## PDHonline Course L105D (12 PDH)

---

# GPS Surveying

*Instructor: Jan Van Sickle, P.L.S.*

**2020**

## PDH Online | PDH Center

5272 Meadow Estates Drive
Fairfax, VA 22030-6658
Phone: 703-988-0088
www.PDHonline.com

An Approved Continuing Education Provider

# Module 2

## Errors

Why doesn't a GPS receiver that costs a couple of hundred bucks deliver the best available accuracy to you? It's the error budget; sounds like you ought to be able to buy something with it, doesn't it? It's just a breakdown of the sources of errors affecting GPS positioning [the figures are in 2drms meters (95%)]:

| | | |
|---|---|---|
| *Selective Availability (SA)* | *46.5* | *(This error source has been eliminated, hooray!)* |
| Ionosphere | 13.6 | |
| Satellite clock and ephemeris | 7.0 | |
| Average DOP | 3.9 | |
| Receiver clock and noise | 2.9 | |
| Typical Multipath | 2.3 | |
| Troposphere | 1.3 | |
| TOTAL without SA | 31.0 | |

Let's look at these error sources one at a time.

### Selective Availability

First, a bit of history, I'll keep it short. The intentional dithering of the satellite clocks by the Department of Defense called Selective Availability, or SA, was instituted right after the first Block II GPS satellites were launched. The accuracy of the C/A point positioning was too good! The accuracy was supposed to be ±100 meters, horizontally, 95% of the time with a vertical accuracy of about ±175 meters.  But in fact, it turned out that the C/A-code point positioning gave civilians access to accuracy much better than that.  That wasn't according to plan, so they degraded the satellite clocks accuracy on the C/A code on purpose until ±100 meters was all you could get.  The good news is this error source is gone now!

Selective Availability was switched off on May 2, 2000 by presidential order.  The intentional degradation of the satellite clocks is a thing of the past.  To tell you the truth Selective Availability never did hinder the surveying application of GPS much anyway, more about that later.  But don't think that the satellite clocks don't contribute error to GPS positioning any more, they do.

Oh, by the way, remember the Navigation Code from the previous module? Instructions by which receivers can make some corrections for most of the errors discussed here are actually built into that Navigation message.  It is modulated onto carriers, L1 and L2.  But some of the information in the Navigation message can get outdated pretty quickly so it's renewed by government upload facilities around the world which are known, along with their tracking and computing counterparts, as the Control Segment.  The information sent to each satellite from the Control Segment makes its way through the satellites and back to the users in the NAV message.  In fact, there are new NAV messages coming into play.  There are four of them.  The content and format of the three new civil messages, L2-CNAV, CNAV-2, L5-CNAV and one military message, MNAV, are improved compared with the legacy NAV.  In general these NAV messages are more flexible and robust.  They are also transmitted at a higher rate than the legacy NAV, but back to error sources the first big error source is based on an effect of the ionosphere, that's still with us.

# Ionosphere

The GPS signal does just fine in space, but when it hits the atmosphere, oh boy. From about 50 km to 1000 km above the earth, the ionosphere is the first layer it comes to. This layer appears to delay the GPS signal. The magnitude of the delay depends on the density and stratification of the ionosphere when the signal passes through it. Actually it is the codes, the modulations on the carrier waves appear to be delayed, the carrier wave itself appears to be advanced.

The density of the ionosphere changes with the number and dispersion of free electrons released when gas molecules are ionized by the sun's ultraviolet radiation. This density is measured by something called the total electron content or TEC. That’s the number of free electrons in a column through the ionosphere with a cross-sectional area of 1 square meter. You see the ionosphere is pretty inconsistent. It changes from layer to layer, it changes with the time of day and even the season. During the daylight hours in the midlatitudes the ionospheric delay may be as much as five times greater than it is at night, and it’s usually least between midnight and early morning. When the earth is nearing its perihelion in November, that’s its closest approach to the sun, the ionospheric delay is nearly four times greater than it is in July near the earth's aphelion, the farthest point from the sun.

