CHAPTER 9. GLOBAL POSITIONING SYSTEM

SECTION I. CONFIGURATION, SIGNALS, AND CODE

The Navigation Satellite Timing and Ranging (NAVSTAR) GPS is configured into space, control, and user segments. Each segment depends upon the other. See figure 9-1.

Originally, the complete space segment was to consist of 24 Block II satellites. Block I satellites were considered developmental. Block IIR (Replacement) satellites are being developed to provide system operations through the year 2025. More than 24 satellites are currently in orbit. They are arranged into six orbital planes, each inclined 55° from the Equator. Each orbital plane contains at least four unevenly spaced satellites orbiting the Earth twice a day at an altitude of 10,898 miles. Satellites move continuously through their orbit in the same direction as the Earth's rotation. They orbit the earth twice in 23 hours, 56 minutes, and 04.091 seconds solar time or 1 24-hour sidereal day.

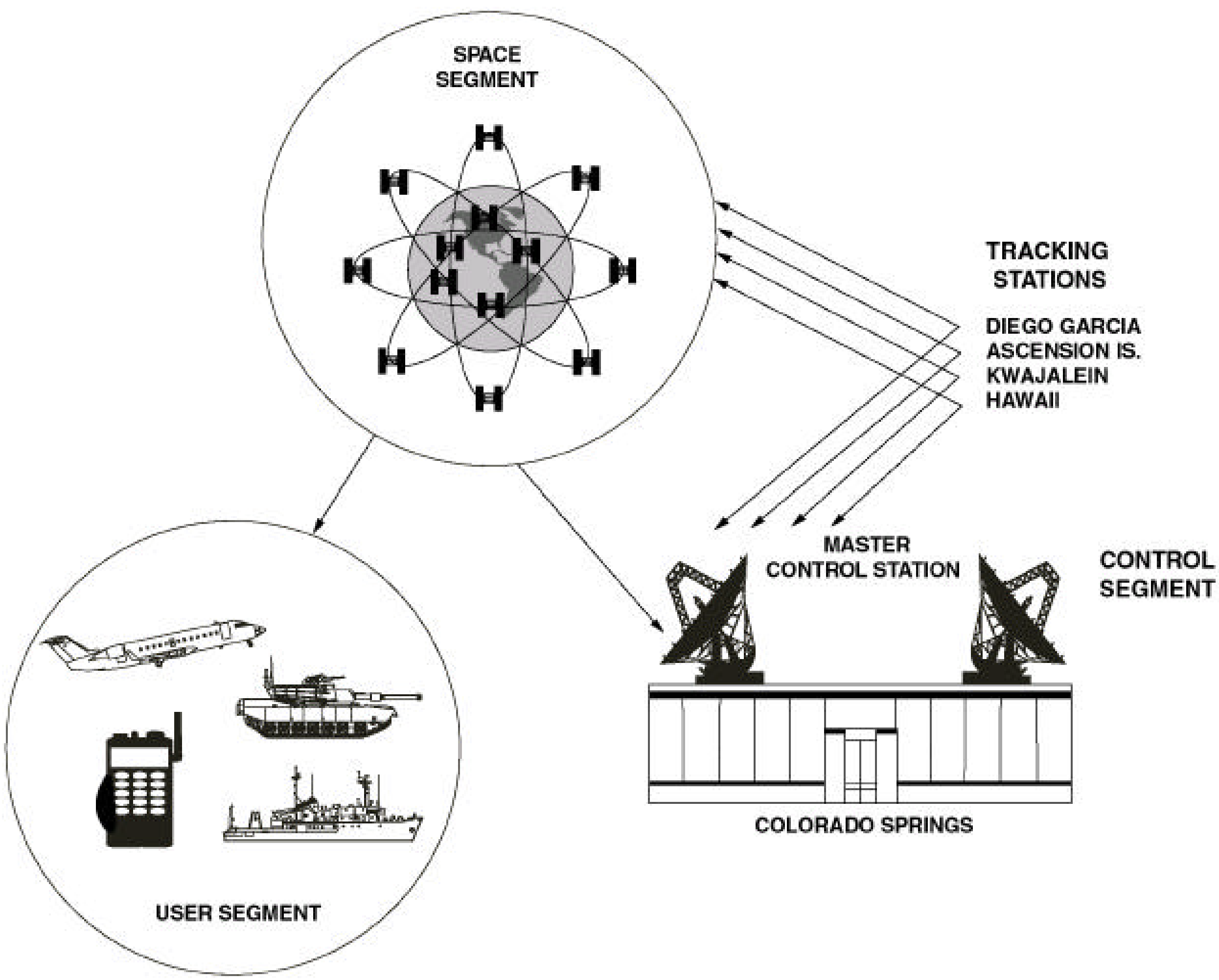

Figure 9-1. NAVSTAR System Configuration: Space, Control, and User Segments.

The control segment consists of five unmanned monitor/tracking stations located in Hawaii, Colorado, Ascension Island, Diego Garcia, and Kwajalein. Tracking stations use special receivers to track each satellite individually. The information from tracking the satellites helps control the satellites and predict their orbits. Three of the stations transmit information back to the satellites. All data collected at the tracking stations is transmitted to the master control station, located at Colorado Springs, Colorado where it is processed and analyzed. Ephemerides, clock corrections, and other message data are then transmitted back to the three stations for subsequent transmittal back to the satellites. The master control station is also responsible for the daily management and control of the satellites and the overall control segment.

The user segment consists of any one with a GPS receiver. Military and civilian personnel (including the enemy) use these receivers.

Satellite Signals

See figures 9-2 through 9-5.

Carrier Frequencies

Each GPS satellite broadcasts signals on two spread-spectrum radio frequencies (RFs). These are termed carrier frequencies because they are modulated with signal codes "carried" on the radio wave. The satellite's onboard atomic clocks generate a

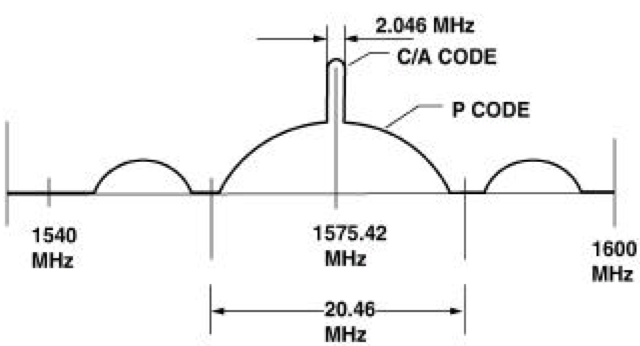

Figure 9-2. L1 Carrier Frequency.

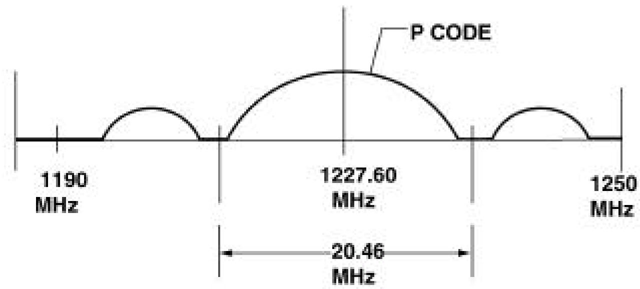

Figure 9-3. L2 Carrier Frequency.

fundamental frequency of 10.23 megahertz (10,230,000 cycles per second) multiplied by a factor that produces the actual carrier frequency.

The Link 1 (L1) RF carrier frequency is generated by multiplying the fundamental frequency by 154. It is centered at 1575.42 megahertz and has a bandwidth of 20.46 megahertz. The majority of the intensity of the signal lies at 1575.42 megahertz (±10.23 megahertz). Signal wavelength is 19 centimeters.

The Link 2 (L2) RF carrier frequency is generated by multiplying the fundamental frequency by 120. It is centered at 1227.60 megahertz and has a bandwidth of 20.46 megahertz. The majority of the intensity of the signal lies at 1227.60 megahertz (±10.23 megahertz). Signal wavelength is 24 centimeters.

Data Sequences

Each GPS satellite develops several binary data sequences transmitted from the GPS control segment. These sequences are the coarse/acquisition (C/A) code, the precise (P) code, and the Navigation Data Message (Nav Data).

