Discrete Electronic Components - Semiconductors

A. Discrete Electronic Components - Semiconductors

1. Semiconductor Diode

   a. In this discussion the term diode and rectifier will be used interchangeably; however, the term diode usually implies a small signal device with current typically in the milliamp range; and a rectifier, a power device, conducting from1 to 1000 amps or even higher. Many diodes or rectifiers are identified as 1NXXXX . A semiconductor diode consists of a PN junction and has two(2) terminals, an anode(+) and a cathode(-). Current flows from anode to cathode within the diode. A diode and schematic representation are shown In Figure "Semiconductors-001".

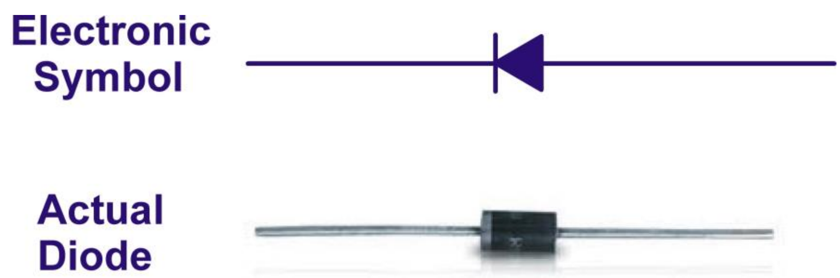


Figure 1, Semiconductors-001 Diode and Symbol

   b. An ideal diode is like a light switch in your home. When the switch is closed, the circuit is completed; and the light turns on.When the switch is open, there is no current and the light is off.

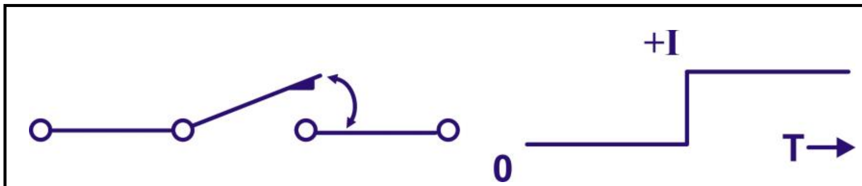


Figure 2, Semiconductors-002 Ideal Diode

   c. However, the diode has an additional property; it is unidirectional, i.e.current flows in only one direction(anode to cathode internally).When a forward voltage is applied, the diode conducts; and when a reverse voltage is applied, there is no conduction. A mechanical analogy is a ratchat, which allows motion in one direction only. An ideal diode characteristic would be as shown in Figure "Semiconductors-003".

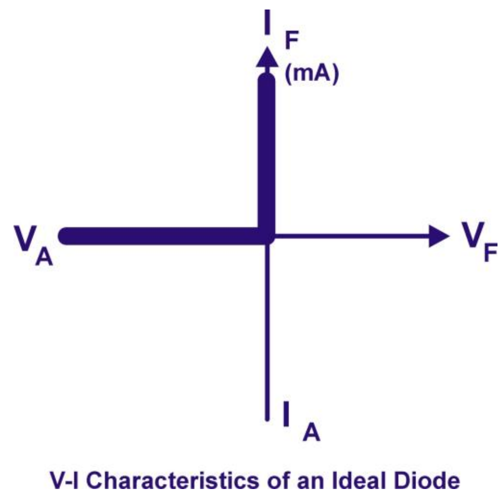


Figure 3, Semiconductors-003 Ideal Diode Characteristics

d. However, a typical diode characteristic is more like that of Figure "Semiconductors-004"

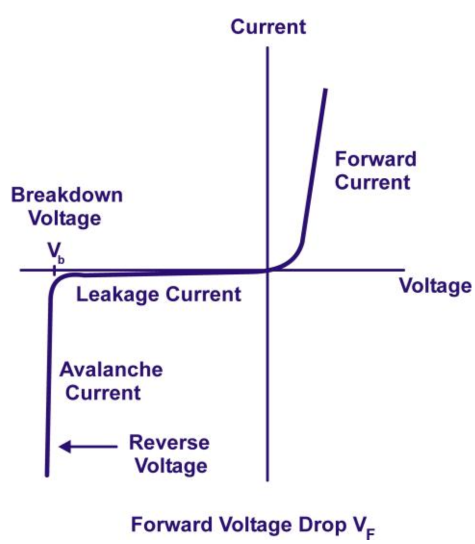


Figure 4, Semiconductors-004 Realistic Diode Characteristics

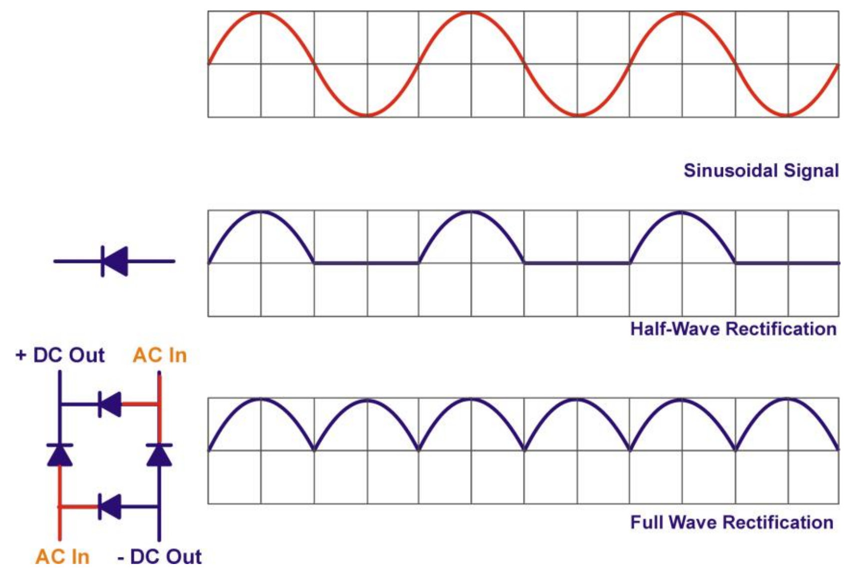


Figure 5, Semiconductors-005 Diode Wave Forms

e. Notice that the diode conducts a small current in the forward direction up to a threshold voltage, 0.3 for germanium and 0.7 for silicon; after that it conducts as we might expect. The forward voltage drop, V_f, is specified at a forward current, I_f.

f. Leakage Current

   1) In the reverse direction there is a small leakage current up until the reverse breakdown voltage is reached. This leakage is undesireable, obviously the lower the better, and is specified at a voltage less the than breakdown; diodes are intended to operate below their breakdown voltage.

g. Current Rating

   1) The current rating of a diode is determined primarily by the size of the diode chip, and both the material and configuration of the package, Average Current is used, not RMS current. A larger chip and package of high thermal conductivity are both conducive to a higher current rating.

h. Switching

1) The switching speed of a diode depends upon its construction and fabrication. In general the smaller the chip the faster it switches, other things being equal. The chip geometry, doping levels, and the temperature at nativity determine switching speeds . The reverse recovery time, trr, is usually the limiting parameter; trr is the time it takes a diode to switch from on to off.

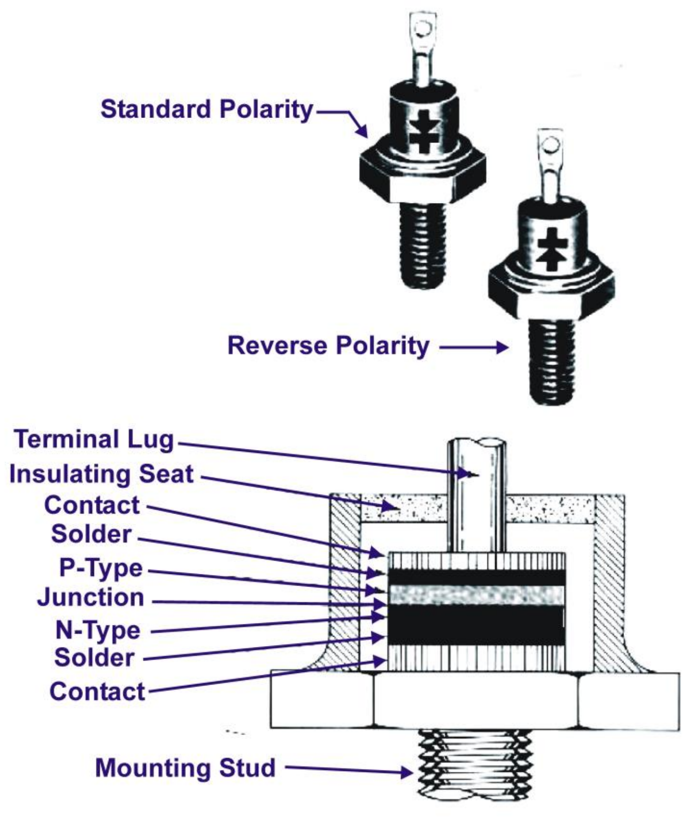


Figure 6, Semiconductors-006 Diode Cross-Section

2. Zener Diode

Figure 7, Semiconductors-007 Zener Diode Electronic Symbol

a. A Zener diode is a type of diode that permits current to flow in the forward direction like a normal diode, but also in the reverse direction if the voltage is larger (not equal to, but larger) than the rated breakdown voltage known as "Zener knee voltage" or "Zener voltage".

b. A conventional solid-state diode will not let significant current flow if it is reverse-biased below its reverse breakdown voltage. By exceeding the reverse bias breakdown voltage, a conventional diode is subject to high current flow due to avalanche breakdown. Unless this current is limited by external circuitry, the diode will be permanently damaged. In case of large forward bias (current flow in the direction of the arrow), the diode exhibits a voltage drop due to its junction built-in voltage and internal resistance. The amount of the voltage drop depends on the semiconductor material and the doping concentrations.

c. A Zener diode exhibits almost the same properties, except the device is especially designed so as to have a greatly reduced breakdown voltage, the so-called Zener voltage. A Zener diode contains a heavily doped p-n junction allowing electrons to tunnel from the valence band of the p-type material to the conduction band of the n-type material. In the atomic model, this tunneling corresponds to the ionization of covalent bonds. The Zener effect was discovered by the American physicist Clarence Melvin Zener. A reverse-biased Zener diode will exhibit a controlled breakdown and let the current flow to keep the voltage across the Zener diode at the Zener voltage. For example, a diode with a Zener breakdown voltage of 3.2 V will exhibit a voltage drop of 3.2 V if reverse bias voltage applied across it is more than its Zener voltage. However, the current is not unlimited, so the Zener diode is typically used to generate a reference voltage for an amplifier stage, or as a voltage stabilizer for low-current applications.

d. The breakdown voltage can be controlled quite accurately in the doping process. Tolerances to within 0.05% are available though the most widely used tolerances are 5% and 10%.

e. Another mechanism that produces a similar effect is the avalanche effect as in the avalanche diode. The two types of diode are in fact constructed the same way and both effects are present in diodes of this type. In silicon diodes up to about 5.6 volts, the zener effect is the predominant effect and shows a marked negative temperature coefficient. Above 5.6 volts, the avalanche effect becomes predominant and exhibits a positive temperature coefficient.

f. In a 5.6 V diode, the two effects occur together and their temperature coefficients neatly cancel each other out, thus the 5.6 V diode is the part of choice in temperature critical applications.

g. Modern manufacturing techniques have produced devices with voltages lower than 5.6 V with negligible temperature coefficients, but as higher voltage devices are encountered, the temperature coefficient rises dramatically. A 75 V diode has 10 times the coefficient of a 12 V diode.

h. All such diodes, regardless of breakdown voltage, are usually marketed under the umbrella term of 'zener diode'.

i. Zener diodes are widely used to regulate the voltage across a circuit. When connected in parallel with a variable voltage source so that it is reverse biased, a zener diode conducts when the voltage reaches the diode's reverse breakdown voltage. From that point it keeps the voltage at that value.

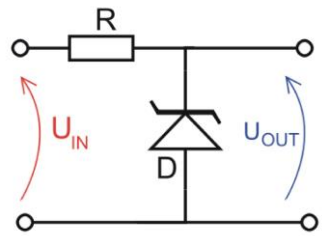


Figure 8, Semiconductors-008 Zener Diode Typical Circuit

j. In the circuit shown (Figure "Semiconductors-008"), resistor R provides the voltage drop between U_{IN} and U_{Out}. The value of R must satisfy two conditions:

k. R must be small enough that the current through D keeps D in reverse breakdown. The value of this current is given in the data sheet for D. For example, the common BZX79C5V6 device, a 5.6 V 0.5 W zener diode, has a recommended reverse current of 5 mA. If insufficient current flows through D, then U_{OUT} will be unregulated, and less than the nominal breakdown voltage (this differs to voltage regulator tubes where the output voltage will be higher then nominal and could rise as high as U_{IN}). When calculating R, allowance must be made for any current flowing through the external load, not shown in this diagram, connected across U_{Out}.

l. R must be large enough that the current through D does not destroy the device. If the current through D is I_D, its breakdown voltage V_B and its maximum power dissipation P_{MAX}, then $I_D V_B < P_{MAX}$.

m. A zener diode used in this way is known as a shunt voltage regulator (shunt, in this context, meaning connected in parallel, and voltage regulator being a class of circuit that produces a stable voltage across any load).

n. These devices are also encountered, typically in series with a base/emitter junction, in transistor stages where selective choice of a device centred around the avalanche\zener point can be used to introduce compensating temperature co-efficient balancing of the transistor PN junction. An example of this kind of use would be a d.c. error amplifier used in a stabilized power supply circuit feedback loop system.

3. Varactor

a. The varactor diode symbol is shown below with a diagram representation.

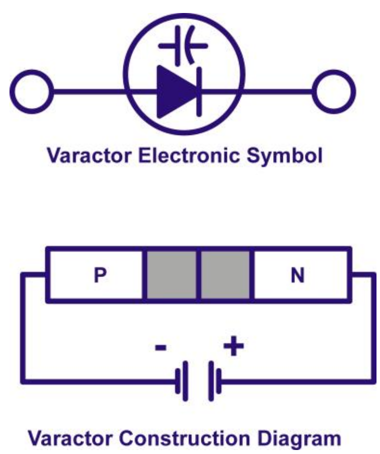


Figure 9, Semiconductors-009 Varactor Symbol and Construction Diagram

b. When a reverse voltage is applid to a PN junction , the holes in the p-region are attracted to the anode terminal and electrons in the n-region are attracted to the cathode terminal creating a region where there is little current.This region ,the depletion region, is essentially devoid of carriers and behaves as the dielectric of a capacitor.

c. The depletion region increases as reverse voltage across it increases; and since capacitance varies inversely as dielectic thickness, the junction capacitance will decrease as the voltage across the PN junction increases. So by varying the reverse voltage across a PN junction the junction capacitance can be varied .This is shown in the typical varactor voltage-capacitance curve below.