I’m sure all of this is wildly interesting, but now here is some really practical information. The severity of the ionospheric effect varies with the amount of time the GPS signal spends traveling through it. A signal originating from a satellite near the observer's horizon passes through more of the ionosphere than a signal coming in from a satellite straight overhead. The longer the signal is in the ionosphere, the greater the ionospheric effect. So if you want to avoid receiving GPS signals with severe delay you set a mask angle on your GPS receiver so it just ignores any signal within, say, 15° of the horizon.

How am I doing? Ok, here's some more practical information. The apparent ionospheric delay affects higher frequencies less than lower frequencies. This is called the ionosphere's dispersive property. That means that L1, 1575.42 MHz, is not affected as much as L2, 1227.60 MHz, and L2 is not affected as much as L5, 1176.45MHz. And right there is one of the greatest advantages of a multi-frequency receiver over the single-frequency receivers. By tracking all carriers, a multi-frequency receiver can remove not all, but a significant portion of the ionospheric error.

But as I mentioned there is an ionospheric correction available to the single frequency receiver in the Navigation message. The Control Segment's monitoring stations find the apparent delay by looking at the different propagation rates of the carrier frequencies. A correction is calculated and uploaded to the satellites to broadcast to GPS receivers. Well that's fine, but the atmosphere over Kwajalein in the Pacific probably isn't much like the atmosphere where you are, so this broadcast correction should not be expected to remove all of the ionospheric effect.

## Satellite Clock

GPS clocks keep GPS Time. The rate of the GPS Time is kept within 1 microsecond of Coordinated Universal Time, UTC, and UTC is determined by the more than 150 atomic clocks around the globe.

UTC is actually more stable than the rotation of the earth itself. Believe it or not, there is a discrepancy between UTC and the earth's actual motion, so leap seconds are put in once in a while to keep it from getting too far out of whack with the planet. But GPS doesn't use leap seconds, so UTC and GPS Time keep getting further apart. They started off together back on midnight January 5, 1990, since then many leap seconds have been added to UTC but none have been added to GPS Time. Confused yet? Ok, even though their <u>rates</u> are the same, the numbers expressing a particular instant in GPS Time are always different by some seconds from the numbers expressing the same instant in UTC.

To make it even more interesting each GPS satellite carries its own onboard clocks in the form of very stable and accurate atomic clocks regulated by the vibration frequencies of the atoms of two elements. Onboard clocks are regulated by cesium or rubidium. Since the clocks in any one satellite are completely independent from those in any other, they are allowed to drift up to one millisecond from the strictly controlled GPS Time standard. That might seem a little strange at first, but the alternative would be to have the Control Segment constantly tweaking the satellite's onboard clocks. That is the only way they could keep them all in lockstep with each other and with GPS Time. Instead, their individual drifts are carefully monitored. And the government stations record each satellite clock's deviation from GPS Time. That drift is uploaded into each satellite's Navigation message, it is known as the broadcast clock correction.

In other words, there are three kinds of time are involved here. The first is UTC per the United States Naval Observatory (USNO). The second is GPS time. The third is the time determined by each independent GPS satellite.

Here is how they work together. There is a Master Control Station (MCS) at Schriever (formerly Falcon) Air Force Base near Colorado Springs, Colorado gathers the GPS satellites' data from monitoring stations around the world. After processing, this information is uploaded back to each satellite to become the broadcast clock correction.

The actual specification for GPS Time demands that it be within one microsecond of UTC as determined by USNO, without consideration of leap seconds. In practice, GPS Time is much closer to UTC than the microsecond specification; it is usually within about 40 nanoseconds of UTC, minus leap seconds. The system also makes sure that the time broadcast by each independent satellite in the GPS constellation is no farther than one millisecond from GPS Time. But the drift of each satellite's clock is not constant, nor can the broadcast clock correction be updated frequently enough to completely define the drift. So the satellite clocks make a contribution to the errors in a GPS point position.

Now, there is one more issue regarding the GPS satellite clocks you might have thought of already, relativistic effects. Albert Einstein's special and general theories of relativity predicted that a clock in orbit around the earth would appear to run faster than a clock on its surface. And they do indeed, due to their greater speed and the weaker gravity around them, the clocks in the GPS satellites do appear to run faster than the clocks in GPS receivers. There are actually two parts to the effect.

