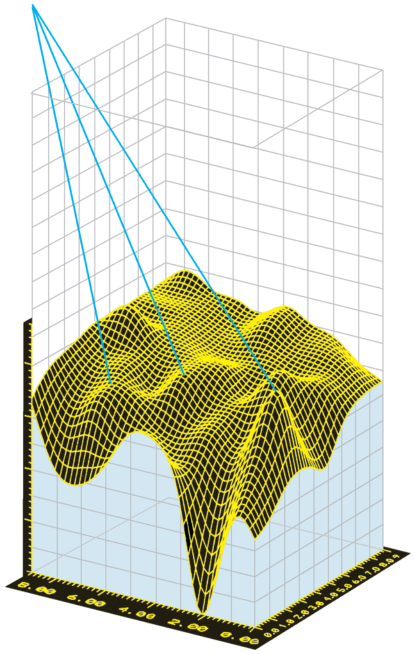


A Primer on GPS Antennas

Richard B. Langley

University of New Brunswick

The GPS receiver is a marvel of modern electronic engineering. By processing the signals transmitted by the constellation of orbiting Navstar satellites, its sophisticated circuitry can deliver position, velocity, and time information to a user anywhere on or near the earth's surface, 24 hours a day, every day. But before the receiver can use the signals, they must first be captured. This is the task of the receiver's antenna.

GPS signals are relatively weak compared with the signals from broadcasting stations and terrestrial communications services, and a GPS antenna is specially designed to work with these feeble signals — a coat hanger will not do! In this month's column, we'll take a look at the GPS antenna. This will only be an introduction to the complex subject of antenna design and construction, but it should enable you to better understand antenna specifications and how your receiver's antenna works.

"Innovation" is a regular column featuring discussions about recent advances in GPS technology and its applications as well as the fundamentals of GPS positioning. The column is coordinated by Richard Langley of the Department of Geodesy and Geomatics Engineering at the University of New Brunswick, who appreciates receiving your comments as well as topic suggestions for future columns. To contact him, see the "Columnists" section on page 4 of this issue.

How does a GPS receiver's antenna, or any antenna for that matter, manage to snare radio signals out of the ether and pass them on to the receiver? And what are the important characteristics of antennas? These are questions that we will attempt to answer in this month's column. Before we can understand how antennas work, however, we need a quick refresher on electromagnetic waves.

FIELDS AND WAVES

An electromagnetic wave is a self-propagating wave with both electric and magnetic field components generated by the acceleration of a charged particle. If we connect an oscillator — a generator of alternating voltage — to the center of a piece of wire, the outer electrons of the atoms making up the wire move back and forth, generating an alternating current and an oscillating electric field. The oscillating electric field produces an oscillating magnetic field, which, in turn, gives rise to a new oscillating electric field, and so on. These coupled fields form a unified electromagnetic field that continually spreads ever outward from the antenna in the form of an electromagnetic wave.

Unlike sound waves, electromagnetic waves do not require a medium for their transmission. The ether, the medium once thought to permeate all of space and to permit the waves to propagate, is now considered imaginary. When a receiving antenna intercepts an electromagnetic wave, the wave's associated field will induce currents in the antenna that are then fed to the receiver through a transmission line.

Although it is the wave's electric field that stimulates the currents in most receiving antennas, the current in some antennas — notably loop antennas of the kind used in AM radios and some antennas used to receive marine differential GPS radiobeacons, for example — comes primarily from the magnetic field.

In many practical situations, an antenna used for reception has the same properties as the identical antenna when used for transmission. This characteristic of reciprocity proves useful when describing how antennas work.

It can be shown that in free space (a vacuum) — or in any homogeneous, isotropic, linear, and stationary medium — the electric and magnetic fields are transverse to the direction of propagation, and the fields are mutually perpendicular. If we introduce a coordinate system whose positive z-axis is aligned with the direction of propagation of the wave, then the electric and magnetic field vectors lie in the x-y plane.

