(1)

AGARD-LS-161

AGARD-LS-161

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARD LECTURE SERIES No.161

The NAVSTAR GPS System

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY ON BACK COVER

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARD Lecture Series No.161

THE NAVSTAR GPS SYSTEM

This material in this publication was assembled to support a Lecture Series under the sponsorship of the Avionics Panel of AGARD and the Consultant and Exchange Programme of AGARD presented on 19—20 September 1988 in Wessling, Germany, on 22—23 September 1988 in Athens, Greece, on 26—27 September 1988 in Ankara, Turkey and 29—30 September 1988 in Lisbon, Portugal.

THE MISSION OF AGARD

According to its Charter, the mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community;
- Providing scientific and technical advice and assistance to the Military Committee in the field of aerospace research and development (with particular regard to its military application);
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Exchange of scientific and technical information;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the authors.

Published September 1988

ISBN 92-835-0477-1

SPS

Printed by Specialised Printing Services Limited
40 Chigwell Lane, Loughton, Essex IG10 3TZ

ABSTRACT

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

* * *

Ce Cycle de Conférences présente de façon exhaustive le système de positionnement à couverture mondiale NAVSTAR. Les divers sujets traités portent sur la philosophie du système, les principes de navigation, la conception du système, les applications pour les utilisateurs, ainsi que les technologies utilisés. Les conférences traitent de part égale de la theorie, des applications et des performances enregistrées.

Ce Cycle de Conférences est presenté dans le cadre du Programme des Consultants et d'Echanges, sous l'égide du Panel AGARD d'Avionique.

LIST OF AUTHORS/SPEAKERS

Lecture Series Director: Mr Edward M.Lassiter
The Aerospace Corp/M5-647
POB 92957
Los Angeles, CA 90009-2957
USA

Lecture Series Authors: Dr Mohan P.Ananda
The Aerospace Corp/M5-647
POB 92957
Los Angeles, CA 90009-2957
USA

Major Elio Bottari, IAF
NATO Representative
GPS Joint Program Office
Space Division — CWNI
P.O. Box 92960
Los Angeles, CA 90009
USA

Lt Cdr Per Nieuwejaar
Norwegian Naval Materiel Command K3
POB 3
5078 Haakonsvern
Norway

Mr Robert P.Denaro
TAU Corporation
485 Alberto Way
Los Gatos, CA 95070
USA

Lt Cmdr H.Kunze SD/CWNG
Dept. of the Air Force
Hqs Space Div. (AFSC)
Los Angeles, CA 90009-2960
USA

CONTENTS

*This paper was prepared from the material presented as Lectures 1 and 2.

THE GPS OVERVIEW AND NAVIGATION SYSTEM CONCEPT

EDWARD LASSITER AND MOHAN ANANDA
The Aerospace Corporation
El Segundo, California, U.S.A.

ABSTRACT

The Global Positioning System (GPS) objectives for both military and civil applications and the historical evolutions of the GPS are presented. The GPS concept validation phase, full-scale engineering development, system test phase and operational system phase are described. The fundamental principles of the GPS concept are briefly reviewed. The GPS radiometric measurements of pseudo-range and accumulated delta range and how these measurements are used in forming the navigation equations to solve for the user position parameters are described. The current program status and future plans for system enhancements are outlined.

1. INTRODUCTION

The use of heavenly bodies--i.e., the sun, the moon, stars, and planets--for purposes of navigation was started centuries ago and was utilized extensively by the early Portugese navigators to explore this planet. These early explorers made "position fixes" by combining the known positions of the heavenly bodies with on-board position measurements made with an instrument called an "astrolabe". The results of this simple technique were sufficiently accurate to allow navigators to find their approximate positions even when far from land. In the 18th century, the sextant, compass, and star and sun tables were integrated with the clock to improve navigation performance significantly. Later, a technological breakthrough in the use of radio signal direction-finding for navigation produced significant advances in the accuracy of position fixing; many of today's radio navigation systems, such as Loran and Omega, utilize this basic concept.

The use of artificial earth satellites for purposes of navigation originated with Sputnik I in October 1957. Satellite navigation combines the methods of celestial navigation, as used by the early explorers, with those of radio navigation to achieve systems having revolutionary improvements in accuracy and performance. The fundamental difference between artificial satellite navigation using radio signals and other radio navigation methods is simply the geometry. Space offers the opportunity for line-of-sight signal propagation over vast areas of the world, so the usual tradeoff of less accuracy for greater range is not involved. Also, since satellite signals penetrate the ionosphere rather than being reflected by it, difficulties encountered with "sky waves" are eliminated. Artificial earth satellites are obviously desirable platforms from which to provide navigational services, but these advantages have been gained at the price of increased sophistication. In the 30 years since Sputnik I, space technology has generated positioning systems, now under development, which promise to provide users with position accuracies of a few meters or so, velocity measurements to within a few tenths of a meter per second, and time readings within a few billionths of a second

The development of Transit I, the first navigation satellite system, was triggered by observations made on signals from the first Sputnik. Officially begun in December 1958, Transit I resulted in a worldwide navigation system which has been in continuous operation since January 1964.

It became obvious that a global navigation satellite system had much to offer military users in terms of accurate all-weather, continuous, worldwide navigation capability. Consequently, the Defense Department established requirements for a DoD tri-service worldwide navigation satellite system. During the 1967-1969 time frame, preliminary concept formulations and system design studies were conducted by the U.S. Air Force for such a system, which was designated System 621B. As a result of these efforts, combined with mission analyses and parametric studies, a space-based navigation system was developed that called for 20 satellites, deployed in synchronous orbits, whose ground tracks formed four "eggbeater"-shaped clusters extending to 60 deg north and south latitudes. Satellite tracking and control were to be maintained from ground stations in the continental United States through inter-satellite links. This inter-satellite tracking approach minimized the vulnerability of the system to physical attack on ground stations. System 621B was designed to make direct, simultaneous range measurements from at least three satellites and instantly compute a position fix at the intersection of three spheres with centers at the satellites.

Simultaneous range measurements from a fourth satellite eliminated the need for synchronization of satellite user clocks, since the time bias could be calculated in the navigation solution, which consisted of using four measurements to solve for the three unknown positions and one time. The demonstration and devlopment program for this early global navigation satellite system called for four synchronous repeater satellites that would provide for test and evaluation experiments and for the development of user equipment. Although no System 621B satellites were launched, nine-month test of basic navigation satellite techniques and user equipment demonstrated the merit of the basic concept and confirmed the signal structure design.

Concurrently with the USAF space-based navigation studies, the Naval Research Laboratory (NRL) conceived the idea of a timing/navigation satellite system (TIMATION). Development of the TIMATION system was to consist of two phases. Phase I (TIMATION I, II, and III) was initiated primarily as a technology effort to investigate the behavior of high-stability crystal oscillators in low-altitude orbits and to verify the TIMATION technique. Phase II involved the development and deployment of the operational system. The TIMATION concept involved making direct range measurements from the satellite to the user, with time-delay readings being taken each minute during a satellite pass. The direct range measurements were made by making phase measurements on several side tones modulated on a carrier signal. TIMATION I and II, which were launched in 1967 and 1969, performed precise time transfer, navigation, and geodesy experiments, transmitted both side-tone-ranging (STR) and pseudo-random noise (PRN) signals, and were used to conduct navigation and time transfer experiments. TIMATION III, subsequently identified as the NTS-1, was launched in mid-1974. As a result of these studies, a TIMATION global investigation system was proposed by the Navy, utilizing 21 to 27 satellites in medium (8-hr) orbits and both STR and PRN signals.

The USAF system 621B and the Navy TIMATION system were both candidates for the DoD Navigation Satellite System. Budgetary constraints would not permit the deployment of two independent systems. The compromise configuration consisted essentially of orbits proposed by the Navy to permit evolutionary deployment and the signal structure and frequencies proposed by the USAF for maximum user performance. The program resulting from this composite effort is the Navstar/Global Positioning System.

Since the early 1970s, the Navstar/Global Positioning System has evolved significantly. There exists a large number of publications in the literature throughout various stages in the growth of the GPS. Some of the papers were published by the Institute of Navigation in the special issues of Navigation Vol I (1980), Vol II (1984), and Vol III (1986). Some specific papers of interest were contributed by Parkinson (1976); Fried (1977); Easton (1978); Leondes (1979); Milliken and Zoller (1980); Payne (1982); Parkinson and Gilbert (1983); Porter, et al. (1984); Ananda, et al. (1984); Bowen, et al. (1985); Kalafus, et al. (1986); and Ananda, et al. (1988).

2. PROGRAM OVERVIEW

The GPS program overview is shown in Figure 1. The program has three phases: Phase I, concept and validation; Phase II, full-scale development and system test; and Phase III, production and deployment. The Phase I program began in 1973 and ended in June 1979. The first DSARC (Defense System Acquisition Review Council) approved the program in December 1973. During Phase I, two navigation technology satellites (NTS-1 and NTS-2) were launched in 1974 and in 1977 as part of the Navy's program. The GPS satellite Block I contract was awarded in September 1974. During the concept validation phase, four Block I satellites were launched. After the successful completion of DSARC II in June 1979, the full-scale development and systems test phase began. Immediately after DSARC II, the Phase IIB full-scale engineering development contract for the user equipment was awarded in July 1979. In 1980 two more Block I satellites were launched. The Block II production satellite development contract was awarded in December 1980. The control system development contract was also awarded in 1980. The GPS program lost one of its satellites during the launch phase. In 1982 the GPS Block II production contract was awarded for 28 operational satellites.

Additional Block I satellites were launched: one in 1983, two in 1984, and the last one in 1985. Altogether, 11 Block I satellites were built and 10 were successfully launched. The operational control segment became operational in 1985 and at the same time the master control station located at the Consolidated Satellite Operational Center (CSOC), Falcon Air Force Base, Colorado Springs, began supporting and maintaining the Block I satellites. The second phase ended in 1985 and Phase III, production and deployment of operational satellites, began. In 1986 the Joint Requirements and Management Board (JRMB) approved the limited production of user equipments. But for the Space Shuttle Challenger disaster, the first Block II operational satellite would have been launched in early 1987. Since Challenger, the GPS launch strategy has been significantly changed. All the GPS satellites were planned for launch using the Space Shuttle. All of the GPS Block II satellites, except two, will be launched using a Delta II launch vehicle. The first launch is expected in October 1988. By the end of 1989, a constellation of nine operational satellites will be available, and in 1990 a global two-dimensional navigation capability is expected. Global three-dimensional navigation will be available by the end of 1991.

Of 28 Block II production satellites, the first 9 are known as Block II satellites and the rest (19 satellites) are classified as Block IIA satellites. Since the Block IIA satellites carry an additional sensor, these satellites are required to function continuously for a period of 6 months without any ground support. The Block II satellites need upload from the ground periodically for its momentum management. The Block II satellites may start to tumble between 28 to 45 days after the last ground contact. The Block IIA satellites have been designed to have autonomous momentum management capability for a period of 180 days without any ground contact.

The GPS full system accuracy is only available when the operational control system is functioning properly and navigation messages are uploaded on a daily basis. Once the control system is inoperable, the Block II satellites would provide navigation service to the users for a period of 14 days with gradually degraded accuracy and the

Block IIA satellites would provide navigation service to users for a period of 180 days with gradually degraded accuracy.

3. GPS AND OTHER NAVIGATION SYSTEMS

There exists a number of navigation systems. Comparisons between the various available navigation systems are made in Tables 1, 2, and 3. Transit is the only other functioning space-based global navigation system. The Soviet Union is developing and fielding a global satellite-based navigation system (GLONASS), which is extremely similar to the GPS. The accuracy of GLONASS is not known; however, it is anticipated that the GLONASS would provide navigation accuracy comparable to that of the GPS. Transit is a system developed by the U.S. Navy primarily to support marine navigation. As the GPS becomes operational, this dependance on Transit will become less and eventually the U.S. Navy will phase out of Transit. However, there is a large number of civilian users for the Transit system and it is not clear who would take the responsibility of maintaining and operating this system.

Navigation systems such as Loran C and D are regional and Omega is near global. However, Omega accuracy is on the order of kilometers and not useful for many military applications. Most systems other than inertial navigation are line-of-sight limited. A detailed comparison among the various navigation systems (in particular for military applications) is given in Tables 2 and 3.

4. GPS FUNDAMENTAL PRINCIPLES

Even though the GPS navigation concept is rather simple, it may be beneficial to review the underlying principles to understand how the GPS really works. The GPS is a one-way ranging system and, therefore, the ranging accuracy is extremely sensitive to the stability of the frequency oscillator at the transmitter. In order to understand the GPS, a simple case can be examined. At first, a two-dimensional problem can be constructed in which the user who needs to navigate wants to estimate the longitude and latitude of its location. From first principles, in order to solve for two parameters, one needs two independent measurements that are orthogonal, so that one can construct two linearly independent equations. In this simple case, let there be two widely separated transmitters with highly stable frequency oscillators (preferably atomic clocks) transmitting ranging signals, carrying some information regarding the time of transmission of the ranging signal. Let the user be a ship on the surface of the ocean interested only in estimating longitude and latitude and not the altitude. Let the user also carry a receiver capable of tracking this transmitted ranging signal to make range measurements from the transmitter. If the user also carries an atomic clock and both transmitting clock and user clock are a priori synchronized by some means, then the measured transit time from the transmitter to the rescurer would be the true range between the transmitter and the receiver.

The simple case discussed here is illustrated in Figure 2. The range measurements made by the receiver can be written as $D_1 = C \times \Delta T_1$ and $D_2 = C \times \Delta T_2$, where C is the the speed of light and ΔT_1, and ΔT_2 are the time delays for the ranging signal to travel from the transmitter to the receiver. Since the transmitters are permanently located, its coordinates in an earth-fixed reference system are available. Knowing the locations of the transmitters, measurement equations can be constructed and from which the two parameters--namely, the latitude and longitude of the user can be estimated.

However, the GPS is not based on transmitters permanently fixed at known locations. The GPS is a space-based system and the transmitters are on satellites that revolve around the earth. The spaceborne ranging system is illustrated in Figure 3. Even though the GPS satellites move continuously, it is possible by conventional tracking methods to estimate the orbit parameters of the satellites by which one can compute the position of the satellites in an earth-fixed reference system (just like the permanently-fixed ground transmitters), as a function of time. This position information, generally known as the ephemeris of the satellites, has to be continuously transmitted to the users. In this example, instead of characterizing a two-dimensional user, a three-dimensional user (such as a helicopter) is used. If there are three parameters to be estimated, one needs three independent measurements. Three satellites transmitting to the user, similar to the previous case, are shown in Figure 3. the user can make three measurements by computing the transit time for the ranging signal from each transmitter to the user. In this case also, it is assumed that the transmitters, as well as the user, carry an atomic clock and they are all a priori synchronized to each other. From the three range measurements, knowing the transmitter positions, three linear equations can be constructed and by solving these three equations the position components of the user can be estimated.

As discussed earlier, in order to make a true range measurement, both transmitter and the user should carry atomic clocks and the clocks should be synchronized independently. This is extremely difficult for a normal user. A typical user may not have access to an atomic clock and, moreover, it may be difficult to synchronize the clock with the transmitters. Therefore, it is assumed that the user only has a crystal clock and its initial time is not synchronized with the transmitter time. Now, when a measurement is made, the transmitter and the receiver have different reference times. The measured time delay from the transmitter to the receiver has two components. The first component is the transit time of the ranging signal and the second component is

Table 1. Navigation System Comparisons

Navigation System	Method Used: Coordinates Provided	Coverage Provided	Status of the System
Navstar GPS	Spherical Ranging: 3D Position 3D Velocity Precise Time	Global (24 hrs/day)	6 Satellites in 12-hr Orbits. Available Worldwide 1 to 4 hrs/day
Transit	Doppler Shift: Longitude Latitude	Global Except at the Poles (Periodic fixes only. Typically 1/2 to 2 hrs apart)	5 Satellites in Polar "Birdcage" Orbits >10,000 Sets in Use 80% Civilian Users
Loran C/D	Hyperbolic Ranging: Longitude Latitude	Regional Coverage: ~10% of Earth	8 Loran C Chains with 34 Transmitters Cover about 10% of the Earth
Omega	Hyperbolic Ranging: Longitude Latitude	Essentially Global 88% Coverage by Day 98% by Night	8 Transmitting Stations in Operation Worldwide
VOR/DME Tacan	Lighthouse Signal + Spherical Ranging: Heading Slant Range	Line of Sight Along Present Air Routes	More than 1000 Transmitters in Operation. At least 250,000 Users
ILS/MLS	Beam Steering: Heading Elevation Range	Line of Sight: 17 to 35 nmi Available Only At Properly Equipped Airports	Hundreds of Systems Operating Worldwide. 120,000 + Domestic Users
Inertial Navigation	Integrating Accelerometers: 3D Position 3D Velocity	Global wi[illegible] Periodic Updates	Thousands of Self-Contained Units in Use on Civilian and Military Planes and Ships
JTIDS RELNAV	Active and Passive Spherical Ranging: 2D Position 2D Velocity (Altitude if Geometry is Favorable)	Line of Sight: Available Only in Local Battlefield Areas	Developmental System for Use by A/F Vehicles in Conjunction with JTIDS COM NETS
PLRS	Active and Passive Spherical Ranging: 2D Position	Line of Sight: Available only in Local Battlefield Areas	Developmental System for Use by Army Users in Conjunction with JTIDS COM NETS

Table 2. Military Navigation System Comparison

System	Coverage	Accuracy	Common Grid	Passive User	User No. Limitation	Ambiguity	Selective Denial
Inertial	Worldwide	Time Variable	Singular	Yes	None	None	NA
Doppler	Worldwide	Time Variable	Singular	No	None	None	NA
Loran C	Limited ~14%	Moderate - 150m	None	Yes	None	Slight	No
Loran D	Theater Only	Moderate	None	Yes	None	Slight	No
Decca	Limited	Moderate	None	Yes	None	Slight	No
Omega	Worldwide	2000 m	None	Yes	None	Severe	No
Transit	Worldwide	50 m	WGS-72	Yes	None	Velocity Aid	No
Radio Beacon	Regional	Low	Singular	Yes	None	Slight	No
VOR/DME	Regional	Low	Singular	No	Limited	None	No
TACAN	Regional	Moderate	Singular	No	Limited	None	No
MLS/ILS	Terminal Area	Range Dependent	Singular	Yes	None	None	No
PLRS	Sector Net	30 m	Singular	No	Limited	Slight	Code Key
JTIDS RELNAV	Sector Net	30 m	Singular	No	Limited	Slight	Code Key
NAVSTAR GPS	Worldwide	16 m	WGS-84	Yes	None	None	Yes

------- IMPORTANT FOR COMBINED FORCES OPERATIONS -------

Table 3. Military Navigation System Comparison

System	Continuous Navigation	3 Dimension	Velocity Data	High Dynamic Operation	All Weather	Propagation Limitation	Autonomous U.S. Ground Control	Time Dissemination
Inertial	Yes	With Baro	Yes	Yes	Yes	None	NA	None
Doppler	Yes	Yes	Yes	Yes	Rain/Altitude	Altitude	NA	None
Loran C	Yes	No	Crude	Degradation	Moderate	Warp	Yes	Moderate
Loran D	Yes	No	Crude	Degradation	Moderate	Warp	Yes	None
Decca	Yes	No	No	Degradation	Moderate	Warp	No	None
Omega	Yes	No	No	Degradation	Moderate	Ionosphere	No	Slight
Transit	Cyclic	No	No	No	Moderate	Iono 2 Freq	Yes	Moderate
Radio Beacon	Relative	No	No	Degradation	Slight	Terrain	No	None
VOR/DME	Relative	No	Crude	Degradation	Slight	Terrain	No	None
TACAN	Relative	No	Crude	Degradation	Slight	Terrain	Yes	None
MLS/ILS	Relative	Yes	Crude	Yes	Yes	Terrain	No	None
PLRS	Relative	No	No	Slight	Yes	L.O.S.	Yes	Relative
JTIDS RELNAV	Relative	Limited	Crude	Slight	Yes	L.O.S.	Yes	Relative
NAVSTAR GPS	Yes	Yes	Yes	Yes	Yes	None	Yes	Yes

------- IMPORTANT FOR COMBINED FORCES OPERATIONS -------

the time offset between the transmitter clock and the receiver clock due to the non-synchronization of the clocks. The measured range can be written as $R_1 = C(\Delta t_1 + \Delta T)$, where Δt_1 is the transit time and ΔT is the time offset. This measured range is known as pseudo range rather than true range. The difference between pseudo range and true range is the apparent range error caused by the non-synchronization of the transmitter and receiver clocks.

Because it is assumed that all the transmitter clocks are synchronized among each other, the time offset between the receiver and any one of the transmitters is the same. This time offset is often known as timing bias and this is also an additional unknown parameter. Now, the navigation user has three position parameters and a time bias parameter to be estimated. When one has four parameters to be solved, these require four independent measurements. This requirement is illustrated in Figure 4. From the four measurements four linearly independent equations can be constructed and the four unknown parameters can be estimated.

As discussed earlier, the GPS requires all the transmitter clocks to be synchronized. In reality, the GPS satellite clocks are slowly but steadily drifting away from each other. Therefore, at any given time, the transmitter clocks are not per se synchronized. However, these clocks can be mathematically synchronized by external means. The GPS has defined a GPS master time as its reference time. The master time is maintained at the master control station. This GPS time is continuously monitored and related to the Universal Time Coordinate (UTC) maintained by the United States Naval Observatory (USNO). Each satellite time is related to the GPS time by a mathematical expression. The user corrects the satellite time to the GPS time by using the equation

$$t = t_{s/c} - \Delta t_{s/c}$$

where t is the GPS time in seconds; $t_{s/c}$ is the effective satellite time at signal transmission in seconds, and $\Delta t_{s/c}$ is the time offset between the satellite and the GPS master time. The time offset $\Delta t_{s/c}$ can be computed from the equation below.

$$\Delta t_{s/c} = a_0 + a_1 (t - toc) + a_2 (t - toc)^2 + \Delta t_r$$

where a_0, a_1, and a_2 are the polynomial coefficients representing the phase offset, frequency offset, and aging term of the satellite atomic clock with respect to the GPS master time, and Δt_r is the relativistic correction term (seconds). The parameter t is again the GPS time used in the previous equation and t_{oc} is the epoch time at which the polynomial coefficients are referenced and, generally, the t_{oc} is chosen at the midpoint of the fit interval. The polynomial coefficients a_0, a_1, and a_2 are estimated by the control segment for each satellite clock and periodically uplinked to the satellite. These coefficients are transmitted along with satellite ephemeris parameters (discussed below) to the navigation user as navigation messages. In Figure 5 the satellite clock correction parameters are designated as τ_i. By utilizing these clock correction terms, all satellite clocks are synchronized to the GPS master time. However, the error in synchronization will grow if polynomial coefficients a_0, a_1, and a_2 are not updated periodically. The operational baseline assumes an update rate of these parameters three times a day.

In addition to the clock parameters, a navigation user needs the instantaneous position values of the GPS satellites from which range measurements are made. These position values are provided to the user in the form of ephemeris parameters. These parameters are defined in Table 4. The control segment, by processing the tracking data acquired from the monitor stations, generates the orbit estimates for the GPS satellite. Thus, by integrating the equations of motion of the GPS satellites, predicted estimates of the satellite position coordinates are generated. These Cartesian position components are fit over a specified interval of time to compute the ephemeris parameters defined in Table 4. The control segment at each and every upload is required to uplink 14 days of navigation messages to satisfy the 14-day autonomy requirement for the Block II satellites and, for Block IIA satellites, the control segment is required to upload 180 days of navigation messages to meet the survivability requirement.

The navigation message will be uploaded such that a new message is provided to the user once every hour for the first day and once every 4 hr for the next 13 days. The navigation message for the first day is fit over a 4-hr fit interval such that there exists a 3-hr overlap of message and, for days 2 through 14, the fit interval is 6 hr so that the overlap period is 2 hr. The navigation message fit interval increases rapidly beyond 14 days and the only requirement is that the error introduced by the fitting process should be less than the error due to prediction.

A navigation user can compute the Cartesian position coordinates of the GPS satellites by employing the equations given in Table 5 (see also interface control document GPS-ICD-200). Both clock and ephemeris parameters are downlinked to the user at 50-bps data rate, modulated on both C/A and P-code navigation signals. The navigation message utilizes a basic format consisting of a 1500-bit-long frame made up of five subframes, each subframe being 300 bits long. Subframes 4 and 5 are subcommutated 25 times each, so that a complete data message takes a transmission of 25 full frames. Subframe 1 contains the clock parameters and subframe 2 and 3 contain the ephemeris parameters. Since subframes 1, 2, and 3 are repeated every 30 sec, it is possible for a user to update the clock and ephemeris parameters every 30 sec. Since subframes 4 and 5 have each 25 pages, these subframes are repeated only once in every 12.5 minutes.

Table 4. Ephemeris Data Definition

M_o	MEAN ANOMALY AT REFERENCE TIME
Δn	MEAN MOTION DIFFERENCE FROM COMPUTED VALUE
e	ECCENTRICITY
$(A)^{1/2}$	SQUARE ROOT OF THE SEMI-MAJOR AXIS
$(OMEGA)_o$	LONGITUDE OF ASCENDING NODE OF ORBIT PLANE AT WEEKLY EPOCH
i_o	INCLINATION ANGLE AT REFERENCE TIME
ω	ARGUMENT OF PERIGEE
OMEGADOT	RATE OF RIGHT ASCENSION
IDOT	RATE OF INCLINATION ANGLE
C_{uc}	AMPLITUDE OF THE COSINE HARMONIC CORRECTION TERM TO THE ARGUMENT OF LATITUDE
C_{us}	AMPLITUDE OF THE SINE HARMONIC CORRECTION TERM TO THE ARGUMENT OF LATITUDE
C_{rc}	AMPLITUDE OF THE COSINE HARMONIC CORRECTION TERM TO THE ORBIT RADIUS
C_{rs}	AMPLITUDE OF THE SINE HARMONIC CORRECTION TERM TO THE ORBIT RADIUS
C_{ic}	AMPLITUDE OF THE COSINE HARMONIC CORRECTION TERM TO THE ANGLE OF INCLINATION
C_{is}	AMPLITUDE OF THE SINE HARMONIC CORRECTION TERM TO THE ANGLE OF INCLINATION
t_{oe}	REFERENCE TIME EPHEMERIS

In addition to the precision ephemeris and clock parameters of the satellite, each satellite transmits almanac data of all satellites to the user, primarily to facilitate satellite acquisition and to compute geometric dilution of precision (GDOP) values to assist selection of satellites to achieve better accuracy. Pages 1 through 24 of subframe 5 contain the almanac data for each satellite through 24 and page 25 of subframe 5 contains the satellite health data for each satellite 1 through 24. The subframe 4, pages 2 through 5 and 7 through 10 contain almanac data for satellites 25 through 32. The page 18 of subframe 4 contains both the ionospheric data for a single frequency user and the conversion parameters from GPS time to UTC. The remaining pages of subframe 4 are reserved for other functions. For details of the navigation message format and bit structure the interface control document (GPS-ICD-200) should be consulted.

5. NAVIGATION SOLUTION

The basic navigation equations are non-linear and can be written as

$$[(X-X_i)^2 + (Y-Y_i)^2 + (Z-Z_i)^2]^{1/2} + T = R_i, \quad i = 1, 2, 3, 4$$

where X, Y, Z are user position components; X_i, Y_i, Z_i are the satellite position coordinates; and T is the range equivalent of the user clock offset.

Let $X = X_u + \Delta X$; $Y = Y_u + \Delta Y$; $Z = Z_u + \Delta Y$; and $T = T_u + \Delta T$ and corresponding $R_i = R_{ni} + \Delta R_i$, such that

$$R_{ni} = [(X_u - X_i)^2 + (Y_u - Y_i)^2 + (Z_u - Z_i)^2]^{1/2} + T_u$$

where,

X_u, Y_u, Z_u, T_u are nominal (a priori best estimate) values of X,Y, Z and T;

ΔX, ΔY, ΔZ, and ΔT are the corrections to these nominal values;

R_{ni} is the nominal pseudo-range measurements from the i^{th} satellite;

Table 5. Elements of Coordinate Systems

$\mu = 3.986005 \times 10^{14} \frac{\text{meters}^3}{\text{sec}^2}$	WGS 84 VALUE OF THE EARTH'S UNIVERSAL GRAVITATIONAL PARAMETER
$\dot{\Omega}_e = 7.2921151467 \times 10^{-5} \frac{\text{rad}}{\text{sec}}$	WGS 84 VALUE OF THE EARTH'S ROTATION RATE
$A = (\sqrt{A})^2$	SEMI-MAJOR AXIS
$n_o = \sqrt{\frac{\mu}{A^3}}$	COMPUTED MEAN MOTION - radians/second
$t_k = t - t_{oe}$*	TIME FROM EPHEMERIS REFERENCE EPOCH
$n = n_o + \Delta n$	CORRECTED MEAN MOTION
$M_k = M_o + nt_k$	MEAN ANOMALY
$M_k = E_k - e \sin E_k$	KEPLER'S EQUATION FOR ECCENTRIC ANOMALY (MAY BE SOLVED BY ITERATION) - radians
$v_k = \tan^{-1}\left\{\frac{\sin v_k}{\cos v_k}\right\} = \tan^{-1}\left\{\frac{\sqrt{1 - e^2} \sin E_k/(1 - e \cos E_k)}{(\cos E_k - e)/(1 - e \cos E_k)}\right\}$	TRUE ANOMALY
$E_k = \cos^{-1}\left\{\frac{e + \cos v_k}{1 + e \cos v_k}\right\}$	ECCENTRIC ANOMALY
$\Phi_k = v_k + \omega$	ARGUMENT OF LATITUDE
$\delta u_k = C_{us} \sin 2\Phi_k + C_{uc} \cos 2\Phi_k$	Argument of Latitude Correction — SECOND HARMONIC PERTURBATIONS
$\delta r_k = C_{rc} \cos 2\Phi_k + C_{rs} \sin 2\Phi_k$	Radius Correction — SECOND HARMONIC PERTURBATIONS
$\delta i_k = C_{ic} \cos 2\Phi_k + C_{is} \sin 2\Phi_k$	Correction to Inclination — SECOND HARMONIC PERTURBATIONS
$u_k = \Phi_k + \delta u_k$	CORRECTED ARGUMENT OF LATITUDE
$r_k = A(1 - e \cos E_k) + \delta r_k$	CORRECTED RADIUS
$i_k = i_o + \delta i_k + (\text{IDOT})t_k$	CORRECTED INCLINATION
$x_k' = r_k \cos u_k$ $y_k' = r_k \sin u_k$	POSITIONS IN ORBITAL PLANE
$\Omega_k = \Omega_o + (\dot{\Omega} - \dot{\Omega}_e) t_k - \dot{\Omega}_e t_{oe}$	CORRECTED LONGITUDE OF ASCENDING NODE
$x_k = x_k' \cos \Omega_k - y_k' \cos i_k \sin \Omega_k$ $y_k = x_k' \sin \Omega_k + y_k' \cos i_k \cos \Omega_k$ $z_k = y_k' \sin i_k$	EARTH FIXED COORDINATES

* t is GPS system time at time of transmission, i.e., GPS time corrected for transit time (range/speed of light). Furthermore, t_k shall be the actual total time difference between the time t and the epoch time t_{oe}, and must account for beginning or end of week crossovers. That is, if t_k is greater than 302,400 seconds, subtract 604,800 seconds from t_k. If t_k is less than 302,400 seconds, add 604,800 seconds to t_k.

and ΔR_i is the difference between the actual and nominal measurements.