Concerning the first part, time dilation is taken into account before the satellite's clocks are sent into orbit. To ensure the clocks will actually achieve the correct fundamental frequency of 10.23 MHz in space, their frequency is set a bit slow before launch to 10.22999999545 MHz.

The second part is attributable to the eccentricity of the orbit of GPS satellites. The orbital effect can be as much as 45.8 nanoseconds. Fortunately, the offset is eliminated by a calculation in the GPS receiver itself, thereby avoiding a ranging error of about 14 meters. In other words, both relativistic effects on the satellite clocks can be accurately computed and are removed from the system, so don't fret.

## Ephemeris

Remember that the satellites are the control points of the system. If you didn't know where they are, ranges to them wouldn't be of much use. Since they are constantly moving an ephemeris is the best way to define their location at a particular instant. It is very much like using an ephemeris to calculate the position of the sun at a particular moment of time. For GPS satellites the ephemeris information is contained in the Navigation message. It is called the broadcast ephemeris and it has all the information the user's receiver needs to calculate earth-centered, earth-fixed coordinates of any GPS satellite at any moment. But the broadcast ephemeris is far from perfect.

It is given in a right ascension, RA, system of coordinates. There are six orbital elements, they are: the semimajor axis of the orbit, the eccentricity of the orbit, the right ascension of its ascending node, the inclination of its plane, the argument of its perigee and the true anomaly. Now these parameters appear Keplerian, named for the 17th century German astronomer Johannes Kepler. But in this case, they really aren't.

The orbital motion of GPS satellites is subject to a bunch of disturbing forces, for example, the non-spherical nature of the earth's gravity, the attractions of the sun and the moon, and solar radiation pressure. Actually the best way to know what all these forces are doing to the satellites is to watch the motion of the satellites themselves. That's why government facilities distributed around the world, the Control Segment, carefully track the satellites and using least squares and curve-fitting analysis they produce the broadcast ephemeris from they data they collect.

This might be a good time to say just a bit more about the Control Segment. As I've said before, the Master Control Station *MCS* is located at Schriever (formerly Falcon) Air Force Base in Colorado Springs, Colorado. The 2nd Space Operations Squadron mans the station. They compute updates for the Navigation message, generally, and the broadcast ephemeris, in particular, based on about one week of tracking information they collect from monitoring stations around the world.

It's a good thing that the Control Segment exists, because the GPS system requires constant maintenance. Orbital and clock adjustments and other data uploads are necessary to keep the constellation from degrading. Due to recent upgrades in the system every GPS satellite is tracked by at least three monitoring stations at all times and the orbital tracking data gathered by monitoring stations are then passed on to the Master Control Station. There, new ephemerides are computed. This tabulation of the anticipated locations of the satellites with respect to time is then transferred to four uploading stations, where it is transmitted back to the satellites themselves.

## DOP

### Dilution of Precision

Here's the question, "Are the satellites crowded together in one part of the sky, or are they spread out?" If they're crowded together the DOP, dilution of precision, number is high and that's bad. If they're spread out the DOP number is low and that's good. In other words, this number is like the strength of figure consideration in the design of a network. DOP is all about the geometric strength of the described by the positions of the satellites with respect to one another (Figure 2.1).