The CA code is sometimes referred to as the standard (S) code. It has also been called the clear access or civilian access code. It is broadcast by all GPS satellites on the L1 carrier wave. Transmission of the data sequence is centered at 1575.42 megahertz (L1 frequency). It is modulated at ±1.023 megahertz providing a bandwidth of 2.046 megahertz. The code contains a sequence of 1,023 pseudo-random binary biphase modulations on the GPS carrier at a chipping rate of 1.023 megahertz, thus having a repetition period of 1 microsecond. The C/A code is a 300-meter measurement wave.

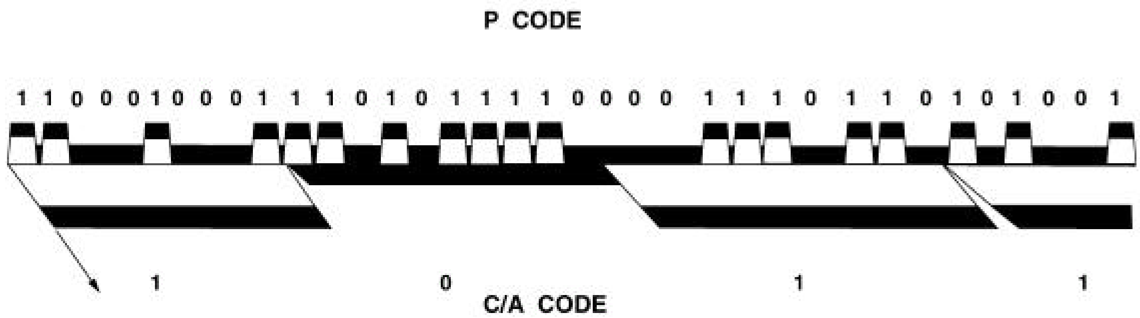

Figure 9-4. P Code and C/A Code Data Sequence.

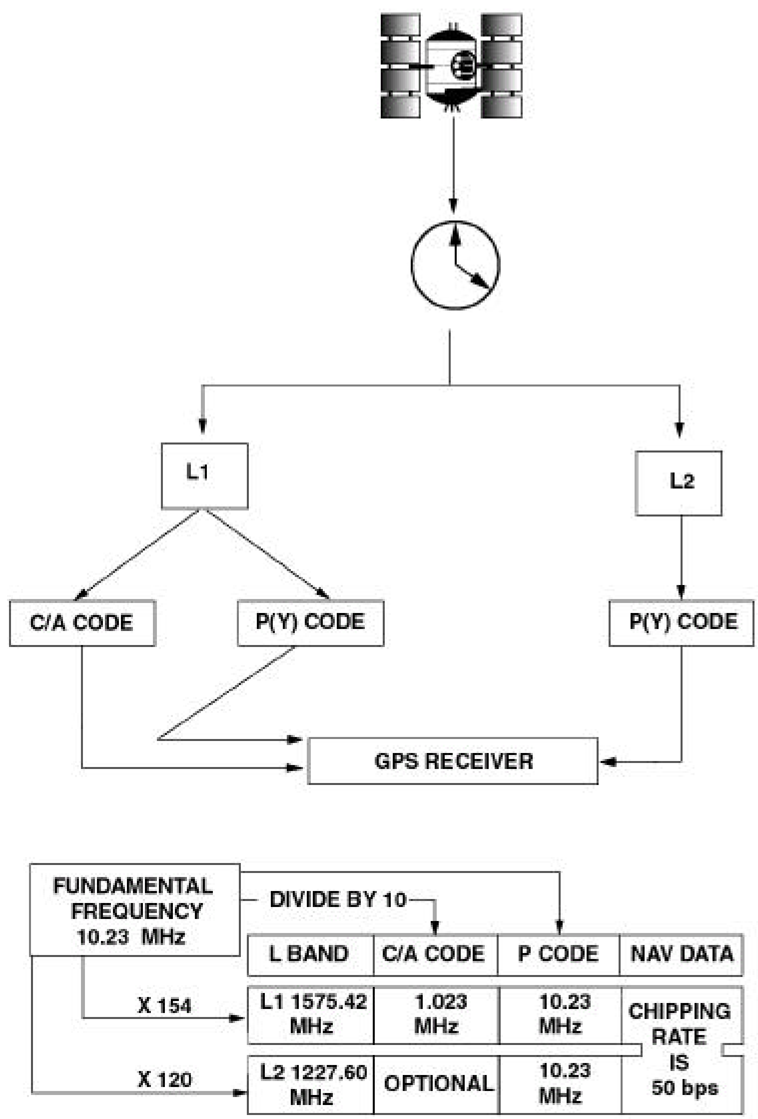

Figure 9-5. GPS Signal Data Flow.

The P code is sometimes referred to as the protected code. It is broadcast by all GPS satellites on both the L1 and L2 carrier. Transmission of the data sequence is centered at 1575.42 megahertz on the L1 carrier and at 1227.60 megahertz on the L2 carrier. It modulates at ±10.23 megahertz on carrier frequencies providing a bandwidth of 20.46 megahertz. The P code is a 30-meter measurement wave and can be encrypted by the satellite creating a Y code.

The overall P code is a mathematically derived binary sequence that is 267 days (approximately 37 weeks) long. It is broken into 1-week segments for operational use. Five of these 1-week segments are reserved for the GPS control segment. The other 32 segments are available for satellite vehicles (SVs). Each SV has a unique 1-week segment code that is a subset of the overall P code sequence. It is generally accepted that the P code repeats every week.

The Nav Data is a 1,500 bit navigation message broadcast on both L1 and L2 carriers at a rate of 50 bits per second or 50 hertz. The Nav Data contains system time, clock correction parameters, ionospheric delay model parameters, and the almanac (ephemeris and health data) on the entire constellation. It is broadcast once each hour by each GPS satellite and can be referred to as the D code.

The Nav Data is a separate binary data sequence in the satellite. But it is modulated over the C/A and P codes for transmission. Each satellite develops the binary code sequence of the C/A code, the P code, and the Nav Data. Since these sequences are all 1's and 0's, the satellite combines the Nav Data with the C/A code and P code to form two data streams: one a combination of the C/A code and the Nav Data, the other a combination of the P code and the Nav Data. These codes are actually transmitted on the carrier frequencies. The ground receiver then extracts the Nav Data from the broadcast C/A or P code, whichever it was receiving.

Ephemeris and Almanac Data

Each GPS satellite transmits almanac data once an hour. The almanac data is the position and health status of all satellites in the constellation. The ephemeris is the position data for each individual satellite. There are two types of ephemeris data to be considered: broadcast and precise.

The broadcast ephemerides are actually predicted satellite positions transmitted as part of the Nav Data. Ephemerides can be acquired in real time by any receiver capable of acquiring the C/A or P codes. Broadcast ephemerides are computed by the master control station using past tracking data provided from the five tracking stations. The new orbital parameters are then transmitted back to the satellites once every 24 hours for subsequent transmission to the user segment.

Precise ephemerides are computed from actual tracking data post-processed to obtain more accurate satellite positions. Precise ephemerides are available later and are more accurate than broadcast ephemerides because they are based on actual tracking data and not predicted data.

SECTION II. SYSTEM SAFEGUARDS, ERROR SOURCES, AND SURVIVABILITY

Selective Availability

When the concept of GPS was initially developed, it was planned that the P code would be reserved for military use, while the less accurate C/A code would be authorized for use by anyone. During initial system testing, it was discovered that while P code measurements provided the expected 10-20 meter accuracy the C/A code provided accuracy as high as 20 meters (much better than the expected 100 meters). The DOD, expecting much lower accuracy, determined that a method of introducing errors into the satellite signals was needed to ensure that enemy forces would not be able to obtain high position and timing accuracy from GPS. Selective availability (SA) was the outcome.

Methods

The DOD uses SA to deny precise position and timing accuracy to unauthorized users. SA uses two methods to intentionally introduce errors into the signals transmitted to the user segment.

The dither method alters or manipulates the satellite clocks. This method intentionally introduces timing errors, which ultimately produces position errors at the receiver because of the importance of accurate time to the computation of the pseudo-range.

The epsilon method alters the orbital parameters (satellite position) that are broadcast in the Almanac portion of the Nav Data. Position error is then created because the receiver is positioned based on the satellite location.

Accuracy Levels

The level of accuracy achieved by a GPS receiver now depends on if the receiver is equipped with an encryption device that allows the receiver to accept and store crypto variables referred to as a key. This key allows the receiver to decrypt SA correction data that is transmitted in the Nav Data message. This key also allows the receiver to use the encrypted P code. The two accuracy levels are the precise positioning service (PPS) and the standard positioning service (SPS).