One can decompose the vector describing the electric field into two orthogonal vectors, one parallel to the positive x-axis and one parallel to the positive y-axis. If the x- and y-components have the same phase (or are different by an integer multiple of π), the wave is said to be linearly polarized, as the electric field vector is always directed along a fixed line. Most terrestrial radio signals are linearly polarized with the electric field oriented either horizontally or vertically. If the two components differ in phase, their sum describes an ellipse about the z-axis. This is an elliptically polarized wave. If the two components have the same amplitude but are $\pi/2$ (or an odd multiple of $\pi/2$) out of phase, the ellipse becomes a circle and the wave is said to be circularly polarized (see Figure 1).

Circular polarization implies a "handedness." If, at a fixed point in space, the electric (and magnetic) field vectors rotate clockwise (counterclockwise) for an observer looking from the source toward the direction of wave propagation, the polarization is right-handed (left-handed). (This is the electrical engineer's convention; in classical physics literature, we sometimes find the opposite convention.) The signals emitted by GPS satellites are right-hand circularly polarized (RHCP). Some other satellites, notably spin-stabilized ones, also use circular polarization. For maximum signal strength, the polarization of the receiving antenna must match the polarization of the signals.

The antenna described in our opening paragraph is perhaps the simplest kind of antenna — the dipole. At its microscopic limit, it consists of a pair of spaced charges of equal magnitude and of opposite signs. When this antenna is driven, the charges oscillate in magnitude and polarity, generating a uniform current between the charges and radiating an electromagnetic wave. On the macroscopic scale, one can approximate such an idealized antenna using a center-driven piece of wire whose length is a fraction of the wavelength of its radiation, with metal plates or spheres attached to the ends of the dipole that increase the antenna's capacitance. The capacitive ends allow the antenna to store more energy, which helps maintain a nearly uniform charging current in the wire.

We can easily calculate the radiating properties of this so-called Hertzian dipole. We

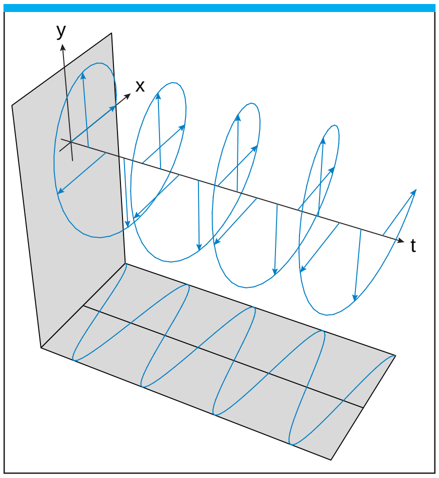


Figure 1. At a fixed point in space, the electric field vector of a right-hand circularly polarized wave rotates clockwise as seen from the wave's source.

can consider a real antenna to be equivalent to a number of Hertzian dipoles in series, and we can deduce its properties by superimposing the effects of the individual Hertzian dipoles.

ANTENNA CHARACTERISTICS

The GPS antenna's job is to convert the energy in the electromagnetic waves arriving from the satellites into an electric current that can be processed by the electronics in the receiver. The antenna's size and shape are very important, as these characteristics govern, in part, the antenna's ability to pick up and pass on to the receiver the very weak GPS signals. The antenna may be required to operate at just the L1 frequency or, for dual-frequency receivers, at both the L1 and L2 frequencies. Also, because the GPS signals are RHCP, all GPS antennas must be RHCP as well. Despite these restrictions, several different types of antennas have been and are currently used with GPS receivers. These include monopole and dipole configurations, quadrifilar helices (also known as volutes), spiral helices, slots, and microstrips.

Perhaps the most common GPS antenna is the microstrip because of its ruggedness and relative ease of construction. It can be circular or rectangular in shape and is roughly similar in appearance to a small piece of copper-clad printed circuit board. Made up of one or more patches of metal separated from a ground plane by a dielectric sheet (referred to as the substrate), microstrips are often referred to as patch antennas. They may have either single- or dual-frequency capability, and their exceptionally low profile makes them ideal for many applications.

Axial Ratio. To be maximally sensitive to GPS signals, the ideal GPS antenna should be perfectly right-hand circularly polarized. However, a real antenna will actually be elliptically polarized. The more elliptically polarized it is, the lower its RHCP sensitivity, or gain. The degree of ellipticity is given by the antenna's axial ratio (essentially the ratio of the axes of the polarization ellipse). An axial ratio of unity, or 0 decibels (dB), implies circular polarization. Good GPS antennas have an axial ratio in the zenith direction of 2 dB or better.