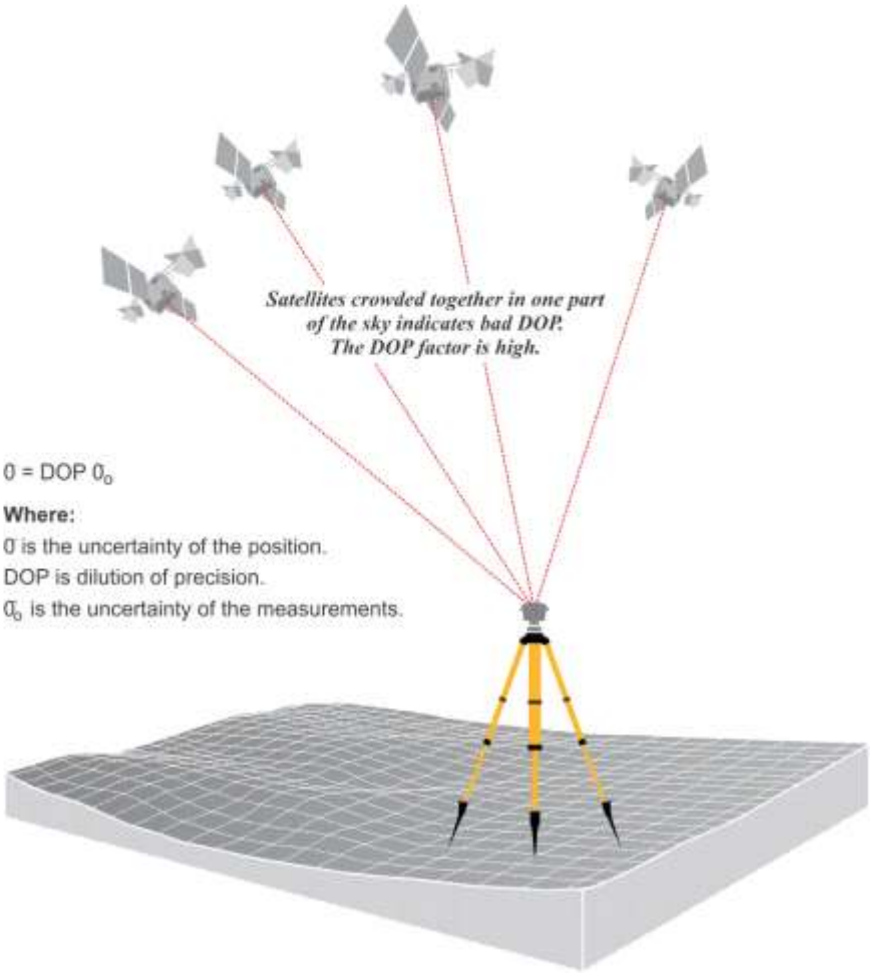

Figure 2.1

Four or more satellites must be above the observer's mask angle for the simultaneous solution of the clock error and three dimensions of the receiver’s position. But if all of those satellites are crowded together in one part of the sky, it’s not going to work very well.