The PPS is a precise positioning and timing service that is reserved for the US and allied military, as well as, specific authorized civilian users as long as their receiver accepts the crypto key discussed above. The technical specification is listed at 16 meters spherical error probable (SEP).

The SPS is the less accurate positioning and timing service offered to all GPS users. The DOD has stated that this service will be accurate to 100 meters in horizontal position and 150 meters vertical, under normal conditions, 95% of the time. The DOD does have the ability to increase the errors created by SA based upon national security needs.

Antispoofing

Antispoofing (S) is a method used by the DOD to prevent possible hostile imitations of the GPS signal. Encrypting the P code creates the Y code, which can only be processed by GPS receivers with a valid crypto key. This encrypted code is very difficult to imitate. It is important to understand that the P code and the Y code are not two separate codes; one is the encrypted version of the other. A GPS receiver without a valid crypto key cannot process the Y Code and will be limited to measurements from the C/A code.

Cryptovariables

As stated before, cryptovariables are necessary for a GPS receiver to access the Precise Positioning Service, allowing the receiver to correct for errors caused by SA and AS. It is unauthorized to use a military GPS receiver without a valid crypto fill.

There are three types of key materials available for use with a GPS receiver: operational, maintenance, and simulator. There are three formats for key loading: the KOI-18 General Purpose Tape Reader, the KYK-13 Electronic Transfer Device (capable of loading multiple keys), and the AN/CYZ-10 Data Transfer Device (DTD).

Operational Key Material

Two operational cryptographic keys (group unique variable [GUV] key and cryptovariable weekly [CVW] key) are available for issue to a GPS user. Both keys can be used by a receiver to obtain a daily cryptovariable key (CVd). All operational keys are classified CONFIDENTIAL and are marked CRYPTO.

The GUV key is an annual key. It is a key encryption key (KEK) that decrypts previously encrypted daily keys. A GPS receiver loaded with a GUV key takes longer to begin processing navigational data than the weekly key because it must first acquire and decrypt the CVd key being broadcast by any GPS satellite. This process could take as long as 12.5 minutes after initial GPS signal acquisition. The GUV key is not a years worth of daily keys. It is merely the data needed by the receiver to decrypt the broadcast daily key.

The cryptovariable weekly (CVW) key is sometimes referred to as the crypto key weekly (CKW). It is a key production key (KPK) that automatically generates daily keys within the user equipment. Obtaining the daily key from a satellite downlink is not necessary for receivers loaded with a CVW key. A user with a CVW key starts processing navigational data in less time than those with a GUV key. Because of this special capability, distribution of the CVW key is limited to those users who demonstrate a valid need for initial GPS acquisition in a minimal amount of time.

The GUV key and the CVW key will produce the same CVd key. Once the receiver determines the current working CVd, processing navigational data may commence and the effects of SA and AS can be removed from the GPS signal and full navigational accuracy is restored. A CVd cannot be entered directly into the receiver, only a CVW or a GUV key can be entered. The same CVd is used for both SA and AS.

Maintenance Key Material

A maintenance key is available to users for troubleshooting GPS user equipment. It does not allow a user to gain access to the daily encryption key. Maintenance keys are unclassified and may be reused until they are unusable.

Simulator Key Material

A simulator key is available to users for testing receivers. The simulator and the equipment must be keyed with the simulator key. The simulator key does not allow a user to gain access to the daily encryption key. Simulator keys are unclassified and may be reused until they are physically unusable.

User Equipment Security

Because of the security classification of the CVW and GUV keys, GPS receivers designed specifically for military uses are equipped with a special certified security module that prevents the extraction of cryptographic information from the receiver. These receivers can then remain unclassified even when loaded with a cryptographic key. If classified information other than crypto is stored in the receiver, the receiver becomes classified at the level of the stored information.

GPS Error Sources

There are many sources of measurement error that influence GPS performance. The sum of all systematic errors or biases contributing to the measurement error is referred to as range bias.

The observed GPS range (the range from the satellite to the receiver) without the removal of biases is called a pseudo-range. Principal contributors to the final range error that also contribute to overall GPS error are ephemeris error, satellite clock and electronic inaccuracies, tropospheric and ionospheric refraction, atmospheric absorption, receiver noise, and multipath effects. Other errors include those induced by the DOD (SA and AS). GPS also contains random observation errors, such as unexplainable and unpredictable time variation. Due to their random nature, these errors cannot be modeled and corrected. The following paragraphs discuss these errors as they associate with GPS positioning and navigation. Most are eliminated or their effects significantly reduced when GPS is used in a differential mode. (The same errors are common to both receivers during simultaneously observed sessions.)

Ephemeris Errors and Orbit Perturbations

Satellite ephemeris errors are errors in the prediction of the satellite position that are transmitted to the user in the Nav Data. Ephemeris errors are satellite-dependent and very difficult to predict and compensate. The many forces acting on the predicted orbit of a satellite are difficult to measure directly.

Clock Error

GPS relies heavily on accurate time measurements. GPS satellites carry rubidium and cesium time standards that are usually accurate to 1 part in 10^{12} and 1 part in 10^{13}, respectively. Most receiver clocks are actuated by a quartz time standard accurate to 1 part in 10^8. The difference between the satellite time and the receiver time is called the time offset. The product of the time offset and the speed of light equal the possible error due to clock bias as—

$R_E = T_O x c$

when R_E is the range error due to clock bias.

T_O is the time offset.

C is the speed of light (299,792,458 m/s).

For example, if the time offset is 1 microsecond (10^{-6}) then the $R_E = 10^{-6} \times 299{,}792{,}458 = 299.8$ meters.

Ionospheric Delays

GPS signals are electromagnetic signals that are dispersed nonlinearly and refracted when transmitted through a highly charged environment, such as the ionosphere. Dispersion and refraction of the GPS signal is referred to as an ionospheric range effect. (The dispersion and refraction of the signal results in an error in the GPS range value.) Ionospheric range effects are frequency dependent. L1 and L2 frequencies are affected differently even though they follow the same path through the ionosphere.

The error effect of ionospheric refraction on the GPS range value depends on sunspot activity, time of day, and satellite geometry. Periods of high sunspot activity produce greater range errors than periods of low sunspot activity because of the effects of the Sun's gravity on the ionosphere. Daylight GPS operations will produce greater range errors than night operations. GPS operations with satellites near the horizon will have larger range errors than those with satellites near the zenith. (The signal must pass through a larger portion of the ionosphere when the satellite is near the horizon.)

Resolution of ionospheric refraction can be accomplished with the use of a dual frequency receiver (L1/L2). During a period of uninterrupted observation of the L1 and L2 signals, the signals can be continuously counted and differenced. The resultant difference reflects the variable effects of the ionosphere delay on the GPS signal. Single frequency receivers in the absolute and differential positioning modes normally rely on an ionospheric model that model the typical ionosphere. Use of these models can remove a significant amount of the ionospheric delay.

Tropospheric Delays

GPS signals are not dispersed by the troposphere but they are refracted. Tropospheric conditions causing this refraction can be modeled by measuring the dry and wet components.

The dry component can be modeled easily through measuring the surface pressure using the equation—

$D_c = (2.27 \times 0.001) \times P_o$

when D_c is the dry term range contribution in zenith direction in meters.

P_o is the surface pressure in millibars.

Example: if the surface pressure is 765 millibars then the dry term range error is $(2.27 \times 0.001) \times 765 = 1.73655$ or 1.7 meters.

The wet component cannot be so easily modeled. The wet component is approximated by not only surface conditions, but by atmospheric conditions along the entire path of the GPS signal (water vapor content, temperature, altitude, and the angle of the signal path above the horizon).

Multipath

Multipath is a positioning error caused by the signal arriving at the receiver from more than one path. Generally, this is due to the receiver being located near a reflective surface such as a metal building or structure. Newer antenna designs have some filtering capabilities to reduce the effects of multipath. However, proper mission planning and site reconnaissance is the best way to reduce this type of error. Averaging of GPS signals over a period can also reduce multipath effects.

Receiver Noise

This error source includes a variety of errors associated with the receiver's ability to measure a finite time difference. Errors include signal processing and filtering, clock/signal synchronization and correlation methods, receiver resolution, signal noise, and electronic interference. Most errors cannot be modeled or accounted.