Applying Taylor Series approximations, the basic equations can be linearized about the nominal values to obtain:

$$H X = r$$

Where $H = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & 1 \\ \alpha_{21} & \alpha_{22} & \alpha_{23} & 1 \\ \alpha_{31} & \alpha_{32} & \alpha_{33} & 1 \\ \alpha_{41} & \alpha_{42} & \alpha_{43} & 1 \end{bmatrix}$, α_{ij} $(i = 1,2,3,4_{ij} = 1,2,3)$

is the direction cosine of the angle between the range to i^{th} satellite and the j^{th} coordinate;

$$X = [\Delta X\ \Delta Y\ \Delta Z\ \Delta T]^T$$

and

$$r = [\Delta R_1\ \Delta R_2\ \Delta R_3\ \Delta R_4]^T$$

The linear equation Hx = r can be solved by

$$X = H^{-1} r$$

This would be an instantaneous solution or is often known as an estimate for point positioning.

For a dynamic user, a sequential processing of the measurements is best because better estimates can be obtained by utilizing the previous estimates and their associated uncertainties. Kalman filters are often used in user equipment for dynamic platforms.

6. SYSTEM STATUS AND FUTURE PLANS

Since the Space Shuttle Challenger disaster, the Department of Defense has reevaluated the need for expendable launch vehicles for DoD space missions and it was decided that most of the GPS satellites would be placed into orbit using expendable launch vehicles. The McDonnell Douglas Delta 2, as the medium launch vehicle (MLV), has been selected as the primary launch vehicle for the GPS (see Figure 6). The operational constellation of Block II GPS satellites will be placed in orbit beginning with the first launch in October 1988. The current plan shows that in 1989 there will be six launches, in 1990 another six launches, and in 1991 an additional six launches, out of which two would be launched using the Space Shuttle and all the rest of the satellites would be launched by Delta 2.

Reliability studies have shown that three to four additional launches, every year, are required to maintain a constellation of 18 satellites with 98 percent availability. There are, however, indications that the GPS constellation may consist of 24 satellites, rather than 21. In the event of a 24-satellite constellation, additional launches are required to maintain the desired number of satellites.

In addition to the procurement and deployment of GPS satellites, the GPS program office is also responsible for procuring user equipments. Currently, the program has a contract with Rockwell Collins for Limited Rate Initial Production. One-channel, two-channel, and five-channel user equipments are being built for the Air Force, Navy, and Army platforms. Over 2,000 units will be produced during this initial production phase and will be integrated into various platforms. During the next few years, the projected full rate production of user equipments will involve over 24,000 units.

The GPS program office has begun efforts to procure replenishment satellites for the Block II satellites. The replenishment satellites, classified as Block IIR, will have all the features of Block II satellites, with some additional capabilities. One of the primary enhancements planned for the Block IIR satellites is the autonomous navigation of GPS satellites utilizing crosslink ranging. The Block II satellites have crosslink communications capability, but no ranging. The Block IIR satellites will be modified to enable crosslink ranging on the same crosslink frequency and, by processing the crosslink range measurements, the GPS navigation message can be generated onboard the satellite without daily upload from the ground. Analysis has shown that this automous navigation capability can maintain a user navigation accuracy without any significant degradation for a period of about 6 months with no ground contact (see Ananda, et al., 1984).

In addition to the autonomous navigation feature, the Block IIR satellites will have increased survivability, will be more reliable, and will incorporate state-of-the-art technology for better satellite structure and subsystem components, such as microprocessors and ASICS (Application Specific Integrated Circuits). The GPS program has awarded contracts to both Rockwell and General Electric for the Phase I (system design) effort and one of the contractors will be selected for the Phase II (development and production) effort. In Phase II, it is expected that 20 Block IIR satellites will be bought with an option to buy six more satellites. The first Block IIR satellite is to be launched in 1995, unless the Block II satellites survive beyond their expected design life.

7. SUMMARY

The paper summarizes the historical evolution of the GPS and both military and civil applications. The GPS concept validation phase, full-scale engineering development, and system test phase and operational system phase have been described. A brief review of the GPS principles and navigation solution has been given. The current system status and future plans have been discussed.

REFERENCES

1. Fried, W. R. "A Comparative Performance Analysis of Modern Ground-Based, Air-Based, and Satellite-Based Radio Navigation Systems," Navigation, Vol 24, No. 1, Spring 1977.

2. Leondes, C. T., Editor, "Principles and Operational Approach of Precision Position Determination Systems," NATO, AGARDograph No. 245, July 1979.

3. Milliken, R. J., and Zoller, C. J. "Principles of Operation of Navstar and System Characteristics," the ION, GPS Special Issues, 1980.

4. Easton, R. G. "The Navigation Technology Program," J. of ION, Vol 25, 2, 1978.

5. Parkinson, B. W., "Navstar Global Positioning System (GPS)," Proc. National Telecomm. Conference, 1976.

6. Parkinson, B. W. and S. W. Gilbert, "Navstar: Global Positioning System - Ten Years Later," Proc. IEEE Conf., Vol 71, 10, 1983.

7. Payne, C. R., "Navstar: Global Positioning Systems: 1982," Proc. National Telesystems Corp., Galveston, Texas, 1982.

8. Porter, J., et al., "GPS System Overview," paper presented at Royal Inst. of Navigation, London, 1984.

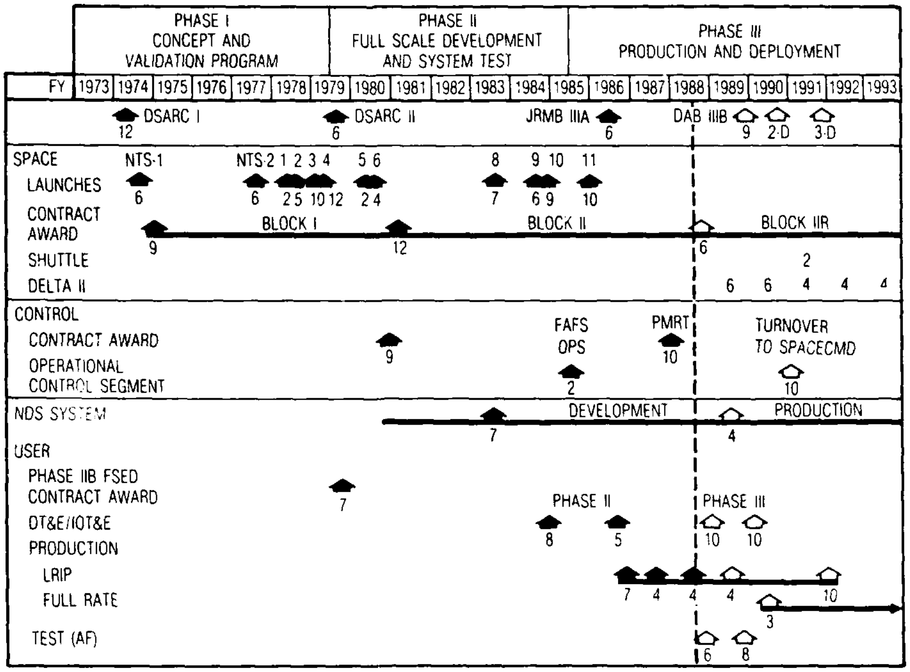
9. Institute of Navigation GPS Special Issues, Vol I, 1980; Vol II, 1984; Vol III, 1986.

10. Ananda, M., et al., "Autonomous Navigation of the Global Positioning System Satellite," paper presented at AIAA Guidance and Control Conference, Seattle, Washington, August 20-23, 1984.

11. Bowen, R., et al., "Global Positioning Systems Operational Control Systems Accuracies," Navigation: Journal of the Inst. of Navigation, Vol 32, 2, 1985.

12. Kalafus, R. M., et al., "Special Committee, 104 Recommendations for Differential GPS Service," Journal of the ION, Vol 33, 1, Spring 1986.

13. Ananda, M., et al., "The Global Positioning System (GPS) Autonomous User System," paper presented at ION National Technical Meeting, Santa Barbara, 1988.



DSARC – DEFENSE SYSTEM ACQUISITION REVIEW COUNCIL DAB – DEFENSE ACQUISITION BOARD JRMB – JOINT REQUIREMENTS AND MANAGEMENT BOARD

Figure 1. NAVSTAR Program Overview

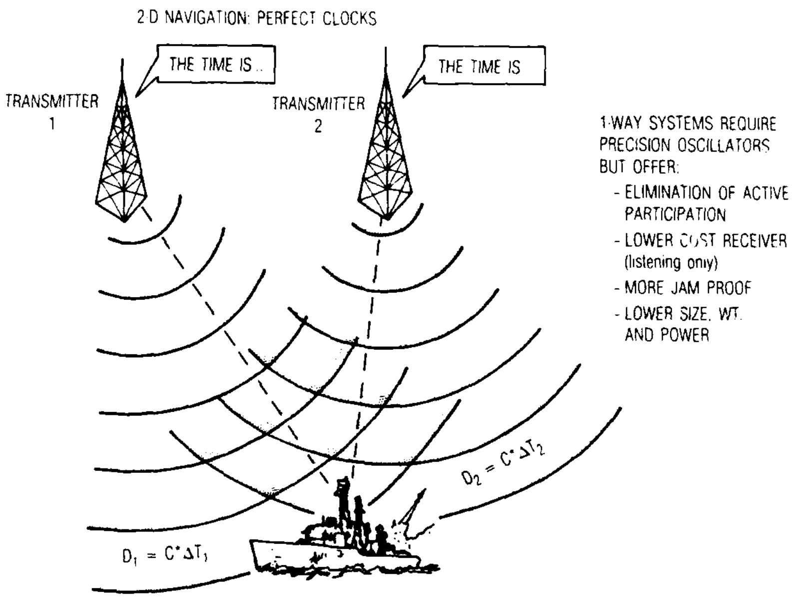


Figure 2. Electronic One-Way Ranging System

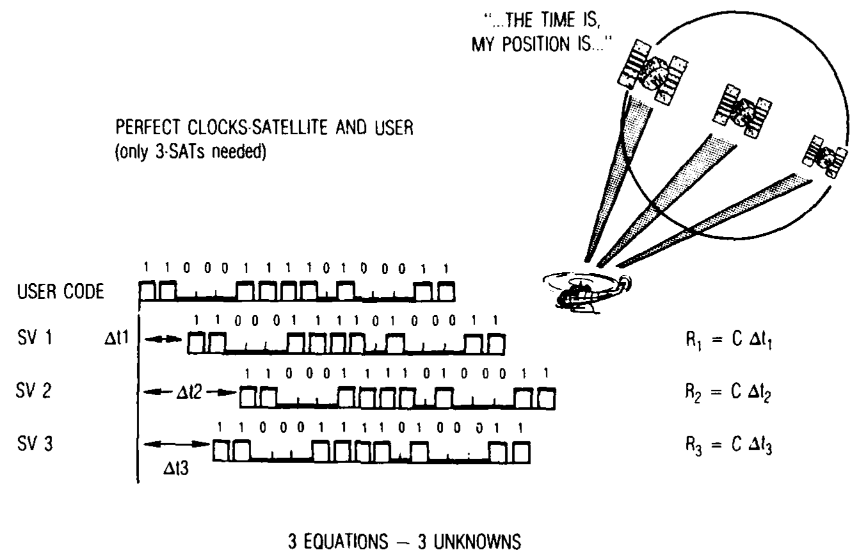


Figure 3. Spaceborne One-Way Ranging

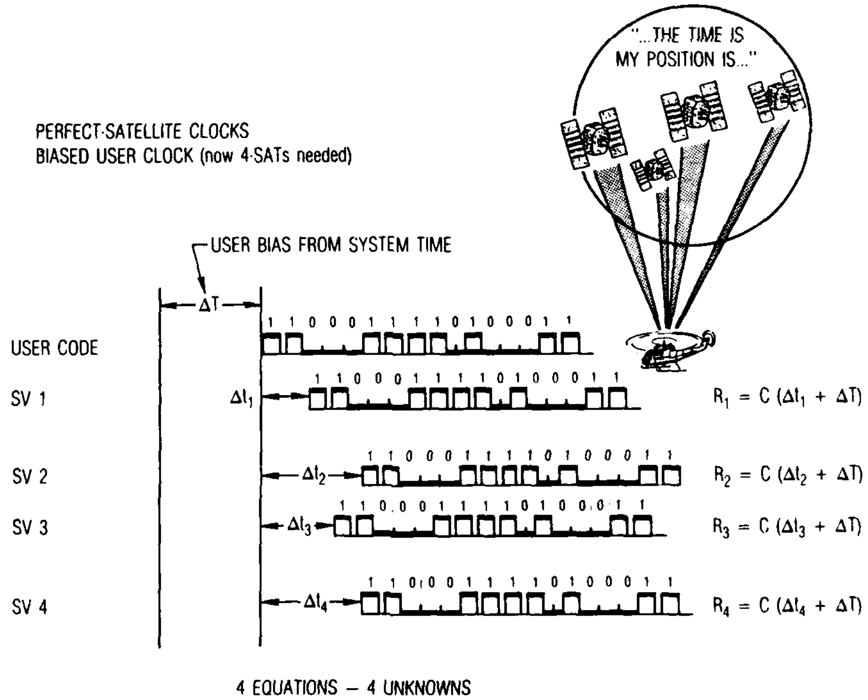


Figure 4. Ephemeris Parameters

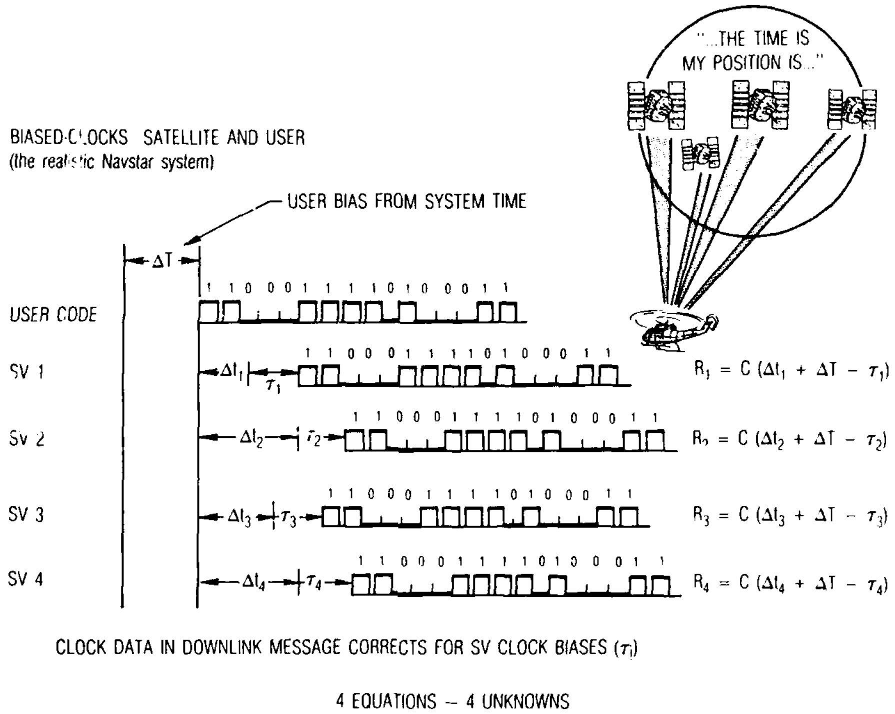


Figure 5. Clock and Ephemeris Parameters

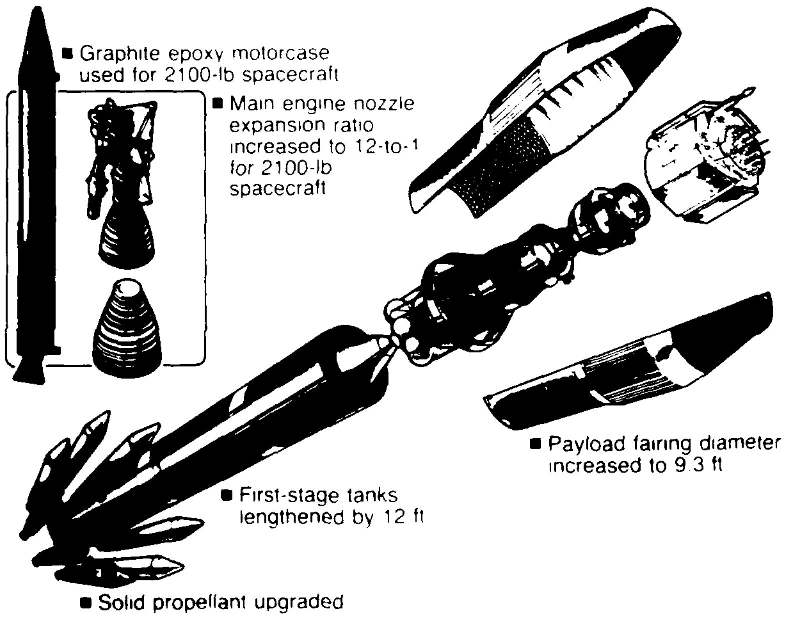


Figure 6. GPS Launch

THE GLOBAL POSITIONING SYSTEM (GPS) CONSTELLATION AND COVERAGE

MOHAN ANANDA
The Aerospace Corporation
El Segundo, California, U.S.A.

ABSTRACT

The Navstar Global Positioning System (GPS) is a space-based system that will be operational in 1990 and will enable a user to determine position, velocity, and time with greater accuracy on a worldwide basis than it has ever been achieved. The GPS is currently undergoing full-scale engineering development and testing. This paper reviews the test constellation and available coverage for testing the various system segments. This paper also discusses the 18-, 21-, and 24-satellite operational constellations. The current baseline constellation consists of 18 satellites with three active spare satellites. The concept of constellation value and various dilution-of-precision parameters due to geometry are discussed, such as geometrical dilution of precision (GDOP) and position dilution of precision (PDOP). The constellation design issues related to minimizing the reduced accuracy regions due to geometry problems are also reviewed in this paper. Coverage maps over the globe are presented for the test constellation, as well as for the operational constellations.

1. *INTRODUCTION*

Navstar Global Positioning System (GPS) is a space-based navigation system that provides accurate position, velocity, and time information to any user located at any region of the earth by acquiring and processing the continuously emitted radio signals from the GPS satellites. The system concept was the result of research and development over the years by the Department of Defense in the area of navigation technology. The GPS program is managed by the United States Air Force as a joint program of all the United Stated Armed services, NATO, the Defense Mapping Agency, and the Department of Transportation.

A review of the history and evolution of the navigation technology program of the Department of Defense can be found in the paper by Easton (1978). The Navstar GPS is a result of research and development efforts of the Navy on the Timation program and the Air Force on the project 621B. In the early years of the navigation technology program, navigation technology satellites (NTS) were launched primarily to provide information regarding the performance of atomic clocks in the space environment. There exists a number of publications giving details of the GPS concept and a GPS special issue of the Journal of the Institute of Navigation (summer, 1978) contains some of these publications. An overview of the Global Positioning System concept is provided by Parkinson (1976). Details of principles of operation and system characteristics are given by Milliken and Zoller (1978). Review of enhanced system capabilities is given by Ananda (1981). Payne (1982) provides details regarding the GPS full scale engineering development phase. Parkinson and Gilbert (1983) looks at the GPS program over the 10 years from its inception and provides the significant evolutionary changes over that time period. Even today the system capability and system configuration are continuously evolving.

An overall system review is given by Porter, et al. (1984). A brief discussion of the system is given here to facilitate the reader in understanding the issues related to constellation design and coverage requirements. Details of the system requirements regarding navigation accuracy are not included in this paper; accuracy requirements will be discussed in a companion paper entitled GPS Accuracy, System Error Budget, Space and Control Segment Overview.

The baseline GPS program is divided into three phases: I. concept validation phase, II. system validation phase, and III. production phase. The concept validation phase and the system validation phase have been completed. The program is currently in its production phase. However, because of the Space Shuttle Challenger disaster, the production phase has been delayed by a couple of years to achieve global navigation capability.

2. SYSTEM OVERVIEW

The basic system elements of the GPS are the space segment, the control segment, and the user segment. The system elements and the overall system concept are illustrated in Figure 1. The space segment consists of satellites placed in a 12-hr period orbit. The placement of satellites is designed to provide global coverage for a user. The satellite radiates two spread spectrum pseudo-random noise (PRN) radio signals. The signal consists of a C/A (coarse acquisition) code at 1.023 MHz and a P (precision) code at 10.23 MHz bandwidths. The signals are transmitted at two frequencies, L_1 (1575 MHz) and L_2 (1227 MHz). Both are coherently derived from highly stable onboard atomic clocks. Both C/A and P-codes are transmitted on the L_1 frequency, whereas either C/A or P-code is transmitted on the L_2 frequency. The selection is achieved by ground command. The two frequencies are utilized primarily to calibrate the delay in range measurements due to ionospheric effects.

The C/A code is available to all users, however the P-code may be available to only authorized users because of the anti-spoof (AS) and selective availability (SA) features. Therefore, an authorized user will have access to C/A code on the L_1 frequency because, generally, only the P-code is transmitted on the L_2 frequency. In addition to the PRN range codes, 50-bps data, which consist of the navigation message comprising both ephemeris and clock parameters (Van Dierendock, et al., 1976), are modulated onto the PRN sequence on both L_1 and L_2 frequencies.

The control segment consists of the master control station, which is located in Colorado Springs and is responsible for all the data processing, monitor stations, and ground antennas. The five monitor stations are located globally at Ascension Island, Diego Garcia, Kawajalein, Colorado Springs, and Hawaii. The radiometric data are tracked by the monitor stations and transmitted to the master control station by communication lines. Accurate ephemeris and clock parameters for the GPS satellites are estimated by extensive data processing at the master control station. Then predicted clock and ephemeris information in the form of navigation message is transmitted uplink to the satellites via the ground antennas. The ground antennas are co-located with three of the monitor stations at Ascension Island, Diego Garcia, and Kwajalein. To ensure the system accuracy, the predicted navigation message is uplinked to each satellite three times daily. The control segment is also responsible for maintaining the health and welfare of satellites .

The user segment consists of equipment with antenna, receiver, signal processing, and data processing capabilities. The user equipment generates the pseudo-range measurement by tracking the satellite navigation signal by generating an identical code of the transmitted signal and slewing the code until it correlates with the incoming signal. The amount of slew is a measure of the transit time between the time the signal is transmitted by the satellite and the time the signal is received by the user equipment. This transmit time is not identically equivalent to the true range between the satellite and receiver because the transit time is measured by the satellite clock, whereas the received time is measured by the clock in the user equipment. Since both satellite clock and the user clock are not a priori synchronized, there exists a clock bias between the satellite clock and the user clock. Therefore, the measured transit time includes the true time of travel between the satellite and the receiver, which represents the true range between the satellite and the receiver, and the clock bias between the satellite clock and the user clock. Thus, the measured quantity represents a pseudo range rather than the true range.

The conventional near-earth user equipment measures pseudo-range and pseudo range rate or often called accumulated delta range (ADR) from the navigation signal. The data in Figure 2 illustrate how user equipment can estimate navigation parameters by tracking four satellites. In order to estimate the position component and the user clock bias, the user equipment has to observe four GPS satellites because four orthogonal measurements are required to solve for four parameters. Similarly, by observing pseudo-range rate measurements, one can estimate velocity components as well.

3. TEST CONSTELLATION

Both the concept validation phase (Phase I) and the system validation phase (Phase II) utilized the Block I space vehicle as shown in the Figure 3. Eleven Block I space vehicles were launched, however, one of the space vehicles was lost due to launch vehicle malfunction. Ten Block I vehicles have been launched into orbit over the last 9 years. The test constellation was designed such that best coverage is available for a maximum period over the Yuma Proving Ground for testing. The graphics of Figure 4 indicate various locations of the space vehicles in the test constellation. The test constellation consists of two planes separated by 120 deg; and the inclination of the orbital plane is 63 deg. Navstars 1, 4, 5, 8, 9, and 11 are placed in one orbital plane. The Navstar 1 navigation payload failed after about one year of launch. Navstar 4 operated for over 7 years and recently the last atomic clock failed; however, the crystal clock is functioning. The control system periodically uploads the navigation message and maintains a user range error of about 100 m. Navstar 5 operated for about 4 years until the reaction control wheels failed. Navstars 8, 9, and 11 are functioning well.

Navstars 2, 3, 6, and 10 are placed in another orbital plane separated by 120 deg from the first orbital plane. Navstar 2 operated for about 2.5 years. Navstar 3 has been operating for about 9 years. The battery and the power subsystem manifests problems when the satellite goes in and out of eclipses. During the next eclipse season the navigation payload will be probably switched off, depending on power system performance. Navstar 6 has been operating for about 6 years and it is performing well. Navstar 10 has been in operation for about three years and it is functioning very well.

The design life of Block I space vehicles is 5 years and the estimated mean mission duration is 4 years. However, the average active life of the Block I space vehicle is over 5 years.

The correct status of the test constellation as of the date of this paper (October 1987) is given in Figure 5. The space vehicles are numbered as Navstar 1 through Navstar 11. However, each space vehicle is also identified by a PRN number that is different from the Navstar number. The information in Figure 5 also provides the age of each space vehicle and the status of the various subsystem elements, such as the navigation payload, electrical power system, attitude control system, and telemetry and control systems. It is clear that most of the space vehicles have operated beyond the Block I space vehicle design life of 5 years.

Of the ten space vehicles in orbit, Navstars 6, 8, 9, 10, and 11 are in reasonably good health and Navstar 3 is in marginal health. Recently, Navstar 11 has been moved to facilitate better coverage at the Yuma Proving Ground, where the user equipments are currently being tested. The coverage on May 31, 1987, for a test constellation without and with Navstar 4, is shown in Figures 5 and 6. The coverage on September 13, 1987, with Navstar 4, is shown in Figure 7. Because the space vehicle Navstar 11 was moving in relation to the rest of the constellation, the coverage region changed over the time period.

The data of Figure 8 provide coverage for the test constellation without Navstar 4 during October 1987. Coverage including Navstar 4 in the test constellation is given in Figure 9. It can be seen that coverage improves significantly when Navstar 4 is added to the constellation. Coverage when both Navstars 3 and 4 are excluded from the constellations is displayed in Figure 10. Coverage becomes poor and this may happen before the next launch.

4. OPERATIONAL CONSTELLATION

The problem of constellation design for optimum coverage has been extensively studied by several authors. The papers by Walker (1977); Book, et al. (1980); Jorgensen (1980); Brady and Jorgensen (1981); Kruh (1981); Kruh (1982); Kruh, et al. (1983); Porter et al. (1984); and Stein and Wheaton (1986) all provide detailed analysis of the constellation design issues. A brief review of the essential features related to designing an optimal constellation is outlined here.

5. SYSTEM AVAILABILITY AND CONSTELLATION VALUE

System availability is defined as the probability of a particular number of space vehicles (18 or more) operating within the specified requirements at any given time. The GPS Program Management Directive requires that system availability for the GPS be 0.98 with 18 or more satellites functioning within specifications. Therefore, the total number of satellites to be launched and maintained in operation is derived from the system availability requirement.

In order to define constellation value, one has to first understand position dilution of precision (PDOP), a parameter which reflects the geometry of the space vehicles in relation to the user. The PDOP parameter can be derived from the basic navigation equations. Let the user be at x, y, z in an earth-fixed, earth-centered coordinate system and let the space vehicles be at x_i, y_i, z_i, i = 1, 2, 3, 4 in the same coordinate system as the user; the measurement equations can be written as follows.

$$(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2 + T = R_i, \quad i = 1, 2, 3, 4$$

where T is the user clock bias and R_i, i = 1, 2, 3, 4 are the pseudo-range measurements from each satellite.

Even though, the above equations are non-linear, it can be linearized without loss of accuracy. The linearization is accomplished by assuming $x = x_n + \Delta x$, where x_n denotes nominal and Δx is the linear correction. By proper substitution and algebraic manipulation, these linearized equations can be conveniently written in matrix notation,

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & 1 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & 1 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & 1 \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & 1 \end{bmatrix} \times \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta T \end{bmatrix} = \begin{bmatrix} \Delta R_1 \\ \Delta R_2 \\ \Delta R_3 \\ \Delta R_y \end{bmatrix}$$

where σ_{ij} is the direction cosine of the angle between the range to the i^{th} coordinate. The above matrix equation can be rewritten as

$$Ax = r$$

where A is the information matrix, x is the vector of estimates, and r is the measurement residual vector.

It is possible to estimate the covariance of the state (estimated parameters) vector by

$$Cov(x) = A^{-1} \; Cov(r) \; A^{-T}$$

and, rewritten, is

$$Cov(x) = [A^T \; Cov(r)^{-1} \; A]^{-1}$$

It is also possible to assume that the pseudo-range measurement has an error (1σ) of unity and the expected means of the measurement error is zero, as well as that the correlation of errors between the satellite measurements is zero. Now, the covariance equation reduces to

$$Cov(x) = [A^T A]^{-1}$$

For a navigation user, the above equation represents the fact that if the measurements are uncorrelated with unit error, the error in the estimation is purely dependent upon the geometry of the satellites with respect to the user because the A matrix consists of only direction cosines of the angle between the range to the satellites and coordinates. Therefore, a measure of the "goodness" of the geometry is determined by the term geometrical dilution of precision (GDOP), computed as the square root of the trace of the covariance matrix.

$$GDOP = [Trace \; [A^T A]^{-1} \;]^{1/2}$$

let σx^2, σy^2, σz^2, and σT^2 be diagonal components of the coverage matrix, then

$$GDOP = [\sigma x^2 + \sigma y^2 + \sigma z^2 + \sigma T^2]^{1/2}$$

There are other related dilutions of precision parameters:

$$PDOP = [\sigma x^2 + \sigma y^2 + \sigma z^2]^{1/2}$$

$$HDOP = [\sigma x^2 + \sigma y^2]^{1/2}$$

$$VDOP = \sigma z$$

$$TDOP = \sigma T$$

where PDOP is the position dilution of precision (this parameter is often used in constellation design problems), HDOP is the horizontal dilution of precision, VDOP is the vertical dilution of precision, and TDOP is the time dilution of precision. All DOPs are, in effect, the amplification factors of pseudo-range measurement errors due to the effect of satellite geometry.

For defining constellation value, the parameter PDOP has been used as a geometric measure of performance. From the analysis above, it is clear that the lower the PDOP, the higher the accuracy that can be achieved. Constellation value is defined as the percentage of occurrences for which PDOP is less than or equal to 6 for all sample points on the earth over a 24 hr period. The constellation value is computed by generating sample points worldwide at regular time intervals during a 24-hr period. Sample points can be generated either by random points or a fixed grid of points. These samples are taken every few minutes. Equal area sample points can be generated by properly weighing the longitude at higher latitudes. Once the equal area sample points are generated, the best PDOP from the satellites in view for the particular constellation under consideration is determined for each sample point. This process is repeated for all sample points over the 24-hr period. Then the ratio of the number of sample points that have equal or less than the threshold PDOP value (e.g., 6) to the total number of sample points is the constellation value. For an ideal constellation, the constellation value should be one. If the constellation value is less than one, there exist regions where the PDOP is greater than the threshold value, indicating degraded performance.

6. 18-SATELLITE CONSTELLATION

In the early period of the GPS program, a 24-satellite constellation was considered for continuous global coverage. However, because of funding constraints, studies were made to evaluate whether near-continuous global coverage can be achieved with a minimum number of satellites. Extensive analysis resulted in a 18-satellite constellation that provides reasonable coverage. This constellation is denoted by Walker (1977) notation N/P/V (18/6/2), where N is the number of satellites (18), P is the number of orbital planes (6), and V is the parameter that defines the phasing angle between two adjacent satellites in the adjacent planes. The phasing angle $\theta = 360/N \times V$ (V = 0, 1, 2...). The phasing angle θ becomes 40 deg for V = 2. The constellation also assures that the orbital planes are uniformly distributed in longitude.