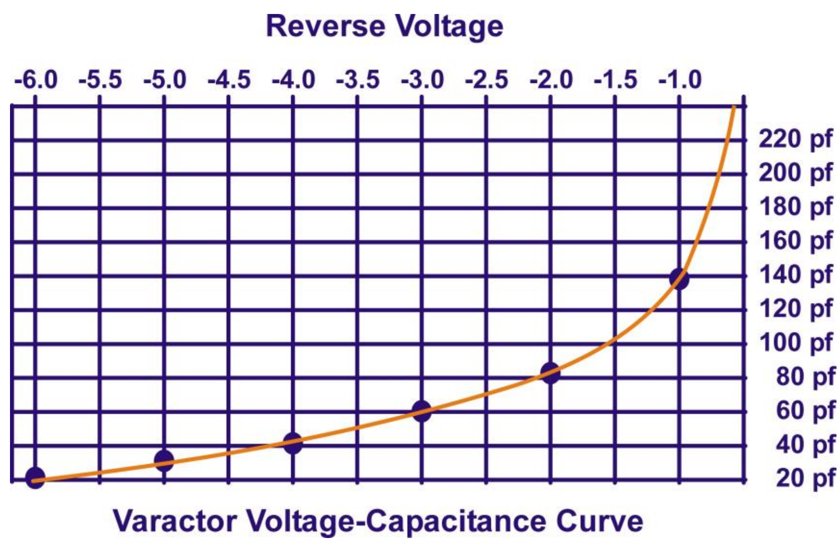


Figure 10, Semiconductors-010 Varactor Voltage-Capacitance Curve

d. Notice the nonlinear increase in capacitance as the reverse voltage is decreased. This nonlinearity allows the varactor to be used also as a harmonic generator.

4. Tunnel Diode

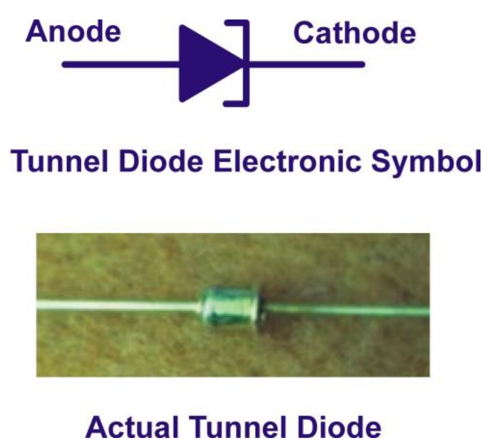


Figure 11, Semiconductors-011 Tunnel Diode Electronic Symbol

a. A tunnel diode is a semiconductor with a negative resistance region that results in very fast switching speeds , up to 5 GHz. The operation depends upon a quantum mechanic principle known as "tunneling" wherein the intrinsic voltage barrier (0.3 Volt for Germanium junctions) is reduced due to doping levels which enhance tunneling. Refering to the curves below, superimposing the tunneling characteristic upon a conventional P-N junction, we have:

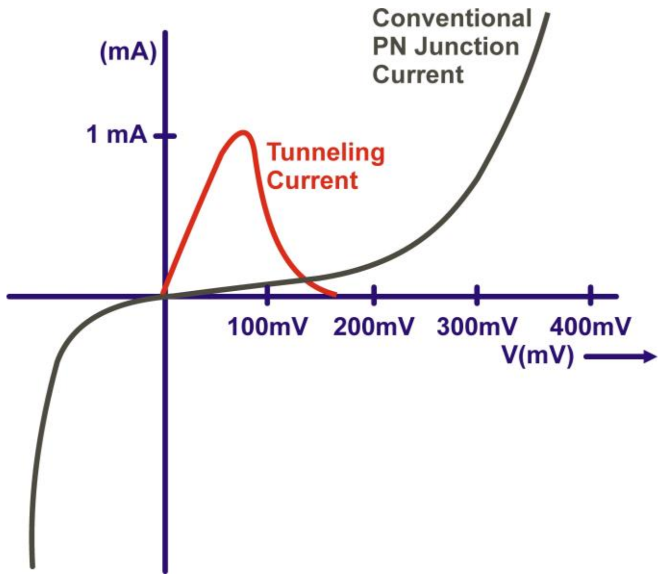


Figure 12, Semiconductors-012 Combination of Tunneling Current and Conventional PN Junction Current

b. Resulting in a composite characteristic which is the tunnel diode characteristic curve:

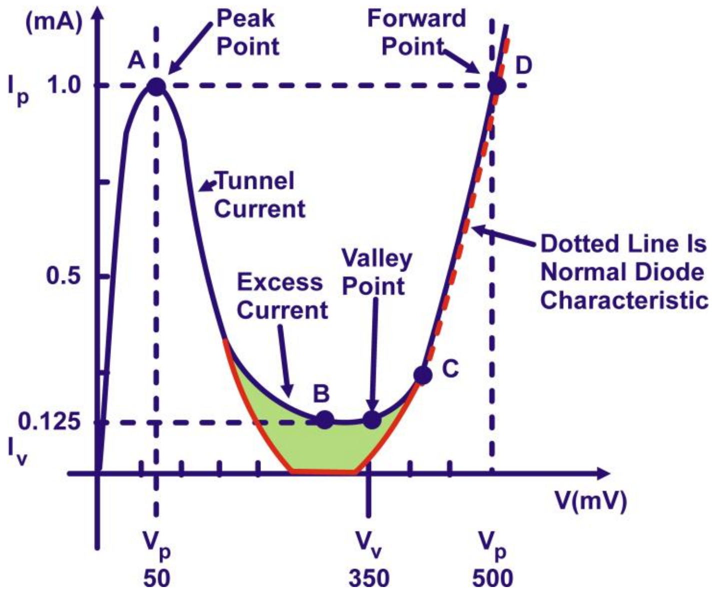


Figure 13, Semiconductors-013 Composite Characteristic Representing Overall Operational Curve Of A Tunnel Diode

c. The negative resistance region (between points A and B) is the important characteristic for the tunnel diode. In this region, as the voltage is increased, the current decreases; just the opposite of a conventional diode. The most important specifications for the tunnel diode are the Peak Voltage (Vp), Peak Current (Ip) , Valley Voltage (Vv), and Valley Current (Iv).

5. Back Diode

a. A Back diode is a tunnel diode with a suppressed Ip and so approximates a conventional diode characteristic See the comparison below:

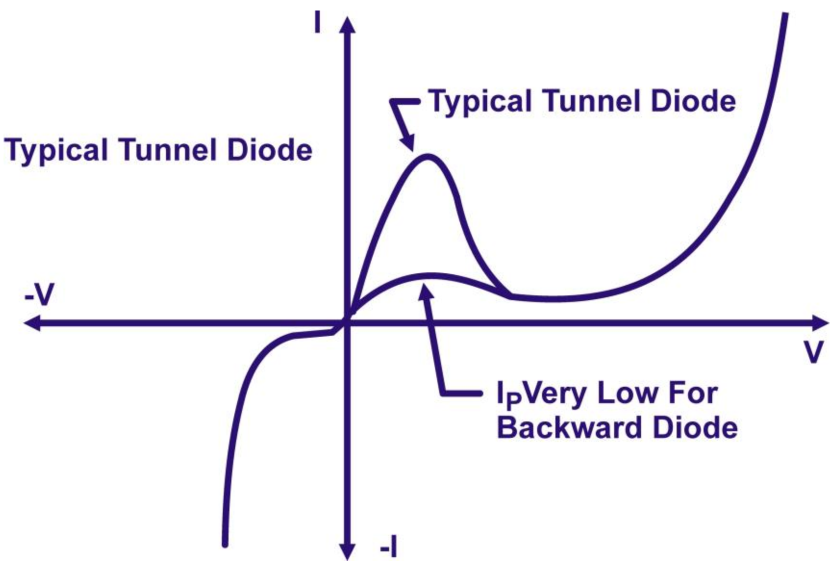


Figure 14, Semiconductors-014 Comparison of Forward Characteristics of Tunnel and Backward Diodes

b. The reverse breakdown for tunnel diodes is very low, typically 200mV, and the TD conducts very heavily at the reverse breakdown voltage. Referring to the BD curve the back diode conducts to a lesser degree in a forward direction . It is the operation between these two points that makes the back diode important. Forward conduction begins at 300 mV (for germanium) and a voltage swing of only 500mV is required for full range operation.

6. Thyristor

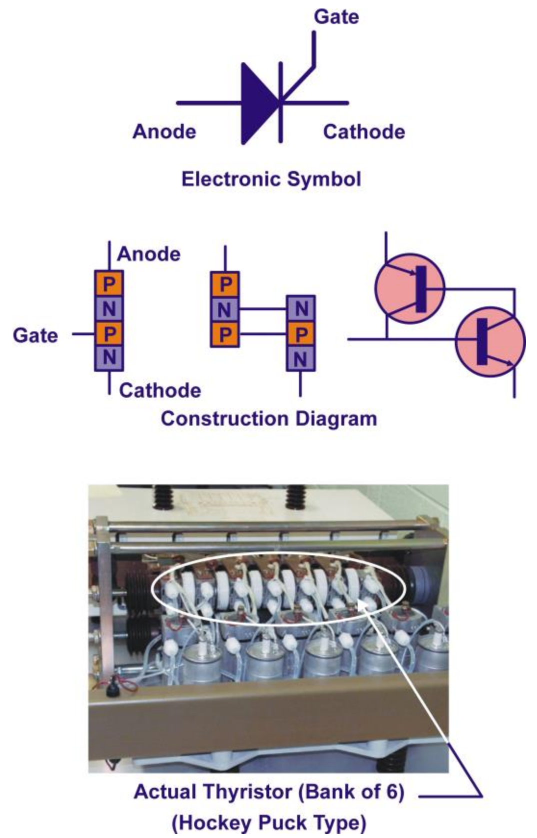


Figure 15, Semiconductors-015 Thyristor Electronic Symbol and Construction Diagram

a. The thyristor is a solid-state semiconductor device with four layers of alternating N and P-type material. They act as a switch, conducting when their gate receives a current pulse, and continue to conduct for as long as they are forward biased.

b. Some sources define silicon controlled rectifiers and thyristors as synonymous; others define SCRs as a subset of thyristors. Among the latter, the International Electrotechnical Commission 60747-6 standard stands out.

c. Non-SCR thyristors include devices with more than four layers, such as triacs and DB-GTOs.

d. The thyristor is a four-layer semiconducting device, with each layer consisting of an alternately N-type or P-type material, for example P-N-P-N. The main terminals, labeled anode and cathode, are across the full four layers, and the control terminal, called the gate, is attached to p-type material near to the cathode. (A variant called a SCS Silicon Controlled Switch brings all four layers out to terminals.) The operation of a thyristor can be understood in terms of a pair of tightly coupled Bipolar Junction Transistors, arranged to cause the self-latching action.

e. Thyristors have three states:

1) Reverse blocking mode -- Voltage is applied in the direction that would be blocked by a diode

2) Forward blocking mode -- Voltage is applied in the direction that would cause a diode to conduct, but the thyristor has not yet been triggered into conduction

3) Forward conducting mode -- The thyristor has been triggered into conduction and will remain conducting until the forward current drops below a threshold value known as the "holding current"

f. The thyristor has three p-n junctions (serially named J1, J2, J3 from the anode).

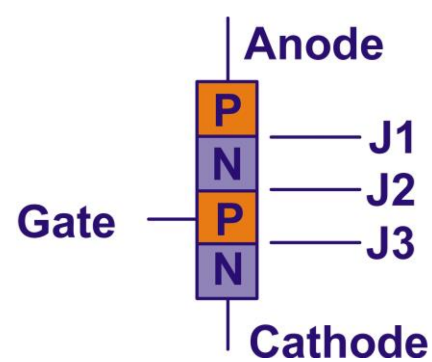


Figure 16, Semiconductors-016 Thyristor Layer Diagram

g. When the anode is at a positive potential V_{AK} with respect to the cathode with no voltage applied at the gate, junctions J1 and J3 are forward biased, while junction J2 is reverse biased. As J2 is reverse biased, no conduction takes place (Off state). Now if V_{AK} is increased beyond the breakdown voltage V_{BO} of the thyristor, avalanche breakdown of J2 takes place and the thyristor starts conducting (On state).

h. If a positive potential V_G is applied at the gate terminal with respect to the cathode, the breakdown of the junction J2 occurs at a lower value of V_{AK}. By selecting an appropriate value of V_G, the thyristor can be switched into the on state immediately.

i. It must be noted that V_G need not be applied after the avalanche breakdown has occurred. Hence V_G can be a voltage pulse, such as the voltage output from a UJT relaxation oscillator.

j. These gate pulses are characterized in terms of gate trigger voltage (V_{GT}) and gate trigger current (I_{GT}). Gate trigger current varies inversely with gate pulse width in such a way that it is evident that there is a minimum gate charge required to trigger the thyristor.

k. In a conventional thyristor, once it has been switched on by the gate terminal, the device remains latched in the on-state (i.e. do not need a continuous supply of gate current to conduct), providing the anode current has exceeded the latching current (I_L). As long as the anode remains positively biased, it cannot be switched off until the anode current falls below the holding current (I_H).

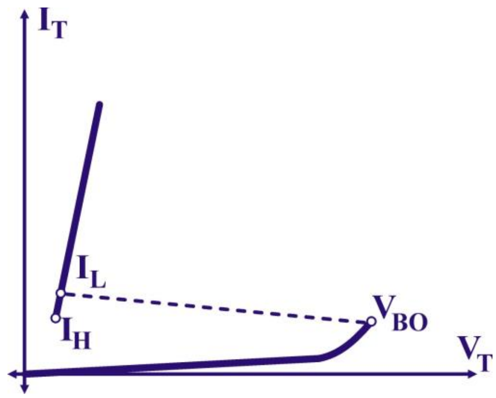


Figure 17, Semiconductors-017 Thyristor V-I Characteristics

l. A thyristor can be switched off if the external circuit causes the anode to become negatively biased. In some applications this is done by switching a second thyristor to discharge a capacitor into the cathode of the first thyristor. This method is called forced commutation.

m. After a thyristor has been switched off by forced commutation, a finite time delay must have elapsed before the anode can be positively biased in the off-state. This minimum delay is called the circuit commutated turn off time (tQ). Attempting to positively bias the anode within this time causes the thyristor to be self-triggered by the remaining charge carriers (holes and electrons) that have not yet recombined.

n. For applications with frequencies higher than the domestic AC mains supply (e.g. 50Hz or 60Hz), thyristors with lower values of tQ are required. Such fast thyristors are made by diffusing into the silicon heavy metals ions such as gold or platinum which act as charge combination centers. Alternatively, fast thyristors may be made by neutron irradiation of the silicon.

