How GPS Works
Extracted from the USFS Website
http://www.fs.fed.us/database/gps/aboutgps/begincor.htm#top
February 21, 2008

Trying to figure out where you are and where you're going is probably one of man's oldest pastimes. Navigation and positioning are crucial to so many activities and yet the process has always been quite cumbersome. Over the years, all kinds of technologies have tried to simplify the task but every one has had some disadvantage.

Landmarks:	Only work in local area. Subject to movement or destruction by environmental factors.
Dead Reckoning:	Very complicated. Accuracy depends on measurement tools which are usually relatively crude. Errors accumulate quickly.
Celestial:	Complicated. Only works at night in good weather. Limited precision.
OMEGA:	Based on relatively few radio direction beacons. Accuracy limited and subject to radio interference.
LORAN:	Limited coverage (mostly coastal). Accuracy variable, affected by geographic situation. Easy to jam or disturb.
SatNav:	Based on low-frequency doppler measurements so it's sensitive to small movements at receiver. Few satellites so updates are infrequent.

Finally, the U.S. Department of Defense decided that the military had to have a super precise form of worldwide positioning. Fortunately they had the kind of money ($12 billion!) it took to build something really good.

Why did the Department of Defense develop GPS?
In the latter days of the arms race the targeting of ICBMs became such a fine art that they could be expected to land right on an enemy's missile silos. Such a direct hit would destroy the silo and any missile in it. The ability to take out your opponent's missiles had a profound effect on the balance of power. But you could only expect to hit a silo if you knew exactly where you were launching from. That's not hard if your missiles are on land, as most of them were in the Soviet Union. Most of the U.S. nuclear arsenal was at sea on subs. To maintain the balance of power the U.S. had to come up with a way to allow those subs to surface and fix their exact position in a matter of minutes anywhere in the world. The result is the Global Positioning System, a system that's changed navigation forever.

The Global Positioning System (GPS) is a worldwide radio-navigation system formed from a constellation of 24 satellites and their ground stations. GPS uses these "man-made stars" as reference points to calculate positions accurate to a matter of meters. In fact, with advanced forms of GPS you can make measurements to better than a centimeter!

Here's how GPS works in five logical steps:
1. The basis of GPS is "triangulation" from satellites.
2. To "triangulate," a GPS receiver measures distance using the travel time of radio signals.
3. To measure travel time, GPS needs very accurate timing which it achieves with some tricks.
4. Along with distance, you need to know exactly where the satellites are in space. High orbits and careful monitoring are the secret.
5. Finally, you must correct for any delays the signal experiences as it travels through the atmosphere.

Improbable as it may seem, the whole idea behind GPS is to use satellites in space as reference points for locations here on earth.

By very, very accurately measuring our distance from three satellites we can " triangulate " our position anywhere on earth. Forget for a moment how our receiver measures this distance. We'll get to that later. First consider how distance measurements from three satellites can pinpoint you in space.

The Big Idea Geometrically:
Suppose we measure our distance from a satellite and find it to be 11,000 miles.
Knowing that we're 11,000 miles from a particular satellite narrows down all the possible locations we could be in the whole universe to the surface of a sphere that is centered on this satellite and has a radius of 11,000 miles. Next, say we measure our distance to a second satellite and find out that it's 12,000 miles away. That tells us that we're not only on the first sphere but we're also on a sphere that's 12,000 miles from the second satellite. Or in other words, we're somewhere on the circle where these two spheres intersect.
If we then make a measurement from a third satellite and find that we're 13,000 miles from that one, that narrows our position down even further, to the two points where the 13,000 mile sphere cuts through the circle that's the intersection of the first two spheres.

By ranging from three satellites we can narrow our position to just two points in space.
To decide which one is our true location we could make a fourth measurement. But usually one of the two points is a ridiculous answer (either too far from Earth or moving at an impossible velocity) and can be rejected without a measurement. A fourth measurement does come in very handy for another reason however, but we'll tell you about that later.

Next, we'll see how the system measures distances to satellites.
But how can you measure the distance to something that's floating around in space? We do it by timing how long it takes for a signal sent from the satellite to arrive at our receiver.

The Big Idea Mathematically
In a sense, the whole thing boils down to those “velocity times travel time” math problems we did in high school. Remember the old: “If a car goes 60 miles per hour for two hours, how far does it travel?”