An illustration of how the operational constellation would look is provided in Figure 12. The operational constellation configuration is shown in Figure 13. There are six orbital planes, each inclined to 55 deg. The current test constellation of Block I satellites consists of orbital planes with an inclination of 63 deg. For the operational constellation with Block II satellites, the inclination of 55 deg was selected because of launch constraints imposed by the Space Shuttle, even though the optimum inclination for coverage is not 55 deg and the current baseline is 55 deg. because of changes in the launch strategies since the Space Shuttle disaster, studies are under way to increase the orbital inclination from 55 deg to 60 deg.

There are three satellites in each orbital plane in the 18 satellite constellation. The closed circles shown in Figure 13 are the operational satellite locations and they are identified by their position numbers. Also shown in Figure 13 are the locations of the active spare satellites (indicated by open circles). These three spare satellites are located in every other plane to increase system survivability, as well as to generate a 98 percent availability of 18 or more satellites (Program Management Directive requirement). The locations of spare satellites are chosen to improve coverage in the continental United States. The location of Block I satellites in the test constellation are also given in Figure 13, as indicated by the diamonds. There are six Block I satellites shown in the figure. However, based on reliability analysis, there may be only four working Block I satellites available when the operational constellation buildup begins. The inclination of the test constellation satellites is 63 deg and, therefore, different from that of the operational satellites.

The strategy for the buildup of the operational constellation is such that the constellation would give maximum coverage at Yuma, where the user equipment tests are conducted, as well as optimum worldwide coverage. This is difficult to achieve because of competing objectives. However, the current baseline for the buildup of the operational satellites is to launch the Block II satellites first to fill the test locations that are vacant and then proceed with the rest of the buildup. Studies have been conducted to evaluate when the test constellation satellites should be rephased. Analysis shows that optimum coverage is achieved when the constellation is rephased after launching the ninth satellite. The current baseline buildup strategy is shown in the Figure 14. The rephasing of the test constellation satellites is shown by the arrow in the Figure 13. The satellite will be moved from its current location to the new location, which is dedicated to the operational constellation. The launch sequence and the location of each satellite are shown in the Figure 14. The worldwide constellation value at each time a new satellite is added to the constellation is also shown in Figure 14.

Composite coverage for the 18-satellite constellation, for a 24-hr period, is displayed in Figure 15. This is a composite of all the regions of degraded coverage, where the PDOP is greater than the threshold value of six. At any given time there are only a few such regions for the entire globe. In such regions, the PDOP becomes unbounded for a very few minutes. There will be no such region for a period of about 40 min and, then again, a few such regions occur somewhere else on the globe. Therefore, Figure 15 is a composite of all such regions over a period of 24 hours. This coverage map is generated by assuming an elevation cutoff angle of 5 deg, this assumption limits the number of satellites in view for a user. The degraded coverage region increases considerably when the elevation cutoff angle is increased and, similarly, the degraded region becomes smaller and smaller if the mask (elevation cutoff) angle is reduced. If the mask angle is about 2 deg, there are no such degraded regions at all. However, from a practical point of view, it is often difficult to track the satellite at such low elevation angles.

7. 18-SATELLITE CONSTELLATION WITH 3 ACTIVE SPARES

As was discussed earlier, the program management directive (PMD) requires a system availability of 18 or more satellites 98 percent of the time. Therefore, analyses were done to establish the minimum number of satellites needed in the constellation to meet the PMD requirement. If it was determined that, with an 18-satellite constellation, the system availability of 18 satellites would be about 50 percent and it would take a 21-satellite constellation to generate a 98-percent availability of 18 or more satellites. Based on this study a 18/6/2 plus 3 active spares constellation became the baseline constellation. However, efforts are under way to improve this constellation and it is possible that a new revised constellation may be defined by the time the operational constellation buildup begins.

The composite coverage of the degraded regions, where the PDOP value exceeds 6 for the 18/6/2 + 3 active spares constellation with an orbital inclination of 55 deg, is shown in Figure 16. A comparison of Figures 15 and 16 would reveal that the coverage provided by the additional three satellites has eliminated the degraded region from the continental United States. In most cases, the duration of degradation is about half an hour. Both the region of degradation and the time of degradation can be precomputed and, therefore, missions can be planned with this a priori knowledge. However, certain missions that cannot be planned do suffer due to the degraded coverage (reduced accuracy regions).

A sensitivity analysis has been performed to examine whether the coverage is sensitive to the orbital inclination. The result of the study is shown in the Figure 17, in which the constellation value as a function of the orbital inclination is shown. It is clear that both the 18-satellite constellations (with and without the spares) show some sensitivity to the orbital inclinations. The constellation value is maximum for the orbital inclinations of about 60.5 to 63 deg. The graphics of Figure 18 show the composite coverage of the degraded regions for the 18 + 3 constellation with an orbital inclination of 60.5 deg. It is clear, by comparing Figure 18 and 16, that much of the high latitude degradation regions are eliminated by changing the inclination. In addition, the duration of the degraded coverage is reduced significantly by increasing the inclination. Therefore, it is certainly desirable to achieve an orbital inclination of about 60 deg. Launch vehicle constraints are currently being evaluated as to whether a 60-deg inclination is achievable.

The composite coverage of the baseline 18 + 3 constellation, with an inclination of 55 deg and assuming a PDOP cutoff of 10 and 20, respectively, is given in Figures 19 and 20. These figures show slight reduction in the area of the degraded regions. However, most of the degraded regions are dominated by high PDOP values.

A rephased 21-satellite constellation has been designed by Massatt (personal communications) for optimization of the coverage. Composite coverage for the rephased 21 satellites with inclinations of 55 and 60 deg, respectively, is shown in Figures 21 and 22.

The rephased constellation removes 95 percent of the degraded coverage that was found in the baseline constellation and has an improved robustness in the event of satellite failure. Only 3.3 percent of the earth experiences any degraded coverage (PDOP > 5) for some amount of time. Only 1 percent of the earth experiences degraded coverage for more than 5 min, and the maximum degraded coverage lasts for only 12 min. Even when the coverage is degraded, the largest PDOP value is less than 9 and, therefore, the geometry never becomes singular. This rephased 21-satellite constellation, developed by Massatt, seems extremely interesting to the GPS Joint Program Office; it is possible that this newly rephased 21 satellite constallation may soon become the baseline constellation for GPS operations.

8. SUMMARY AND CONCLUSIONS

A brief review of the test constellation, constellation design techniques, and operational constellations has been presented. It has been shown that navigation accuracy is sensitive to satellite geometry and the primary objective of constellation design is to optimize constellation value with the minimum number of satellites. Examination of various constellations would reveal that the newly rephased 21-satellite constellation is the optimal constellation to date to achieve continuous global navigation coverage. It is possible to further refine this constellation, and efforts to optimize this constellation will continue.

REFERENCES

1. N. Ananda. "Navstar Global Positioning System (GPS): Future Enhanced Capabilities," (1981), Proc. International Symposium, Space Flight Dynamics, Darmstadt, FRG.

2. S. A. Book, et al., "The Navigation GPS Constellations," (1980), Proc. PLANS, IEEE.

3. W. F. Brady and P. S. Jorgensen, "Worldwide Coverage of the Phase II Navstar Satellite Constellation," (1981), ION National Aerospace Meeting.

4. R. L. Easton, "The Navigation Technology Program," (1978), J. Institute of Navigation, Vol. 25, No. 2.

5. P. Jorgensen, "Navstar Global Positioning System: 18 Satellite Constellation," (1980), Proceedings ION 36th Annual Meeting, Monterey, California.

6. P. Kruh, "Coverage and Buildup of the Navstar Constellation," (1982), NATO AGARD Conference Proceedings, No. 344.

7. P. Kruh, "The Navstar Global Positioning System Six-Plane, 18 Satellite Constellation," (1981), Navational Telecommunications Conference.

8. P. Kruh, et al., "A Strategy for Buildup to the Operational Navstar GPS Constellation," (1983), ION National Aerospace Meeting.

9. R. J. Millikan and C. J. Zoller, "Principles of Operating Navstar and System Characteristics," (1978), J. Institute of Navigation, Vol. 25, No. 2.

10. J. Parker, et al., "GPS System Overview," (1984), paper presented at Royal Institute of Navigation Meeting, London.

11. B. W. Parkinson and S. W. Gilbert, "Navstar: Global Positioning System — Two Years Later," (1983), Proc. IEEE Conference, Vol. 71, No. 10.

12. B. W. Parkinson, "Navstar Global Positioning (GPS)," (1976), Proc. National Telecomm. Conference.

13. C. R. Payne, "Navstar Global Positioning System: 1982," (1982), Proc. National Telesystems Conference, Galveston, Texas.

14. B. Stain and E. Wheaton, "Equatorial Spacing of the Navstar/GPS 18-Satellite Constellation," (1986), Proc. of PLANS, IEEE.

15. A. T. Van Dierendock, et al., "The GPS Navigation Message," (1978), J. Institute of Navigation, Vol. 25, No. 2.

16. J. G. Wolker, "Continual Whole-Earth Coverage by Circular Orbit Satellite Patterns," (1977), Technical Report 77044, Royal Aircraft Establishment Farnborough., England.

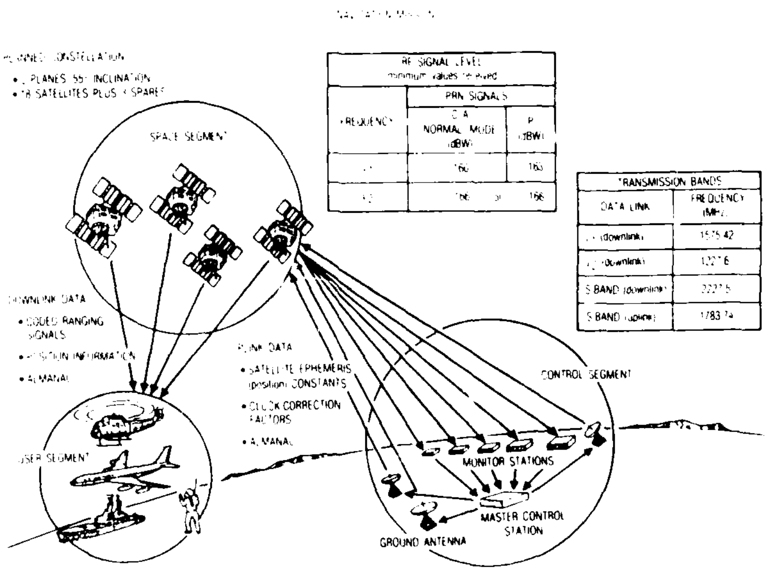


Figure 1

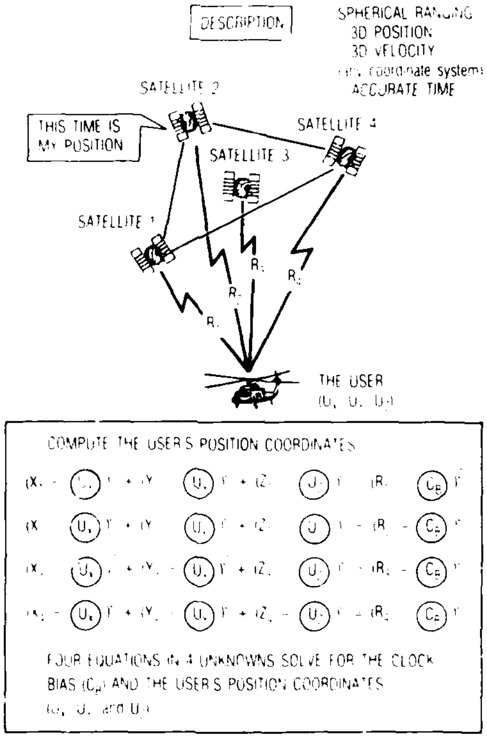


Figure 2

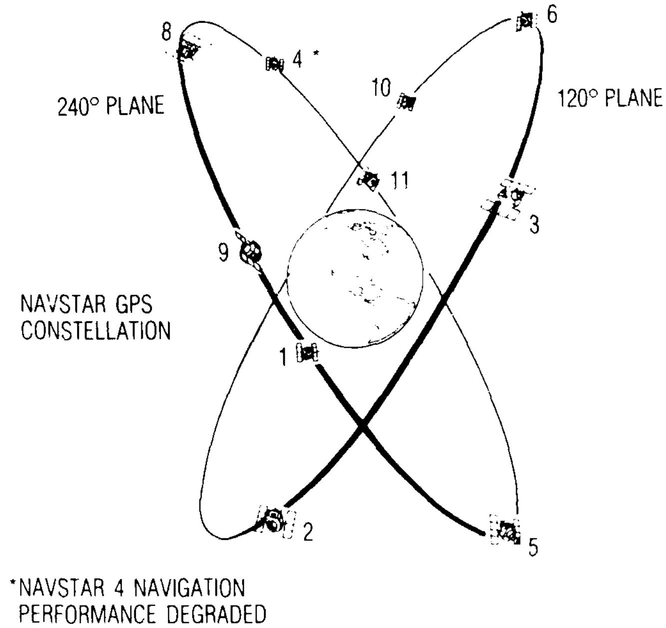
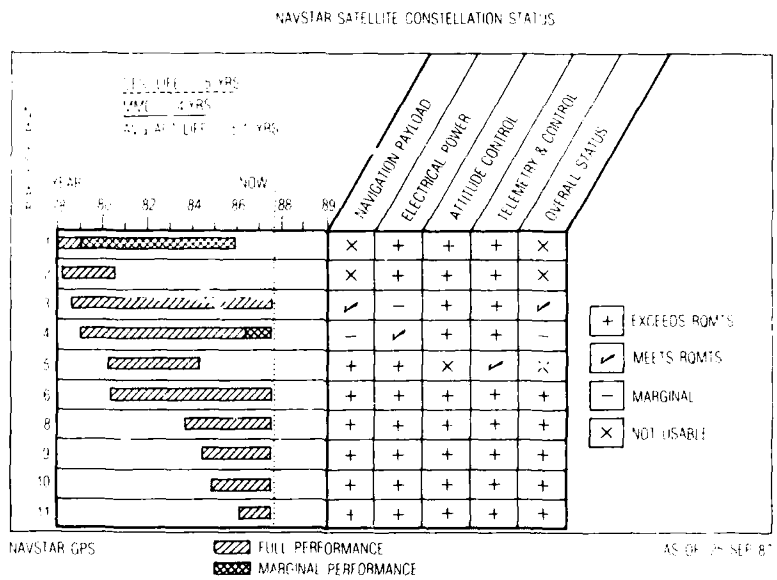


Figure 3



	Navigation Payload	Electrical Power	Attitude Control	Telemetry & Control	Overall Status
1	×	+	+	+	×
2	×	+	+	+	×
3	✓	—	+	+	✓
4	—	✓	+	+	—
5	+	+	×	✓	×
6	+	+	+	+	+
8	+	+	+	+	+
9	+	+	+	+	+
10	+	+	+	+	+
11	+	+	+	+	+

Figure 4

7 SV CONSECUTIVE COVERAGE ON 5/31/87

EPOCH = 1987. 5. 31. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - SVN3 SVN4 SVN6 SVN8 SVN9 SVN10 SVN11

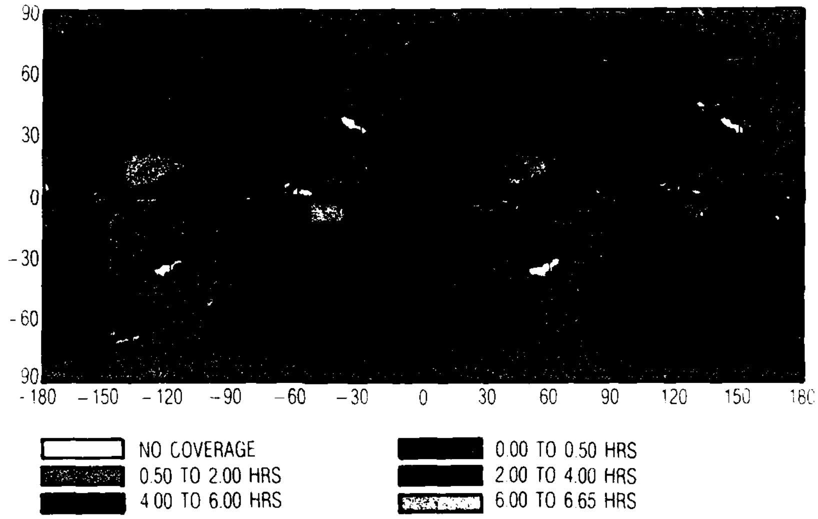


Figure 5

6 SV CONSECUTIVE COVERAGE ON 5/31/87

EPOCH = 1987. 5. 31. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - SVN3 SVN6 SVN8 SVN9 SVN10 SVN11

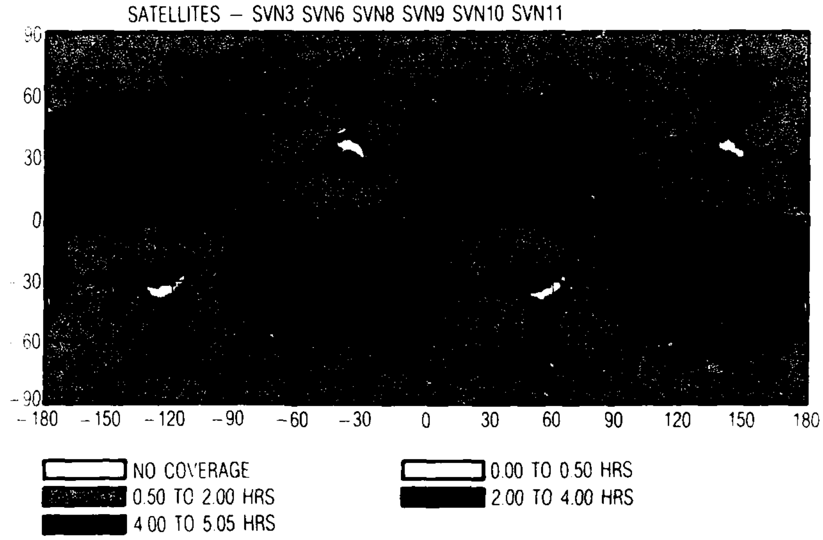


Figure 6

7 SV CONSECUTIVE COVERAGE ON 9/13/87

EPOCH = 1987. 9. 13. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - SVN3 SVN4 SVN6 SVN8 SVN9 SVN10 SVN11

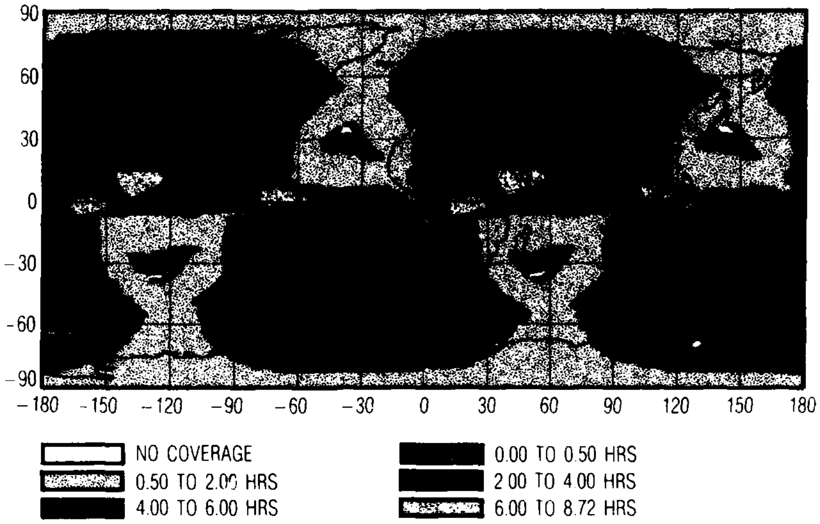


Figure 7

7 SV CONSECUTIVE COVERAGE ON 12/26/87

EPOCH = 1987. 12. 26. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - SVN3 SVN4 SVN6 SVN8 SVN9 SVN10 SVN11

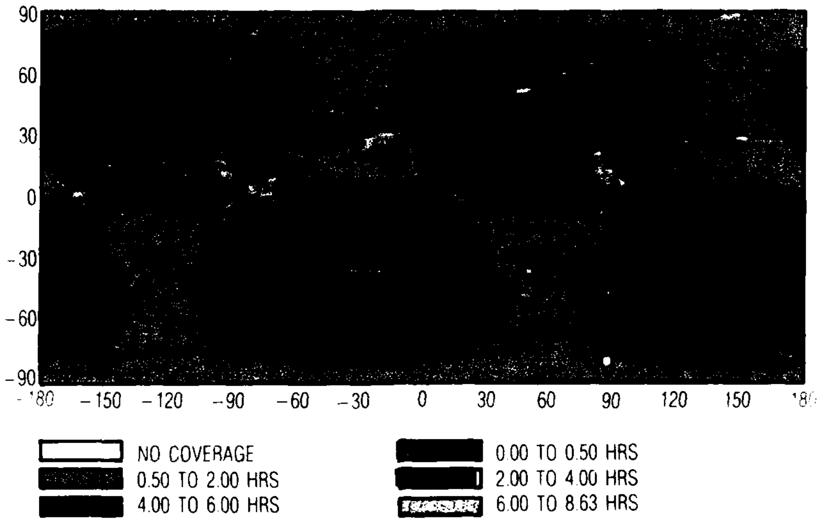


Figure 8

6 SV CONSECUTIVE COVERAGE ON 12/26/87

EPOCH = 1987. 12. 26. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM −90.0 TO 90.0 LONGITUDE FROM −180.0 TO 180.0
SATELLITES — SVN3 SVN6 SVN8 SVN9 SVN10 SVN11

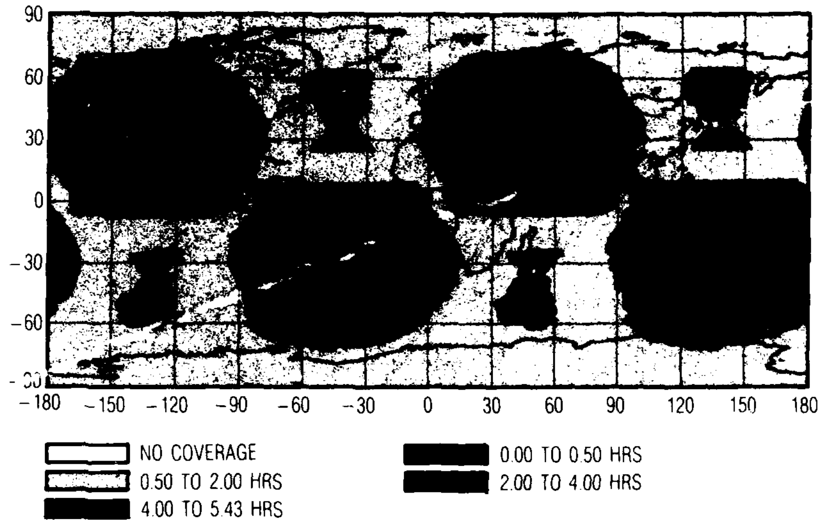


Figure 9

5 SV CONSECUTIVE COVERAGE ON 12/26/87

EPOCH = 1987. 12. 26. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM −90.0 TO 90.0 LONGITUDE FROM −180.0 TO 180.0
SATELLITES — SVN6 SVN8 SVN9 SVN10 SVN11

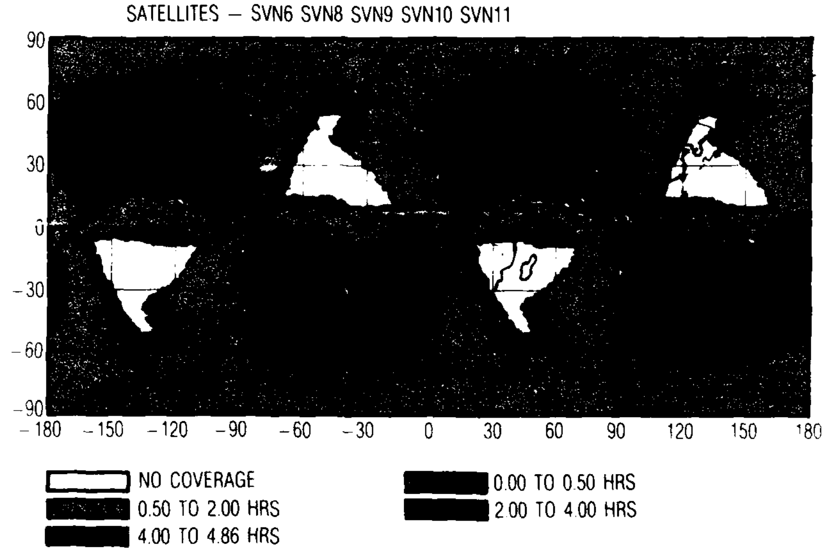


Figure 10

THE NAVSTAR OPERATIONAL CONSTELLATION
18 SATELLITES PLUS 3 ACTIVE SPARES

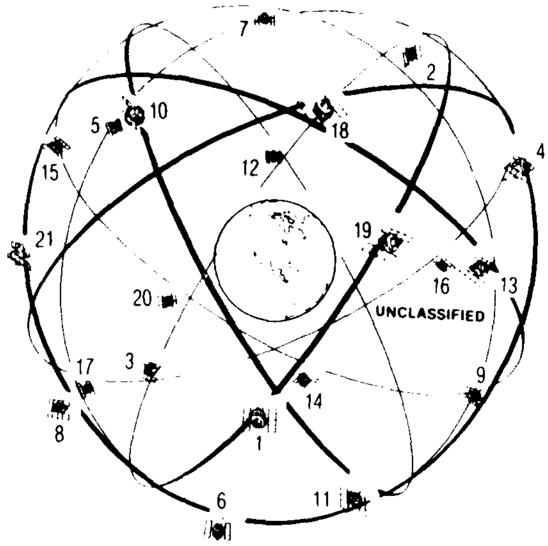
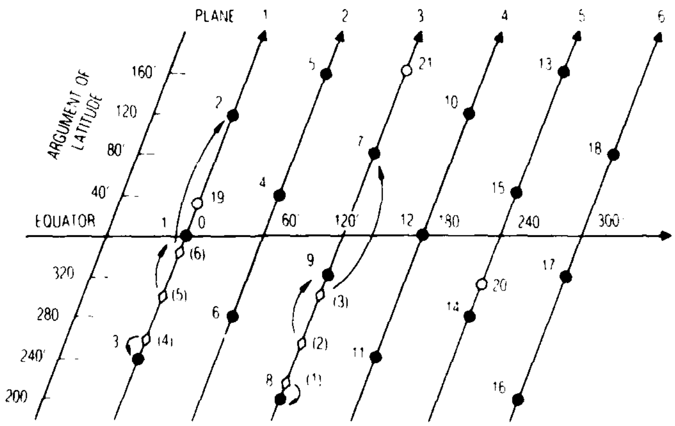


Figure 11

OPERATIONAL CONSTELLATION WITH TEST PHASE LOCATIONS SUPERIMPOSED



● OPERATIONAL SATELLITE AND POSITION NUMBER
○ SPARE
◇ (N) NOMINAL TEST PHASE SATELLITE AND TEST POSITION NUMBER
◇● DENOTES PROPOSED SATELLITE MOVE WHEN REPHASING TEST CONFIGURATION
ALL BLOCK I SATELLITES ARE AT 63° INCLINATION
ALL BLOCK II SATELLITES ARE AT 55° INCLINATION

Figure 12

CANDIDATE BUILDUP
START FROM 5 SATELLITE TEST CONFIGURATION

NUMBER OF SATELLITES	NODE deg	ARG OF LATITUDE deg	ORBIT INCLINATION deg	WORLDWIDE CONSTELLATION VALUE PERCENT
	120	80 (299)*	63	
	120	320 (257)*	63	
	120	200 (214)*	63	
	0	120 (346)*	63	
	0	0 (304)*	63	
6	0	240 (261)*	55	14.7
7	300	80	55	19.7
8	240	160	55	25.7
9	180	120	55	32.4
10**	240	40	55	39.1
11	300	320	55	49.4
12	300	200	55	59.4
13	240	280	55	74.1
14	180	240	55	83.9
15	180	0	55	92.3
16	60	160	55	95.0
17	60	40	55	97.4
18	60	280	55	99.5
SPARES				
19	0	30	55	99.6
20	240	310	55	99.8
21	120	170	55	99.9

( )* Location before rephasing **Rephase test configuration at 10 satellites

Figure 13

COMPOSITE AREAS OF DEGRADATION
18/6/2 AT 55 DEGREE INCL
5-DEGREE MASK ANGLE, PDOP ≤ 6

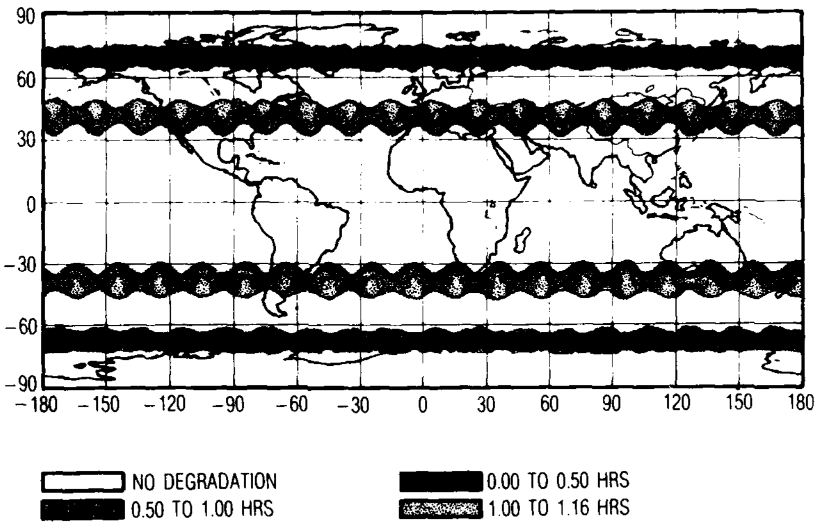


Figure 14

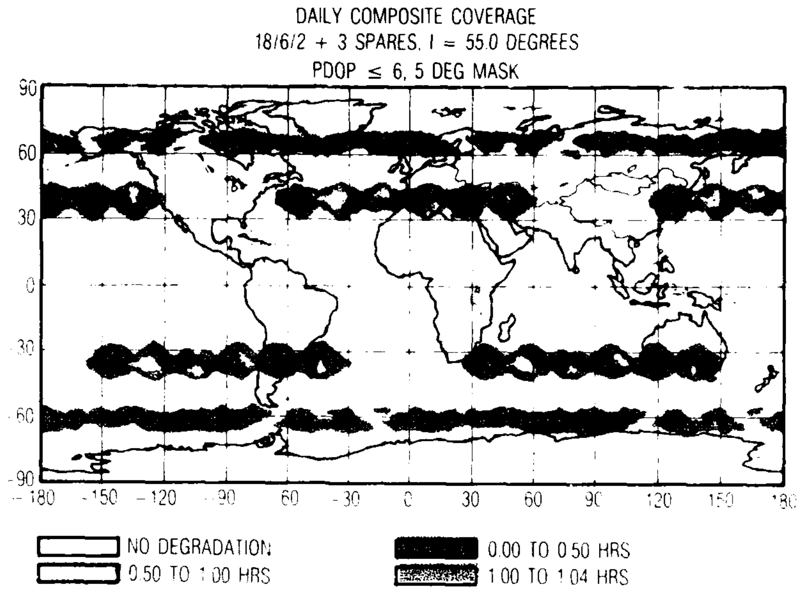


Figure 15

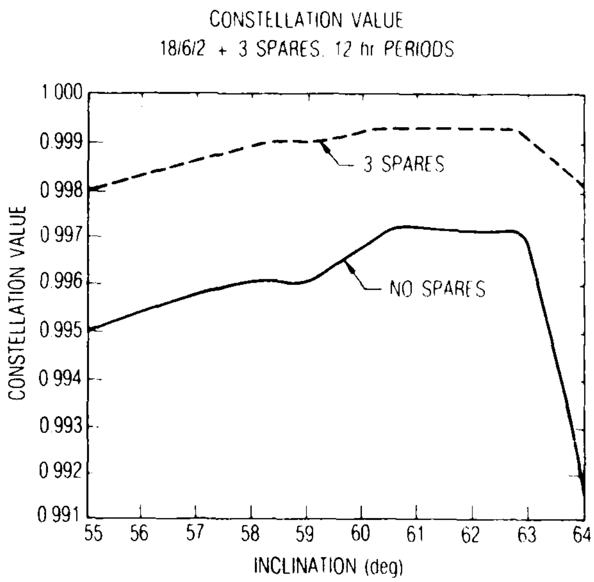


Figure 16

DAILY COMPOSITE COVERAGE
18/6/2 + 3 SPARES, I = 60.5 DEGREES
PDOP ≤ 6, 5 DEG MASK

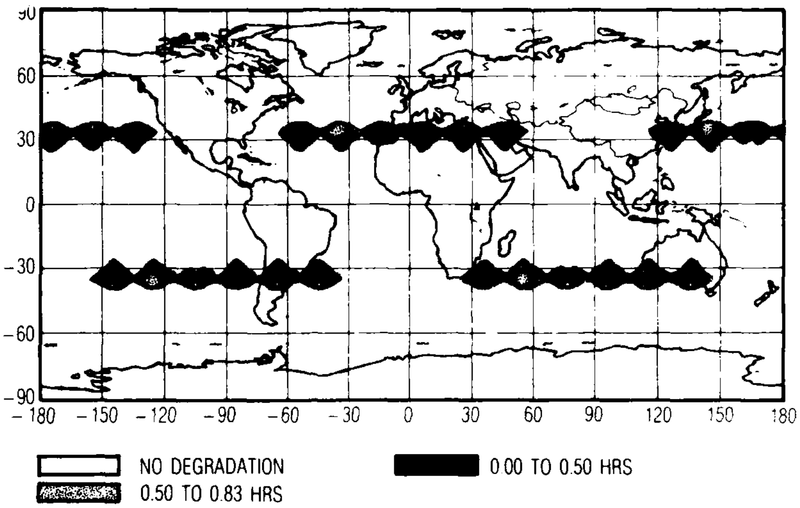


Figure 17

BASELINE 21 SVS (PDOP > 10)

