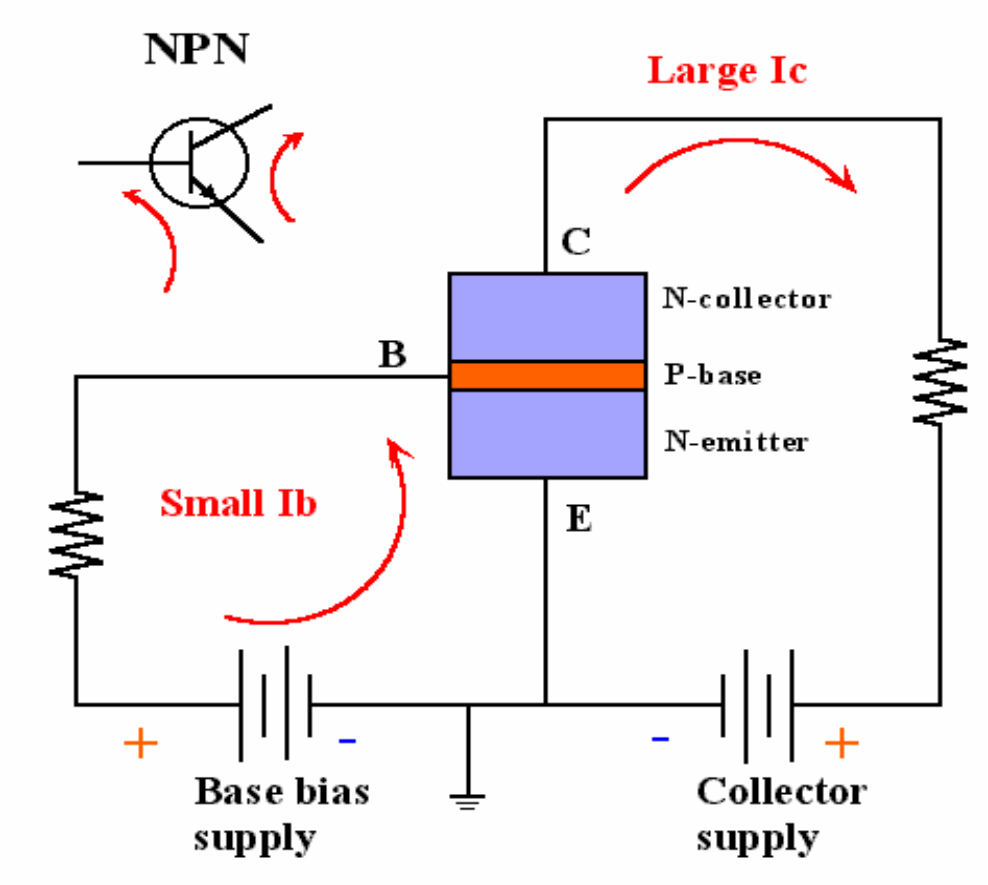
The diode is normally reversed biased and open circuit. The wrong polarity causes forward bias blowing the fuse and often destroying the diode - the radio is saved from disaster.

Power diodes are made from silicon. Besides being used as rectifiers, one very useful application for a single power diode is reverse voltage protection in a radio. All mobile (used in vehicles) radio equipment is very much prone to being connected to the source of power the wrong way. I have first hand experience of doing this myself. Connecting the power supply or battery to a mobile radio (or any radio) the wrong way around would be devastating to the radio if some sort of reverse polarity protection did not exist. All mobile radio equipment (CB, amateur etc) has an inline fuse. Where the negative and positive leads of the power supply enter the equipment there is a reverse biased silicon diode. When the power supply polarity is connected the correct way this diode is reverse biased and acts like an open circuit. If the user accidentally connects the radio to the wrong polarity, the diode becomes forward biased and conducts heavily, blowing the fuse. Most times when this happens the diode is destroyed and remains as a short circuit across positive and negative. Backyard technicians will often fix the problem by just cutting one lead of the diode removing it from the circuit and replacing the fuse. Of course they will charge you $50 for this two minute job! With the diode out of circuit, the next time the radio is connected to reverse polarity you can say goodbye to the radio for good.

TRANSISTORS

The basic transistor can be thought of as two diode junctions constructed in series. From the bottom there is a contact against an N-type emitter element. Next to this is a thin P-type base element, with a metal electrode connected to it, forming the first PN junction. A second N-type collector element is added, with a contact on it, forming the second PN junction. This produces an NPN transistor.

Figure 17.