Velocity (60 mph) x Time (2 hours) = Distance (120 miles)

In the case of GPS we're measuring a radio signal so the velocity is going to be the speed of light or roughly 186,000 miles per second. The problem is measuring the travel time. The timing problem is tricky. First, the times are going to be awfully short. If a satellite were right overhead the travel time would be something like 0.06 seconds. So, we're going to need some really precise clocks. We'll talk about those soon. But, assuming we have precise clocks, how do we measure travel time? To explain it let's use a goofy analogy: Suppose there was a way to get both the satellite and the receiver to start playing "The Star Spangled Banner" at precisely 12 noon. If sound could reach us from space (which, of course, is ridiculous) then standing at the receiver we'd hear two versions of the Star Spangled Banner, one from our receiver and one from the satellite.

These two versions would be out of sync. The version coming from the satellite would be a little delayed because it had to travel more than 11,000 miles. If we wanted to see just how delayed the satellite's version was, we could start delaying the receiver's version until they fell into perfect sync. The amount we have to shift back the receiver's version is equal to the travel time of the satellite's version. We just multiply that time times the speed of light and BINGO! we've got our distance to the satellite. That's basically how GPS works. Only instead of the Star Spangled Banner the satellites and receivers use something called a "Pseudo Random Code" - which is probably easier to sing than the Star Spangled Banner.The Pseudo Random Code (PRC, shown above) is a fundamental part of GPS. Physically it's just a very complicated digital code, or in other words, a complicated sequence of "on" and "off" pulses as shown here: The signal is so complicated that it almost looks like random electrical noise. Hence the name "Pseudo-Random."

GPS Signals in Detail:

Carriers
The GPS satellites transmit signals on two carrier frequencies. The L1 carrier is 1575.42 MHz and carries both the status message and a pseudo-random code for timing. The L2 carrier is 1227.60 MHz and is used for the more precise military pseudo-random code.

Pseudo-Random Codes
There are two types of pseudo-random code (see tutorial for explanation of pseudo random codes in general). The first pseudo-random code is called the C/A (Coarse Acquisition) code. It modulates the L1 carrier. It repeats every

1023 bits and modulates at a 1MHz rate. Each satellite has a unique pseudo-random code. The C/A code is the basis for civilian GPS use.
The second pseudo-random code is called the P (Precise) code. It repeats on a seven day cycle and modulates both the L1 and L2 carriers at a 10MHz rate. This code is intended for military users and can be encrypted. When it's encrypted it's called "Y" code. Since P code is more complicated than C/A it's more difficult for receivers to acquire. That's why many military receivers start by acquiring the C/A code first and then move on to P code.

Navigation Message There is a low frequency signal added to the L1 codes that gives information about the satellite's orbits, their clock corrections and other system status.
There are several good reasons for that complexity: First, the complex pattern helps make sure that the receiver doesn't accidentally sync up to some other signal. The patterns are so complex that it's highly unlikely that a stray signal will have exactly the same shape.
Since each satellite has its own unique Pseudo-Random Code this complexity also guarantees that the receiver won't accidentally pick up another satellite's signal. All the satellites can use the same frequency without jamming each other. And it makes it more difficult for a hostile force to jam the system. In fact, the Pseudo Random Code gives the DoD a way to control access to the system.

Encrypted GPS

GPS was developed by the Defense Department primarily for military purposes. And even though it's been estimated that there are ten times as many civilian receivers as military ones the system still has considerable military significance.
To that end the military maintains exclusive access to the more accurate "P-code" pseudo random code. It's ten times the frequency of the civilian C/A code (and so potentially much more accurate) and much harder to jam. When it's encrypted it's called "Y-code" and only military receivers with the encryption key can receive it. Because this code is modulated on two carriers, sophisticated games can be played with the frequencies to help eliminate errors caused by the atmosphere. But there's another reason for the complexity of the Pseudo Random Code, a reason that's crucial to making GPS economical. The codes make it possible to use "information theory" to " amplify " the GPS signal. And that's why GPS receivers don't need big satellite dishes to receive the GPS signals.

In Review: **Measuring Distance**

1. Distance to a satellite is determined by measuring how long a radio signal takes to reach us from that satellite.
2. To make the measurement we assume that both the satellite and our receiver are generating the same pseudo-random codes at exactly the same time.
3. By comparing how late the satellite's pseudo-random code appears compared to our receiver's code, we determine how long it took to reach us.
4. Multiply that travel time by the speed of light and you've got distance

If measuring the travel time of a radio signal is the key to GPS, then our stop watches had better be darn good, because if their timing is off by just a thousandth of a second, at the speed of light, that translates into almost 200 miles of error! On the satellite side, timing is almost perfect because they have incredibly precise atomic clocks on board.