EPOCH = 1985. 7. 1. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - VEH1 VEH2 VEH3 VEH4 VEH5 VEH6 VEH7 VEH8 VEH9 VEH10 VEH11
VEH12 VEH13 VEH14 VEH15 VEH16 VEH17 VEH18 VEH19 VEH20 VEH21

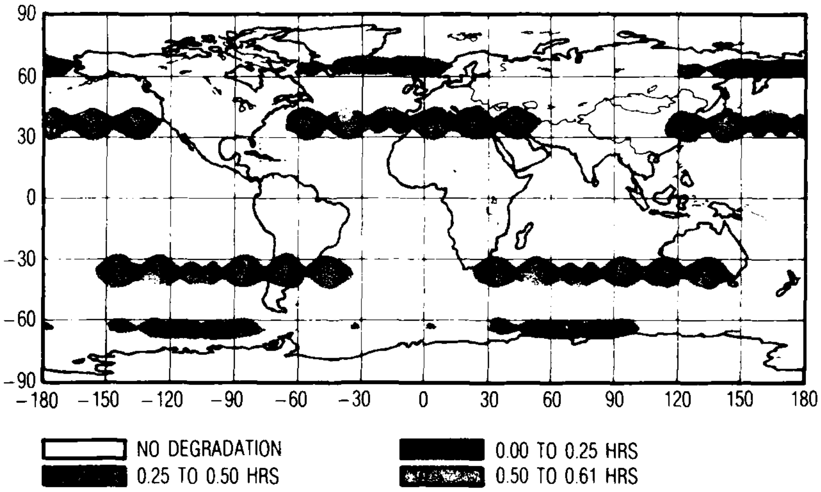


Figure 18

GPS 21-SATELLITE CONSTELLATION (PDOP ≥ 20)

EPOCH = 1985. 7. 1. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - A1 A2 A3 B1 B2 B3 C1 C2 C3 D1 D2 D3
E1 E2 E3 F1 F2 F3 A4 E4 C4

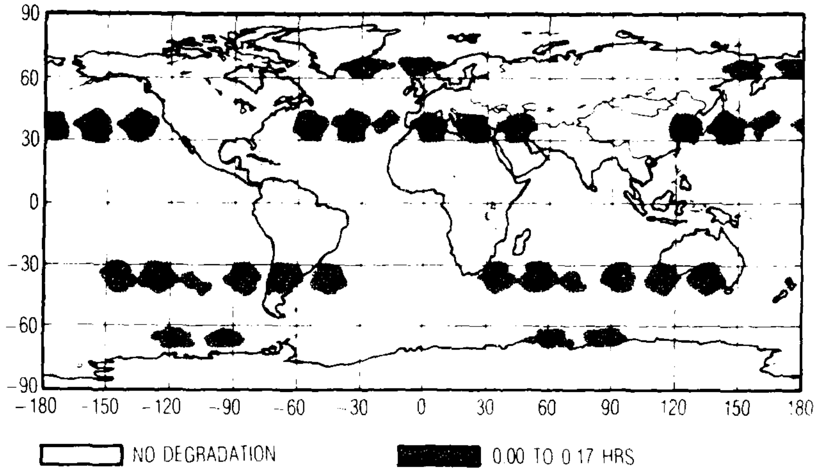


Figure 19

REVISED 21-SATELLITE DEGRADED COVERAGE

EPOCH = 1989. 4. 1. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM -90.0 TO 90.0 LONGITUDE FROM -180.0 TO 180.0
SATELLITES - A1 A2 A3 B1 B2 B3 C1 C2 C3 D1 D2 D3
E1 E2 E3 F1 F2 F3 A4 E4 C4

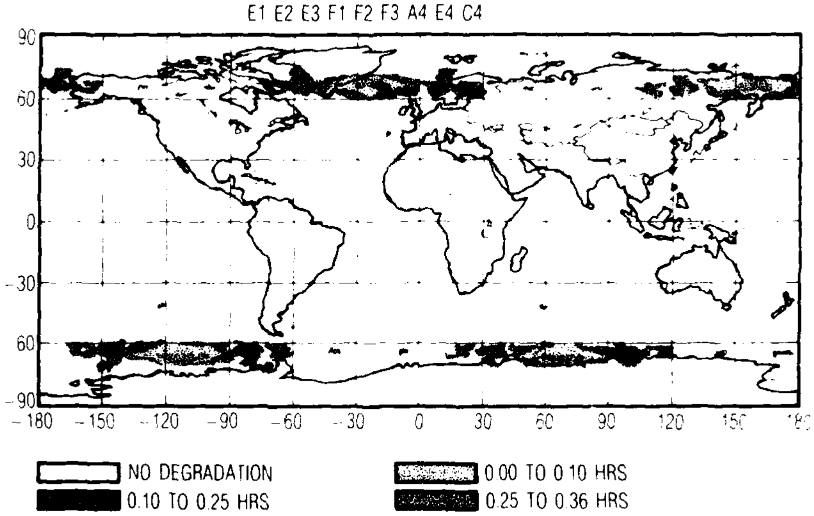


Figure 20

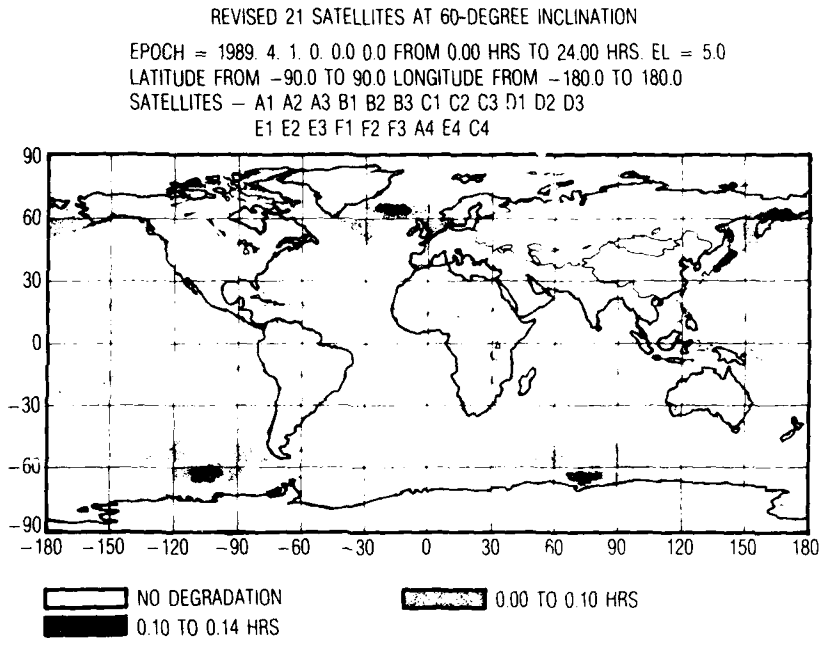

REVISED 21 SATELLITES AT 60-DEGREE INCLINATION
EPOCH = 1989. 4. 1. 0. 0.0 0.0 FROM 0.00 HRS TO 24.00 HRS. EL = 5.0
LATITUDE FROM −90.0 TO 90.0 LONGITUDE FROM −180.0 TO 180.0
SATELLITES − A1 A2 A3 B1 B2 B3 C1 C2 C3 D1 D2 D3
E1 E2 E3 F1 F2 F3 A4 E4 C4
90
60
30
0
−30
−60
−90
−180
−150
−120
−90
−60
−30
0
30
60
90
120
150
180
NO DEGRADATION
0.00 TO 0.10 HRS
0.10 TO 0.14 HRS

Figure 21

THE GLOBAL POSITIONING SYSTEM (GPS) ACCURACY, SYSTEM ERROR BUDGET, SPACE AND CONTROL SEGMENT OVERVIEW

MOHAN ANANDA
The Aerospace Corporation
El Segundo, California, U.S.A.

ABSTRACT

The Navstar Global Positioning System is a space-based navigation system that will provide continuous, all weather, global navigation capability to properly equipped users with high accuracy. This paper discusses the position, velocity, and time accuracy that can be achieved by a user of the GPS. The system accuracy will be examined in light of the Geometric Dilution of Precision parameter and measurement error. The paper also reviews the system error budget for the three subsystem segments; space, control, and user. A discussion of the user range error (URE) and its relationship to three-dimensional position error expressed in spherical error probable (SEP) and horizontal position error expressed in circular error probable (CEP) is presented. The role of the World Geodetic System reference frame in the GPS is discussed. A brief overview of space and control systems is also presented.

1. INTRODUCTION

The Navstar Global Positioning System, which will be operational in early 1990, is the most ambitious global navigation system ever attempted. The GPS is a satellite-based navigation system that will provide extremely accurate three-dimensional position and velocity information to properly equipped users anywhere on or near the earth. It is a Joint Service Program managed by the U. S. Air Force with deputies from the Navy, Army, Marines, Defense Mapping Agency, Department of Transportation, and North Atlantic Treaty Organization. General systems engineering and integration is provided by The Aerospace Corporation.

The history and evolution of the navigation technology program of the Department of Defense has been discussed in detail by Eastson (1978). An overview of the GPS and the system concept have been described by Parkinson (1976) and Parkinson and Gilbert (1983). The capability and specific configuration of the system are continuously evolving.

The current baseline orbital configuration for the operational phase GPS employs 18 satellites in 55-deg inclined, circular, 12-hr orbits to transmit navigational signals. The continuous four-satellite global coverage (Porter, et al., 1984) is provided by placing three equally spaced satellites in each of six orbit planes. In addition to the baseline 18 satellites, there will be three active spare satellites in orbit. The six planes of the operational constellation are 60 deg apart in longitude, and in each plane the three satellites are spaced 120 deg apart. The phasing from plane to plane is 40 deg, so that a satellite in one plane will have another satellite 40 deg ahead (north) of it in the adjacent plane to the east. A companion paper by Ananda on GPS constellation and coverage provides the details of constellation design.

The space segment consists of the previously described satellites. The satellites radiate two spread-spectrum pseudo random noise (PRN) radio signals. The navigation message, which consists of onboard clock and satellite ephemeris information, is modulated onto the PRN sequence. The navigation signals are transmitted at two frequencies L_1 (1575.42 MHz) and L_2 (1227.6 MHz). Both are coherently derived from a highly stable onboard atomic clock.

The control segment consists of a master control station, monitor stations, and ground antennas. The radiometric data from the satellites are tracked by the monitor stations. Accurate ephemeris and clock parameters are estimated by extensive data processing. Subsequently predicted ephemeris and clock information in the form of navigation messages are periodically transmitted (uplinked) to the satellites for later transmission to users. The control segment is also responsible for maintaining the health of the satellites. An overview of both space segment and control segment is given in subsequent sections.

The user segment consists of equipment with antenna, receiver, signal processing, and data processing capabilities. The satellite-transmitted radio signal is first received by the user. The user, knowing the signal PRN code, obtains pseudorange data and demodulates the navigation message. Data from four satellites allow the user state vector (consisting of position, velocity, and time) to be computed.

The primary purpose of this paper is to review the navigation accuracy that can be achieved using the GPS. The concept of user range error (URE) will be defined and its relationship to navigation accuracy will be discussed. The system error budget including error contribution due to the space segment, control segment, and user segment will be analyzed. The concept of differential navigation and how certain systematic errors will be cancelled when used in a differential mode will be examined. Concepts of extended navigation, user autonomy, and autonomous navigation capability will be discussed in reference to the achievable navigation accuracy.

2. USER RANGE ERROR

The user range error (URE) is an error vector along the line of sight between a user of the GPS and a GPS satellite. In principle, the user range error is the projection of all system errors, errors due to the inability to model the space segment parameters that affect the range measurement, errors due to the inability to estimate the spacecraft ephemeris and clock parameters by the operational control system, and errors due to modeling the measurements that can be made by the user equipment, on the line of sight between the user and the GPS satellite. Therefore, every single range measurement will have different user range error. However, in order to establish a system accuracy one has to statistically define a user range error that would be valid for every measurement over the entire globe. In addition, the user navigation accuracy is a function of both user range error and the system geometry. Even though the user range error and the system geometry applicable to a particular user may be different from anybody else, the system capability and performance can only be expressed in a statistical sense that would be valid both spatially and temporally.

The user range error component due to both space segment and control segment can not be improved by a user by designing better user equipment. Careful design and more accurate modeling in the user equipment can reduce the user range error component due to the user equipment. However, the user navigation accuracy will be limited by the inherent user range error due to contributions by the space segment and control segment. The details of deriving the user range error due to both space segment and control segment are given by Bernstein (1983). A brief overview is given here for completeness.

The user range error due to space segment and control segment is primarily due to the inability of the control segment to predict the ephemeris and space vehicle clock. The user range error of concern is caused by ephemeris errors and clock errors associated with the ephemeris and clock messages (navigation message) broadcast by the satellite. The navigation message is modulated on the navigation signal (both P and C/A code) at a 50-bps rate. The navigation user, when estimating the user location, utilizes the navigation message for computing the satellite state and navigation signal transmission time with respect to GPS time. A deterministic user range error has been derived by Bernstein (1983) and is expressed as

$$URE = \epsilon_R + \epsilon_C + 0.24\,(\epsilon_{IT}\,\mathrm{Cos}\,\alpha - \epsilon_{CT}\,\mathrm{Sin}\,\alpha)\,\mathrm{Cos}\,E$$

where ϵ_R is the radial trajectory error, ϵ_C is the satellite clock error projected along the line of sight, and ϵ_{IT} and ϵ_{CT} represents the intrack and crosstrack ephemeris error, respectively. As shown in Figure 1, the angle α is the angle between the direction of the in track ephemeris error and the plane formed by the line of sight and the radial vectors. The angle E is the user elevation angle.

Bernstein has also derived the user range error variance averaged over all users having visibility to the satellite. The expression for the variance is

$$\sigma^2_{URE} = \sigma^2_R + \sigma^2_C - 2\,\rho_{R/C}\,\sigma_R\sigma_C + \frac{(0.24)^2}{3}\,(\sigma^2_{IT} + \sigma^2_{CT})$$

where σ^2_R is the variance of the radial ephemeris error

σ^2_C is the variance of the clock error

σ^2_{IT} is the variance of the intrack ephemeris error

σ^2_{CT} is the variance of the crosslink ephemeris error

ρ_{RC} is the correlation coefficient between the radial ephemeris error and the clock error.

The data in Figure 2 show the orbital geometry identifying the various component of the user range error.

System Error Budget

The GPS Program Management Directive specifically requires the GPS to provide a user navigation accuracy of 16 m (SEP). Therefore, the system specifications are developed to meet the program management directive. The error budget allocated for the various subsystem elements is given in Table 1. The error budget is shown in terms of user range error. The user range error has been defined previously. The error budget assumes two conditions. Column 1 in Table 1 reflects the error budget during the normal operations of the operational control system. When the Operational Control System (OCS) is functioning properly, in order to meet the specifications, the OCS is expected to upload the satellites with new navigation messages consisting of predicted ephemeris and clock parameters three times daily. This corresponds to a prediction error for both satellite trajectory and the space vehicle atomic clock of 8 hr.

The ephemeris prediction error for an 8-hr period is negligible compared to the orbit estimation error. Therefore, the error budget related to the ephemeris is dominated by the orbit knowledge error rather than the modeling error, which causes the orbit prediction error. However, the clock error is totally due to the inability to predict the random variations of the clock frequency. The clock specifications for the Block II space vehicle were developed based on the available clock technology in the late 1970s. The predicted error for the 8-hr period for the space vehicle clock is specified to be about 3.3 m. However, the performance of the Block I cesium clocks are much better than the Block I specifications. Therefore, it is expected that Block II satellite atomic clocks would also outperform the Block II specifications. A more detailed discussion of clock performance is given in subsequent sections.

Column 2 in Table 1 reflects the error budget for the system 14 days after the destruction or inoperability of the control system. The program management directive requires the system to be capable of providing navigation support for a period of 14 days after the control system has been made inoperable--however, with a degraded navigation accuracy. This requirement is the limited autonomous feature of the Block II space vehicle. Since no more navigation upload is feasible due to the elimination of the operational control system, the navigation message is based on predicted ephemeris and clock data and the total error over a 14-day period is specified to be 200 m in user range. Both clock prediction and ephemeris prediction contribute significantly to the total user range error. However, Block I cesium data indicate that the clock contribution to the user range error is considerably smaller than what it is specified to be. As can be seen in Table 1, the user contribution to the total user range error is independent of whether there is an operational control system or not.

The Program Management Directive has been recently updated to have the system provide continuous navigation support with reduced accuracy for 180 days after the failure of the operational control system, instead of 14 days. A more detailed discussion of this is given in subsequent sections.

Table 1. GPS User Range Error Budget

Source of Error and Responsibility	Error Sources	Error Quantities in Meters (1 σ) Normal Operations	14-Day Autonomous Operations
Space	Clock and Navigation Subsystem Stability	3.3	131.00
	L-Band Phase Uncertainty	0.5	0.5
	Predictability of SV Purturbations	1.0	108.0
	Other	0.5	8.0
	Maximum Total Segment URE	3.5	170.0
Control	Ephemeris Prediction and Model Implementation	4.2	105.0
	Other	0.9	10.0
	Maximum Total Segment URE	4.3	105.0
Navigation User	Ionospheric Delay Compensation	2.3	2.3
	Tropospheric Delay Compensation	2.0	2.0
	Receiver Noise and Resolution	1.5	1.5
	Multipath	1.2	1.2
	Other	0.5	0.5
	Maximum Total Segment URE	3.6	3.6
System	Total System URE	6.66	200.0

3. STATISTICAL RELATIONSHIPS OF NAVIGATION ERRORS

Navigation errors are represented in various forms. In military applications, often the errors are represented in CEP for two-dimensional navigation errors and in SEP for three-dimensional navigation errors. In addition, for civilian applications such as for aircraft navigation, a horizontal error representation is often made using 2 d_{rms}. It should be noted however, that 2 d_{rms} does not mean 2D (two dimensional) rms. The accurate definition, as given by Bowditch (1977), is two times d_{rms}. An effort has been made to relate these statistical parameters CEP, SEP, and 2 d_{rms}.

In order to develop the relationships, one has to assume that the errors along the three mutually orthogonal axes are 1) unbiased, 2) normally distributed with equal variance, and 3) uncorrelated. This type of error distribution is generally known as spherical distribution. Thus, if only the horizontal plane is considered, and if the errors along orthogonal axes x and y have the statistical properties defined above, the radial error has a Rayleigh distribution. The probability that the radial error is equal to or less than r is given by

$$P(r) = 1 - e^{-r^2/2\sigma^2}$$

where σ is the standard deviation for both x and y (i.e., $\sigma = \sigma_x = \sigma_y$) (Burington and May).

The circular error probable is obtained from the previous equation with $P(r) = 0.5$.

$$CEP = 1.1774\ \sigma$$

When σ_x and σ_y are not equal, the relationship for CEP is given in Bowditch.

The horizontal radial error can also be expressed in the form of 2 d_{rms}. A detailed discussion of d_{rms} is given in appendix to Bowditch. The basic relationship is

$$2\ d_{rms} = 2\ \left[\sigma_x^2 + \sigma_y^2\right]^{1/2}$$

For equal standard deviation in x and y ($\sigma = \sigma_x = \sigma_y$)

$$2\ d_{rms} = 2.828\ \sigma$$

The probabilities associated with d_{rms} are shown in Table 2 (Bowditch).

Table 2. d_{rms} Error Probabilities

σx	σy	Length of 1 d_{rms}	Probability 1 d_{rms}	2 drms
0.0	1.0	1.000	0.683	0.954
0.1	1.0	1.005	0.682	0.955
0.2	1.0	1.020	0.682	0.957
0.3	1.0	1.042	0.676	0.961
0.4	1.0	1.077	0.671	0.966
0.5	1.0	1.118	0.662	0.969
0.6	1.0	1.166	0.650	0.973
0.7	1.0	1.220	0.641	0.977
0.8	1.0	1.280	0.635	0.980
0.9	1.0	1.345	0.632	0.981
1.0	1.0	1.414	0.632	0.982

The horizontal radial error with 0.95 probability can be obtained from the probability equation:

$$H(0.95) = 2.0789\sigma$$

In a normal distribution, 0.95 probability is generally associated with 2 σ error.

When a three-dimensional error is computed for the spherical error distribution specified above, the vertical error is normally distributed and has a standard deviation of σ.

In Burington and May, the spherical error probability relationships can be found. The SEP error, which is a 50 percent error, is given by

$$SEP = 1.538\sigma$$

When GPS navigation estimation is expressed in terms of a three-dimensional radial error, it is computed by multiplying the position dilution of precision (PDOP) value by the standard deviation of the user range error. In this computation, the navigation error varies directly in proportion to either the standard deviation of the user range error or the PDOP.

4. WORLD GEODETIC SYSTEM

The World Geodetic System (WGS) is a coordinate reference system developed by the Defense Mapping Agency (DMA) of the Department of Defense. The purpose of the World Geodetic System is to provide a single, consistent, and accurate reference for positional, digital, mapping, charting, and gravimetric products produced for the Department of Defense by the Defense Mapping Agency. There have been WGS 60, WGS 66, WGS 72, and, most recently, WGS 84. Each reference system represents the best model of the earth from a geometric, geodetic, and gravitational standpoint using data, techniques, and technology available in that year.

One of the attributes of the GPS is its ability to provide navigation in a globally referenced coordinate system, rather than with respect to local geodetic datums. The GPS makes use of the World Geodetic System as its reference system and, in particular, WGS 84 is the GPS earth-centered, earth-fixed (ECEF) coordinate system. A detailed discussion of WGS 84 is found in the report by the Defense Mapping Agency (1987).

The origin of the WGS 84 coordinate system is the center of mass of the earth; the WGS 84 z-axis is parallel to the direction of the conventional terrestrial pole (CTP) for polar motion as defined by the Bureau International de l'Heure (BIH) on the basis of the coordinates adopted for the BIH stations; the x-axis is the intersection of the WGS 84 reference meridian plane and the plane of the CTP equator; the reference meridian is parallel to the zero meridian defined by the BIH on the basis of the coordinates adopted by the BIH stations; and the y-axis completes a right-handed, earth-fixed orthogonal coordinate system measured in the plane of the CTP equator, 90 deg east of the x-axis. The WGS 84 coordinate system origin and axes also serve as the geometric center and x, y, and z axis of the WGS 84 ellipsoid. Therefore, the WGS 84 coordinate system z-axis is the rotational axis of the WGS 84 ellipsoid.

The WGS 84 ellipsoid has been defined as the geometric equipotential ellipsoid of revolution and the ellipsoid is defined by the semimajor axis (a), the earth gravitational constant (GM), the normalized second-degree zonal gravitational constant (C_{20}), and the angular velocity (ω) of the earth (Table 3). The WGS 84 gravitational model is a spherical harmonic expansion of the gravitational potential (Table 4). Even though the WGS 84 gravitational model consists of both degree and order of 180, for the orbit determination of the GPS satellites, the gravitational model of both degree and order of only 8 is utilized because of the insensitivity of the higher degree and order gravitational coefficients on the GPS satellites orbits.

Table 3. WGS 84 Ellipsoid

Parameters	Notation	Magnitude	Accuracy (1σ)
Semimajor Axis	a	6378137 m	± 2 m
Normalized Second Degree Zonal Harmonic Coefficient of the Gravitational Potential	$\overline{C}_{2,0}$	$-484.16685 \times 10^{-6}$	$\pm 1.30 \times 10^{-9}$
Angular Velocity of the Earth	ω	7292115×10^{-11} rad s^{-1}	$\pm 0.1500 \times 10^{-11}$ rad s^{-1}
The Earth's Gravitational Constant (Mass of Earth's Atmosphere Included)	GM	3986005×10^{8} m^{3} s^{-2}	$\pm 0.6 \times 10^{8}$ m^{3} s^{-2}

Table 4. WGS 84 Earth Gravitational Model

$$V = \frac{GM}{r}\left[1 + \sum_{n=2}^{n_{max}} \sum_{m=0}^{n} \left(\frac{a}{r}\right)^n \bar{P}_{n,m} (\sin \phi') (\bar{C}_{n,m} \cos m\lambda + \bar{S}_{n,m} \sin m\lambda)\right]$$

Parameter	Definition
V	= Gravitational potential function
GM	= Earth's gravitational constant
r	= Radius vector from the earth's center of mass
a	= Semimajor axis of the WGS 84 Ellipsoid
n, m	= Degree and order, respectively
ϕ'	= Geocentric latitude
λ	= Geocentric longitude = geodetic longitude
$\bar{C}_{n,m}, \bar{S}_{n,m}$	= Normalized gravitational coefficients*
$\bar{P}_{n,m} (\sin \phi')$	= Normalized associated Legendre function
	$= \left[\frac{(n-m)!\,(2n+1)k}{(n+m)!}\right]^{1/2} P_{n,m} (\sin \phi')$
$P_{n,m} (\sin \phi')$	= Associated Legendre function
$P_{n,m} (\sin \phi')$	$= (\cos\phi')^m \frac{d^m}{d(\sin \phi')^m} [P_n(\sin \phi')]$
$P_n (\sin \phi')$	= Legendre polynomial
$P_n (\sin \phi')$	$= \frac{1}{2^n n!} \frac{d^n}{d(\sin \phi')^n} (\sin^2 \phi' - 1)^n$

* Note:

$$\begin{matrix}\bar{C}_{n,m} \\ \bar{S}_{n,m}\end{matrix} = \left[\frac{(n+m)!}{(n-m)!\,(2n+1)k}\right]^{1/2} \begin{matrix}C_{n,m} \\ S_{n,m}\end{matrix}$$

where

$C_{n,m}$, $S_{n,m}$ = Conventional gravitational coefficients

$m=0$, $k=1$;

$m\neq0$, $k=2$.

In addition to the WGS 84 ellipsoid, another earth figure that is often used is the WGS 84 geoid. The geoid is defined as that particular equipotential surface of the earth that coincides with mean sea level over the oceans and extends hypothetically beneath all land surfaces. The WGS 84 geoid or mean sea surface reference is applicable to GPS specifically for those user equipments developed for ocean surface vehicles.

5. NAVIGATION ACCURACY

As far as the GPS performance is concerned, the GPS Program Management Directive (PMD) provides the following accuracy requirements.

1. Three-dimensional positioning accuracy requirement is to be 16-m spherical error probable.
2. The accuracy of the GPS provided to the civil sector is to be 100 m (2 d_{rms}) in the horizontal plane.
3. Three-dimensional positioning accuracy requirement after 14 days from the last navigation upload by the control segment is to be 425 m (SEP).

The GPS P-code accuracy is available to the authorized user, whereas the C/A-code accuracy has been further degraded by selective availability (S/A) features which is available to the unauthorized users. The C/A code has been degraded such that the horizontal position error is 100 m (2 d_{rms}). The corresponding three-dimensional position error would be 171.0 m and the vertical error would be 156.0 m with the same statistics. In Table 5 the various representations of GPS accuracy for both C/A-code and P-code are provided. In order to compute these accuracy values, a 21-satellite constellation with 18/6/2 distribution, with 3 active spares with viewing constraint of 5 deg elevation has been assumed. In Table 6 the various dilution of precision values in order to compute the corresponding GPS accuracy are provided. The spherical or circular error probable accuracies are computed from 50 percent values, the 1-σ accuracies are computed by using RMS values, and the 2 d_{rms} accuracies are computed by using 2 d_{rms} DOP values.

Table 5. GPS Accuracy

	C/A - Code			P-Code		
	50%	1 σ	2 d_{rms}	50%	1 σ	2 d_{rms}
Position (3D) (m)	75.7 (SEP)	94.6	171.0	13.5 (SEP)	16.8	30.4
Horizontal (m)	43.0 (CEP)	55.6	100.0	7.7 (CEP)	9.9	17.8
Vertical (m)	49.7	72.0	156.0	8.8	12.8	27.7
Time (ns)		100.0*	294.0		87.2	172.0

*Unsurveyed location without S/A.

Table 6. Dilution of Precision Values

	Horizontal	Vertical	3-Dimensional	Time
50%	1.16	1.34	2.04	1.09
rms (1σ)	1.50	1.94	2.55	1.20
2 d_{rms}	2.70	4.20	4.60	1.96

In addition, a user range error value of 6.6 m is assumed for P-code and 13.9 m is assumed for C/A code without S/A. This is primarily due to inability to model the ionosphere accurately. Since C/A code is available, most of the time, only on L_1 frequency, two-frequency calibration of ionosphere is not possible and, therefore, ionospheric calibration has to be performed by using the model provided by the satellite (Klobuchar, 1976; ICD-GPS-200, 1984). Analysis of ionospheric time delay algorithms for single frequency users has shown that better than 50 percent accuracy in calibrating the ionospheric delay can be achieved by the single frequency model (Feess and Stephens, 1986). It has also been shown that ionospheric model error affects only when three-dimensional position determination is needed and the sensitivity of ionospheric model error is almost negligible to horizontal position determination. The ionospheric model error directly affects the vertical component of the position. Analysis and tracking data processing provides results showing that both C/A-code with single frequency measurements, and P-code with ionospheric calibration utilizing two frequency measure- ments, give similar accuracy in the horizontal position determination. However, if the C/A-code measurements are degraded by S/A, the horizontal position accuracy with C/A -code will be significantly worse than what could be achieved by P-code. In Table 5 the navigation accuracy that is attainable by tracking C/A-code and P-code when S/A has been implemented is provided. When the C/A-code is degraded by S/A, the resulting user range error for C/A-code is 37.1 m; whereas the user range error for P-code is only 6.6 m.