Discrete Electronic Components - Semiconductors

7. Silicon Controlled Rectifier (SCR)

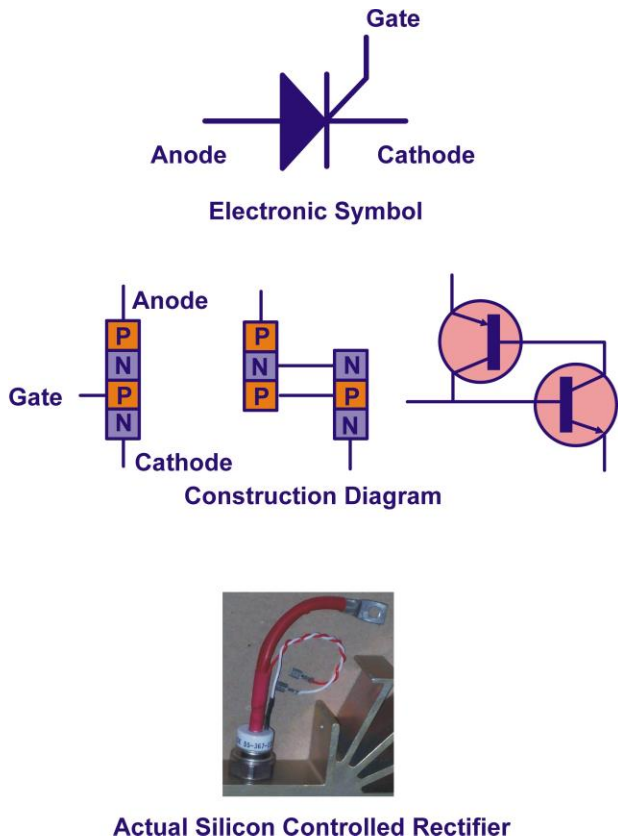


Figure 18, Semiconductors-018 Silicon Controlled Rectifier Electronic Symbol

a. The Silicon Controlled Rectifier (SCR) is simply a conventional rectifier controlled by a gate signal. The main circuit is a rectifier, however the application of a forward voltage is not enough for conduction. A gate signal controls the rectifier conduction. The schematic representation is:

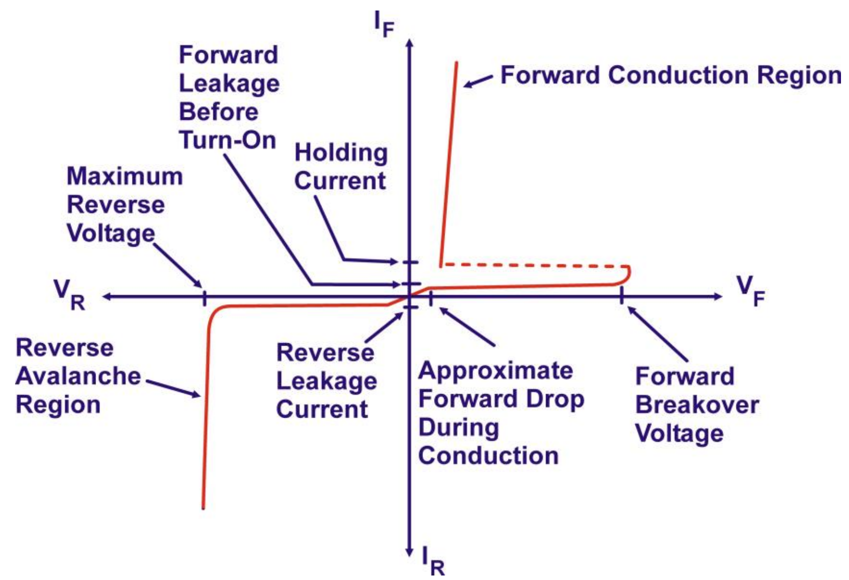


Figure 19, Semiconductors-019 Characteristics of SCR With Gate Open

b. Notice the reverse characteristics are the same as discussed previously for the rectifier or diode, having a breakover voltage with its attending avalanche current; and a leakage current for voltages less than the breakover voltage. However, in the forward direction with open gate, the SCR remains essentially in an off condition (notice though that there is a small forward leakage) up until the forward breakover voltage is reached. At that point the curve snaps back to a typical forward rectifier characteristic. The application of a small forward gate voltage switches the SCR onto its standard diode forward characteristic for voltages less than the forward breakover voltage.

c. Obviously, the SCR can also be switched by exceeding the forward breakover voltage, however this is usually considered a design limitation and switching is normally controlled with a gate voltage. One serious limitation of the SCR is the rate of rise of voltage with respect to time, dV/dt. A large rate of rise of circuit voltage can trigger an SCR into conduction. This is a circuit design concern. Most SCR applications are in power switching, phase control, chopper, and inverter circuits.

d. Thyristors and SCRs are used to control at what point in a cycle current is allowed to flow. This in turn controls effective power. In Figure "Semiconductors-020" , as the gate is triggered, the SCR (or Thyristor) conducts. The gate is triggered at different points in the cycle, thus controlling how much of the cycle the SCR (or Thyristor) conducts.

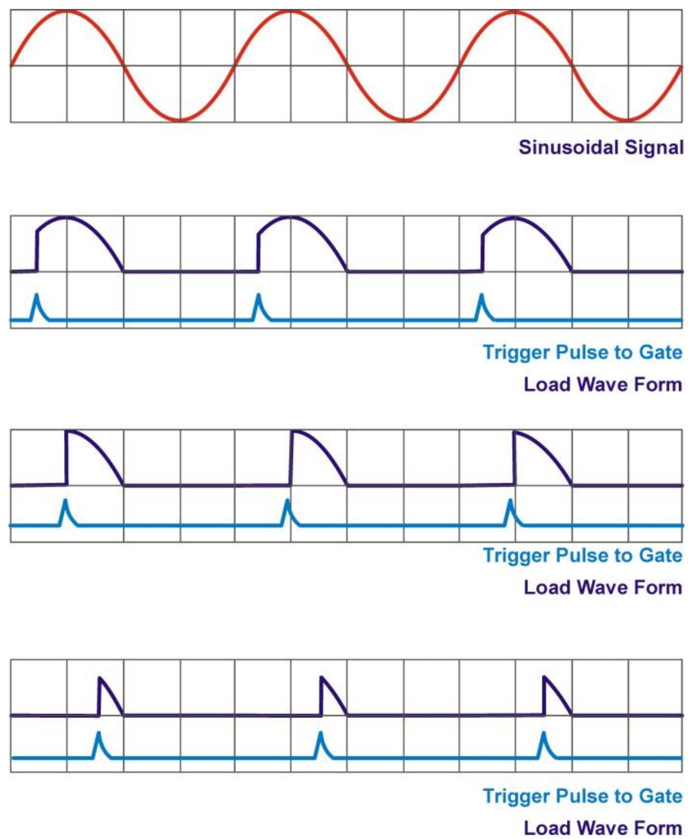


Figure 20, Semiconductors-020 Thyristor Wave Form as Triggered at Different Times In a Cycle

8. DIAC

Figure 21, Semiconductors-021 DIAC Electronic Symbol

a. The DIAC is a bidirectional trigger diode which is designed specifically to trigger a TRIAC or SCR. Basically the DIAC does not conduct (except for a small leakage current) until the breakover voltage is reached. At that point the diac goes into avalanche conduction also at that point the device exhibits a negative resistance characteristic, and the voltage drop across the DIAC snaps back, typically about 5 volts, creating a breakover current sufficient to trigger a TRIAC or SCR. A typical characteristic is shown in Figure "Semiconductors-022".

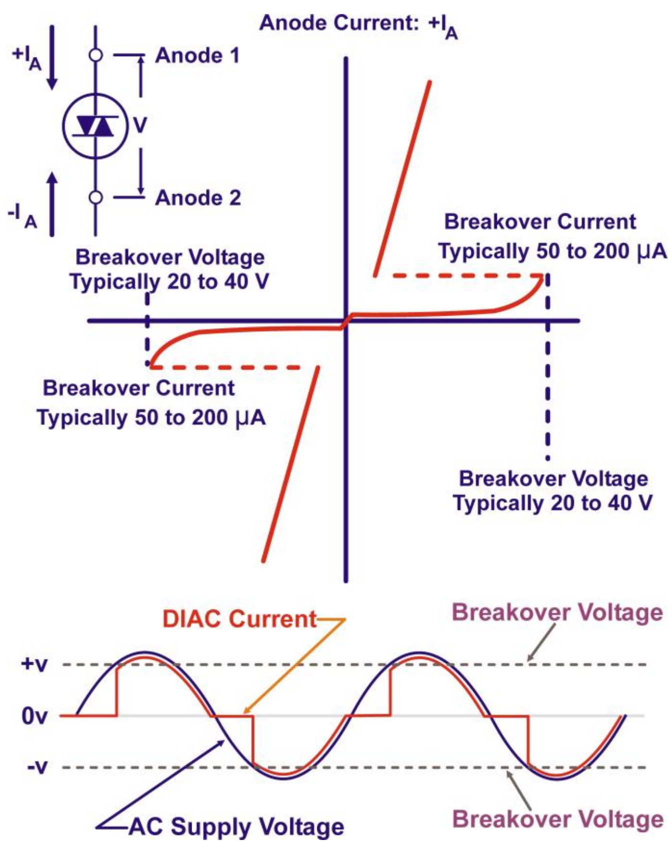


Figure 22, Semiconductors-022 DIAC Characteristic Curve

b. Although most DIACs have symmetric switching voltages, asymmetric DIACs are available. Typical DIACsa power dissipations ranging from 1/2 to 1 watt.

c. DIACs are typically used to trigger SCRs or Thyristors.

9. TRIAC

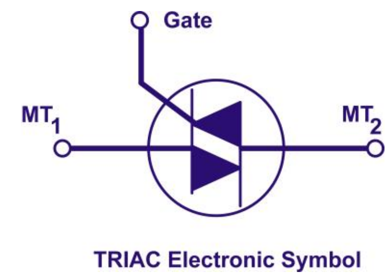


Figure 23, Semiconductors-023 TRIAC Electronic Symbol

a. The TRIAC is a three terminal semiconductor for controlling current in either direction. Below is the schematic symbol for the TRIAC. Notice the symbol looks like two SCRs in parallel(opposite direction) with one trigger or gate terminal.The main or power terminals are designated as MT_1 and MT_2 . (See the schematic representation below) When the voltage on the MT_2 is positive with regard to MT_1 and a positive gate voltage is applied, the left SCR conducts. When the voltage is reversed and a negative voltage is applied to the gate, the right SCR conducts. Minimum holding current, I_h, must be maintained in order to keep a TRIAC conducting.

b. A TRIAC operates in the same way as the SCR however it operates in both a forward and reverse direction. To get a quick understanding of its operation refer to its characteristic curve below and compare this to the SCR characteristic curve. It can be triggered into conduction by either a PLUS (+) or MINUS (-) gate signal.

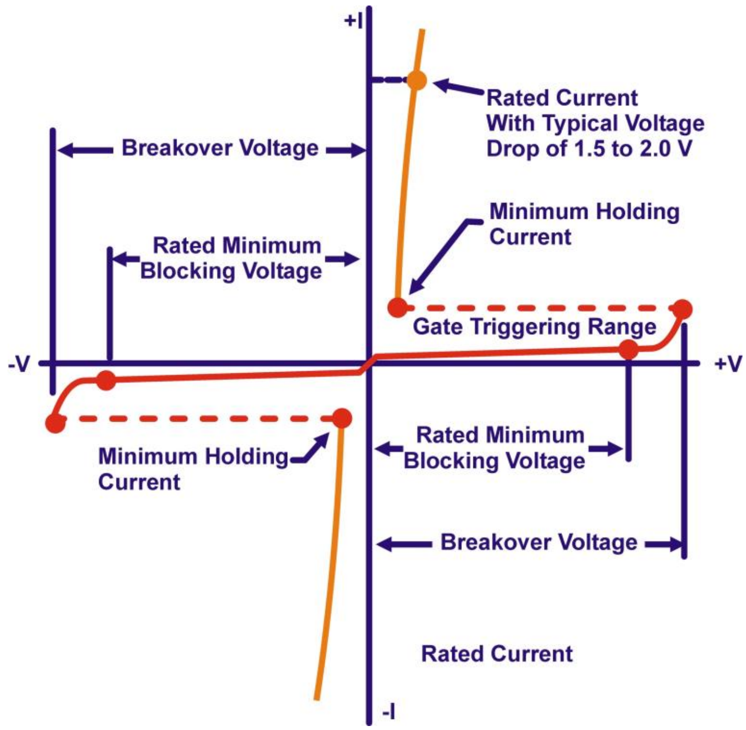


Figure 24, Semiconductors-024 TRIAC V-I Characteristic Curve

c. Obviously a TRIAC can also be triggered by exceeding the breakover voltage. This is not normally employed in triac operation. The breakover voltage is usually considered a design limitation. One other major limitation, as with the SCR, is dV/dt, which is the rate of rise of voltage with respect to time. A triac can be switched into conduction by a large dV/dt. Typical applications are in phase control, inverter design, AC switching, relay replacement, etc.

d. TRIAC waveforms are similar to SCR waveforms except that a single TRIAC will control both halves of a cycle.

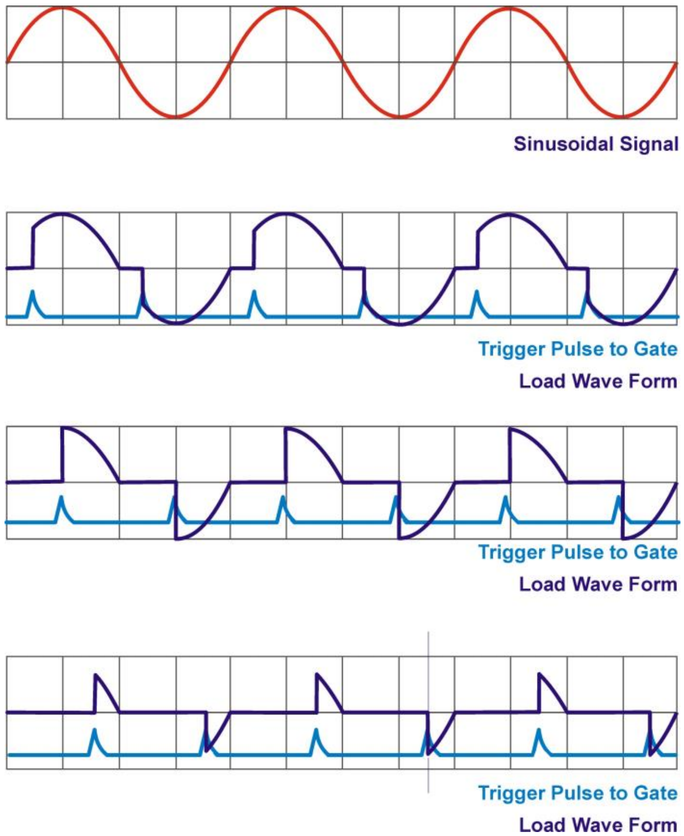


Figure 25, Semiconductors-025 TRIAC Waveforms

e. As shown in Figure "Semiconductors-025", the further into the cycle that the TRIAC is triggered, the less time current flows through the load, thus transferring less and less power to the load.