Atomic Clocks

Atomic clocks don't run on atomic energy. They get the name because they use the oscillations of a particular atom as their "metronome." This form of timing is the most stable and accurate reference man has ever developed. But what about our receivers here on the ground? Remember that both the satellite and the receiver need to be able to precisely synchronize their pseudo-random codes to make the system work. If our receivers needed atomic clocks (which cost upwards of $50K to $100K) GPS would be a lame duck technology. Nobody could afford it. Luckily the designers of GPS came up with a brilliant little trick that lets us get by with much less accurate clocks in our receivers. This trick is one of the key elements of GPS and as an added side benefit it means that every GPS receiver is essentially an atomic-accuracy clock.

Using GPS for Timing

We generally think of GPS as a navigation or positioning resource but the fact that every GPS receiver is synchronized to universal time makes it the most widely available source of precise time. This opens up a wide range of applications beyond positioning. GPS is being used to synchronize computer networks, calibrate other navigation systems, synchronize motion picture equipment and much more.The secret to perfect timing is to make an *extra* satellite measurement. That's right, if three perfect measurements can locate a point in 3-dimensional space, then four *imperfect* measurements can do the same thing.

Extra Measurement Cures Timing Offset
If our receiver's clocks were perfect, then all our satellite ranges would intersect at a single point (which is our position). But with imperfect clocks, a fourth measurement, done as a cross-check, will NOT intersect with the first three. So the receiver's computer says "Uh-oh! There is a discrepancy in my measurements. I must not be perfectly synced with universal time."

Since any offset from universal time will affect all of our measurements, the receiver looks for a single correction factor that it can subtract from all its timing measurements that would cause them all to intersect at a single point. That correction brings the receiver's clock back into sync with universal time, and bingo! - you've got atomic accuracy time right in the palm of your hand. Once it has that correction it applies to all the rest of its measurements and now we've got precise positioning. One consequence of this principle is that any decent GPS receiver will need to have at least four channels so that it can make the four measurements simultaneously. With the pseudo-random code as a rock solid timing sync pulse, and this extra measurement trick to get us perfectly synced to universal time, we have got everything we need to measure our distance to a satellite in space.
But for the triangulation to work we not only need to know distance, we also need to know exactly where the satellites are.

In Review: **Getting Perfect Timing**
1. Accurate timing is the key to measuring distance to satellites.
2. Satellites are accurate because they have atomic clocks on board.
3. Receiver clocks don't have to be too accurate because an extra satellite range measurement can remove errors.

In this tutorial we've been assuming that we know where the GPS satellites are so we can use them as reference points. But how do we know *exactly* where they are? After all they're floating around 11,000 miles up in space.

A high satellite gathers no moss.
That 11,000 mile altitude is actually a benefit in this case, because something that high is well clear of the atmosphere. And that means it will orbit according to very simple mathematics. The Air Force has injected each GPS satellite into a very precise orbit, according to the GPS master plan.

GPS Master Plan
The launch of the 24th block II satellite in March of 1994 completed the GPS constellation. Four additional satellites are in reserve to be launched "on need." The spacings of the satellites are arranged so that a minimum of five satellites are in view from every point on the globe. On the ground all GPS receivers have an almanac programmed into their computers that tells them where in the sky each satellite is, moment by moment.
The basic orbits are quite exact but just to make things perfect the GPS satellites are constantly monitored by the Department of Defense. {These stations monitor the GPS satellites, checking both their operational health and their exact position in space. The master ground station transmits corrections for the satellite's ephemeris constants and clock offsets back to the satellites themselves. The satellites can then incorporate these updates in the signals they send to GPS receivers.
There are five monitor stations: Hawaii, Ascension Island, Diego Garcia, Kwajalein, and Colorado Springs. They use very precise radar to check each satellite's exact altitude, position and speed. The errors they're checking for are called "ephemeris errors" because they affect the satellite's orbit or "ephemeris." These errors are caused by gravitational pulls from the moon and sun and by the pressure of solar radiation on the satellites. The errors are usually very slight but if you want great accuracy they must be taken into account.

Getting the message out
Once the DoD has measured a satellite's exact position, they relay that information back up to the satellite itself. The satellite then includes this new corrected position information in the timing signals it's broadcasting. So a GPS signal

is more than just pseudo-random code for timing purposes. It also contains a navigation message with ephemeris information as well. With perfect timing and the satellite's exact position you'd think we'd be ready to make perfect position calculations. But there's trouble afoot. Check out the next section to see what's up.