Impedance. An antenna will have some *resistance* to current flowing in it. In addition, it may have some *reactance* because of antenna capacitance or inductance, which also affect the alternating current flow. The combined opposition to current is known as *impedance*. It can be expressed as a complex number whose real part is the resistance and the imaginary part is the reactance. The relationship between impedance and the voltage and current in an antenna or any circuit element is given by the extension of Ohm's law to alternating current: It is the ratio of the voltage to the current and is measured in ohms.

For antennas, impedance is typically measured at the feed point where the antenna is connected to the transmission line. An antenna's impedance depends on many factors: how it is constructed, how it is fed, and, to some degree, the surrounding environment. For example, an isolated center-fed dipole with a length equal to one-half its operating free-space wavelength has a characteristic impedance of 73 + j42.5 ohms. By making the antenna shorter by about 5 percent to account for the current speed in the antenna being slightly smaller than the vacuum speed of light, the reactance disappears and the antenna becomes resonant at the free-space wavelength.

The folded dipole, commonly used as the driving element in multi-element Yagi antennas (often used for point-to-point signal transmission and reception at very- and ultra-high frequencies, for example), has a resonant impedance of about 300 ohms. Most GPS antennas are designed with a characteristic impedance of 50 ohms.

If one places an antenna inside an enclosure, its impedance and resonant frequency may change. A microstrip patch antenna placed in a plastic enclosure, for example, can have its resonant frequency shifted downward by several megahertz (MHz), depending on the thickness of the plastic and its dielectric constant. Antenna manufacturers, therefore, purposely design their patch antennas to resonate at a higher frequency than the actual operating value.

Standing Wave Ratio. In a resonant antenna, the current and voltage distribution is a standing or stationary wave. A standing wave is formed by the superposition of waves of the same frequency traveling in opposite directions. In the ideal half-wavelength dipole, for example, the current traveling from the feed point undergoes a reflection at the ends of the wire. The reflected wave, when combined with the incident wave, creates a sinusoidal standing wave that has a constant zero amplitude at each end — a null or node — and a maximum amplitude in the center — a loop.

An important consideration is the antenna–receiver connection. This is achieved with a transmission line, usually a coaxial cable. To maximize signal transfer from the antenna to the receiver, we must minimize power loss. Power may be lost if the coupling between the antenna and the cable is imperfect and also within the cable itself. (We'll discuss cable loss later.) To prevent power loss at the interface between the cable and the antenna, the impedances of the cable and the antenna must be the same.

Antennas are designed to be coupled to a coaxial cable with a certain impedance. If the antenna's characteristic impedance is different from that of the cable, then one must incorporate a matching circuit of some kind within the antenna so that the cable and antenna impedances match at the antenna connection terminals. As we've mentioned, most GPS antennas have a 50-ohm impedance because they are designed to work with 50-ohm cable. If there is a mismatch between antenna and matching circuit or matching circuit and cable, signal reflections can occur, which give rise to standing waves in the cable.

The impedance mismatch can be quantified by measuring the peak signal voltages along the cable. The ratio of the maximum to minimum peak voltages is called the *voltage standing wave ratio* (VSWR). (We can also measure the VSWR by expressing it as a function of the forward and reflected signal powers.) In the absence of reflections (in other words, a perfect match), the VSWR is 1:1. Such an ideal VSWR is essentially impossible to obtain in practice, and most consider a value of 1.5:1 to be quite good. The corresponding signal loss for a VSWR of 1.5:1 is only 0.18 dB.

How an antenna is fed can also determine its polarization. Several techniques exist for feeding microstrip patch antennas, for example, to produce circular polarization.

Bandwidth. Another important characteristic of an antenna is its bandwidth. This is the frequency band over which the antenna's performance, as measured by one or more parameters (such as input impedance, pattern, polarization, and so on) is acceptably good. The bandwidth needs to be large enough so that the antenna functions well over the range of frequencies for which it is intended.

Antennas may be narrowband or broadband. Resonant antennas are characteristically narrowband, although their bandwidth can be increased through certain construction techniques. If one uses the antenna for a system that employs only a narrow frequency band, then keeping the antenna's bandwidth narrow is advantageous, as potentially interfering signals on adjacent frequencies will then be attenuated somewhat. Most single-frequency GPS antennas are narrowband devices.