User Equivalent Range Error

User equivalent range error (UERE) can be referred to as the total budgeted error caused by the error sources listed above. Many of these error sources can be reduced through planning or using L1 or L2 antennas. Differential techniques can even eliminate some of these error sources. Figure 9-6 lists these errors and biases by associating them with their source segment. Error values in this figure do not include the effects of SA.

Absolute GPS Accuracy

Absolute positions are those that are established with no reference or tie to any other station. They are sometimes referred to as autonomous. For GPS purposes, this is generally accomplished by use of code-phase measurements to determine a pseudo-range. The accuracy of these ranges depends largely on the code (C/A or P(Y)) being used to determine the position. This range accuracy (UERE) when coupled with the geometrical relationships of the satellites results in a 3-dimensional confidence ellipsoid that depicts uncertainties in all three coordinates. Since satellites are constantly moving, the geometry constantly changes. GPS accuracy is time/position-dependent.

Segment Source	Error Source	Absolute, C/A code Pseudo-range, m	Absolute, P(Y) code Pseudo-range, m	Differential Positioning, m P(Y) code
Space	Clock Stability	3.0	3.0	Negligible
	Orbit Perturbations	1.0	1.0	Negligible
	Other	0.5	0.5	Negligible
Control	Ephemeris Predictions	4.2	4.2	Negligible
	Other	0.9	0.9	Negligible
User	Ionosphere	3.5	2.3	Negligible
	Troposphere	2.0	2.0	Negligible
	Receiver Noise	1.5	1.5	1.5
	Multipath	1.2	1.2	1.2
	Other	0.5	0.5	0.5
I_0 UERE		±12.1	±6.5	±2.0

Figure 9-6. User Equivalent Range Error.

Root-Mean-Square Error Measures

The two-dimensional (2-D) (horizontal) GPS position accuracy is normally estimated using a root mean square (RMS) radial error statistic called standard deviation or sigma (σ). A 1-RMS (one sigma) error equates to the radius of a circle in which the position has a 63 percent probability of falling. A circle twice this radius represents an approximate probability of 97 percent. This is a 2-σRMS or 2DRMS (2-deviations RMS) and is the most commonly used accuracy statistic in GPS survey. In some instances, a 3-σRMS (3DRMS) depicts a circle three times the radius of the 1-σcircle. This circle has a 99.7 percent probability. An RMS error statistic represents the radius of a circle and is not listed with a ±.

Figure 9-7 depicts RMS on an error ellipse at 2-σ. This ellipse represents a normal distribution of GPS position errors and is centered at the indicated position of the receiver. The radii of an error ellipse are expressed in standard deviation (sigma (σ)) of the position distribution and usually provide a direction such as Sigma North, Sigma East, and Sigma Up. Each sigma is a probability estimate of how close the actual position is to the displayed position as discussed.

Probable Error Measures

In 2-D horizontal positioning, a circular error probable (CEP) statistic is most commonly used, especially in military targeting. CEP refers to the radius of a circle with a 50 percent probability of position confidence. A measured or calculated position will fall inside a circle of some radius at least 50 percent of the time.

Three-dimensional (3-D) GPS accuracy is most commonly expressed as a spherical error probable (SEP). This value represents the radius of a sphere with a 50 percent confidence level or probability. It is important to understand that this sphere only approximates an actual 3-D-error ellipsoid that represents the uncertainties in the geocentric coordinate system.

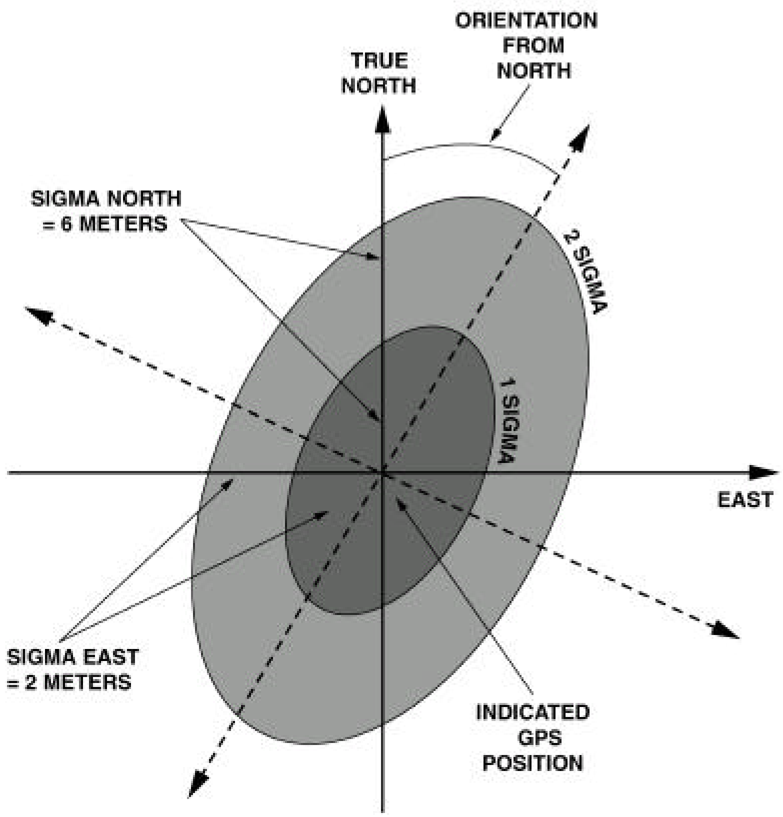

Figure 9-7. Error Ellipse at 2-σ RMS.

Dilution of Precision

Dilution of precision (DOP) is a scalar quantity representing the contribution or effect of the satellite geometry to the GPS accuracy. It is the ratio of the standard deviation of one coordinate to the measurement accuracy. It could be said that DOP is the measure of the strength of the satellite geometry. In general, the more satellites that are observed and used in the final solution, the better the solution. Since DOP can be a measure of geometric strength, it can also determine the satellites a receiver uses to determine the most accurate position.

Satellite geometry and DOP can best be visualized by the following analogy. A rubber ball is suspended from five strings. All five strings are attached at the other end to the ceiling within a couple of meters of each other. Because of poor geometry of the strings' attachment to the ceiling, the ball can be moved easily. A large portion of the possible error in the position of the ball is due to poor geometry or high DOP. On the other hand, if the ball was still attached to the same strings but four of the strings were attached to the ceiling at the four corners of the room and the fifth string attached directly above the ball, the ball would not move so easily. This would be due to strong geometry. The DOP would be small.

The main form of DOP used in measuring absolute accuracy is the geometric DOP (GDOP). GDOP is the measure of accuracy of 3-D position and time. GDOP is related to actual range error by stating that the actual range error equals the GDOP multiplied by the UERE.

Position dilution of precision (PDOP) is the measure of the accuracy in 3-D position only. PDOP values are generally developed from satellite ephemerides prior to the conduct of survey operations. The surveyor can more adequately plan sessions and occupations for his equipment.

PDOP represents the position recovery at a particular instance in time and is not representative of the entire session. PDOP error is generally given in units of meters of error per 1-meter error in pseudo-range measurement; i.e., m/m. If the pseudo-range measurement due to clock errors, atmospheric conditions, etc., is in error by 2 meters and the PDOP is 5.5, the possible position error due to satellite geometry is 11 meters. When using pseudo-ranging techniques for absolute positioning (code phase), PDOP values lower than 5 m/m are considered very good; values greater than 10 m/m are very poor. For static-type surveys, it is desirable to make observations during periods of rapidly changing PDOP.

When the values of PDOP and GDOP are observed over time, high values (> 10 m/m) can be associated with poor geometry. The higher the PDOP the poorer the solution for that instant in time. Poor geometry can be the result of satellites being in the same plane, orbiting near each other or at similar elevations.

Horizontal dilution of precision (HDOP) is a measurement of the accuracy in 2-D horizontal position. It is significant in evaluating surveys intended for horizontal control. Basically, HDOP is the RMS position error divided by the standard error in the range measurements. It roughly indicates the effects of satellite range geometry on a resultant position. HDOP values lower than 3 indicate the best geometry.

Vertical dilution of precision (VDOP) is a measurement of the accuracy in standard deviation (σ) in vertical height. Mathematically, it is the σ_u (Sigma Up) divided by UERE (around 6 meters with P(Y) code and 12 meters with C/A code). A VDOP value lower than 3 indicates a strong vertical component in the geometry.

Time dilution of precision (TDOP) is the measurement of the accuracy of the time determined by the GPS receiver.

GPS vulnerabilities are generally grouped by the segment they threaten.