In Review: **Satellite Positions**

1. To use the satellites as references for range measurements we need to know exactly where they are.
2. GPS satellites are so high up their orbits are very predictable.
3. Minor variations in their orbits are measured by the Department of Defense.
4. The error information is sent to the satellites, to be transmitted along with the timing signals.

Up to now we've been treating the calculations that go into GPS very abstractly, as if the whole thing were happening in a vacuum. But in the real world there are lots of things that can happen to a GPS signal that will make its life less than mathematically perfect.

To get the most out of the system, a good GPS receiver needs to take a wide variety of possible errors into account. Here's what they've got to deal with.

First, one of the basic assumptions we've been using throughout this tutorial is not exactly true. We've been saying that you calculate distance to a satellite by multiplying a signal's travel time by the speed of light. But the speed of light is only constant in a vacuum.

As a GPS signal passes through the charged particles of the ionosphere and then through the water vapor in the troposphere it gets slowed down a bit, and this creates the same kind of error as bad clocks.

There are a couple of ways to minimize this kind of error. For one thing we can predict what a typical delay might be on a typical day. This is called modeling and it helps but, of course, atmospheric conditions are rarely exactly typical.

Error Modeling

Much of the delay caused by a signal's trip through our atmosphere can be predicted.

Mathematical models of the atmosphere take into account the charged particles in the ionosphere and the varying gaseous content of the troposphere.

On top of that, the satellites constantly transmit updates to the basic ionospheric model.

A GPS receiver must factor in the angle each signal is taking as it enters the atmosphere because that angle determines the length of the trip through the perturbing medium.}

Another way to get a handle on these atmosphere-induced errors is to compare the relative speeds of two different signals. This "dual frequency" measurement is very sophisticated and is only possible with advanced receivers.

Dual Frequency Measurements

Physics says that as light moves through a given medium, low-frequency signals get "refracted" or slowed more than high-frequency signals. By comparing the delays of the two different carrier frequencies of the GPS signal, L1 and L2, we can deduce what the medium (i.e. atmosphere) is, and we can correct for it. Unfortunately, this requires a very sophisticated receiver since only the military has access to the signals on the L2 carrier. Civilian companies have worked around this problem with some tricky strategies. Unfortunately, they're so secret if we told you how they work we'd have to kill you.

Trouble for the GPS signal doesn't end when it gets down to the ground. The signal may bounce off various local obstructions before it gets to our receiver.

This is called multipath error and is similar to the ghosting you might see on a TV. Good receivers use sophisticated signal rejection techniques to minimize this problem.

Multipath error

The whole concept of GPS relies on the idea that a GPS signal flies straight from the satellite to the receiver.

Unfortunately, in the real world the signal will also bounce around on just about everything in the local environment and get to the receiver that way too.

The result is a barrage of signals arriving at the receiver: first the direct one, then a bunch of delayed reflected ones.

This creates a messy signal. If the bounced signals are strong enough they can confuse the receiver and cause erroneous measurements. Sophisticated receivers use a variety of signal processing tricks to make sure that they only consider the earliest arriving signals (which are the direct ones).}

Problems at the satellite

Even though the satellites are very sophisticated they do account for some tiny errors in the system.
The atomic clocks they use are very, very precise but they're not perfect. Minute discrepancies can occur, and these translate into travel time measurement errors. And even though the satellites positions are constantly monitored, they can't be watched every second. So slight position or " ephemeris" errors can sneak in between monitoring times.

Ephemeris errors

Ephemeris (or orbital) data is constantly being transmitted by the satellites. Receivers maintain an "almanac" of this data for all satellites and they update these almanacs as new data comes in.
Typically, ephemeris data is updated hourly. }
Basic geometry itself can magnify these other errors with a principle called "Geometric Dilution of Precision" or GDOP. It sounds complicated but the principle is quite simple. There are usually more satellites available than a receiver needs to fix a position, so the receiver picks a few and ignores the rest. If it picks satellites that are close together in the sky the intersecting circles that define a position will cross at very shallow angles. That increases the gray area or error margin around a position. If it picks satellites that are widely separated the circles intersect at almost right angles and that minimizes the error region. Good receivers determine which satellites will give the lowest GDOP.

Intentional Errors!

As hard as it may be to believe, the same government that spent $12 billion to develop the most accurate navigation system in the world intentionally degraded its accuracy. The policy was called "Selective Availability" or "SA" and the idea behind it was to make sure that no hostile force or terrorist group can use GPS to make accurate weapons.