10. NPN and PNP Transistors (Bipolar Transistors)

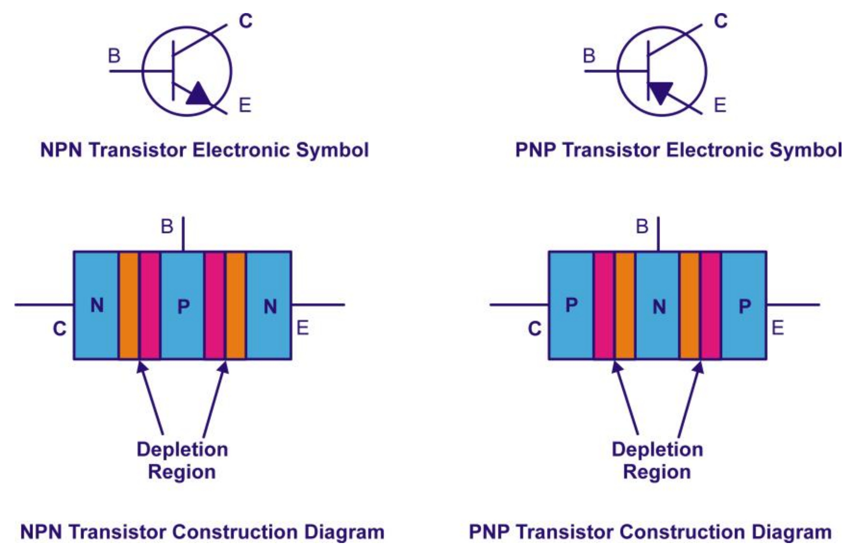


Figure 26, Semiconductors-026 NPN Transistor Electronic Symbol

a. An NPN transistor is made by sandwiching a 'P' type wafer between two 'N' type wafers.

b. An PNP transistor is made by sandwiching a 'N' type wafer between two 'P' type wafers

c. The 'P' and 'N' do not denote polarity as in positive and negative, but the type of material from which the wafer is made.

d. Note that the direction of the emitter arrow defines the type transistor. Biasing and power supply polarity are positive for NPN and negative for PNP transistors. The transistor is primarily used as an current amplifier. When a small current signal is applied to the base terminal, it is amplified in the collector circuit. This current amplification is referred to as H_{FE} or beta and equals I_c/I_b.

e. As with all semiconductors, breakdown voltage is a design limitation. There are breakdown voltages that must be taken into account for each combination of terminals. i.e. V_{ce}, V_{be},and V_{cb}. However, V_{ce} (collector-emitter voltage) with open base, designated as V_{ceo}, is usually of most concern and defines the maximum circuit voltage.

f. Also as with all semiconductors there are undesireable leakage currents, notably I_{cbo} ,collector junction leakage; and I_{ebo}, emitter junction leakage. A typical collector characteristic curve is shown below:

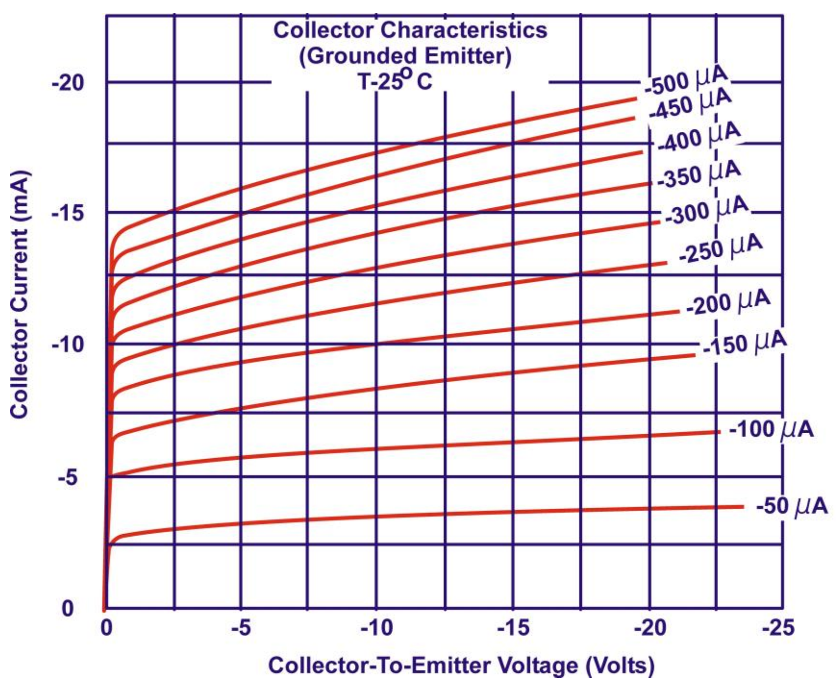


Figure 27, Semiconductors-027 Typical Transistor Collector Characteristic Curves

g. Note that the negative collector-emitter voltage tells you that the transistor is PNP. Also that the output current increases with input or base current and varies very little with collector-emitter voltage.

h. A transistor can also be thought of simply as a small current variable resistor controlling a larger current variable resistor.

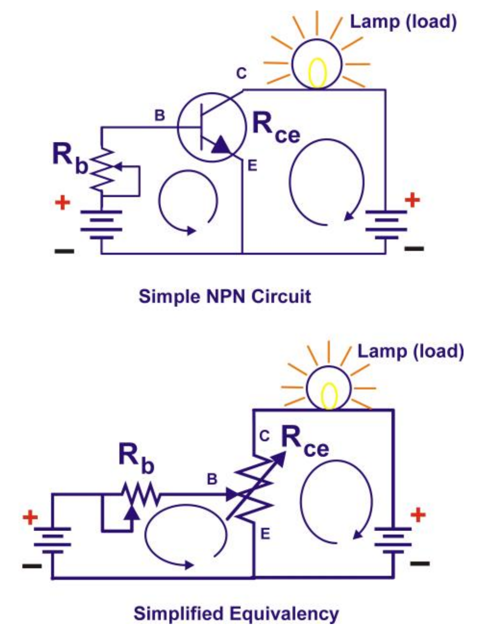


Figure 28, Semiconductors-028 Simple NPN Transistor Circuit

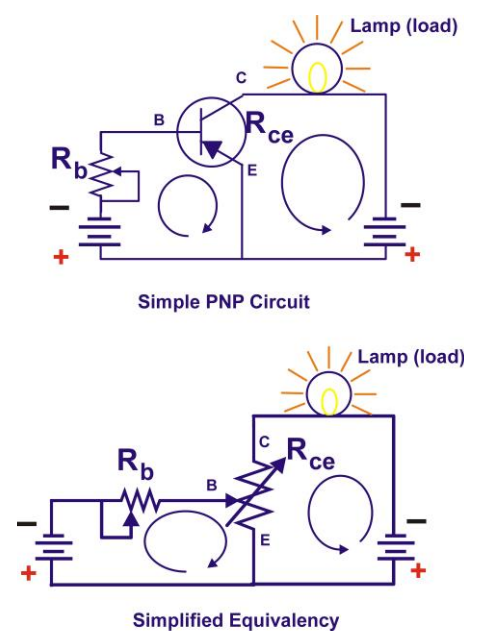


Figure 29, Semiconductors-029 Simple PNP Transistor Circuit

i. Refer to Figures "Semiconductors-028 & 29"

j. As the base circuit resistance (R_b) decreases, current through the base increases. The resistance of the collector - emitter circuit (R_{ce}) also decreases, but by a greater amount. Lowering the resistance of the base circuit increases the current through the base circuit (ohms law). Lowering the collector-emitter circuit resistance increases the collector - emitter current, but by a larger amount. This is known as amplification.

k. Within operating parameters, the voltage drop between the base of the transistor and the emitter of the transistor remains constant (about 0.7 volts) regardless of the current through the base-emitter junction. However, the voltage between the collector and the emitter of the transistor will vary as the current through the collector-emitter circuit varies. The transistor is inherently a current device.

11. Field Effect Transistor (FET)

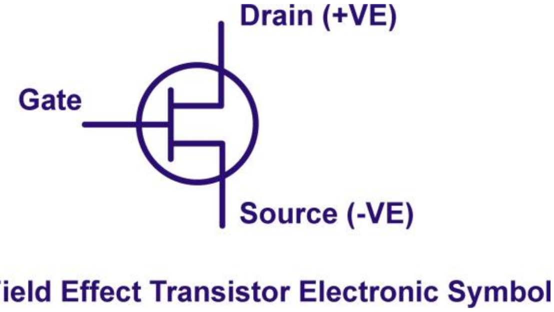


Figure 30, Semiconductors-030 Field Effect Transistor Electronic Symbol

a. The FET can be constructed from a number of semiconductors, silicon being by far the most common. Most FETs are made with conventional bulk semiconductor processing techniques, using the single crystal semiconductor wafer as the active region, or channel.

b. Among the more unusual body materials are amorphous silicon, polycrystalline silicon or other amorphous semiconductors in thin-film transistors or organic field effect transistors that are based on organic semiconductors and often apply organic gate insulators and electrodes.

c. All FETs except J-FETs have four terminals, which are known as the gate, drain, source and body/base/bulk. Compare these to the terms used for Bipolare Junction Transistors (BJTs): base, collector and emitter. BJTs and J-FETs have no body terminal.

d. The names of the terminals refer to their function. The gate terminal may be thought of as controlling the opening and closing of a physical gate. This gate permits electrons to flow through or blocks their passage. Electrons flow from the source terminal towards the drain terminal if influenced by an applied voltage. The body simply refers to the bulk of the semiconductor in which the gate, source and drain lie. Usually the body terminal is connected to the highest or lowest voltage within the circuit, depending on type. The body terminal and the source terminal are sometimes connected together since the source is also sometimes connected to the highest or lowest voltage within the circuit, however there are several uses of FETs which do not have such a configuration, such as transmission gates and cascode circuits.

e. The body of a FET is either doped to produce an N-type semiconductor or a P-type semiconductor. The drain and source may be doped of opposite type to the body, in the case of enhancement mode FETs, or doped of similar type to the body as in depletion mode FETs. Field-effect transistors are also distinguished by the method of insulation between body and gate. Types of FETs are:

  1) The MOSFET (Metal-Oxide-Semiconductor Field-Effect Transistor) utilizes an insulator (typically SiO2) between the gate and the body

  2) The JFET (Junction Field-Effect Transistor) uses a reverse biased p-n junction to separate the gate from the body.

3) The MESFET (Metal-Semiconductor Field-Effect Transistor) substitutes the p-n junction of the JFET with a Schottky barrier; used in GaAs and other III-V semiconductor materials.

4) Using bandgap engineering in a ternary semiconductor like AlGaAs gives a HEMT (High Electron Mobility Transistor), also called an HFET (heterostructure FET). The fully depleted wide-band-gap material forms the isolation between gate and body.

5) The MODFET (Modulation-Doped Field Effect Transistor) uses a quantum well structure formed by graded doping of the active region.

6) The IGBT (insulated-gate bipolar transistor) is a device for power control. It has a structure akin to a MOSFET coupled with a bipolar-like main conduction channel. These are commonly used for the 200-3000 V drain-to-source voltage range of operation. Power MOSFETs are still the device of choice for drain-to-source voltages of 1 to 200 V.

f. The FET controls the flow of electrons from the source to drain by affecting the size and shape of a "conductive channel" created and influenced by voltage (or lack of voltage) applied across the gate and source terminals. (For ease of discussion, this assumes body and source are connected). This conductive channel is the "stream" through which electrons flow from source to drain.

g. Consider an n-channel "depletion-mode" device. A negative gate-to-source voltage causes a depletion region to expand in width and encroach on the channel from the sides, narrowing the channel. If the depletion region expands to completely close the channel, the resistance of the channel from source to drain becomes large, and the FET is effectively turned off like a switch. Likewise a positive gate-to-source voltage increases the channel size and allows electrons to flow easily.

h. Now consider an n-channel "enhancement-mode" device. A positive gate-to-source voltage is necessary to create a conductive channel, since one does not exist naturally within the transistor. The positive voltage attracts free-floating electrons within the body towards the gate, forming a conductive channel. But first, enough electrons must be attracted near the gate to counter the dopant ions added to the body of the FET; this forms a region free of mobile carriers called a depletion region, and the phenomenon is referred to as the threshold voltage of the FET. Further gate-to-source voltage increase will attract even more electrons towards the gate which are able to create a conductive channel from source to drain; this process is called inversion.

i. For either enhancement- or depletion-mode devices, at drain-to-source voltages much less than gate-to-source voltages, changing the gate voltage will alter the channel resistance. In this mode the FET operates like a variable resistor and the FET is said to be operating in a linear mode. This mode is not employed when amplification is needed.

j. If drain-to-source voltage is increased, this creates a significant asymmetrical change in the shape of the channel due to a gradient of voltage potential from source to drain. The shape of the inversion region becomes "pinched-off" near the drain end of the channel. If drain-to-source voltage is increased further, the pinch-off point of the channel begins to move away from the drain towards the source. The FET is said to be in saturation mode.

k. Even though the conductive channel formed by gate-to-source voltage no longer connects source to drain during saturation mode, carriers are not blocked from flowing. Considering again an n-channel device, a depletion region exists in the p-type body, surrounding the conductive channel and drain and source regions. The electrons which comprise the channel are free to move out of the channel through the depletion region if attracted to the drain by drain-to-source voltage. The depletion region is free of carriers and has a resistance similar to silicon. Any increase of the drain-to-source voltage will increase the distance from drain to the pinch-off point, increasing resistance due to the depletion region proportionally to the applied drain-to-source voltage. This proportional change causes the drain-to-source current to remain relatively fixed independent of changes to the drain-to-source voltage and quite unlike the linear mode operation. Thus in saturation mode, the FET behaves as a constant-current source rather than as a resistor and can be used most effectively as a voltage amplifier. In this case, the gate-to-source voltage determines the level of constant current through the channel.

l. The most commonly used FET is the MOSFET. The CMOS (complementary-symmetry metal oxide semiconductor) process technology is the basis for modern digital integrated circuits. This process technology uses an arrangement where the (usually "enhancement-mode") p-channel MOSFET and n-channel MOSFET are connected in series such that when one is on, the other is off.

m. The fragile insulating layer of the MOSFET between the gate and channel makes it vulnerable to electrostatic damage during handling. This is not usually a problem after the device has been installed.

n. In FETs electrons can flow in either direction through the channel when operated in the linear mode, and the naming convention of drain terminal and source terminal is somewhat arbitrary, as the devices are typically (but not always) built symmetrically from source to drain. This makes FETs suitable for switching analog signals between paths (multiplexing). With this concept, one can construct a solid-state mixing board, for example.