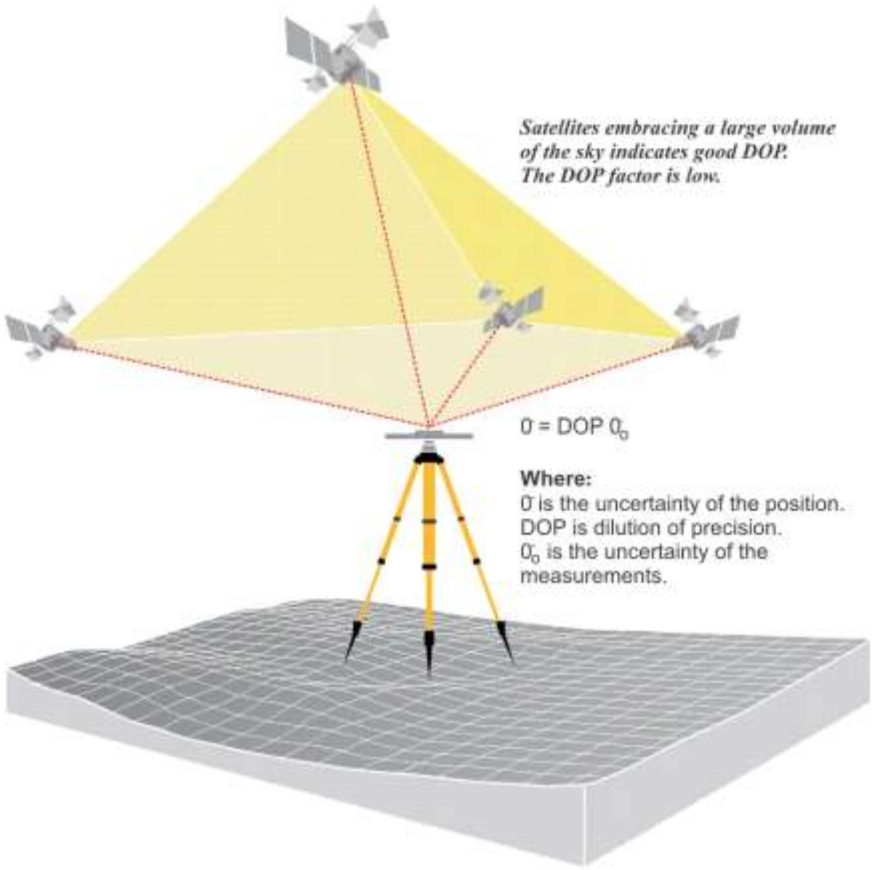

Figure 2.2

The larger the volume of the body defined by the lines from the receiver to the satellites, the better the satellite geometry and the lower the DOP (Figure 2.2). An ideal arrangement of four satellites would be one directly above the receiver, the others 120 degrees from one another in azimuth near, but not too close, to the horizon. With that distribution the DOP would be nearly 1, the lowest possible value. In practice, the lowest DOPs are generally around 2.

There are many DOP factors used to evaluate the uncertainties in the components of a receiver's position. For example, there is horizontal dilution of precision *(HDOP)* and vertical dilution of precision *(VDOP)* where the uncertainty of a solution for positioning has been isolated into its horizontal and vertical components, respectively. When both horizontal and vertical components are combined, the uncertainty is called *PDOP*, position dilution of precision. There is also *TDOP*, time dilution of precision, that indicates only the clock offset; and *RDOP*, relative dilution of precision, that includes the number of receivers, the number of satellites they can handle, the length of the observing session as well as the geometry of the satellites configuration.

When a DOP factor exceeds a maximum limit in a particular location, indicating an unacceptable level of uncertainty exists over a period of time, that period is known as an *outage*. Of course, since the satellites are always moving an outage of this kind is temporary.

## The Receiver Clock

A receiver's measurement of phase differences and its generation of replica codes is only as good as its clock that is its oscillator. You can think of it as the internal frequency standard for a receiver.

GPS receivers are usually equipped with quartz crystal clocks. They’re relatively inexpensive and compact. They have low power requirements and long life spans. These clocks work by the piezoelectric effect in an oven-controlled quartz crystal disk. You’ll see this type of clock symbolized by *OCXO* sometimes. Their reliability is about equal to a quarter of a second over a human lifetime. Even so they are sensitive to temperature changes, shock, and vibration.

## Multipath

Multipath occurs when part of the signal from the satellite reaches the receiver after reflecting from the ground, a building, or another object. These reflected signals interfere with the signal that reaches the receiver directly from the satellite (Figure 2.3).