Basically the DoD introduced some "noise" into the satellite's clock data which, in turn, added noise (or inaccuracy) into position calculations. The DoD may have also been sending slightly erroneous orbital data to the satellites which they transmitted back to receivers on the ground as part of a status message. Together these factors made SA the biggest single source of inaccuracy in the system. Military receivers used a decryption key to remove the SA errors and so they're much more accurate.

Turning Off Selective Availability

On May 1, 2000 the White House announced a decision to discontinue the intentional degradation of the GPS signals to the public beginning at midnight. Civilian users of GPS are now able to pinpoint locations up to ten times more accurately. As part of the 1996 Presidential Decision Directive goals for GPS, President Clinton committed to discontinuing the use of SA by 2006. The announcement came six years ahead of schedule. The decision to discontinue SA was the latest measure in an on-going effort to make GPS more responsive to civil and commercial users worldwide.

The bottom line

Fortunately all of these inaccuracies still don't add up to much of an error. And a form of GPS called "Differential GPS" can significantly reduce these problems. We'll cover this type of GPS later. To get an idea of the impact of these errors see a typical error budget:

Summary of GPS Error Sources

Typical Error in Meters	(per satellite)	
	Standard GPS	Differential GPS
Satellite Clocks	1.5	0
Orbit Errors	2.5	0
Ionosphere	5.0	0.4
Troposphere	0.5	0.2
Receiver Noise	0.3	0.3
Multipath	0.6	0.6

SA	30	0
Typical Position Accuracy		
Horizontal	50	1.3
Vertical	78	2.0
3-D	93	2.8

In Review: **Correcting Errors**

1. The earth's ionosphere and atmosphere cause delays in the GPS signal that translate into position errors.
2. Some errors can be factored out using mathematics and modeling.
3. The configuration of the satellites in the sky can magnify other errors.
4. Differential GPS can eliminate almost all error.

Why We Need Differential GPS:

Basic GPS is the most accurate radio-based navigation system ever developed. And for many applications it's plenty accurate. But it's human nature to want MORE!

So some crafty engineers came up with "Differential GPS," a way to correct the various inaccuracies in the GPS system, pushing its accuracy even farther.

Differential GPS or "DGPS" can yield measurements good to a couple of meters in moving applications and even better in stationary situations.

That improved accuracy has a profound effect on the importance of GPS as a resource. With it, GPS becomes more than just a system for navigating boats and planes around the world. It becomes a universal measurement system capable of positioning things on a very precise scale.

Differential GPS involves the cooperation of two receivers, one that's stationary and another that's roving around making position measurements. The stationary receiver is the key. It ties all the satellite measurements into a solid local reference.

Here's how it works:

The problem

Remember that GPS receivers use timing signals from at least four satellites to establish a position. Each of those timing signals is going to have some error or delay depending on what sort of perils have befallen it on its trip down to us. Since each of the timing signals that go into a position calculation has some error, that calculation is going to be a compounding of those errors.

An extenuating circumstance

Luckily, the sheer scale of the GPS system comes to our rescue. The satellites are so far out in space that the little distances we travel here on earth are insignificant. So if two receivers are fairly close to each other, say within a few hundred kilometers, the signals that reach both of them will have traveled through virtually the same slice of atmosphere, and so will have virtually the same errors.That's the idea behind differential GPS: We have one receiver measure the timing errors and then provide correction information to the other receivers that are roving around. That way virtually all errors can be eliminated from the system, even the pesky Selective Availability error that the DoD puts in on purpose.The idea is simple. Put the reference receiver on a point that's been very accurately surveyed and keep it there.

This reference station receives the same GPS signals as the roving receiver but instead of working like a normal GPS receiver it attacks the equations ***backwards***. Instead of using timing signals to calculate its position, it uses its known position to calculate timing. It figures out what the travel time of the GPS signals *should* be, and compares it with what they actually *are*. The difference is an "error correction" factor.

The receiver then transmits this error information to the roving receiver so it can use it to correct its measurements. Since the reference receiver has no way of knowing which of the many available satellites a roving receiver might be using to calculate its position, the reference receiver quickly runs through all the visible satellites and computes each of their errors.

Then it encodes this information into a standard format and transmits it to the roving receivers.
It's as if the reference receiver is saying: "OK everybody, right now the signal from satellite #1 is ten nanoseconds delayed, satellite #2 is three nanoseconds delayed, satellite #3 is sixteen nanoseconds delayed...." and so on.
The roving receivers get the complete list of errors and apply the corrections for the particular satellites they're using.

