and nano-technologies continue to this day. Trends are toward lighter, smaller devices with remarkable capability and reliability. Integration of the wide range of communication and navigational aids is a focus.

Fundamentals of Electronics

Analog Versus Digital Electronics

Electronic devices represent and manipulate real-world phenomenon through the use of electrical signals. Electronic circuits are designed to perform a wide array of manipulations. Analog representations are continuous. Some aspect of an electric signal is modified proportionally to the real-world item that is being represented. For example, a microphone has electricity flowing through it that is altered when sound is applied. The type and strength of the modification to the electric signal is characteristic of the sound that is made into the microphone. The result is that sound, a real-world phenomenon, is represented electronically. It can then be moved, amplified, and reconverted from an electrical signal back into sound and broadcast from a speaker across the room or across the globe.

Since the flow of electricity through the microphone is continuous, the sound continuously modifies the electric signal. On an oscilloscope, an analog signal is a continuous curve. *[Figure 11-3]* An analog electric signal can be modified by changing the signal's amplitude, frequency, or phase.

A digital electronic representation of a real-world event is discontinuous. The essential characteristics of the continuous event are captured as a series of discrete incremental values. Electronically, these representative samplings are successive chains of voltage and non-voltage signals. They can be transported and manipulated in electronic circuits. When the samples are sufficiently small and occur with high frequency, real-world phenomenon can be represented to appear continuous.

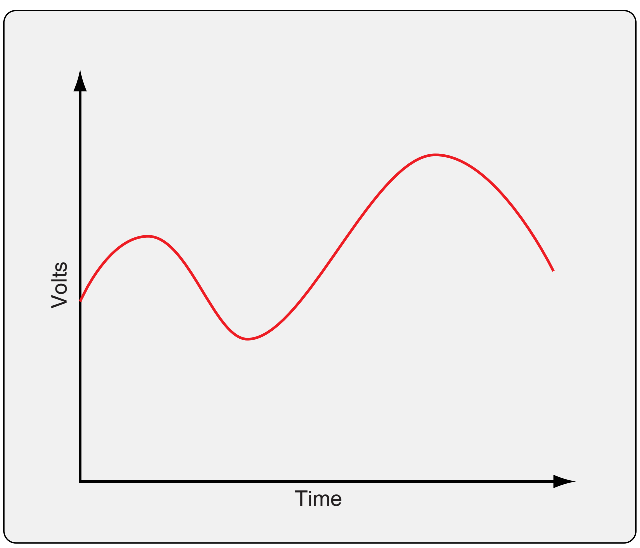


Figure 11-3. *An analog signal displayed on an oscilloscope is a continuous curve.*

Noise

A significant advantage of digital electronics over analog electronics is the control of noise. Noise is any alteration of the represented real-world phenomenon that is not intended or desired. Consider the operation of a microphone when understanding noise. A continuous analog voltage is modified by a voice signal that results in the continuous voltage varying in proportion to the volume and tone of the input sound. However, the voltage responds and modifies to any input. Thus, background sounds also modify the continuous voltage as will electrostatic activity and circuitry imperfections. This alteration by phenomenon that are not the intended modifier is noise.

During the processing of digitized data, there is little or no signal degradation. The real-world phenomenon is represented in a string of binary code. A series of ones and zeros are electronically created as a sequence of voltage or no voltage and carried through processing stages. It is relatively immune to outside alteration once established. If a signal is close to the set value of the voltage, it is considered to be that voltage. If the signal is close to zero, it is considered to be no voltage. Small variations or modifications from undesired phenomenon are ignored. *Figure 11-4* illustrates an analog sine wave and a digital sine wave. Any unwanted voltage will modify the analog curve. The digital steps are not modified by small foreign inputs. There is either voltage or no voltage.

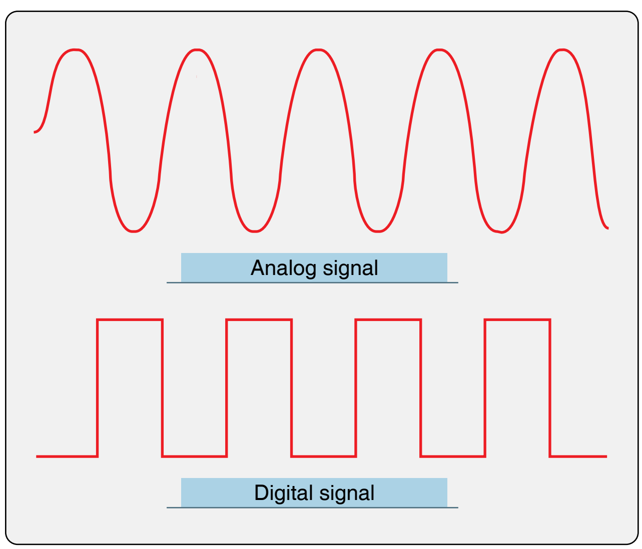


Figure 11-4. *Analog signals are continuous voltage modified by all external events including those that are not desired called noise. Digital signals are a series of voltage or no voltage that represent a desired event.*

Analog Electronics

Early aircraft were equipped with radio communication and navigational devices that were constructed with analog electronic circuits. They used vacuum tubes that functioned as electron control valves. These were later replaced by solid-state devices. Today, digital electronic circuits dominate modern avionics. A brief look at various electron control valves used on aircraft follows.

Electron Control Valves

Electron control valves are an essential part of an electronic circuit. Control of electron flow enables the circuit to produce the desired outcome. Early aircraft made use of vacuum tubes to control electron flow. Later, transistors replaced vacuum tubes. Semiconductors used in transistors and integrated circuits have enabled the solid-state digital electronics found in aircraft today.

Vacuum Tubes

Electron control valves found in the analog circuits of early aircraft electronics are constructed of vacuum tubes. Only antique aircraft retain radios with these devices due to their size and inability to withstand the harsh vibration and shock of the aircraft operating environment. However, they do function, and a description is included here as a foundation for the study of more modern electronic circuits and components.

Diodes

A diode acts as a check valve in an alternating current (AC) circuit. It allows current to flow during half of the AC cycle but not the other half. In this manner, it creates a pulsating direct current (DC) with current that drops to zero in between pulses. A diode tube has two active electrodes: the cathode and the plate. It also contains a heater. All of this is housed in a vacuum environment inside the tube. *[Figure 11-5]* The heater glows red hot while heating the cathode. The cathode is coated with a material whose electrons are excited by the heat. The excited electrons expand their orbit when heated. They move close enough to the plate, which is constructed around the cathode and heater arrangement, that they are attracted to the positively-charged plate. When the AC current cycles, the plate becomes negatively charged and the excited cathode electrons do not flow to the plate. In a circuit, this causes a check valve effect that allows current to only flow in one direction, which is the definition of a diode. *[Figure 11-6]* The various symbols used to depict diodes are shown in *Figure 11-7*.

Triodes

A triode is an electron control valve containing three elements. It is often used to control a large amount of current with a smaller current flow. In addition to the cathode, plate, and

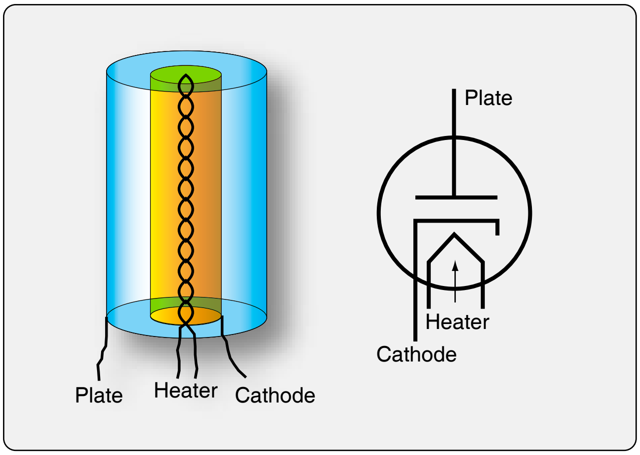


Figure 11-5. *A vacuum tube diode contains a cathode, heater, and plate. Note that the arrow formed in the symbol for the heater points to the direction of electron flow.*

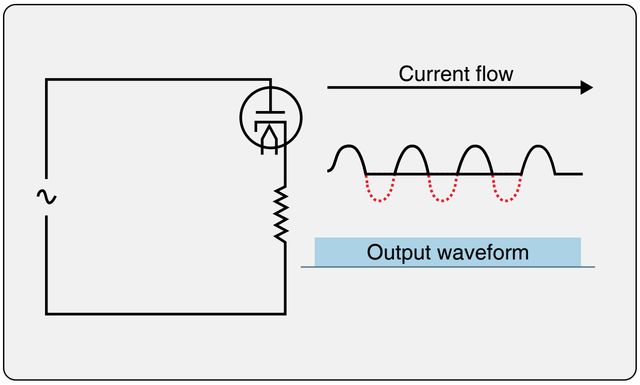


Figure 11-6. *A vacuum tube diode in a circuit allows current to flow in one direction only. The output waveform illustrates the lack of current flow as the AC cycles.*

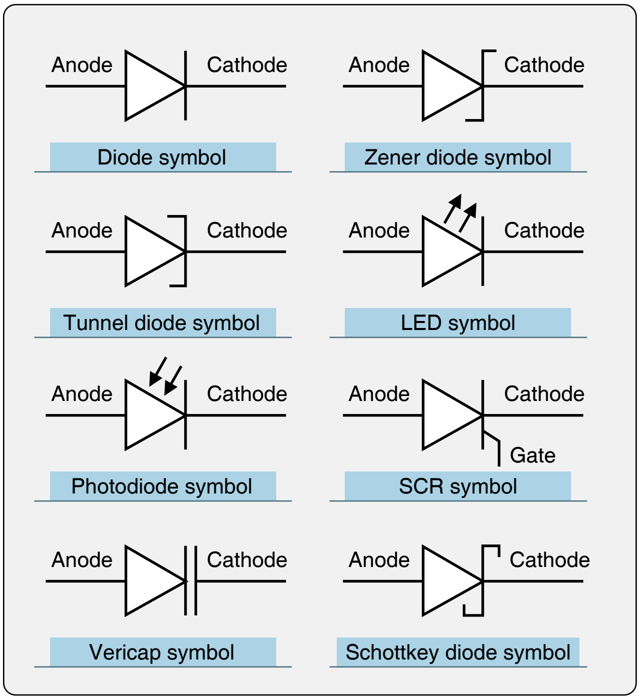


Figure 11-7. *Diode symbols.*

heater present in the diode, a triode also contains a grid. The grid is composed of fine wire spiraled between the cathode and the plate but closer to the cathode. Applying voltage to the grid can influence the cathode's electrons, which normally flow to the plate when the cathode is heated. Changes in the relatively small amount of current that flows through the grid can greatly impact the flow of electrons from the cathode to the plate. *[Figure 11-8]*

Figure 11-9 illustrates a triode in a simple circuit. AC voltage input is applied to the grid. A high-resistance resistor is used so that only minimum voltage passes through to the grid. As this small AC input voltage varies, the amount of DC output in the cathode-plate circuit also varies. When the input signal is positive, the grid is positive. This aids in drawing electrons from the cathode to the plate. However, when the AC input signal cycles to negative, the grid becomes negatively charged and flow from the cathode to the plate is cut off with the help of the negatively charged grid that repels the electrons on the cathode.

Tetrodes

A tetrode vacuum tube electron control valve has four elements. In addition to the cathode, plate, and grid found in a triode, a tetrode also contains a screen grid. The cathode and plate of a vacuum tube electron control valve can act as a capacitor. At high frequencies, the capacitance is so low that feedback occurs. The output in the plate circuit feeds back into the control grid circuit. This causes an oscillation generating AC voltage that is unwanted. By placing a screen grid between the anode and the control grid windings, this feedback and the inter-electrode capacitive effect of the anode and cathode are neutralized. *[Figure 11-10]*

Figure 11-11 illustrates a tetrode in a circuit. The screen grid is powered by positive DC voltage. The inter-electrode capacitance is now between the screen grid and the plate. A capacitor is located between the screen grid and ground. AC feedback generated in the screen grid goes to ground and does not oscillate. This allows use of the tetrode at higher frequencies than a triode.

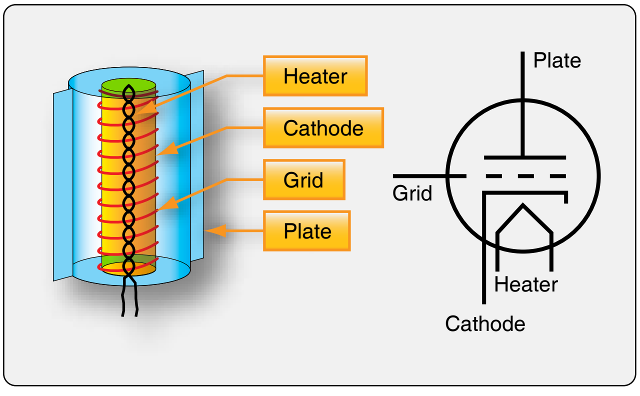


Figure 11-8. *A triode has three elements: the cathode, plate, and a grid.*

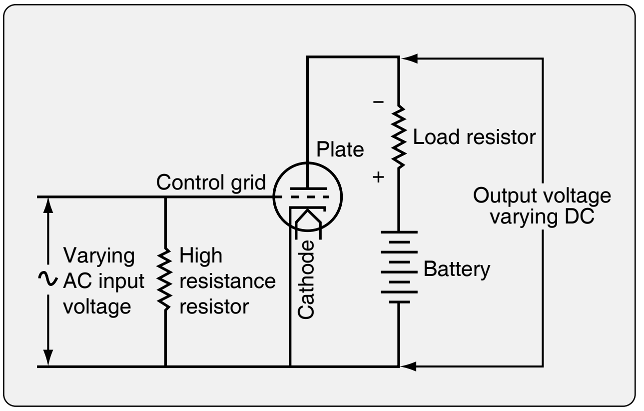


Figure 11-9. *Varying AC input voltage to the grid circuit in a triode produces a varying DC output.*

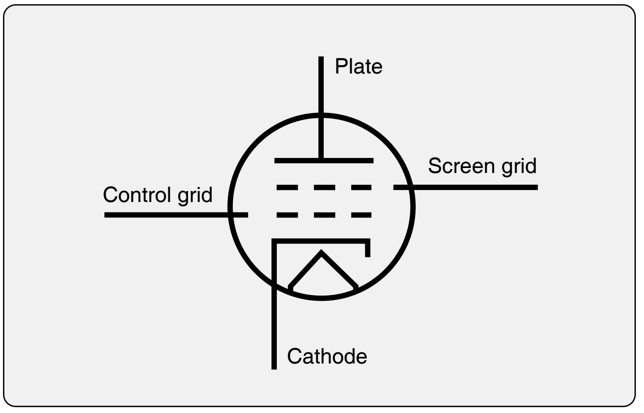


Figure 11-10. *A tetrode is a four element electron control valve vacuum tube including a cathode, a plate, a control grid, and a screen grid.*

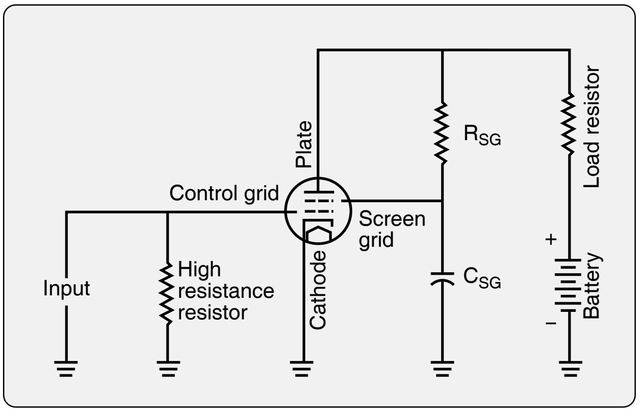


Figure 11-11. *To enable a triode to be used at high frequencies, a screen grid is constructed between the plate and the control grid.*

Pentodes

The plate in a vacuum tube can have a secondary emission that must be controlled. When electrons flow from the cathode through the control grid and screen grid to the plate, they can arrive at such high velocity that some bounce off. Therefore, the tendency is for those electrons to be attracted to the positively charged screen grid. The screen grid is not capable of handling large amounts of current without burning up. To solve this problem, a third grid is constructed between the plate and the screen grid. Called a suppression grid, it is charged negatively so that secondary electron flow from the plate is repelled by the negative charge back toward the plate and is not allowed to reach the screen grid. The five-element pentode is especially useful in high-power circuits where secondary emissions from the plate are high. *[Figure 11-12]*

Solid-State Devices

Solid-state devices began replacing vacuum tube electron control valves in the late 1950s. Their long life, reliability, and resilience in harsh environments make them ideal for use in avionics.

Semiconductors

The key to solid-state electronic devices is the electrical behavior of semiconductors. To understand semiconductors, a review of what makes a material an insulator or a conductor follows. Then, an explanation for how materials of limited conductivity are constructed and some of their many uses is explained. Semiconductor devices are the building blocks of modern electronics and avionics.

An atom of any material has a characteristic number of electrons orbiting the nucleus of the atom. The arrangement of the electrons occurs in somewhat orderly orbits called rings or shells. The closest shell to the nucleus can only contain two electrons. If the atom has more than two electrons, they are found in the next orbital shell away from the nucleus. This second shell can only hold eight electrons. If the atom has more than eight electrons, they orbit in a third shell farther out from the nucleus. This third shell is filled with eight electrons and then a fourth shell starts to fill if the element still has more electrons. However, when the fourth shell contains eight electrons, the number of electrons in the third shell begins to increase again until a maximum of 18 is reached. *[Figure 11-13]*

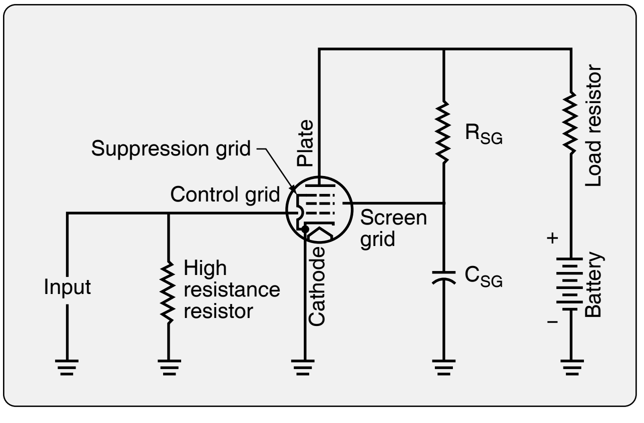


Figure 11-12. *A pentode contains a suppression grid that controls secondary electron emissions from the plate at high power. This keeps the current in the screen grid from becoming too high.*

The outer most orbital shell of any atom's electrons is called the valence shell. The number of electrons in the valence shell determines the chemical properties of the material. When the valence shell has the maximum number of electrons, it is complete, and the electrons tend to be bound strongly to the nucleus. Materials with this characteristic are chemically stable. It takes a large amount of force to move the electrons in this situation from one atom valence shell to that of another. Since the movement of electrons is called electric current, substances with complete valence shells are known as good insulators because they resist the flow of electrons (electricity). *[Figure 11-14]*

In atoms with an incomplete valence shell, that is, those without the maximum number of electrons in their valence shell, the electrons are bound less strongly to the nucleus. The material is chemically disposed to combine with other materials or other identical atoms to fill in the unstable valence configuration and bring the number of electrons in the valence shell to maximum. Two or more substances may share the electrons in their valence shells and form a covalent bond. A covalent bond is the method by which atoms complete their valence shells by sharing valence electrons with other atoms.

Electrons in incomplete valence shells may also move freely from valence shell to valence shell of different atoms or compounds. In this case, these are known as free electrons. As stated, the movement of electrons is known as electric

Shell or Orbit Number	1	2	3	4	5
Maximum number of electrons	2	8	18	32	50

Figure 11-13. *Maximum number of electrons in each orbital shell of an atom.*

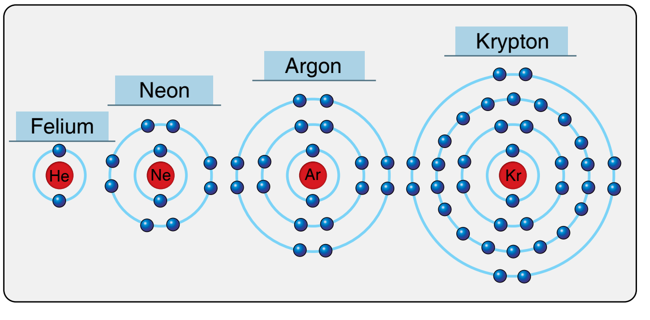


Figure 11-14. *Elements with full valence shells are good insulators. Most insulators used in aviation are compounds of two or more elements that share electrons to fill their valence shells.*

current or current flow. When electrons move freely from atom to atom or compound to compound, the substance is known as a conductor. *[Figure 11-15]*

Not all materials are pure elements, that is, substances made up of one kind of atom. Compounds occur when two or more different types of atoms combine. They create a new substance with different characteristics than any of the component elements. When compounds form, valence shells and their maximum number of electrons remain the rule of physics. The new compound molecule may either share electrons to fill the valence shell or free electrons may exist to make it a good conductor.

Silicon is an atomic element that contains four electrons in its valence shell. It tends to combine readily with itself and form a lattice of silicon atoms in which adjacent atoms share electrons to fill out the valance shell of each to the maximum of eight electrons. *[Figure 11-16]* This unique symmetric alignment of silicon atoms results in a crystalline structure.

Once bound together, the valence shells of each silicon atom are complete. In this state, movement of electrons does not occur easily. There are no free electrons to move to another atom and no space in the valence shells to accept a free electron. Therefore, silicon in this form is a good insulator.

Silicon is a primary material used in the manufacture of semiconductors. Germanium and a few other materials are also used.

Since silicon is an insulator, it must be modified to become a semiconductor. The process often used is called doping. Starting with ultra-pure silicon crystal, arsenic, phosphorus, or some other element with five valence electrons in each atom is mixed into the silicon. The result is a silicon lattice with flaws. *[Figure 11-17]* The elements bond, but numerous free electrons are present in the material from the 5th electron that is part of the valence shell of the doping element atoms. These free electrons can now flow under certain conditions. Thus, the silicon becomes semiconductive. The conditions required for electron flow in a semiconductor are discussed in the following paragraphs.

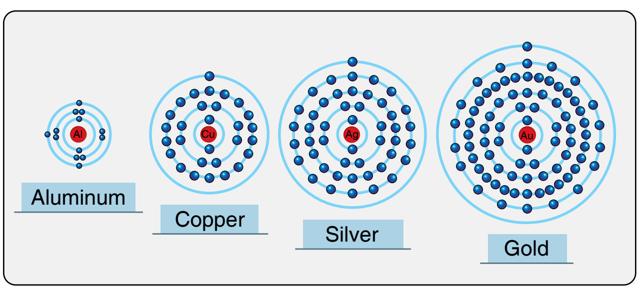


Figure 11-15. *The valence shells of elements that are common conductors have one (or three) electrons.*

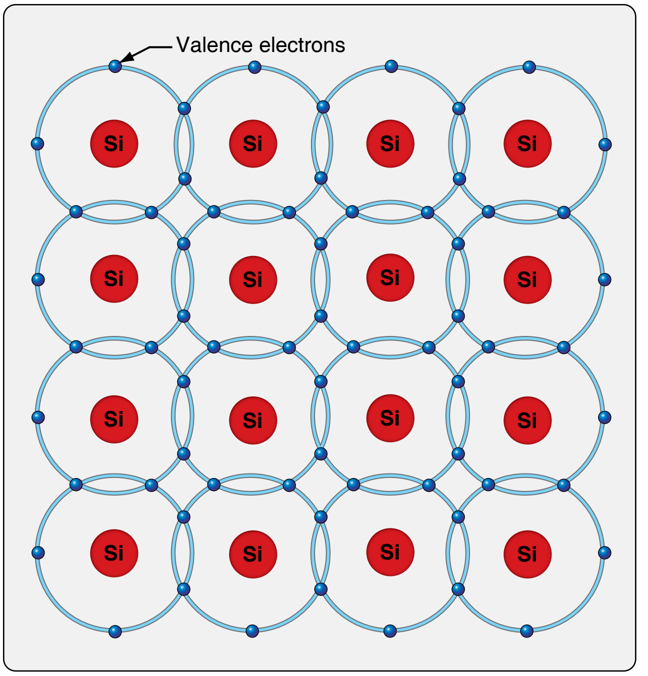


Figure 11-16. *The silicon atoms with just the valence shell electrons share these valence electrons with each other. By sharing with four other silicon atoms, the number of electrons in each silicon atom valence shell becomes eight, which is the maximum number. This makes the substance stable and it resists any flow of electrons.*

When silicon is doped with an element or compound containing five electrons in its valence shell, the result is a negatively charged material due to the excess free electrons, and the fact that electrons are negatively charged. This is known as an N-type semiconductor material. It is also known as a donor material because, when it is used in electronics, it donates the extra electrons to current flow.

Doping silicon can also be performed with an element that has only three valence electrons, such as boron, gallium, or indium. Valence electron sharing still occurs, and the silicon atoms with interspersed doping element atoms form a lattice molecular structure. However, in this case, there are many valence shells where there are only seven electrons and not eight. This greatly changes the properties of the material. The absence of the electrons, called holes, encourages electron flow due to the preference to have eight electrons in all valence shells. Therefore, this type of doped silicon is also semiconductive. It is known as P-type material or as an acceptor since it accepts electrons in the holes under certain conditions. *[Figure 11-18]*

Combining N- and P-type semiconductor material in certain ways can produce very useful results. A look at various semiconductor devices follows.

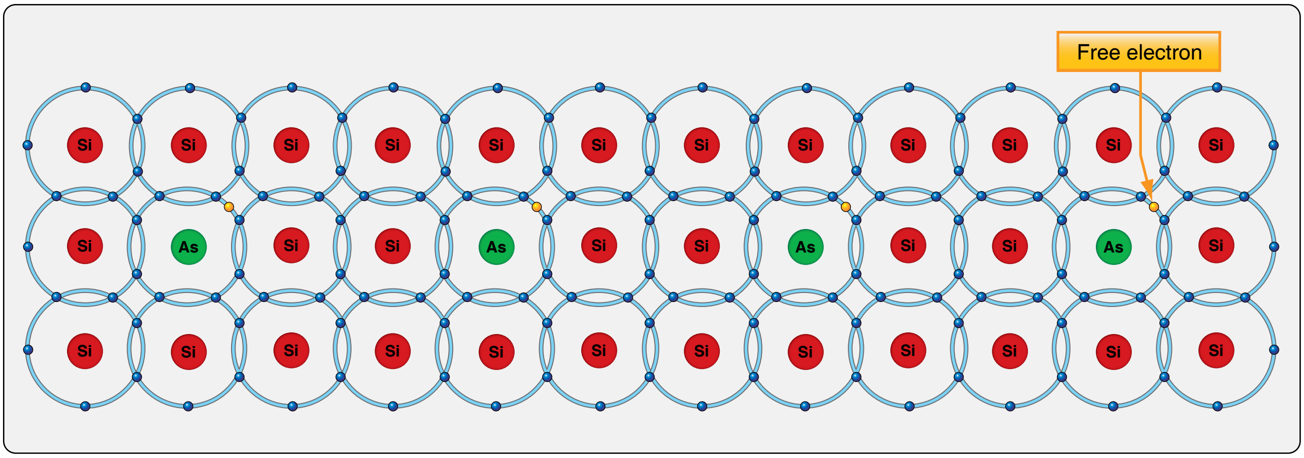


Figure 11-17. *Silicon atoms doped with arsenic form a lattice work of covalent bonds. Free electrons exist in the material from the arsenic atom's 5th valence electron. These are the electrons that flow when the semiconductor material, known as N-type or donor material, is conducting.*

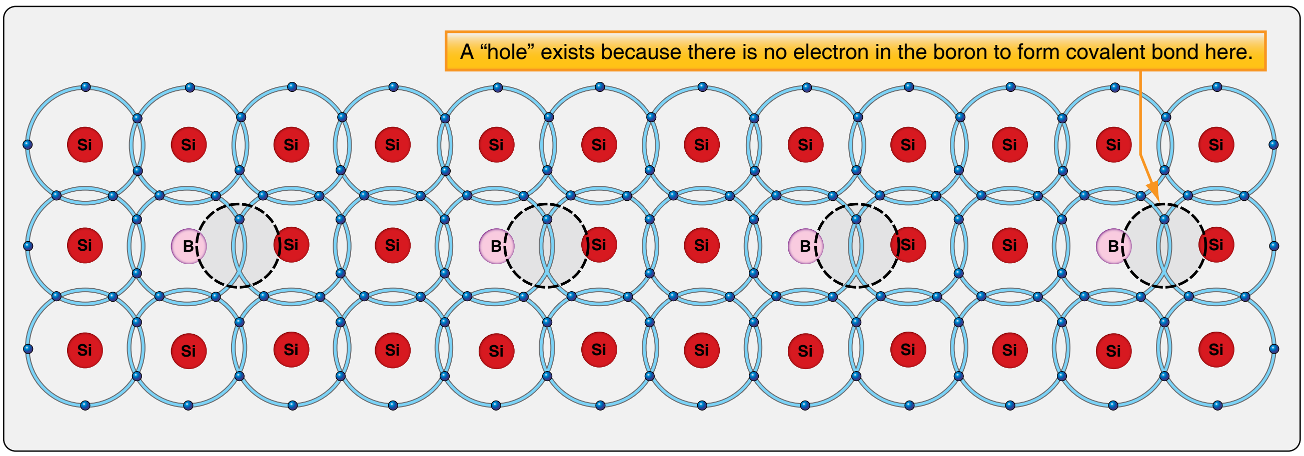


Figure 11-18. *The lattice of boron doped silicon contains holes where the three boron valence shell electrons fail to fill in the combined valence shells to the maximum of eight electrons. This is known as P-type semiconductor material or acceptor material.*

Semiconductor Diodes

A diode is an electrical device that allows current to flow in one direction through the device but not the other. A simple device that can be made from N- and P-type semiconductors is a semiconductor diode. When joined, the junction of these two materials exhibits unique properties. Since there are holes in the P-type material, free electrons from the N-type material are attracted to fill these holes. Once combined, the area at the junction of the two materials where this happens is said to be depleted. There are no longer free electrons or holes. However, having given up some electrons, the N-type material next to the junction becomes slightly positively charged, and having received electrons, the P-type material next to the junction becomes slightly negatively charged.

The depletion area at the junction of the two semiconductor materials constitutes a barrier or potential hill. The intensity of the potential hill is proportional to the width of the depletion area (where the electrons from the N-type material have filled holes in the P-type material). *[Figure 11-19]*

The two semiconductors joined in this manner form a diode that can be used in an electrical circuit. A voltage source is attached to the diode. When the negative terminal of the battery is attached to the N-type semiconductor material and the positive terminal is attached to the P-type material, electricity can flow in the circuit. The negative potential of the battery forces free electrons in the N-type material toward the junction. The positive potential of the battery forces holes in the P-type material toward the other side of the junction. The holes move by the rebinding of the doping agent ions closer to the junction. At the junction, free electrons continuously arrive and fill the holes in the lattice. As this occurs, more room is available for electrons and holes to move into the area. Pushed by the potential of the battery, electrons and holes continue to combine. The depletion area becomes extremely narrow under these conditions. The potential hill or barrier is, therefore, very small. The flow of current in the electrical circuit is in the direction of electron movement shown in *Figure 11-20*.

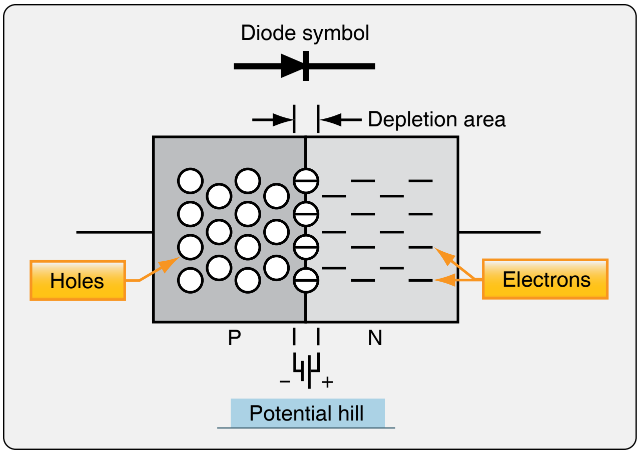


Figure 11-19. *A potential hill.*

In similar circuits where the negative battery terminal is attached to the N-type semiconductor material and the positive terminal is attached to the P-type material, current flows from N-type, or donor material, to P-type receptor material. This is known as a forward-biased semiconductor. A voltage of approximately 0.7 volts is needed to begin the current flow over the potential hill. Thereafter, current flow is linear with the voltage. However, temperature affects the ease at which electrons and holes combine given a specific voltage.