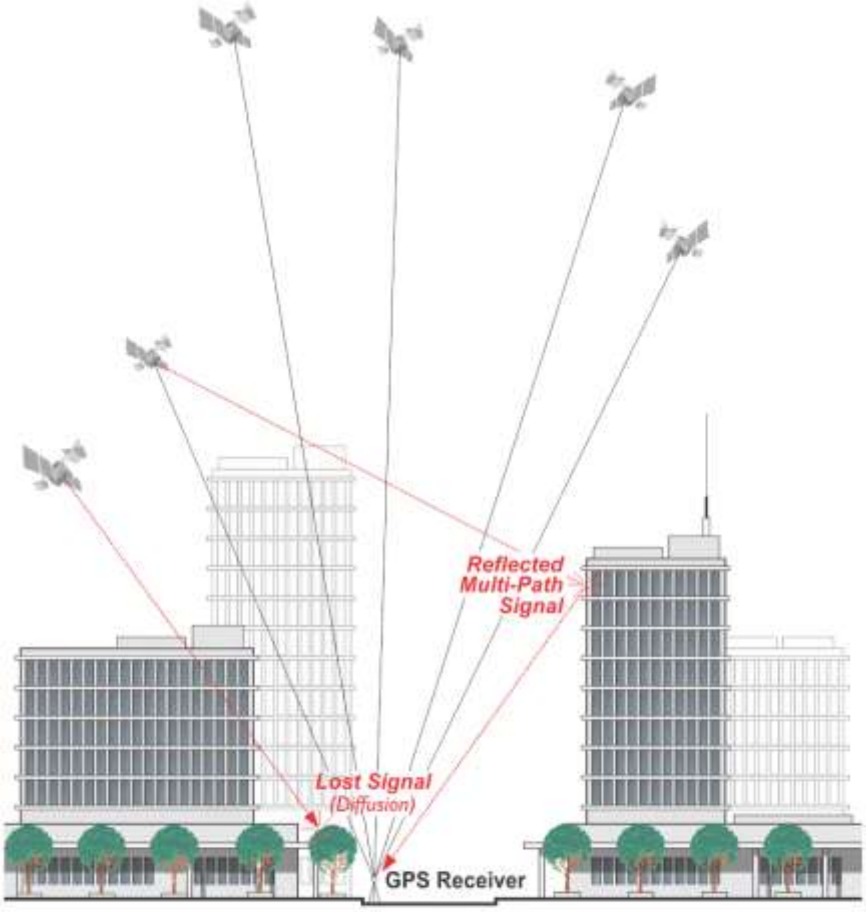

Figure 2.3

The high frequency of the GPS codes tends to limit the field over which multipath can contaminate pseudorange observations. Once a receiver has achieved lock; that is, its replica code is correlated with the incoming signal from the satellite; signals outside the expected chip length can be rejected.

There are other factors that distinguish reflected multipath signals from direct signals. For example, reflected signals at the frequencies used for the carriers tend to be more diffuse than the directly received signals. Another difference involves the circular polarization of the GPS signal. The polarization is actually reversed when the signal is reflected. These characteristics allow some multipath signals to be identified and rejected at the receiver’s antenna.

GPS antenna design can play a role in minimizing the effect of multipath. Ground planes, usually metal sheets, are used with many antennas to reduce multipath interference by eliminating signals from low elevation angles.

Choke ring antennas, based on a design first introduced by the Jet Propulsion Laboratory (JPL), can reduce antenna gain at low elevations. This design contains a series of concentric circular troughs that are a bit more than a quarter of a wavelength deep. When a GPS signal’s wavefront arrives at the edge of an antenna’s ground plane from below it can induce a surface wave on the top of the plane that travels horizontally. A choke ring antenna can prevent the formation of these surface waves.