A microstrip patch antenna designed for use with a standard C/A-code receiver might have a quoted bandwidth of only ±2 MHz. Antennas for receivers that use more than the central lobe of the C/A-code spectrum or that use the encrypted P-code — whether full P(Y)-code or codeless or semicodeless tracking — need to have a wider bandwidth, say ±10 MHz. Dual-frequency, L1/L2 antennas typically have two patches, one for each frequency, each one of which has a bandwidth of about ±10 MHz. Back in GPS's early days, some dual-frequency receivers used conical spiral antennas whose bandwidths stretched all the way from the L2 to the L1 frequency.

Gain Pattern. Other important characteristics of a GPS antenna are its gain pattern, which describes its sensitivity over some range of elevation and azimuth angles; its ability to discriminate against multipath signals, that is, signals arriving at the antenna after being reflected off nearby objects; and, for antennas used in very precise positioning applications, the stability of its phase center — the antenna's electrical center, to which the position given by a GPS receiver actually refers.

A GPS antenna is typically omnidirectional. Such an antenna has an essentially nondirectional pattern in azimuth and a directional pattern in elevation angle. At the zenith, a flat microstrip patch, for example, might have a few dB of gain with respect to a circularly polarized isotropic radiator (dBic), a hypothetical ideal reference antenna. The gain gradually drops down to a few dB below that of a circularly polarized isotropic radiator at an elevation angle of 5 degrees or so. (It is possible to improve the microstrip patch's low-elevation-angle response by bending down the apexes of a polygonally-shaped patch to produce a three-dimensional dome-like structure.)

Ground Planes. Some antennas, such as the microstrip patch, require a ground plane to make them work properly. This is usually a flat or shaped piece of metal on which the actual antenna element sits. Up to a certain size, the larger the ground plane, the higher the antenna's zenith gain. For example, a microstrip patch element might have a zenith gain of only 0.75 dBic. By placing the element on a 10 × 10-centimeter ground plane, one can increase the gain to about 5 dBic.

In geodetic surveying, a metal plate or plates often further extend the antenna's ground plane to enhance its performance in the presence of multipath. This is done through beam shaping (reducing the device's gain at low elevation angles) and enhancing the attenuation of LHCP signals. On reflection, the polarization of RHCP signals changes. Depending on the nature of the reflector, the reflected signals can be linearly polarized or LHCP. An antenna designed to have a low sensitivity to LHCP signals offers some protection from single-bounce reflections.

One very useful form of ground plane is the choke ring. It consists of several concentric metal hoops, or thin-walled hollow cylinders, mounted on a circular base at the center of which is placed a microstrip patch antenna. Choke rings have proven particularly effective in reducing the effects of ground-bounce multipath.

One antenna not requiring a ground plane is the quadrifilar helix. This antenna is used, for example, with some popular handheld GPS receivers. It consists of two bifilar helical loops, orthogonally oriented on a common axis. As with the microstrip patch, this is an RHCP antenna with an omnidirectional pattern.

Depending on its construction, the quadrifilar helix can have a gain pattern similar to that of a flat microstrip patch or its low- to medium-elevation-angle gain can be enhanced at the expense of its zenith gain. This can be particularly useful in acquiring and tracking low-elevation-angle satellites, although the susceptibility to mulitpath is increased.

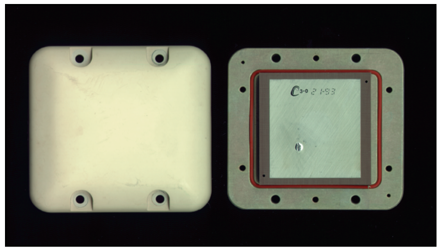
Sometimes, one wants the antenna gain pattern to have a null in a particular direction, say, in the direction of a jammer. Antennas have been developed that can be electronically steered to afford them an antijamming capability.

Phase-Center Variation. Ideally, a GPS antenna's electrical phase center is independent of a signal's direction of arrival. In practice, however, small (subcentimeter in the case of well-designed, geodetic-quality antennas) displacements of the phase center may occur with changing azimuth or elevation angle.