If the battery terminals are reversed, the semiconductor diode circuit is said to be reversed biased. *[Figure 11-21]* Attaching the negative terminal of the battery to the P-type material

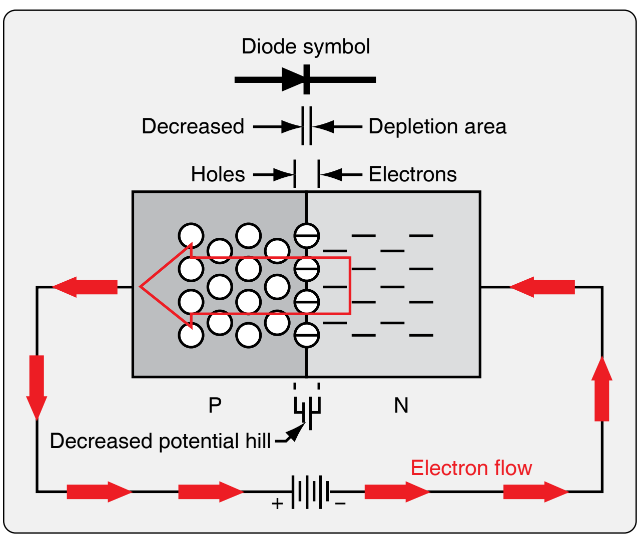


Figure 11-20. *The flow of current and the P-N junction of a semiconductor diode attached to a battery in a circuit.*

attracts the holes in the P-type material away from the junction in the diode. The positive battery terminal attached to the N-type material attracts the free electrons from the junction in the opposite direction. In this way, the width of the area of depletion at the junction of the two materials increases. The potential hill is greater. Current cannot climb the hill; therefore, no current flows in the circuit. The semiconductors do not conduct.

Semiconductor diodes are used often in electronic circuits. When AC current is applied to a semiconductor diode, current flows during one cycle of the AC but not during the other cycle. The diode, therefore, becomes a rectifier. When it is forward biased, electrons flow; when the AC cycles, electrons do not flow. A simple AC rectifier circuit containing a semiconductor diode and a load resistor is illustrated in *Figure 11- 22*. Semiconductor diode symbols and examples of semiconductor diodes are shown in *Figure 11-23*.

Note: Electron flow is typically discussed in this text. The conventional current flow concept where electricity is thought to flow from the positive terminal of the battery through a circuit to the negative terminal is sometimes used in the field.

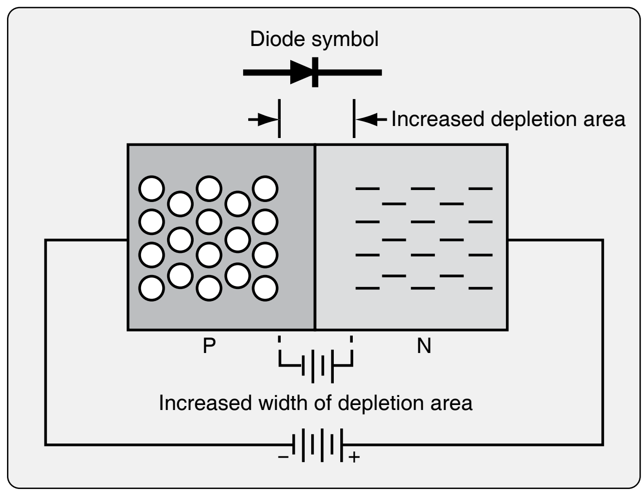


Figure 11-21. *A reversed biased condition.*

Semiconductor diodes have limitations. They are rated for a range of current flow. Above a certain level, the diode overheats and burns up. The amount of current that passes through the diode when forward biased is directly proportional to the amount of voltage applied. But, as mentioned, it is affected by temperature.

Figure 11-24 indicates the actual behavior of a semiconductor diode. In practice, a small amount of current does flow through a semiconductor diode when reversed biased. This is known as leakage current and it is in the micro amperage range. However, at a certain voltage, the blockage of current flow in a reversed biased diode breaks down completely. This voltage is known as the avalanche voltage because the diode can no longer hold back the current and the diode fails.

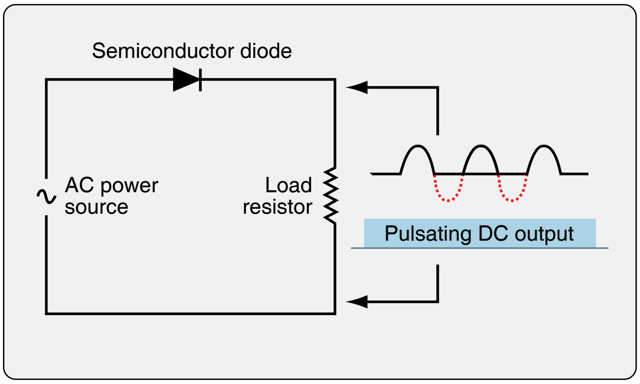


Figure 11-22. *A semiconductor diode acts as a check valve in an AC circuit resulting in a pulsating DC output.*

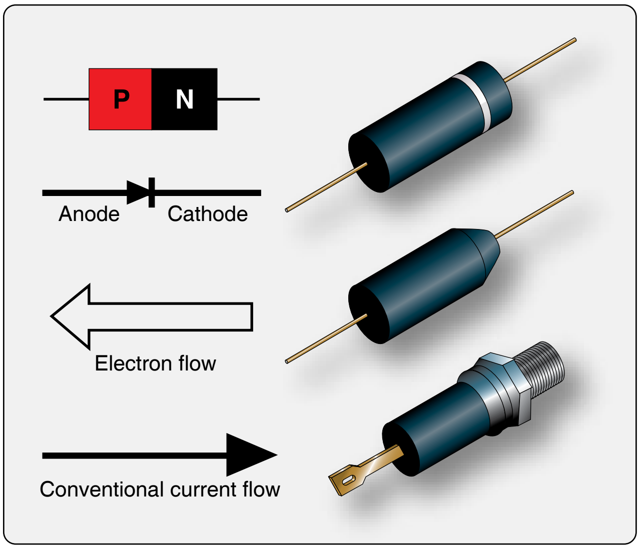


Figure 11-23. *Symbols and drawings of semiconductor diodes.*

Zener Diodes

Diodes can be designed with a zener voltage. This is similar to avalanche flow. When reversed biased, only leakage current flows through the diode. However, as the voltage is increased, the zener voltage is reached. The diode lets current flow freely through the diode in the direction in which it is normally blocked. The diode is constructed to be able to handle the zener voltage and the resulting current, whereas avalanche voltage burns out a diode. A zener diode can be used as means of dropping voltage or voltage regulation. It can be used to step down circuit voltage for a particular application but only when certain input conditions exist. Zener diodes are constructed to handle a wide range of voltages. *[Figure 11-25]*

Transistors

While diodes are very useful in electronic circuits, semiconductors can be used to construct true control valves known as transistors. A transistor is little more than a sandwich

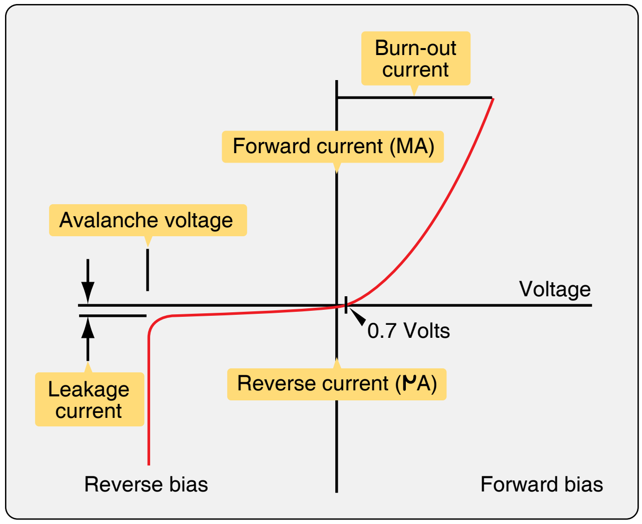


Figure 11-24. *A semiconductor diode.*

of N-type semiconductor material between two pieces of P-type semiconductor material or vice versa. However, a transistor exhibits some remarkable properties and is the building block of all things electronic. *[Figure 11-26]* As with any union of dissimilar types of semiconductor materials, the junctions of the P- and N- materials in a transistor have depletion areas that create potential hills for the flow of electrical charges.

Like a vacuum tube triode, the transistor has three electrodes or terminals, one each for the three layers of semiconductor material. The emitter and the collector are on the outside of the sandwiched semiconductor material. The center material

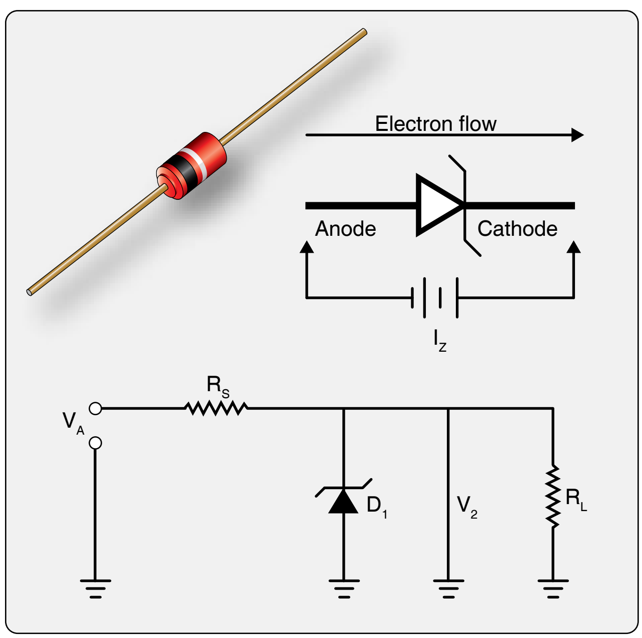


Figure 11-25. *A zener diode, when reversed biased, will break down and allow a prescribed voltage to flow in the direction normally blocked by the diode.*

is known as the base. A change in a relatively small amount of voltage applied to the base of the transistor allows a relatively large amount of current to flow from the collector to the emitter. In this way, the transistor acts as a switch with a small input voltage controlling a large amount of current.

If a transistor is put into a simple battery circuit, such as the one shown in *Figure 11-27*, voltage from the battery (EB) forces free electrons and holes toward the junction between the base and the emitter just as it does in the junction of a semiconductor diode. The emitter-base depletion area becomes narrow as free electrons combine with the holes at the junction. Current (IB) (solid arrows) flows through the junction in the emitter-base battery circuit. At the same time, an emitter-collector circuit is constructed with a battery (EC) of much higher voltage in its circuit. Because of the narrow depletion area at the emitter-base junction, current IC is able to cross the collector base junction, flow through emitter-base junction, and complete the collector-emitter battery circuit (hollow arrows).

To some extent, varying the voltage to the base material can increase or decrease the current flow through the transistor as the emitter-base depletion area changes width in response to the base voltage. If base voltage is removed, the emitter-base depletion area becomes too wide and all current flow through the transistor ceases.

Current in the transistor circuit illustrated has a relationship as follows: $I_E = I_B + I_C$. It should be remembered that it is the voltage applied to the base that turns the collector-emitter transistor current on or off.

Controlling a large amount of current flow with a small independent input voltage is very useful when building electronic circuits. Transistors are the building blocks from which all electronic devices are made, including Boolean gates that are used to create microprocessor chips. As production techniques have developed, the size of reliable transistors has shrunk. Now, hundreds of millions and even billions of transistors may be used to construct a single chip such as the one that powers your computer and various avionic devices.

Silicon Controlled Rectifiers

Combination of semiconductor materials is not limited to a two-type, three-layer sandwich transistor. By creating a four-layer sandwich of alternating types of semiconductor material (i.e., PNPN or NPNP), a slightly different semiconductor diode is created. As is the case in a two-layer diode, circuit current is either blocked or permitted to flow through the diode in a single direction.

Within a four-layer diode, sometimes known as a Shockley diode, there are three junctions. The behavior of the junctions and the entire four-layer diode can be understood by considering it to be two interconnected three-layer transistors. *[Figure 11-28]* Transistor behavior includes no current flow until the base material receives an applied voltage to narrow the depletion area at the base-emitter junction. The base materials in the four-layer diode transistor model receive charge from the other transistor's collector. With no other means of reducing any of the depletion areas at the junctions, it appears that current does not flow in either direction in this device. However, if a large voltage is applied to forward bias the anode or cathode, at some point the ability to block flow breaks down. Current flows through whichever transistor is charged. Collector current then charges the base of the other transistor and current flows through the entire device.

Some caveats are necessary with this explanation. The transistors that comprise this four-layer diode must be constructed of material similar to that described in a zener diode. That is, it must be able to endure the current flow without burning out. In this case, the voltage that causes the diode to conduct is known as breakover voltage rather than breakdown voltage. Additionally, this diode has the unique characteristic of allowing current flow to continue until the applied voltage is reduced significantly, in most cases, until it is reduced to zero. In AC circuits, this would occur when the AC cycles.

While the four-layer, Shockley diode is useful as a switching device, a slight modification to its design creates a silicon-controlled rectifier (SCR). To construct a SCR, an additional terminal known as a gate is added. It provides more control and utility. In the four-layer semiconductor construction, there are always two junctions forward biased and one junction reversed biased. The added terminal allows the momentary application of voltage to the reversed biased junction. All three junctions then become forward biased and current at the anode flows through the device. Once voltage is applied to the gate, the SCR become latched or locked on. Current continues to flow through it until the level drops off significantly, usually to zero. Then, another applied voltage through the gate is needed to reactivate the current flow. *[Figures 11-29* and *11-30]*

SCRs are often used in high voltage situations, such as power switching, phase controls, battery chargers, and inverter circuits. They can be used to produce variable DC voltages for motors and are found in welding power supplies. Often, lighting dimmer systems use SCRs to reduce the average voltage applied to the lights by only allowing current flow during part of the AC cycle. This is controlled by controlling the pulses to the SCR gate and eliminating the massive heat

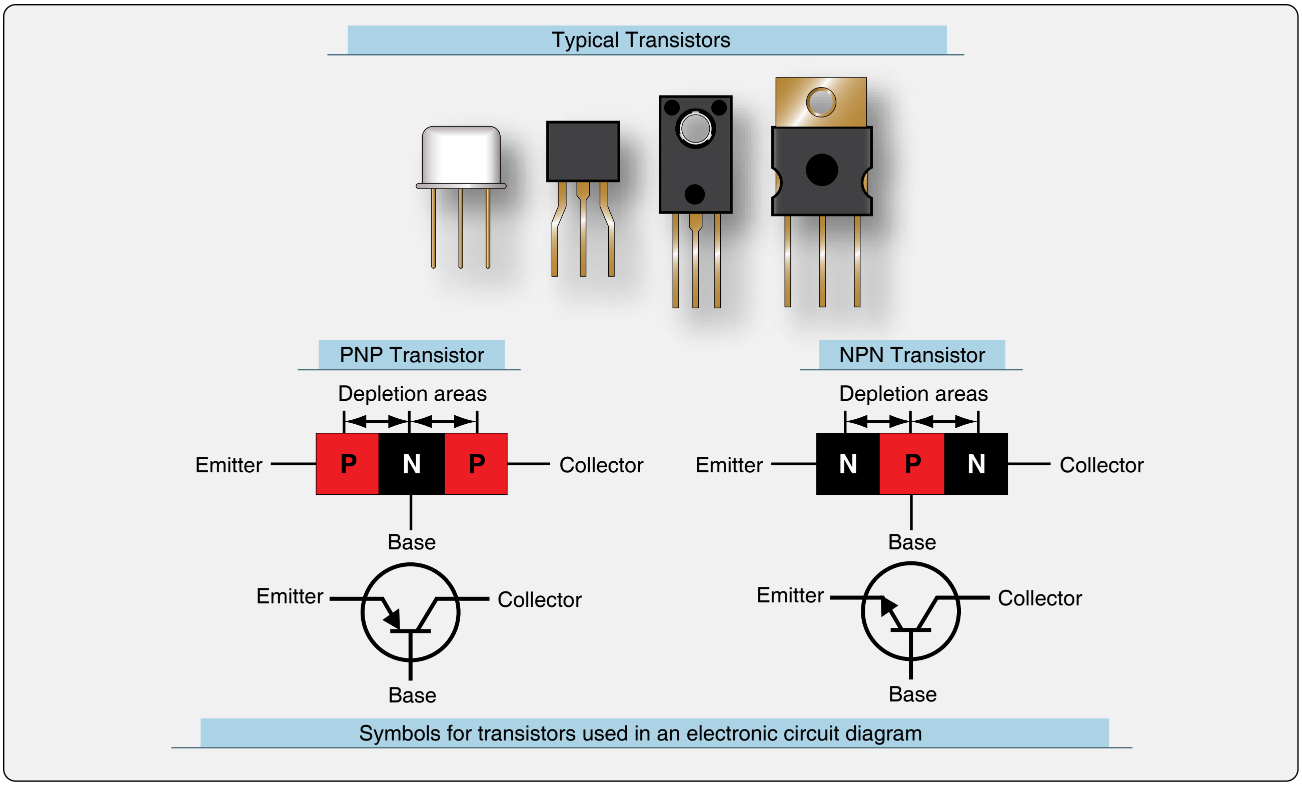


Figure 11-26. *Typical transistors, diagrams of a PNP and NPN transistor, and the symbol for those transistors when depicted in an electronic circuit diagram.*

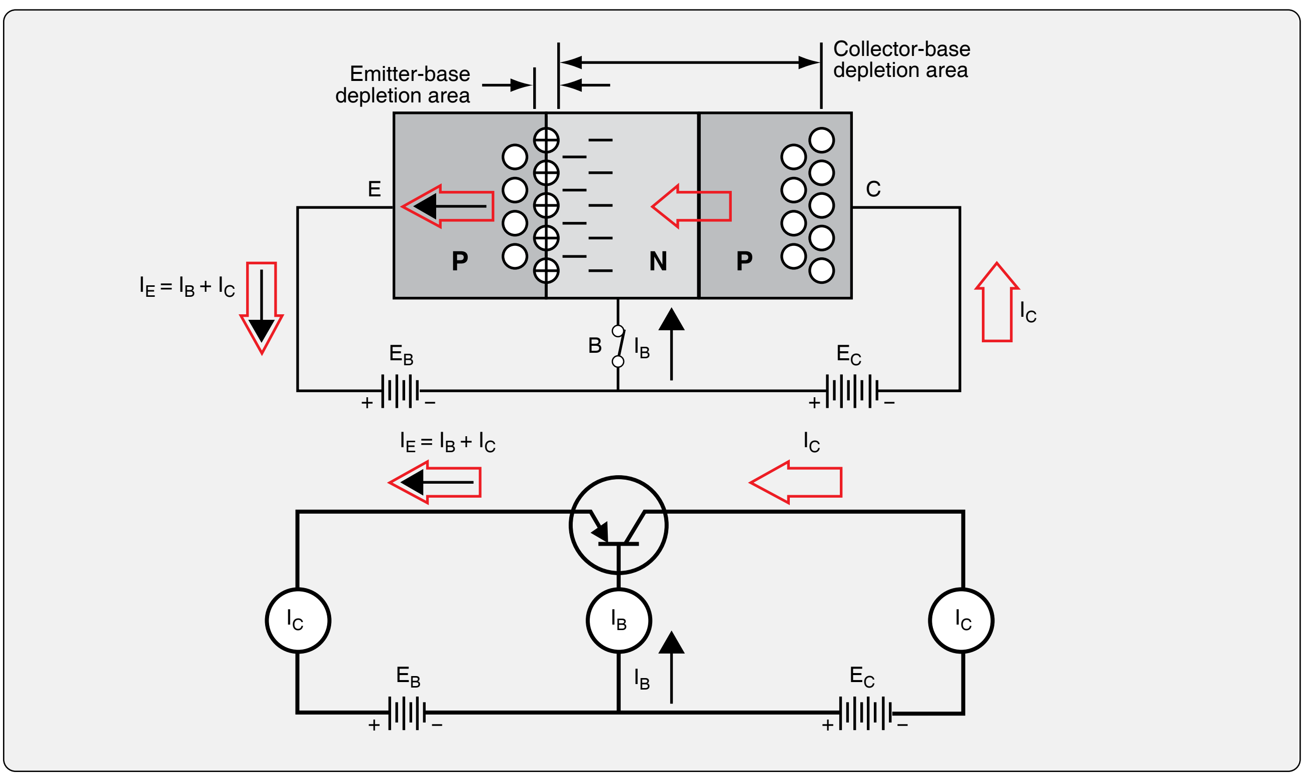


Figure 11-27. *The effect of applying a small voltage to bias the emitter-base junction of a transistor (top). A circuit diagram for this same transistor (bottom).*

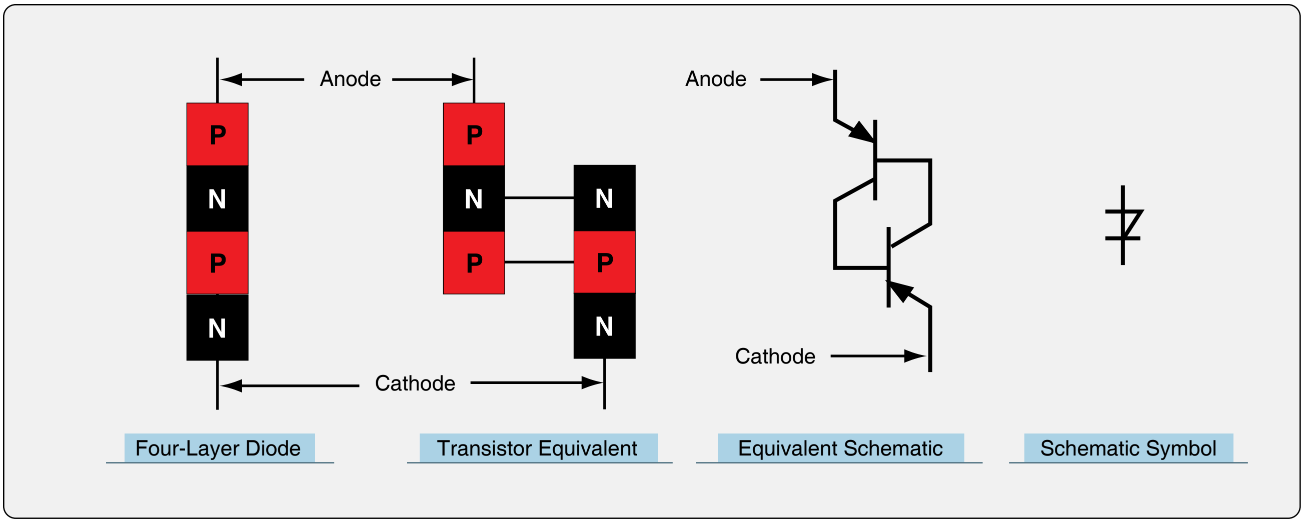


Figure 11-28. *A four-layer semiconductor diode behaves like two transistors. When breakover voltage is reached, the device conducts current until the voltage is removed.*

dissipation caused when using resistors to reduce voltage. *Figure 11-31* graphically depicts the timing of the gate pulse that limits full cycle voltage to the load. By controlling the phase during which, time the SCR is latched, a reduced average voltage is applied.

Triacs

SCRs are limited to allowing current flow in one direction only. In AC circuitry, this means only half of the voltage cycle can be used and controlled. To access the voltage in the reverse cycle from an AC power source, a triac can be used. A triac is also a four-layer semiconductor device. It differs from an SCR in that it allows current flow in both directions. A triac has a gate that works the same way as in a SCR; however, a positive or negative pulse to the gate triggers current flow in a triac. The pulse polarity determines the direction of the current flow through the device.

Figure 11-32 illustrates a triac and shows a triac in a simple circuit. It can be triggered with a pulse of either polarity and remains latched until the voltage declines, such as when the AC cycles. Then, it needs to be triggered again. In many ways, the triac acts as though it is two SCRs connected side by side only in opposite directions. Like an SCR, the timing of gate pulses determines the amount of the total voltage that is allowed to pass. The output waveform if triggered at 90° is shown in *Figure 11-32*. Because a triac allows current to flow in both directions, the reverse cycle of AC voltage can also be used and controlled.

When used in actual circuits, triacs do not always maintain the same phase firing point in reverse as they do when fired with a positive pulse. This problem can be regulated somewhat through the use of a capacitor and a diac in the gate circuit. However, as a result, where precise control is

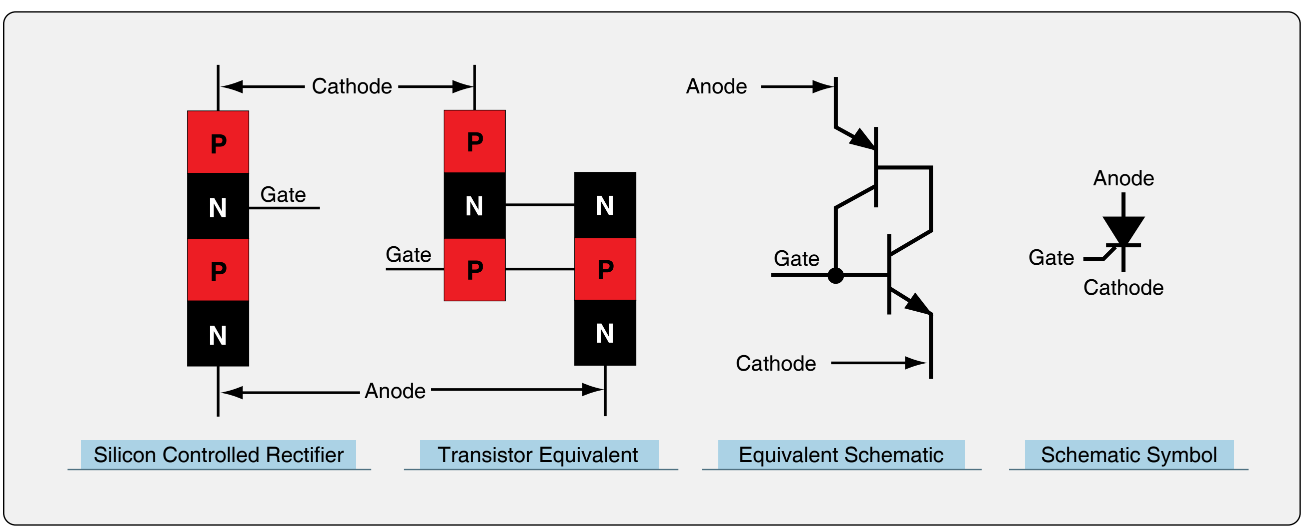


Figure 11-29. *A silicon controlled rectifier (SCR) allows current to pass in one direction when the gate receives a positive pulse to latch the device in the on position. Current ceases to flow when it drops below holding current, such as when AC current reverses cycle.*

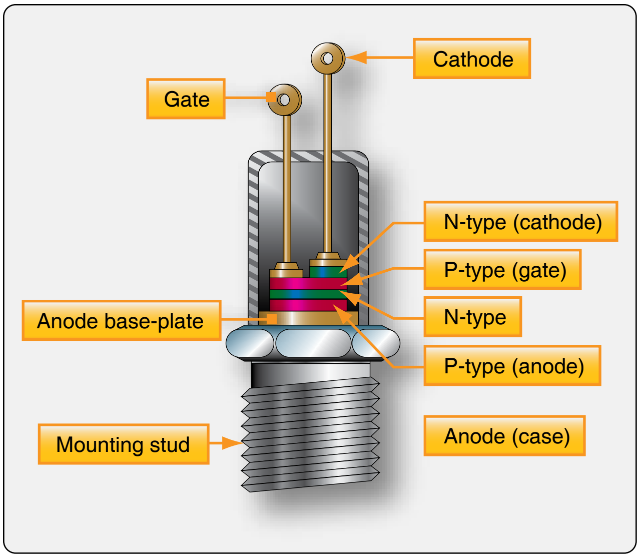


Figure 11-30. *Cross-section of a medium-power SCR.*

required, two SCRs in reverse of each other are often used instead of the triac. Triacs do perform well in lower voltage circuits. *Figure 11-33* illustrates the semiconductor layering in a triac.

Note: The four layers of N- and P-type materials are not uniform as they were in previously described semiconductor devices. None the less, gate pulses affect the depletion areas at the junctions of the materials in the same way allowing current to flow when the areas are narrowed.

Unijunction Transistors (UJT)

The behavior of semiconductor materials is exploited through the construction of numerous transistor devices containing various configurations of N-type and P-type materials. The physical arrangement of the materials in relation to each other yields devices with unique behaviors and applications. The transistors described above having two junctions of P-type and N-type materials (PN) are known as bipolar junction transistors. Other more simple transistors can be fashioned with only one junction of the PN semiconductor materials. These are known as unijunction transistors (UJT). [*Figure 11-34]*

The UJT contains one base semiconductor material and a different type of emitter semiconductor material. There is no collector material. One electrode is attached to the emitter and two electrodes are attached to the base material at opposite ends. These are known as base 1 (B1) and base 2 (B2). The electrode configuration makes the UJT appear physically the same as a bipolar junction transistor. However, there is only one PN junction in the UJT and it behaves differently.

The base material of a UJT behaves like a resistor between the electrodes. With B2 positive with respect to B1, voltage gradually drops as it flows through the base. *[Figure 11-35]* By placing the emitter at a precise location along the base material gradient, the amount of voltage needed to be applied to the emitter electrode to forward bias the UJT base-emitter junction is determined. When the applied emitter voltage

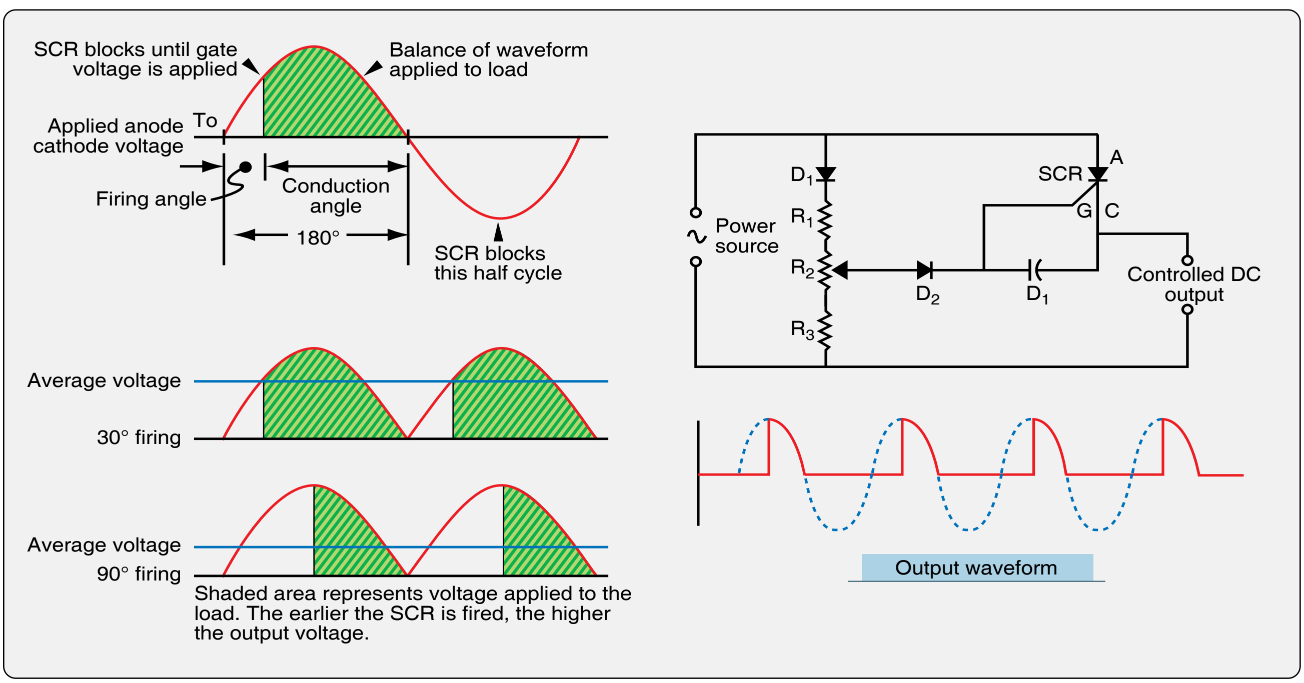


Figure 11-31. *Phase control is a key application for SCR. By limiting the percentage of a full cycle of AC voltage that is applied to a load, a reduced voltage results. The firing angle or timing of a positive voltage pulse through the SCR's gate latches the device open allowing current flow until it drops below the holding current, which is usually at or near zero voltage as the AC cycle reverses.*

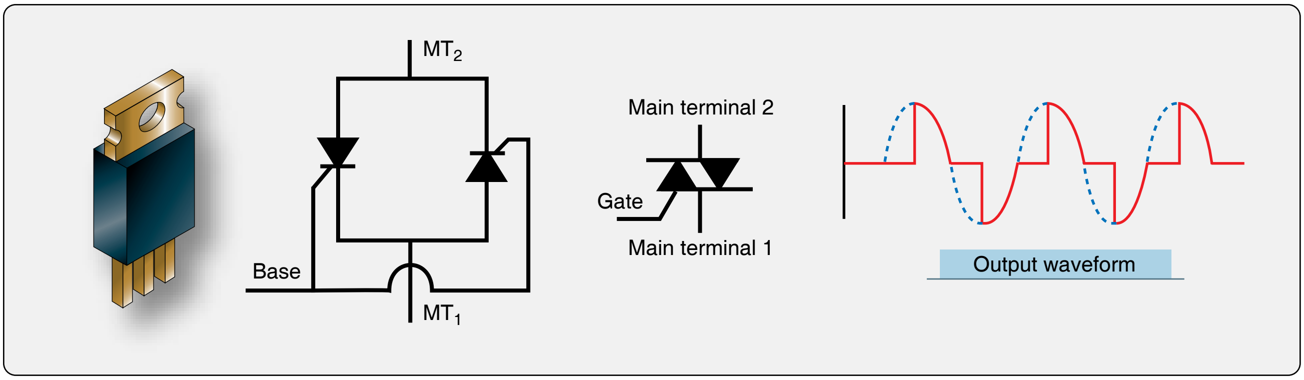


Figure 11-32. *A triac is a controlled semiconductor device that allows current flow in both directions.*

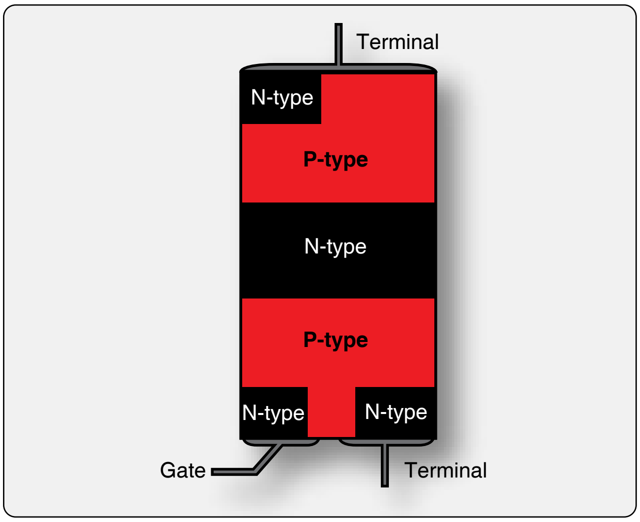


Figure 11-33. *The semiconductor layering in a triac. A positive or negative gate pulse with respect to the upper terminal allows current to flow through the devise in either direction.*

exceeds the voltage at the gradient point where the emitter is attached, the junction is forward biased and current flows freely from the B1 electrode to the E electrode. Otherwise, the junction is reversed biased and no significant current flows although there is some leakage. By selecting a UJT with the correct bias level for a particular circuit, the applied emitter voltage can control current flow through the device.