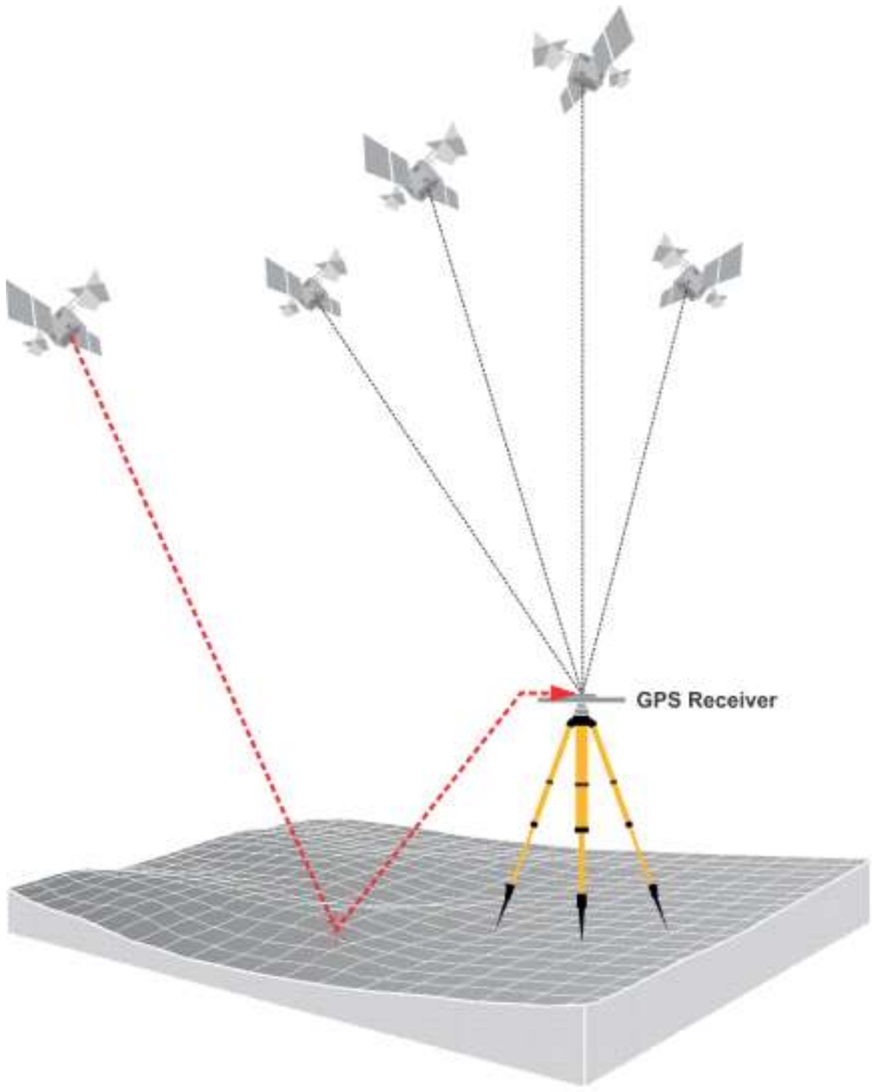

Figure 2.4

But neither ground planes nor choke rings mitigate the effect of reflected signals from above the antenna very effectively. There are signal processing techniques that can reduce multipath, but when the reflected signal originates less than a few meters from the antenna, this approach is not as effective.

One of the best ways to limit multipath is the 15 degree cut-off or *mask angle*. This idea was mentioned in limiting the effect of the ionosphere. Tracking satellites only after they are more than 15 degrees above the receiver’s horizon limits multipath too.

## The Troposphere

The troposphere is that part of the atmosphere closest to the earth. Including all its layers it extends up to about 50 km above the surface.

Like the ionosphere, the troposphere appears to delay the GPS signal too. But the troposphere is *electrically neutral*; meaning it is neither ionized nor dispersive for frequencies below 30 GHz. In other words, the delay of a GPS satellite’s signal in the troposphere has nothing to do with its frequency. Therefore, both the carriers are equally refracted.

The density of the troposphere does govern the severity of its effect on the GPS signal. Once again a satellite close to the horizon will be more delayed than a signal from a satellite at zenith.

Modeling the troposphere is one technique used to reduce the bias in GPS data processing, and it can be up to 95 percent effective. However, the residual 5 percent can be quite difficult to remove. Refraction in the troposphere has a dry component and a wet component. The dry component is closely correlated to the atmospheric pressure and can be more easily estimated than the wet component. It is fortunate that the dry component contributes the larger portion of range error in the troposphere since the high cost of water vapor radiometers and radiosondes generally restricts their use to only the most high-precision GPS work.

## Answering the Question

At the top I asked, “Why doesn’t a GPS receiver that costs a couple of hundred bucks deliver the best available accuracy to you?” The answer is this. Each of the errors mentioned here, along with a few more, are fully present in the code pseudorange point positioning such receivers offer. They are mitigated a little by the corrections available from the Navigation message, but high accuracy is just not in the cards with point positioning. High accuracy begins with Relative, also known as differential GPS.

Relative GPS involves the use of two or more GPS receivers simultaneously observing the same satellites. This approach attains much higher accuracy than point positioning because of the extensive correlation of errors. It is not that all the errors mentioned here are not present at all; it is that they are virtually the same for each of the receivers. For example, consider signals traveling from four satellites to three receivers that are close together, and please consider that even distances normally be considered large are short compared with the 20,000-km altitude of the GPS satellites. These three receivers are operating simultaneously and are collecting signals from the same satellites. They will record errors, yes, but the same errors. For example, The signal from each satellite would pass through virtually the same atmosphere on its way to each receiver. The ionospheric delay will be present, but it will be almost identical for each particular signal when it arrives at each receiver. It is at this point that we can begin to talk about centimeter, and even millimeter accuracy, more about that in the next module.