Space Segment

The height of the GPS satellites (10,898 miles) is outside the range of current antisatellite weapons. Such space weapons would be deployed at lower altitudes where they could possibly be used against more attractive space-based systems. The constellation generally keeps the satellites approximately 44,000 kilometers apart in each orbital plane. The control segment manages the system so that no two satellites will orbit within 8,100 kilometers of each other. This spatial separation ensures that a single nuclear burst in space at half the closest distance will have little effect, forcing a more direct assault against individual satellites. A ground-launched attack by antisattelites would be detected early enough in the 3-hour flight time to allow for maneuvering of satellites.

It is unlikely that current or projected technology would allow a nation to launch a direct attack or even detonate a nuclear device in the GPS orbit. A nuclear detonation in space would equally affect other space-based systems with a blackout or scintillation effect lasting 10 minutes or longer in the L-band. Detonation would disturb the atmosphere and the ionosphere with subsequent effects on propagation. Radiation effects could incapacitate the functioning of the erasable read only memories (EROMS) and random access memories (RAMS) of these systems if the blast is close enough. GPS satellites have built-in protection against the electromagnetic pulse (EMP) caused by nuclear detonations and can restore their erased memories through the control segment.

Two other factors exist that add an unplanned edge of survivability against antisatellites. The Russian Global Navigation Satellite System (GLONASS) orbits at an altitude close to the GPS orbit. A nuclear assault against GPS would have the same effect on GLONASS. The increased use of GPS by former Eastern Bloc nations including the former Soviet Union is also an unplanned edge.

Since the introduction of the Strategic Defense Initiative (SDI), lasers have become a planned vulnerability to orbiting satellites. Survivability of GPS against laser technology is enhanced by limited laser hardening of the satellites. A space-based laser at low altitudes would be an extremely heavy device to launch. Tracking and targeting of a GPS satellite is extremely difficult due to its relatively small size and its orbital rate (14,500 kilometers per hour). Technology is available for a high-power ground-based laser-to-target and has a limited effect on GPS satellites.

Should there be a loss of any GPS satellite, replenishment from the ground can be accomplished within 2 months. The system can operate fully even with the loss of 9 or more satellites, still providing 3-D coverage for 12 hours and 2-D coverage for more than 20 hours.

Control Segment

This segment is considered by many to be the most vulnerable GPS segment. The master control station (Colorado Springs) is very well protected by being collocated with several military bases. The tracking stations do not have this protection. All stations are susceptible to espionage and natural disasters, but tracking stations are much more vulnerable to nuclear and conventional assaults. These factors have been considered, and although the ability exists to exert complete control over the satellites from the master control station, the uplink to each satellite will be less frequent than the current 8 hours. Next-generation satellites are planned to provide a cross-link ranging capability that will allow satellites to communicate with each other, making offshore tracking stations redundant. Jamming the control segment between the tracking stations and the master control station is a real possibility.

User Segment

Only two major vulnerabilities exist with the user segment: cryptographic key security and jamming. Several types of jamming can be used against GPS receivers.

Cryptographic Key Security

Should the enemy capture a GPS receiver with a valid crypto fill, the PPS-security module prevents reverse engineering so that the enemy cannot gain access to the cryptographic data stored inside the receiver. Also, the regular change of crypto keys adds to the security and survivability of the system.

Spoofing

Spoofing is classified as the enemy's attempts to duplicate or imitate GPS signals. Spoofing requires that the enemy have a knowledge of the received satellite phase and frequency at the targeted receiver antenna. The proper carrier frequency and timing code phase plus a sufficiently higher power output will allow a deceptive jammer to establish a false lock with the receiver. The enemy must know which satellites are being tracked and the position and velocity of the receiver to create a false signal with the correct Doppler shift. Once false lock is established, the spoofer can perturb the duplicated navigation signals to cause navigation errors.

When GPS antispoofing protection is enabled, spoofers cannot autonomously generate the signals needed to deceive the receiver that is looking for the encrypted P code. This technique transforms the P code by cryptographic means. The resultant bit stream is a cypher text called the Y code that replaces the P code in its entirety. Receivers with a valid crypto key will encrypt their own generated P code to produce a Y code inside the receiver needed for correlation with the satellites transmitted Y code. The C/A code is totally unaffected by the P code encryption. This is why the C/A code is very susceptible to deception jamming.

Continuous Wave Jamming

A continuous wave (CW) jammer (or spot jammer) concentrates its jamming power in a very narrow band around the L1 or L2 carrier. In the GPS receiver the jamming power spreads out according to the receiver processing gain using the modulated C/A or P code. After spreading, only a small portion of the jamming power enters the tracking loops of the receiver. Consequently, a CW (spot) jammer needs to be very powerful to effectively jam a GPS receiver.

Wide Band Jamming

A wide band jammer needs a sufficiently wide spectrum to cover the frequency band of the GPS signals. Usually, wide band jamming is performed in either a sweeping mode or a quasi-random noise (barrage) mode. In a sweeping mode the jammer carrier moves rapidly and possibly randomly through the band to be jammed. In barrage jamming the spectral power density is low compared to spot jamming.

Pulse and Amplitude Modulated Jamming

A pulse jammer switches its power on and off during an operation cycle. The modulation rate can be chosen as needed (pulse repetition frequency and pulse duration). An amplitude modulated jammer varies its power linearly; modulation depth and frequency can be chosen as needed. Besides jamming effects already described, this type of jamming could have a negative impact on the GPS receiver by disturbing the internal operating cycles of the receiver processor unit or the electronics of the antenna.

Jamming Effects on Code Acquisition

The GPS user is at a severe disadvantage when having to use C/A code acquisition in a hostile region. Current technology receivers have a limited tolerance to jamming during C/A code acquisition (usually around 25 decibels). Although only 1,023 chips must be searched in the C/A code, a large frequency window must be searched to account for satellite and receiver Doppler shift. The receiver must maintain a large predetection bandwidth to acquire the C/A code at the expense of reduced jamming tolerance (J/S). The C/A code is widely disseminated. A smart jammer can broadcast a false signal and lock up the receiver with very low power. Tests indicate that a very modest 1-watt noise jammer prevents a receiver from acquiring the GPS signal using the C/A code out to 85 kilometers. If the jammer were to transmit a spoofing signal, the receiver would not be able to discriminate between the desired GPS signal and the spoofing signal. The receiver's jamming tolerance would be on the order of 0 decibels, a 1-watt spoofer could deny C/A code acquisition past 1,000 kilometers or to the jammer's horizon.

The P(Y) code is virtually impossible for an adversary to spoof because of its long length and the encryption of the P code into the Y code. Also, the P(Y) code has 10 decibel more antijam protection than the C/A code because of its 10 times larger bandwidth. Some current technology receivers using direct P(Y) code acquisition (hot start) can tolerate a J/S level in excess of 35 decibels. However, because the P(Y) code is so long, (6×10^{12} chips) much time or more correlators operating in parallel are needed for the two-dimensional search over code timing and Doppler frequency. Thus, up to date satellite ephemerides and accurate code timing must be available to effectively perform a hot start.

Reacquisition typically has very accurate initial code timing. If this accuracy is available during direct P(Y) code acquisition, the reacquisition problem is the same as the acquisition problem with increased J/S tolerance (in excess of 50 decibels depending on the receiver).

Antijam Capabilities of GPS

GPS vulnerability to jamming can never be completely eliminated. With constantly improving techniques and technology, higher power outputs by the enemy jammers is required to deny access to GPS signals.

The GPS satellites broadcast in the ultra-high frequency (UHF) domain. UHF signals cannot “go around” obstacles such as buildings, hilltops, large rocks, etc. the way low domain frequencies can. For this reason, ground based jammers are restricted by severe line-of-sight limitations. Air or space-borne jammers, obviously, are not so restricted. A smart user will position his GPS receiver to limit the effects of a line-of-sight jammer. Defilade areas that still allow open skies for satellite signals can be used. Handheld receivers can be placed in a “cat hole” or the user can simply turn his back to the suspected direction of the jammer. Once the receiver has a lock on the P(Y) code, the user should be able to operate normally in most jamming environments.

The satellite signal is a continuous wave type signal. This signal is spread by modulation and enters the receiver with the jamming signal. The original signal is recovered as the jammer energy is spread over the modulation and most of the jammer power is filtered out.