In the early days of GPS, reference stations were established by private companies with big projects demanding high accuracy - groups like surveyors or oil drilling operations. And that is still a very common approach. You buy a reference receiver and set up a communication link with your roving receivers.

But now there are enough public agencies transmitting corrections that you might be able to get them for free!
The United States Coast Guard and other international agencies are establishing reference stations all over the place, especially around popular harbors and waterways.
These stations often transmit on the radio beacons that are already in place for radio direction finding (usually in the 300kHz range).

Anyone in the area can receive these corrections and radically improve the accuracy of their GPS measurements.
Most ships already have radios capable of tuning the direction finding beacons, so adding DGPS will be quite easy.
Many new GPS receivers are being designed to accept corrections, and some are even equipped with built-in radio receivers.

Post Processing DGPS
Not all DGPS applications are created equal. Some don't need the radio link because they don't need precise positioning immediately. It's one thing if you're trying to position a drill bit over a particular spot on the ocean floor from a pitching boat, but quite another if you just want to record the track of a new road for inclusion on a map.
For applications like the later, the roving receiver just needs to record all of its measured positions and the exact time it made each measurement. Then later, this data can be merged with corrections recorded at a reference receiver for a final clean-up of the data. So you don't need the radio link that you have to have in real-time systems.
If you don't have a reference receiver there may be alternative source for corrections in your area. Some academic institutions are experimenting with the Internet as a way of distributing corrections.

There's another permutation of DGPS, called "inverted DGPS," that can save money in certain tracking applications.
Let's say you've got a fleet of buses and you'd like to pinpoint them on street maps with very high accuracy (maybe so you can see which side of an intersection they're parked on or whatever).
Anyway, you'd like this accuracy but you don't want to buy expensive "differential-ready" receivers for every bus.
With an inverted DGPS system the buses would be equipped with standard GPS receivers and a transmitter and would transmit their standard GPS positions back to the tracking office. Then at the tracking office the corrections would be applied to the received positions. It requires a computer to do the calculations, a transmitter to transmit the data but it gives you a fleet of very accurate positions for the cost of one reference station, a computer and a lot of standard GPS receivers. Such a deal! If you want to know where DGPS might be headed, take a look at your hand, because soon DGPS may be able to resolve positions that are no farther apart than the width of your little finger.
Imagine the possibilities. Automatic construction equipment could translate CAD drawings into finished roads without any manual measurements. Self-guided cars could take you across town while you quietly read in the back seat.

To understand how this kind of GPS is being developed you need to understand a little about GPS signals. If two receivers are fairly close to each other, say within a few hundred kilometers, the signals that reach both of them will have traveled through virtually the same slice of atmosphere, and so will have virtually the same line.
The words "Code-Phase" and "Carrier-Phase" may sound like electronic mumbo-jumbo but, in fact, they just refer to the particular signal that we use for timing measurements. Using the GPS carrier frequency can significantly improve the accuracy of GPS. The concept is simple but to understand it let's review a few basic principles of GPS.
Remember that a GPS receiver determines the travel time of a signal from a satellite by comparing the "pseudo random code" it's generating, with an identical code in the signal from the satellite. The receiver slides its code later and later in time until it syncs up with the satellite's code. The amount it has to slide the code is equal to the signal's travel time. The problem is that the bits (or cycles) of the pseudo random code are so wide that even if you do get synced up there's still plenty of slop.

Survey receivers beat the system by starting with the pseudo random code and then move on to measurements based on the carrier frequency for that code. This carrier frequency is much higher so its pulses are much closer together and therefore more accurate. If you're rusty on the subject of carrier frequencies consider your car radio. When you tune to 94.7 on the dial you're locking on to a carrier frequency that's 94.7 MHz.

Obviously, we can't hear sounds at 94 million cycles a second. The music we hear is a modulation (or change) in this carrier frequency. So when you hear someone sing an "A" note on the radio you're actually hearing the 94.7 MHz carrier frequency being varied at a 440 cycle rate.
GPS works in the same way. The pseudo random code has a bit rate of about 1 MHz but its carrier frequency has a cycle rate of over a GHz (which is 1000 times faster!) At the speed of light the 1.57 GHz GPS signal has a wavelength of roughly twenty centimeters, so the carrier signal can act as a much more accurate reference than the pseudo random code by itself. And if we can get to within one percent of perfect phase like we do with code-phase receivers we'd have 3 or 4 millimeter accuracy.