Table 7. GPS Characteristics (Signal-In-Space)**

SYSTEM: Global Positioning System (GPS)

SYSTEM DESCRIPTION: GPS is a space-based radio positioning navigation system that will provide three-dimensional position, velocity and time information to suitably equipped users anywhere on or near the surface of the earth. The space segment will consist of 18 satellites plus 3 operational spares in 12-hour orbits. Each satellite will transmit navigation data and time signals on 1575.4 and 1227.6 MHz.

Accuracy: Predictable	Accuracy: Repeatable	Accuracy: Relative	Availability	Coverage	Reliability	Fix Rate	Fix Dimension	Capacity	Ambiguity Potential
PPS* Horz - 17.8m Vert - 27.7m Time - 90ns	Horz - 17.8m Vert - 27.7m	Horz - 7.6m Vert - 11.7m	*Expected* to approach 100%	Worldwide *continuous*	98% *probability* that an 18-satellite constellation will be available	Essentially *continuous*	3D + Velocity + Time	Unlimited	None
SPS Horz - 100m Vert - 156m Time - 175ns	Horz - 100m Vert - 156m	Horz - 28.4m Vert - 44.5m							

*For US and Allied military, US Government, and selected civil users specifically approved by the US Government.
**Table A-9 taken from Federal Radio Navigation Plan 1986.

The time transfer accuracy is determined by assuming that a C/A-code receiver without S/A implementation can provide 100 ns (1σ) time synchronization accuracy with the universal time coordinate (UTC). In order to achieve 100 ns (1σ) time accuracy with the UTC, the allowable GPS time error with the UTC is 83.1 ns (1σ). The corresponding time transfer accuracy with the P-code would be 87.2 ns (1σ), and the equivalent 2 d_{rms} time transfer accuracy is 172.0 ns. When the C/A code is degraded with S/A, the resulting $2d_{rms}$ time transfer accuracy is 294.0 ns. The data of Table 7 are taken from the Federal Radio Navigation Plan (1986). Identical accuracies, except for the time transfer accuracy are given in Tables 5 and 7. In Table 7 PPS refers to Precise Positioning Service using the P-code for authorized users, and SPS refers to Standard Positioning Service using the C/A-code with S/A implementation to unauthorized users. The time transfer accuracy shown in Table 7 for both PPS and SPS is different from that shown in Table 5. The Table 7 time transfer accuracy values are not consistent with the rest of the accuracy values.

6. NAVIGATION ACCURACY IN A GPS DIFFERENTIAL MODE

It is well known that navigation accuracy can be improved significantly in a GPS differential mode. (Beser and Parkinson, 1982; Kalafus, et al., 1986; Quill, 1986; Denaro, et al., 1987; Knight and Rhodes, 1987). A brief review of differential navigation and achievable accuracy is included here for completeness. In a differential system, a reference receiver tracks the GPS satellites from a known location and computes range biases and transmits them to the navigation users. The navigation user estimates navigation parameters by processing navigation signals from the GPS satellites in conjunction with transmitted data from the reference stations. A pictorial representation of differential GPS operation is shown in Figure 3.

Major error sources such as the ionospheric errors, tropospheric errors, ephemeris errors, and satellite clock errors that are common to both the reference receiver and the navigation user equipment, are cancelled out in a differential GPS mode, and improved navigation accuracy is achieved. The user GPS relative ranging error budget is shown in Table 8 and the resulting error is 2.0 m (1σ) for P-code, and corresponding

Table 8. User GPS Relative Ranging Error Budget

Segment	Error Magnitude One User, meters (1σ)
Space	
Clock Error, etc.	0
Control	
Ephemeris, etc.	0
User	
Ionospheric	0
Tropospheric	0
Receiver	1.5
Multipath	1.2
Other	0.5
	2.0 meters

instantaneous navigation accuracy for a moving platform is 4.2 m (SEP) and 5.5 m (2 d_{rms}) in horizontal plane.

Differential GPS operation is enhanced by utilizing a pseudolite. A differential GPS with pseudolite is shown in Figure 4. The RTCM special committee on differential Navstar/GPS service has defined a pseudolite in the following way. The pseudolite receives the GPS satellite signals and computer pseudorange and range rate corrections as needed to satisfy the requirements of a differential GPS reference station. The pseudolite also transmits the correction information at 50 bps on an L-band frequency--preferably, but not necessarily, on L_1. The transmitted signal is GPS-like, in that a pseudorandom noise code is used to permit local user equipment to obtain an additional pseudorange measurement to the transmitting antenna. The transmitted signal is also diffused to prevent interference with other equipments. The use of pseudolites is often suggested for improving navigation accuracy for aircraft landing at noninstrumented air fields and also for precision landing.

7. EXTENDED NAVIGATION

The GPS Program Management Directive requires the GPS to provide continuous navigation service for a period of 14 days after the ground control segment has become inoperative. However, recently this requirement has been extended to a period of 180 days after the failure of the control system. The operational GPS satellites achieve the current baseline 14-day autonomy by storing the predicted navigation message for a period of 14 days onboard the satellite. This 14 days of navigation message is uploaded by the ground control system three times daily. As discussed earlier, the navigation accuracy degrades gracefully during the 14-day period. The requirement of navigation service for a period of 180 days after the control system has become inoperative has introduced changes to the navigation message upload procedure. This extended navigation message capability will be only available for GPS operational satellites 11 through 28. The first ten satellites (known as Block II GPS satellites) will have only 14-days stored navigation messages, whereas satellites 11 through 28 (known as Block IIA satellites) will have 180-days stored navigation messages.

The navigation messages for the first 14 days are generated by fitting the ephemeris (Cartesian position components of the orbital state vector) by 15 coefficients (ICD-GPS-200) over the fit span of either 4 or 6 hr. The current baseline scheme utilizes a 4-hr fit span during the first day after the upload and a 6-hr fit span for the days 2 through 14. New navigation messages are used every hour during the first day and every 4 hr during the days 2 through 14. In the extended navigation message scheme due to onboard storage limitation, all of the messages beyond 14 days of prediction will be fitted over extended periods of time, thus requiring fewer messages. This scheme would not significantly decrease the navigation accuracy since the prediction error is much larger than that caused by the large fitting span. The navigation accuracy degrades gracefully over the 180-days time interval to about 10.0 km.

8. USER AUTONOMY

User autonomy (or autonomous user system) is a scheme by which a navigation user can achieve full navigation accuracy during the period when the control system has become inoperable (Ananda, et al., 1988). Certain critical users may not be able to meet the mission objectives by utilizing the extended navigation capability of the GPS. The concept of autonomous user system has been developed and tested primarily to satisfy the requirements of such users. An autonomous user has a GPS user equipment set augmented with an atomic clock, an autonomous user algorithm, and a prestored data base that consists of 6 months of ephemeris and clock parameters of each GPS satellite. The autonomous user utilizes tracking measurements (pseudo-range and delta-range), while stationary from a presurveyed location, to estimate differences of the true satellite ephemeris and clock from the prestored reference. The corrected reference can then be utilized for navigation by the autonomous user. Field test and demonstration of the autonomous user system has shown that navigation accuracy comparable to what can be achieved with a fully operational control system can be obtained with the user equipment, augmented with autonomous user system algorithm.

9. AUTONOMOUS NAVIGATION SYSTEM

The GPS program is currently in the process of procuring replenishment satellites for the GPS Block II. The replenishment satellites are known as Block IIR. The primary objectives of the Block IIR satellites are to provide improved navigation accuracy and increased autonomy and survivability. Major emphasis will be placed on reducing satellite dependence on ground support. The objective is to provide full system accuracy for a period of 6 months without any ground support.

Ananda, et al. (1984) describes an autonomous navigation system wherein the GPS satellites would make crosslink ranging measurements to each other and exchange data via a crosslink communication system. Each satellite would use onboard processors to compute satellite ephemeris and clock parameters using the crosslink range measurements. Study results show that such a system could operate for a period of 6 months without ground contact and achieve system accuracies comparable of operating the system with the control segment.

The crosslink ranging system is based on a time division multiple access (TDMA) scheme. Each satellite has a specific time slot during which the satellite transmits a pseudorandom noise ranging code similar to the C/A or P code, which can be received by satellites that are within the viewing geometry. The viewing geometry is determined by the crosslink antenna gain pattern. The Block II antenna design shows that 12 to 14 satellites will be able to make the ranging measurements. The Block II design has 24 time slots and each slot time period is 1.5 sec; a complete ranging frame would take 36 sec. The ranging measurements are then exchanged during the next crosslink frame. By processing the crosslink range measurements, each satellite will be able to update a prestored reference navigation message. The update scheme would be totally transparent to the navigation user and could be able to maintain the navigation accuracy without any degradation for a period of about 6 months.

The GPS program office has currently funded efforts to build brassboards of Block IIR satellite functions that include the autonomous navigation system. It is expected that the Block IIR satellite would be used to replace failed Block II and IIA satellites with the first launch scheduled in 1995.

10. CONTROL SEGMENT

The primary function of the control segment is to track the GPS satellites and provide them with periodic updates correcting their ephemeris and clock parameters. In addition, the control segment is also responsible for monitoring and maintaining the health and welfare of the satellites. The control segment has a master control station at Colorado Springs at the Consolidated Satellite Operational Center (CSOC), which is responsible for processing all the downlink data including tracking and telemetry data and performing estimation of navigation updates, which will be periodically uploaded to the satellites. The GPS tracking data are collected by the monitor stations distributed around the world. The monitor stations are located in Hawaii, Colorado Springs, Ascension Island, Diego Garcia, and Kwajalein. The commanding and upload functions are performed by the ground antenna and these ground antennas are located with the monitor stations at Ascension Island, Diego Garcia, and Kwajalein. The monitor stations are globally dispersed in longitude such that the maximum separation is less than 90 deg.

The control segment major operations are shown in Figure 6. The various GPS signal links are shown in Figure 7. The control segment will be operated by the Air Force Space Command. Currently the systems command is responsible for testing and validating the operational segment. Analyses by Bowen, et al. (1985) and Feess, et al. (1987) have shown that control system accuracy is well beyond the specified requirements. This is also due to better performance of the satellite atomic clocks.

The data in Figure 8 show the Allan variance plot of the satellite atomic clocks for the Block I satellites. The specifications for the Block I clocks are $\sigma\left(\frac{\Delta F}{F}\right) = 10^{-12}$ at an integration time of 10^5 sec. Ground testing of 24 satellite clocks has shown that the clocks would perform two to five times better than the specification. On-orbit performances for Navstar 3, 4, 5, and 6 are shown in Figure 9. Navstar 3, 4, and 5 data are based on rubidium clocks, and Navstar 6 data are based on a cesium clock. Data clearly show that on-orbit performance is significantly better than the specification. The Block II cesium clock has a specification of $\sigma\left(\frac{\Delta F}{F}\right) = 2 \times 10^{-13}$ for an integration time of 10^5 secs. On-orbit data indicate that a cesium stability of 1×10^{-13} to 5×10^{-14} can be achieved. The atomic clock stability specifications for Block IIR satellites has been increased to 1×10^{-13} for an integration time of 10^5 secs.

Figure 10 shows the prediction error expressed in user range error due to the clock over a 14-day period for Navstar 3, 4, 5, and 6. The results clearly show that the cesium clock is performing better by a factor of 2 to 3 over the rubidium clock.

11. SPACE SEGMENT

The Block II GPS space vehicle has an on-orbit weight of about 1862 lb and it is designed to have 700 watts of power at the end of design life. The satellite is designed with a 7.5-year life including 10 years of consumables. The Block II satellite carries two cesium and two rubidium clocks; at any given time only one clock would be operating. The clocks will be sequentially turned on after the failure of the operating clock. Figure 11 shows the various subsystem elements of the satellites.

Originally the GPS satellites were planned to be launched only with the space shuttle. However, after the Challenger disaster it was decided to use expendable launch vehicles in addition to the space shuttle to place the GPS satellites in orbit. Except for a few GPS satellites, all other GPS satellites will be launched using Delta II launch vehicles. The satellite block diagram is provided in Figure 12. Major subsystems are orbit injection subsystem, attitude and velocity control subsystem, reaction control subsystem, thermal control subsystem, electrical power subsystem, tracking telemetry and command subsystem, integrated transfer subsystem, navigation payload subsystem, and L-band subsystem.

The first ten Block II satellites will continue to function without any ground contact for a period of about 45 days. Satellites 11 through 28, known as Block IIA satellites, are designed to function without any ground contact for a period of 180 days. However, if any subsystem element fails, redundant systems can only be switched on by ground command. There exists no autonomous redundancy management system in the Block II satellite design. However, analysis has shown that, over a period of 180 days, only a maximum of two satellites out of 21 satellites would be subject to any significant anomaly that would take the satellite out of operation. The Block IIR design will have some other autonomous features that will further reduce dependence on the ground control system.

12. SUMMARY

A brief review of user range error and system error budget has been accomplished. A discussion of statistical relationships of navigation errors has been given. The role of the World Geodetic System in the GPS has been briefly examined. A detailed analysis of navigation accuracy achievable using the GPS has been provided. A review of achievable navigation accuracy in a GPS differential mode has also been given. A discussion of extended navigation, user autonomy and autonomous navigation system has been provided. Finally, an overview of the control segment and the space segment was also achieved.

References

1. Easton, R.L., "The Navigation Technology Program", J. Institute of Navigation, Vol. 25, No. 2, 1978.

2. Parkinson, B. W., "Navstar Global Positioning System (GPS)", Proc. National Telecommunication Conference, (November 1976).

3. Parkinson, B. W., and Gilbert, S. W., "Navstar: Global Positioning System - Ten Years Later", Proc. IEEE Conference, Vol. 71, No. 10, October 1983.

4. Porter, J., Kruh, P., and Sprosen, B., "GPS Navstar System Overview", paper presented at Royal Institute of Navigation Meeting, London, May 1984.

5. Bernstein, H., Calculations of User Range Error (URE), Variance from a Global Positioning System (GPS) Satellite, Aerospace Report No. TOR-0083 (3476-02)-1, 1983.

6. Bowditch, N., American Practical Navigation, Vol. I., Defense Mapping Agency Hydrographic Center, 1977.

7. Burington and May, Handbook of Probability and Statistics with Tables, Second Edition, McGraw-Hill Book 6, New York.

8. Department of Defense World Geodetic Systems 1984, Defense Mapping Agency, TR 8350.2, 1987.

9. Klobuchar, J. A., ('Ionospheric Time Delay Corrections for Advanced Satellite Ranging Systems'), NATO, AGARD Conference Proceeding No. 209, 1976.

10. Navstar GPS Space Segment/Navigation User Interfaces, ICD-200, 26 September 1984.

11. Feess, B., and Stephens, S., "Evaluation of GPS Ionospheric Time Delay Algorithm for Single Frequency Users", Proceedings of IEEE PLANS, Las Vegas, Nevada, 1986.

12. 1986 Federal Radio Navigation Plan, Department of Defense and Department of Transportation, DOD-4650.4, DOT-TSC-RSPA-87-3.

13. Beser, J., and Parkinson, B. W., "The Application of NAVSTAR Differential GPS in the Civilian Community", Journal of ION, Vol. 29, No. 2, Summary 1982.

14. Kalafus, R. M., et al., "Special Committee 104 Recommendations for Differential GPS Service", Journal of the ION, Vol. 33, No. 1, Spring 1986.

15. Quill, J., "U.S. Coast Guard Differential GPS System Development", Radio Technical Commission for Maritime Service Assembly Meeting, Boston, Massachusetts, April/May 1986.

16. Denaro, R. P., et al., "Differential GPS Reference Stations Design and Development", paper presented at the ION National Technical Meeting, Anaheim, CA., January 1987.

17. Knight, J. E., and Rhodes, K. W., "Differential GPS Static and Dynamic Test Results", Proceedings of the ION Satellite Division First Technical Meeting, Colorado Springs, 1987.

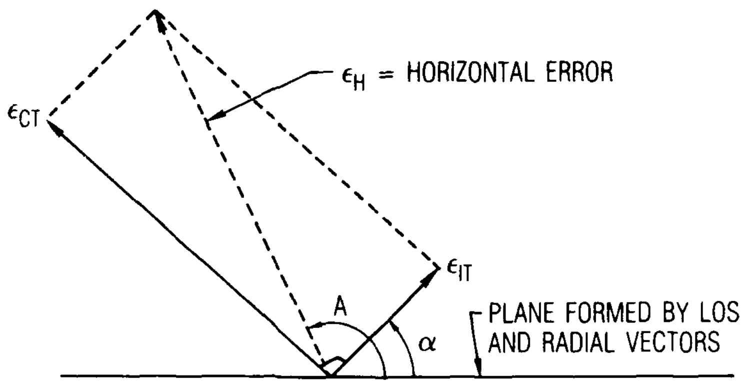
18. Ananda, M., et al., "The Global Positioning System (GPS) Autonomous User System", paper presented at ION National Technical Meeting, Santa Barbara, 1988.

19. Ananda, M., et al., "Autonomous Navigation of the Global Positioning System Satellite", paper presented at AIAA Guidance and Control Conference, Seattle, Washington, August 20-22, 1984.

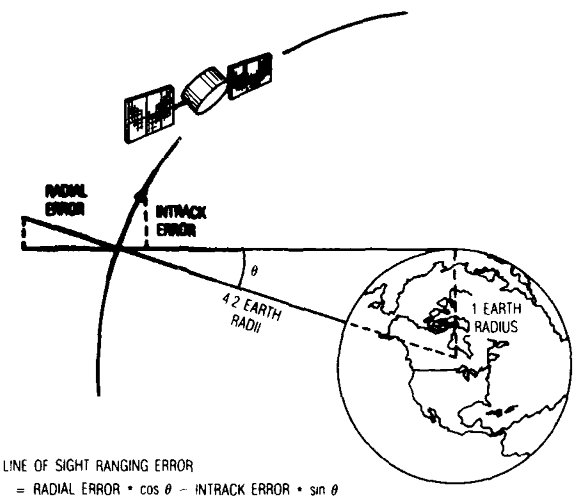
20. Bowen, R., et al., "Global Positioning Systems Operational Control Systems Accuracies", Navigation:, Journal of the Institute of Navigation, Vol. 32, No. 2, 1985.

21. Feess, B., et al., "GPS Satellite-to-User Range Accuracies:, A calibration experiment: 20 October to 17 November 1986", Proceedings of the ION National Technical Meeting, Anaheim, California, 1987.

22. Ananda, M., "GPS Constellation and Coverage", AGARD Presentation of Lectures on 'NAVSTAR GPS', 1988.

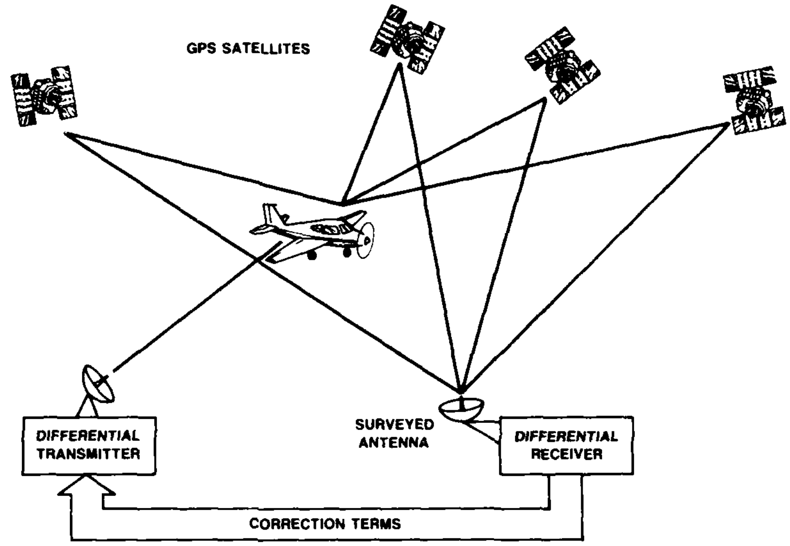


1. *Ephemeris Error Geometry*

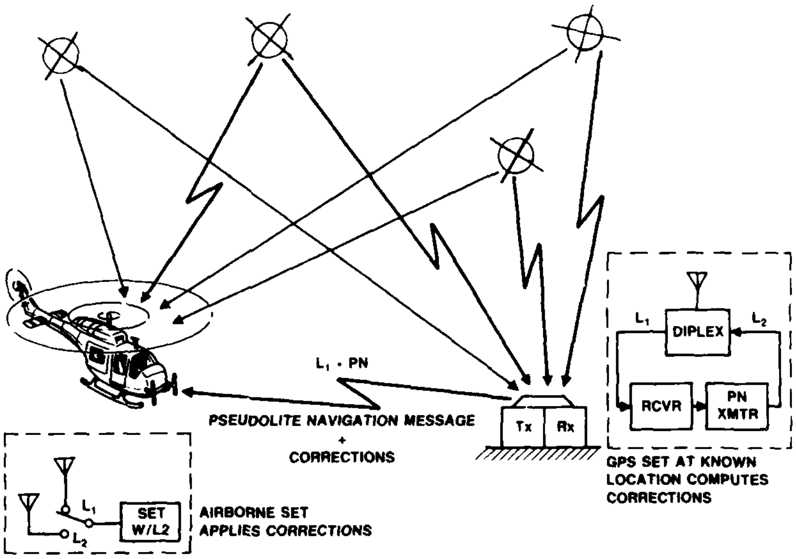


$$\sigma_{URE} = \left[\sigma_R^2 + \frac{0.24^2}{3}\left(\sigma_{IT}^2 + \sigma_{CT}^2\right) + \sigma_C^2 + 2\rho_{R/C}\sigma_R\sigma_C\right]^{1/2}$$

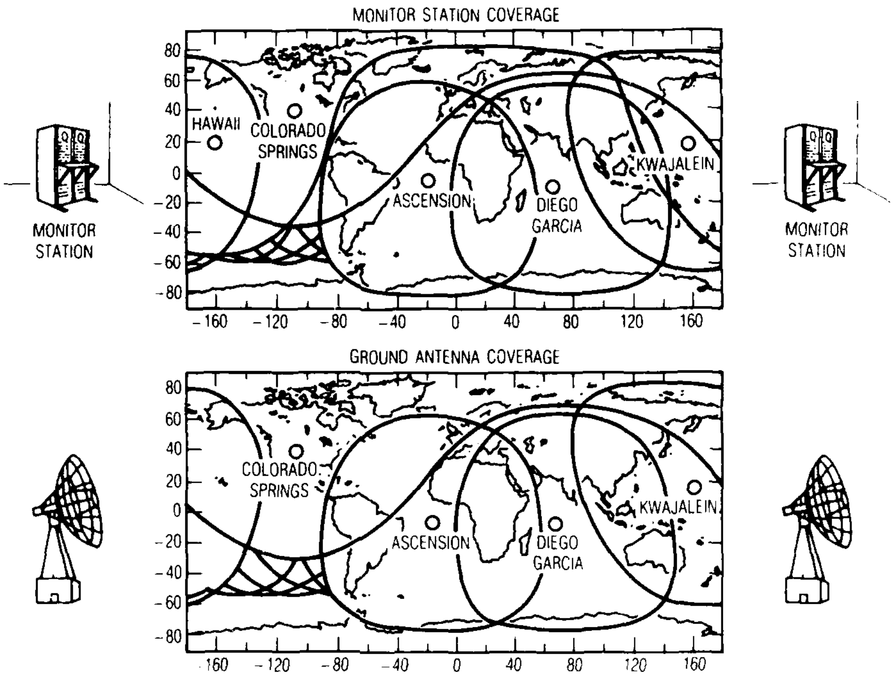
2. Statistical Definition of User Range Error



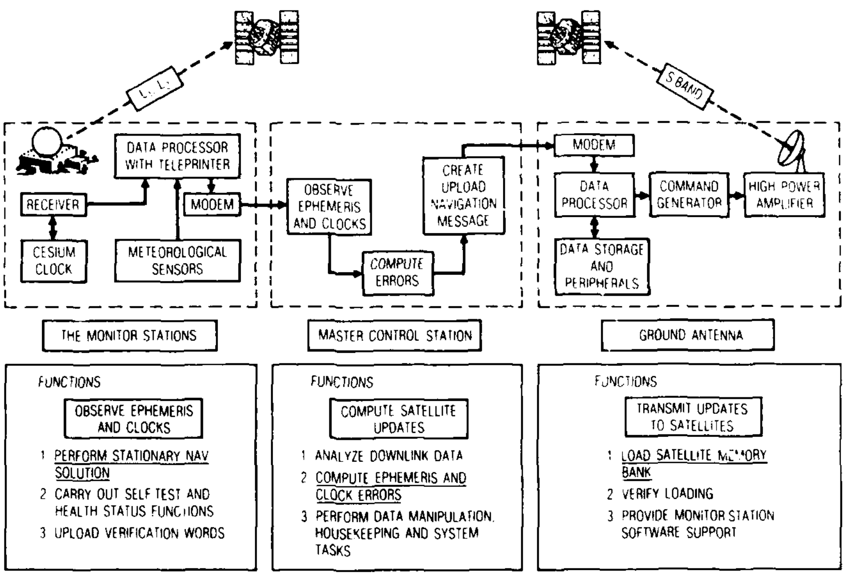
3. Differential GPS Operation



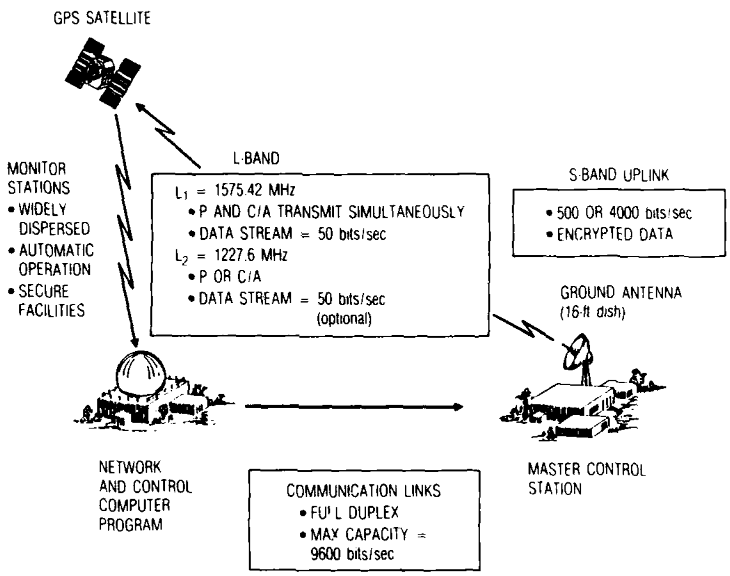
4. Differential GPS with Pseudolite



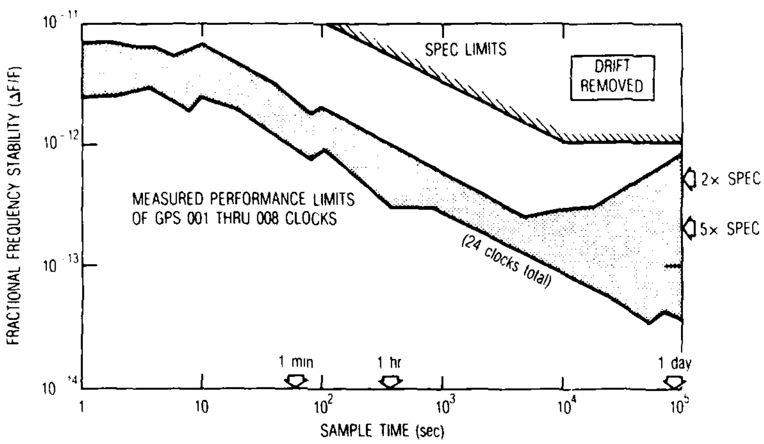
5. Ground Segment Coverage



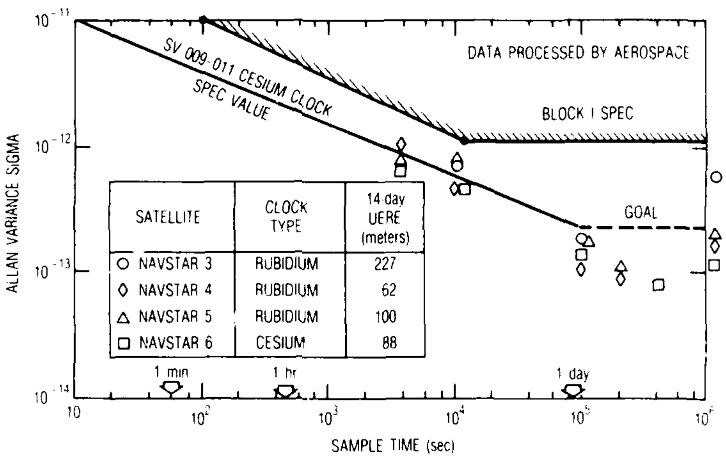
6. Control Segment Operations



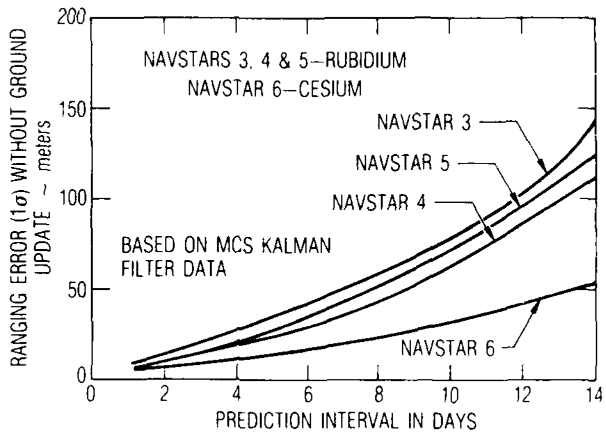
7. GPS Signal Linkages



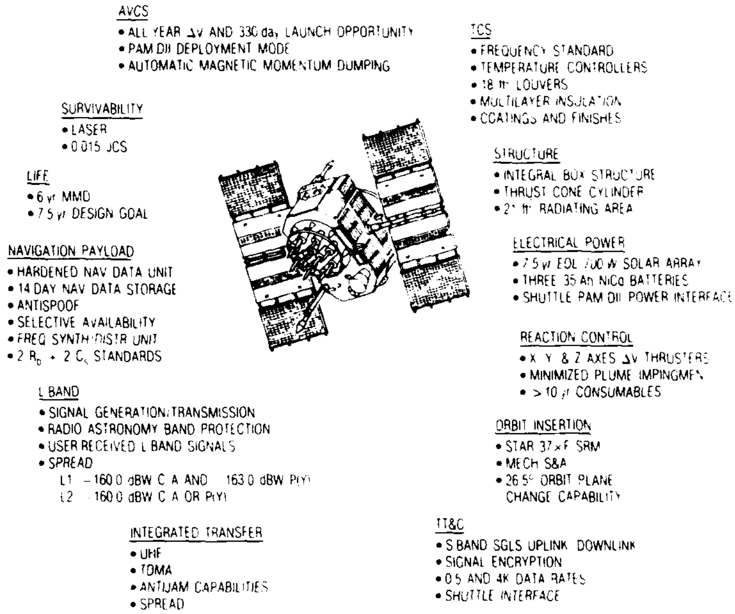
8. Frequency Standard Stability Performance



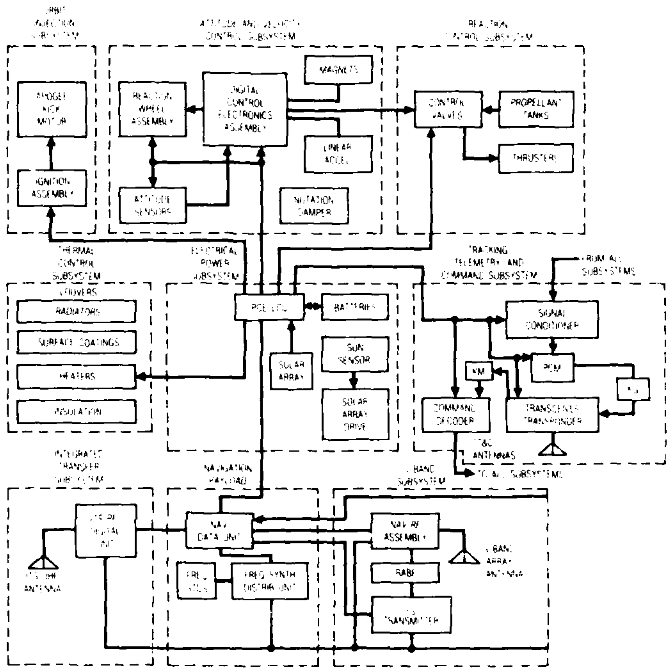
9. Frequency Standard On-Orbit Performance



10. Clock Prediction Error



11. Block II Subsystem Features



12. Satellite Block Diagram

GPS SIGNAL STRUCTURE

by

Lt. Cdr Per W.Nieuwejaar
Naval Materiel Command K3
POB 3
5078 Haakonsvern, Norway

SUMMARY

The characteristics of the GPS signal structure, frequencies, codes, and navigation-message are described. The content of the navigation-message and the user algorithms associated with the navigation-message parameters are described in some detail.