UJTs of a wide variety of designs and characteristics exist. A description of all of them is beyond the scope of this discussion. In general, UJTs have some advantages over bipolar transistors. They are stable in a wide range of temperatures. In some circuits, use of UJTs can reduce the overall number of components used, which saves money and potentially increases reliability. They can be found in switching circuits, oscillators, and wave shaping circuits. However, four-layered semiconductor thyristors that function the same as the UJT just described are less expensive and most often used.

Field Effect Transistors (FET)

As shown in the triac and the UJT, creative arrangement of semiconductor material types can yield devices with a variety of characteristics. The field effect transistor (FET) is another such device which is commonly used in electronic circuits. Its N- and P-type material configuration is shown in *Figure 11-36*. A FET contains only one junction of the two types of semiconductor material. It is located at the gate where it contacts the main current carrying portion of the device. Because of this, when an FET has a PN junction, it is known as a junction field effect transistor (JFET). All FETs operate by expanding and contracting the depletion area at the junction of the semiconductor materials.

One of the materials in a FET or JFET is called the channel. It is usually the substrate through which the current needing to be controlled flows from a source terminal to a drain terminal. The other type of material intrudes into the channel and acts as the gate. The polarity and amount of voltage applied to the gate can widen or narrow the channel due to expansion or shrinking of the depletion area at the junction of the semiconductors. This increases or decreases the amount of current that can flow through the channel. Enough reversed biased voltage can be applied to the gate to prevent the flow of current through the channel. This allows the FET to act as a switch. It can also be used as a voltage-controlled resistance.

FETs are easier to manufacture than bipolar transistors and have the advantage of staying on once current flow begins without continuous gate voltage applied. They have higher impedance than bipolar transistors and operate cooler. This makes their use ideal for integrated circuits where millions of FETs may be in use on the same chip. FETs come in N-channel and P-channel varieties.

Metal Oxide Semiconductor Field Effect Transistors (MOSFETs) & Complementary Metal Oxide Semiconductor (CMOS)

The basic FET has been modified in numerous ways and continues to be at the center of faster and smaller electronic

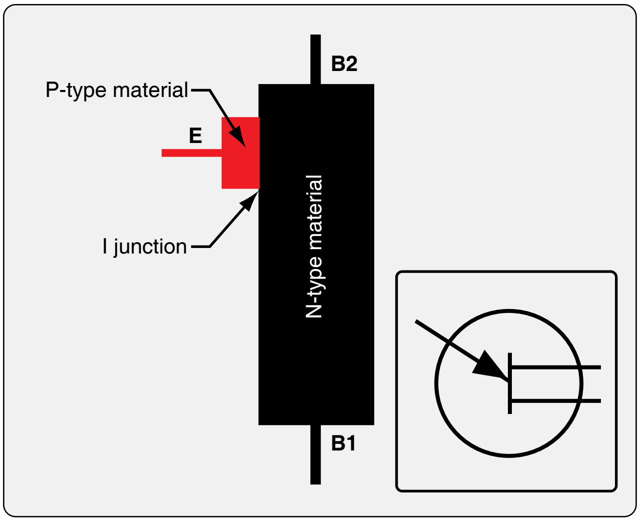


Figure 11-34. *A unijunction transistor (UJT).*

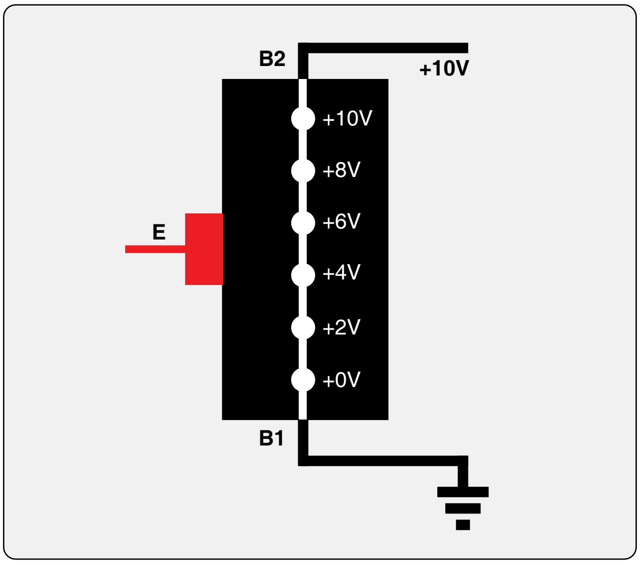


Figure 11-35. *The voltage gradient in a UJT.*

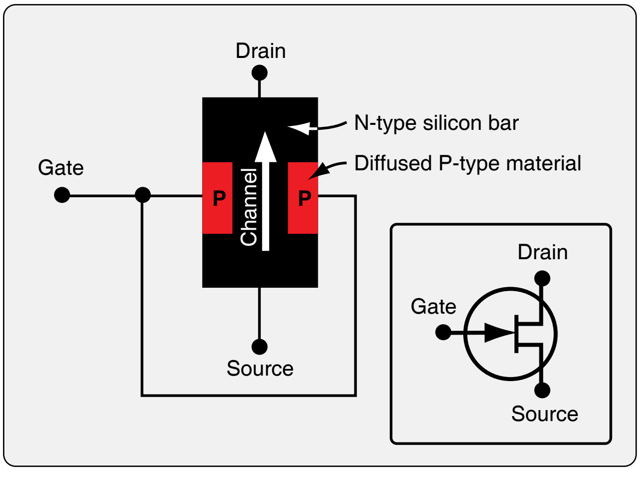


Figure 11-36. *The basic structure of a field effect transistor and its electronic symbol.*

component development. A version of the FET widely used is the metal oxide semiconductor field effect transistor (MOSFET). The MOSFET uses a metal gate with a thin insulating material between the gate and the semiconductor material. This essentially creates a capacitor at the gate and eliminates current leakage in this area. Modern versions of the MOSFET have a silicon dioxide insulating layer and many have poly-crystalline silicon gates rather than metal, but the MOSFET name remains and the basic behavioral characteristic are the same. *[Figure 11-37]*

As with FETs, MOSFETs come with N-channels or P-channels. They can also be constructed as depletion mode or enhancement mode devices. This is analogous to a switch being normally open or normally closed. Depletion mode MOSFETs have an open channel that is restricted or closed when voltage is applied to the gate (i.e., normally open). Enhancement mode MOSFETs allow no current to flow at zero bias but create a channel for current flow when voltage is applied to the gate (normally closed). No voltage is used when the MOSFETs are at zero bias. Millions of enhancement mode MOSFETs are used in the construction of integrated circuits. They are installed in complimentary pairs such that when one is open, the other is closed. This basic design is known as complementary MOSFET (CMOS), which is the basis for integrated circuit design in nearly all modern electronics. Through the use of these transistors, digital logic gates can be formed, and digital circuitry is constructed.

Other more specialized FETs exist. Some of their unique characteristics are owed to design alterations and others to material variations. The transistor devices discussed above use silicon-based semiconductors. But the use of other semiconductor materials can yield variations in performance. Metal semiconductor FETs (MESFETS) for example, are often used in microwave applications. They have a combined metal and semiconductor material at the gate and are typically made from gallium arsenide or indium phosphide. MESFETs are used for their quickness when starting and stopping current flows especially in opposite directions. High electron mobility transistors (HEMT) and pseudomorphic high electron mobility transistors (PHEMT) are also constructed from gallium arsenide semiconductor material and are used for high-frequency applications.

Photodiodes & Phototransistors

Light contains electromagnetic energy that is carried by photons. The amount of energy depends on the frequency of light of the photon. This energy can be very useful in the operation of electronic devices since all semiconductors are affected by light energy. When a photon strikes a semiconductor atom, it raises the energy level above what is needed to hold its electrons in orbit. The extra energy frees an electron enabling it to flow as current. The vacated position

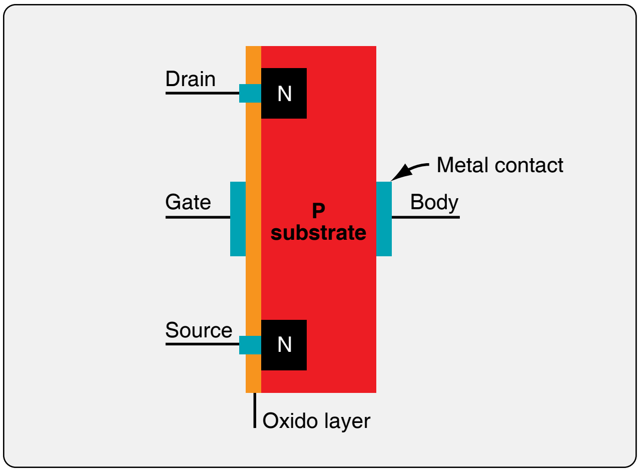


Figure 11-37. *A MOSFET has a metal gate and an oxide layer between it and the semiconductor material to prevent current leakage.*

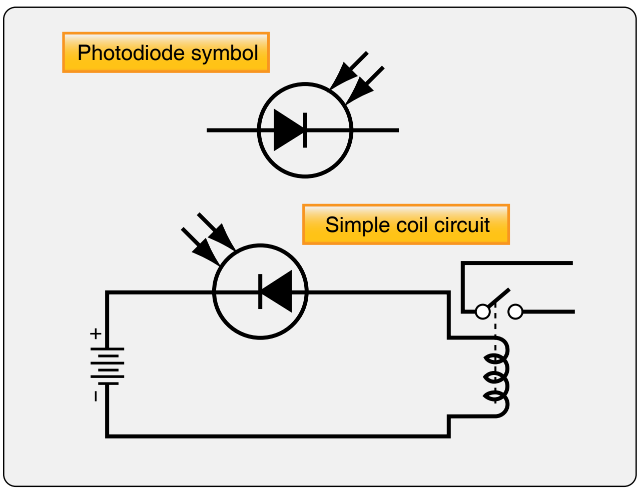


Figure 11-38. *The symbol for a photodiode and a photodiode in a simple coil circuit.*

of the electron becomes a hole. In photodiodes, this occurs in the depletion area of the reversed biased PN junction turning on the device and allowing current to flow.

Figure 11-38 illustrates a photodiode in a coil circuit. In this case, the light striking the photodiode causes current to flow in the circuit whereas the diode would have otherwise blocked it. The result is the coil energizes and closes another circuit enabling its operation.

A photon activated transistor could be used to carry even more current than a photodiode. In this case, the light energy is focused on a collector-base junction. This frees electrons in the depletion area and starts a flow of electrons from the base that turns on the transistor. Once on, heavier current flows from the emitter to the collector. *[Figure 11-39]* In practice, engineers have developed numerous ways to use the energy in light photons to trigger semiconductor devices in electronic circuits. *[Figure 11-40]*

Light Emitting Diodes

Light emitting diodes (LEDs) have become so commonly used in electronics that their importance may tend to be overlooked. Numerous avionics displays and indicators use LEDs for indicator lights, digital readouts, and backlighting of liquid crystal display (LCD) screens.

LEDs are simple and reliable. They are constructed of semiconductor material. When a free electron from a semiconductor drops into a semiconductor hole, energy is given off. This is true in all semiconductor materials. However, the energy released when this happens in certain materials is in the frequency range of visible light. *Figure 11-41* is a table that illustrates common LED colors and the semiconductor material that is used in the construction of the diode.

Note: When the diode is reversed biased, no light is given off. When the diode is forward biased, the energy given off is visible in the color characteristic for the material being used. *Figure 11-42* illustrates the anatomy of a single LED, the symbol of an LED, and a graphic depiction of the LED process.

Basic Analog Circuits

The solid-state semiconductor devices described in the previous section of this chapter can be found in both analog and digital electronic circuits. As digital electronics evolve, analog circuitry is being replaced. However, many aircraft still make use of analog electronics in radio and navigation equipment, as well as in other aircraft systems. A brief look at some of the basic analog circuits follows.

Rectifiers

Rectifier circuits change AC voltage into DC voltage and are one of the most commonly used type of circuits in aircraft electronics. *[Figure 11-43]* The resulting DC waveform

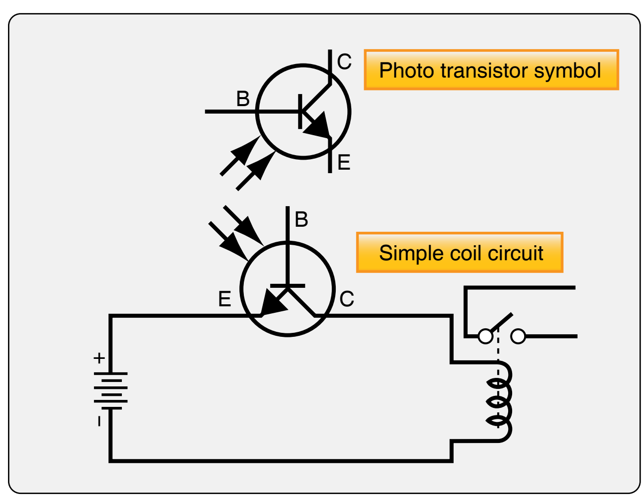


Figure 11-39. *A photo transistor in a simple coil circuit (bottom) and the symbol for a phototransistor (top).*

Figure 11-40. *Phototransistors.*

output is also shown. The circuit has a single semiconductor diode and a load resistor. When the AC voltage cycles below zero, the diode shuts off and does not allow current flow until the AC cycles through zero voltage again. The result is pronounced pulsating DC. While this can be useful, half of the original AC voltage is not being used.

A full wave rectifier creates pulsating DC from AC while using the full AC cycle. One way to do this is to tap the secondary coil at its midpoint and construct two circuits with the load resistor and a diode in each circuit. *[Figure 11-44]* The diodes are arranged so that when current is flowing through one, the other blocks current.

When the AC cycles so the top of the secondary coil of the transformer is positive, current flows from ground, through the load resistor (V_{RL}), Diode 1, and the upper half of the coil. Current cannot flow through Diode 2 because it is blocked. *[Figure 11-44A]* As the AC cycles through zero, the polarity

Color	Wavelength (nm)	Voltage (V)	Semiconductor Material
Infrared	λ > 760	ΔV < 1.9	Gallium arsenide (GaAs) Aluminium gallium arsenide (AlGaAs)
Red	610 < λ < 760	1.63 < ΔV < 2.03	Aluminium gallium arsenide (AlGaAs) Gallium arsenide phosphide (GaAsP) Aluminium gallium indium phosphide (AlGaInP) Gallium(III) phosphide (GaP)
Orange	590 < λ < 610	2.03 < ΔV < 2.10	Gallium arsenide phosphide (GaAsP) Aluminium gallium indium phosphide (AlGaInP) Gallium(III) phosphide (GaP)
Yellow	570 < λ < 590	2.10 < ΔV < 2.18	Gallium arsenide phosphide (GaAsP) Aluminium gallium indium phosphide (AlGaInP) Gallium(III) phosphide (GaP)
Green	500 < λ < 570	1.9[32] < ΔV < 4.0	Indium gallium nitride (InGaN) / Gallium(III) nitride (GaN) Gallium(III) phosphide (GaP) Aluminium gallium indium phosphide (AlGaInP) Aluminium gallium phosphide (AlGaP)
Blue	450 < λ < 500	2.48 < ΔV < 3.7	Zinc selenide (ZnSe) Indium gallium nitride (InGaN) Silicon carbide (SiC) as substrate Silicon (Si) as substrate — (under development)
	400 < λ < 450	2.76 < ΔV < 4.0	Indium gallium nitride (InGaN)
	multiple types	2.48 < ΔV < 3.7	Dual blue/red LEDs, blue with red phosphor, or white with purple plastic
Ultraviolet	λ < 400	3.1 < ΔV < 4.4	Diamond (235 nm)[33] Boron nitride (215 nm)[34][35] Aluminium nitride (AlN) (210 nm)[36] Aluminium gallium nitride (AlGaN) Aluminium gallium indium nitride (AlGaInN) — (down to 210 nm)[37]
White	Broad spectrum	ΔV = 3.5	Blue/UV diode with yellow phosphor

Figure 11-41. *LED colors and the materials used to construct them as well as their wavelength and voltages.*

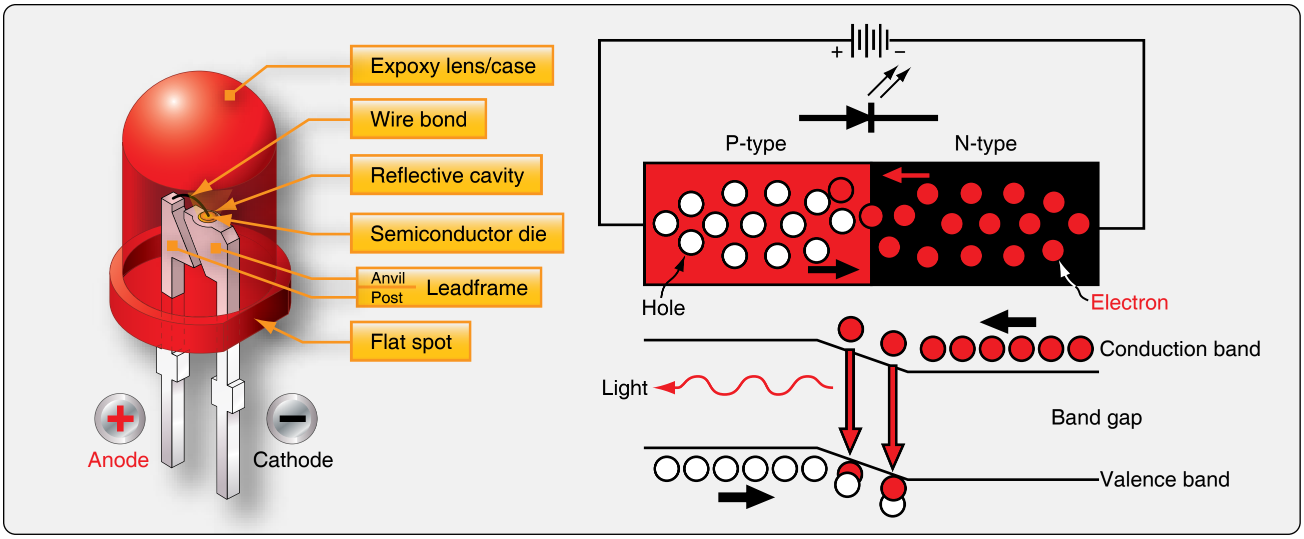


Figure 11-42. *A close up of a single LED (left) and the process of a semi-conductor producing light by electrons dropping into holes and giving off energy (right). The symbol for a light emitting diode is the diode symbol with two arrows pointing away from the junction.*

of the secondary coil changes. *[Figure 11-44B]* Current then flows from ground, through the load resistor, Diode 2, and the bottom half of the secondary coil. Current flow through Diode 1 is blocked. This arrangement yields positive DC from cycling AC with no wasted current.

Another way to construct a full wave rectifier uses four semiconductor diodes in a bridge circuit. Because the secondary coil of the transformer is not tapped at the center, the resultant DC voltage output is twice that of the two-diode full wave rectifier. *[Figure 11-45]* During the first half of the AC cycle, the bottom of the secondary coil is negative. Current flows from it through diode (D_1), then through the load resistor, and through diode (D_2) on its way back to the top of the secondary coil. When the AC reverses its cycle, the polarity of the secondary coil changes. Current flows from the top of the coil through diode (D_3), then through the load resistor, and through diode (D_4) on its way back to the bottom of the secondary coil. The output waveform reflects the higher voltage achieved by rectifying the full AC cycle through the entire length of the secondary coil.

Use and rectification of three-phase AC is also possible on aircraft with a specific benefit. The output DC is very smooth and does not drop to zero. A six-diode circuit is built to rectify

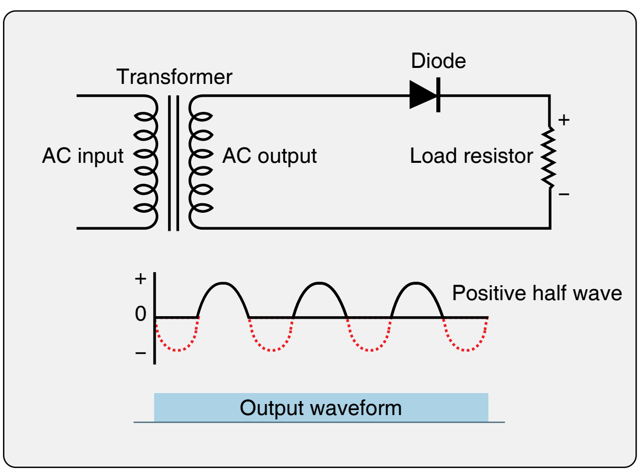


Figure 11-43. *A half wave rectifier uses one diode to produce pulsating DC current from AC. Half of the AC cycle is wasted when the diode blocks the current flow as the AC cycles below zero.*

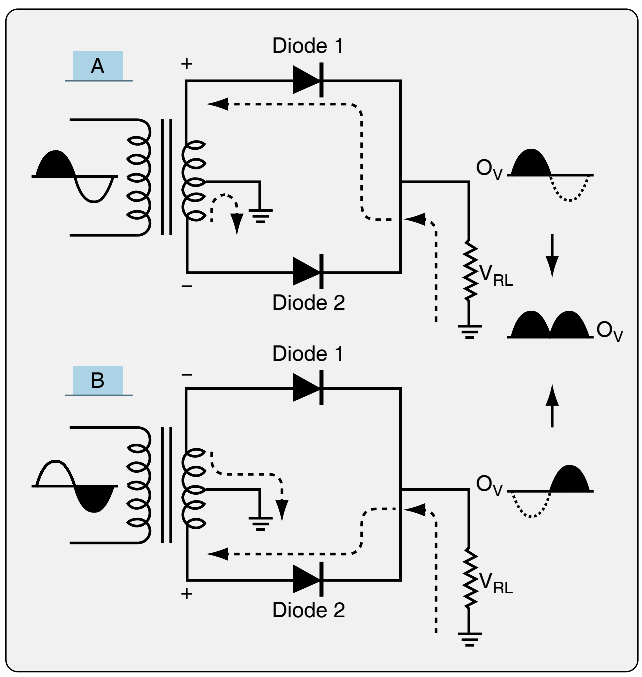


Figure 11-44. *A full wave rectifier can be built by center tapping the secondary coil of the transformer and using two diodes in separate circuits. This rectifies the entire AC input into a pulsating DC with twice the frequency of a half wave rectifier.*

the typical three-phase AC produced by an aircraft alternator. *[Figure 11-46]*

Each stator coil corresponds to a phase of AC and becomes negative for 120° of rotation of the rotor. When stator 1 or the first phase is negative, current flows from it through diode (D_1), then through the load resistor and through diode (D_2) on its way back to the third phase coil. Next, the second phase coil becomes negative and current flows through diode (D_3). It continues to flow through the load resistor and diode (D_4) on its way back to the first phase coil. Finally, the third stage coil becomes negative causing current to flow through diode (D_5), then the load resistor and diode (D_6) on its way back to the second phase coil. The output waveform of this three-phase rectifier depicts the DC produced. It is a relatively steady, non-pulsing flow equivalent to just the tops of the individual curves. The phase overlap prevents voltage from falling to zero producing smooth DC from AC.

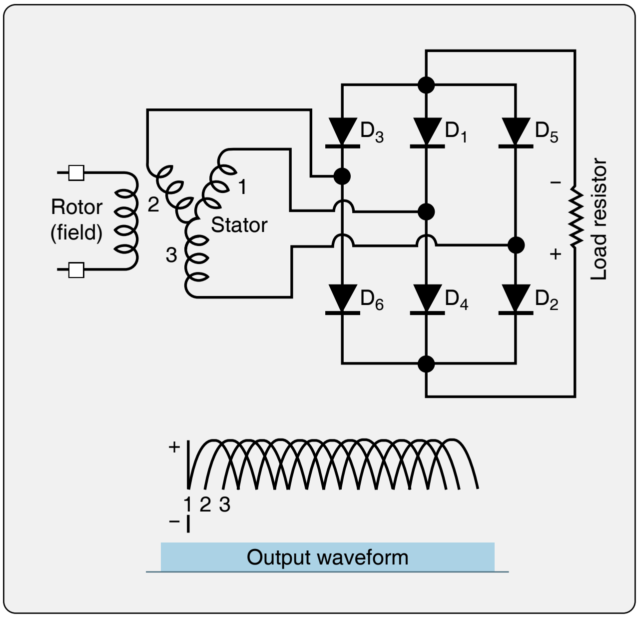


Figure 11-46. *A six-diode, three-phase AC rectifier.*

Amplifiers

An amplifier is a circuit that changes the amplitude of an electric signal. This is done through the use of transistors. As mentioned, a transistor that is forward biased at the base-emitter junction and reversed biased at the collector-base junction is turned on. It can conduct current from the collector to the emitter. Because a small signal at the base can cause a large current to flow from collector to emitter, a transistor in itself can be said to be an amplifier. However, a transistor properly wired into a circuit with resistors, power sources, and other electronic components, such as capacitors, can precisely control more than signal amplitude. Phase and impedance can also be manipulated.

Since the typical bipolar junction transistor requires a based circuit and a collector-emitter circuit, there should be four terminals, two for each circuit. However, the transistor only has three terminals (i.e., the base, the collector, and the emitter). Therefore, one of the terminals must be common to both transistor circuits. The selection of the common terminal affects the output of the amplifier.

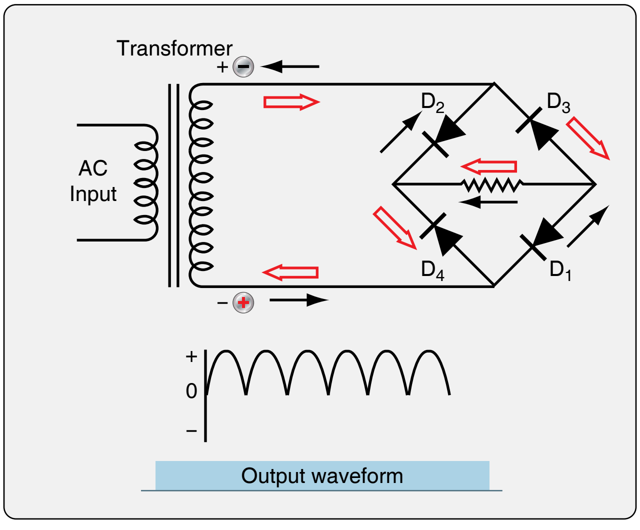


Figure 11-45. *The bridge-type four-diode full wave rectifier circuit is most commonly used to rectify single-phase AC into DC.*

The three basic amplifier types, named for which terminal of the transistor is the common terminal to both transistor circuits, include:

1. Common-emitter amplifier.
2. Common-collector amplifier.
3. Common-base amplifier.

Common-Emitter Amplifier

The common-emitter amplifier controls the amplitude of an electric signal and inverts the phase of the input signal. *Figure 11-47* illustrates a common-emitter amplifier for AC using a NPN transistor and its output signal graph. Common emitter circuits are characterized by high current gain and a 180° voltage phase shift from input to output. It is for the amplification of a microphone signal to drive a speaker. As always, adequate voltage of the correct polarity to the base puts the transistor in the active mode or turns it on. Then, as the base input current fluctuates, the current through the transistor fluctuates proportionally. However, AC cycles through positive and negative polarity. Every 180°, the transistor shuts off because the polarity to the base-emitter junction of the transistor is not correct to forward bias the junction. To keep the transistor on, a DC biasing voltage of the correct polarity (shown as a 2.3 volts (V) battery) is placed in series with the input signal in the base circuit to hold the transistor in the active mode as the AC polarity changes. This way the transistor stays in the active mode to amplify an entire AC signal.

Transistors are rated by ratio of the collector current to the base current, or Beta (β). This is established during the manufacture of the unit and cannot be changed. A 100 β transistor can handle 100 times more current through a collector-emitter circuit than the base input signal. This current in *Figure 11-47* is provided from the 15V battery, V_1. So, the amplitude of amplification is a factor of the beta of the transistor and any inline resistors used in the circuits. The fluctuations of the output signal, however, are entirely controlled by the fluctuations of current input to the transistor base.

If measurements of input and output voltages are made, it is shown that as the input voltage increases, the output voltage decreases. This accounts for the inverted phase produced by a common-emitter circuit. *[Figure 11-47]*

Common-Collector Amplifier

Another basic type of amplifier circuit is the common-collector amplifier. Common-collector circuits are characterized by high current gain, but virtually no voltage gain. The input circuit and the load circuit in this amplifier share the collector terminal of the transistor used. Because the load is in series with the emitter, both the input current and output current run through it. This causes a directly proportional relationship between the input and the output. The current gain in this circuit configuration is high. A small amount of input current can control a large amount of current to flow from the collector to the emitter. A common collector amplifier circuit is illustrated in *Figure 11-48*. The base current needs to flow through the PN junction of the transistor, which has about a 0.7V threshold to be turned on. The output current of the amplifier is the beta value of the transistor plus 1.