12. Junction Field Effect Transistor (JFET)

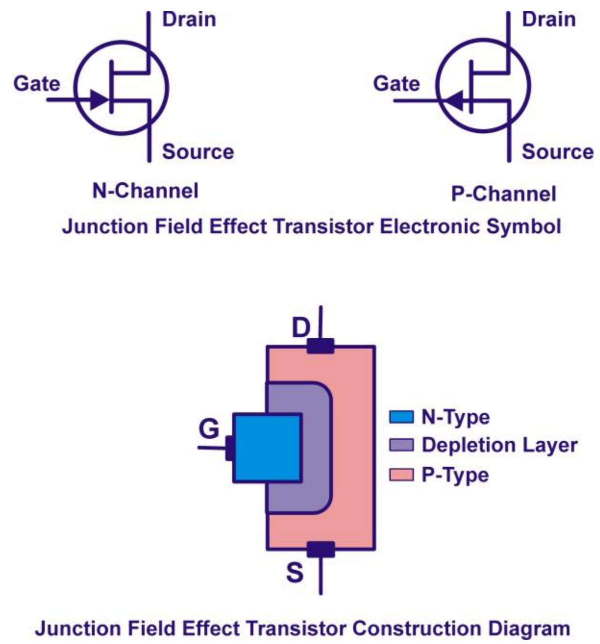


Figure 31, Semiconductors-031 Junction Field Effect Transistor Electronic Symbol

a. Sometimes the JFET gate is drawn in the middle of the channel instead of at the drain/source electrode as in these examples. The US style of the symbol is to draw the whole component inside a circle, whilst the european style is to draw it without a circle.

b. The symmetric variation is hinting at that the channel is indeed symmetric in sense that drain and source are interchangeable physical terminals.

c. In every case the arrow head is telling the direction, where the P-N-junction of the gate is in relationship to the channel. In order to pinch off the channel, one must produce around 2 volts in reverse direction (VGS) of that junction.

d. In N-type channel that is hinting to us that negative voltage at the gate in comparison of the source is called for. In P-type channel the hint is for positive VGS.

e. The junction gate field-effect transistor (JFET or JUGFET) is the simplest type of field effect transistor. Like other transistors, it can be used as an electronically-controlled switch. They are also used as voltage-controlled resistances. An electric current flows from one connection, called the source, to a second connection, called the drain. A third connection, the gate, determines how this current flows. By applying an increasing negative (for an n-channel JFET) bias voltage to the gate, the current flow from source to drain can be impeded by pinching off the channel, in effect switching

f. The JFET consists of a long channel of semiconductor material. This material is doped so that it contains an abundance of positive charge carriers (p-type), or of negative charge carriers (n-type). There is a contact at each end; these are the source and drain. The third control terminal, the gate, surrounds the channel, and is doped opposite to the doping-type of the channel.Then, a pn junction is formed at the interface of the two types of the material and one has to make sure that the terminal made with the semiconductor are usually made Ohmic.

g. The operation of a JFET can easily be understood by considering a garden hose. The flow of water through a garden hose can be controlled by squeezing it and reducing its cross section; the flow of electric charge through a JFET is controlled by constricting the cross section of the current-carrying channel.

h. The JFET gate presents a small current load which is the reverse leakage of the gate-to-channel junction. The MOSFET has the advantage of extremely low gate current (measured in picoamperes) because of the insulating oxide between the gate and channel. However, compared to the base current of a bipolar junction transistor the JFET gate current is much lower, and the JFET has higher transconductance than the MOSFET. Therefore JFETs are used to advantage in some low-noise, high input-impedance op-amps and sometimes used in switching applications.

i. The JFET had been predicted as early as 1925 by Julius Lilienfeld, and the theory of operation of the device was sufficiently well known by the mid 1930's for a patent to be issued for it. However, technology at the time was not sufficiently advanced to produce doped crystals with enough precision for the effect to be seen until many years later. In 1947, researchers John Bardeen, Walter Houser Brattain, and William Shockley were attempting to construct a JFET when they discovered the bipolar junction transistor. The first practical JFETs were thus constructed many years after the first bipolar junction transistors, in spite of having been invented much earlier.

13. Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

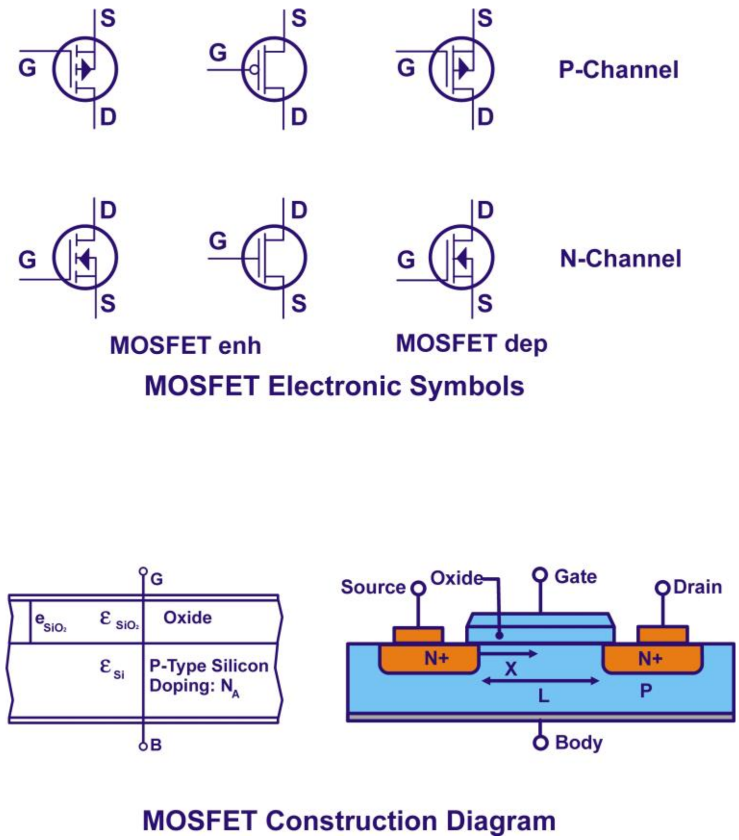


Figure 32, Semiconductors-032 Metal Oxide Field Effect Transistor Electronic Symbol

a. A FET with an oxide coating between gate and channel is called a MOSFET (metal- oxide semiconductor field effect transistor). As a result, the MOSFET has very high input resistance, higher than the JFET ; and as with the JFET, the gate controls the main or channel current, I_{ds}.

b. Notice that a postive gate voltage will induce a negative charge in the n-channel, enhancing the drain-source current, Ids; while a negative gate voltage will induce a depletion region in the n- channel, thereby reducing the drain-source current, I_{ds}. A MOSFET so constructed is a depletion/enhancement MOSFET. A typical n-channel MOSFET curve is shown Figure "Semiconductors-033".

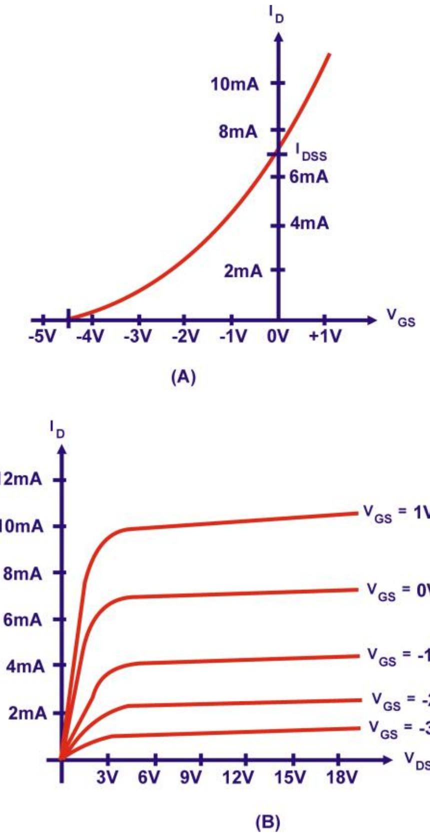


Figure 33, Semiconductors-033 MOSFET N-Channel Characteristic Curves

c. Figure "Semiconductors-033" (A), shows the transconductance, Gm, or effect of the gate voltage upon the drain to source current. Notice the gate-source voltage can be positive or negative. This would not be possible with a JFET. In Figure "Semiconductors-033" (B), again notice the effect of increasing Vgs on Ids; notice also that for any given gate current, the drain-source voltage has little effect upon the drain current above 3 volts, since the MOSFET is in saturation.

d. The MOSFET is also made in an enhancement-only mode, where a gate signal only induces or enhances channel current ,the gate signal never depletes the channel current. Naturally there are p-channel enhancement MOSFETS, where a negative gate voltage enhances channel conductivity; and n-channel enhancement mode MOSFETS where a positive gate voltage enhances channel conductivity. One final note, breakdown voltage in MOS devices do not depend upon p-n junction stress but rather upon the thickness and quality of the insulating oxide.When breakdown does occur, the oxide is punctured and the device is destroyed. The specifications for MOSFETS or any transistor are multitudinous, and so only the basics are listed. They are:

1) Breakdown Voltages
2) Forward transconductance
3) Drain- source on resistance, $R_{ds}(on)$
4) Switching characteristics
5) Zero gate voltage drain current, I_{dss}
6) Input capacitance, C_{iss}

14. Unijunction Transistor (UJT)

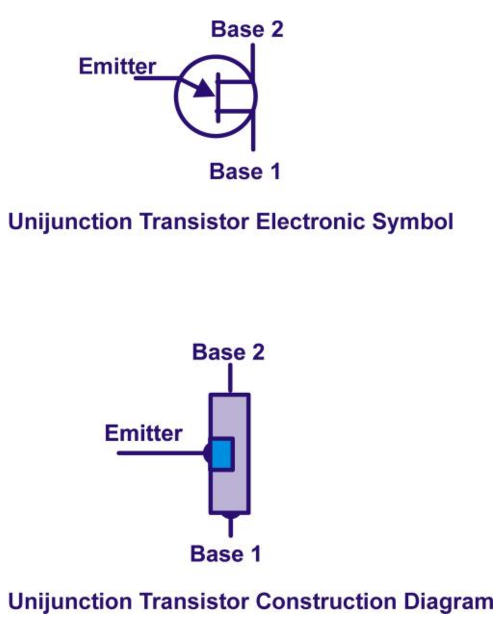


Figure 34, Semiconductors-034 Unijunction Transistor Electronic Symbol

a. The unijunction transistor (UJT) is a three terminal device with characteristics very different from the conventional 2 junction, bipolar transistor. It is a pulse generator with the trigger or control signal applied at the emitter . This trigger voltage is a fraction (n) of interbase voltage, Vbb. Figure "Semiconductors-035" illustrates a UJT characteristic curve.

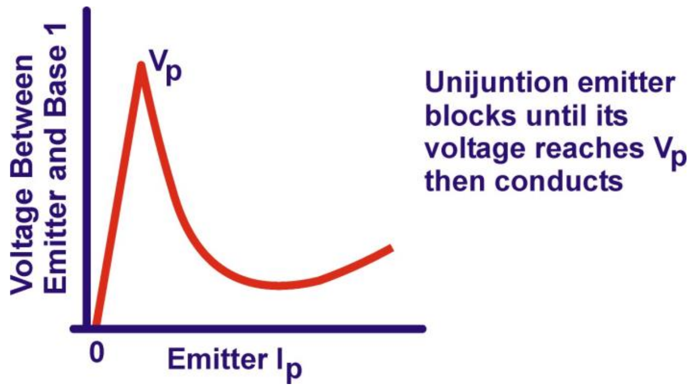


Figure 35, Semiconductors-035 Unijunction Transistor Characteristic Curve

b. The emitter terminal does not inject current into the base region until its voltage reaches V_p. Once V_p is reached the base circuit conducts and a postive pulse appears at the B_1 terminal and a negative pulse at B_2. The UJT incorporates a negative resistance region, a low emitter current, and a high output pulse current at terminals B_1 and B_2, making it an ideal pulse trigger. A simple RC timer circuit using a UJT is shown in Figure "Semiconductors-036".

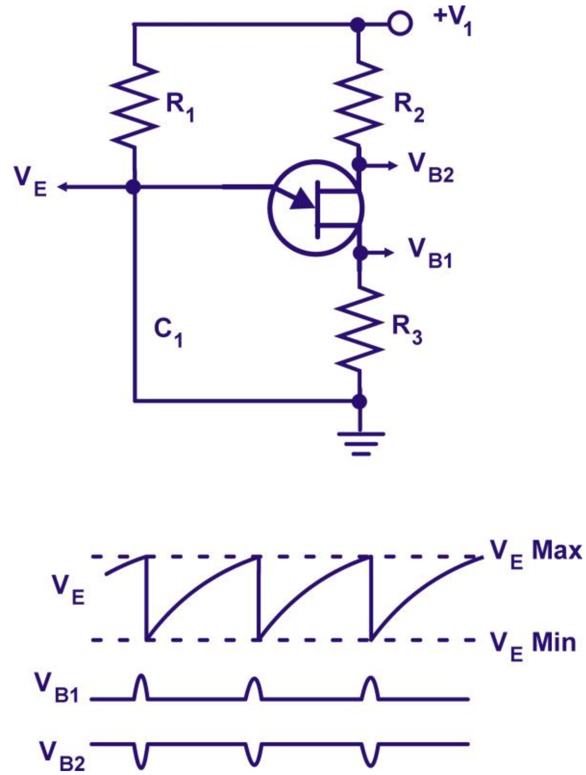
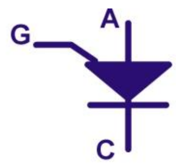
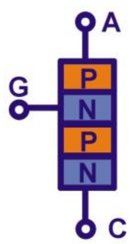


Figure 36, Semiconductors-036 Basic UJT Relaxation Oscillator

15. Programmable Unijunction Transistor (PUT)



Programmable Unijunction Transistor
Electronic Symbol



Programmable Unijunction Transistor
Construction Diagram

Figure 37, Semiconductors-037 Programmable Unijunction Transistor Electronic Symbol

a. The Programmable Unijunction Transistor behaves much like a unijunction transistor (UJT), but is "programmable" via external resistors (that is, you can use two resistors to set a PUT's peak voltage). Note that the name is a bit of a misnomer -- as a thyristor, it is a four layer device, unlike a true unijunction transistor which has but two layers.

b. Like other thyristors, a PUT looks much like a junction transistor with a fourth layer and therefore a total of three P-N junctions. Meanwhile, a third terminal, the gate (G), makes a PUT function like a hybrid of transistor and diode.

c. PUTs are not often used in BEAM (BEAM is an acronym standing for Biology, Electronics, Aesthetics, Mechanics); they're essentially special-purpose devices in electronics, used for lighting control, motor speed control and other variable power applications. In combination with an SCR they can, though, make a mean solar engine.

d. In a pinch, you can build up something much like a PUT from discrete transistors wired as a complementary feedback pair:

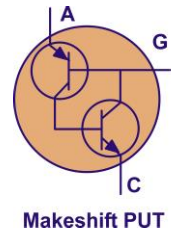


Figure 38, Semiconductors-038 Makeshift PUT

e. Here, as soon as any current flows in either transistor, this current becomes base current for the other transistor, and both transistors turn on hard. This means you can only build up this circuit using low-leakage transistors ('though this should be the case with any decent-quality modern transistor).

f. As part of a larger circuit, the pseudo-UJT would look like this, including its "programming" resistors:

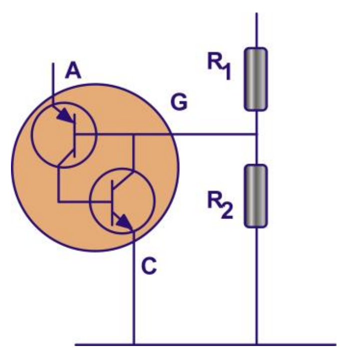


Figure 39, Semiconductors-039 Programmed Makeshift PUT

16. Photoresistor (Photoconductive cell)

a. A photoresistor is an electronic component whose resistance decreases with increasing incident light intensity. It can also be referred to as a light-dependent resistor (LDR), photoconductor, or photocell.

b. A photoresistor is made of a high-resistance semiconductor. If light falling on the device is of high enough frequency, photons absorbed by the semiconductor give bound electrons enough energy to jump into the conduction band. The resulting free electron (and its hole partner) conduct electricity, thereby lowering resistance.

c. A photoelectric device can be either intrinsic or extrinsic. An intrinsic semiconductor has its own charge carriers and is not an efficient semiconductor, eg. silicon. In intrinsic devices, the only available electrons are in the valence band, and hence the photon must have enough energy to excite the electron across the entire bandgap. Extrinsic devices have impurities added, which have a ground state energy closer to the conduction band — since the electrons don't have as far to jump, lower energy photons (i.e. longer wavelengths and lower frequencies) are sufficient to trigger the device. If a sample of silicon has some of its atoms replaced by phosphorus atoms(impurities), there will be extra electrons available for conduction. This is an example of an extrinsic semiconductor.