1. INTRODUCTION

The GPS satellite signals, signal generation, and navigation message information are described herein.

2. GPS FREQUENCIES

The GPS satellites transmit on two frequencies, L_1 and L_2. The L_1 frequency, as the primary frequency, is 1575.42 MHz; L_2, the secondary frequency, is 1227.6 MHz. The L_2 is a backup frequency in case L_1 is lost or the receiver, all cases, is being jammed on L_1. It also enables the receiver to perform a dual frequency compensation for signal delay due to ionospheric refraction. In addition, the satellites operate S-band links for communication between the control segment installations and the satellites.

3. CODES

The two codes used in the GPS system are called C/A code and P code. Each GPS satellite is assigned unique codes such that the receiver can distinctively discriminate among the satellite signals it hears and select those signals it wants to use for navigation, even though they all operate on the same frequencies. The codes have the characteristics discussed below.

3.1 P Code/Y Code

The P code is a 288-day-long code sequence, and each of the GPS satellites is assigned a unique 1-week segment of this code. The P-code bit rate is 10.23 MHz, and each satellite will have a 7-day long portion that restarts every Saturday/Sunday at midnight. The P code will normally be transmitted on both L_1 and L_2 and will be for users authorized by the U.S. Department of Defense (DoD) only. The code is protected against spoofing (i.e., the deliberate transmission of incorrect GPS information) by encryption of the P code. The encrypted P code is called Y code.

3.2 C/A Code

The C/A code consists of a 1023-bit code with a clock rate of 1.023 MHz; hence it takes only 1 msec to run through the whole code. A different C/A code is assigned to each GPS satellite and is chosen from a set of codes called Gold codes. The C/A code will normally be transmitted on L_1 only, but it can also be transmitted on the L_2 rather than the P code. The C/A code is available to any user of GPS, and is used by P-code users to assist the receiver in reducing the time to acquire the longer P code.

4. THE NAVIGATION MESSAGE

4.1 Introduction

The Navigation message is superimposed on both the P code and the C/A code with a data rate of 50 bits/sec. The message format is a 1500-bit frame made up of five subframes, each subframe being 300 bits long. Subframes 4 and 5 will be subcommunicated 25 times each, so that a complete data message requires the transmission of 25 full frames.

Each subframe consists of 10 words, each 30-bits long (Figure 1). It will, therefore, take 30 sec to receive one data page and 12-1/2 min to receive all 25 data pages. Subframes 1, 2, and 3 have identical data content on all 25 pages. This allows the receiver to obtain critical NAV-data in 30 sec. Each subframe or page of a subframe in both contains a telemetry (TLM) word and a handover word (HOW), both generated by the satellite, and will start with the TLM/HOW pair. The frame will also contain eight data words generated by the control segment. Each word contains parity. The satellite calculates parity for the TLM and HOW and the control segment calculates parity for all other words in each frame.

Each satellite will have sufficient memory capacity for storing 14 days of uploaded navigation data. The data broadcast in the NAV-msg are normally valid for a 4-hr period and are "updated" every hour to reflect more up-to-date clock and ephemeris information. The control segment uploads the satellites on average three times per day.

4.2 Navigation Message Content

The navigation message, basically, contains four sets of information: time and satellite clock information, correction data to compensate for signal delay, satellite orbit information, and satellite health/status. A detailed description of the navigation message and the standard algorithms used in conjunction with the navigation message parameters are described in STANAG 4294, Navstar Global Positioning System (GPS)-System Characteristics.

4.2.1 Telemetry Word (TLM)

The TLM, which is the first word in every subframe, contains information necessary for the control segment and will therefore not be read by a standard GPS receiver.

4.2.2 Handover Word (HOW)

The HOW is the second word in each subframe. It contains time information that "time tags" the transmission time of the following subframe. The HOW also allows the GPS receiver to transition from C/A-code tracking to P-code tracking, because it tells the receiver what part of the P code is used to generate the following subframe.

4.2.3 Subframe 1

Subframe 1 contains the following information:

- Week Number. Represents the week number of the current GPS time week.
- Code on L_2. Indicates if P code or C/A code is transmitted on L_2.
- Satellite Accuracy. Gives the predicted user range accuracy (URA) of the satellite, using a two-frequency ionospheric correction. The URA is based on tracking data collected and processed by the control segment during the last couple of hours of satellite tracking before satellite upload. The C/A-code users must take into account the additional error due to use of an ionospheric correction model instead of a two-frequency measurement.
- Satellite Health. Indicates the health status of the transmitting satellite. Additional health data are given in Subframe 4 and 5.
- Issue of Data Clock (IODC). Indicates the issue number of the transmitted data set and gives a convenient means of detecting any change in correction data.
- Data Flag for L_2 P-Code. Indicates whether navigation-message is modulated on the P code on L_2.
- Estimated Group Delay Differential. Correction terms to be used in a given algorithm to determine the satellite PRN code phase offset referenced to the phase center of the satellite antennas with respect to the GPS time at the time of data transmission.
- Satellite Clock Correction. Contains the parameters to be used in a given algorithm used to calculate the apparent satellite clock correction.

4.2.4 Subframe 2 and 3

Subframes 2 and 3 contain ephemeris information for the transmitting satellite. The ephemeris parameters are an extension to Keplerian orbital parameters that describe the orbit during the interval of time in which the parameters are transmitted.

Subframes 2 and 3 include information such as:

- Ephemeris Data. Contains all the necessary data to compute the transmitting satellite orbits using a given algorithm.
- Issue of Data Ephemeris (IODE). Provides the user with a convenient means for detecting any change in the ephemeris representation parameters.
- User Range Error (URE). A figure of merit used to measure the quality of the user range data provided to the user by the satellite.

4.2.4.1 Reference System

All the GPS satellite position information is defined in the World Geodetic System 1984 (WGS 84) reference coordinates.

4.2.5 Subframes 4 and 5

These subframes have 25 pages each and the data contained on each page are as follows:

For Subframe 4:

1. Pages 1, 6, 11, 16, and 21: Reserved.

2. Pages 2, 3, 4, 5, 7, 8, 9 and 10: Almanac data for satellites 25 through 32, respectively.

3. Pages 12, 19, 20, 22, 23 and 24: Reserved.

4. Pages 13, 14 and 15: Spare.

5. Page 17: Special messages.

6. Page 18: Ionospheric and UTC data

7. Page 25: A-S flags/satellites configurations for 32 satellites plus satellite health for satellites 25 through 32.

For Subframe 5:

1. Pages 1 through 24: Almanac data for satellites 1 through 24.

2. Page 25: Satellite health data for satellites 1 through 24.

4.2.5.1 Almanac Data

Almanac data is a reduced-precision subset of the clock and ephemeris parameters. The almanac data will be updated by the control segment at least once every 6 days, but they are usually valid for a very long period of time (weeks or months) if the satellites experience no anomalies.

4.2.5.2 Ionospheric Data

The ionospheric data allows an L_1 or L_2 only user to compensate for ionospheric delay using a given mode.

4.2.5.3 Universal Time Coordinated Data (UTC)

The UTC data, used in a given algorithm, enables the user to convert his GPS-given reference time to UTC time.

4.2.5.4 Special Messages

Special messages will be broadcast at the discretion of Space Command and can contain up to 22 eight-bit ASCII characters.

4.2.5.5 Anti-Spoofing (A-S) Data

Indicates the A-S status (A-S on or off) and the configuration code of each satellite (Block I or Block II).

4.2.5.6 Satellite Health

Subframes 4 and 5 contain satellite health data for each satellite in the constellation. The health data in Subframes 4 and 5 may vary from that in Subframe 1, because the data may be updated at different times.

5. SATELLITE SIGNAL GENERATION

Before the P and C/A codes are modulated for transmission on L_1, they are both combined with navigation-message data by Modulo-2 addition. Therefore, since the P and C/A codes are chosen not to interfere with each other (i.e., minimum correlation among the codes), they can be modulated onto the same carrier frequencies in the satellites. It is known that transmitters are most efficient with constant amplitude signals. To produce the effect, the P- and C/A-code carriers, although derived from the same source, are phase shifted 90 deg apart, modulated by the P and C/A code, and then combined. This addition, called phase quadrature, produces a composite continuous wave signal. On L_1, the navigation-message is superimposed on both the P-and the C/A-code carrier and the navigation-message is also superimposed on the P-code signal on L_2 (Figure 2).

5.1 Signal Modulation

The GPS satellites use a type of signal modulation called pseudo-random noise (PRN) bi-phase shift keying (BPSK) of the carrier. The BPSK technique reverses the carrier phase when the digital PRN code transitions from 0 to 1 or from 1 to 0 (Figure 2). The very long sequences of ones and zeros (which constitute the C/A and P-codes) are called PRN codes since, to a casual observer, the ones and zeros appear to occur in a random fashion. The resulting frequency spectrum for the carrier, due to the BPSK, equals 20 MHz for the P code and 2 MHz for the C/A code (Figure 3).

The carrier frequency is suppressed. In actuality, the C/A and the P codes generated are precisely predictable relative to the start time of the code sequence. The user can therefore replicate the same code as the satellite. The amount the user must offset his code generator to match the incoming code from the satellite is directly proportional to the range between the GPS receiver antenna and the satellite. Phase shifting of the carrier results in a spreading of carrier power between ±10.23 MHz of center frequency due to the P code BPSK and ±1.023 MHz due to the C/A-code BPSK. The resulting waveform (Figure 3) is P-code and C/A-code modulation, or chipping rates. When the spread spectrum signal is received at the GPS receiver, the signal power is spread out over such a large bandwidth that the signal is below the thermal noise level. When the satellite signal is multiplied with the GPS-receiver-generated P codes and C/A codes, the satellite signal will be collapsed into the original carrier frequency band. Signal power is then again concentrated into a very narrow frequency band and is well above the thermal noise level (Figure 4).

5.2 RF Signal Levels

The satellites will transmit the signals with the following power levels:

	PRN Signal	
Frequency	C/A (dBW)	P (dBW)
L_1	+ 26.8	+ 23.8
L_2	-	+ 19.7

The received signal values for a receiver close to the earth will be:

	PRN Signal	
Frequency	C/A (dBW)	P (dBW)
L_1	- 160	- 163
L_2	-	- 166

Space loss is therefore on the order of -186 dBW for the satellite signals. The transmission beam for each downlink is shaped to minimize transmission loss and to provide uniform signal strength over the visible portion of the earth.

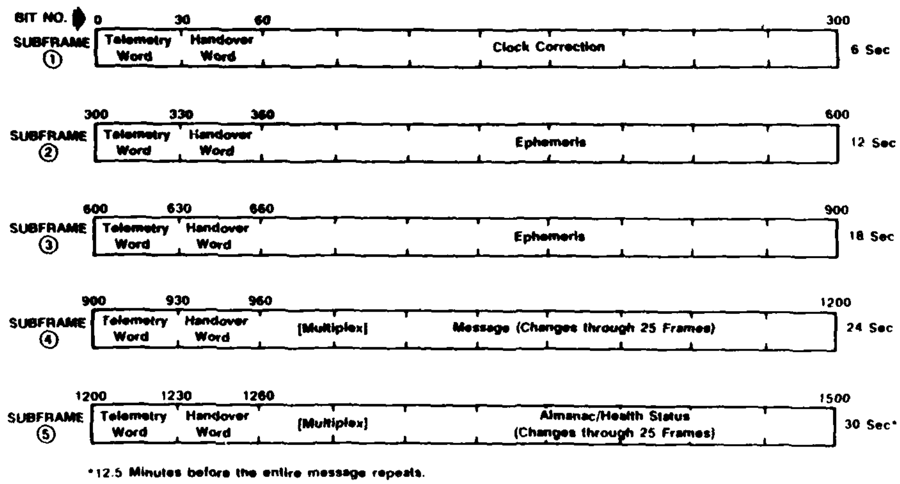


Figure 1. The Navigation Message

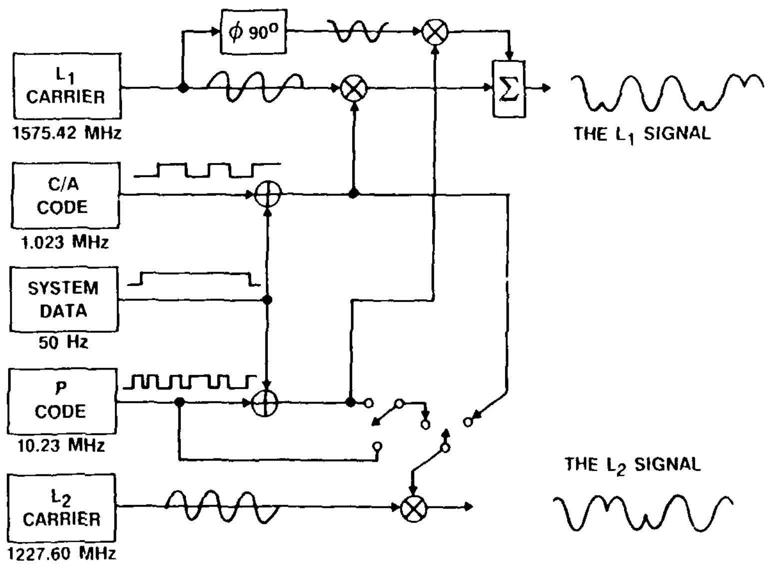


Figure 2. Satellite Signal Modulation

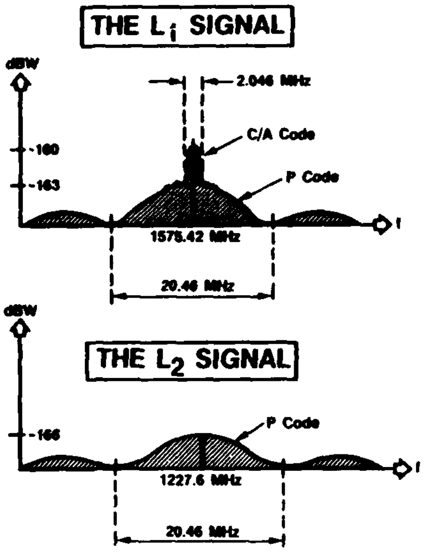


Figure 3. GPS Signal Frequency Spectrum

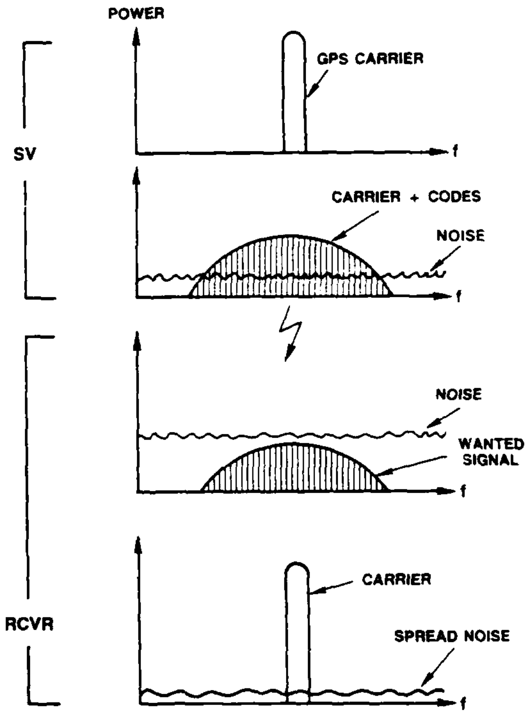


Figure 4. Spread Spectrum Used in GPS

USER EQUIPMENT OVERVIEW

Elio Bottari, Maj.
Italian Air Force
NATO Representative in the
GPS Joint Program Office
Space Division - CWNI
P.O. Box 92960
Los Angeles, California, USA 90009

1. INTRODUCTION

The GPS User System includes all the hardware and software needed to determine the user's position, velocity, time data, and other derived parameters, as required. Additionally, the GPS User System can be integrated with other navigation systems to provide accurate navigation under severe dynamics and hostile environmental conditions. During the development phase, various classes of users have been defined based on user requirements and characteristics such as desired accuracy, user maneuvering, antijamming capability, and cost. The application of GPS user equipment (UE) in various types of host vehicles, used under a wide variety of operational conditions, has led to the development of three types of UE sets:

- o LD = Low Dynamic (one channel)
- o MD = Medium Dynamic (two channel)
- o HD = High Dynamic (four or more channels).

Table 1 is presented to show the operational dynamic limitations for each class of equipment.

Table 1. Set Dynamic Operational Capabilities

	LINEAR DYNAMICS*			ANGULAR DYNAMICS*		
	VELOCITY** (m/sec)	ACCEL** (M/sec²)	JERK** (m/sec³)	YAW RATE (rad/sec) / ACC (rad/sec²)	PITCH RATE (rad/sec) / ACC (rad/sec²)	ROLL RATE (rad/sec) / ACC (rad/sec²)
LOW DYNAMIC SET	25	6	6	NOT SPECIFIED	NOT SPECIFIED	NOT SPECIFIED
MEDIUM DYNAMIC SET	150	40	40	±1.0 / ±1.0	±1.0 / ±1.0	±1.0 / ±1.0
HIGH DYNAMIC SET	1200	90	100	±1.0 / ±3.0	±1.0 / ±6.0	±5.5 / ±17.5

*Values shown represent maximum dynamic limitations
**Along any axis

In Table 2, a chart correlating UE set type versus possible applications is shown.

As can be seen in Table 2, the choice of set type, number of channels, and complexity of the receiver structure is primarily dependent upon the dynamics of the host vehicle, and, secondarily, upon the accuracy and interference resistance required. A typical UE architecture is shown in Figure 1, but other architectures are possible.

The control display unit (CDU) is an interface between the receiver and the operator. It is used to control the GPS set and to selectively display navigation/system status data. Two types of CDU are being developed: one for aircraft/shipboard application and one for manpack/vehicular applications. A typical version of the first type, for airborne and shipborne receivers, is shown in Figure 2.

A cathode ray tube displays the navigation/system status data with alphanumeric characters. The alphanumeric keyboard allows for data entry, and rotary switches are used to select operation mode and display pages. Typical CDU control functions are: receiver mode commands, waypoint definition and activation, mark definition, destination/desired track/desired vertical angle selections, altitude hold and stationary mode activation, map datum, and rendezvous mode selection. The manpack/vehicular CDU is designed for handheld operation, and it is typically composed of a liquid crystal display and a simple alphanumeric keyboard for data entry (see Fig. 3).

A brief description of GPS signals is given to better describe the functions and operating concept of user equipment. The two carrier signals broadcast by the satellites are designated L1 (1575.42 MHz) and L2 (1227.6 MHz). In order to determine position, the two carrier frequencies are biphase modulated by pseudorandom sequences,

Table 2. User Equipment Set Type versus Host Vehicle Type

	LOW DYNAMIC SET	MEDIUM DYNAMIC SET	HIGH DYNAMIC SET
PHASE III PRODUCTION APPLICATIONS	FIXED BASE OPERATIONS PERSONNEL PORTABLE GROUND VEHICLES WATERCRAFT	ROTARY WING (Army) FIXED WING (Army)	ROTARY WING (Air Force Navy) FIXED WING (Air Force Navy) SUBMARINES (Navy) SURFACE SHIPS (Navy)
REPRESENTATIVE PHASE II HOST VEHICLES	SOLDIER CARRIED MANPACK M 60 TANK	UH 60 HELICOPTER CV 64 AIRCRAFT CARRIER	A 6E AIRCRAFT F 16A AIRCRAFT B 52G AIRCRAFT SSN 701 SUBMARINE

providing a spread spectrum modulation. The L1 carrier is modulated by a C/A code (coarse acquisition code) and a P code (precision code). The L2 carrier is also biphase modulated by the same P code, or, as a ground controlled option, by the C/A code. The P code is a long sequence, repeating every 280 days and each satellite is assigned a week-long portion of this sequence. This long duration, in addition to a high chipping rate of 10.23 MHz, makes the P code appear to be random noise to an observer and, hence, is described as pseudorandom noise. The C/A code is short and it repeats every millisecond. The C/A code uniquely identifies a satellite, because each satellite broadcasts a different C/A code sequence. Due to the short duration and lower chipping rate, the C/A code is comparatively easy to acquire and can be used for initial acquisition of the P code. It can also be used as a medium accuracy navigation signal. A navigation message is superimposed on both P and C/A codes at a data rate of 50 bits/sec. The navigation message contains GPS system time, satellite ephemeris and clock data, satellite health status, coefficients for the ionospheric delay model, coefficients to calculate the universal time coordinates (UTC) and the HOW (hand-over word). The HOW is used in the receiver to transit from C/A to P code tracking.

2. ANTENNA SUBSYSTEM

An antenna subsystem consists of an antenna that performs signal reception and an antenna electronics (AE) unit that performs antenna pattern control, preamplification, and frequency down conversion (see Fig. 4).

The frequency down conversion within the antenna subsystem is a feature that allows for significant physical separation between the antenna subsystem and the receiver, without suffering performance degradation. This feature means installation flexibility and affordability, and it is particularly important in those applications where the receiver cannot be located close to the antenna (i.e., aircraft carrier, submarines, or airplanes). Nevertheless, some manufacturers use a different approach, such as submitting antenna input directly into the receiver, and, where remote installation is required, a radio frequency (RF) amplifier is used to regain attenuation loss in the cable (see Fig. 5).

Antenna placement on the host vehicle is important in two respects. First, there should be a clear view of the whole sky; shadowing of some satellites can result in degraded performance. Second, a potentially significant error source is multipath signals, particularly of the stationary sort produced by reflection from surfaces near the antenna. Today, all the GPS antenna subsystems can be classified into two wide categories: Fixed Radiation Pattern Antenna (FRPA) Subsystem, and Controlled Radiation Pattern Antenna (CRPA) Subsystem.

2.1 Fixed Radiation Pattern Antenna Subsystem

In the FRPA subsystem, the antenna, in most cases, is a relatively simple element providing approximate isotropic gain from the zenith to the horizon at one or both of the GPS frequencies and the pattern cannot be changed by the AE unit. Therefore, the AE for the FRPA will perform only filtering, preamplification, and radio frequency to intermediate frequency (IF) down conversion. In Figure 6, a functional block diagram shows signal processing performed by the antenna electronics for the FRPA.

Preselector filters for the L1 and L2 signals at the input of the preamplifier provide out-of-band signal rejection. The limiter preceding the RF amplifier serves to protect the circuits that follow, as well as the receiver, from damage by high magnitude interference. The L1 and L2 signals from the preamplifier then pass through two band pass filters (BPF) centered around L1 and L2, respectively. A local oscillator signal is conveyed to the antenna electronics from the receiver, in order to down convert the RF signals into IF signals. After the down conversion in the diplexers, the two IF signals are sent to the receiver. To date, many types of FRPA have been developed by many companies for low, medium, and high dynamics applications. Because the GPS signal is circularly polarized, quadrifilar or bifilar helix and spiral helix or crossed monopoles are among the most used techniques for the FRPA.

2.2 Controlled Radiation Pattern Antenna Subsystem

In the CRPA subsystem the AE unit takes on the extra processing function of changing the antenna reception pattern in order to reject interference signals and/or emphasize the GPS satellite signals. Therefore, the CRPA is a more sophisticated subsystem, microprocessor-controlled with phased array beam-steering and/or adaptive array null-steering capability. At this time, the CRPA developed for the U.S. program is a seven-element adaptive array antenna that can form nulls against six jamming sources simultaneously. Figure 7 is a typical functional block diagram showing the null steering process for a seven-element array.

A complex gain (W = Weight), amplitude and phase, is assigned to the signals coming from each of six antenna auxiliary elements, whereas the reference element will not be modified (W = 1). The total output of the array is fed back to the control loop. This is a negative feedback in order to minimize the total output power. Since the power of the jammer will normally be much greater than the power of the GPS signals, a null must be formed towards the jammer to minimize the output power. Adjusting the adaptive weights of the auxiliary elements, a null will be placed in the antenna pattern toward the jammer, as shown in Figure 8.

When no interfering or jamming signals are present, the CRPA subsystem performs as an FRPA subsystem.

3. RECEIVER SUBSYSTEM

The GPS receiver performs the tasks of satellite signal acquisition, signal processing, coordinate transformation, area navigation, and host vehicle interface (if required). Subsequently, it is appropriate to consider the receiver as the heart of the GPS User System.

In Figure 1 the receiver subsystem is shown divided into two parts: receiver/ processor and interface. This division is not rigid, but it is representative of two different tasks performed in the GPS receiver.

The interface consists of the portion of hardware and software that adapts the PVT (position, velocity, time) solution to the specific host vehicle requirements and converts the vehicle input data to internal computer format. Several different interfaces have been developed for U.S. programs to integrate the GPS user equipment into different platforms.

To meet the requirements of low, medium, and high dynamic applications, the receiver/processor, which performs the real tracking and signal processing, can be designed with different numbers of channels (such as one, two, or five).

3.1 GPS Receiver Types

A receiver may be defined by the type of satellite signal tracking used. There are basically two different types of receivers: continuous tracking and sequential tracking. In a continuous tracking set, each receiver channel is dedicated to a particular satellite signal; therefore, at least four channels are required to solve for the three unknown position components and the time. An additional channel can perform many functions, including dual frequency measurement for ionospheric delay, determining interchannel bias correction, and keeping track of all the satellites in the sky in order to select the next satellite to be tracked in a changing constellation (see Fig. 9).

The resultant "five-channel" set provides the highest accuracy under high dynamics, maximum antijamming capability, and the lowest TTFF (time to first fix). The TTFF is the elapsed time from the initial demand, on a set that has been turned on for a minimum of 7 min, to the subsequent display and output of position, velocity, and time with specified accuracy. These features make the five-channel set suitable for high dynamic vehicles such as fighter aircraft, for submarines (because of low TTFF), and for all users requiring good antijamming performance. In sequential tracking, the receiver channel is sequentially switched through the satellites. The receiver will track one satellite at a time, dwelling on each satellite for a relatively short period of time. Generally, the sequential receiver has the advantages of being cheaper, smaller, lighter, and requiring less power. But it has the disadvantages of poorer dynamic tracking capability, poorer jamming immunity, and longer acquisition time. To date, there are mainly two types of sequential receiver: the single-channel set for manpack and low dynamic platforms, and the two-channel set for medium dynamic applications. The single channel set measures four pseudoranges on both the L1 and L2 frequencies to compensate for ionospheric delay.

If the set uses only one frequency, the ionospheric delay will be estimated from a mathematical model, but the accuracy may be degraded (see Fig. 10).

In addition, the channel will read the navigation message to obtain ephemeris information. All these activities take a relatively long period of time. If the single-channel set is moved during the measurements of the four pseudoranges, the accuracy can degrade, and this is the main reason for restricting the single-channel set to low dynamic or stationary applications (see Fig. 11).

The two-channel sequential receiver uses one channel part of the time to perform various service functions (navigation messages, ionospheric compensation, etc.) and uses both channels the rest of the time to sequentially track four satellites. The use of two channels decreases the time needed for the receiver to generate the navigation solution.

Actually, there is another type of receiver that can be considered part of the family of sequential tracking sets: the multiplex receivers.

In a multiplex receiver, one hardware channel switches at a fast rate among the satellites being tracked and sampled data are collected continuously and processed in software. This is basically a time sharing technique applied to a single hardware channel that is also tasked to continuously read the navigation message from all four satellites being tracked.

3.2 Signal Processing and Receiver Operations

3.2.1 Acquisition Process

As mentioned in the introduction, the receiver/processor processes satellite signals, collects navigation message data, and makes pseudorange and deltarange measurements to compute position/velocity/time (PVT). To perform these tasks, the GPS proceeds through several steps, called states, before attaining steady-state operation.

The first state is acquisition, in which the receiver determines the satellites available for tracking. The receiver seeks a satellite C/A code using satellite position estimates residing in memories and enters the user's approximate present position, velocity, and time.

If no stored satellite almanac information exists, or only very poor estimates of position and time are available, the receiver will start a search-the-sky mode, attempting to locate and lock on to any satellite in view. When the first satellite is tracked, the receiver demodulates the navigation message and reads the almanac information about all satellites in the constellation.

If the receiver platform is equipped with a very precise time reference, such as an atomic clock, the receiver can acquire the P code directly without first going through the C/A code tracking state. This faster acquisition is extremely valuable for submarines, because it reduces the time the submarine must remain on the surface.

3.2.2 Tracking Process

The received satellite signal level near the earth is less than the background noise level. Correlation techniques are used by the receiver to obtain the navigation signal. A carrier tracking loop is used to track the carrier frequency, whereas a code tracking loop is used to track the C/A code or the P code.

The two tracking loops will work together in an iterative process, aiding each other to acquire and track the satellite signals (see Fig. 12).

The received signal is code correlated by the code tracking loop so that the carrier tracking loop can track the carrier frequency. The carrier tracking loop then compares a locally-generated L1 or L2 frequency with the received signal.

The two carrier frequencies are generally different because there is:

a. A doppler offset of the satellite carrier signal caused by the relative velocity, along the line of sight, between the user equipment and the satellite.

b. A bias residual in the receiver frequency generator.

The carrier tracking loop adjusts the receiver-generated frequency until it matches the incoming carrier frequency and, thereby, determines the relative velocity between the receiver and the satellite. Using the four relative velocities with four satellites, the receiver then calculates its velocity in earth-centered-earth-fixed (ECEF) system coordinates.

Similarly, the code tracking loop generates a replica of the satellite C/A code with an estimated ranging delay. To match the received signal with the internally-generated replica, the two signals must have the same center frequency and the same phase. The center frequency of the replica is set using the doppler estimate from the carrier tracking loop.

The phase-offset between the receiver code and the internally generated code is directly proportional to the pseudorange, which can be easily calculated. The word pseudorange is used because the measurement of range is made by the user, measuring code delay time with an imprecise clock, and, therefore, a bias of fixed magnitude (for a given set of measurements) is included in each range estimate due to the clock error.