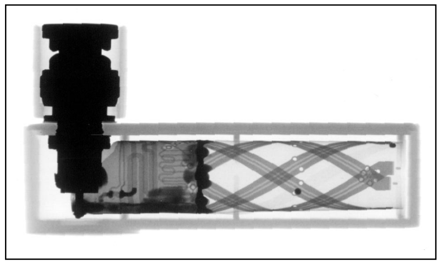
Antennas of the same make and model will typically show similar variations so that their effects in relative positioning can be minimized by orienting antennas on short baselines to the same direction, say magnetic north. For a well-designed antenna, the phase center's mean horizontal position usually coincides with the antenna's physical center. The phase center's vertical position with respect to an accessible physical plane through the antenna must be established by anechoic chamber measurements. Note that the L1 and L2 phase centers of dual-frequency antennas may be different.

Now, as long as one is using the same make and model of antenna at both ends of a baseline, the phase center's actual position is usually unimportant; only the vertical heights of a specific point on the exterior of the antennas (say on the preamplifier housing's base) above the geodetic markers must be measured. If one employs a mixture of antennas of different make and/or model on a baseline or in a network, however, then the data-processing software must know the heights of the antennas' mean phase centers with respect to the physical reference points on the antennas so that it can make the appropriate corrections.

The effects of the variation in phase center position on some geodetic positioning can be important. Observation site, length of observation session, use of ground planes, choice of elevation cut-off angle, antenna orientation, and frequency all can affect the an-tenna's estimated coordinates. The maximum sizes of the effects can range from a few millimeters to more than a centimeter.



TERRY ARSENAULT



GERHARD WOLTERS, DJ6YB

This (top) 6.2 × 6.2–centimeter microstrip patch GPS antenna is diagonally fed and housed in a protective radome.

This (bottom) x-ray of a quadrifilar helix GPS antenna shows its four, half-turn helices and feed structure, which are printed on flexible foil material. The antenna boresight, or zenith direction, is to the right.

Some users are applying azimuth and/or elevation angle–dependent phase center corrections in processing GPS data using different or widely spaced antennas.

For some applications, users may also need to consider the carrier-phase windup caused by rotating antennas.

Other Factors. There are a number of environmental factors that can affect a GPS antenna's performance and may need to be considered, depending on application. These include the effects of temperature, moisture (including high relative humidity), salt, vibration, and mechanical shock.

LOW NOISE PREAMP

Usually, a plastic housing (radome) — which is designed to minimally attenuate the signals — protects GPS antennas from possible damage by the elements or other means. As we know, GPS signals are very weak. The reason a GPS receiver does not need a large antenna has to do with the GPS signal's structure and the GPS receiver's ability to de-spread it.

So, the power to extract a GPS signal out of the ether's general background is concentrated in the receiver rather than the antenna. Nevertheless, one must generally combine a GPS antenna with a low-noise preamplifier that boosts the signal's level before feeding it to the receiver itself. In systems where the antenna is a separate unit, the preamplifier is housed in the base of the antenna and receives power from the same coaxial cable along which the signal travels to the receiver. Such devices are called active antennas.

The gain needed by a GPS antenna preamplifier depends on several factors, including the gain of the antenna element itself, cable-run length, and the requirements of the receiver's front end. Active antennas are available with gains of around 20, 26, 40, and even 50 dB, among others.

The preamplifier should have a low noise figure as well as high gain, as its noise figure has a dominating effect on the overall signal-to-noise level performance of the complete GPS receiving apparatus: Subsequent stages in the receiver will amplify the preamplifier noise. Typical noise figures range from about 1.2 to 2.5 dB. Although a GPS antenna's narrowband design helps to protect the preamplifier and the receiver from interfering signals, such as those from a nearby cellular telephone, the system can be further aided by employing filters either before or after the preamplifier.

Placing a filter after the preamplifier helps preserve the antenna assembly's low noise characteristics. Placing a filter ahead of the preamplifier, while likely increasing the noise figure of the antenna assembly, helps to protect the preamplifier from potentially overloading interference by preselecting the frequency band to be amplified.