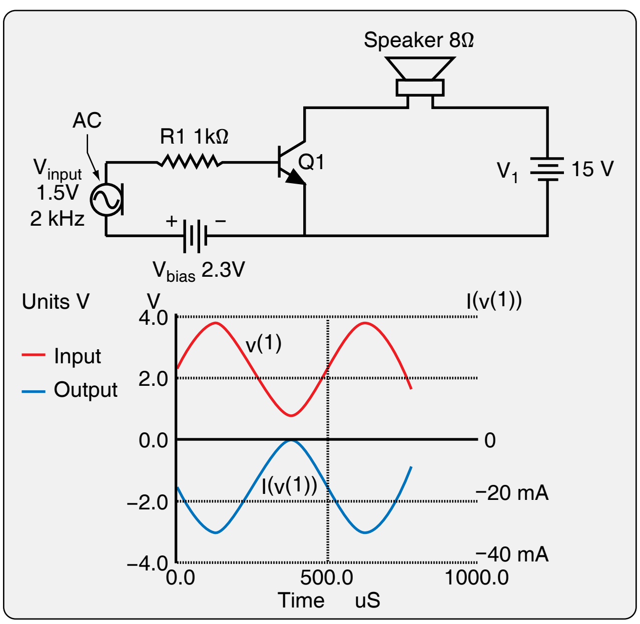


Figure 11-47. *A common-emitter amplifier circuit for amplifying an AC microphone signal to drive a speaker (top) and the graph of the output signal showing a 180 degree shift in phase (bottom).*

During AC amplification, the common-collector amplifier has the same problem that exists in the common-emitter amplifier. The transistor must stay on or in the active mode regardless of input signal polarity. When the AC cycles through zero, the transistor turns off because the minimum amount of current to forward bias the transistor is not available. The addition of a DC biasing source (battery) in series with the AC signal in the input circuit keeps the transistor in the active mode throughout the full AC cycle. *[Figure 11-49]*

A common-collector amplifier can also be built with a PNP transistor. *[Figure 11-50]* It has the same characteristics as the NPN common-collector amplifier shown in *Figure 11-50*. When arranged with a high resistance in the input circuit and a small resistance in the load circuit, the common-collector amplifier can be used to step down the impedance of a signal. *[Figure 11-51]*

Common-Base Amplifier

A third type of amplifier circuit using a bipolar transistor is the common-base amplifier. In this circuit, the shared transistor terminal is the base terminal. *[Figure 11-52]* This causes a unique situation in which the base current is actually larger than the collector or emitter current. As such, the common-base amplifier does not boost current as the other amplifiers do. It attenuates current but causes a high gain in voltage. A very small fluctuation in base voltage in the input circuit causes a large variation in output voltage. The effect on the circuit output is direct, so the output voltage phase is the same as the input signal but much greater in amplitude.

As with the other amplifier circuits, when amplifying an AC signal with a common-base amplifier circuit, the input signal to the base must include a DC source to forward bias the transistor's base-emitter junction. This allows current to flow from the collector to the emitter during both cycles of the AC. A circuit for AC amplification is illustrated in *Figure 11-53* with a graph of the output voltage showing the large increase produced. The common-base amplifier is limited in its use since it does not increase current flow. This makes it the least used configuration. However, it is used in radio frequency amplification because of the low input Z. *Figure 11-54* summarizes the characteristics of the bipolar amplifier circuits discussed above.

Note: There are many variations in circuit design. JFETs and MOSFETS are also used in amplifier circuits, usually in small signal amplifiers due to their low noise outputs.

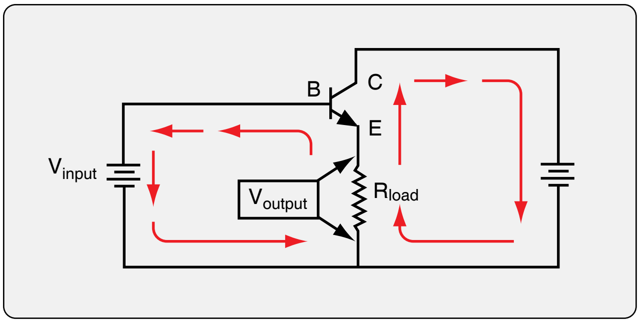


Figure 11-48. *A basic common-collector amplifier circuit. Both the input and output circuits share a path through the load and the emitter. This causes a direct relationship of the output current to the input current.*

Oscillator Circuits

Oscillators function to make AC from DC. They can produce various waveforms as required by electronic circuits. There are many different types of oscillators and oscillator circuits. Some of the most common types are discussed below.

A sine wave is produced by generators when a conductor is rotated in a uniform magnetic field. The typical AC sine wave is characterized by a gradual build-up and decline of voltage in one direction, followed by a similar smooth build-up to peak voltage and decline to zero again in the opposite direction. The value of the voltage at any given time in the cycle can be calculated by taking the peak voltage and

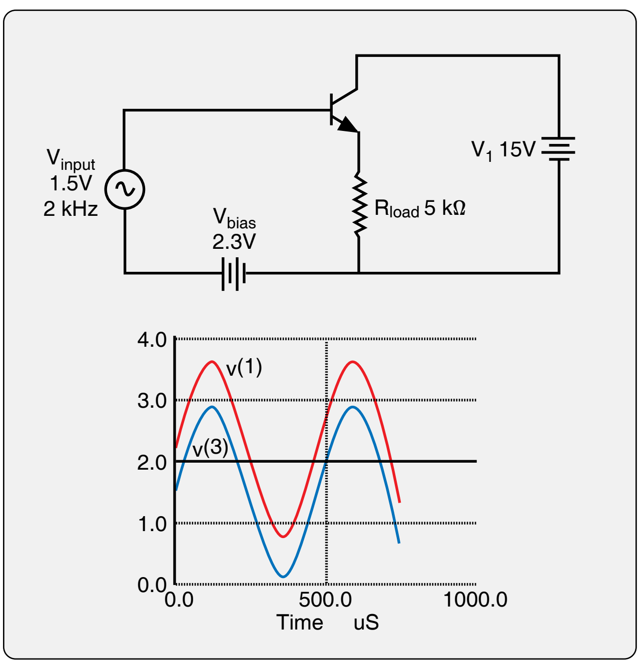


Figure 11-49. *A DC biasing current is used to keep the transistor of a common-collector amplifier in the active mode when amplifying AC (top). The output of this amplifier is in phase and directly proportional to the input (bottom). The difference in amplitude between the two is the 0.7V used to bias the PN junction of the transistor in the input circuit.*

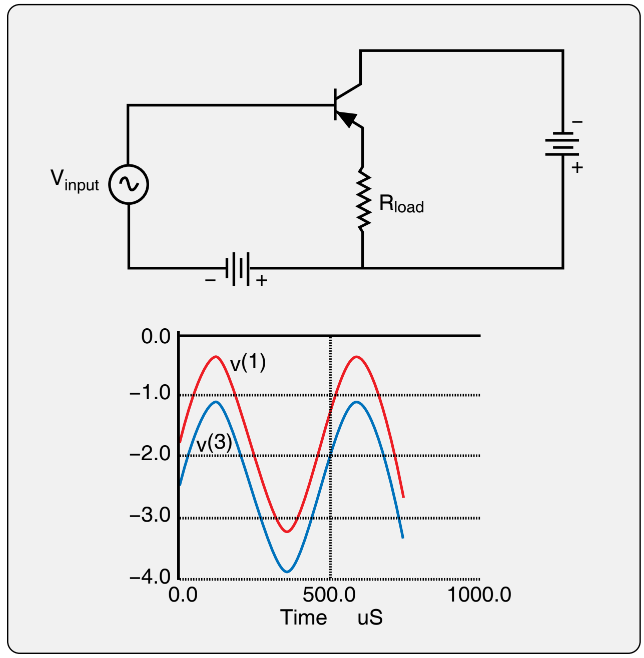


Figure 11-50. *A common-collector amplifier circuit with a PNP transistor has the same characteristics as that of a common-collector amplifier with a NPN transistor except for reversed voltage polarities and current direction.*

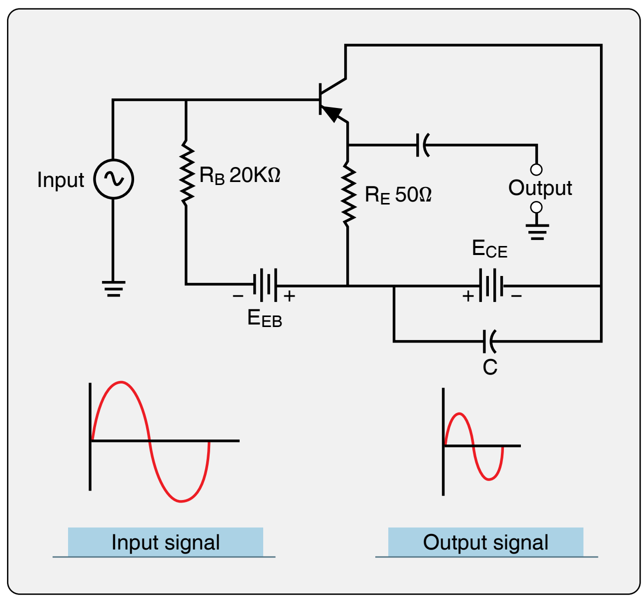


Figure 11-51. *This common-collector circuit has high input impedance and low output impedance.*

multiplying it by the sine of the angle through which the conductor has rotated. *[Figure 11-55]*

A square wave is produced when there is a flow of electrons for a set period that stops for a set amount of time and then repeats. In DC current, this is simply pulsing DC. *[Figure 11-56]* This same wave form can be of opposite polarities when passed through a transformer to produce AC. Certain oscillators produce square waves.

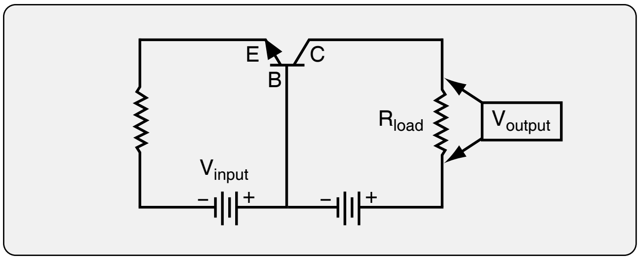


Figure 11-52. *A common-base amplifier circuit for DC current.*

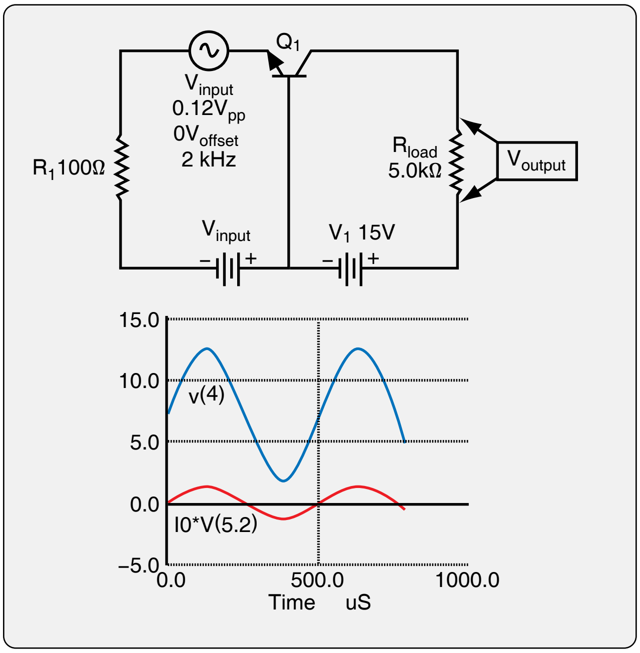


Figure 11-53. *In a common-base amplification circuit for AC (top), output voltage amplitude is greatly increased in phase with the input signal (bottom).*

An oscillator known as a relaxation oscillator produces another kind of wave form, a sawtooth wave. A slow rise from zero to peak voltage is followed by a rapid drop-off of voltage back to almost zero. Then it repeats. *[Figure 11-57]* In the circuit, a capacitor slowly charges through a resistor. A neon bulb is wired across the capacitor. When its ignition voltage is reached, the bulb conducts. This short-circuits the charged capacitor, which causes the voltage to drop to nearly zero and the bulb goes out. Then, the voltage rises again as the cycle repeats.

Electronic Oscillation

Oscillation in electronic circuits is accomplished by combining a transistor and a tank circuit. A tank circuit is comprised of a capacitor and coil parallel to each other. *[Figure 11-58]* When attached to a power source by closing switch A, the capacitor charges to a voltage equal to the battery voltage. It stays charged, even when the circuit to the battery is open (switch in position B). When the switch is put in position C, the capacitor and coil are in a closed circuit. The capacitor discharges through the coil. While receiving the energy from the capacitor, the coil stores it by building up an electromagnetic field. When the capacitor is fully discharged, the coil stops conducting. The magnetic field collapses, which induces current flow. The current charges the opposite plate of the capacitor. When completely charged, the capacitor discharges into the coil again. The magnetic field builds again and stops when the capacitor is fully discharged. The magnetic field collapses again, which induces current that charges the original plate of the capacitor and the cycle repeats.

This oscillation of charging and discharging the capacitor through the coil would continue indefinitely if a circuit could be built with no resistance. This is not possible. However, a circuit can be built using a transistor that restores losses due to resistance. There are various ways to accomplish this. The Hartley oscillator circuit in *Figure 11-59* is one. The circuit can oscillate indefinitely as long as it is connected to power.

When the switch is closed, current begins to flow in the oscillator circuit. The transistor base is supplied with biasing current through the voltage divider R_A and R_B. This allows current to flow through the transistor from the collector to the emitter, through R_E and through the lower portion of the center tapped coil that is labeled L_1. The current increasing through this coil builds a magnetic field that induces current in the upper half of the coil labeled L_2. The current from L_2 charges capacitor C_2, which increases the forward bias of the transistor. This allows an increasing flow of current through the transistor, R_E, and L_1 until the transistor is saturated and capacitor C_1 is fully charged. Without force to add electrons to capacitor C_1, it discharges and begins

Type of Amplifier	Impedance	Voltage Gain	Current Gain	Power Gain	Phase
Common-emitter	Input: fairly high Output: fairly high	Relatively large	Relatively large	Large	Inverts phase
Common-collector	Input: high Output: low	Always less than one	Relatively large	Relatively large	Output same as input
Common-base	Input: low Output: high	Large	Always less than one	Relatively large	Output same as input

Figure 11-54. *PN junction transistor amplifier characteristics.*

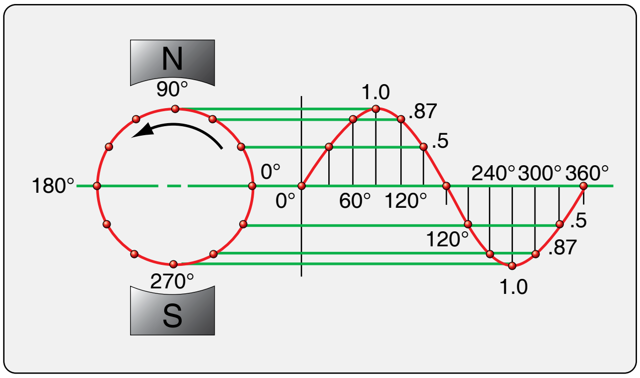


Figure 11-55. *Voltage over time of sine waveform electricity created when a conductor is rotated through a uniform magnetic field, such as in a generator.*

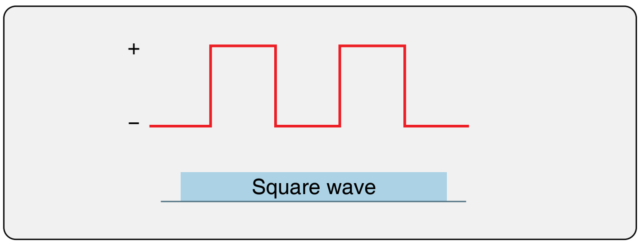


Figure 11-56. *The waveform of pulsing DC is a square wave.*

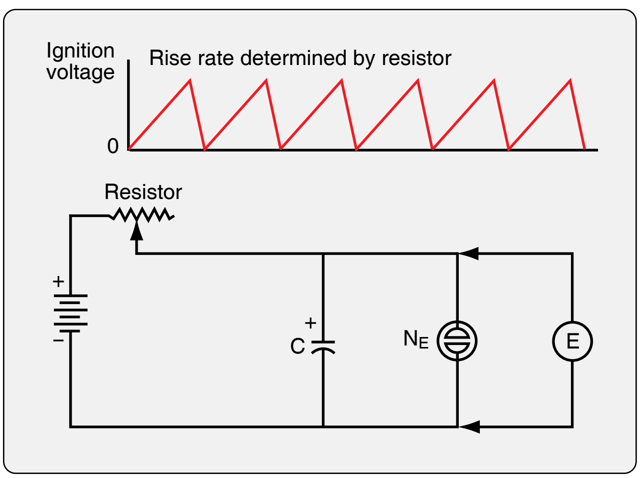


Figure 11-57. *A relaxation oscillator produces a sawtooth wave output.*

the oscillation in the tank circuit described in the previous section. As C_1 becomes fully charged, current to charge C_2 reduces and C_2 also discharges. This adds the energy needed to the tank circuit to compensate for resistance losses. As C_2 is discharging, it reduces forward biasing and eventually the transistor becomes reversed biased and cuts off. When the opposite plate of capacitor C_1 is fully charged, it discharges, and the oscillation is in progress. The transistor base becomes forward biased again, allowing for current flow through the resistor R_E, coil L_1, etc.

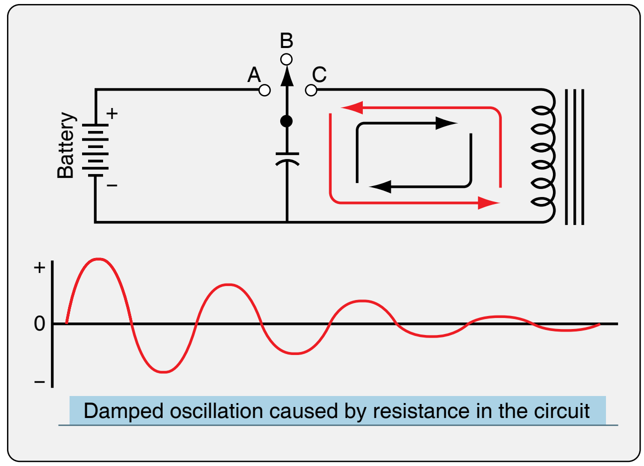


Figure 11-58. *A tank circuit alternately charges opposite plates of a capacitor through a coil in a closed circuit. The oscillation is an alternating current that diminishes due to resistance in the circuit.*

The frequency of the AC oscillating in the Hartley oscillator circuit depends on the inductance and capacitance values of the components used. Use of a crystal in an oscillator circuit can control the frequency more accurately. A crystal vibrates at a single, consistent frequency. When flexed, a small pulse of current is produced through the piezoelectric effect. Placed in the feedback loop, the pulses from the crystal control the frequency of the oscillator circuit. The tank circuit component values are tuned to match the frequency of the crystal. Oscillation is maintained as long as power is supplied. *[Figure 11-60]*

Other types of oscillator circuits used in electronics and computers have two transistors that alternate being in the active mode. They are called multi-vibrators. The choice of oscillator in an electronic device depends on the exact type of manipulation of electricity required to permit the device to function as desired.

Digital Electronics

The above discussion of semiconductors, semiconductor devices, and circuitry is only an introduction to the electronics found in communications and navigation avionics. In-depth maintenance of the interior electronics on most avionics devices is performed only by certified repair stations and trained avionics technicians. The airframe technician is responsible for installation, maintenance, inspection, and proper performance of avionics in the aircraft.

Modern aircraft increasingly employs digital electronics in avionics rather than analog electronics. Transistors are used in digital electronics to construct circuits that act as digital logic gates. The purpose and task of a device is achieved by manipulating electric signals through the logic gates. Thousands, and even millions, of tiny transistors can be

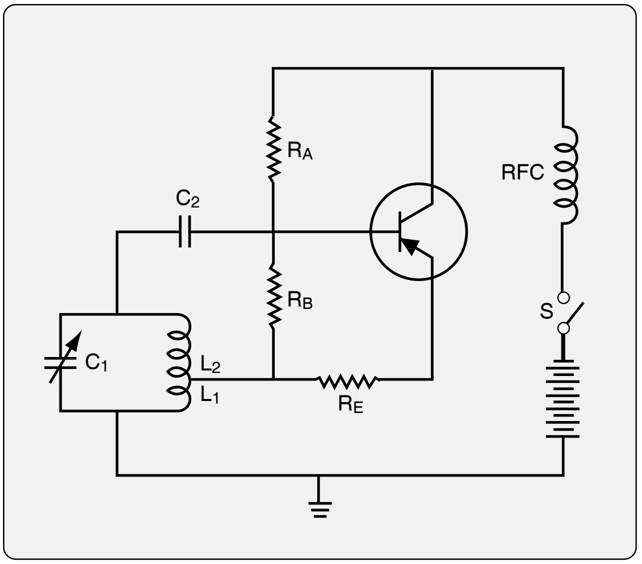


Figure 11-59. *A Hartley oscillator uses a tank circuit and a transistor to maintain oscillation whenever power is applied.*

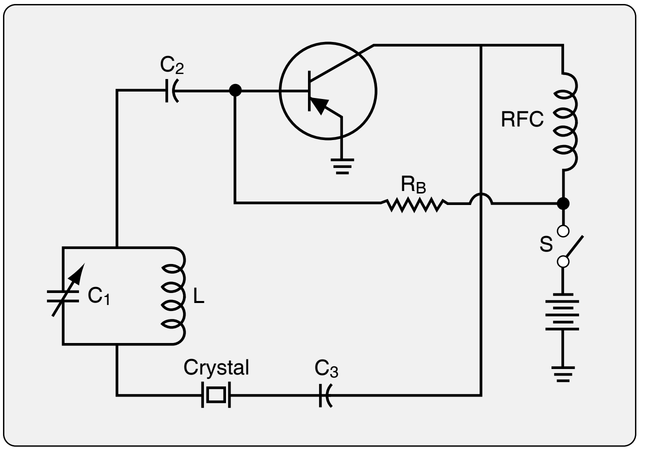


Figure 11-60. *A crystal in an electronic oscillator circuit is used to tune the frequency of oscillation.*

placed on a chip to create the digital logic landscape through which a component's signals are processed.

Digital Building Blocks

Digital logic is based on the binary number system. There are two conditions that may exist, 1 or 0. In a digital circuit, these are equivalent to voltage or no voltage. Within the binary system, these two conditions are called Logic 1 and Logic 0. Using just these two conditions, gates can be constructed to manipulate information. There are a handful of common logic gates that are used. By combining any number of these tiny solid-state gates, significant memorization, manipulation, and calculation of information can be performed.

The NOT Gate

The NOT gate is the simplest of all gates. If the input to the gate is Logic 1, then the output is NOT Logic 1. This means that it is Logic 0, since there are only two conditions in the binary world. In an electronic circuit, a NOT gate would invert the input signal. In other words, if there was voltage at the input to the gate, there would be no output voltage. The gate can be constructed with transistors and resistors to yield this electrical logic every time. (The gate or circuit would also have to invert an input of Logic 0 into an output of Logic 1.)

To understand logic gates, truth tables are often used. A truth table gives all of the possibilities in binary terms for each gate containing a characteristic logic function. For example, a truth table for a NOT gate is illustrated in *Figure 11-61*. Any input (A) is NOT present at the output (B). This is simple, but it defines this logic situation. A tiny NOT gate circuit can be built using transistors that produce these results. In other words, a circuit can be built such that if voltage arrives at the gate, no voltage is output or vice-versa.

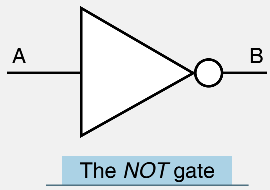
When using transistors to build logic gates, the primary concern is to operate them within the circuits so the transistors are either OFF (not conducting) or fully ON (saturated). In this manner, reliable logic functions can be performed. The variable voltage and current situations present during the active mode of the transistor are of less importance.

Figure 11-62 illustrates an electronic circuit diagram that performs the logic NOT gate function. Any input, either a no voltage or voltage condition, yields the opposite output. This gate is built with bipolar junction transistors, resistors, and a few diodes. Other designs exist that may have different components.

When examining and discussing digital electronic circuits, the electronic circuit design of a gate is usually not presented. The symbol for the logic gate is most often used. *[Figure 11-61]* The technician can then concentrate on the configuration of the logic gates in relation to each other. A brief discussion of the other logic gates, their symbols, and truth tables follow.

Buffer Gate

Another logic gate with only one input and one output is the buffer. It is a gate with the same output as the input. While this may seem redundant or useless, an amplifier may be considered a buffer in a digital circuit because if there is voltage present at the input, there is an output voltage. If there is no voltage at the input, there is no output voltage. When used as an amplifier, the buffer can change the values of a signal. This is often done to stabilize a weak or varying signal. All gates are amplifiers subject to output fluctuations. The buffer steadies the output of the upstream device while maintaining its basic characteristic. Another application of a buffer that is two NOT gates, is to use it to isolate a portion of a circuit. *[Figure 11-63]*



A	B
1	0
0	1

Figure 11-61. *A NOT logic gate symbol and a NOT gate truth table.*

AND Gate

Most common logic gates have two inputs. Three or more inputs are possible on some gates. When considering the characteristics of any logic gate, an output of Logic 1 is sought and a condition for the inputs is stated or examined. For example, *Figure 11-64* illustrates an AND gate. For an AND gate to have a Logic 1 output, both inputs have to be Logic 1. In an actual electronic circuit, this means that for a voltage to be present at the output, the AND gate circuit has to receive voltage at both of its inputs. As pointed out, there are different arrangements of electronic components that yield this result. Whichever is used is summarized and presented as the AND gate symbol. The truth table in *Figure 11-64* illustrates that there is only one way to have an output of Logic 1 or voltage when using an AND gate.

OR Gate

Another useful and common logic gate is the OR gate. In an OR gate, to have an output of Logic 1 (voltage present), one of the inputs must be Logic 1. As seen in *Figure 11-65,* only one of the inputs needs to be Logic 1 for there to be an output of Logic 1. When both inputs are Logic 1, the OR gate has a Logic 1 output because it still meets the condition of one of the inputs being Logic 1.

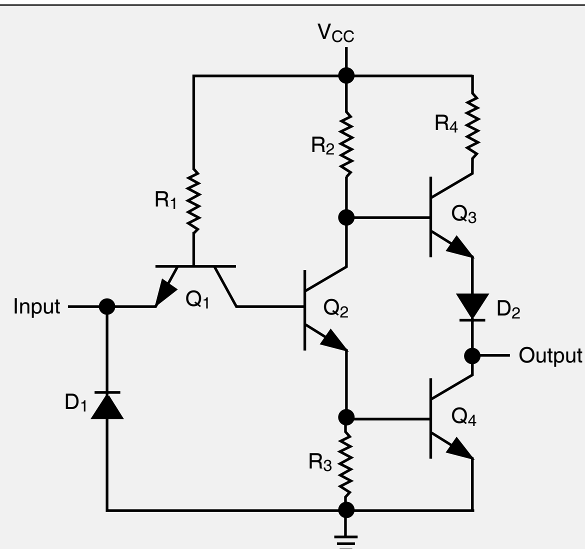
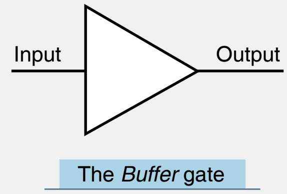


Figure 11-62. *An electronic circuit that reliably performs the NOT logic function.*



Input	Output
0	0
1	1

Figure 11-63. *A buffer or amplifier symbol and the truth table of the buffer, which is actually two consecutive NOT gates.*

NAND Gate

The AND, OR, and NOT gates are the basic logic gates. A few other logic gates are also useful. They can be derived from combining the AND, OR, and NOT gates. The NAND gate is a combination of an AND gate and a NOT gate. This means that AND gate conditions must be met and then inverted. So, the NAND gate is an AND gate followed by a NOT gate. The truth table for a NAND gate is shown in *Figure 11-66* along with its symbol. If a Logic 1 output is to exist from a NAND gate, inputs A and B must not both be Logic 1. Or, if a NAND gate has both inputs Logic 1, the output is Logic 0. Stated in electronic terms, if there is to be an output voltage, then the inputs cannot both have voltage or, if both inputs have voltage, there is no output voltage.

Note: The values in the output column of the NAND gate table are exactly the opposite of the output values in the AND gate truth table.

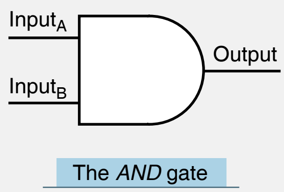
NOR Gate

A NOR gate is similarly arranged except that it is an inverted OR gate. If there is to be a Logic 1 output, or output voltage, then neither input can be Logic 1 or have input voltage. This is the same as satisfying the OR gate conditions and then putting output through a NOT gate. The NOR gate truth table in *Figure 11-67* shows that the NOR gate output values are exactly the opposite of the OR gate output values.

The NAND gate and the NOR gate have a unique distinction. Each one can be the only gate used in circuitry to produce the same output as any of the other logic gates. While it may be inefficient, it is testimonial to the flexibility that designers have when working with logic gates, the NAND and NOR gates in particular.

EXCLUSIVE OR Gate

Another common logic gate is the EXCLUSIVE OR gate. It is the same as an OR gate except for the condition where both inputs are Logic 1. In an OR gate, there would be Logic



Input A	Input B	Output
0	0	0
0	1	0
1	0	0
1	1	1

Figure 11-64. *An AND gate symbol and its truth table.*

1 output when both inputs are Logic 1. This is not allowed in an EXCLUSIVE OR gate. When either of the inputs is Logic 1, the output is Logic 1. But, if both inputs are logic 1, the Logic 1 output is excluded or Logic 0. *[Figure 11-68]*

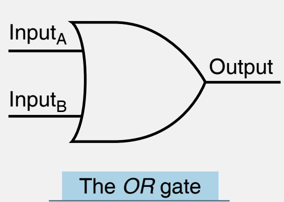
Negative Logic Gates

There are also negative logic gates. The negative OR and the negative AND gates are gates wherein the inputs are inverted rather than inverting the output. This creates a unique set of outputs as seen in the truth tables in *Figure 11-69*. The negative OR gate is not the same as the NOR gate as is sometimes misunderstood. Neither is the negative AND gate the same as the NAND gate. However, as the truth tables reveal, the output of a negative AND gate is the same as a NOR gate, and the output of a negative OR gate is the same as a NAND gate.

In summary, electronic circuits use transistors to construct logic gates that produce outputs related to the inputs shown in the truth tables for each kind of gate. The gates are then assembled with other components to manipulate data in digital circuits. The electronic digital signals used are voltage or no voltage representations of Logic 1 or Logic 0 conditions. By using a series of voltage output or no voltage output gates, manipulation, computation, and storage of data takes place.

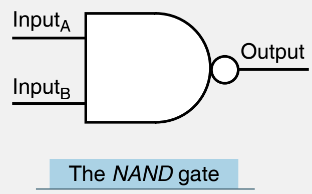
Digital Aircraft Systems

Digital aircraft systems are the present and future of aviation. From communication and navigation to engine and flight controls, increased proliferation of digital technology increases reliability and performance. Processing, storing, and transferring vital information for the operation of an aircraft in digital form provides a usable common language



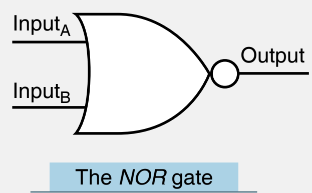
Input A	Input B	Output
0	0	0
0	1	1
1	0	1
1	1	1

Figure 11-65. *An OR gate symbol and its truth table.*



Input A	Input B	Output
0	0	1
0	1	1
1	0	1
1	1	0

Figure 11-66. *A NAND gate symbol and its truth table illustrating that the NAND gate is an inverted AND gate.*



Input A	Input B	Output
0	0	1
0	1	0
1	0	0
1	1	0

Figure 11-67. *A NOR gate symbol and its truth table illustrating that the NOR gate is an inverted OR gate.*

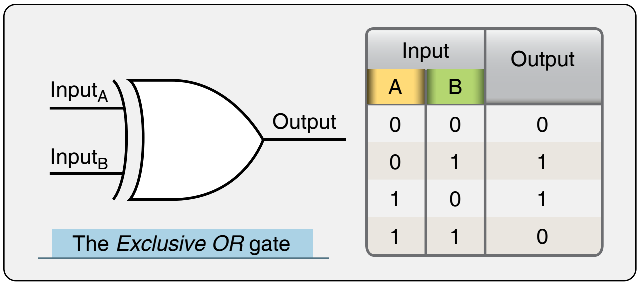
for monitoring, control, and safety. Integration of information from different systems is simplified. Self-monitoring, built-in test equipment (BITE) and air-to-ground data links increase maintenance efficiency. Digital buss networking allows aircraft system computers to interact for a coordinated comprehensive approach to flight operations.

Digital Data Displays

Modern digital data displays are the most visible features of digital aircraft systems. They extend the functional advantages of state of the art digital communication and navigation avionics and other digital aircraft systems via the use of an enhanced interface with the pilot. The result is an increase in situational awareness and overall safety of flight. Digital data displays are the glass of the glass flight deck. They expand the amount, clarity, and proximity of the information presented to the pilot. *[Figure 11-70]*

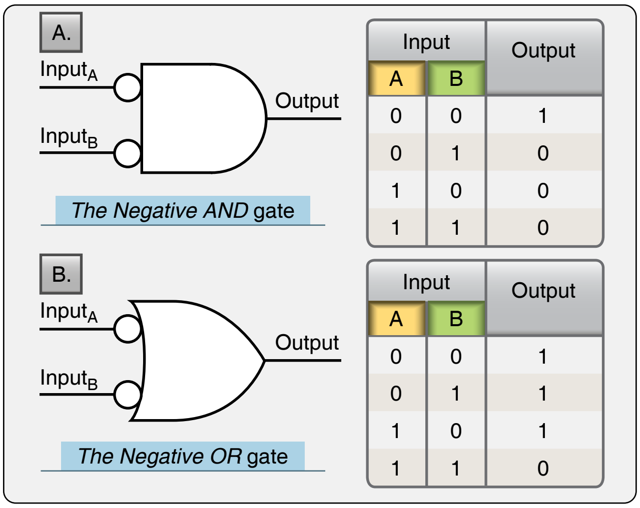
Many digital data displays are available from numerous manufacturers as original equipment in new aircraft, or as retrofit components or complete retrofit systems for older aircraft. Approval for retrofit displays is usually accomplished through supplementary type certificate (STC) awarded to the equipment manufacturer.