This type of transistor is called a bipolar junction transistor (BJT). The important characteristics of construction are:

- **The base region is very thin and light doped.**
- **The emitter region is heavily doped.**
- **The collector is large and usually connected to the case as a heat sink, so that heat can be removed from the transistor.**

With just the collector supply connected, no current will flow in the collector circuit, as the top PN junction - the one between collector and base - is reverse biased.

Consider when the base-emitter junction is forward biased as shown in figure 17. The heavily dope emitter region floods the thin base region with charge carriers (electrons). The base region is lightly doped compared to the emitter. All of the electrons passing across the forward biased base-emitter junction are looking for a hole to fall into. There are more electrons crossing the lower junction than there are holes available on the other side (in the base) to meet them. I like to think of the base region as becoming saturated with charge carriers (which are electrons for an NPN transistor). The excess of electrons come

under the influence of the collector voltage (which is higher than the base voltage too) and consequently electrons flow in the collector circuit.

The important aspect of the transistor is that small amounts of base-emitter current can control large amounts of collector current. If the base current was made to vary, say by the insertion of a carbon microphone (which is a sound dependent resistor), then the collector will be an amplified version of the base current. The transistor is an amplifier.

It is interesting to note that the name transistor comes from "transfer resistor". Another way of looking at the operation is that without the base-emitter junction being forward biased there is no collector current. When the base-emitter junction is forward biased the resistance between collector and emitter decreases from infinity (or some very high value) and current flows in the emitter circuit. Small variations in the amount of base-emitter current causes the resistance between collector and emitter to vary greatly, but in proportion. Therefore, the collector current faithfully follows the base current but is much larger, and supplied with a higher voltage. The power is greater, so the transistor amplifies.

The amount by which an amplifier amplifies is called the gain. The amount by which a transistor amplifies is called Beta and has the Greek symbol β. The beta of a transistor is calculated from:

$$\beta = \Delta I_C \, / \, \Delta I_B$$

The triangle symbol Δ is the Greek letter delta, and is the mathematical shorthand for 'change in'. Thus, ΔI_C and ΔI_B are the change in collector and base currents respectively.

So the Beta (β) or gain of a transistor, is the change in collector current divided by the change in base current.

A small variation of base current can control 50 to 150 times as much collector current. Thus, the transistor is an ideal control and amplifying device. A junction transistor of this type can be called a bipolar junction transistor (BJT) to differentiate it from a field-effect transistor (FET) discussed later.

If a Junction transistor has its base emitter current changed from 10 to 30 milliamps and this causes the collector current to change from 50 to 250 milliamps, what is the current gain or Beta of the transistor? The change in base emitter current is from 10mA to 30mA = 20 mA. The change in collector current is from 50mA to 250mA = 200 mA. The Beta is therefore 200/20 = 10. The transistor has a current gain of 10.

COMPARING A TRIODE AND BJT

A triode electron tube is also an amplifier as we have learnt. There is one significant difference between a triode and the BJT which I would like to mention. Firstly, just to refresh your memory. A triode amplifies by adjusting the negative voltage on the control grid, which in turn is able to control the large cathode to plate current, resulting in amplification.

A triode is often called a *voltage amplifier*, because it is voltage on the control grid which controls the anode current. A BJT is called a *current amplifier* because it is base current which controls the much larger collector current.

The control-grid cathode circuit of a triode does not have any current flowing in it (there are exceptions). This is important. Since a triode can amplifier with voltage alone, a triode consumes no power from the input source. From P=EI, if you have no I you have no P. A triode is a high input impedance amplifier, whereas a BJT is a low input impedance amplifier.

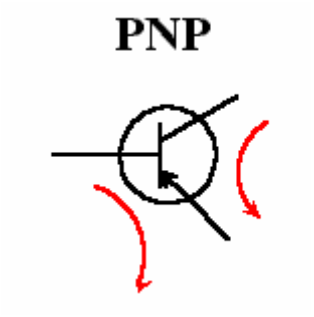
Whether a device is a voltage amplifier or current amplifier is irrelevant to the final amplification, though sometimes the input impedance is important. The triode has a significant advantage in being able to amplify a weak signal from a low power source without taking any power from it.

The advantages of the BJT though, are enormous: size, low heat, lower voltages, easier construction, and many more.

Memory Jogger:

Always remember, if you are trying to work out, or asked to work out, if a transistor has the correct polarity voltages to operate:

- The base must have forward bias.
- If the bias voltage is correct, current will flow *against* the arrow in the symbol.
- The collector voltage must also permit current flow against the arrow.