Caution should be taken when using a GPS receiver with an active antenna not supplied by the receiver manufacturer. Not only must its noise figure and gain be within an acceptable range for proper receiver operation, but the voltage (including its polarity) and current supplied by the receiver to the antenna's preamplifier must be compatible with the antenna's requirements. An active antenna should include diode protection to prevent damage to the preamplifier from reverse or overvoltage connections.

TRANSMISSION LINES

As mentioned, the signals received by the antenna are typically passed to the receiver along a coaxial transmission line. The signals are attenuated with the degree of attenuation, referred to as insertion loss, dependent on the type and length of coaxial cable used. RG-58C cable has an insertion loss of about 0.8 dB per meter at a frequency of 1575 MHz. The thicker Belden 9913, on the other hand, has an insertion loss of only 0.2 dB per meter. Even lower loss cables are available. For long cable runs, one should use low-loss cable or place an additional preamplifier in line between the antenna and cable. These insertion losses assume a perfect match between the cable and antenna. As previously noted, a mismatch produces reflections in the cable, which increases signal loss.

A minor complication in connecting an antenna to a receiver is matching the cable connectors to the antenna and receiver connectors. A variety of connector types exist, including BNC, F, MCX, Type N, OSX, SMA, SMB, and TNC in both male and female varieties.

The signals traveling from the antenna to the receiver experience a small delay. This delay, however, is the same for the signals simultaneously received from different satellites and so acts like a receiver clock offset. In positioning applications, therefore, this delay is immaterial and is absorbed in the estimate of the clock offset or differenced away in between-receiver relative measurements. In timing applications, on the other hand, this delay must be carefully calibrated.

LOOSE ENDS

GPS signals suffer attenuation when they pass through most structures. Some antenna/receiver combinations are sensitive enough to work with signals received inside wooden-frame houses, on automobile dashboards, and in the window recesses of aircraft, for example, but it is generally recommended that antennas be mounted in the open air and with a clear view of the satellites. Even outdoors, dense foliage, particularly when it is wet, can attenuate the GPS signals sufficiently that receivers may have difficulty tracking them.

Two or more GPS receivers can share the same antenna if an antenna splitter is used. The splitter must block the preamplifier DC voltage supplied by all but one of the receivers. The splitter should provide a degree of isolation between the receiver ports so that no mutual interference between receivers occurs. Unless the splitter contains an active preamplifier, there will be at least a 3 dB loss each time the signal from the antenna is split.

CONCLUSION

Although the GPS receiver derives most of its amazing capabilities from its digital circuitry and firmware, it cannot begin to function until its antenna picks up the GPS signals. In this brief article, we have overviewed most of the important characteristics of this critical component of a GPS receiving system. Armed with this information, you should be better able to interpret antenna specifications and be better armed next time you go antenna shopping. ■

Further Reading

For a general introduction to the interesting but sometimes arcane subject of antennas by the American dean of antenna engineering, see

■ "Antennas: Our Electronic Eyes and Ears," by J.D. Kraus, in *Microwave Journal*, Vol. 32, No. 1, 1989, pp. 77–92.

For the arguably most-thorough discussion about the theory of antenna design and operation, see

■ *Antennas*, 2nd edition, by J.D. Kraus, published by McGraw-Hill, Inc., New York, 1988.

For a review of antenna design, including microstrip patch antennas, see

■ *Antenna Engineering Handbook*, 3rd edition, edited by R.C. Johnson and published by McGraw-Hill, Inc., New York, 1993.

For discussions about how antennas affect GPS observations, see

■ "How Different Antennas Affect the GPS Observable," by B.R. Schupler and T.A. Clark, in *GPS World*, Vol. 2, No. 10, 1991, pp. 32–36.

■ "Characterizations of GPS User Antennas: Reanalysis and New Results," by B.R. Schupler, T.A. Clark, and R.L. Allshouse, in *GPS Trends in Precise Terrestrial, Airborne, and Spaceborne Applications*, the Proceedings of the International Association of Geodesy Symposium No. 115, pp. 328–332, Springer-Verlag, Berlin, 1996.

For an overview of some low-cost GPS antennas, see

■ <http://callisto.worldonline.nl/~samsvl/>

For the National Geodetic Survey's measurements of the phase-center variations of commercially available, survey-quality GPS antennas, see

■ <http://www.grdl.noaa.gov/GRD/GPS/Projects/ANTCAL/>