Early digital displays presented scale indication in digital or integer format readouts. Today's digital data displays are analogous to computer screen presentations. Numerous aircraft and flight instrument readouts and symbolic presentations are combined with communication and navigational information on multifunctional displays (MFD).



Input A	Input B	Output
0	0	0
0	1	1
1	0	1
1	1	0

Figure 11-68. *An EXCLUSIVE OR gate symbol and its truth table, which is similar to an OR gate but excludes output when both inputs are the same.*



A.

Input A	Input B	Output
0	0	1
0	1	0
1	0	0
1	1	0

B.

Input A	Input B	Output
0	0	1
0	1	1
1	0	1
1	1	0

Figure 11-69. *The NEGATIVE AND gate symbol and its truth table (A) and the NEGATIVE OR gate symbol and truth table (B). The inputs are inverted in the NEGATIVE gates.*

Often a display has a main function with potential to back-up another display should it fail. Names, such as primary flight display (PFD), secondary flight display, navigational display (ND), etc., are often used to describe a display by its primary use. The hardware composition of the displays is essentially the same. Avionics components and computers combine to provide the different information portrayed on the displays.

Controls on the instrument panel or on the display unit itself are used for selection. Some screens have limited display capability because they are not part of a totally integrated system; however, they are extremely powerful electronic units with wide capability. *[Figure 11-71]*

The basis of the information displayed on what is known as a PFD, is usually an electronic flight instrument system (EFIS) like representation of the aircraft attitude indicator in the upper half of the display, and an electronic horizontal situation indicator display on the lower half. Numerous ancillary readouts are integrated or surround the electronic attitude indicator and the horizontal situation indicator (HSI). On full glass flight deck PFDs, all of the basic T instrument indications are presented and much more, such as communication and navigation information, weather data, terrain features, and approach information. Data displays for engine parameters, hydraulics, fuel, and other airframe systems are often displayed on the secondary flight display or on an independent display made for this purpose. *[Figure 11-72]*

As with other avionics components, repair and maintenance of the internal components of digital data displays is reserved for licensed repair stations only.

Digital Tuners & Audio Panels

Numerous communication and navigation devices are described in the following sections of this chapter. Many of these use radio waves and must be tuned to a desired frequency for operation. As a flight progresses, retuning and changing from one piece of equipment to another can occur frequently. An audio panel or digital tuner consolidates various communication and navigation radio selection controls into a single unit. The pilot can select and use, or

Figure 11-70. *A modern glass flight deck on a general aviation aircraft. Digital data displays replace many older instruments and indicators of the past.*

Figure 11-71. *A retrofit digital data display.*

select and tune, most of the aircraft's avionics from this one control interface. *[Figure 11-73]*

Radio Communication

Much of aviation communication and navigation is accomplished through the use of radio waves. Communication by radio was the first use of radio frequency transmissions in aviation.

Radio Waves

A radio wave is invisible to the human eye. It is electromagnetic in nature and part of the electronic spectrum of wave activity that includes gamma rays, x-rays, ultraviolet rays, infrared waves, and visible light rays, as well all radio waves. *[Figure 11-74]* The atmosphere is filled with these waves. Each wave occurs at a specific frequency and has a corresponding wavelength. The relationship between frequency and wavelength is inversely proportional. A high frequency wave has a short wavelength and a low frequency wave has a long wavelength.

In aviation, a variety of radio waves are used for communication. *Figure 11-75* illustrates the radio spectrum that includes the range of common aviation radio frequencies and their applications.

Note: A wide range of frequencies are used from low frequency (LF) at 100 kHz (100,000 cycles per second) to super high frequency (SHF) at nearly 10gHz (10,000,000,000 cycles per second). The Federal Communications Commission (FCC) controls the assignment of frequency usage.

AC power of a particular frequency has a characteristic length of conductor that is resonant at that frequency. This length is the wavelength of the frequency that can be seen on an oscilloscope. Fractions of the wavelength also resonate, especially half of a wavelength, which is the same as half of the AC sine wave or cycle.

The frequency of an AC signal is the number of times the AC cycles every second. AC applied to the center of a radio antenna, a conductor half the wavelength of the AC frequency, travels the length of the antenna, collapses, and travels the

Figure 11-72. *A digital data display dedicated to the depiction of engine and airframe system parameter status.*

length of the antenna in the opposite direction. The number of times it does this every second is known as the radio wave signal frequency or radio frequency as shown in *Figure 11-75*. As the current flows through the antenna, corresponding electromagnetic and electric fields build, collapse, build in the opposite direction, and collapse again. *[Figure 11-76]*

To transmit radio waves, an AC generator is placed at the midpoint of an antenna. As AC current builds and collapses in the antenna, a magnetic field also builds and collapses around it. An electric field also builds and subsides as the voltage shifts from one end of the antenna to the other. Both fields, the magnetic and the electric, fluctuate around the antenna at the same time. The antenna is half the wavelength of the AC signal received from the generator. At any one point along the antenna, voltage and current vary inversely to each other.

Because of the speed of the AC, the electromagnetic fields and electric fields created around the antenna do not have time to completely collapse as the AC cycles. Each new current flow creates new fields around the antenna that force the not-totally-collapsed fields from the previous AC cycle out into space. These are the radio waves. The process is continuous as long as AC is applied to the antenna. Thus, steady radio waves of a frequency determined by the input AC frequency propagate out into space.

Radio waves are directional and propagate out into space at 186,000 miles per second. The distance they travel depends on the frequency and the amplification of the signal AC sent to the antenna. The electric field component and the electromagnetic field component are oriented at 90° to each other, and at 90° to the direction that the wave is traveling. [*Figure 11-77*]

Types of Radio Waves

Radio waves of different frequencies have unique characteristics as they propagate through the atmosphere. Very low frequency (VLF), LF, and medium frequency (MF) waves have relatively long wavelengths and utilize correspondingly long antennas. Radio waves produced at these frequencies ranging from 3 kHz to 3 mHz are known as ground waves or surface waves. This is because they follow the curvature of the earth as they travel from the broadcast antenna to the receiving antenna. Ground waves are particularly useful for long distance transmissions. Automatic direction finders (ADF) and LORAN navigational aids use these frequencies. *[Figure 11-78]*

High frequency (HF) radio waves travel in a straight line and do not curve to follow the earth's surface. This would limit transmissions from the broadcast antenna to receiving antennas only in the line-of-sight of the broadcast antenna except for a unique characteristic. HF radio waves bounce off of the ionosphere layer of the atmosphere. This refraction

Figure 11-73. *An audio panel in a general aviation aircraft integrates the selection of several radio-based communication and navigational aids into a single control panel (left). A digital tuner (right) does the same on a business class aircraft and allows the frequency of each device to be tuned from the same panel as well.*

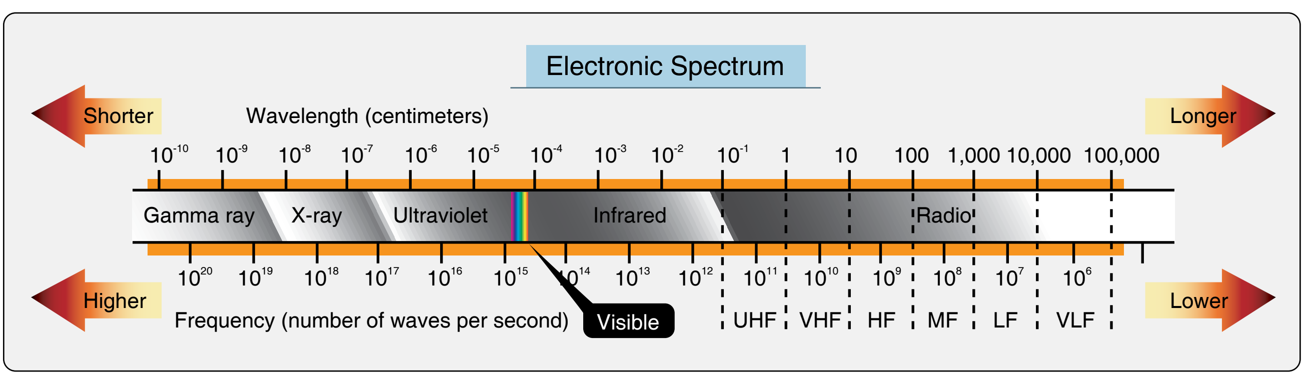


Figure 11-74. *Radio waves are just some of the electromagnetic waves found in space.*

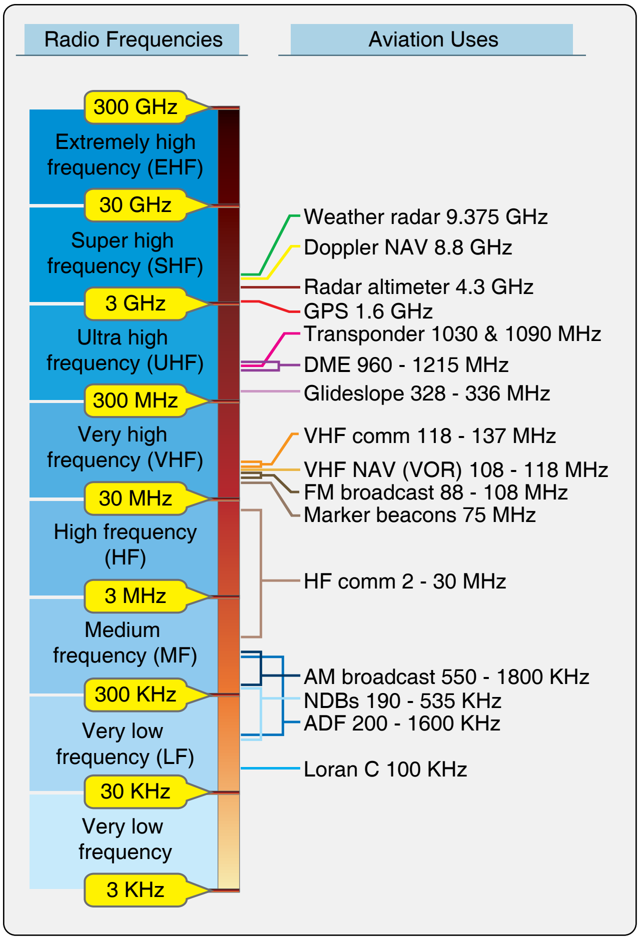


Figure 11-75. *There is a wide range of radio frequencies. Only the very low frequencies and the extremely high frequencies are not used in aviation.*

extends the range of HF signals beyond line-of-sight. As a result, transoceanic aircraft often use HF radios for voice communication. The frequency range is between 2 to 25 MHz. These kinds of radio waves are known as sky waves. *[Figure 11-78]*

Above HF transmissions, radio waves are known as space waves. They are only capable of line-of-sight transmission and do not refract off of the ionosphere. *[Figure 11-78]* Most aviation communication and navigational aids operate with space waves. This includes VHF (30 - 300 MHz), UHF (300 MHz - 3 GHz), and super high frequency (SHF) (3 - 30 Ghz) radio waves.

VHF communication radios are the primary communication radios used in aviation. They operate in the frequency range from 118.0 MHz to 136.975 MHz. Seven hundred and twenty separate and distinct channels have been designated in this range with 25 kilohertz spacing between each channel. Further division of the bandwidth is possible, such as in Europe where 8.33 kilohertz separate each VHF communication channel. VHF radios are used for communications between aircraft and air traffic control (ATC), as well as air-to-air communication between aircraft. When using VHF, each party transmits and receives on the same channel. Only one party can transmit at any one time.

Loading Information onto a Radio Wave

The production and broadcast of radio waves does not convey any significant information. The basic radio wave discussed above is known as a carrier wave. To transmit and receive useful information, this wave is altered or modulated by an information signal. The information signal contains the unique voice or data information desired to be conveyed. The modulated carrier wave then carries the information from the transmitting radio to the receiving radio via their respective antennas. Two common methods of modulating carrier waves are amplitude modulation and frequency modulation.

Amplitude Modulation (AM)

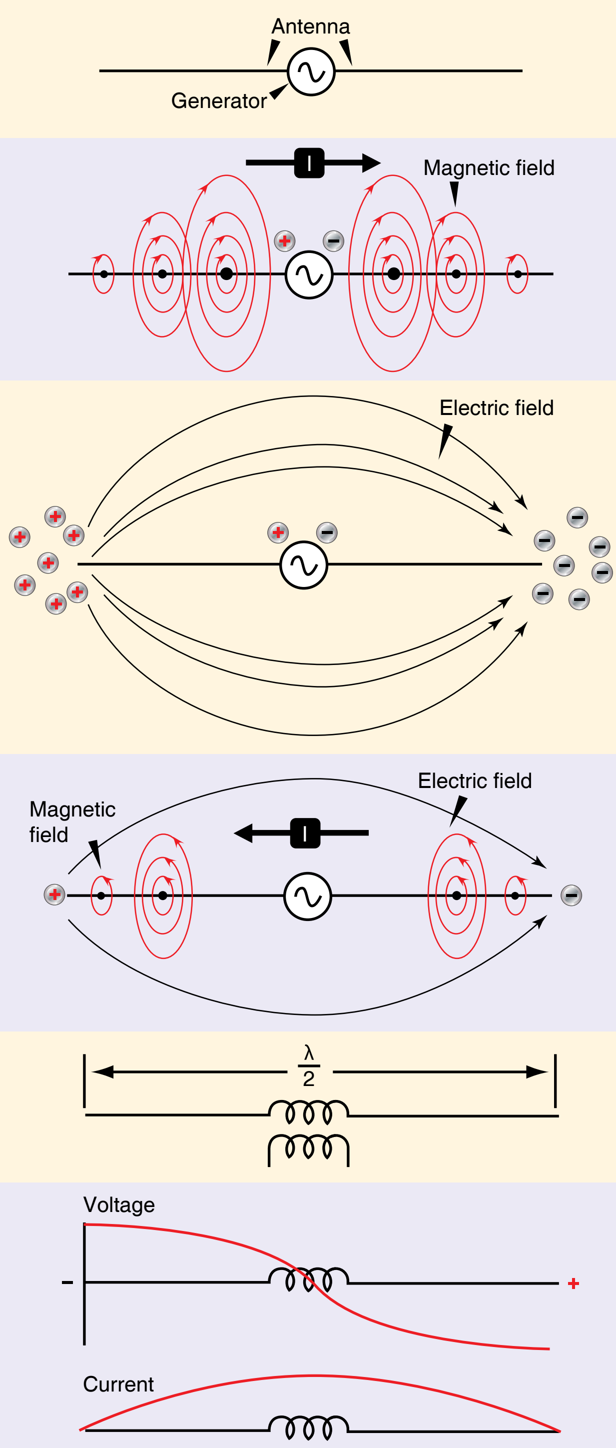
A radio wave can be altered to carry useful information by modulating the amplitude of the wave. A DC signal, for example from a microphone, is amplified and then superimposed over the AC carrier wave signal. As the varying DC information signal is amplified, the amplifier output current varies proportionally. The oscillator that creates the carrier wave does so with this varying current. The oscillator frequency output is consistent because it is built into the oscillator circuit. But the amplitude of the oscillator output varies in relation to the fluctuating current input. *[Figure 11-79]*

When the modulated carrier wave strikes the receiving antenna, voltage is generated that is the same as that which was applied to the transmitter antenna. However, the signal is weaker. It is amplified so that it can be demodulated. Demodulation is the process of removing the original information signal from the carrier wave. Electronic circuits containing capacitors, inductors, diodes, filters, etc., remove all but the desired information signal identical to the original input signal. Then, the information signal is typically amplified again to drive speakers or other output devices. *[Figure 11-80]*

AM has limited fidelity. Atmospheric noises or static alter the amplitude of a carrier wave making it difficult to separate the intended amplitude modulation caused by the information signal and that which is caused by static. It is used in aircraft VHF communication radios.

Frequency Modulation (FM)

Frequency modulation (FM) is widely considered superior to AM for carrying and deciphering information on radio waves. A carrier wave modulated by FM retains its constant amplitude. However, the information signal alters the



To transmit radio waves, an AC generator is placed at the midpoint of an antenna.

As AC current builds and collapses in the antenna, a magnetic field also builds and collapses around it.

An electric field also builds and subsides as the voltage shifts from one end of the antenna to the other.

Both fields, the magnetic and the electric, fluctuate around the antenna at the same time.

The antenna is ½ the wavelength of the AC signal received from the generator.

At any one point along the antenna, voltage and current vary inversely to each other.

Figure 11-76. *Radio waves are produced by applying an AC signal to an antenna. This creates a magnetic and electric field around the antenna. They build and collapse as the AC cycles. The speed at which the AC cycles does not allow the fields to completely collapse before the next fields build. The collapsing fields are then forced out into space as radio waves.*

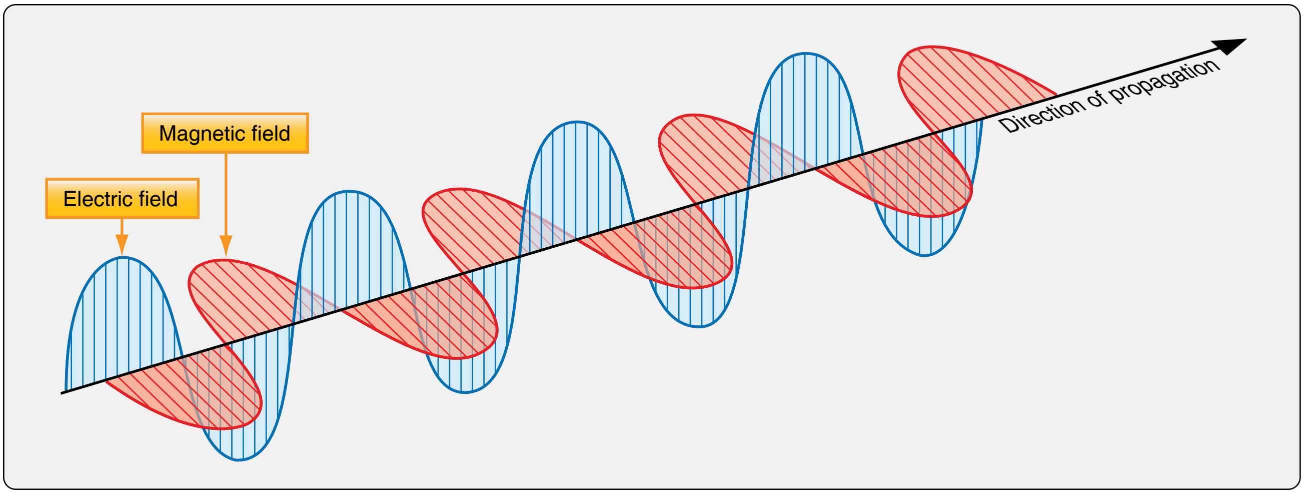


Figure 11-77. *The electric field and the magnetic field of a radio wave are perpendicular to each other and to the direction of propagation of the wave.*

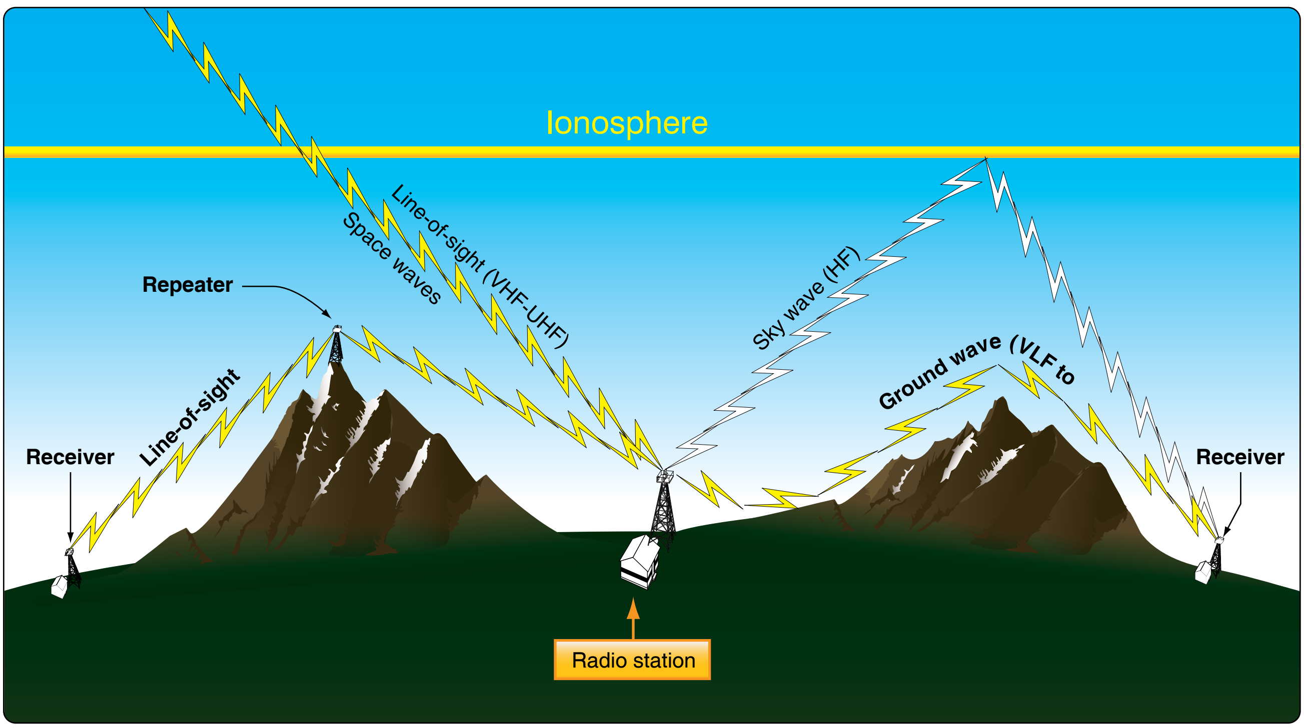


Figure 11-78. *Radio waves behave differently in the atmosphere depending in their frequency.*

frequency of the carrier wave in proportion to the strength of the signal. Thus, the signal is represented as slight variations to the normally consistent timing of the oscillations of the carrier wave. *[Figure 11-81]*

Since the transmitter oscillator output fluctuates during modulation to represent the information signal, FM bandwidth is greater than AM bandwidth. This is overshadowed by the ease with which noise and static can be removed from the FM signal. FM has a steady current flow and requires less power to produce since modulating an oscillator producing a carrier wave takes less power than modulating the amplitude of a signal using an amplifier.

Demodulation of an FM signal is similar to that of an AM receiver. The signal captured by the receiving antenna is usually amplified immediately since signal strength is lost as the wave travels through the atmosphere. Numerous circuits are used to isolate, stabilize, and remove the information from the carrier wave. The result is then amplified to drive the output device.

Single Sideband (SSB)

When two AC signals are mixed together, such as when a carrier wave is modulated by an information signal, three main frequencies result:

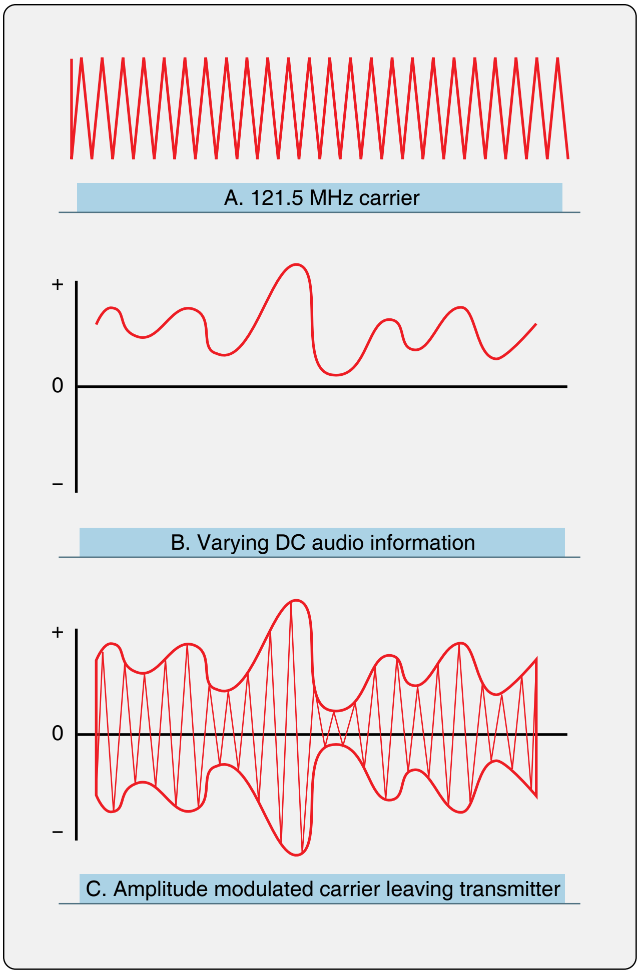


Figure 11-79. *A DC audio signal modifies the 121.5 MHz carrier wave as shown in C. The amplitude of the carrier wave (A) is changed in relation to modifier (B). This is known as amplitude modulation (AM).*

1. Original carrier wave frequency;
2. Carrier wave frequency plus the modulating frequency; and
3. Carrier wave frequency minus the modulating frequency.

Due to the fluctuating nature of the information signal, the modulating frequency varies from the carrier wave up or down to the maximum amplitude of the modulating frequency during AM. These additional frequencies on either side of the carrier wave frequency are known as sidebands. Each sideband contains the unique information signal desired to be conveyed. The entire range of the lower and upper sidebands including the center carrier wave frequency is known as bandwidth. *[Figure 11-82]*

There are a limited number of frequencies within the usable frequency ranges (i.e., LF, HF, and VHF). If different

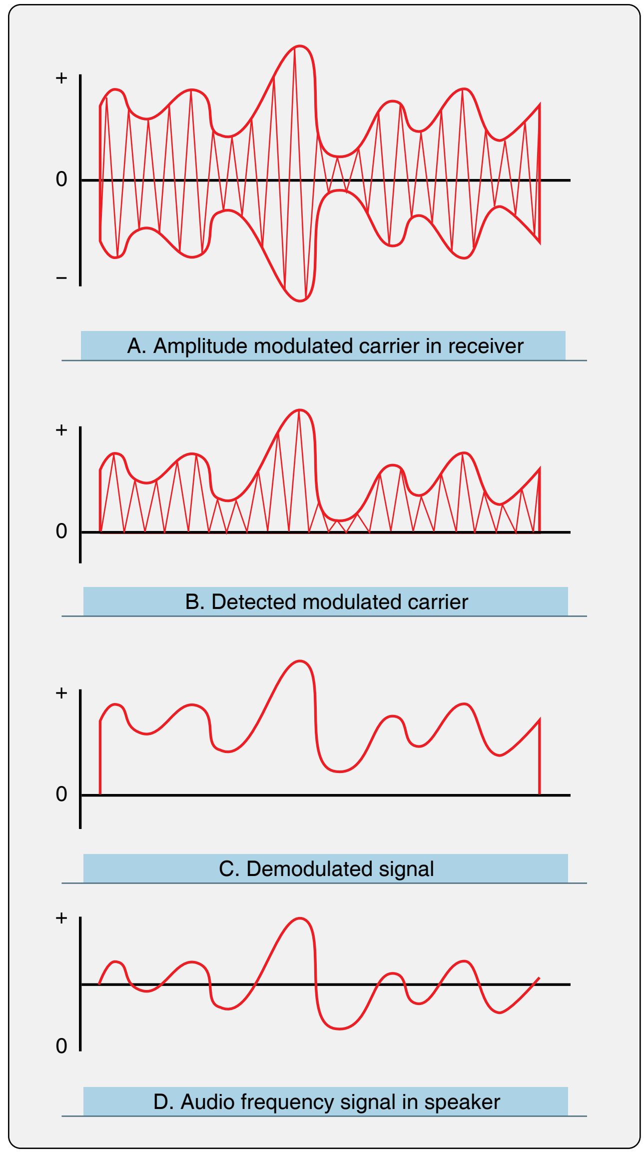


Figure 11-80. *Demodulation of a received radio signal involves separating the carrier wave from the information signal.*

broadcasts are made on frequencies that are too close together, some of the broadcast from one frequency interferes with the adjacent broadcast due to overlapping sidebands. The FCC divides the various frequency bands and issues rules for their use. Much of this allocation is to prevent interference. The spacing between broadcast frequencies is established so that a carrier wave can expand to include the upper and lower sidebands and still not interfere with a signal on an adjacent frequency.

As use of the radio frequencies increases, more efficient allocation of bandwidth is imperative. Sending information via radio waves using the narrowest bandwidth possible is the focus of engineering moving forward. At the same time, fully representing all of the desired information or increasing

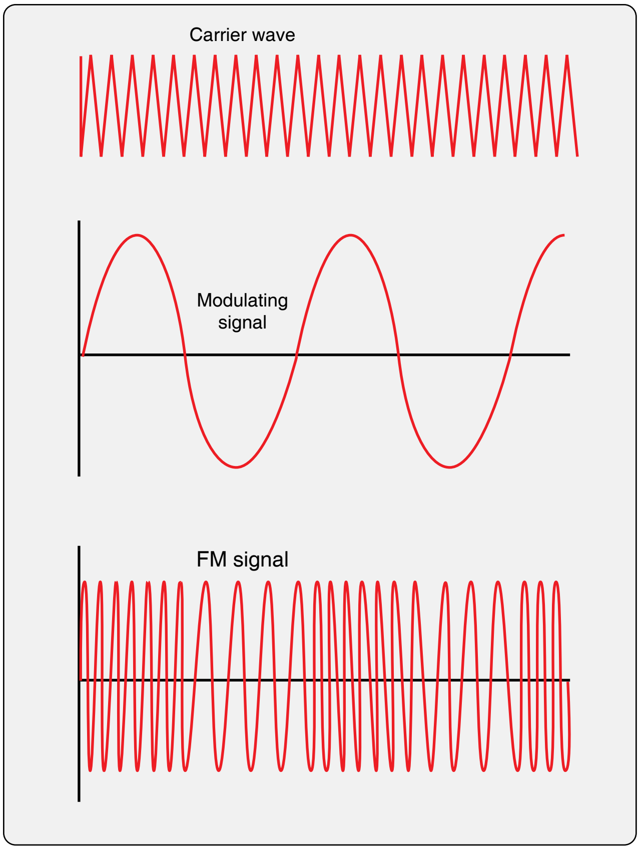


Figure 11-81. *A frequency modulated (FM) carrier wave retains the consistent amplitude of the AC sine wave. It encodes the unique information signal with slight variations to the frequency of the carrier wave. These variations are shown as space variations between the peaks and valleys of the wave on an oscilloscope.*

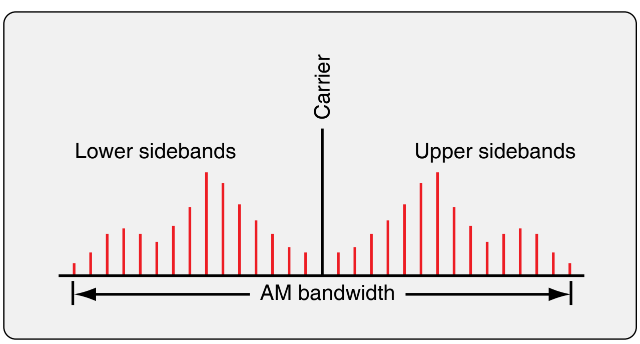


Figure 11-82. *The bandwidth of an AM signal contains the carrier wave, the carrier wave plus the information signal frequencies, and the carrier wave minus the information signal frequencies.*

the amount of information conveyed is also desired. Various methods are employed to keep bandwidth to a minimum, many of which restrict the quality or quantity of information able to be transmitted.

In lower frequency ranges, such as those used for ground wave and some sky wave broadcasts, SSB transmissions are a narrow bandwidth solution. Each sideband represents the initial information signal in its entirety. Therefore, in an SSB broadcast, the carrier wave and either the upper or lower sidebands are filtered out. Only one sideband with its frequencies is broadcast since it contains all of the needed information. This cuts the bandwidth required in half and allows more efficient use of the radio spectrum. SSB transmissions also use less power to transmit the same amount of information over an equal distance. Many HF long-distance aviation communications are SSB. *[Figure 11-83]*

Radio Transmitters & Receivers

Radio transmitters and receivers are electronic devices that manipulate electricity resulting in the transmission of useful information through the atmosphere or space.