At the beginning of the tracking phase, the satellite code will not correlate with the locally-generated code due to the time delay for the satellite signal to reach the receiver, and the receiver clock offset. The locally-generated code is shifted until maximum correlation is achieved between the two codes. The magnitude of the shift determines the value of the pseudorange.

When the two tracking loops are locked on the received signal, the navigation message is demodulated from the carrier frequency to obtain the data for accurate pseudorange calculations. In addition, the HOW contained in the navigation message is read to obtain P code phase information for the transfer of C/A code to P code.

Once the receiver has acquired four satellite signals, achieved carrier and code tracking, and read the navigation message, the GPS receiver accurately calculates position, velocity, and time, and navigation can start.

The position and the time are derived by solving the four equations in Figure 13, inputting the four measured pseudoranges. The receiver solves the four simultaneous equations for U_x, U_y, U_z (user position) and CB (user clock bias).

The velocity is determined, using the same types of equations as in Figure 13, by inputting relative velocities rather than pseudoranges. In practice, these computations are performed by the UE navigation software in a Kalman filter to allow continuous navigation even if some measurements are missing or not time-synchronized and to exploit a priori knowledge of user dynamics. Calculating velocity by differencing the pseudoranges would also be feasible, but it is somewhat less accurate than carrier-loop-based velocity computation.

3.3 Receiver Operating Modes and States

During the development phase of the GPS User Equipment Program, special terminology was introduced to describe operational requirements of the receiver. The navigation modes and the operating states were categorized in d'fferent classes so that all contractors involved with the GPS user equipment Program could orient design, development, set performance, and test requirements in accordance with the Government specifications.

To date, this system of classification is generally accepted in the GPS community, and will be presented to describe how GPS receivers operate in different integration architectures and under various jamming and dynamic conditions.

3.3.1 Navigation Modes

The navigation modes are divided into the following categories:

- Unaided Modes U1, U2, U3
- Aided Modes A1, A2, A3, A4, A5

In the unaided modes, no host sensor data are available and the UE set operates only on GPS receiver data.

Mode U1 (Full Accuracy): Four or more GPS transmitters are tracked by the receiver, and the optimal constellation of four satellites is selected on a "weighted" GDOP optimality criterion.

Mode U2 (Degraded Accuracy): In this mode, the number of GPS transmitters that the receiver tracks is less than four due to poor visibility and/or host vehicle dynamics and/or jamming constraints. The set continues to navigate but with degraded accuracy.

Mode U3 (Great-Circle Navigation): In this mode, no satellite measurements are available and the set computes present position along a constant altitude, great-circle path, based upon the best known position and velocity. Under these conditions, the operator can manually update position and horizontal velocity (e.g., ground speed and ground track). Of course, the navigation accuracy of the set is undefined for this mode.

In an aided host vehicle, the set can support the following aided navigation modes.

Mode A1: In this mode, the GPS receiver and an inertial measurement sensor are the only sensors supplying inputs to the navigation data processor.

Mode A2: In A2 mode, the GPS receiver and an altitude sensor are the only sensors providing inputs to the navigation data processor.

Mode A3: In this mode, the GPS receiver and a velocity (speed and heading) sensor are the only sensors supplying inputs to the navigation data processor.

Mode A4: In A4 mode, the GPS receiver and a heading and attitude sensor are the only sensors supplying inputs to the navigation data processor.

Mode A5: In this mode, combinations of all available data from the host vehicle (not previously defined) and the GPS receiver are processed.

3.3.2 Receiver Operating States

The GPS receiver can operate in the following seven different states.

State 1 (Normal Acquisition): In this state the receiver tries to acquire the C/A signal (L1 or L2), using doppler estimates derived from almanac data, plus present position, velocity, and time inputs from the operator or host vehicle system. Subsequent to reading and verifying the hand-over word (HOW) from the C/A signal, the receiver will acquire and track the associated P signals.

State 2 (Direct Acquisition): In this state, the receiver directly acquires the P signal without first acquiring its associated C/A signal. The P code phase and frequency estimates for the acquisition are derived from almanac or current ephemeris data, present position and velocity estimates, and a precise input of universal time coordinates (UTC).

State 3 Code Lock: In State 3 the receiver is able to maintain code lock (e.g., detect the signal and be able to make coarse pseudorange measurements), but is unable to main tain precise carrier tracking (i.e., carrier phase delta-range measurements might be inaccurate). The receiver will revert automatially to State 4 or 5 when dynamic excursion or jamming levels do not exceed the carrier tracking thresholds of the receiver.

State 4 Carrier Lock: In this state, the receiver maintains carrier lock (i.e., makes delta-range measurements to less than full accuracy). In addition, pseudorange measurements may be made to less than full accuracy and data may be demodulated.

State 5 Carrier Track/Data Demodulation: In State 5, the receiver precisely tracks the carrier and is able to demodulate system data from the carrier. In addition, pseudorange and pseudodelta-range measurements are made to full accuracy. If L2 is being tracked as the primary navigation signal and does not have biphase data, the requirement to demodulate data is not applicable.

State 6 Sequential Resynchronization: In this state, the receiver serially measures pseudorange and delta range to the GPS satellites. Of course, the five-channel "continuous tracking" does not have this state. Dependent on dynamics and jamming levels, this state may be subdivided to include States 3, 4, 5, and 7 as substates.

State 7 Signal Reacquisition: In State 7, the receiver tries to reacquire the signal, after it has been in a tracking state but has subsequently encountered a loss of signal.

4. CONCLUSIONS

Results from Phases I and II demonstrate that GPS user equipment can provide the superb accuracy anticipated by early studies. A family of user equipment has been developed as part of the U.S. GPS program and a limited initial production has started.

The user equipment will continue to evolve as rapid advances in electronics make more demanding user requirements feasible. Factors that will influence further development in user equipment are demands for reduced volume, weight, cost, and power consumption; increased ease of calibration and maintenance; and improved performance.

Manufacturers not formally under contract to the U.S. Government Joint Program Office are independently developing their own GPS equipment and, in some cases, are already providing enhancements. For instance, some companies are developing a receiver that will replace some existing analog functions with digital processes. This "digital receiver" will reduce the complex analog portion of the receiver, significantly reduce cost, and yet ease maintenance and calibration requirements.

Gallium arsenide (GaAs) technology will dominate the architecture of future receivers, thus improving both processing speed and power consumption.

In summary, technological advancements in hardware components and packaging, and improved software programming techniques, will make GPS user equipment smaller, lighter, and cheaper. These enhancements, in conjunction with improved performance, will lead to more tractable integration and many more applications of GPS.

REFERENCES

1. The Nato Team, Introduction to Navstar GPS User Equipment, USAF, Space Division/LAAFB, CWNI (June 1987).

2. USAF, User's Overview, GPS Joint Program Office, Space Division/LAAFB (September 1987).

3. Collins Government Avionics Division, The Global Positioning System Papers. (March 1985).

4. USAF, SS-US-200 User System Segment Specification, GPS Joint Program Office, USAF, Space Division/LAAFB (January 1979).

5. Institute of Navigation, Global Positioning System Papers, Vol I and Vol II, Washington, DC.

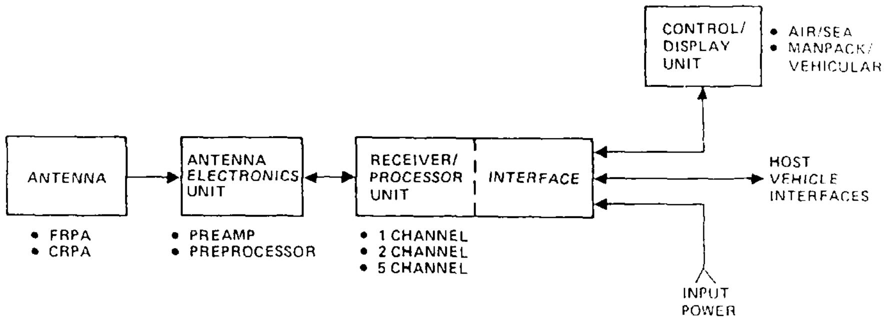


Figure 1. Description of a Typical GPS User Equipment Set

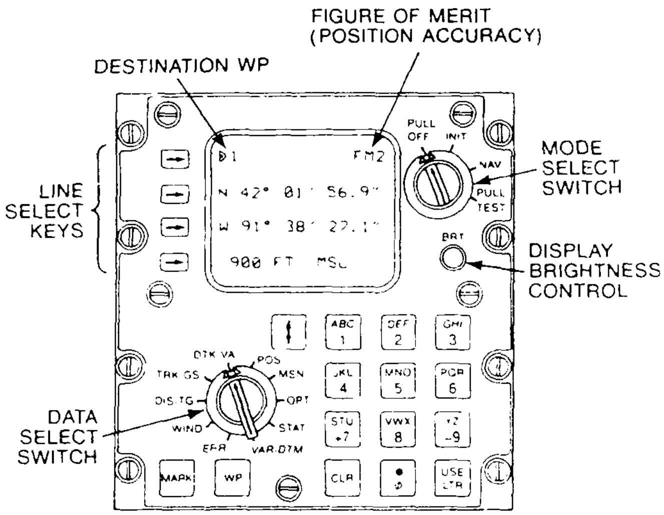


Figure 2. Typical Control Display Unit for Aircraft/Shipboard Applications

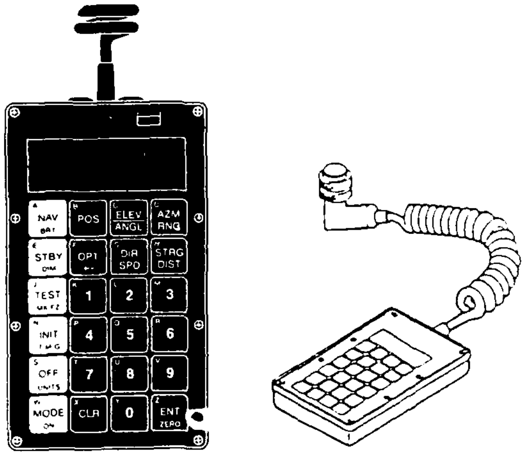


Figure 3. Typical Manpack/Vehicular Control Display

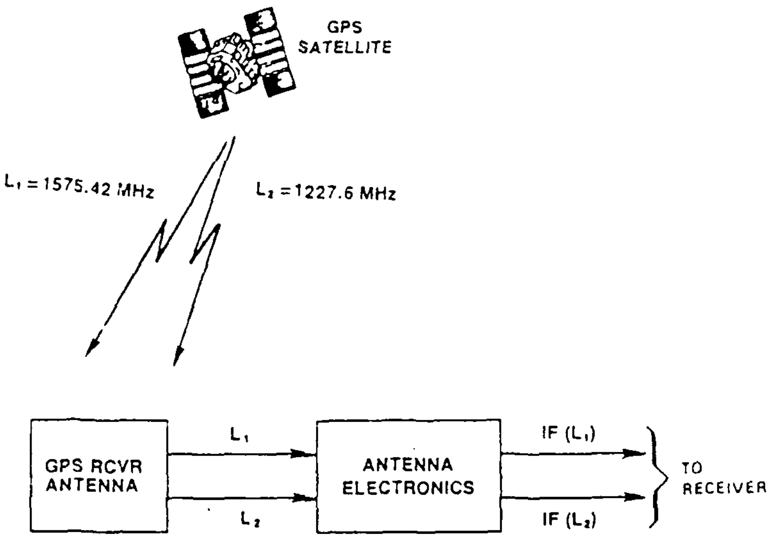


Figure 4. Simplified Functional Diagram of GPS Antenna Subsystem

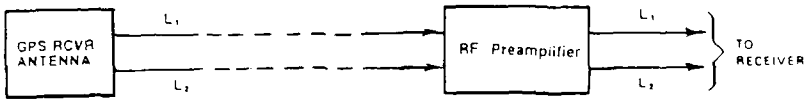


Figure 5. Remote Antenna Subsystem Functional Diagram

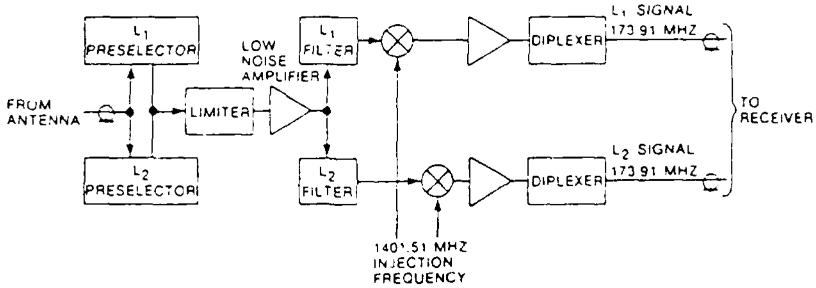


Figure 6. Functional Diagram of Antenna Electronics for FRPA

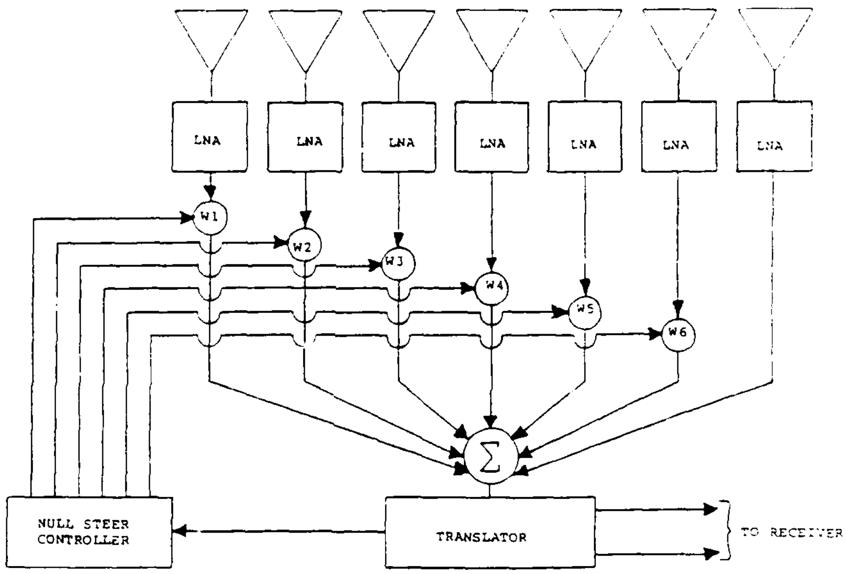


Figure 7. Functional Block Diagram of Null Steering Process for a 7-Element Antenna

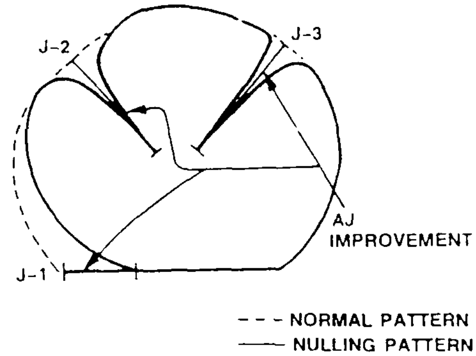


Figure 8. CRPA-Adaptive Array Antenna Performance Pattern

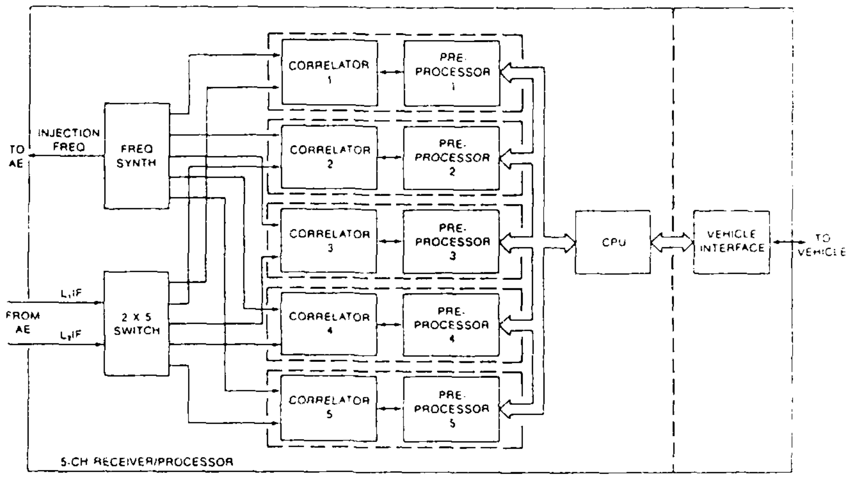


Figure 9. Block Diagram of a Typical 5-Channel GPS Receiver

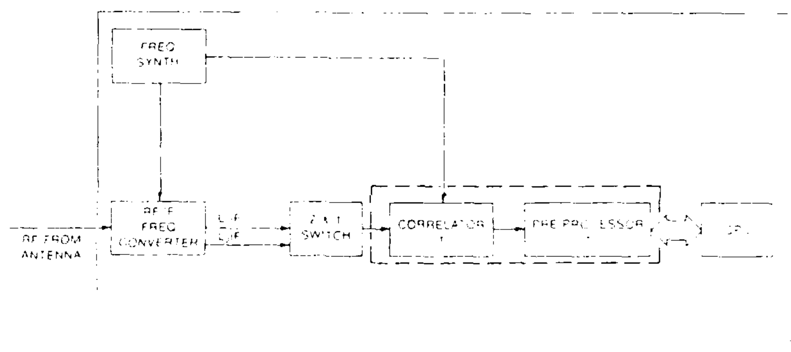


Figure 10. Block Diagram of a Typical 1-Channel GPS Receiver

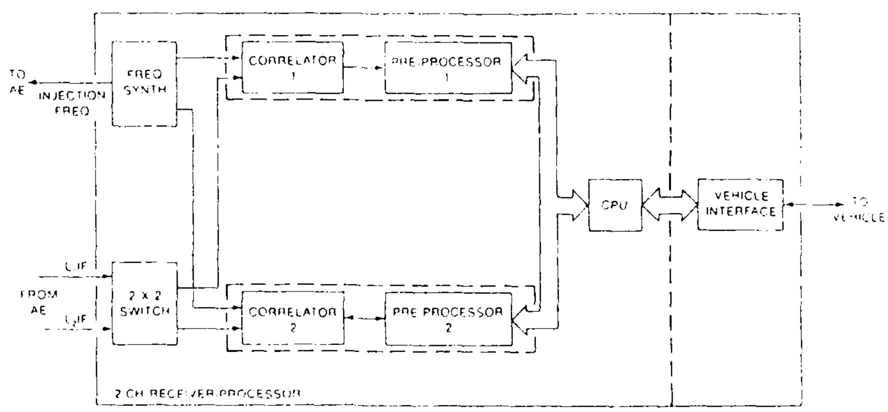


Figure 11. Block Diagram of a Typical 2-Channel GPS Receiver

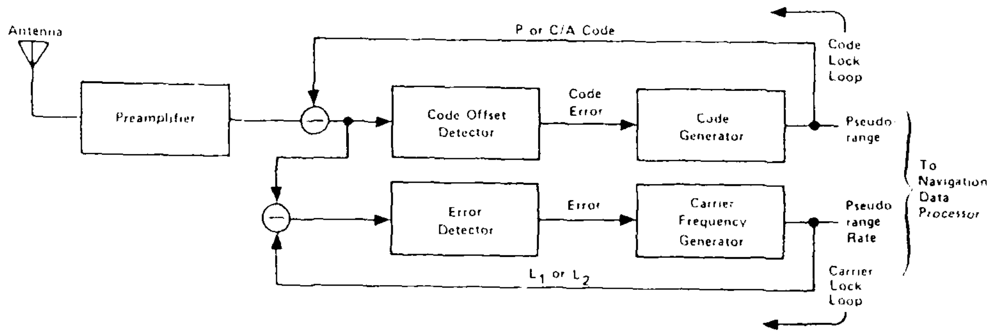
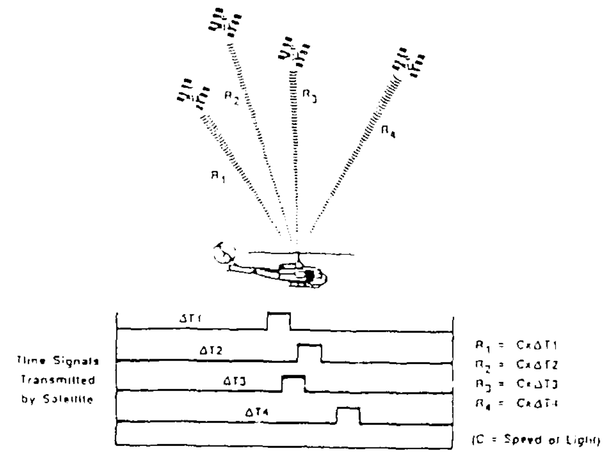


Figure 12. Generic Receiver Operation Block Diagram



USER SET PERFORMS THE NAV SOLUTION FOR POSITION

PSEUDO RANGES:	POSITION EQUATIONS:
$R_1 = C\,\Delta t_1$	$(X_1 - U_X)^2 + (Y_1 - U_Y)^2 + (Z_1 - U_Z)^2 = (R_1 - C_B)^2$
$R_2 = C\,\Delta t_2$	$(X_2 - U_X)^2 + (Y_2 - U_Y)^2 + (Z_2 - U_Z)^2 = (R_2 - C_B)^2$
$R_3 = C\,\Delta t_3$	$(X_3 - U_X)^2 + (Y_3 - U_Y)^2 + (Z_3 - U_Z)^2 = (R_3 - C_B)^2$
$R_4 = C\,\Delta t_4$	$(X_4 - U_X)^2 + (Y_4 - U_Y)^2 + (Z_4 - U_Z)^2 = (R_4 - C_B)^2$

R_i = PSEUDO RANGE (i = 1,2,3,4)

- o PSEUDO RANGE INCLUDES ACTUAL DISTANCE BETWEEN SATELLITE AND USER PLUS SV CLOCK BIAS, USER CLOCK BIAS, ATMOSPHERIC DELAYS, AND RECEIVER NOISE
- o SV CLOCK BIAS AND ATMOSPHERIC DELAYS ARE COMPENSATED FOR BY INCORPORATION OF DETERMINISTIC CORRECTIONS PRIOR TO INCLUSION INTO NAV SOLUTION

Figure 13. Receiver Navigation Solution Concept

GPS/INERTIAL NAVIGATION SYSTEM INTEGRATION FOR ENHANCED NAVIGATION PERFORMANCE AND ROBUSTNESS

Robert P. Denaro
G. Jeffrey Geier
TAU Corporation
485 Alberto Way
Los Gatos, California 95032

INTRODUCTION

The emergence of GPS from the laboratory and the test range into the operational environment, with affordable, high-performance receivers now available, has heralded a widespread study of the integration of GPS with the current standard in precise positioning, inertial navigation systems. Once considered by some to be a potential competitor of inertial navigation, GPS is now recognized by most navigation system developers as having tremendous potential for system performance improvement over either system in isolation by integrating with inertial systems. So powerful is this synergism that deep integration of the two systems often dictates potential major design modifications in the inertial measurement unit (IMU) as well as the GPS receiver. These design modifications are necessary to exploit the performance potential to the maximum, particularly in the military environment, and at the same time to realize substantial cost savings in the system as an additional benefit.

MOTIVATION FOR GPS/INS INTEGRATION

The benefits of integration of GPS with an inertial navigation system (INS) or inertial measurement unit (IMU) are significant and diverse. Basically, each system has important shortcomings when used in a stand-alone mode (although either can be integrated with other sensors to alleviate some of these shortcomings).

The inertial navigation system is subject to an ever growing drift in position accuracy caused by various instrument error sources that cannot be eliminated in manufacturing, assembly, calibration or initial system alignment. A later section on Inertial Navigation System Fundamentals discusses typical instrument errors in various classes of inertial sensors. Inertial navigation systems require a somewhat lengthy static alignment period or, alternatively, continuous operation in a powered "alert" status, for best operation. Operation of an aircraft from an aircraft carrier, for example, requires some means of transfer alignment from the ship's inertial system. Furthermore, high quality inertial systems are complex electro-mechanical devices with significant risk of component failure. Mission or flight critical implementations of inertial navigation require substantial redundancy in installation to detect, isolate, and recover from such failures for fail-operational performance.

Stand-alone GPS has its shortcomings as well. The system is somewhat vulnerable to loss of signals due to hostile jamming, antenna shadowing, aircraft attitude maneuvers, or other causes. With the requirement to process radio frequency signals and the complex processing required to formulate a position or velocity solution, the update rates of solution are usually at 1 Hz, or at best 10 Hz. While sufficient for most navigation applications, some avionics reference functions and interfaces require much higher rates.

Used in concert through some adequate integration configuration, the integrated GPS/IMU solves all of these problems and more. Basically, the marriage is between a short term, very precise measurement source with poor long term stability, the IMU, and a very long term stable reference of less precision, GPS. The benefits of this integration are:

- Long term drift bounding

  The GPS provides a more-than-adequate reference for bounding or resetting the inherent drift characteristic of the IMU.

- Improved Jamming Resistance

  Use of "velocity aiding" from the IMU by the GPS allows narrowing of signal tracking loops and thereby significant rejection of hostile jamming signals [1].

- Improved fault-tolerance

  GPS and the IMU provide dissimilar redundancy for detecting hard and soft failures in either system.

- Inertial System Calibration

  With GPS present, continuous calibration of conventional inertial error sources is possible, yielding much improved inertial solution stability if GPS is temporarily lost for any reason [1].

- Improved GPS acquisition and reacquisition

  The inertial operates as a rough initialization device for initial GPS signal acquisition as well as for rapid reacquisition should signals be temporarily lost [2].

- High-rate Position and Velocity Solutions

  If the IMU is used as the basic process in the navigation solution, the accuracy of GPS is essentially extrapolated with great precision by the IMU between GPS solution updates.

- Incomplete GPS Measurement Set Use

  With a healthy IMU, error bounding and accurate navigation can continue with only two or three satellites in track, instead of the nominal 4 required for stand-alone GPS solutions [3,4].

- Attitude Solutions

  The IMU provides precise attitude and heading measurements to complement the precise position and velocity solutions of the GPS.

- Dynamic Alignment

  GPS enables, and perhaps improves, the initial alignment of the IMU through maneuvers once airborne, eliminating the static alignment pre-mission requirement [5].

The principles of operation and integration concepts that enable these performance enhancements in the integrated GPS/IMU system will be described in the following sections. Before describing these concepts, a foundation of relevant GPS and inertial principles of operation are discussed.

GPS FUNDAMENTALS

The general design and operation of GPS receivers has been extensively addressed in the literature and will not be repeated here [1,6,7,8,9]. Instead, we will focus on those characteristics that are critical to the understanding and analysis of the GPS/IMU integration task. The critical design components are the tracking loop design and reference oscillator.

A generic block diagram of a GPS receiver is shown in Figure 1. The GPS signal, after reception and intermediate frequency (IF) downconversion, is input to the receiver baseband processing. Many modern receivers convert the final IF stage to a digital signal through analog-to-digital convertors, performing the correlation function digitally [9,10]. Early and late replicas of the code generated internally are used for code tracking (delay lock loop) while the prompt version is used for carrier tracking and data demodulation. Some receivers employ extended range detection and tracking which enables generation of a tracking error signal over 1 chip wide (29 meters P-code, 293 meters C/A-code) which is useful for maintaining lock under conditions of high error dynamics and narrow tracking loop bandwidths [9,10]. Since phase tracking is commonly performed prior to data demodulation, the Costas form of a phased lock loop is usually used for carrier tracking [7]. The code tracking process yields the pseudorange measurement for navigation processing. The input to the Numerically Controlled Oscillator (NCO) of the carrier tracking loop is typically integrated over a predetermined interval (e.g., 1 second or less) to produce the deltarange measurement for navigation processing.

The code tracking loop is usually mechanized as some form of a Delay Lock Loop (DLL); the error signal can be generated coherently or noncoherently, depending on the application. When carrier tracking is maintained, many receiver designs aid the code tracking loop with the very precise range rate information derived by carrier tracking. For applications in which an INS is present, INS-derived velocity replaces the carrier aiding signal when carrier tracking cannot be maintained. This simplifies the design of the code loop filter, since the code loop does not have to track host vehicle dynamics. Carrier tracking is usually performed by a Costas tracking loop, but the more classical Phase Lock Loop (PLL) is also used in some applications. Selection of appropriate bandwidths for the code and carrier tracking loops is, of course, driven by the expected range of dynamics of the host and the anticipated noise environment. Usually, the lowest possible bandwidth which can tolerate the worst-case dynamic transients and maintain signal lock is chosen, to permit generation of the most accurate (i.e., minimum noise content) pseudo and deltarange measurements for navigation processing. However, the stability of the local oscillator can also establish lower limits on the possible tracking loop bandwidth, particularly in very low dynamics, or when the dynamic tracking is provided external to the receiver (e.g., when INS-aided) [11]. The accuracies with which pseudorange and deltarange measurements are derived vary significantly, depending upon the application. Pseudorange errors can vary from the submeter level to several meters, one sigma, while deltarange errors can vary from the subcentimeter level to several centimeters. Generally speaking, carrier tracking will be lost when the signal to noise ratio (C/N_0) falls to 27 dB-Hz (Costas mechanization) or 21 dB-Hz (PLL mechanization). After carrier lock has been lost, code tracking can continue to roughly 12 dB-Hz if INS aiding is present.

The source for the local replica code and carrier timing and other signal processing timing functions is the local oscillator, usually a high quality, temperature-controlled quartz oscillator. The oscillator can approach 1 part in 10^9 in stability. In addition to its random walk error characteristics, most quartz oscillators have a strong acceleration-induced error in each of their orthogonal axes.

The g-sensitivity of the crystal results in an error in signal phase input to the receiver tracking loops. This is interpreted as vehicle dynamic motion, and may be of sufficient magnitude to cause loss on carrier lock in a narrow bandwidth receiver. Hence, for high acceleration cases, the g-sensitivity of the quartz oscillator may be the primary constraint for minimum bandwidth selection. The result is higher susceptibility to jamming and dynamics-induced noise in the solution [11].

Since vehicle dynamics and other factors can cause problems in maintaining carrier lock compared with the benign case, there is an incentive to design adaptive tracking loop bandwidths so that the maximum antijam performance can be maintained under all conditions. When velocity aiding is available, loop bandwidths can be reduced to minimize interference. Many modern receivers employ such functionality to achieve performance optimization [8,12].

INERTIAL NAVIGATION SYSTEM FUNDAMENTALS

As with the GPS Fundamentals discussion in the previous section, we will not document the entire design of the inertial navigation system, but rather discuss those design features and variations that are significant to the problem of design of an integrated GPS/IMU package.

Obviously there are a multitude of possibilities for inertial components for integration with GPS. In aircraft with existing inertial navigation capability, the inertial system may be a very high quality gimballed local-level system with electrostatically suspended gyros or ring laser gyros [13]. At the other end of the spectrum, an integration in a small vehicle or weapon may use a coarse strapdown system with dry-tuned rotor gyros or even piezoelectric sensors [2,14]. Future developments are anticipating the use of hemispherical resonator gyros and solid-state fiber optic gyros [15,16]. The error characteristics and magnitudes of these various sensors vary as widely as the technology employed.

In all cases the role of the inertial system in the integrated GPS/INS system is to provide a near-instantaneous measure of the vehicle acceleration to anticipate signal doppler shifts and aid Kalman filter state (process) modeling. With GPS operational, this is a short time interval requirement. If GPS is temporarily lost, then this requirement may be for a longer period of time, perhaps minutes or significant portions of an hour. The mission requirements for accuracy will determine the required stability of the inertial system for these two roles, although the analysis is not completely straightforward since very often the critical performance parameters are with GPS-computed calibration factors for the inertial components.