In essence, this method is counting the exact number of carrier cycles between the satellite and the receiver.
The problem is that the carrier frequency is hard to count because it's so uniform. Every cycle looks like every other. The pseudo random code on the other hand is intentionally complex to make it easier to know which cycle you're looking at. So the trick with "carrier-phase GPS" is to use code-phase techniques to get close. If the code measurement can be made accurate to say, a meter, then we only have a few wavelengths of carrier to consider as we try to determine which cycle really marks the edge of our timing pulse.

Resolving this "carrier phase ambiguity" for just a few cycles is a much more tractable problem and as the computers inside the receivers get smarter and smarter it's becoming possible to make this kind of measurement without all the ritual that surveyors go through. You've got to hand it to the FAA. They think big!
They realized the great benefits GPS could bring to aviation, but they wanted ***more***. They wanted the accuracy of Differential GPS and they wanted it across the whole continent. Maybe the whole world. Their plan is called the "Wide Area Augmentation System" or "WAAS," and it's basically a *continental* DGPS system.
The idea grew out of some very specific requirements that basic GPS just couldn't handle by itself. It began with "system integrity." GPS is very reliable but every once in a while a GPS satellite malfunctions and gives inaccurate data.

The GPS monitoring stations detect this sort of thing and transmit a system status message that tells receivers to disregard the broken satellite until further notice. Unfortunately, this process can take many minutes which could be too late for an airplane in the middle of a landing. So the FAA got the idea that they could set up their own monitoring system that would respond much quicker. In fact, they figured they could park a geosynchronous satellite somewhere over the U.S. that would instantly alert aircraft when there was a problem. Then they reasoned that they could transmit this information right on a GPS channel so aircraft could receive it on their GPS receivers and wouldn't need any additional radios. But wait a second! If we've got the geosynchronous satellite already transmitting on the GPS frequency, why not use it for positioning purposes too? Adding another satellite helps with positioning accuracy and it ensures that plenty of satellites are always visible around the country. But wait another second! Why not use that satellite to relay differential corrections too?

The FAA figured that with about 24 reference receivers scattered across the U.S. they could gather pretty good correction data for most of the country. That data would make GPS accurate enough for "Category 1" landings (i.e. very close to the runway but not zero visibility) This system is underway. Specifications have been drafted and approved and it's expected that the system could be working as early as 1997. The ramifications of this go well beyond aviation, because the system guarantees that DGPS corrections will be raining out of the sky for everyone to use. To complete the system the FAA wants to eventually establish "Local Area Augmentation Systems" near runways.
These would work like the WAAS but on a smaller scale. The reference receivers would be near the runways and so would be able to give much more accurate correction data to the incoming planes.
With a LAAS aircraft would be able to use GPS to make Category 3 landings. (zero visibility)

"Where is everything else?"

It's a big world out there, and using GPS to survey and map it precisely saves time and money in this most stringent of all applications. Today, Trimble GPS makes it possible for a single surveyor to accomplish in a day what used to take weeks with an entire team. And they can do their work with a higher level of accuracy than ever before. Trimble pioneered the technology which is now the method of choice for performing control surveys, and the effect on surveying in general has been considerable. You've seen how GPS pinpoints a position, a route, and a fleet of vehicles. Mapping is the art and science of using GPS to locate items, then create maps and models of everything in the world. And we do mean everything. Mountains, rivers, forests and other landforms. Roads, routes, and city streets. Endangered animals, precious minerals and all sorts of resources. Damage and disasters, trash and archeological treasures. GPS is mapping the world.

For example, Trimble GPS helped fire fighters respond with speed and efficiency during the 1991 Oakland/Berkeley fire to plot the extent of the blaze and to evaluate damage. In a less urgent yet equally important situation, the city of Modesto, California improved their efficiency and job performance by using GPS and mountain bikes to create a precise map of its network of water resources and utilities.

The Ideas Behind the Jargon:

Anywhere fix The ability of a receiver to start position calculations without being given an approximate location and approximate time.

Bandwidth The range of frequencies in a signal.

C/A code The standard (Course/Acquisition) GPS code. A sequence of 1023 pseudo-random, binary, biphase modulations on the GPS carrier at a chip rate of 1.023 MHz. Also known as the "civilian code."

Carrier A signal that can be varied from a known reference by modulation.

Carrier-aided tracking a signal processing strategy that uses the GPS carrier signal to achieve an exact lock on the pseudo random code.

Carrier frequency The frequency of the unmodulated fundamental output of a radio transmitter.