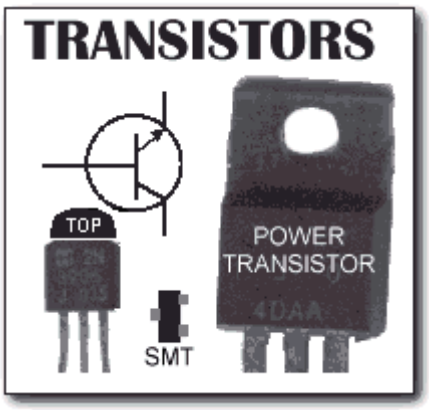


Figure 18.

A LITTLE ABOUT NOISE

We tend to think of electricity (electron flow) as being fluid – smooth. This is not correct, electricity is made up of lumps – very small lumps called electrons. Noise is produced whenever an electron does not do what it is supposed to do, when it is supposed to do it. In the Electron tube, electrons might collide with secondary electrons emitted from the anode. Imagine a situation in a PN junction where an electron is ready to fall into a hole! But no hole is to be found! For that very small instant that electron represents noise. Any random or unwanted electron motion, or lack of motion, in semiconductor devices, is noise. Don't get the wrong idea – PN junctions in transistors are wonderful low noise amplifiers, however when it comes to super sensitive receivers like those used for radio astronomy, PN junctions and hole-electron recombination is just too noisy. We shall see later that there are semiconductor devices that amplify and don't have a hole-electron recombination, making them much quieter devices.

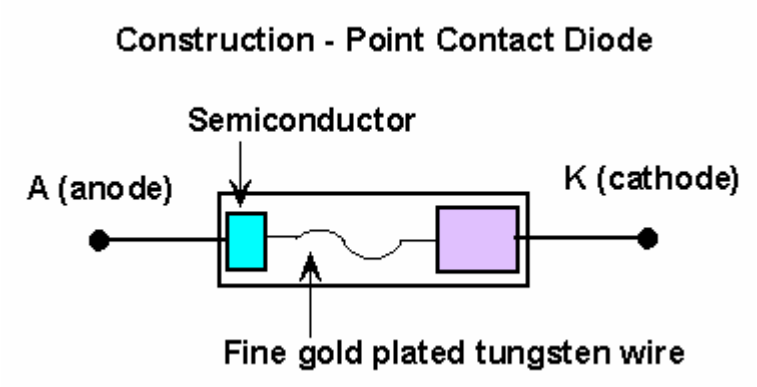
APPENDIX - OTHER DIODES

There are many other types of diodes that have special properties and uses. There is just no need to discuss them all here. However, I am going to mention two more types. They are not in the syllabus. So why am I discussing them? Well after a number of students sat their exams, I received email saying, heh! what are these things?

One of the problems about multiple-choice exams is that the candidate can be thrown off by an obscure term. The two diodes are the Point Contact and the Schottky Diode. They are never the right answer. This appendix will just broaden you knowledge a little and you won't be put off by these names.

THE POINT CONTACT DIODE

Figure 19.



The construction of a typical point-contact silicon diode is shown in figure 19. This diode consists of a (usually) brass base on which a small pellet of silicon, germanium, gallium arsenide or even indium phosphide is mounted (labelled semiconductor in the drawing). A fine gold-plated tungsten wire with a diameter of about 80 to 400 microns (millionths of a metre) and a sharp point, makes contact with the polished top of the semiconductor pellet, and is pressed down on it slightly from a spring contact. This cat's whisker, as it is known, is connected on the right hand side to a brass plate which is the cathode. The semiconductor injects electrons into the metal. The energy level between the valence electrons in the semiconductor pallet and the tip of the wire produces a diode action. The contact area exhibits extremely low capacitance. Because of the low capacitance, point contact diodes can be used for applications in excess of 100 GHz – other types of diode have too much junction capacitance for use at these high frequencies.

SCHOTTKY DIODE

The examiner seems to have a fetish for mentioning this diode. The Schottky diode (named after the inventor) uses a metal such as gold, silver, or platinum on one side of the junction and doped silicon (usually N-type) on the other side. When the Schottky diode is unbiased, free electrons on the N side are in smaller orbits than the free electrons in the metal. This difference in orbit size is called the Schottky barrier.

When the diode is forward biased, free electrons on the N side gain enough energy to travel to larger orbits (it takes energy to make an electron move to a larger orbit). Because of this, free electrons can cross the junction and enter the metal, producing a large current.

Because the metal has no holes, there is no depletion zone. In an ordinary diode the depletion zone must be overcome before a diode can conduct – this takes time - a very short time, but time nonetheless. The Schottky diode can switch on and off faster than an ordinary PN junction. In fact, a Schottky diode easily rectifies frequencies above 300 MHz.

End of Reading 24.
Last revision: July 2006

E-mail: manager@radioelectronicschool.com
http://www.radioelectronicschool.net