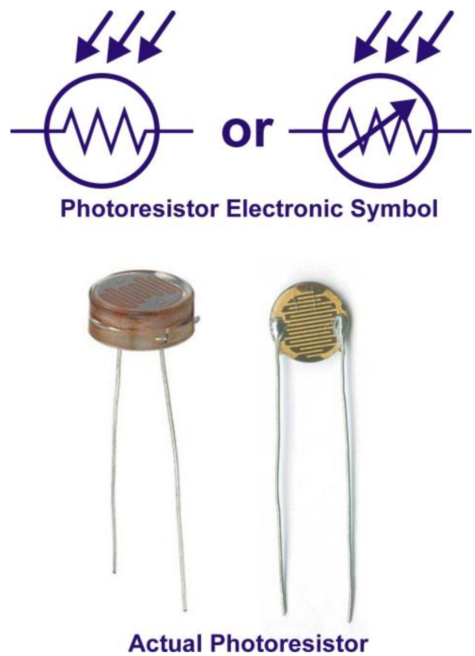


Figure 40, Semiconductors-040 Photoresistor Electronic Symbol

d. Cadmium sulphide or cadmium sulphide (CdS) cells rely on the material's ability to vary its resistance according to the amount of light striking the cell. The more light that strikes the cell, the lower the resistance. Although not accurate, even a simple CdS cell can have a wide range of resistance from less than 100 Ω in bright light to in excess of 10 MΩ in darkness. Many commercially available CdS cells have a peak sensitivity in the region of 500nm - 600nm. The cells are also capable of reacting to a broad range of frequencies, including infrared (IR), visible light, and ultraviolet (UV). They are often found on street lights as automatic on/off switches. They were once even used in heat-seeking missiles to sense for targets.

Figure 41, Semiconductors-041 Various Photoresistors

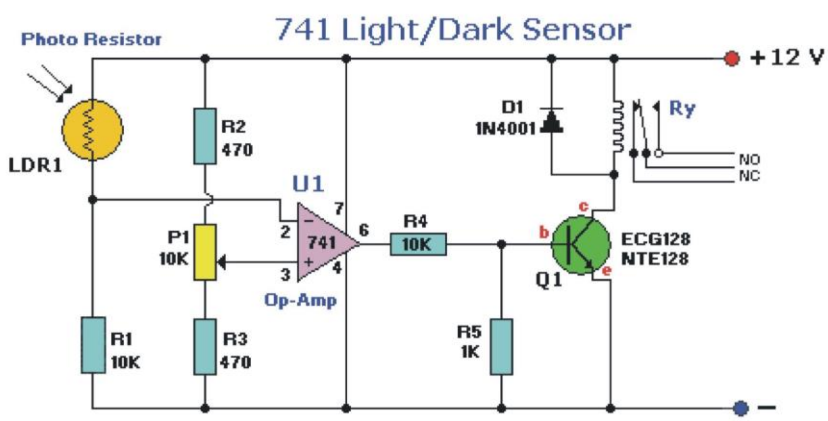


Figure 42, Semiconductors-042 Typical Circuit Using Photoresistor

17. Photodiodes

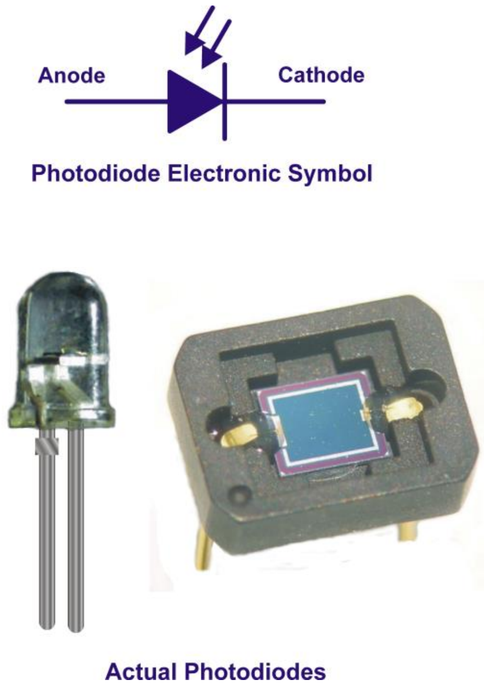


Figure 43, Semiconductors-043 Photodiode Electronic Symbol

a. A photodiode is a semiconductor diode that functions as a photodetector. Photodiodes are packaged with either a window or optical fiber connection, in order to let in the light to the sensitive part of the device. They may also be used without a window to detect vacuum UV or X-rays.

b. A photodiode is a p-n junction or p-i-n structure. When light of sufficient photon energy strikes the diode, it excites an electron thereby creating a mobile electron and a positively charged electron hole. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in field of the depletion region, producing a photocurrent.

c. Photodiodes can be used under either zero bias (photovoltaic mode) or reverse bias (photoconductive mode). In zero bias, light falling on the diode causes a current across the device, leading to forward bias which in turn induces "dark current" in the opposite direction to the photocurrent. This is called the photovoltaic effect, and is the basis for solar cells — in fact, a solar cell is just a large number of big photodiodes.

d. Reverse bias induces only little current (known as saturation or back current) along its direction. But a more important effect of reverse bias is widening of the depletion layer (therefore expanding the reaction volume) and strengthening the photocurrent. Circuits based on this effect are more sensitive to light than ones based on the photovoltaic effect and also tend to have lower capacitance, which improves the speed of their time response. On the other hand, the photovoltaic mode tends to exhibit less electronic noise.

e. Avalanche photodiodes have a similar structure, but they are operated with much higher reverse bias. This allows each photo-generated carrier to be multiplied by avalanche breakdown, resulting in internal gain within the photodiode, which increases the effective responsivity of the device.

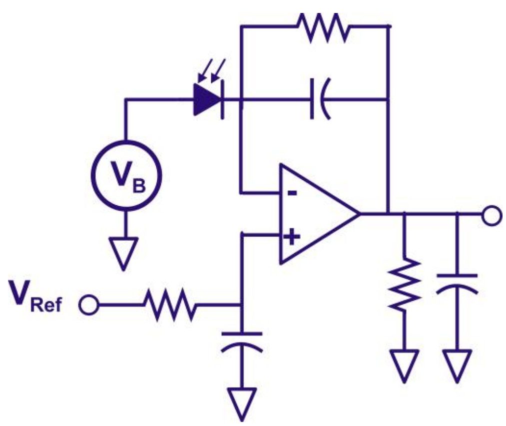


Figure 44, Semiconductors-044 Typical Photodiode Circuit

18. Phototransistors

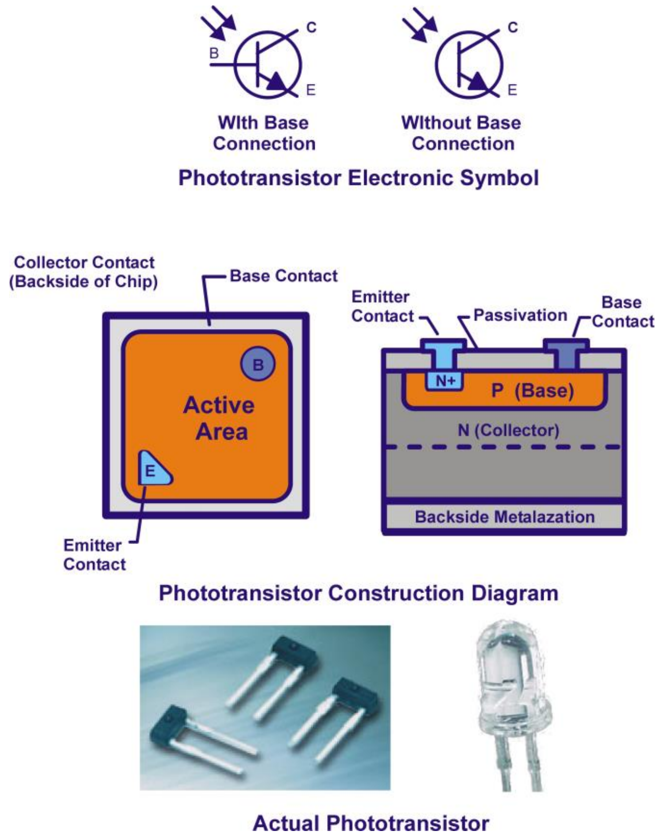


Figure 45, Semiconductors-045 Phototransistor Electronic Symbol

a. The structure of a phototransistor is very similar to that of a photodiode. In fact, while not optimized for this mode of operation, the collector-base junction of a phototransistor can be used as a photodiode with fairly good results. The major structural difference is that the phototransistor has two junctions compared with one for the photodiode.

b. Like diodes, all transistors are light-sensitive. Phototransistors are designed specifically to take advantage of this fact. The most-common variant is an NPN bipolar transistor with an exposed base region. Here, light striking the base replaces what would ordinarily be voltage applied to the base -- so, a phototransistor amplifies variations in the light striking it. Note that phototransistors may or may not have a base lead (if they do, the base lead allows you to bias the phototransistor's light response.

c. Phototransistors are transistors designed to capture light and are assembled in a transparent package. They are often more convenient than photodiodes because they have built in gain: the absorbed light creates a current in the base region of the phototransistor, resulting in current gains from 100 to several thousands. Photodarlingtons have two stages of gain, with net gains that can be greater than 100,000.

d. The built in gain allows the phototransistor to be coupled with a load resistor to accommodate TTL level voltages for a wide range of light levels. Because of their ease of use, low cost and TTL compatible signal levels, phototransistors have become popular for applications where there is more than a few hundred nanowatts of available optical power. These devices however, do have some drawbacks compared to photodiodes. The frequency bandwidth and linearity are relatively limited and spectral response is restricted to between 350 and 1100 nm. In addition, there are very large variations in sensitivity between individual devices and few standard package options.

e. To demonstrate the relative sensitivity of these different types of detectors, compare the output currents that could be expected from a .025" x .025" detector chip exposed to .05 mW/cm2 of illumination.

DETECTOR	GAIN	OUTPUT CURRENT
Photodiode	1x	50 µA
Phototransistor	500x	100 nA

f. The current-voltage characteristics of the phototransistor are similar to NPN signal transistors, with the major exception that incident light replaces current base drive.

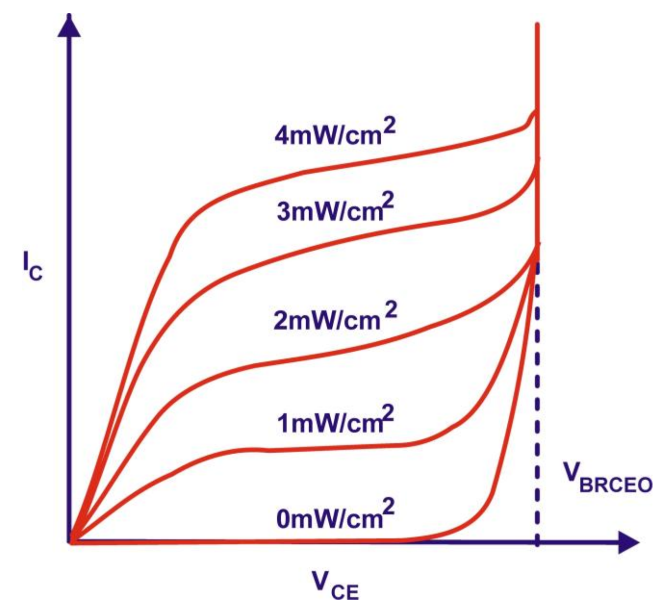


Figure 46, Semiconductors-046 Typical Phototransistor Curves

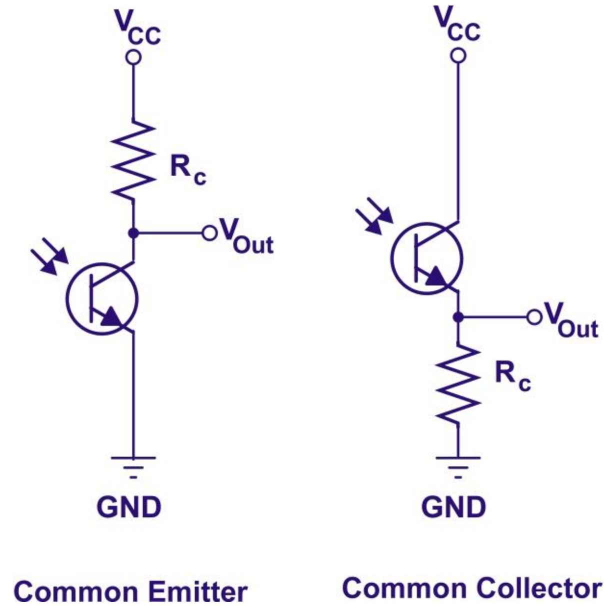


Figure 47, Semiconductors-047 Typical Phototransistor Circuits

19. Light Emitting Diodes

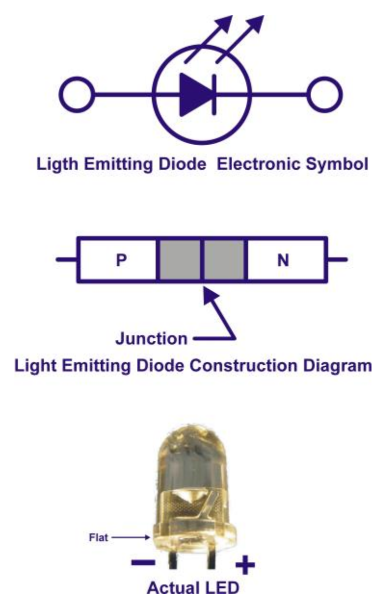


Figure 48, Semiconductors-048 Light Emitting Diode Electronic Symbol

a. LED's are special diodes that emit light when connected in a circuit. They are frequently used as "pilot" lights in electronic appliances to indicate whether the circuit is closed or not. A a clear (or often colored) epoxy case enclosed the heart of an LED, the semi-conductor chip.

b. The two wires extending below the LED epoxy enclosure, or the "bulb" indicate how the LED should be connected into a circuit. The negative side of an LED lead is indicated in two ways: 1) by the flat side of the bulb, and 2) by the shorter of the two wires extending from the LED. The negative lead should be connected to the negative terminal of a battery. LED's operate at relative low voltages between about 1 and 4 volts, and draw currents between about 10 and 40 milliamperes. Voltages and currents substantially above these values can melt a LED chip.

c. The most important part of a light emitting diode (LED) is the semi-conductor chip located in the center of the bulb as shown at the right. The chip has two regions separated by a junction. The p region is dominated by positive electric charges, and the n region is dominated by negative electric charges. The junction acts as a barrier to the flow of electrons between the p and the n regions. Only when sufficient voltage is applied to the semi-conductor chip, can the current flow, and the electrons cross the junction into the p region.

d. In the absence of a large enough electric potential difference (voltage) across the LED leads, the junction presents an electric potential barrier to the flow of electrons.