To locate, track, and demodulate a satellite signal, the receiver requires a certain minimum ratio between received signal power and the noise power. This value (or threshold) is referred to as the signal-to-noise ratio (SNR). The noise power is dissipated over a wide bandwidth, while the satellite signal power is concentrated in a very narrow frequency band. Using filters, the GPS receiver narrows its bandwidth to the absolute minimum in order to receive the maximum signal power and minimum noise. A receiver can usually recognize the jammer power as noise; therefore, much of its effects are filtered out. Depending n the power output of the jammer, the SNR may be too high for the receiver to filter out enough noise to determine data.

Many other antijam capabilities exist for GPS. Among them is the inertial navigation system (INS) integration with GPS and high cost antennas. Both are used in military applications but generally with missile, rocket, and aircraft navigation technologies. For the most part, antijam capabilities for Marine ground units will be enhanced by proper planning and positioning of GPS assets.

The surveyor must be aware of the ambient conditions of the GPS satellites and their surrounding environment. Solar flares, occurring in 11-year cycles, send great fountains of electromagnetic energy and radiation deep into space, causing interferences in GPS signals. Meteor showers, such as the Lenoid event, occur close to the Earth, and could potentially damage satellites in the GPS constellation. GPS satellites are hardened against such possibilities, but are not invulnerable.

SECTION III. GPS MEASUREMENTS

GPS Reference System

The GPS satellites reference their own position to the WGS 84 coordinate system. This system is based on the WGS 84 ellipsoid. To fix the Earth in time and space for the development of the WGS 84 ellipsoid, the Conventional Terrestrial Pole of 1984 (CTP 84) was chosen. The position of the Earth’s polar axis at 1984.0 as defined by the Bureau International De l’Heure (BIH) is used to define the z-axis of the WGS 84 cartesian coordinate system. The x- and y-axes are then referenced to that z-axis. All positions determined by GPS are originally in this format, WGS 84 cartesian. Receivers and software applications for GPS have the capability to provide positions in other coordinate systems and other datum/ellipsoid references.

Code Phase Measurements

The primary purpose of code phase measurements is to determine approximate ranges from satellites to the

GPS receiver that allow the receiver to determine its position. Since this position is not referenced or relative to any other position or receiver, it is referred to as an absolute position. Clock biases, atmospheric absorption and refraction, and other inherent errors make determination of a true range virtually impossible. The actual range determined by the GPS receiver is referred to as a pseudo-range. See figure 9-8.

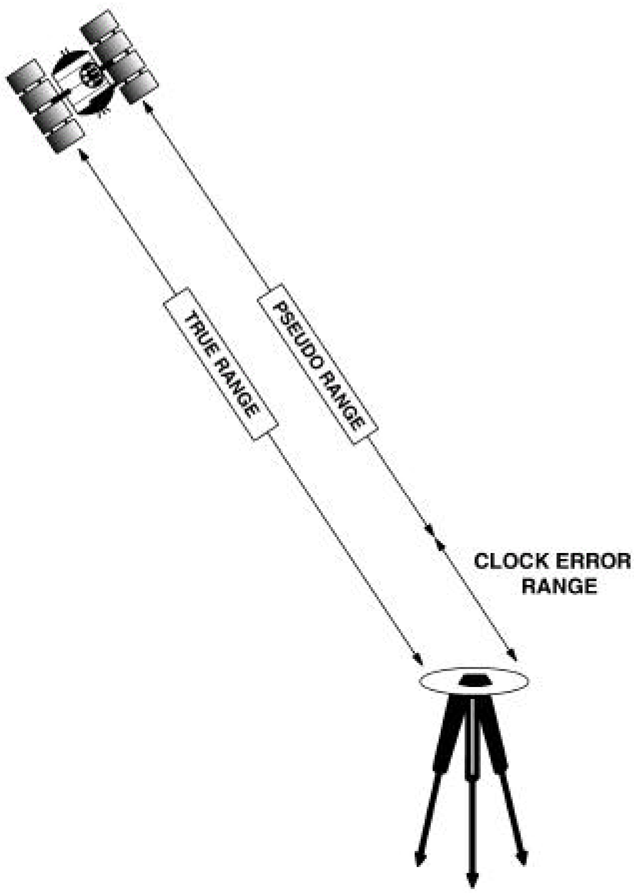

Figure 9-8. Pseudo-range.

Pseudo-random Noise Code

The Pseudo-random noise (PRN) code is a binary data string of digital ones and zeros that is unique to the satellite broadcasting it. This code is used by the receiver to identify the satellite it is tracking. This code is repeated every millisecond (C/A) and every week (P/Y). See figure 9-9.

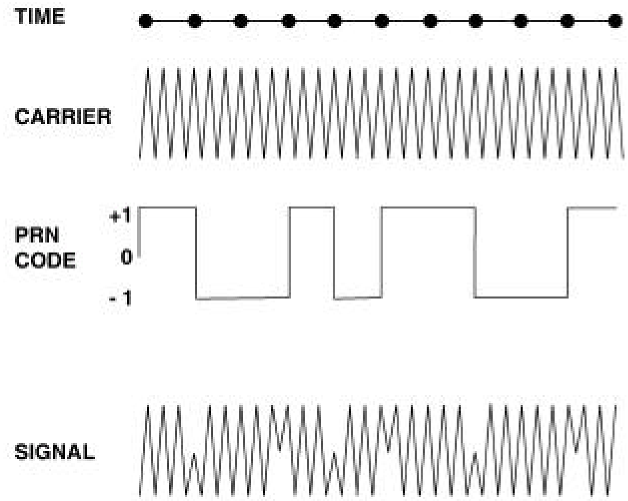

Figure 9-9. Carrier Modulated by PRN.

Time Delay

Before the receiver can compute the pseudo-range to the satellite, it must determine the time required for the signal to travel from the satellite to the antenna. The receiver stores a replica of each satellite's PRN code. When the receiver detects a satellite signal, it identifies the satellite by its PRN that has been replicated from memory. The code received from the satellite is then compared to the replicated code. The receiver slides the replicated code in time until it lines up with the satellite's transmitted code. The amount of time that was needed to slide the code is the time delay of the transmission or travel time. See figure 9-10.

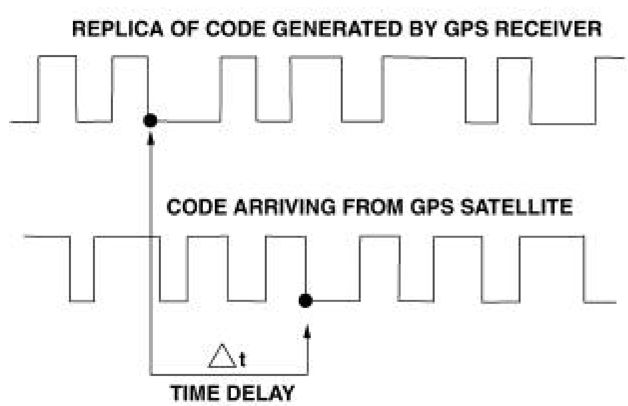

Figure 9-10. Time Delay of a GPS Signal.

Satellite Ranging

The only measurement made by a code-phase receiver is the time delay or transmission time. Radio waves travel at the speed of light. This constant value is stored in the memory of the GPS receivers. A receiver capable of making code-phase measurements will compute the pseudo-range using the formula pseudo-range = $\Delta t \times$ speed of light.

This procedure can be simultaneously performed on many satellites. The number of satellites depends on the receiver used, or more specifically, how many channels are available in the receiver.

Once a satellite is tracked and the receiver determines a pseudo-range, the receiver basically knows it is located on a sphere whose radius equals the pseudo-range with the satellite at the center of the sphere. See figure 9-11.

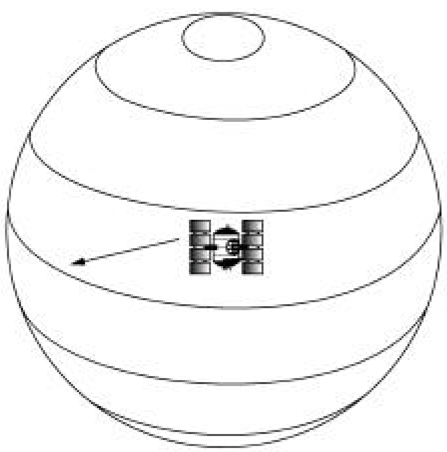

Figure 9-11. Ranging One Satellite.

When the second satellite is acquired, the same ranging technique is used creating a second sphere. The intersection of the two spheres is a circle. The receiver is located somewhere along the edge of that circle. See figure 9-12.