Transmitters

A transmitter consists of a precise oscillating circuit or oscillator that creates an AC carrier wave frequency. This is combined with amplification circuits or amplifiers. The distance a carrier wave travels is directly related to the amplification of the signal sent to the antenna.

Other circuits are used in a transmitter to accept the input information signal and process it for loading onto the carrier wave. Modulator circuits modify the carrier wave with the processed information signal. Essentially, this is all there is to a radio transmitter.

Note: Modern transmitters are highly refined devices with extremely precise frequency oscillation and modulation. The circuitry for controlling, filtering, amplifying, modulating, and oscillating electronic signals can be complex.

A transmitter prepares and sends signals to an antenna that, in the process described above, radiates the waves out into the atmosphere. A transmitter with multiple channel (frequency) capability contains tuning circuitry that enables the user to select the frequency upon which to broadcast. This adjusts

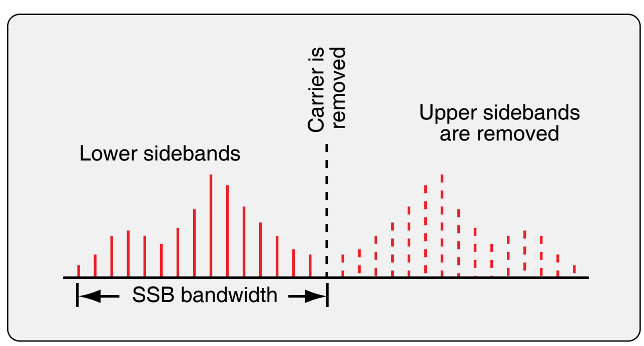


Figure 11-83. *The additional frequencies above and below the carrier wave produced during modulation with the information signal are known as sidebands. Each sideband contains the unique information of the information signal and can be transmitted independent of the carrier wave and the other sideband.*

the oscillator output to the precise frequency desired. It is the oscillator frequency that is being tuned. *[Figure 11-84]*

As shown in *Figure 11-84*, most radio transmitters generate a stable oscillating frequency and then use a frequency multiplier to raise the AC to the transmitting frequency. This allows oscillation to occur at frequencies that are controllable and within the physical working limits of the crystal in crystal-controlled oscillators.

Receivers

Antennas are simply conductors of lengths proportional to the wavelength of the oscillated frequency put out by the transmitter. An antenna captures the desired carrier wave as well as many other radio waves that are present in the atmosphere. A receiver is needed to isolate the desired carrier wave with its information. The receiver also has circuitry to separate the information signal from the carrier wave. It prepares it for output to a device, such as speakers or a display screen. The output is the information signal originally introduced into the transmitter.

A common receiver is the super heterodyne receiver. As with any receiver, it must amplify the desired radio frequency captured by the antenna since it is weak from traveling through the atmosphere. An oscillator in the receiver is used to compare and select the desired frequency out of all of the frequencies picked up by the antenna. The undesired frequencies are sent to ground.

A local oscillator in the receiver produces a frequency that is different than the radio frequency of the carrier wave. These two frequencies are mixed in the mixer. Four frequencies result from this mixing. They are the radio frequency, the local oscillator frequency, and the sum and difference of these two frequencies. The sum and difference frequencies contain the information signal.

The frequency that is the difference between the local oscillator frequency and the radio frequency carrier wave frequency is used during the remaining processing. In VHF aircraft communication radios, this frequency is 10.8 MHz. Called the intermediate frequency, it is amplified before it is sent to the detector. The detector, or demodulator, is where the information signal is separated from the carrier wave portion of the signal. In AM, since both sidebands contain the useful information, the signal is rectified leaving just one sideband with a weak version of the original transmitter input signal. In FM receivers, the varying frequency is changed to a varying amplitude signal at this point. Finally, amplification occurs for the output device. *[Figure 11-85]*

Over the years, with the development of transistors, micro-transistors, and integrated circuits, radio transmitters and receivers have become smaller. Electronic bays were established on older aircraft as remote locations to mount radio devices simply because they would not fit in the flight deck. Today, many avionics devices are small enough to be mounted in the instrument panel, which is customary on most light aircraft. Because of the number of communication and navigation aids, as well as the need to present an uncluttered interface to the pilot, most complicated aircraft retain an area away from the flight deck for the mounting of avionics. The control heads of these units remain on the flight deck.

Transceivers

A transceiver is a communication radio that transmits and receives. The same frequency is used for both. When transmitting, the receiver does not function. The push to talk (PTT) switch blocks the receiving circuitry and allows the transmitter circuitry to be active. In a transceiver, some of the circuitry is shared by the transmitting and receiving functions of the device. So is the antenna. This saves space and the number of components used. Transceivers are half duplex systems where communication can occur in both directions but only one party can speak while the other must listen. VHF aircraft communication radios are usually transceivers. *[Figure 11-86]*

Antennas

As stated, antennas are conductors that are used to transmit and receive radio frequency waves. Although the airframe technician has limited duties in relation to maintaining and repairing avionics, it is the responsibility of the technician to install, inspect, repair, and maintain aircraft radio antennas.

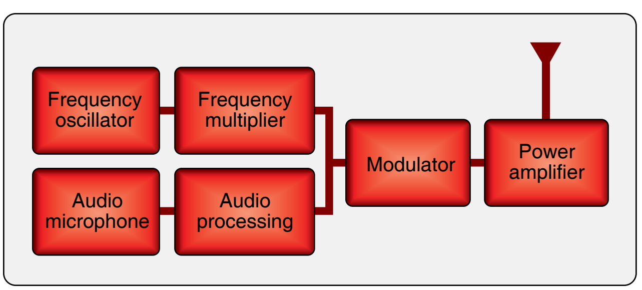


Figure 11-84. *Block diagram of a basic radio transmitter.*

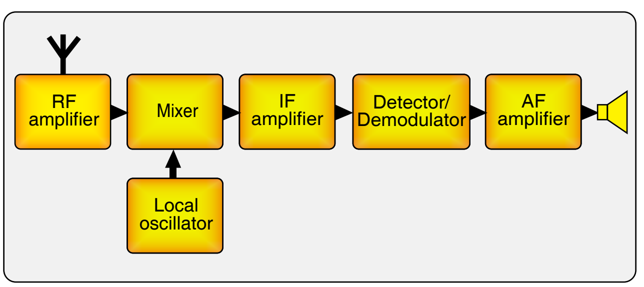


Figure 11-85. *The basic stages used in a receiver to produce an output from a radio wave.*

Three characteristics are of major concern when considering antennas:

1. Length.
2. Polarization.
3. Directivity.

The exact shape and material from which an antenna is made can alter its transmitting and receiving characteristics. Also note that some non-metallic aircraft have antennas imbedded into the composite material as it is built up.

Length

When an AC signal is applied to an antenna, it has a certain frequency. There is a corresponding wavelength for that frequency. An antenna that is half the length of this wavelength is resonant. During each phase of the applied AC, all voltage and current values experience the full range of their variability. As a result, an antenna that is half the wavelength of the corresponding AC frequency is able to allow full voltage and full current flow for the positive phase of the AC signal in one direction. The negative phase of the full AC sine wave is accommodated by the voltage and current simply changing direction in the conductor. Thus, the applied AC frequency flows through its entire wavelength, first in one direction and then in the other. This produces the strongest signal to be radiated by the transmitting antenna. It also facilitates capture of the wave and maximum induced voltage in the receiving antenna. *[Figure 11-87]*

Most radios, especially communication radios, use the same antenna for transmitting and receiving. Multichannel radios could use a different length antenna for each frequency, however, this is impractical. Acceptable performance can exist from a single antenna half the wavelength of a median frequency. This antenna can be made effectively shorter by placing a properly rated capacitor in series with the transmission line from the transmitter or receiver. This electrically shortens the resonant circuit of which the antenna is a part. An antenna may be electrically lengthened by adding an inductor in the circuit. Adjusting antenna length in this fashion allows the use of a single antenna for multiple frequencies in a narrow frequency range.

Many radios use a tuning circuit to adjust the effective length of the antenna to match the wavelength of the desired frequency. It contains a variable capacitor and an inductor connected in parallel in a circuit. Newer radios use a more efficient tuning circuit. It uses switches to combine frequencies from crystal controlled circuits to create a resonant frequency that matches the desired frequency. Either way, the physical antenna length is a compromise when using a multichannel communication or navigation device that must be electronically tuned for the best performance.

A formula can be used to find the ideal length of a half wavelength antenna required for a particular frequency as

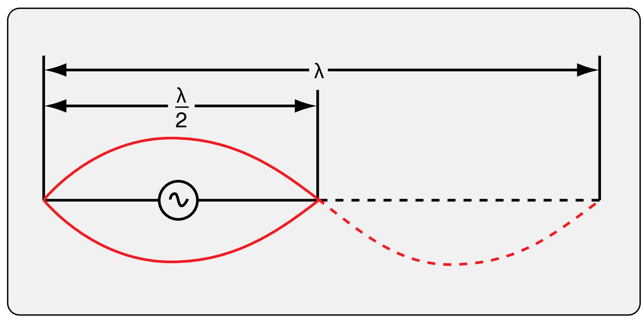


Figure 11-87. *An antenna equal to the full length of the applied AC frequency wavelength would have the negative cycle current flow along the antenna as shown by the dotted line. An antenna that is ½ wavelength allows current to reverse its direction in the antenna during the negative cycle. This results in low current at the ends of the ½ wavelength antenna and high current in the center. As energy radiates into space, the field is strongest 90° to the antenna where the current flow is strongest.*

Figure 11-86. *VHF aircraft communication transceivers.*

follows:

$$\text{Antenna Length (feet)} = \frac{468}{\text{F MHz}}$$

The formula is derived from the speed of propagation of radio waves, which is approximately 300 million meters per second. It takes into account the dielectric effect of the air at the end of an antenna that effectively shortens the length of the conductor required.

VHF radio frequencies used by aircraft communication radios are 118–136.975 MHz. The corresponding half wavelengths of these frequencies are 3.96 – 3.44 feet (47.5–41.2 inches). Therefore, VHF antennas are relatively long. Antennas one-quarter of the wavelength of the transmitted frequency are often used. This is possible because when mounted on a metal fuselage, a ground plane is formed and the fuselage acts as the missing one-quarter length of the half wavelength antenna. This is further discussed in the following antenna types section.

Polarization, Directivity, & Field Pattern

Antennas are polarized. They radiate and receive in certain patterns and directions. The electric field cause by the voltage in the conductor is parallel to the polarization of an antenna. It is caused by the voltage difference between each end of the antenna. The electromagnetic field component of the radio wave is at 90° to the polarization. It is caused by changing current flow in the antenna. These fields were illustrated in *Figure 11-76* and *11-77*. As radio waves radiate out from the antenna they propagate in a specific direction and in a specific pattern. This is the antenna field. The orientation of the electric and electromagnetic fields remains at 90° to each other but radiate from antenna with varying strength in different directions. The strength of the radiated field varies depending on the type of antenna and the angular proximity to it. All antennas, even those that are omnidirectional, radiate a stronger signal in some direction compared to other directions. This is known as the antenna field directivity.

Receiving antennas with the same polarization as the transmitting antenna generate the strongest signal. A vertically polarized antenna is mounted up and down. It radiates waves out from it in all directions. To receive the strongest signal from these waves, the receiving antenna should also be positioned vertically so the electromagnetic component of the radio wave can cross it at as close to a 90° angle as possible for most of the possible proximities. *[Figure 11-88]*

Horizontally polarized antennas are mounted side to side (horizontally). They radiate in a donut-like field. The strongest signals come from, or are received at, 90° to the length of the antenna. There is no field generated off of the end of the antenna. *Figure 11-89* illustrates the field produced by a horizontally polarized antenna.

Many vertical and horizontal antennas on aircraft are mounted at a slight angle off plane. This allows the antenna to receive a weak signal rather than no signal at all when the polarization of the receiving antenna is not identical to the transmitting antenna. *[Figure 11-90]*

Types

There are three basic types of antennas used in aviation:

1. Dipole antenna.
2. Marconi antenna.
3. Loop antenna.

Dipole Antenna

The dipole antenna is the type of antenna referred to in the discussion of how a radio wave is produced. It is a conductor, the length of which is approximately equal to half the wavelength of the transmission frequency. This sometimes is referred to as a Hertz antenna. The AC transmission current is fed to a dipole antenna in the center. As the current alternates, current flow is greatest in the middle of the antenna and gradually less as it approaches the ends. Then, it changes direction and flows the other way. The result is that the largest electromagnetic field is in the middle of the antenna and the strongest radio wave field is perpendicular to the length of the antenna. Most dipole antennas in aviation are horizontally polarized.

A common dipole antenna is the V-shaped VHF navigation

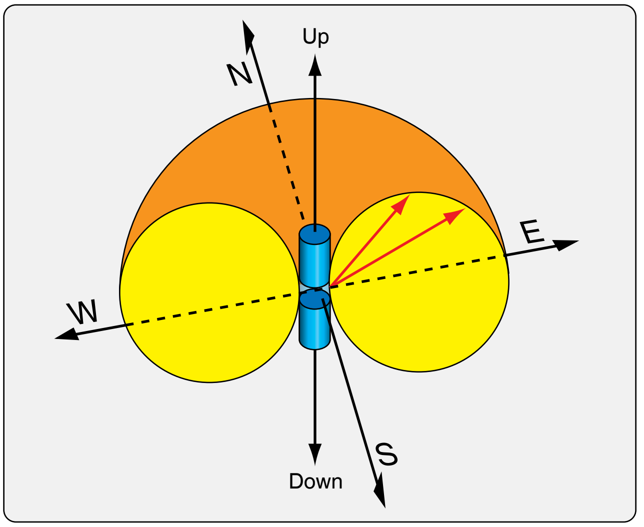


Figure 11-88. *A vertically polarized antenna radiates radio waves in a donut-like pattern in all directions.*

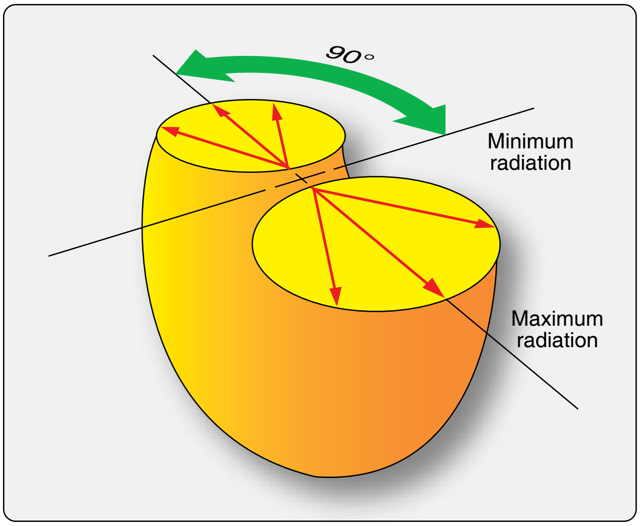


Figure 11-89. *A horizontally polarized antenna radiates in a donut-like pattern. The strongest signal is at 90° to the length of the conductor.*

Figure 11-90. *Many antenna are canted for better reception.*

antenna, known as a VOR antenna, found on numerous aircraft. Each arm of the V is one-fourth wavelength creating a half wave antenna which is fed in the center. This antenna is horizontally polarized. For a dipole receiving antenna, this means it is most sensitive to signals approaching the antenna from the sides rather than head-on in the direction of flight. *[Figure 11-91]*

Marconi Antenna

A Marconi antenna is a one-fourth wave antenna. It achieves the efficiency of a half wave antenna by using the mounting surface of the conductive aircraft skin to create the second one-fourth wavelength. Most aircraft VHF communications antennas are Marconi antennas. They are vertically polarized and create a field that is omnidirectional. On fabric skinned aircraft, the ground plane that makes up the second one-fourth wavelength of the antenna must be fashioned under the skin where the Marconi antenna is mounted. This can be done with thin aluminum or aluminum foil. Sometimes four or more wires are extended under the skin from the base of the vertical antenna that serve as the ground plane. This is enough to give the antenna the proper conductive length. The same practice is also utilized on ground-based antennas. *[Figure 11-92]*

Loop Antenna

The third type of antenna commonly found on aircraft is the loop antenna. When the length of an antenna conductor is fashioned into a loop, its field characteristics are altered significantly from that of a straight-half wavelength antenna. It also makes the antenna more compact and less prone to damage.

Used as a receiving antenna, the loop antenna's properties are highly direction-sensitive. A radio wave intercepting the loop directly broadside causes equal current flow in both sides of the loop. However, the polarity of the current flows is opposite each other. This causes them to cancel out and produce no signal. When a radio wave strikes the loop antenna in line with the plane of the loop, current is generated first in one side, and then in the other side. This causes the current flows to have different phases and the strongest signal can be generated from this angle. The phase difference (and strength) of the generated current varies proportionally to the angle at which the radio wave strikes the antenna loop. This is useful and is discussed further in the section on automatic direction finder (ADF) navigational aids. *[Figure 11-93]*

Transmission Lines

Transmitters and receivers must be connected to their antenna(s) via conductive wire. These transmission lines

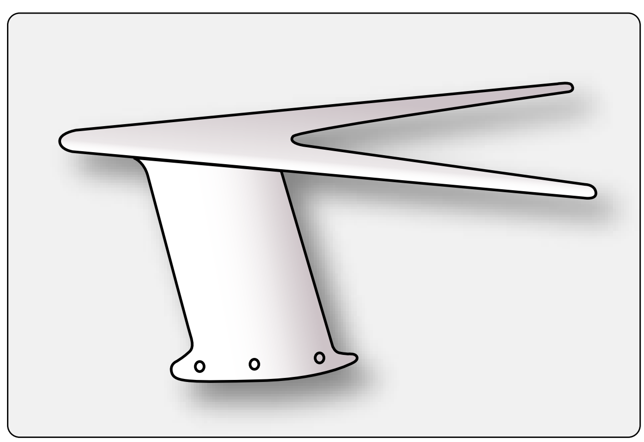


Figure 11-91. *The V-shaped VOR navigation antenna is a common dipole antenna.*

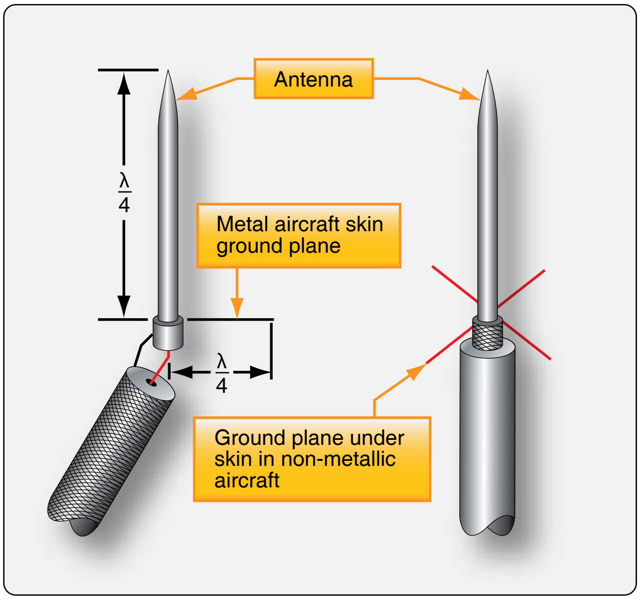


Figure 11-92. *On a metal-skinned aircraft, a ¼ wavelength Marconi antenna is used. The skin is the ground plane that creates the 2nd quarter of the antenna required for resonance (left). On a non-metallic-skinned aircraft, wires, conductive plates or strips equal in length to the antenna must be installed under the skin to create the ground plane (right).*

are coaxial cable, also known as coax. Coax consists of a center wire conductor surrounded by a semirigid insulator. Surrounding the wire and insulator material is a conductive, braided cover that runs the length of the cable. Finally, a waterproof covering is set around the braided shield to protect the entire assembly from the elements. The braided cover in the coax shields the inner conductor from any external fields. It also prevents the fields generated by the internal conductor from radiating. For optimum performance, the impedance of the transmission line should be equal to the impedance of the antenna. In aviation antenna applications, this is often approximately 50 ohms. *[Figure 11-94]* Special connectors are used for coaxial cable. A variety can be seen in Advisory Circular (AC) 43.13-1b, Chapter 11, Section 17, Figure 11-37. The technician should follow all manufacturer's instructions when installing transmission lines and antenna. Correct installation is critical to radio and antenna performance.

Radio Navigation

In the early years of aviation, a compass, a map, and dead reckoning were the only navigational tools. These were marginally reassuring if weather prevented the pilot from seeing the terrain below. Voice radio transmission from someone on the ground to the pilot indicating that the aircraft could be heard overhead was a preview of what electronic navigational aids could provide. For aviation to reach fruition as a safe, reliable, consistent means of transportation, some sort of navigation system needed to be developed.

Early flight instruments contributed greatly to flying when the ground was obscured by clouds. Navigation aids were needed to indicate where an aircraft was over the earth as it progressed toward its destination. In the 1930s and 1940s, a radio navigation system was used that was a low frequency, four-course radio range system. Airports and selected navigation waypoints broadcast two Morse code signals with finite ranges and patterns. Pilots tuned to the frequency of the broadcasts and flew in an orientation pattern until both signals were received with increasing strength. The signals were received as a blended tone of the highest volume when the aircraft was directly over the broadcast area. From this beginning, numerous refinements to radio navigational aids developed.

Radio navigation aids supply the pilot with intelligence that maintains or enhances the safety of flight. As with communication radios, navigational aids are avionics devices, the repair of which must be carried out by trained technicians at certified repair stations. However, installation, maintenance and proper functioning of the electronic units, as well as their antennas, displays, and any other peripheral devices, are the responsibilities of the airframe technician.

VOR Navigation System

One of the oldest and most useful navigational aids is the very high frequency omni-directional range (VOR) navigation system. The four main components of a typical system are: a receiver, a visual indicator, a frequency selector (controller or control panel), and antennas. The system was constructed after WWII and is still in use today. It consists of thousands of land-based transmitter stations, or VORs, that communicate with radio receiving equipment on board aircraft. Many of the VORs are located along airways. The Victor airway system is built around the VOR navigation system. Ground VOR transmitter units are also located at airports where they are known as TVOR (terminal VOR). The U.S. Military has a navigational system known as TACAN that operates similarly to the VOR system. Sometimes VOR and TACAN transmitters share a location. These sites are known as VORTACs.

The position of all VORs, TVORs, and VORTACs are marked on aeronautical charts along with the name of the station, the frequency to which an airborne receiver must be tuned to use the station, and a Morse code designation for the station. Some VORs also broadcast a voice identifier on a separate frequency that is included on the chart. *[Figure 11-95]*

VOR uses VHF radio waves (108–117.95 MHz) with 50 kHz separation between each channel. This keeps atmospheric interference to a minimum but limits the VOR to line-of-

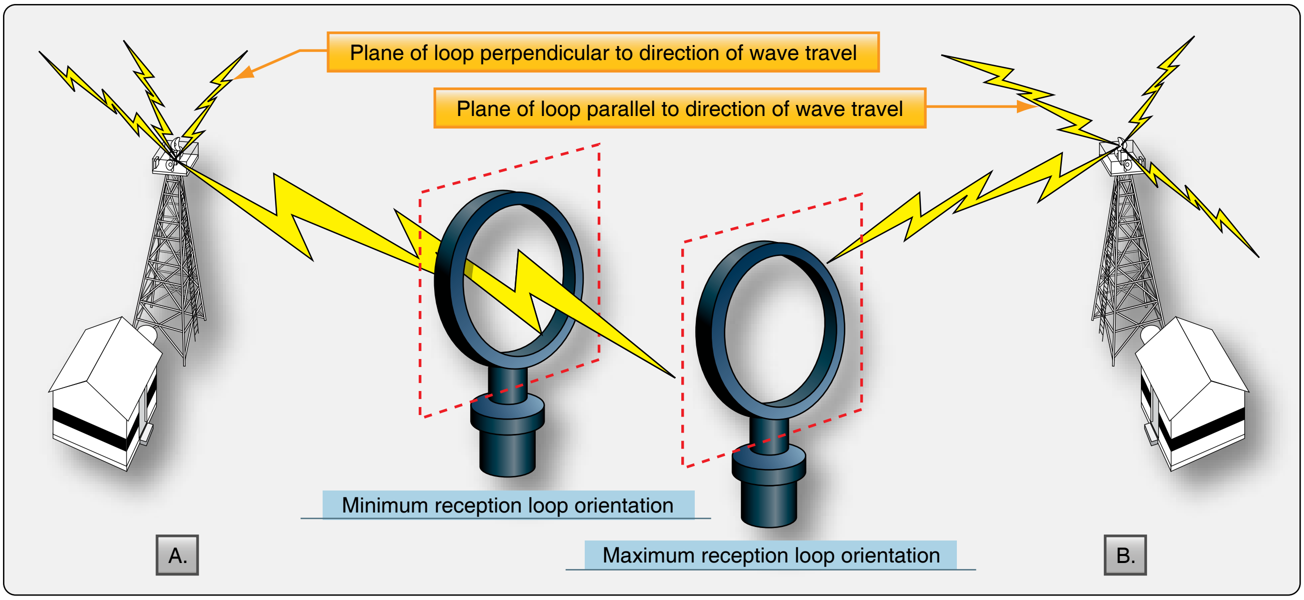


Figure 11-93. *A loop antenna is highly direction-sensitive. A signal origin perpendicular or broadside to the loop creates a weak signal (A). A signal origin parallel or in the plain of the loop creates a strong signal (B).*

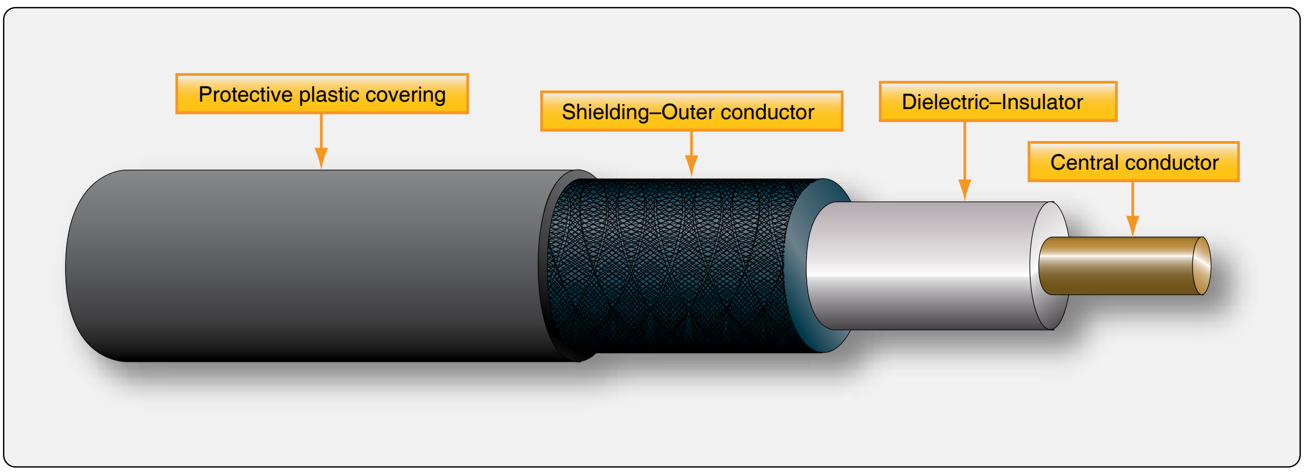


Figure 11-94. *Coaxial cable is used as the transmission line between an antenna and its transmitters and/or receiver.*

sight usage. To receive VOR VHF radio waves, generally a V-shaped, horizontally polarized, bi-pole antenna is used. Other type antennas are also certified. Follow the manufacturer's instructions for installation location. *[Figure 11-96]*

The signals produced by a VOR transmitter propagate 360° from the unit and are used by aircraft to navigate to and from the station with the help of an onboard VOR receiver and display instruments. A pilot is not required to fly a pattern to intersect the signal from a VOR station since it propagates out in every direction. The radio waves are received as long as the aircraft is in range of the ground unit and regardless of the aircraft's direction of travel. *[Figure 11-97]*

A VOR transmitter produces two signals that a receiver on board an aircraft uses to locate itself in relation to the ground station. One signal is a reference signal. The second is produced by electronically rotating a variable signal. The variable signal is in phase with the reference signal when at magnetic north but becomes increasingly out of phase as it is rotated to 180°. As it continues to rotate to 360° (0°), the signals become increasingly in phase until they are in phase again at magnetic north. The receiver in the aircraft deciphers the phase difference and determines the aircraft's position in degrees from the VOR ground based unit. *[Figure 11-98]*

Most aircraft carry a dual VOR receiver. Sometimes, the VOR receivers are part of the same avionics unit as the VHF communication transceiver(s). These are known as NAV/COM radios. Internal components are shared since frequency bands for each are adjacent. *[Figure 11-99]*

Figure 11-95. *A VOR ground station.*

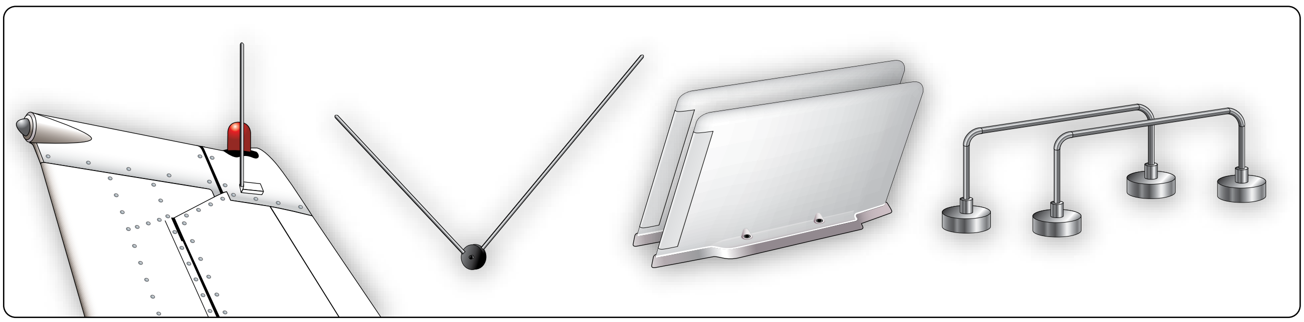


Figure 11-96. *V-shaped, horizontally polarized, bi-pole antennas are commonly used for VOR and VOR/glideslope reception. All antenna shown are VOR/glideslope antenna.*

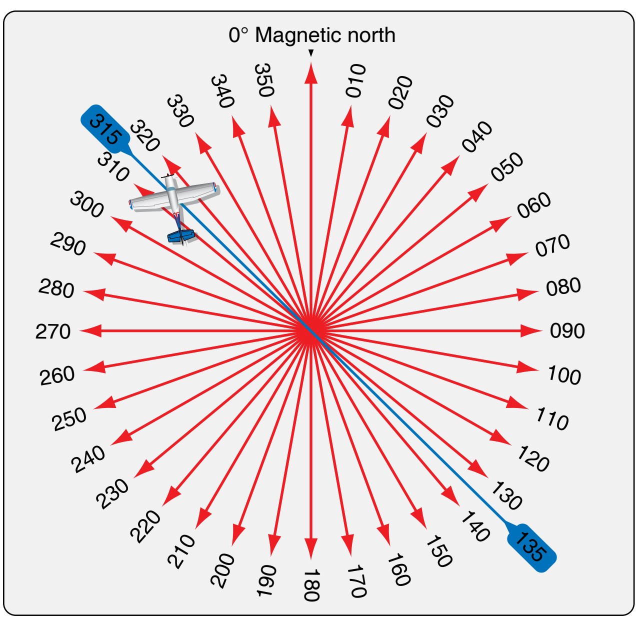


Figure 11-97. *A VOR transmitter produces signals for 360° radials that an airborne receiver uses to indicate the aircraft's location in relation to the VOR station regardless of the aircraft's direction of flight. The aircraft shown is on the 315° radial even though it does not have a heading of 315°.*

Large aircraft may have two dual receivers and even dual antennas. Normally, one receiver is selected for use and the second is tuned to the frequency of the next VOR station to be encountered en route. A means for switching between NAV 1 and NAV 2 is provided as is a switch for selecting the active or standby frequency. *[Figure 11-100]* VOR receivers are also found coupled with instrument landing system (ILS) receivers and glideslope receivers.