Carrier phase GPS GPS measurements based on the L1 or L2 carrier signal.

Channel A channel of a GPS receiver consists of the circuitry necessary to receive the signal from a single GPS satellite.

Chip The transition time for individual bits in the pseudo-random sequence. Also, an integrated circuit. Also a snack food. Also a betting marker.

Clock bias The difference between the clock's indicated time and true universal time.

Code phase GPS measurements based on the pseudo random code (C/A or P) as opposed to the carrier of that code.

Control segment A world-wide network of GPS monitor and control stations that ensure the accuracy of satellite positions and their clocks.

Cycle slip A discontinuity in the measured carrier beat phase resulting from a temporary loss of lock in the carrier tracking loop of a GPS receiver.

Data message A message included in the GPS signal which reports the satellite's location, clock corrections and health. Included is rough information on the other satellites in the constellation.

Differential positioning Accurate measurement of the relative positions of two receivers tracking the same GPS signals.

Dilution of Precision The multiplicative factor that modifies ranging error. It is caused solely by the geometry between the user and his set of satellites. Known as DOP or GDOP

Dithering The introduction of digital noise. This is the process the DoD uses to add inaccuracy to GPS signals to induce Selective Availability.

Doppler-aiding A signal processing strategy that uses a measured doppler shift to help the receiver smoothly track the GPS signal. Allows more precise velocity and position measurement.

Doppler shift The apparent change in the frequency of a signal caused by the relative motion of the transmitter and receiver.

Ephemeris The predictions of current satellite position that are transmitted to the user in the data message.

Fast switching channel A single channel which rapidly samples a number of satellite ranges. "Fast" means that the switching time is sufficiently fast (2 to 5 milliseconds) to recover the data message.

Frequency band A particular range of frequencies.

Frequency spectrum The distribution of signal amplitudes as a function of frequency.

Geometric Dilution of Precision (GDOP) See Dilution of Precision.

Hardover word The word in the GPS message that contains synchronization information for the transfer of tracking from the C/A to P code.

Ionosphere The band of charged particles 80 to 120 miles above the Earth's surface.

Ionospheric refraction

The change in the propagation speed of a signal as it passes through the ionosphere.

L-band The group of radio frequencies extending from 390 MHz to 1550 MHz. The GPS carrier frequencies (1227.6 MHz and 1575.42 MHz) are in the L band.

Multipath error Errors caused by the interference of a signal that has reached the receiver antenna by two or more different paths. Usually caused by one path being bounced or reflected.

Multi-channel receiver A GPS receiver that can simultaneously track more than one satellite signal.

Multiplexing channel A channel of a GPS receiver that can be sequenced through a number of satellite signals.

P-code The Precise code. A very long sequence of pseudo random binary biphase modulations on the GPS carrier at a chip rate of 10.23 MHz which repeats about every 267 days. Each one week segment of this code is unique to one GPS satellite and is reset each week.

Precise Positioning Service (PPS) The most accurate dynamic positioning possible with standard GPS, based on the dual frequency P-code and no SA.

Pseudolite A ground-based differential GPS receiver which transmits a signal like that of an actual GPS satellite, and can be used for ranging.

Pseudo random code A signal with random noise-like properties. It is a very complicated but repeating pattern of 1's and O's.

Pseudorange A distance measurement based on the correlation of a satellite transmitted code and the local receiver's reference code, that has not been corrected for errors in synchronization between the transmitter's clock and the receiver's clock.

Satellite constellation The arrangement in space of a set of satellites.

Selective Availability (SA) A policy adopted by the Department of Defense to introduce some intentional clock noise into the GPS satellite signals thereby degrading their accuracy for civilian users. This policy was discontinued as of May 1, 2000 and now SA is turned off

Slow switching channel A sequencing GPS receiver channel that switches too slowly to allow the continuous recovery of the data message.

Space segment The part of the whole GPS system that is in space, i.e. the satellites.

Spread spectrum A system in which the transmitted signal is spread over a frequency band much wider than the minimum bandwidth needed to transmit the information being sent. This is done by modulating with a pseudo random code, for GPS.

Standard Positioning Service (SPS) The normal civilian positioning accuracy obtained by using the single frequency C/A code.

Static positioning Location determination when the receiver's antenna is presumed to be stationary on the Earth. This allows the use of various averaging techniques that improve accuracy by factors of over 1000.

User interface The way a receiver conveys information to the person using it. The controls and displays.

User segment The part of the whole GPS system that includes the receivers of GPS signals.