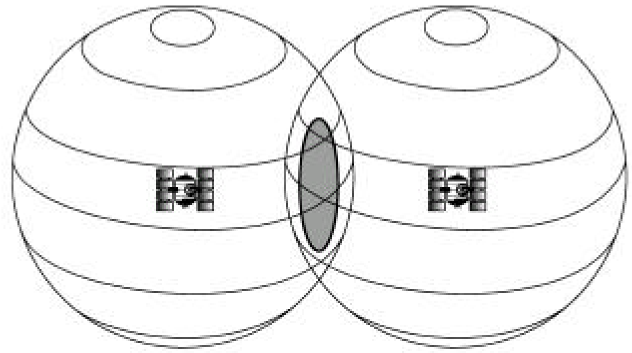

Figure 9-12. Ranging Two Satellites.

The third sphere determined by ranging a third satellite would intersect the circle created above at two points. The receiver knows that its position is referenced to the WGS 84 ellipsoid/datum. Only one of the two points of intersection will be located on this geodetic system. The other point will be out in space, deep inside of the ellipsoid or moving at an extreme velocity. With three satellites, a receiver can provide a 2-D position. To achieve a 3-D position, at least four satellites must be ranged. The fourth satellite provides the timing data for the receiver to resolve timing errors in the system. See figure 9-13.

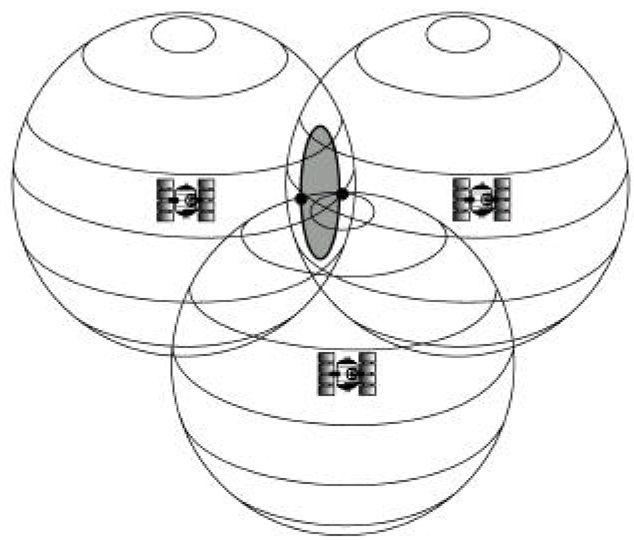

Figure 9-13. Ranging Three Satellites.

Carrier Phase Measurements

The primary purpose of carrier phase measurements is to determine ranges from satellites to receivers that will allow the receiver to position it. Usually these positions are processed relative to another receiver position and are referred to as differential positions. Ranges to the satellites are pseudo-ranges, as in code phase. However, this method of ranging requires the solution of the integer ambiguity of the signal. The determination of this distance requires that the number of whole carrier wavelengths be known.

Integer Ambiguity

The whole number of wavelengths between the satellite and the receiver is known as integer ambiguity or cycle ambiguity. Since we know the L1 carrier wavelength is 19 centimeters long and L2 is 24 centimeters long, and since most carrier phase receivers can determine the partial wavelength to an accuracy around 2 mm; the pseudo-range can be accurately measured as long as we can determine the number of complete wavelengths between the satellite and receiver. This is done by comparing changes in the received frequency (caused by the Doppler effect) to the broadcast frequency over time.

Carrier (Beat) Phase

Carrier phase GPS receivers contain an internal oscillator that generates a carrier signal. This generated carrier signal is compared to the received signal from the satellite. The carrier phase observations (also called carrier beat phase) are determined from these measurements.

Continuous Carrier Phase

When the receiver first locks on to a satellite signal, it can only measure the fractional part of the wavelength. It has no knowledge of the number of full wavelengths at that specific point in time between it and the satellite. After that first measurement, the receiver will count the number of whole wavelengths it observes. This is the continuous carrier phase.

If the satellite signal is interrupted (see next para), the continuous carrier phase is reset. It is set to the next fractional wave measurement (carrier phase observable) immediately following the break.

Cycle Slips

A cycle slip is the interruption or break in the continuous carrier wave.

The wave fronts that are counted by the receiver during continuous phase tracking are called cycles. When the signal is interrupted, the continuous count of those wave fronts or cycles is broken, or the count slips. This cycle slip causes the continuous carrier phase to be reset. The baseline processor in most GPS-S systems can reestablish this count whether in a static or a kinematic mode.

Cycle slips can be caused by any number of barriers between the satellite and the receiver. These barriers can include terrain masks, trees, or even an operator standing between the satellite and the antenna. Cycle slips could cause burst jamming signals. Usually reconnaissance, LOS clearing, and planning can eliminate many sources of cycle slips.

Differencing

Differencing is a method used by the processors to solve for the first estimation of a baseline solution and remove measurement errors.

Single Differences

A single difference can be formed by differencing the measurements acquired by two receivers observing the same satellite at a particular point in time or epoch. Therefore, integer ambiguities associated with each receiver are combined. Single differences between receivers virtually remove all satellite-dependent errors such as satellite clock error, and to a large extent, orbit errors and atmospheric delays. See figure 9-14.

A single difference can be formed by differencing the measurements acquired by one receiver observing two satellites at a particular point in time or epoch. Single differences between satellites reduce most receiver dependent errors. See figure 9-15.

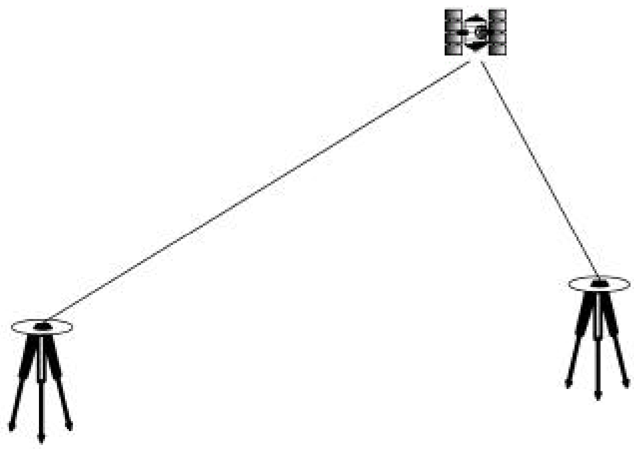

Figure 9-14. Single Differences between Receivers.

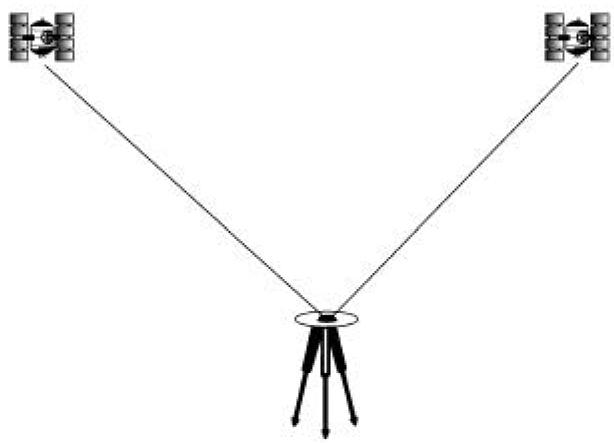

Figure 9-15. Single Differences between Satellites.

Double Differences between Satellites and Receivers

A double difference is formed by differencing two single differences. This involves two receivers observing the same two satellites at the same epoch. Four separate measurements and four separate integer ambiguities are combined to create a difference. The double differencing mode removes most of the effects of satellite and receiver clock drift. See figure 9-16.

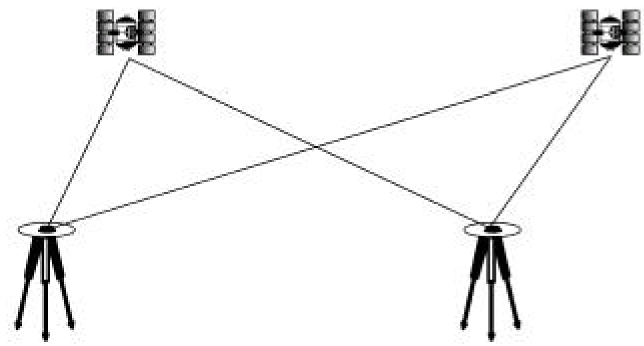

Figure 9-16. Double Differencing.