A VOR receiver interprets the bearing in degrees to (or from) the VOR station where the signals are generated. It also produces DC voltage to drive the display of the deviation from the desired course centerline to (or from) the selected station. Additionally, the receiver decides whether or not the aircraft is flying toward the VOR or away from it. These items can be displayed a number of different ways on various instruments. Older aircraft are often equipped with a VOR gauge dedicated to display only VOR information. This is also called an omni-bearing selector (OBS) or a course deviation indicator (CDI). *[Figure 11-101]*

The CDI linear indicator remains essentially vertical but moves left and right across the graduations on the instrument face to show deviation from being on course. Each graduation represents 2°. The OBS knob rotates the azimuth ring. When

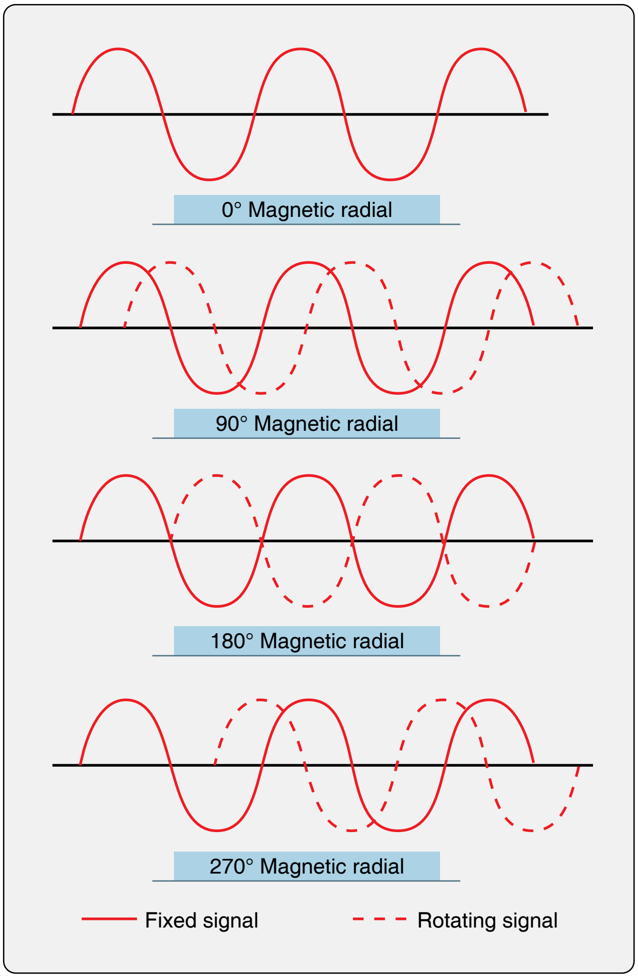


Figure 11-98. *The phase relationship of the two broadcast VOR signals.*

Figure 11-99. *A NAV/COM receiver typically found in light aircraft.*

in range of a VOR, the pilot rotates the OBS until the course deviation indicator centers. For each location of an aircraft, the OBS can be rotated to two positions where the CDI will center. One produces an arrow in the TO window of the gauge indicating that the aircraft is traveling toward the VOR station. The other selectable bearing is 180° from this. When chosen, the arrow is displayed in the FROM window indicating the aircraft is moving away from the VOR on the course selected. The pilot must steer the aircraft to the heading with the CDI centered to fly directly to or from the VOR. The displayed VOR information is derived from deciphering the phase relationship between the two simultaneously transmitted signals from the VOR ground station. When power is lost or the VOR signal is weak or interrupted, a NAV

Figure 11-100. *An airliner VOR control head with two independent NAV receivers each with an active and standby tuning circuit controlled by a toggle switch.*

warning flag comes into view. *[Figure 11-101]*

A separate gauge for the VOR information is not always used. As flight instruments and displays have evolved, VOR navigation information has been integrated into other instruments displays, such as the radio magnetic indicator (RMI), the HSI, an EFIS display or an electronic attitude director indicator (EADI). Flight management systems and automatic flight control systems are also made to integrate VOR information to automatically control the aircraft on its planned flight segments. Flat panel MFDs integrate VOR information into moving map presentations and other selected displays. The basic information of the radial bearing in degrees, course deviation indication, and to/from information remains unchanged however. *[Figure 11-102]*

At large airports, an instrument landing system (ILS) guides the aircraft to the runway while on an instrument landing approach. The aircraft's VOR receiver is used to interpret the radio signals. It produces a more sensitive course deviation indication on the same instrument display as the VOR CDI display. This part of the ILS is known as the localizer and is discussed below. While tuned to the ILS localizer frequency, the VOR circuitry of the VOR/ILS receiver is inactive.

It is common at VOR stations to combine the VOR transmitter with distance measuring equipment (DME) or a nondirectional beacon (NDB) such as an ADF transmitter and antenna. When used with a DME, pilots can gain an exact fix on their location using the VOR and DME together. Since the VOR indicates the aircraft's bearing to the VOR transmitter and a co-located DME indicates how far away the station is, this relieves the pilot from having to fly over the station to

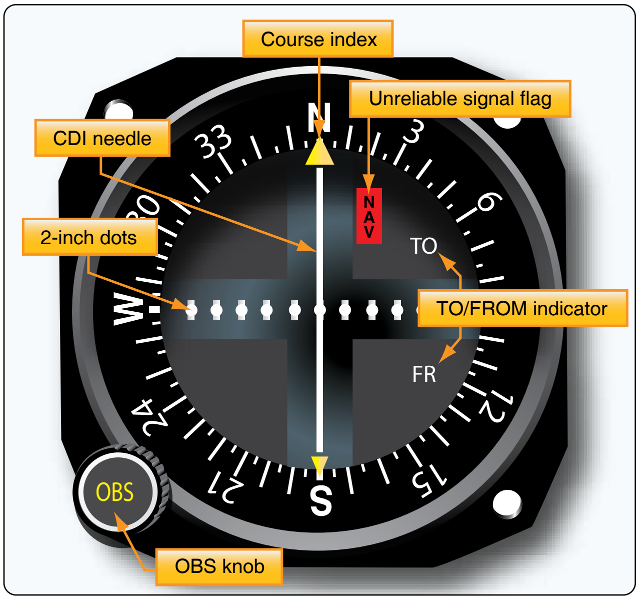


Figure 11-101. *A traditional VOR gauge, also known as a course deviation indicator (CDI) or an omni-bearing selector (OBS).*

know with certainty their location. These navigational aids are discussed separately in the following sections.

Functional accuracy of VOR equipment is critical to the safety of flight. VOR receivers are operationally tested using VOR test facilities (VOT). These are located at numerous airports that can be identified in the Airport Facilities Directory for the area concerned. Specific points on the airport surface are given to perform the test. Most VOTs require tuning 108.0 MHz on the VOR receiver and centering the CDI. The OBS should indicate 0° showing FROM on the indicator or 180° when showing TO. If an RMI is used as the indicator, the test heading should always indicate 180°. Some repair stations can also generate signals to test VOR receivers although not on 108.0 MHz. Contact the repair station for the transmission frequency and for their assistance in checking the VOR system. A logbook entry is required.

Note: Some airborne testing using VOTs is possible by the pilot.

An error of ±4° should not be exceeded when testing a VOR system with a VOT. An error in excess of this prevents the use of the aircraft for IFR fight until repairs are made. Aircraft having dual VOR systems where only the antenna is shared may be tested by comparing the output of each system to the other. Tune the VOR receivers to the local ground VOR station. A bearing indication difference of no more than ±4° is permissible.

Automatic Direction Finder (ADF)

An automatic direction finder (ADF) operates off of a ground signal transmitted from a NDB. Early radio direction finders (RDF) used the same principle. A vertically polarized antenna was used to transmit LF frequency radio waves in the 190 kHz to 535 kHz range. A receiver on the aircraft

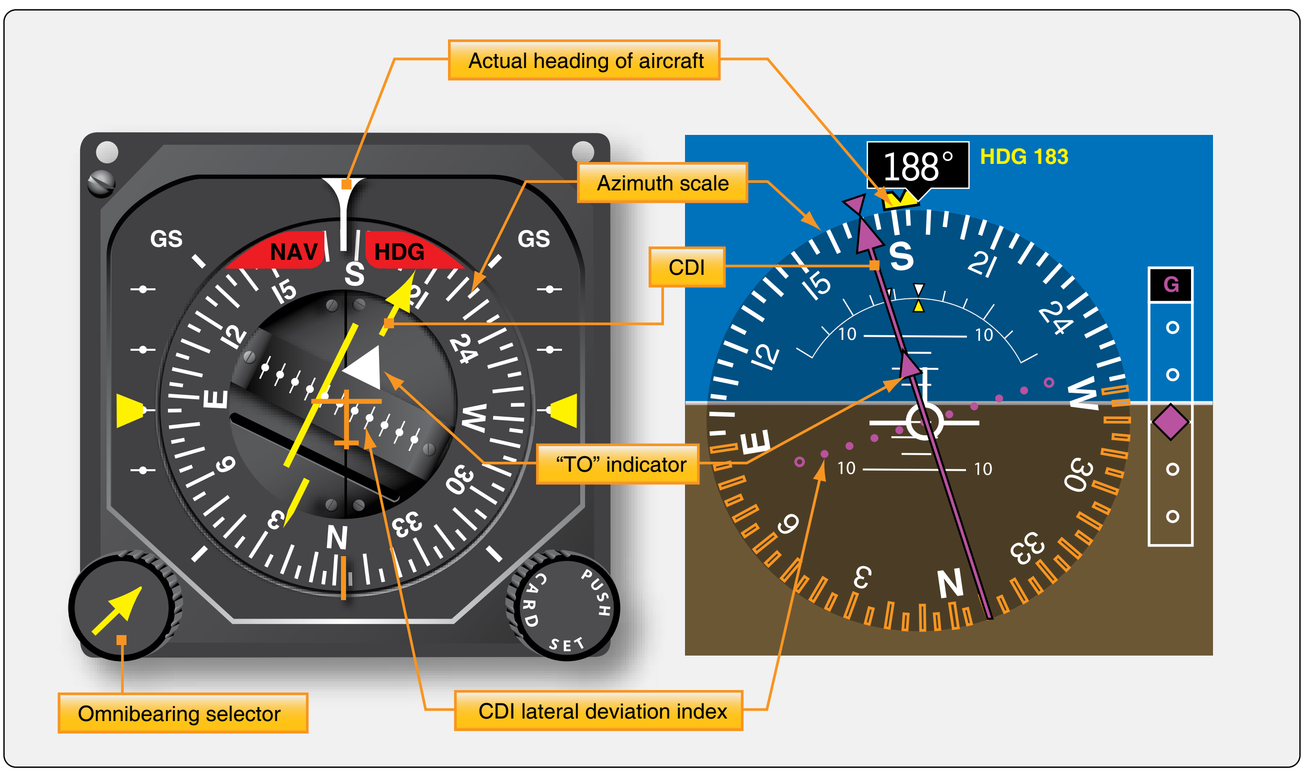


Figure 11-102. *A mechanical HSI (left) and an electronic HSI (right) both display VOR information.*

was tuned to the transmission frequency of the NDB. Using a loop antenna, the direction to (or from) the antenna could be determined by monitoring the strength of the signal received. This was possible because a radio wave striking a loop antenna broadside induces a null signal. When striking it in the plane of the loop, a much stronger signal is induced. The NDB signals were modulated with unique Morse code pulses that enabled the pilot to identify the beacon to which they were navigating.

With RDF systems, a large rigid loop antenna was installed inside the fuselage of the aircraft. The broadside of the antenna was perpendicular to the aircraft's longitudinal axis. The pilot listened for variations in signal strength of the LF broadcast and maneuvered the aircraft so a gradually increasing null signal was maintained. This took them to the transmitting antenna. When over flown, the null signal gradually faded as the aircraft became farther from the station. The increasing or decreasing strength of the null signal was the only way to determine if the aircraft was flying to or from the NDB. A deviation left or right from the course caused the signal strength to sharply increase due to the loop antenna's receiving properties.

The ADF improved on this concept. The broadcast frequency range was expanded to include MF up to about 1800 kHz. The heading of the aircraft no longer needed to be changed to locate the broadcast transmission antenna. In early model ADFs, a rotatable antenna was used instead. The antenna rotated to seek the position in which the signal was null. The direction to the broadcast antenna was shown on an azimuth scale of an ADF indicator in the flight deck. This type of instrument is still found in use today. It has a fixed card with 0° always at the top of a non-rotating dial. A pointer indicates the relative bearing to the station. When the indication is 0°, the aircraft is on course to (or from) the station. *[Figure 11-103]*

As ADF technology progressed, indicators with rotatable azimuth cards became the norm. When an ADF signal is received, the pilot rotates the card so that the present heading is at the top of the scale. This results in the pointer indicating the magnetic bearing to the ADF transmitter. This is more intuitive and consistent with other navigational practices. *[Figure 11-104]*

In modern ADF systems, an additional antenna is used to remove the ambiguity concerning whether the aircraft is heading to or from the transmitter. It is called a sense antenna. The reception field of the sense antenna is omnidirectional. When combined with the fields of the loop antenna, it forms a field with a single significant null reception area on one side. This is used for tuning and produces an indication in the

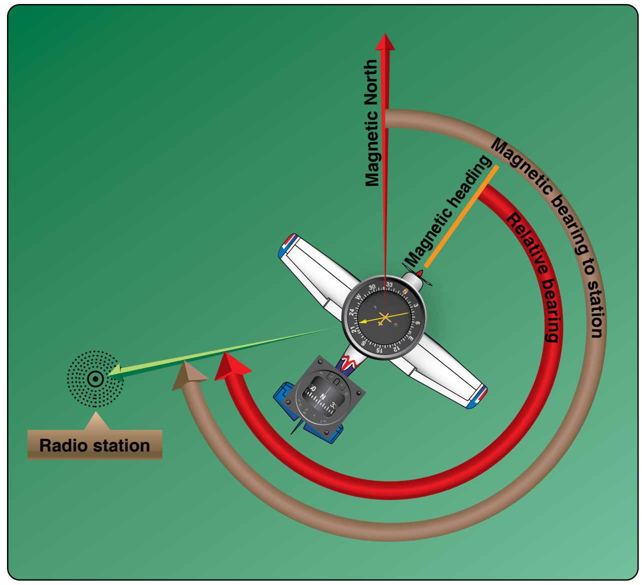


Figure 11-103. *Older ADF indicators have nonrotating azimuth cards. 0° is fixed at the top of the instrument and the pointer always indicates the relative bearing to the ADF transmission antenna. To fly to the station, the pilot turns the aircraft until the ADF pointer indicates 0°.*

direction toward the ADF station at all times. The onboard ADF receiver needs only to be tuned to the correct frequency of the broadcast transmitter for the system to work. The loop and sense antenna are normally housed in a single, low profile antenna housing. *[Figure 11-105]*

Any ground antenna transmitting LF or MF radio waves in

Figure 11-104. *A movable card ADF indicator can be rotated to put the aircraft's heading at the top of the scale. The pointer then points to the magnetic bearing the ADF broadcast antenna.*

range of the aircraft receiver's tuning capabilities can be used for ADF. This includes those from AM radio stations. Audible identifier tones are loaded on the NDB carrier waves. Typically, a two-character Morse code designator is used. With an AM radio station transmission, the AM broadcast is heard instead of a station identifier code. The frequency for an NDB transmitter is given on an aeronautical chart next to a symbol for the transmitter. The identifying designator is also given. *[Figure 11-106]*

ADF receivers can be mounted in the flight deck with the controls accessible to the user. This is found on many general aviation aircraft. Alternately, the ADF receiver is mounted in a remote avionics bay with only the control head in the flight deck. Dual ADF receivers are common. ADF information can be displayed on the ADF indicators mentioned or it can be digital. Modern, flat, multipurpose electronic displays usually display the ADF digitally. *[Figure 11-107]* When ANT is selected on an ADF receiver, the loop antenna is cut out and only the sense antenna is active. This provides better multi-directional reception of broadcasts in the ADF frequency range, such as weather or AWAS broadcasts.

When the best frequency oscillator (BFO) is selected on an

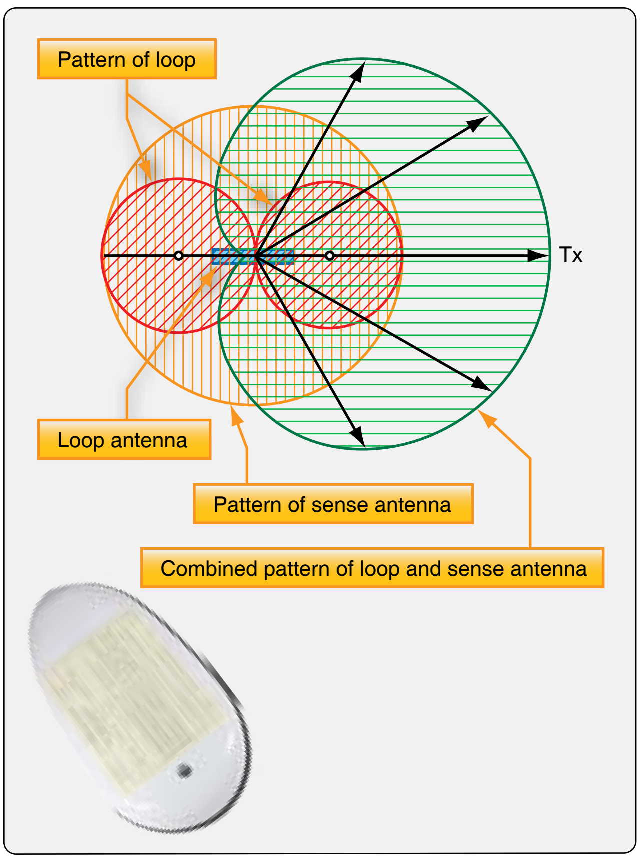


Figure 11-105. *The reception fields of a loop and sense antenna combine to create a field with a sharp null on just one side. This removes directional ambiguity when navigating to an ADF station.*

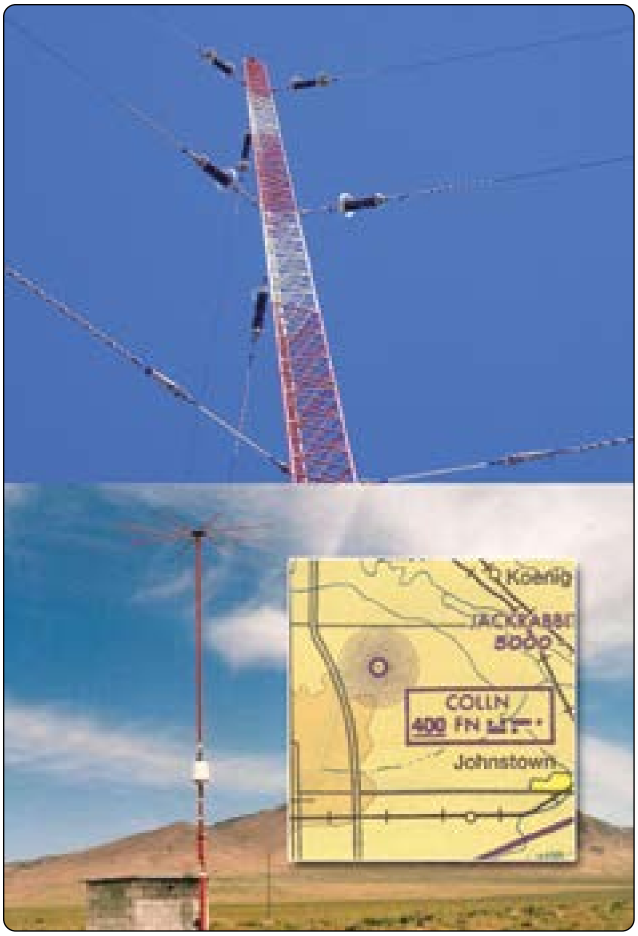


Figure 11-106. *Nondirectional broadcast antenna in the LF and medium frequency range are used for ADF navigation.*

ADF receiver/controller, an internal beat frequency oscillator is connected to the IF amplifier inside the ADF receiver. This is used when an NDB does not transmit a modulated signal.

Continued refinements to ADF technology has brought it to its current state. The rotating receiving antenna is replaced by a fixed loop with a ferrite core. This increases sensitivity and allows a smaller antenna to be used. The most modern ADF systems have two loop antennas mounted at 90° to each other. The received signal induces voltage that is sent to two stators in a resolver or goniometer. The goniometer stators induce voltage in a rotor that correlates to the signal of the fixed loops. The rotor is driven by a motor to seek the null. The same motor rotates the pointer in the flight deck indicator to show the relative or magnetic bearing to the station. *[Figure 11-108]*

Technicians should note that the installation of the ADF antenna is critical to a correct indication since it is a directional device. Calibration with the longitudinal axis of the fuselage or nose of the aircraft is important. A single null reception area must exist in the correct direction. The antenna

Figure 11-107. *A flight deck mountable ADF receiver used on general aviation aircraft.*

must be oriented so the ADF indicates station location when the aircraft is flying toward it rather than away. Follow all manufacturer's instructions.

Radio Magnetic Indicator (RMI)

To save space in the instrument panel and to consolidate related information into one easy to use location, the radio magnetic indicator (RMI) has been developed. It is widely used. The RMI combines indications from a magnetic compass, VOR, and ADF into one instrument. *[Figure 11-109]*

The azimuth card of the RMI is rotated by a remotely located flux gate compass. Thus, the magnetic heading of the aircraft is always indicated. The lubber line is usually a marker or triangle at the top of the instrument dial. The VOR receiver drives the solid pointer to indicate the magnetic direction TO a tuned VOR station. When the ADF is tuned to an NDB, the double, or hollow pointer, indicates the magnetic bearing to the NDB.

Since the flux gate compass continuously adjusts the azimuth card so that the aircraft heading is at the top of the instrument, pilot workload is reduced. The pointers indicate where the VOR and ADF transmission stations are located in relationship to where the aircraft is currently positioned. Push buttons allow conversion of either pointer to either ADF or VOR for navigation involving two of one type of station and none of the other.

Instrument Landing Systems (ILS)

An ILS is used to land an aircraft when visibility is poor. This radio navigation system guides the aircraft down a slope to the touch down area on the runway. Multiple radio transmissions are used that enable an exact approach to landing with an ILS. A localizer is one of the radio transmissions. It is used to provide horizontal guidance to the center line of the runway. A separate glideslope broadcast provides vertical guidance of the aircraft down the proper slope to the touch down point. Compass locator transmissions for outer and middle approach marker beacons aid the pilot in intercepting the

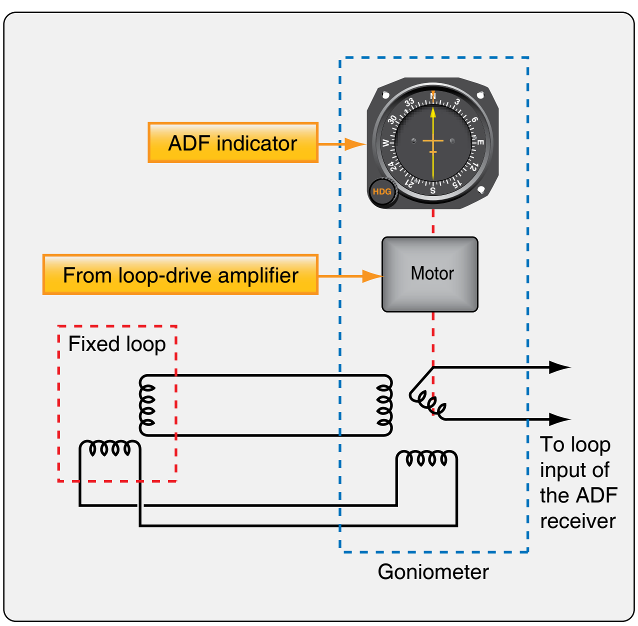


Figure 11-108. *In modern ADF, a rotor in a goniometer replaces a the rotating loop antenna used in earlier models.*

Figure 11-109. *A radio magnetic indicator (RMI) combines a magnetic compass, VOR, and ADF indications.*

approach navigational aid system. Marker beacons provide distance-from-the-runway information. Together, all of these radio signals make an ILS a very accurate and reliable means for landing aircraft. *[Figure 11-110]*

Localizer

The localizer broadcast is a VHF broadcast in the lower range of the VOR frequencies (108 MHz–111.95 MHz) on odd frequencies only. Two modulated signals are produced from a horizontally polarized antenna complex beyond the far end of the approach runway. They create an expanding field that is 2½° wide (about 1,500 feet) 5 miles from the runway. The field tapers to runway width near the landing threshold. The left side of the approach area is filled with a VHF carrier wave modulated with a 90 Hz signal. The right side of the approach contains a 150 MHz modulated signal. The aircraft's VOR receiver is tuned to the localizer VHF frequency that can be found on published approach plates and aeronautical charts.

The circuitry specific to standard VOR reception is inactive while the receiver uses localizer circuitry and components common to both. The signals received are passed through filters and rectified into DC to drive the course deviation indicator. If the aircraft receives a 150 Hz signal, the CDI of the VOR/ILS display deflects to the left. This indicates that the runway is to the left. The pilot must correct course with a turn to the left. This centers the course deviation indicator on the display and centers the aircraft with the centerline of the runway. If the 90 Hz signal is received by the VOR receiver, the CDI deflects to the right. The pilot must turn toward the right to center the CDI and the aircraft with the runway center line. *[Figure 11-111]*

Glideslope

The vertical guidance required for an aircraft to descend for a landing is provided by the glideslope of the ILS. The glideslope provides vertical guidance for correct angle of descent. Radio signals funnel the aircraft down to the touchdown point on the runway at an angle of approximately 3°. The transmitting glideslope antenna is located off to the side of the approach runway approximately 1,000 feet from the threshold. It transmits in a wedge-like pattern with the field narrowing as it approaches the runway. *[Figure 11-112]*

The glideslope transmitter antenna is horizontally polarized. The transmitting frequency range is UHF between 329.3 MHz and 335.0 MHz. The frequency is paired to the localizer frequency of the ILS. When the VOR/ILS receiver is tuned for the approach, the glideslope receiver is automatically tuned. Like the localizer, the glideslope transmits two signals, one modulated at 90 Hz and the other modulated at 150 Hz. The aircraft's glideslope receiver deciphers the signals similar to the method of the localizer receiver. It drives a vertical course deviation indicator known as the glideslope indicator. The glideslope indicator operates identically to the localizer CDI only 90° to it. The VOR/ILS localizer CDI and the glideslope are displayed together on whichever kind of instrumentation is in the aircraft. *[Figure 11-113]*

The UHF antenna for aircraft reception of the glideslope signals comes in many forms. A single dipole antenna mounted inside the nose of the aircraft is a common option. Antenna manufacturers have also incorporated glideslope reception into the same dipole antenna used for the VHS VOR/ILS localizer reception. Blade type antennas are also used. *[Figures 11-114] Figure 11-115* shows a VOR and a glideslope receiver for a GA aircraft ILS.

Compass Locators

It is imperative that a pilot be able to intercept the ILS to enable its use. A compass locator is a transmitter designed for this purpose. There is typically one located at the outer marker beacon 4–7 miles from the runway threshold. Another may be located at the middle marker beacon about 3,500 feet from the threshold. The outer marker compass locator is a 25 watt NDB with a range of about 15 miles. It transmits omnidirectional LF radio waves (190 Hz to 535 Hz) keyed with the first two letters of the ILS identifier. The ADF receiver is used to intercept the locator so no additional equipment is required. If a middle marker compass locator is in place, it is similar but is identified with the last two letters of the ILS identifier. Once located, the pilot maneuvers the aircraft to fly down the glidepath to the runway.

Marker Beacons

Marker beacons are the final radio transmitters used in the ILS. They transmit signals that indicate the position of the aircraft along the glidepath to the runway. As mentioned, an outer marker beacon transmitter is located 4–7 miles from the threshold. It transmits a 75 MHz carrier wave modulated with a 400 Hz audio tone in a series of dashes. The transmission is very narrow and directed straight up. A marker beacon receiver receives the signal and uses it to light a blue light on the instrument panel. This, plus the oral tone in combination with the localizer and the glideslope indicator, positively locates the aircraft on an approach. *[Figure 11-115]*

A middle marker beacon is also used. It is located on approach approximately 3,500 feet from the runway. It also transmits at 75 MHz. The middle marker transmission is modulated with a 1300 Hz tone that is a series of dots and dashes so as to not be confused with the all dash tone of the outer marker. When the signal is received, it is used in the receiver to illuminate an amber-colored light on the instrument panel. *[Figure 11-116]*

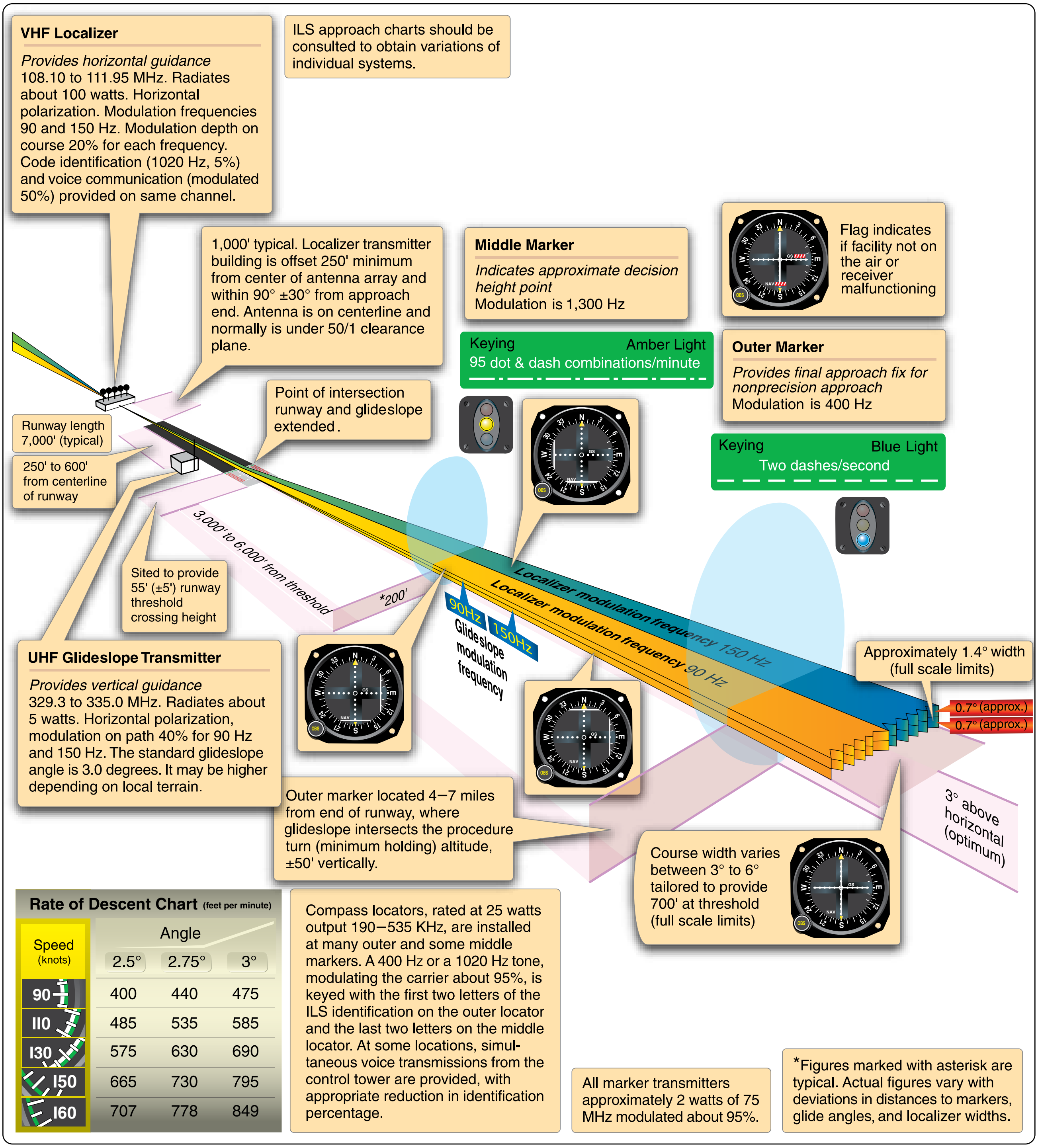


Figure 11-110. *Components of an instrument landing system (ILS).*

Some ILS approaches have an inner marker beacon that transmits a signal modulated with 3000 Hz in a series of dots only. It is placed at the land-or-go-around decision point of the approach close to the runway threshold. If present, the signal when received is used to illuminate a white light on the instrument panel. The three marker beacon lights are usually incorporated into the audio panel of a general aviation aircraft or may exist independently on a larger aircraft. Electronic display aircraft usually incorporate marker lights or indicators close to the glideslope display near attitude director indicator. *[Figure 11-117]*

ILS radio components can be tested with an ILS test unit. Localizer, glideslope, and marker beacon signals are generated to ensure proper operation of receivers and correct display on flight deck instruments. *[Figure 11-118]*

Figure 11-111. *An ILS localizer antenna.*

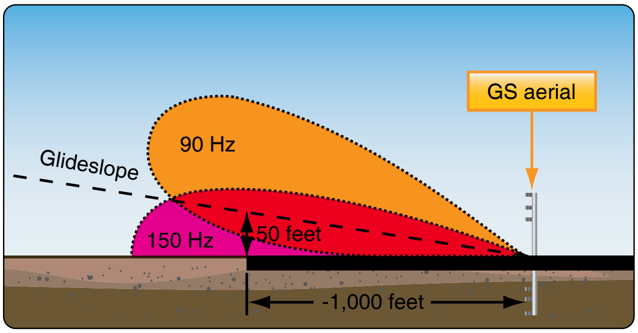


Figure 11-112. *A glideslope antenna broadcasts radio signals to guide an aircraft vertically to the runway.*

Distance Measuring Equipment (DME)

Many VOR stations are co-located with the military version of the VOR station, which is known as TACAN. When this occurs, the navigation station is known as a VORTAC station. Civilian aircraft make use of one of the TACAN features not originally installed at civilian VOR stations–DME. A DME

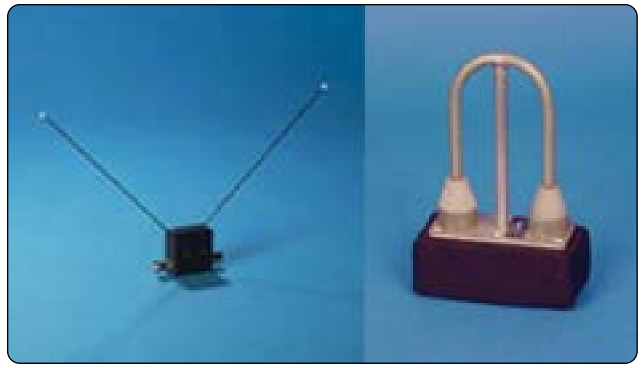


Figure 11-114. *Glideslope antennas—designed to be mounted inside a non-metallic aircraft nose (left), and mounted inside or outside the aircraft (right).*

system calculates the distance from the aircraft to the DME unit at the VORTAC ground station and displays it on the flight deck. It can also display calculated aircraft speed and elapsed time for arrival when the aircraft is traveling to the station.