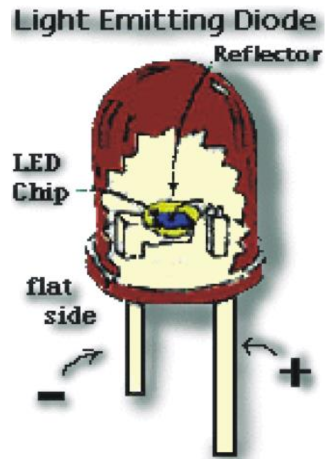


Figure 49, Semiconductors-049 Light Emitting Diode Cutaway View

e. When sufficient voltage is applied to the chip across the leads of the LED, electrons can move easily in only one direction across the junction between the p and n regions. In the p region there are many more positive than negative charges. In the n region the electrons are more numerous than the positive electric charges. When a voltage is applied and the current starts to flow, electrons in the n region have sufficient energy to move across the junction into the p region. Once in the p region the electrons are immediately attracted to the positive charges due to the mutual Coulomb forces of attraction between opposite electric charges. When an electron moves sufficiently close to a positive charge in the p region, the two charges "re-combine".

f. Each time an electron recombines with a positive charge, electric potential energy is converted into electromagnetic energy. For each recombination of a negative and a positive charge, a quantum of electromagnetic energy is emitted in the form of a photon of light with a frequency characteristic of the semi-conductor material (usually a combination of the chemical elements gallium, arsenic and phosphorus). Only photons in a very narrow frequency range can be emitted by any material. LED's that emit different colors are made of different semi-conductor materials, and require different energies to light them.

g. The electric energy is proportional to the voltage needed to cause electrons to flow across the p-n junction. The different colored LED's emit predominantly light of a single color. The energy (E) of the light emitted by an LED is related to the electric charge (q) of an electron and the voltage (V) required to light the LED by the expression: E = qV Joules. This expression simply says that the voltage is proportional to the electric energy, and is a general statement which applies to any circuit, as well as to LED's. The constant q is the electric charge of a single electron, -1.6 x 10-19 Coulomb.

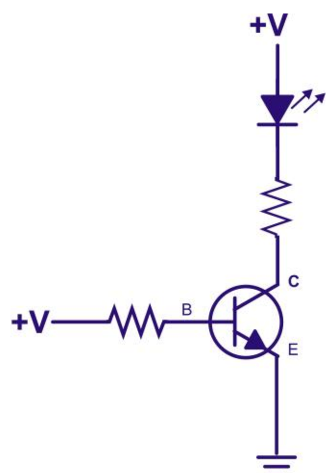


Figure 50, Semiconductors-050 Simple LED Circuit

20. Laser Diodes

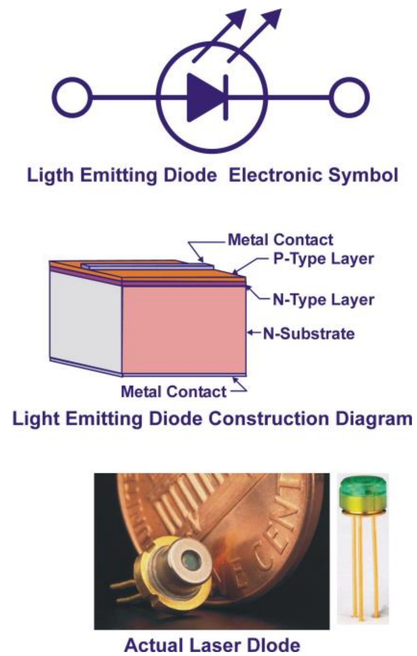


Figure 51, Semiconductors-051 Laser Diode Electronic Symbol

a. A laser diode is a laser where the active medium is a semiconductor similar to that found in a light-emitting diode. The most common and practical type of laser diode is formed from a p-n junction and powered by injected electrical current. These devices are sometimes referred to as injection laser diodes to distinguish them from optically pumped laser diodes, which are more easily produced in the laboratory.

b. A laser diode, like many other semiconductor devices, is formed by doping a very thin layer on the surface of a crystal wafer. The crystal is doped to produce an n-type region and a p-type region, one above the other, resulting in a p-n junction, or diode.

c. As in other diodes, when this structure is forward biased, holes from the p-region are injected into the n-region, where electrons are the majority carrier. Similarly, electrons from the n-region are injected into the p-region, where holes are the majority carrier. When an electron and a hole are present in the same region, they may recombine by spontaneous emission—that is, the electron may re-occupy the energy state of the hole, emitting a photon with energy equal to the difference between the electron and hole states involved. These injected electrons and holes represent the injection current of the diode, and spontaneous emission gives the laser diode below lasing threshold similar properties to an LED. Spontaneous emission is necessary to initiate laser oscillation, but it is a source of inefficiency once the laser is oscillating.

d. Under suitable conditions, the electron and the hole may coexist in the same area for quite some time (on the order of microseconds) before they recombine. Then a nearby photon with energy equal to the recombination energy can cause recombination by stimulated emission. This generates another photon of the same frequency, travelling in the same direction, with the same polarization and phase as the first photon. This means that stimulated emission causes gain in an optical wave (of the correct wavelength) in the injection region, and the gain increases as the number of electrons and holes injected across the junction increases. The spontaneous and stimulated emission processes are vastly more efficient in direct bandgap semiconductors than in indirect bandgap semiconductors, thus silicon is not a common material for laser diodes.

e. As in other lasers, the gain region is surrounded with an optical cavity to form a laser. In the simplest form of laser diode, an optical waveguide is made on that crystal surface, such that the light is confined to a relatively narrow line. The two ends of the crystal are cleaved to form perfectly smooth, parallel edges, forming a Fabry-Perot resonator. Photons emitted into a mode of the waveguide will travel along the waveguide and be reflected several times from each end face before they are emitted. As a light wave passes through the cavity, it is amplified by stimulated emission, but light is also lost due to absorption and by incomplete reflection from the end facets. Finally, if there is more amplification than loss, the diode begins to "lase".

f. Some important properties of laser diodes are determined by the geometry of the optical cavity. Generally, in the vertical direction, the light is contained in a very thin layer, and the structure supports only a single optical mode in the direction perpendicular to the layers. In the lateral direction, if the waveguide is wide compared to the wavelength of light, then the waveguide can support multiple lateral optical modes, and the laser is known as "multi-mode". These laterally multi-mode lasers are adequate in cases where one needs a very large amount of power, but not a small diffraction-limited beam; for example in printing, activating chemicals, or pumping other types of lasers.

g. In applications where a small focused beam is needed, the waveguide must be made narrow, on the order of the optical wavelength. This way, only a single lateral mode is supported and one ends up with a diffraction limited beam. Such single spatial mode devices are used for optical storage, laser pointers, and fiber optics. Note that these lasers may still support multiple longitudinal modes, and thus can lase at multiple wavelengths simultaneously.

h. The wavelength emitted is a function of the band-gap of the semiconductor and the modes of the optical cavity. In general, the maximum gain will occur for photons with energy slightly above the band-gap energy, and the modes nearest the gain peak will lase most strongly. If the diode is driven strongly enough, additional side modes may also lase. Some laser diodes, such as most visible lasers, operate at a single wavelength, but that wavelength is unstable and changes due to fluctuations in current or temperature.

i. Due to diffraction, the beam diverges (expands) rapidly after leaving the chip, typically at 30 degrees vertically by 10 degrees laterally. A lens must be used in order to form a collimated beam like that produced by a laser pointer. If a circular beam is required, cylindrical lenses and other optics are used. For single spatial mode lasers, using symmetrical lenses, the collimated beam ends up being elliptical in shape, due to the difference in the vertical and lateral divergences. This is easily observable with a red laser pointer.

j. The simple diode described above has been heavily modified in recent years to accommodate modern technology, resulting in a variety of types of laser diodes, as described below.

k. The simple laser diode structure, described above, is extremely inefficient. Such devices require so much power that they can only achieve pulsed operation without damage. Although historically important and easy to explain, such devices are not practical.

l. Double Heterostructure Lasers

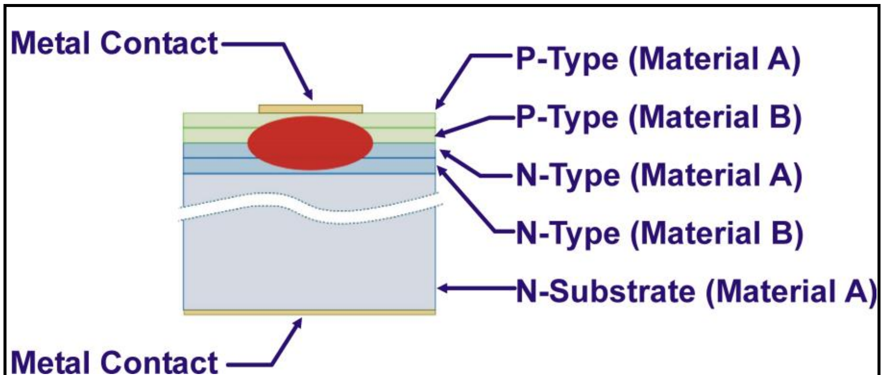


Figure 52, Semiconductors-052 Double Heterostructure Laser Diode

1) The first to demonstrate coherent light emission from a semiconductor diode (the first laser diode), included Robert N. Hall and his team at the General Electric research center in November 1962[1]

2) Other teams at IBM, MIT Lincoln Laboratory, Texas Instruments, and RCA Laboratories were also involved in and receive credit for historic

initial demonstrations of efficient light emission and lasing in semiconductor diodes in 1962 and thereafter.

3) The first laser diode to achieve continuous wave operation was a double heterostructure demonstrated essentially simultaneously by Zhores Alferov and collaborators (including Dmitri Z. Garbuzov) of the Soviet Union, and Morton Panish and Izuo Hayashi working in the United States.

4) In these devices, a layer of low bandgap material is sandwiched between two high bandgap layers. One commonly-used pair of materials is gallium arsenide (GaAs) with aluminium gallium arsenide (AlxGa(1-x)As). Each of the junctions between different bandgap materials is called a heterostructure, hence the name "double heterostructure laser" or DH laser. The kind of laser diode described in the first part of the article may be referred to as a homojunction laser, for contrast with these more popular devices.

5) The advantage of a DH laser is that the region where free electrons and holes exist simultaneously—the "active" region—is confined to the thin middle layer. This means that many more of the electron-hole pairs can contribute to amplification—not so many are left out in the poorly amplifying periphery. In addition, light is reflected from the heterojunction; hence, the light is confined to the region where the amplification takes place.

m. Quantum Well Lasers

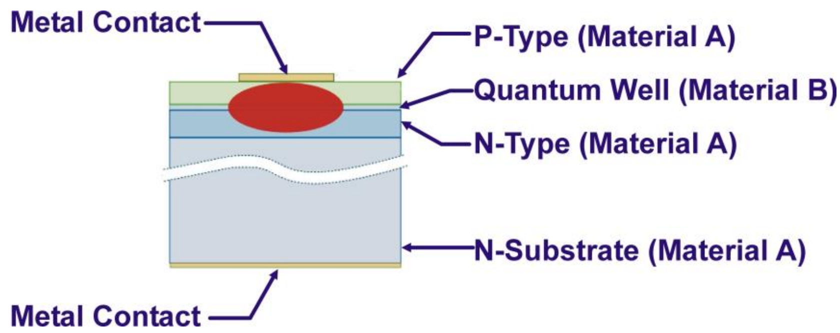


Figure 53, Semiconductors-053 Quantum Well Lasers

1) If the middle layer is made thin enough, it acts as a quantum well. This means that the vertical variation of the electron's wavefunction, and thus a component of its energy, is quantised. The efficiency of a quantum well laser is greater than that of a bulk laser because the density of states function of electrons in the quantum well system has an abrupt edge that concentrates electrons in energy states that contribute to laser action.
2) Lasers containing more than one quantum well layer are known as multiple quantum well lasers. Multiple quantum wells improve the overlap of the gain region with the optical waveguide mode.
3) Further improvements in the laser efficiency have also been demonstrated by reducing the quantum well layer to a quantum wire or to a "sea" of quantum dots.
4) In a quantum cascade laser, the difference between quantum well energy levels is used for the laser transition instead of the bandgap. This enables laser action at relatively long wavelengths, which can be tuned simply by altering the thickness of the layer. As of 2005, quantum cascade lasers have not yet been widely commercialized.

n. Separate Confinement Heterostructure Lasers j

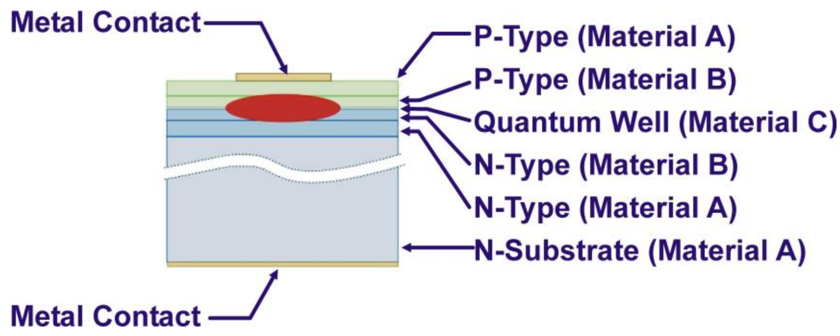


Figure 54, Semiconductors-054 Separate Confinement Heterostructure Lasers

1) The problem with the simple quantum well diode described above is that the thin layer is simply too small to effectively confine the light. To compensate, another two layers are added on, outside the first three. These layers have a lower refractive index than the centre layers, and hence confine the light effectively. Such a design is called a separate confinement heterostructure (SCH) laser diode.