Triple Differences between Satellites, Receivers, and Time

A triple difference is determined by combining two double differences over time. The double difference determined by a set of satellites and receivers at a particular epoch is combined with the double difference form the same satellites and receivers at a different epoch. In this mode, integer ambiguities cancel out of the computations because it does not change over time.

Triple differences are often used to find cycle slips. A cycle slip, in the single differencing mode, causes the receiver to recompute the combined integer; therefore after a cycle slip, the integer ambiguity does change. A large change in the triple difference is a good indicator of a cycle slip at that epoch. See figure 9-17.

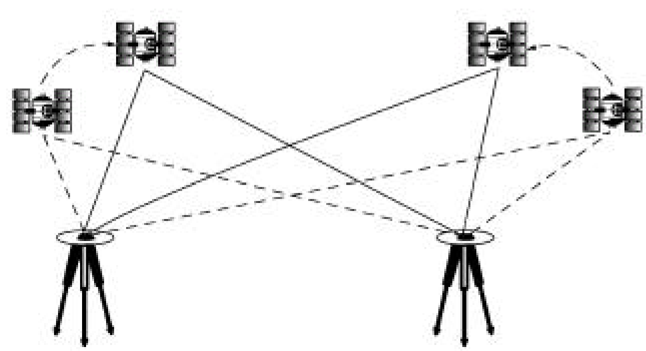

Figure 9-17. Triple Differencing.

Baseline Solutions (Vectors)

A baseline solution (vector) is a straight line defined by its 3-D (ΔX, ΔY, ΔZ) values when one end of the vector is the origin and the other end is the point containing those relative values. The processor uses the differencing methods described above along with code solutions to determine an initial estimate of the baseline vector. This initial estimate is called the triple difference solution.

If the integer ambiguity is known, multiply that value by the wavelength (19 centimeters – L1, 24 centimeters – L2) and add the partial wavelength to obtain the pseudo-range. At this point in the processing, the integer ambiguity is still an unknown value.

Float Solution

Once an initial estimate of the baseline vector has been determined, place that value in the formula—

$$\Delta X, \Delta Y, \Delta Z = (N \times \lambda) + \Delta\phi$$

Whereas: ΔX, ΔY, ΔZ is the baseline vector,

N is the integer ambiguity,

λ is the integer wavelength,

$\Delta\phi$ is the phase change observed in a small portion of the data.

Often in processing, the value for N as determined in the above formula is not an integer. The ambiguity computes to a value such as 500.52. This value is not close enough to a whole integer for the processor to determine if the ambiguity is 500 or 501. It cannot set the value to a whole integer.

The value determined is compared against the remaining observations to see how well it fits. If the residuals (errors) are within a certain tolerance, the processor generates a new baseline vector; i.e., the float solution.

Fixed Solution

A fixed solution is obtained when the processor determines a set of integer values for the ambiguity that is significantly better than the other values.

The processor rounds the ambiguity value determined above to whole numbers for each satellite, each time testing different combinations of whole wavelength values to compute a baseline.

Each time a new set of integers is used, an associated variance (square of the standard deviation) is generated. After all possible combinations of whole wavelengths have been tried, the processor selects the solution with the lowest variance (least error). This is the fixed solution or fixed-integer solution.

The ratio of the errors between the integers used for the last iteration (the last computation of whole integer values) can be determined as—

$$\text{Ratio} = \frac{\textit{(integers giving next least) Errors}}{\textit{(integers giving least) Errors}}$$

With a single frequency receiver, this is the best solution that can be determined. It is sometimes referred to as a double-difference fixed solution.

With a dual-frequency receiver, it is possible to combine carrier phase observables to create other fixed solutions. A wide lane carrier phase is generated when the processor differences the carrier phase observables (L1 – L2). The effective wavelength is 86.2 centimeters. This combination allows for easier resolution of the integer ambiguities so it is often used to solve long baselines. The narrowlane carrier phase is generated when the processor combines the carrier phase observables (L1 + L2). The effective wavelength of the narrow lane carrier phase is 10.7 centimeters. This combination is very effective for canceling out ionospheric errors. This baseline solution uses a combination of the L1 and L2 carrier phases to model and remove the effects of ionospheric interference on the signals. This is the optimal solution, used for high-order control networks and for observing long baselines.

SECTION IV. GPS SURVEY METHODS AND TECHNIQUES

Absolute Positioning

Absolute positioning is a GPS survey method that involves using a single passive receiver; e.g., AN/PSN-11 PLGR, AN/PSN-13 MSGR 4000. The term absolute does not refer to a specific accuracy. It means this method does not rely on any source of information other than what is collected by the receiver at that station. This position is not relative (common) to any other station. The accuracy of this position depends on many different error sources as well as the user's level of authorization (PPS or SPS).

The receiver collects data from multiple satellites and uses this data to determine position, velocity, and timing information. The position is generally determined from code phase measurements. Some receivers can use carrier phase measurements to determine absolute positions.

Differential (Relative) Positioning

Differential positioning requires at least two receivers collect data from at least four common satellites simultaneously to compute a vector between them. The vector is then fixed at one end to a point and the other end is the relative position.

Usually, one receiver is located at a known point. Depending on the differential technique used, more than four common satellites may be necessary. Processing the collected data can be performed in the office or by the receiver in the field, also depending on the differential technique used.

Much of the accuracy achieved from this method is due to the use of common satellites and common epochs (a specific point of time selected for a GPS measurement). Figure 9-16 shows that differential techniques negate most sources of error. This is because the same error exists at each station collecting data from a specific satellite at a specific epoch. The errors broadcast by satellite PRN23 and collected by receiver A at epoch 1 are the same errors collected by receiver B at epoch 1. The errors broadcast from the satellites have no effect on the dimensions of the vector because the errors are equal at each end.

This is actually only true for errors sources in the Space and Control segments. User segment error sources are not always equal at each end of the vector. For distances under 25-30 kilometers, the tropospheric and ionospheric errors are basically the same. Signals from satellite PRN23 to receiver A travel through the same sampling of the atmosphere as the signals from PRN23 to receiver B. Larger distances may add some small errors into these measurements. An L1/L2 antenna will decrease this error.

Code Phase Differential Positioning

Determining differential positions from code phase measurements is performed by applying a correction to the pseudo-range determined from an individual satellite to the receiver.

This process begins with the pseudo-ranges from code phase measurements used to determine the absolute positions of the receivers. Since the errors collected at each receiver are the same for each epoch, a pseudo-range correction (PRC) can be computed. In other words, assume we know the exact position of a satellite at a specific epoch and the surveyed position of a GPS receiver, we can determine a true range. If the measured pseudo-range is 79 meters and the true range is 81 meters, the PRC is +2 meters. A pseudo-range correction can be generated for each satellite being observed. Any receiver that is simultaneously collecting data from at least four common satellites can apply the PRC to its pseudo-range measurements to obtain a relative (common) position; thus the distance between the two points will be relatively accurate (0.5-10 meters) even when the absolute positions are not.

Code phase differential positioning has its primary applications in real-time navigation where relative accuracy is as low as 10 meters are acceptable. Also, some engineering survey applications can tolerate this accuracy. This would not be acceptable for geodetic applications, and does not meet artillery specifications.

Carrier Phase Differential Positioning

Determining differential positions from carrier measurement is as simple as fixing an end of a measured vector to determine the position of the other end. Through processing, other vectors can then be fixed to the end of the first vector to create a network. Kinematic and static surveys are both usually performed using carrier phase differential positioning.

Static and Kinematic Techniques

When GPS receivers are used for surveying purposes, it is generally accepted that the survey will be performed using carrier phase differential (relative) survey methods. Differential survey is usually divided into two techniques: static and kinematic.

Static surveys provide the most accurate results. Receivers must remain stationary (static) for a period of time depending on the type of static survey performed. There are two types of static survey: static and fast static. Both require extensive planning and post-processing. Static survey allows for extremely accurate networking of survey control. Due to planning, field work, and post-processing requirements, this technique should only be used by surveyors whose mission is to provide fourth order control.

Kinematic surveys provide accuracy results sufficient for most artillery survey missions (fourth and fifth order) but does not provide the same networking capabilities as static techniques. There are two types of kinematic surveys: stop-and-go and continuous. Stop-and-go surveys can be post-processed in the office computers or in the field by the receivers using RTK and RTK/OTF procedures. Kinematic survey techniques require that one receiver remains static while another acts as a rover, moving along a route or from station to station collecting data.