DME ground stations have subsequently been installed at civilian VORs, as well as in conjunction with ILS localizers. These are known as VOR/DME and ILS/DME or LOC/DME. The latter aid in approach to the runway during landings. The DME system consists of an airborne DME transceiver, display, and antenna, as well as the ground based DME unit and its antenna. *[Figure 11-119]*

The DME is useful because with the bearing (from the VOR) and the distance to a known point (the DME antenna at the VOR), a pilot can positively identify the location of the aircraft. DME operates in the UHF frequency range from 962 MHz to 1213 MHz. A carrier signal transmitted from

Figure 11-113. *A traditional course deviation indicator is shown on the left. The horizontal white line is the deviation indicator for the glideslope. The vertical line is for the localizer. On the right, a Garmin G-1000 PFD illustrates an aircraft during an ILS approach. The narrow vertical scale on the right of the attitude indicator with the "G" at the top is the deviation scale for the glideslope. The green diamond moves up and down to reflect the aircraft being above or below the glidepath. The diamond is shown centered indicating the aircraft is on course vertically. The localizer CDI can be seen at the bottom center of the display. It is the center section of the vertical green course indicator. LOC1 is displayed to the left of it.*

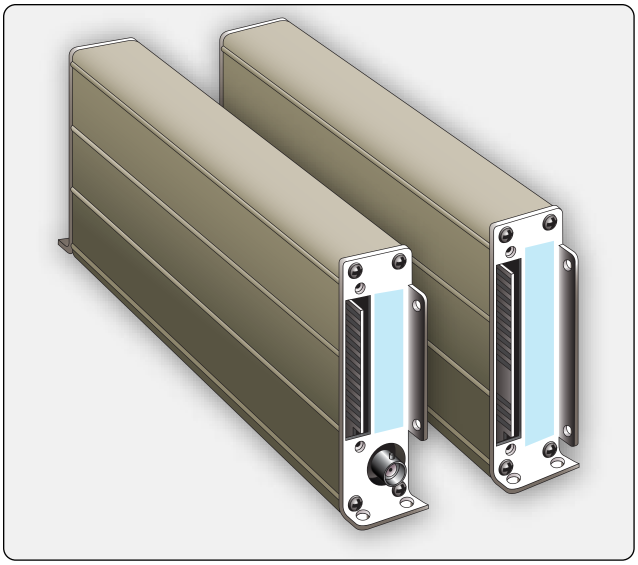


Figure 11-115. *A localizer and glideslope receiver for a general aviation aircraft ILS.*

Figure 11-116. *Various marker beacon instrument panel display lights.*

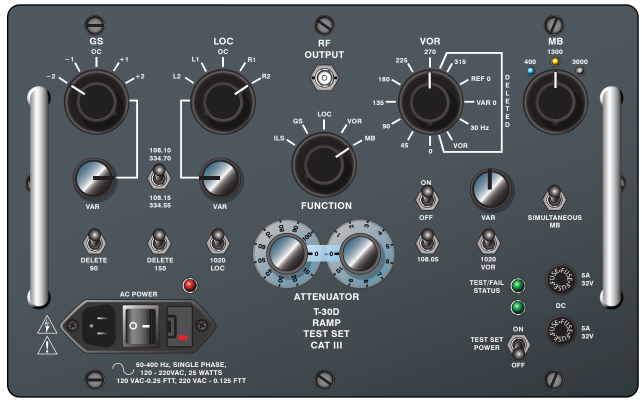


Figure 11-118. *An ILS test unit.*

the aircraft is modulated with a string of integration pulses. The ground unit receives the pulses and returns a signal to the aircraft. The time that transpires for the signal to be sent and returned is calculated and converted into nautical miles for display. Time to station and speed are also calculated and displayed. DME readout can be on a dedicated DME display or it can be part of an EHSI, EADI, EFIS, or on the primary flight display in a glass flight deck. *[Figure 11-120]*

The DME frequency is paired to the co-located VOR or VORTAC frequency. When the correct frequency is tuned for the VOR signal, the DME is tuned automatically. Tones are broadcast for the VOR station identification and then for the DME. The hold selector on a DME panel keeps the DME tuned in while the VOR selector is tuned to a different VOR. In most cases, the UHF of the DME is transmitted and received via a small blade-type antenna mounted to the underside of the fuselage centerline. *[Figure 11-121]*

A traditional DME displays the distance from the DME transmitter antenna to the aircraft. This is called the slant distance. It is very accurate. However, since the aircraft is at altitude, the distance to the DME ground antenna from a

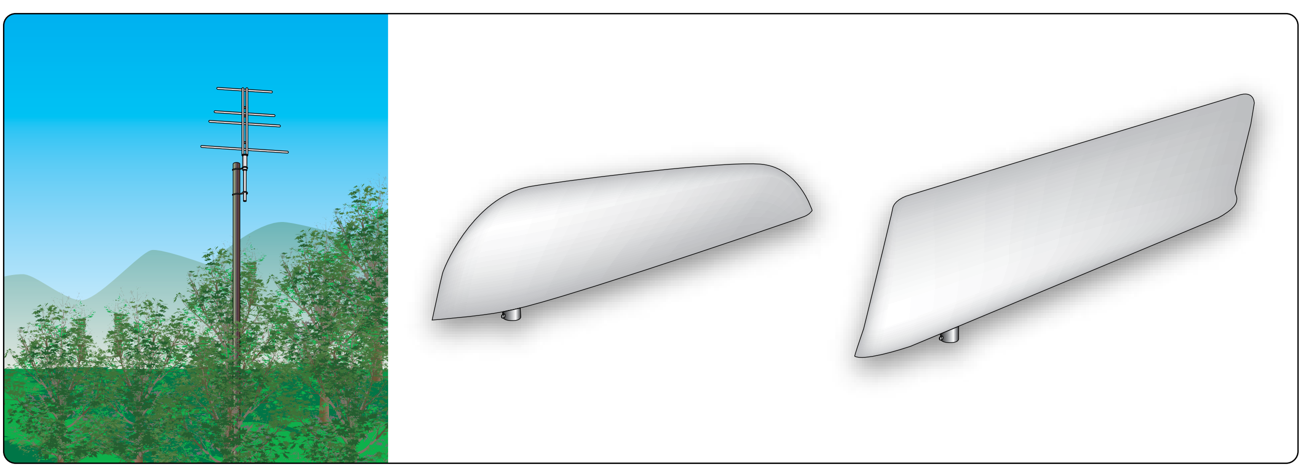


Figure 11-117. *An outer marker transmitter antenna 4–7 miles from the approach runway transmits a 75 MHz signal straight up (left). Aircraft mounted marker beacon receiver antennas are shown (center and right).*

Figure 11-119. *A VOR with DME ground station.*

Figure 11-120. *Distance information from the DME can be displayed on a dedicated DME instrument or integrated into any of the electronic navigational displays found on modern aircraft. A dual display DME is shown with its remote mounted receiver.*

point directly beneath the aircraft is shorter. Some modern DMEs are equipped to calculate this ground distance and display it. *[Figure 11-122]*

Area Navigation (RNAV)

Area navigation (RNAV) is a general term used to describe the navigation from point A to point B without direct over flight of navigational aids, such as VOR stations or ADF non-directional beacons. It includes VORTAC and VOR/DME based systems, as well as systems of RNAV based around LORAN, GPS, INS, and the FMS of transport category aircraft. However, until recently, the term RNAV was most commonly used to describe the area navigation or the process of direct flight from point A to point B using VORTAC and VOR/DME based references which are discussed in this section.

Figure 11-121. *A typical aircraft-mounted DME antenna.*

All RNAV systems make use of waypoints. A waypoint is a designated geographical location or point used for route definition or progress-reporting purposes. It can be defined or described by using latitude/longitude grid coordinates or, in the case of VOR based RNAV, described as a point on a VOR radial followed by that point's distance from the VOR station (i.e., 200/25 means a point 25 nautical miles from the VOR station on the 200° radial).

Figure 11-123 illustrates an RNAV route of flight from airport A to airport B. The VOR/DME and VORTAC stations shown are used to create phantom waypoints that are overflown rather than the actual stations. This allows a more direct route to be taken. The phantom waypoints are entered into the RNAV course-line computer (CLC) as a radial and distance number pair. The computer creates the waypoints and causes the aircraft's CDI to operate as though they are actual VOR stations. A mode switch allows the choice between standard VOR navigation and RNAV.

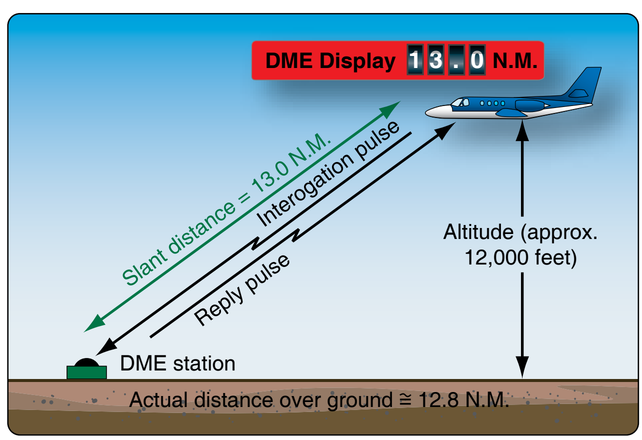


Figure 11-122. *Many DME's only display the slant distance, which is the actual distance from the aircraft to the DME station. This is different than the ground distance due to the aircraft being at altitude. Some DMEs compute the ground distance for display.*

VOR based RNAV uses the VOR receiver, antenna, and VOR display equipment, such as the CDI. The computer in the RNAV unit uses basic geometry and trigonometry calculations to produce heading, speed, and time readouts for each waypoint. VOR stations need to be within line-of sight and operational range from the aircraft for RNAV use. *[Figure 11-124]*

RNAV has increased in flexibility with the development of GPS. Integration of GPS data into a planned VOR RNAV flight plan is possible as is GPS route planning without the use of any VOR stations.

Radar Beacon Transponder

A radar beacon transponder, or simply, a transponder, provides positive identification and location of an aircraft on the radar screens of ATC. For each aircraft equipped with an altitude encoder, the transponder also provides the pressure altitude of the aircraft to be displayed adjacent to the on-screen blip that represents the aircraft. *[Figure 11-125]*

Radar capabilities at airports vary. Generally, two types of radar are used by air traffic control (ATC). The primary radar transmits directional UHF or SHF radio waves sequentially in all directions. When the radio waves encounter an aircraft, part of those waves reflect back to a ground antenna. Calculations are made in a receiver to determine the direction

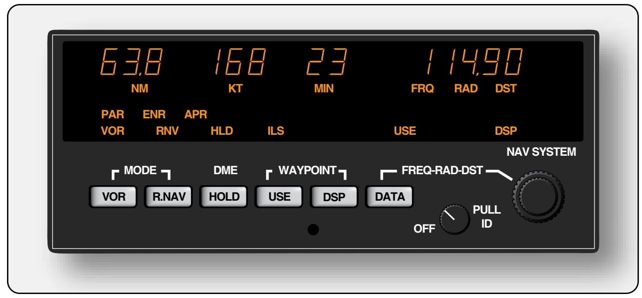


Figure 11-124. *RNAV unit from a general aviation aircraft.*

and distance of the aircraft from the transmitter. A blip or target representing the aircraft is displayed on a radar screen also known as a plan position indicator (PPI). The azimuth direction and scaled distance from the tower are presented giving controllers a two dimensional fix on the aircraft. *[Figure 11-126]*

A secondary surveillance radar (SSR) is used by ATC to verify the aircraft's position and to add the third dimension of altitude to its location. SSD radar transmits coded pulse trains that are received by the transponder on board the aircraft. Mode 3/A pulses, as they are known, aid in confirming the location of the aircraft. When verbal communication is established with ATC, a pilot is instructed to select one of 4,096 discrete codes on the transponder. These are digital octal codes. The ground station transmits a pulse of energy at 1030 MHz and the transponder transmits a reply with

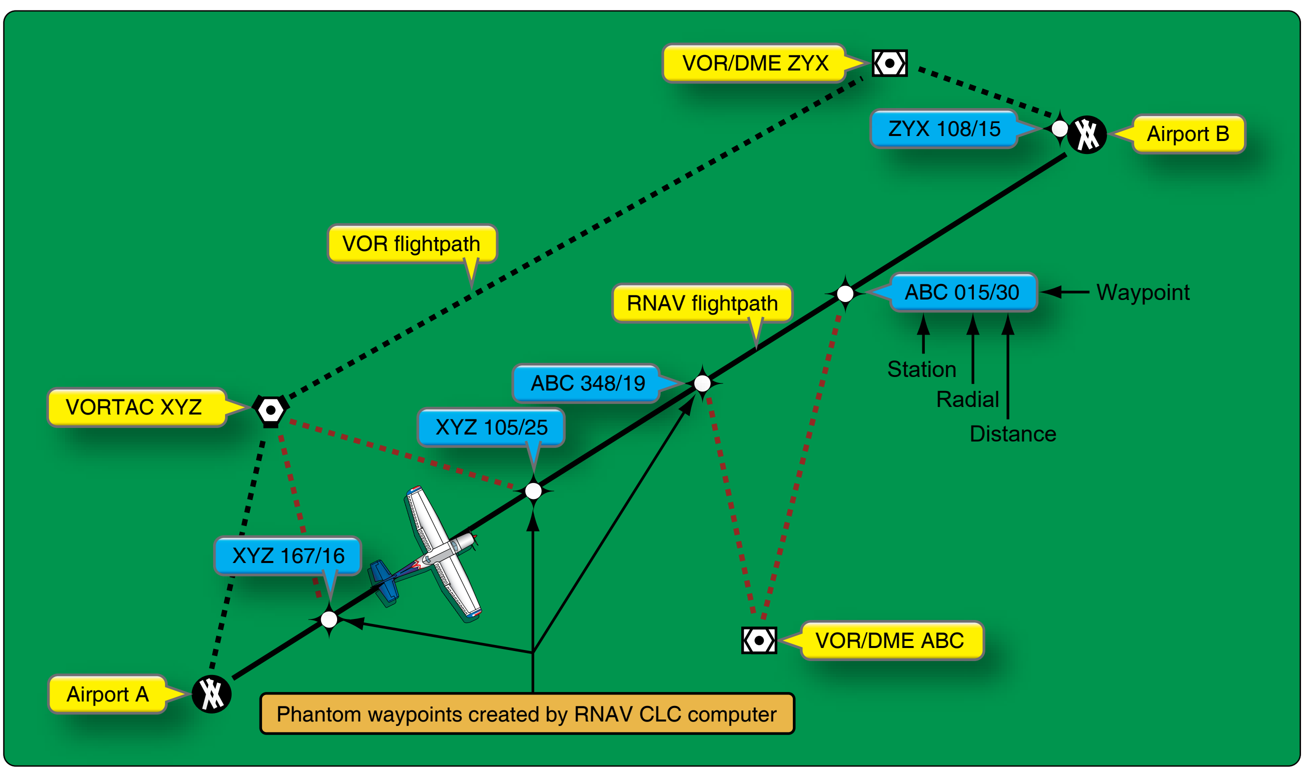


Figure 11-123. *The pilot uses the aircraft's course deviation indicator to fly to and from RNAV phantom waypoints created by computer. This allows direct routes to be created and flown rather than flying from VOR to VOR.*

Figure 11-125. *A traditional transponder control head (A), a lightweight digital transponder (B), and a remote altitude encoder (C) that connects to a transponder to provide ATC with an aircraft's altitude displayed on a PPI radar screen next to the target that represents the aircraft.*

the assigned code attached at 1090 MHz. This confirms the aircraft's location typically by altering its target symbol on the radar screen. As the screen may be filled with many confirmed aircraft, ATC can also ask the pilot to ident. By pressing the IDENT button on the transponder, it transmits in such a way that the aircraft's target symbol is highlighted on the PPI to be distinguishable.

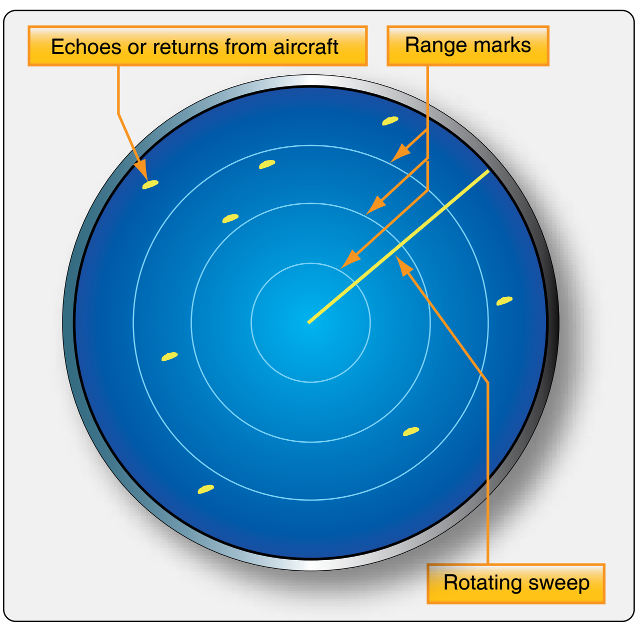


Figure 11-126. *A plan position indicator (PPI) for ATC primary radar locates target aircraft on a scaled field.*

To gain altitude clarification, the transponder control must be placed in the ALT or Mode C position. The signal transmitted back to ATC in response to pulse interrogation is then modified with a code that places the pressure altitude of the aircraft next to the target symbol on the radar screen. The transponder gets the pressure altitude of the aircraft from an altitude encoder that is electrically connected to the transponder. Typical aircraft transponder antennas are illustrated in *Figure 11-127.*

The ATC/aircraft transponder system described is known as Air Traffic Control Radar Beacon System (ATCRBS). To increase safety, Mode S altitude response has been developed. With Mode S, each aircraft is pre-assigned a unique identity code that displays along with its pressure altitude on ATC radar when the transponder responds to SSR interrogation. Since no other aircraft respond with this code, the chance of two pilots selecting the same response code on the transponder is eliminated. A modern flight data processor computer (FDP) assigns the beacon code and searches flight plan data for useful information to be displayed on screen next to the target in a data block for each aircraft. *[Figure 11-128]*

Mode S is sometimes referred to as mode select. It is a data packet protocol that is also used in onboard collision avoidance systems. When used by ATC, Mode S interrogates one aircraft at a time. Transponder workload is reduced by not having to respond to all interrogations in an airspace.

Figure 11-127. *Aircraft radar beacon transponder antennas transmit and receive UHF and SHF radio waves.*

Additionally, location information is more accurate with Mode S. A single reply in which the phase of the transponder reply is used to calculate position, called monopulse, is sufficient to locate the aircraft. Mode S also contains capacity for a wider variety of information exchange that is untapped potential for the future. At the same time, compatibility with older radar and transponder technology has been maintained.

Transponder Tests & Inspections

Title 14 of the Code of Federal Regulations (CFR) part 91, section 91.413 states that all transponders on aircraft flown into controlled airspace are required to be inspected and tested in accordance with 14 CFR part 43, Appendix F, every 24 calendar months. Installation or maintenance that may introduce a transponder error is also cause for inspection and test in accordance with Appendix F. Only an appropriately rated repair station, the aircraft manufacturer (if it installed transponder), and holders of a continuous airworthy program are approved to conduct the procedures. As with many radio-electronic devices, test equipment exists to test airworthy operation of a transponder. *[Figure 11-129]*

Operating a transponder in a hangar or on the ramp does not immunize it from interrogation and reply. Transmission of certain codes reserved for emergencies or military activity must be avoided. The procedure to select a code during ground operation is to do so with the transponder in the OFF or STANDBY mode to avoid inadvertent transmission. Code 0000 is reserved for military use and is a transmittable code. Code 7500 is used in a hijack situation and 7600 and 7700 are also reserved for emergency use. Even the inadvertent transmission of code 1200 reserved for VFR flight not under ATC direction could result in evasion action. All signals received from a radar beacon transponder are taken seriously by ATC.

Altitude Encoders

Altitude encoders convert the aircraft's pressure altitude into a code sent by the transponder to ATC. Increments of 100 feet are usually reported. Encoders have varied over the years. Some are built into the altimeter instrument used in the instrument panel and connected by wires to the transponder. Others are mounted out of sight on an avionics rack or similar out of the way place. These are known as blind encoders. On transport category aircraft, the altitude encoder may be a large black box with a static line connection to an internal aneroid. Modern general aviation encoders are smaller and more lightweight, but still often feature an internal aneroid and static line connection. Some encoders use microtransistors and are completely solid-state including the pressure sensing device from which the altitude is derived. No static port connection is required. Data exchange with GPS and other systems is becoming common. *[Figure 11-130]*

When a transponder selector is set on ALT, the digital pulse message sent in response to the secondary surveillance radar interrogation becomes the digital representation of the pressure altitude of the aircraft. There are 1280 altitude codes, one for each 100 feet of altitude between 1200 feet mean sea level (MSL) and 126,700 feet MSL. Each altitude increment is assigned a code. While these would be 1280 of the same codes used for location and IDENT, the Mode C (or S) interrogation deactivates the 4096 location codes and causes the encoder to become active. The correct altitude code is sent to the transponder that replies to the interrogation. The SSR receiver recognized this as a response to a Mode C (or S) interrogation and interprets the code as altitude code.

Collision Avoidance Systems

The ever increasing volume of air traffic has caused a corresponding increase in concern over collision avoidance. Ground-based radar, traffic control, and visual vigilance are no longer adequate in today's increasingly crowded skies. Onboard collision avoidance equipment, long a staple in larger aircraft, is now common in general aviation aircraft.

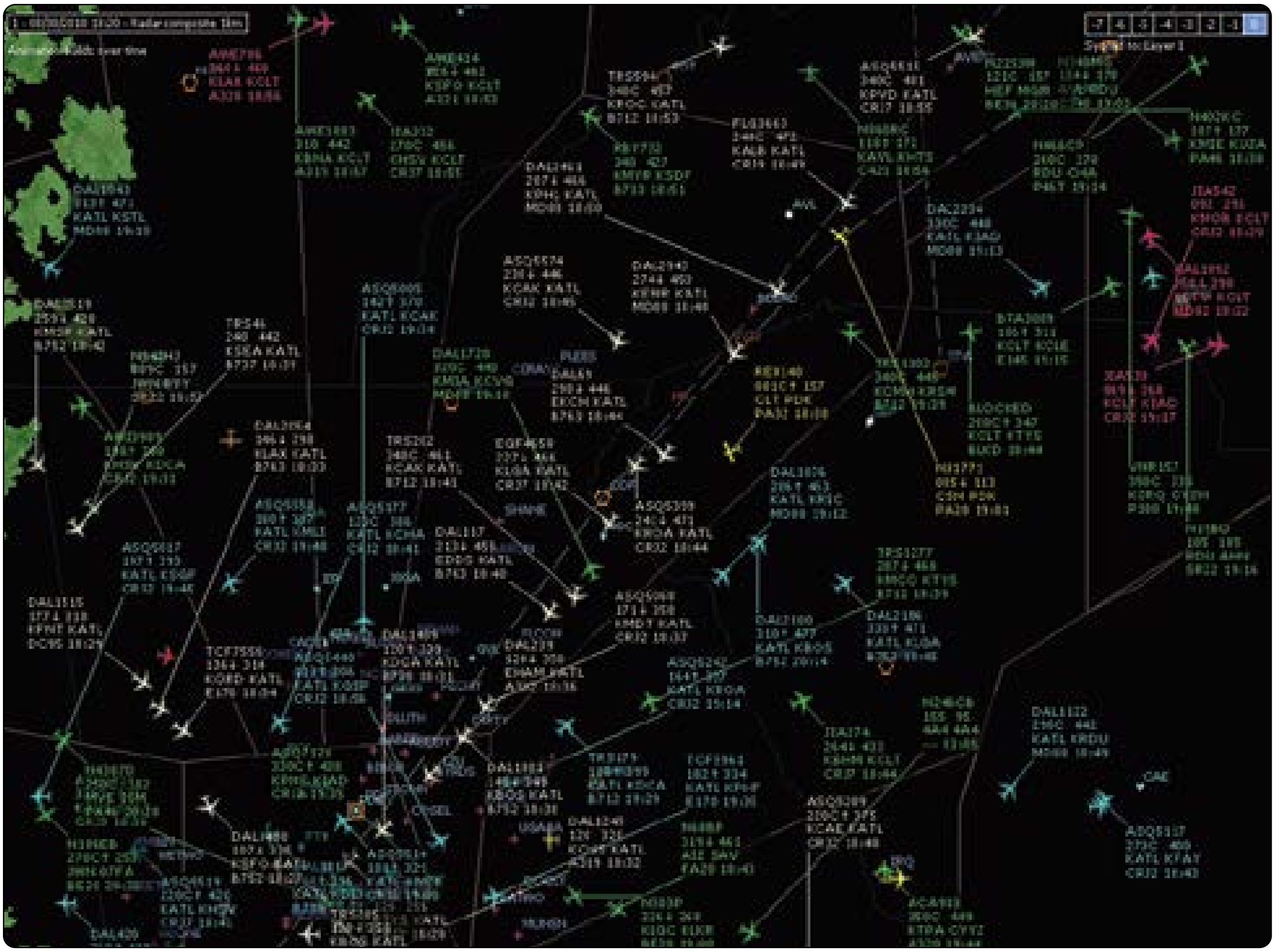


Figure 11-128. *Air traffic control radar technology and an onboard radar beacon transponder work together to convey and display air traffic information on a PPI radar screen. A modern approach ATC PPI is shown. Targets representing aircraft are shown as little aircraft on the screen. The nose of the aircraft indicates the direction of travel. Most targets shown above are airliners. The data block for each target includes the following information either transmitted by the transponder or matched and loaded from flight plans by a flight data processor computer: call sign, altitude/speed, origination/destination, and aircraft type/ETA (ZULU time). A "C" after the altitude indicates the information came from a Mode C equipped transponder. The absence of a C indicates Mode S is in use. An arrow up indicates the aircraft is climbing. An arrow down indicates a descent. White targets are arrivals, light blue targets are departures, all other colors are for arrivals and departures to different airports in the area.*

New applications of electronic technology combined with lower costs make this possible.

Traffic Collision Avoidance Systems (TCAS)

Traffic collision avoidance systems (TCAS) are transponder based air-to-air traffic monitoring and alerting systems. There are two classes of TCAS. TCAS I was developed to accommodate the general aviation community and regional airlines. This system identifies traffic in a 35–40 mile range of the aircraft and issues Traffic Advisories (TA) to assist pilots in visual acquisition of intruder aircraft. TCAS I is mandated on aircraft with 10 to 30 seats.

TCAS II is a more sophisticated system. It is required internationally in aircraft with more than 30 seats or weighing more than 15,000 kg. TCAS II provides the information of TCAS I, but also analyzes the projected flightpath of approaching aircraft. If a collision or near miss is imminent, the TCAS II computer issues a Resolution Advisory (RA). This is an aural command to the pilot to take a specific evasive action (i.e., DESCEND). The computer is programmed such that the pilot in the encroaching aircraft receives an RA for evasive action in the opposite direction (if it is TCAS II equipped). *[Figure 11-131]*

The transponder of an aircraft with TCAS is able to interrogate the transponders of other aircraft nearby using SSR technology (Mode C and Mode S). This is done with a 1030 MHz signal. Interrogated aircraft transponders reply with an encoded 1090 MHz signal that allows the TCAS computer to display the position and altitude of each aircraft. Should the aircraft come within the horizontal or vertical distances shown in *Figure 11-131*, an audible TA is

Figure 11-129. *A handheld transponder test unit.*

announced. The pilot must decide whether to take action and what action to take. TCAS II equipped aircraft use continuous reply information to analyze the speed and trajectory of target aircraft in close proximity. If a collision is calculated to be imminent, an RA is issued.

TCAS target aircraft are displayed on a screen on the flight deck. Different colors and shapes are used to depict approaching aircraft depending on the imminent threat level. Since RAs are currently limited to vertical evasive maneuvers, some stand-alone TCAS displays are electronic vertical speed indicators. Most aircraft use some version of an electronic HSI on a navigational screen or page to display TCAS information. *[Figure 11-132]* A multifunction display may depict TCAS and weather radar information on the same screen. *[Figure 11-133]* A TCAS control panel *[Figure 11-134]* and computer are required to work with a compatible transponder and its antenna(s). Interface with EFIS or other previously installed or selected display(s) is also required.

TCAS may be referred to as airborne collision avoidance system (ACAS), which is the international name for the same system. TCAS II with the latest revisions is known as Version 7. The accuracy and reliability of this TCAS information is such that pilots are required to follow a TCAS RA over an ATC command.

ADS-B

Collision avoidance is a significant part of the FAA's NextGen plan for transforming the National Airspace

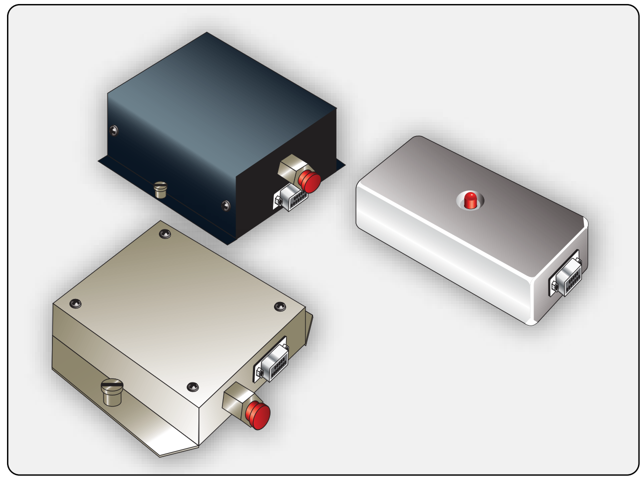


Figure 11-130. *Modern altitude encoders for general aviation aircraft.*

System (NAS). Increasing the number of aircraft using the same quantity of airspace and ground facilities requires the implementation of new technologies to maintain a high level of performance and safety. The successful proliferation of global navigation satellite systems (GNSS), such as GPS, has led to the development of a collision avoidance system known as Automatic Dependent Surveillance-Broadcast (ADS-B). ADS-B is an integral part of NextGen program. The implementation of its ground and airborne infrastructure is currently underway. ADS-B is active in parts of the United States and around the world. *[Figure 11-135]*

ADS-B is considered in two segments: ADS-B OUT and ADS-B IN. ADS-B OUT combines the positioning information available from a GPS receiver with on-board flight status information, i.e., location including altitude, velocity, and time. It then broadcasts this information to other ADS-B equipped aircraft and ground stations. *[Figure 11-136]*

Two different frequencies are used to carry these broadcasts with data link capability. The first is an expanded use of the 1090 MHz Mode-S transponder protocol known as 1090 ES. The second, largely being introduced as a new broadband solution for general aviation implementation of ADS-B, is at 978 MHz. A 978 universal access transceiver (UAT) is used to accomplish this. An omni-directional antenna is required in addition to the GPS antenna and receiver. Airborne receivers of an ADS-B use the information to plot the location and movement of the transmitting aircraft on a flight deck display similar to TCAS. *[Figure 11-137]*

Inexpensive ground stations (compared to radar) are constructed in remote and obstructed areas to proliferate ADS-B. Ground stations share information from airborne ADS-Bs with other ground stations that are part of the air