2) Almost all commercial laser diodes since the 1990s have been SCH quantum well diodes.

o. Distributed Feedback Lasers

1) Distributed feedback lasers (DFB) are the most common transmitter type in DWDM-systems. To stabilize the lasing wavelength, a diffraction grating is etched close to the p-n junction of the diode. This grating acts like an optical filter, causing a single wavelength to be fed back to the gain region and lase. Since the grating provides the feedback that is required for lasing, reflection from the facets is not required. Thus, at least one facet of a DFB is anti-reflection coated. The DFB laser has a stable wavelength that is set during manufacturing by the pitch of the grating, and can only be tuned slightly with temperature. Such lasers are the workhorse of demanding optical communication.

p. VCELs

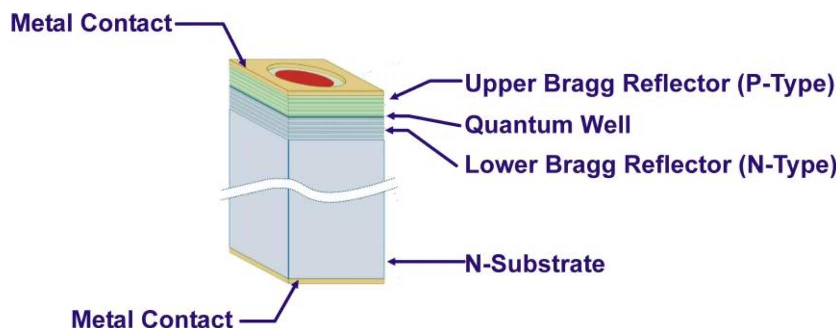


Figure 55, Semiconductors-055 VCELs

1) Vertical cavity surface emitting lasers (VCSELs) have the optical cavity axis along the direction of current flow rather than perpendicular to the current flow as in conventional laser diodes. The active region length is very short compared with the lateral dimensions so that the radiation emerges from the ‘‘surface’’ of the cavity rather than from its edge as shown in Fig. 2. The reflectors at the ends of the cavity are dielectric mirrors made from alternating high and low refractive index quarter-wave thick multilayer.

2) There are several advantages to producing VCSELs when compared with the production process of edge-emitting lasers. Edge-emitters cannot be tested until the end of the production process. If the edge-emitter does not work, whether due to bad contacts or poor material growth quality, the production time and the processing materials have been wasted. Additionally, because VCSELs emit the beam perpendicular to the active region of the laser as opposed to parallel as with an edge emitter, tens of thousands of VCSELs can be processed simultaneously on a three inch Gallium Arsenide wafer. Furthermore, even though the VCSEL production process is more labor and material intensive, the yield can be controlled to a more predictable outcome.

3) Such dielectric mirrors provide a high degree of wavelength-selective reflectance at the required free surface wavelength λ if the thicknesses of alternating layers $d1$ and $d2$ with refractive indices $n1$ and $n2$ are such that $n1d1 + n2d2 = \frac{1}{2}\lambda$ which then leads to the constructive interference of all partially reflected waves at the interfaces. But there is a disadvantage because of the high mirror reflectivities, VCSELs have lower output powers when compared to edge emitting lasers.

q. VECSELs

1) Vertical external-cavity surface-emitting lasers, or VECSELs, are similar to VCSELs. In VCSELs, the mirrors are typically grown epitaxially as part of the diode structure, or grown separately and bonded directly to the semiconductor containing the active region. VECSELs are distinguished by a construction in which one of the two mirrors is external to the diode structure. As a result, the cavity includes a free-space region. A typical distance from the diode to the external mirror would be 1 cm.

2) One of the most interesting features of any VECSEL is the thin-ness of the semiconductor gain region in the direction of propagation, less than 100 nm. In contrast, a conventional in-plane semiconductor laser entails light propagation over distances of from 250 µm upward to 2 mm or longer. The significance of the short propagation distance is that it causes the effect of "antiguiding" nonlinearities in the diode laser gain region to be minimized. The result is a large-cross-section single-mode optical beam which is not attainable from in-plane ("edge-emitting") diode lasers.

r. Laser Diode Applications

1) Laser diodes are numerically the most common type of laser, with 2004 sales of approximately 733 million diode lasers (Steele 2005), as compared to 131,000 of other types of lasers (Kincade and Anderson 2005).

2) Laser diodes find wide use in telecommunication as easily modulated and easily coupled light sources for fiber optics communication. They are used in various measuring instruments, eg. rangefinders. Another common use is in barcode readers. Visible lasers, typically red but recently also green, are common as laser pointers. Both low and high-power diodes are used extensively in the printing industry both as light sources for scanning (input) of images and for very high-speed and high-resolution printing plate (output) manufacturing. Infrared and red laser diodes are common in CD players, CD-ROMs and DVD technology. Blue-violet lasers will find their use in upcoming HD-DVD and Blu-Ray technology. High-power laser diodes are used in industrial applications such as heat treating, cladding, seam welding and for pumping other lasers, such as diode pumped solid state lasers. The use of diode lasers for high-speed, low-cost, combustion spectroscopy is being explored.

3) In general, applications of laser diodes can be categorized in various ways. Most applications of diode lasers can be served by larger solid state lasers or optical parametric oscillators but it is the ability to mass-produce diode lasers at low cost that makes them essential for mass-market applications. Diode lasers have application to virtually every field of endeavor that attracts wide attention today. Since light has many different properties (power, wavelength & spectral quality, beam quality, polarization, etc.) it is interesting to classify applications by these basic properties.

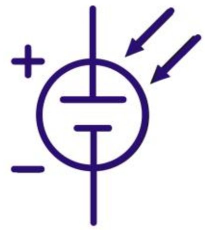
4) Many applications of diode lasers primarily make use of the "directed energy" property of an optical beam. In this category one might include the laser printers, bar-code readers, image scanning, illuminators, designators, optical data recording, combustion ignition, laser surgery, industrial sorting, industrial machining, and directed energy weaponry. Some of these applications are emerging whereas many are familiar to the wider society.

5) Applications which may today or in the future make use of the "coherent" properties of diode-laser-generated light include interferometric distance measurement, holography, coherent communications, and coherent control of chemical reactions.

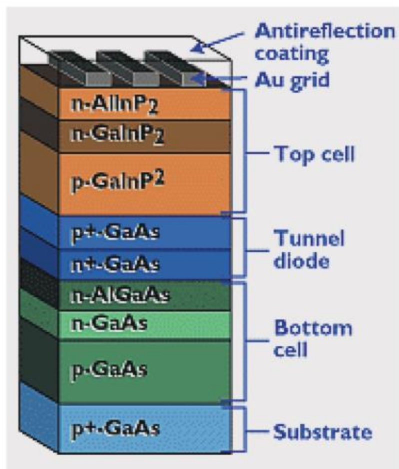
6) Applications which may make use of "narrow spectral" properties of diode lasers include telecommunications, infra-red countermeasures, spectroscopic sensing, generation of radio-frequency or terahertz waves, atomic clock state preparation, quantum key cryptography, frequency doubling and conversion, water purification (in the UV), and photodynamic therapy (where a particular wavelength of light would cause a substance such as porphyrin to become chemically active as an anti-cancer agent only where the tissue is illuminated by light).

7) Applications where the ability to "generate ultra-short pulses of light" by the technique known as "mode-locking" include clock distribution for high-performance integrated circuits, high-peak-power sources for laser-induced breakdown spectroscopy sensing, arbitrary waveform generation for radio-frequency waves, photonic sampling for analog-to-digital conversion, and optical code-division-multiple-access systems for secure communication.

21. Photovoltaic Cell



Photovoltaic Cell Electronic Symbol



Photovoltaic Cell Construction Diagram

ActualPhotovoltaic Cell

Figure 56, Semiconductors-056 Photovoltaic Cell Electronic Symbol

a. Photovoltaics is the direct conversion of light into electricity at the atomic level. Some materials exhibit a property known as the photoelectric effect that causes them to absorb photons of light and release electrons. When these free electrons are captured, an electric current results that can be used as electricity.

b. The photoelectric effect was first noted by a French physicist, Edmund Bequerel, in 1839, who found that certain materials would produce small amounts of electric current when exposed to light. In 1905, Albert Einstein described the nature of light and the photoelectric effect on which photovoltaic technology is based, for which he later won a Nobel prize in physics. The first photovoltaic module was built by Bell Laboratories in 1954. It was billed as a solar battery and was mostly just a curiosity as it was too expensive to gain widespread use. In the 1960s, the space industry began to make the first serious use of the technology to provide power aboard spacecraft. Through the space programs, the technology advanced, its reliability was established, and the cost began to decline. During the energy crisis in the 1970s, photovoltaic technology gained recognition as a source of power for non-space applications.

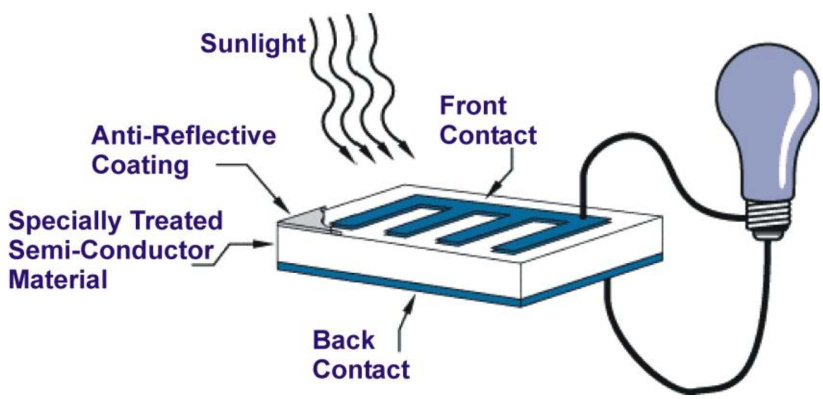


Figure 57, Semiconductors-057 Photovoltaic Cell Operation

c. The diagram above illustrates the operation of a basic photovoltaic cell, also called a solar cell. Solar cells are made of the same kinds of semiconductor materials, such as silicon, used in the microelectronics industry. For solar cells, a thin semiconductor wafer is specially treated to form an electric field, positive on one side and negative on the other. When light energy strikes the solar cell, electrons are knocked loose from the atoms in the semiconductor material. If electrical conductors are attached to the positive and negative sides, forming an electrical circuit, the electrons can be captured in the form of an electric current -- that is, electricity. This electricity can then be used to power a load, such as a light or a tool.

d. A number of solar cells electrically connected to each other and mounted in a support structure or frame is called a photovoltaic module. Modules are designed to supply electricity at a certain voltage, such as a common 12 volts system. The current produced is directly dependent on how much light strikes the module.

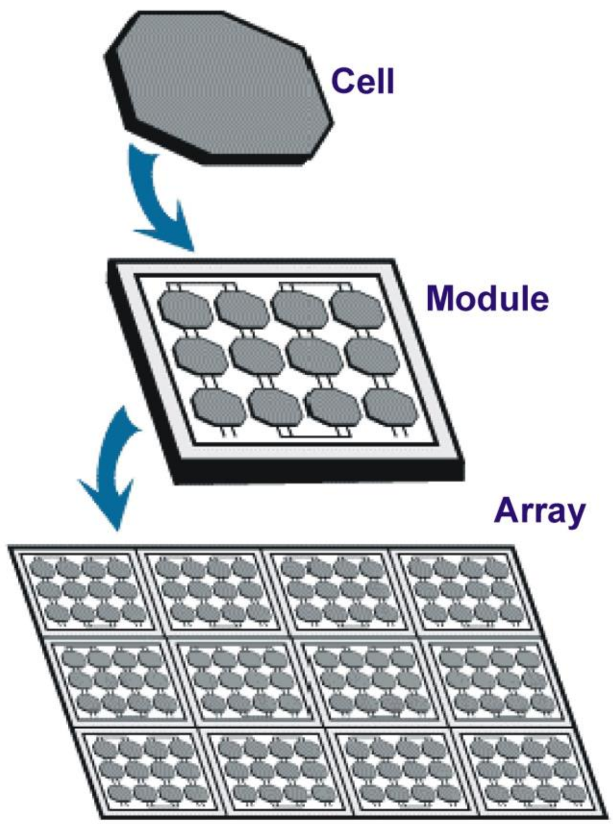


Figure 58, Semiconductors-058 Photovoltaic Cell Configurations

e. Multiple modules can be wired together to form an array. In general, the larger the area of a module or array, the more electricity that will be produced. Photovoltaic modules and arrays produce direct-current (dc) electricity. They can be connected in both series and parallel electrical arrangements to produce any required voltage and current combination.

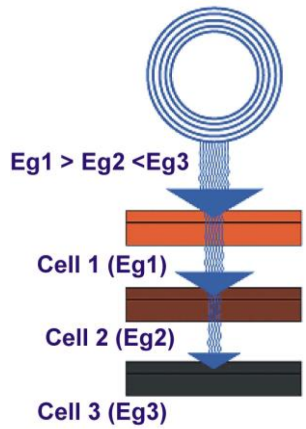


Figure 59, Semiconductors-059 Photovoltaic Cell Layers

f. Today's most common PV devices use a single junction, or interface, to create an electric field within a semiconductor such as a PV cell. In a single-junction PV cell, only photons whose energy is equal to or greater than the band gap of the cell material can free an electron for an electric circuit. In other words, the photovoltaic response of single-junction cells is limited to the portion of the sun's spectrum whose energy is above the band gap of the absorbing material, and lower-energy photons are not used.

g. One way to get around this limitation is to use two (or more) different cells, with more than one band gap and more than one junction, to generate a voltage. These are referred to as "multijunction" cells (also called "cascade" or "tandem" cells). Multijunction devices can achieve a higher total conversion efficiency because they can convert more of the energy spectrum of light to electricity.

h. As shown below, a multijunction device is a stack of individual single-junction cells in descending order of band gap (Eg). The top cell captures the high-energy photons and passes the rest of the photons on to be absorbed by lower-band-gap cells.

i. Much of today's research in multijunction cells focuses on gallium arsenide as one (or all) of the component cells. Such cells have reached efficiencies of around 35% under concentrated sunlight. Other materials studied for multijunction devices have been amorphous silicon and copper indium diselenide.

j. As an example, the multijunction device below uses a top cell of gallium indium phosphide, "a tunnel junction," to aid the flow of electrons between the cells, and a bottom cell of gallium arsenide.

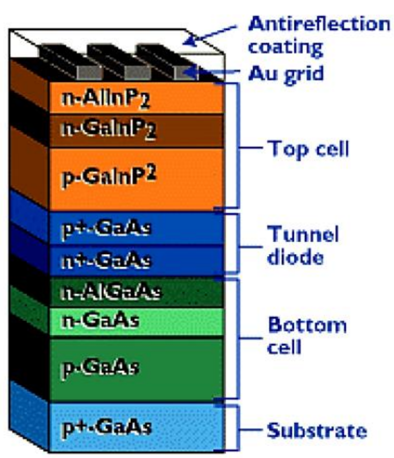


Figure 60, Semiconductors-060 Photovoltaic Cell Multi-Layer Construction