All Inertial Navigation Systems measure acceleration in some frame and integrate it to produce velocity and position information. The frame in which the acceleration is derived can be maintained physically (i.e., as a stabilized platform) or mathematically, by appropriately integrating measured angular velocity. Inertial Navigation Systems can be broadly classified into three categories: Space Stable, Local Level, and Strapdown. In the Space Stable mechanization, the accelerometers are held non-rotating in inertial space by a gyro stabilized platform, while the local level mechanization maintains a platform which is locally level. Both Space Stable and local level INS require three or more gimbals to isolate the stabilized platform from the host vehicle's angular motion. Strapdown systems are therefore simpler, lighter, smaller, and generally less costly, since the accelerometers and gyros are attached directly to the host vehicle. For terrestrial navigation, the Local Level and Strapdown mechanization are most common. (The INS utilized by the Space Shuttle is an example of a Space Stable mechanization.)

The major sources of error for any INS are imperfections in the gyros and accelerometers, and an imperfect knowledge of the earth's gravitational field [17]. A gravity model is used to compensate the accelerometers, which cannot distinguish gravitational forces from the vehicle's inertial acceleration. Significant gyro and accelerometer errors can include non-repeatable instrument biases, residual scale factor errors, g-sensitivities, and g-sensitive gyro drift, in addition to input axis misalignments due to imperfect mounting, and calibration. Depending generally upon the instrument quality, INS error drifts can range from 0.1 nm/hr to 10.0 nm/hr. Although the INS errors can grow unbounded, the temporal behavior has a very well defined frequency behavior; horizontal errors will generally oscillate at the so-called Schuler frequency (with an 84 minute period) and at an earth-rate frequency (24 hour period). Thus, the error behavior can be well modelled as a linear time invariant system.

The inertial error sources that will be of interest in the integration of GPS and inertial are:

- gyro bias
- gyro scale factor error
- accelerometer bias
- accelerometer scale factor error
- gyro and accelerometer misalignment (for strapdown)
- gyro and accelerometer tilt error (for gimballed)

It is important to note that the classical error sources for standalone inertial navigation systems may not be critical for inertial components aided by GPS. For example, turn-on to turn-on bias is quickly calibrated at the beginning of a mission with GPS operational. Similarly, other traditional "first order" error sources may be easily handled in the integrated system while the second order terms dominate the residual error and present the most significant challenges for error compensation.

The magnitude of the error sources can vary widely depending on the quality of the inertial components. Since GPS offers the potential of continuous update and compensation, it is conceivable that very low grade components may be used for the inertial. At the other end of the spectrum, an integration with existing equipment may involve a very high quality, redundant inertial navigator.

LEVELS OF INTEGRATION

Integration of GPS with inertial sensors is not a simple matter of a binary decision to integrate or not to integrate [14,18]. In practice, several levels of integration are possible, each with associated performance benefits and implementation tradeoffs. The negative aspects of some levels of deeper integration are significant; especially when considering the practical aspects of LRU accessibility, data access, processor resources and integration of existing hardware with existing avionics architectures [18]. Therefore, the tradeoff of level of integration must be addressed with care in each application, based on mission needs and system constraints [13].

Criteria for Assessment of Integration Requirements

Integration levels vary from maintaining separate solutions, with perhaps a "vote" selection of the best solution, to deeply integrated mechanisms where major functions typically within either of the subsystems are now performed in a integrated component or software algorithm. The principal criteria by which to compare integration methods in a GPS/inertial system are:

- Jamming immunity for the GPS receiver, which can be significantly enhanced by the use of inertial information
- Navigation performance
- Fault tolerance
- Solution stability, which can be compromised in some integration schemes
- Processor load
- Raw data access
- Data time-tagging ability, a critical requirement for effective integration
- Data update rates and latency, often driven by the needs of other avionics, such as a flight control system
- Use of attitude data, a potential output of the inertial components but not GPS
- Dynamics tolerance, especially coupled with jamming immunity which can vary dramatically with integration technique

Classification of Levels of Integration

The levels of integration can be described in any number of ways. However, for this application certain categories stand out as distinct due to their complexity levels, impact on hardware/software design, quantifiable performance benefits, and satisfaction of mission requirements.

The primary levels of GPS/inertial integration for the aircraft environment are:

- Separate solutions, where the GPS and inertial maintain independent navigation solutions, with perhaps different functions using one or the other, and with possibly a selection of one solution or the other if certain thresholds are exceeded.
- Simple integrated filters, where each solution is still independent but they may be combined in some sort of complementary filtered algorithm.
- Measurement data filtering, which adds weighting of geometric effects of GPS. May also include Kalman filtering which automatically addresses relative "health" of each measurement source.
- Inertial error state estimation, which is the first level of data integration which improves the quality of the INS solution.
- GPS bandwidth control and code loop aiding, which enhances the jamming resistance of the GPS receiver.
- Deep integration with integrated tracking loop closure, an advanced implementation for maximum performance and stability.

Advantages and Disadvantages of Integration Levels

Separate solutions are obviously the simplest to implement, and fit best with a constrained "black box" approach. Such a system provides the first level of redundancy, allowing system failure detection but not necessarily isolation and subsequent operation unless other parameters are monitored or other built-in-test features can resolve the problem. Of course, in most situations the separate solution integration is the simplest to implement in that only off-the-shelf "black boxes," with minimal external software development, are required. Stability of solution is easy to achieve.

The major drawback of separate solutions is that no additional jamming immunity is gained for the GPS receiver and the inertial performance is improved slightly or not at all (other than an upper bound to drift based on GPS accuracy). This integration approach would typically be used only in applications where access to the existing avionics architecture is severely constrained.

Simple integration filters are an obvious improvement on the separate solutions approach where there is an available processor that can input both GPS and inertial solutions and process them together. This may provide complementary filtering or other methods to weigh the input solutions based on known performance expectations. GPS is a good long-term sensor (error does not grow without bound) and the inertial is an excellent short-term sensor (integrating highly accurate accelerometers), which is the ideal situation for complementary filters. Since this method works with the output states of the sensors, as opposed to the raw measurements, the filtering is limited in its ability to observe measurement noise or geometric effects, so the filtering performed is somewhat "ad-hoc." Again, no jamming immunity improvement is possible, and there is no fault tolerance improvement to the system. Navigation performance is improved in nominal conditions with this technique. Stability of the solution must be handled carefully, but at least for simple complementary filters stability is not a problem. Early implementations of integrated navigation systems, using austere processors, often used this approach.

The next level of integration, measurement data filtering, is a major improvement in performance. Now the filter directly observes the measurement geometry and can formulate measurement residuals in a Kalman filter for optimal weighing of the measurements in the final solution. Again, the nature of the GPS long-term accuracy and the inertial short-term accuracy is ideal for the Kalman filter. In this case, the inertial is often used as the "model" of the state dynamics, or fundamental process at work in the navigation problem, while GPS provides fairly frequent (e.g., 1 Hz) updates. The demands on the integration processor are still reasonable for most applications.

This technique, with no further additions, does not help jamming immunity, but performance in dynamics is improved somewhat. Since there is still no feedback from the integration processor to either the GPS receiver or the inertial system, the independent performance of these sensors is not improved. Fault tolerance has a small improvement over the previous level with some additional ability to observe the soft failures, but this impact is minimal. Time-tagging is now important since the inertial data are critical to the process and the filter will expect the inertial measurements and GPS measurements to agree. In practice, this technique is usually implemented only with inclusion of the next level, inertial error state estimation.

The biggest advantage of inertial error state estimation is that the inertial is continuously calibrated by the filter (using the GPS observables), so that when GPS is lost due to jamming, dynamics or satellite shadowing, the inertial can continue the overall solution, but now as a highly precise unit by virtue of its recent calibration. This approach enables the use of far more inferior inertial components which, after calibration, will perform as a maximum performance unit for a small period of time until the nonlinearities take over. This is the first level at which improvements to the inertial attitude solution are gained as well.

The Kalman filter for this application is typically 11, 14 or 17 states, which could be a large load for many existing processors in the avionics architecture. Obviously, at the higher state sizes, required for austere inertial components, update rates may suffer. In addition, barometric altimeter inputs are usually added to ensure stability of the inertial solution in the unstable vertical axis with GPS out. Stability is generally achievable as long as measurement correlations are accounted for.

GPS bandwidth control and code loop aiding constitute the first level at which the GPS receiver performance is enhanced, particularly for jamming and dynamics performance. Actually, code loop aiding can be implemented without bandwidth control. For code loop aiding, the inertial delta-velocities are used to pre-position the doppler effect on code tracking, so that code lock can be maintained after carrier lock is lost (the first to go in a jamming environment). The performance impact is significant, particularly in the presence of dynamics. The typical military high performance GPS receiver implements this technique in its internal Kalman filter by inputting inertial measurements.

Adaptive GPS bandwidth control can be implemented if the receiver provides for such control. This will enable bandwidth reduction for enhanced jamming rejection, and bandwidth increase during periods of higher dynamics uncertainty to prevent loss of lock due an overly narrow tracking window. In practice, however, very few GPS receivers accommodate this level of control.

Both of these techniques have serious potential stability problems due the positive feedback loop of inertial-to-GPS and back to inertial. Many GPS integrations in manned aircraft use this level of integration, and various ad-hoc techniques are employed to prevent solution divergence [19].

The final level of integration is integrated tracking loop closure, which may require major modification to the typical GPS receiver [19]. In this technique, the GPS code loop is not closed separately before input of pseudorange measurements to the Kalman filter, but instead the (averaged) outputs of the code loop error detector are passed directly to the Kalman filter, which closes the code loop using its best estimates of vehicle position and velocity. Since there is not a separate code loop producing pseudorange errors with long (10-100 second) correlation times unmodeled by the Kalman filter, there are no potential stability problems when using this approach. An iteration rate of 1 Hz is generally adequate for the Kalman filter. However, the Kalman filter cannot be designed independently from the code tracking loop, since the two are so tightly coupled.

The tradeoff of levels of integration is obviously a complex issue that must be analyzed in light of the mission requirements and environment, the hardware available, and the navigation processing architecture and flexibility [20].

INTEGRATION DESIGN

GPS Quartz Clock Error Compensation

Besides the integration of the GPS and inertial system, the final system can benefit from additional sensors usually found on the aircraft, such as the barometric altimeter. Also, because of the importance of the quartz oscillator to the GPS solution, compensation of its acceleration-dependent errors can improve performance and possibly enable further tracking loop bandwidth reduction that would otherwise not be possible.

The decision to attempt to estimate the g-sensitive coefficients is strongly driven by two factors: the expected magnitude of the sensitivities, and the stability of the sensitivities. The magnitude may be kept quite small by design, but can still represent a significant error contributor in a high dynamic environment. The sensitivities must be fairly stable to permit estimation: when both code and carrier tracking are maintained, estimates of the g-sensitive coefficients must be sufficiently reliable to allow further bandwidth reduction when carrier tracking has been lost; thus, a minimum stability of several minutes is required.

Given that online estimation is desired, the following analytical treatment should lead to optimal estimation by the Kalman filter. The rate of change of user clock phase error can be written as:

$$d\dot{\Phi} = df + c \cdot ab \quad (1)$$

where:

c is the vector of coefficients to be estimated by the Kalman filter.
ab is the vehicle acceleration vector expressed in the body frame
$d\Phi$ is the nominal, g-insensitive frequency error of the clock

The body frame acceleration is sensed by the accelerometers, once they are gravity-compensated. Since the accelerations can be fairly rapidly varying, Eq. (1) must either be integrated numerically using a small timestep, or, equivalently, the analytic integral of the body frame acceleration components must be calculated. This latter approach is preferred, since the integral is readily found from the delta velocity outputs of the individual accelerometers. Thus, if online estimation of the coefficient vector c is attempted, the state transition matrix utilized by the Kalman filter will include delta velocity dependent terms for propagating the clock phase error estimates.

Barometric Altimeter Aiding

A barometric altimeter is typically included in stand-alone inertial system integration for damping of the otherwise unstable vertical loop. In a GPS/inertial integration, integration of the altimeter is still a good idea for the GPS-out performance, and can aid poor vertical geometry situations under nominal conditions.

Two basic options exist for using the barometric altimeter data:

- conventional, constant gain damping of the INS vertical channel;
- processing the barometric altimeter as a measurement to the Kalman filter.

The second option is generally preferred, since the vertical channel is effectively stabilized through the "optimal" Kalman filter gains, rather than a set of "ad hoc" fixed gains. In addition, the barometric altimeter measurement will only be processed when GPS satellite coverage is incomplete (as determined from the covariance matrix). Otherwise, the barometric altimeter bias error is calibrated using an offline filter, such that calibrated barometric altimeter data is available for measurement processing.

Carrier Loop Aiding

The use of IMU-derived velocity information to aid the GPS carrier tracking loop offers the potential for maintenance of carrier lock under fairly high jamming conditions. Normally, when unaided, the carrier loop will unlock at a J/S ratio of roughly 41 dB under worst-case dynamics. Under fairly ideal circumstances, IMU aiding can extend this threshold by 6-7 dB; however, these ideal circumstances may be difficult to achieve in practice. In order to maintain carrier lock, peak tracking errors must generally be kept to less than one quarter of a wavelength, or roughly 5 centimeters. This constraint places very severe requirements upon the IMU aiding:

- a very high update rate is required
- timing delays must be kept to a few milliseconds
- prediction algorithms, which estimate jerk levels from the delta-velocity histories, are required
- the IMU should be collocated with the GPS antenna

The last bullet arises due to possible uncompensated, flexible motion between the IMU and GPS antenna, which can induce loss of carrier track. This concern may make carrier loop aiding impractical for some applications.

Code Loop Aiding

Unlike carrier loop aiding, rate aiding of the code loop when carrier lock has been lost is typically performed by most integrated GPS/INS sets. Since the code loop can tolerate tracking errors up to roughly fifteen meters, the accuracy requirements on IMU derived velocity are much less stringent. The code loop aiding can be of two general forms: "partitioned" or "integrated." They are illustrated for a single receiver channel in Figures 2 and 3, respectively (from [19]). The so-called partitioned approach basically attempts to separate the signal tracking and navigation functions, rather than considering them as a single integrated function, as in the integrated approach. As illustrated in Figure 2, INS derived velocity, corrected by Kalman filter estimates, is summed with the outputs of the code loop filter (simply a time varying gain), and used to drive the Numerically Controlled Oscillator. Its output, denoted T, is supplied to the error detection equations and passed to the Kalman filter as a pseudorange measurement. The outputs of the integrate and dump circuit are passed to the detector equations every few milliseconds. The detector outputs are then averaged by the prefilter and multiplied by the code loop gain. Note that the loop closure is at a high rate (e.g., every 20 msecs), even though the Kalman filter velocity corrections are updated once per second. The integrated approach, illustrated in Figure 3, closes the code loop entirely through the Kalman filter, offering some immediate advantages over the partitioned approach. The Kalman filter is able to adjust the code loop bandwidth as a function of both the noise and dynamics environment of the receiver, whereas the partitioned design's bandwidth (proportional to k) is adjusted only as a function of the sensed noise environment. In addition, the integrated approach circumvents the well-known stability problem [19] associated with the partitioned approach, since the filter processes (directly) the outputs of the code loop detector, i.e., it does not process pseudoranges with unmodeled long correlation time error components.

Two, specific performance comparisons, abstracted from [19], are germane to the general high performance aircraft applications. The first comparison corresponds to a vehicle flying straight and level subjected to a steadily increasing jamming level input to all channels: the J/S level was increased by roughly 5 dB every 100 seconds. Results for a single, representative channel are plotted in Figure 4. Note that plus and minus one sigma values, computed from the Kalman filter covariance matrix, are superimposed upon the code tracking loop errors. The integrated approach exhibits a modest improvement: lock is lost at a J/S level somewhere between 60 and 65 dB, whereas the partitioned approach loses lock at 60 dB. A much more dramatic improvement is realized by the integrated approach when the vehicle is performing maneuvers, as illustrated in Figure 5. Here, the J/S level is set at 60 dB, and the vehicle performs 5, 7, and 9g turns in succession, spaced roughly 120 seconds apart. As evidenced by Figure 5, the partitioned approach loses lock shortly after the 5g turn, whereas the integrated approach stays in lock through the 9g turn. The significant difference in performance can be attributed to the code loop bandwidth adjustments made by the Kalman filter in response to the turns. To summarize, the integrated approach is preferred in a dynamic environment. Given expected severe dynamics for the a high-performance aircraft application, the integrated method for code loop aiding is the obvious selection.

IMPLEMENTATION CONSIDERATIONS

Failure Tolerance Considerations

Failure tolerance can generally be enhanced through hardware and/or analytic redundancy [21]. Redundant inertial systems, e.g., a "dual quad" arrangement of gyros and accelerometers, are often utilized, together with a parity checking algorithm to enable detection, identification, and removal of failed instruments. Generally, to detect and isolate k failures requires 2k + 3 instruments sensing a three dimensional quantity, e.g., angular velocity. Thus, five gyros are required to detect and remove the effects of a single failure. The "dual quad" mechanization, where two physically separated clusters of inertial instruments (with three input axes along the edges of a cube, and the fourth along the diagonal), is motivated by the potential loss of a single cluster. At any time, only one cluster is used for navigation. Detection of a failure in the active cluster (which requires only 4 gyros) activates the second cluster.

In addition to detecting "hard" failures, where the instrument produces no useful output, it may also be desirable to compensate for "soft" failures (corresponding to degraded instrument operation). Statistical tests on Kalman filter residuals, coupled with hypothesis testing for the failure mode is generally used, and can provide some measure of analytic redundancy.

For smaller and low cost aircraft applications, use of redundant IMU hardware may be too costly and/or occupy too large a volume. If redundant IMU hardware is not available, a backup capability should be supplied, i.e., a GPS only navigation mode [3]. Transitioning to this mode is triggered by Built-In-Test (BIT) detection of a gyro and/or accelerometer failure, and involves a redefinition and restructuring of the Kalman filter. The Kalman filter, up to the downmoding, has been estimating errors in the INS solution. These states must therefore be redefined to represent errors in the current best estimates of vehicle state; the Kalman filter must operate as an extended Kalman filter following the detected IMU failure.

Generally speaking, GPS failures are more easily accommodated than IMU failures. The lock detection schemes are fairly reliable, and the integrated navigation solution will degrade gracefully to a calibrated inertial system when GPS measurements are unavailable. There is a specific, possible soft failure mechanism for GPS worthy of specific attention. In a very heavy jamming state, it is generally very difficult for a GPS receiver to determine if code lock has been maintained with reasonable confidence. Thus, it is possible that invalid pseudorange measurements can be passed to the Kalman filter. As a minimum, the Kalman filter should be made insensitive to this; in addition, with the integrated code loop aiding scheme proposed, the Kalman filter can be utilized in determining the receiver lock status. Several methods can be used to desensitize the filter to suspect pseudoranges, including delaying the correction of the filter state until lock can be determined with certainty, maintaining two parallel Kalman filters (with one filter avoiding the suspect measurements) until lock is determined with some certainty, and delaying measurement updates. Of these three alternatives, the last approach is preferred due to its simplicity: no additional computations are required; rather, the filter needs only to accommodate a variable update interval.

Measurement Timetagging

In an integrated GPS/Inertial system the inertial is informing the GPS receiver tracking loops of sensed dynamics. Obviously, if this information arrives "late," the tracking loop may have already suffered the consequences of the dynamics. The challenge in an operational installation is to achieve near-instantaneous information transfer from the inertial to the GPS receiver so that no latency exists [14].

Inertial System Resets

When an inertial system is calibrated, there exists a tradeoff of whether to "reset" the gyros, in a closed loop fashion, or to simply carry the drift value in software as an open loop implementation. The only problem arises when the cumulative inertial drift is outside of a linear range, after which the system models are inadequate and system performance suffers. This is especially noticeable when the system is force to degrade to (calibrated) inertial standalone mode due to loss of GPS tracking.

Most gyros provide for resetting of the drift, but the penalty is temporary confusion of the integration Kalman filter which does not expect to see discontinuous changes in the error sources.

KALMAN FILTER FORMULATION

A typical state vector for integration of GPS with an inertial system is illustrated in Eq. (2). As indicated, a total of 17 states are modeled.

$$\hat{x} = \begin{bmatrix} \underline{dp} \\ \underline{dv} \\ \underline{d\theta} \\ \underline{da} \\ \underline{d\omega} \\ d\phi \\ df \end{bmatrix} \begin{matrix} \text{INS position errors} \\ \text{INS velocity errors} \\ \text{INS attitude errors} \\ \text{INS accelerometer biases} \\ \text{INS gyro drift rates} \\ \text{GPS user clock phase error} \\ \text{GPS user clock frequency error} \end{matrix} \qquad (2)$$

The first 9 states are required to accurately model the INS error dynamics. Note that barometric altimeter error is not included, since it is assumed that conventional damping of the INS vertical channel is not performed. Instead, the baro altimeter is processed as a measurement to the Kalman filter. The decision to include instrument error sources (i.e., accelerometer and gyro errors) is driven by the expected mission: reliable estimates of the instrument errors can significantly improve navigation performance during periods of GPS outage. Possible additional states include accelerometer and gyro scale factor errors, gyro g-sensitive drift, and receiver clock g-sensitive coefficients (a total of 12 additional states). Their inclusion is driven by the expected dynamic environment. The fundamental INS error states (i.e., position, velocity, and attitude) can generally be represented in any frame. For processing the GPS measurements, an ECEF frame is optimal. This implies that the slowly varying transformation between ECEF and geographic frames be included in the Kalman filter state dynamics and process noise covariance matrices, since instrument errors are usually defined in the geographic (for local level implementation) or body frame (for strapdown implementations). The process noise covariance matrix is intended to represent the effects of unmodeled IMU error sources, so is usually computer dynamically as a function of sensed specific force and angular velocity.

The measurement processing functions performed by the Kalman filter are dependent upon the jamming state of the GPS receiver. In an unjammed state, when both code and carrier tracking loops are in lock, both pseudorange and deltarange measurements are available to the Kalman filter. The pseudorange measurement residual is found by subtracting the range estimated using the best estimate of vehicle position (formed by correcting the INS-based position with the current estimates of INS position error) and the position of the GPS satellite computed from the ephemeris data, and the best estimate of the clock phase error:

$$PR_{RES} = PR_{MEAS} - (\hat{R} + \hat{d\phi})$$

The measurement gradient vector is given by:

$$h_{PR}^{T} = [\underline{u}\ \underline{0}\ 1\ 0]$$

where u is a unit line of sight vector to the satellite of interest. Thus, the INS position errors and the user clock phase error are directly observable in the pseudorange measurement. The deltarange measurement, since it is formed by differencing carrier phase estimates over a predetermined time interval, requires that the best estimate of position be computed at both ends of the interval:

$$DR_{RES} = DR_{MEAS} - [(\hat{R}_{SP} - \hat{R}_{SR}) + (\hat{d\phi}_{SP} - \hat{d\phi}_{SR})]$$

where $\hat{R}_{SR}$ = best estimate of the range to the GPS satellite of interest at the start of the deltarange interval

$\hat{R}_{SP}$ = best estimate of the range to the GPS satellite of interest at the end of the deltarange interval

Computation of the measurement gradient vector for the deltarange measurement requires calculation of the pseudorange gradients at the ends of the interval and the inverse of the state transition matrix over the interval. Thus, INS velocity and attitude error, in addition to modeled instrument errors, and user clock frequency error are directly observable in the deltarange measurement.

In a jammed state, carrier tracking has been lost and the code tracking loop is aided by INS derived velocity. In this case, only pseudorange measurements are available for processing by the Kalman filter, and the measurement processing function is dependent upon the aiding approach. In the so-called partitioned approach, the equations are the same as those used in the unjammed state. However, caution must be exercised, since the pseudorange errors can have very long time constants, owing to the reduced bandwidth of the code tracking loop. This is the source of the well-known stability problem associated with this configuration. As a minimum, the rate of processing the pseudorange measurements should be reduced to avoid contamination of the estimates by the correlated pseudorange error component. In the integrated approach, on the other hand, no potential stability problem exists, so the measurement update rate need not be constrained. However, the pseudorange residual equation must be changed, owing to the different code tracking loop configuration. Since the Kalman filter effectively closes the code loop for the integrated design, the output of the code loop detector (averaged over 1 second or more) represent the Kalman filter residual, and can be input to the measurement update equations as such.

CONCLUSIONS

GPS and inertial integration is a complex subject whose design depends on mission requirements as well as the sensors and processing resources. Jamming immunity and continued accurate navigation once GPS is lost are major objectives of the integration. In spite of the system complexity, the performance potential is a tremendous improvement over either system in isolation, and sophisticated integration schemes with emphasis on thorough data processing can minimize the cost of components and, hence, system cost. With such potential, it is likely that most military and many civil applications will incorporate some level of GPS/inertial integration.

BIBLIOGRAPHY

1. Cox, Jr., D.B., "Integration of GPS with Inertial Navigation Systems," AGARD Conference Proceedings No. 298, France.

2. Knight, D.T., "GPS Integration with Low-Cost Inertial Navigation Unit," The Institute of Navigation, Proceedings of the Satellite Division First Technical Meeting, Colorado Springs, Colorado, September 1987.

3. Schwartz, H.M., Sensitivity Analysis of an Integrated Navstar GPS/INS Navigation System to Component Failure," Journal of the Institute of Navigation, Vol. 30, No. 4, Winter 1983-84.

4. Widnall, W.S., and Diaz Bobillo, I.J., "Partial Inertial Aiding for Low Cost Aircraft GPS Navigators," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

5. Tafel, Jr., R.W., and Krasnjanski, D., "Rapid Alignment of Aircraft Strapdown Inertial Navigation Systems Using Navstar Global Positioning System (GPS)," AGARD Conference Proceedings No. 298, France.

6. Dyment, M.J., and Liang, Dr. D.F., "Dynamic Performance Analysis of Navstar/GPS Navigation Filters," AGARD Conference Proceedings No. 298, France.

7. Krishnamerti, G., Harshbarger, S.A., and Smith, T.N., "The Design and Performance of GPS Phase II User Equipment Navigation Software," Journal of the Institute of Navigation, Vol. 32, No. 3, Fall 1985.

8. Buechler, D., and Foss, M., "Integration of GPS and Strapdown Inertial Subsystems into a Single Unit," Journal of the Institute of Navigation, Vol. 34, No. 2, Summer 1987.

9. Vacherlon, J.F., Huneke, A.C., Kaiser, G.M., Forseth, D.C., and Justice, J.H., "GPS Phase III Multi-Channel User Equipment," The Institute of Navigation, Proceedings of the Satellite Division First Technical Meeting, Colorado Springs, Colorado, September 1987.

10. Du Hart III, J.H., and Semler, J.R., "An Integrated GPS/Strapdown Inertial Navigator Postmission Processor and Initial Test Results," The Institute of Navigation, Proceedings of the Satellite Division First Technical Meeting, Colorado Springs, Colorado, September 1987.

11. Przyjemski, J.M., "A Compensation Technique for Acceleration-Induced Frequency Changes in Crystal Oscillators," Proceedings of the IEEE 1978 National Aerospace and Electronics Conference, Dayton, Ohio, May 1978.

12. Lottman, D., Homb, K., and Brown, A., "Kalman Filter Implementation in the Litton LR-80 AHRS," Proceedings of the IEEE 1984 National Aerospace and Electronics Conference, Dayton, Ohio, May 1984.

13. Divakaruni, S.P., Stensland, R.A., and Brenner, M.B., "Ring Laser Gyro Inertial and GPS Integrated Navigation System for Commercial Aviation," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

14. Teasley, S.P., "Flight Test Results of an Integrated GPS and Strapdown Inertial System," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

15. Loper, E.J., et al, "Projected Performance of Smaller Hemispherical Resonator Gyros," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

16. Pavlath, G.A., and Suman, M.C., "Fiber-Optic Gyroscopes: Advances and Future Developments," Journal of the Institute of Navigation, Vol. 31, No. 2, Summer 1984.

17. Joos, Dr. D.K., and Krogmann, U.K., "Estimation of Strapdown Sensor Parameters for Inertial System Error-Compensation," AGARD Conference Proceedings No. 298, France.

18. Sturza, M.A., Brown, A.K., and Kemp, J.C., "GPS/AHRS: A Synergistic Mix," Proceedings of the IEEE 1984 National Aerospace and Electronics Conference, Dayton, Ohio, May 1984.

19. Copps, E.M., Geier, G.J., Fidler, W.C., and Grundy, P.A., "Optimal Processing of GPS Signals," Journal of the Institute of Navigation, Vol. 27, No. 3, Fall 1980.

20. Hudak, G.J., "Strategies for Optimizing GPS Capabilities in Host Vehicle Integrations."

21. Rounds, S.F., and Casey, J.M., "A Fully Integrated GPS/Doppler/Inertial Navigation System," The Institute of Navigation, Proceedings of the Satellite Division First Technical Meeting, Colorado Springs, Colorado, September 1987.

22. Diesel, J.W., "Integration of GPS/INS for Maximum Velocity Accuracy," Journal of the Institute of Navigation, Vol. 34, No. 3, Fall 1987.

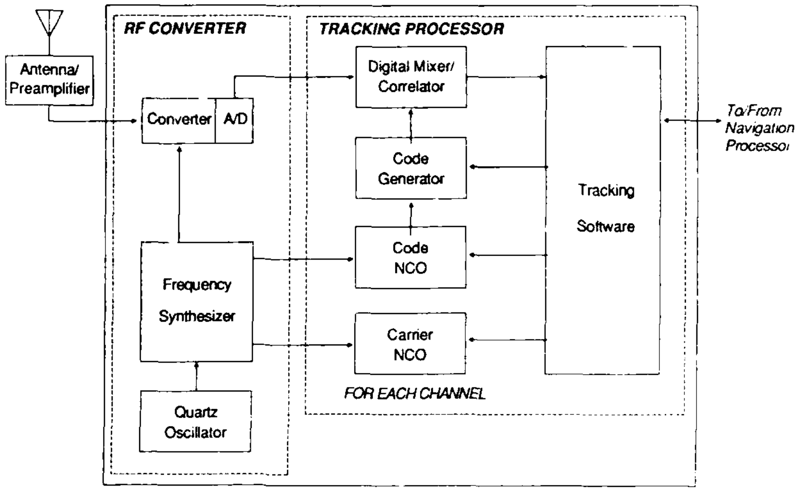


Figure 1. Generic Block Diagram of a GPS Receiver

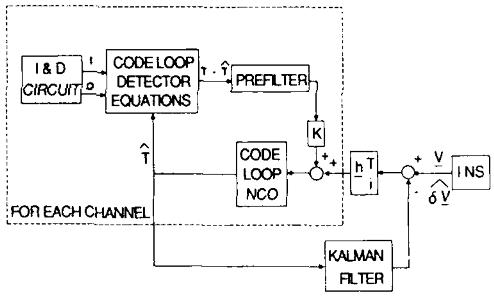


Figure 2. Partitioned Inertial Aiding Scheme (Single Channel)

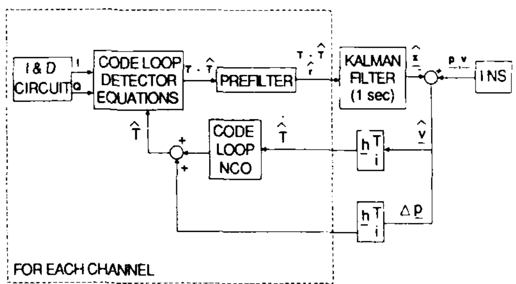


Figure 3. Integrated Inertial Aiding Scheme (Single Channel)

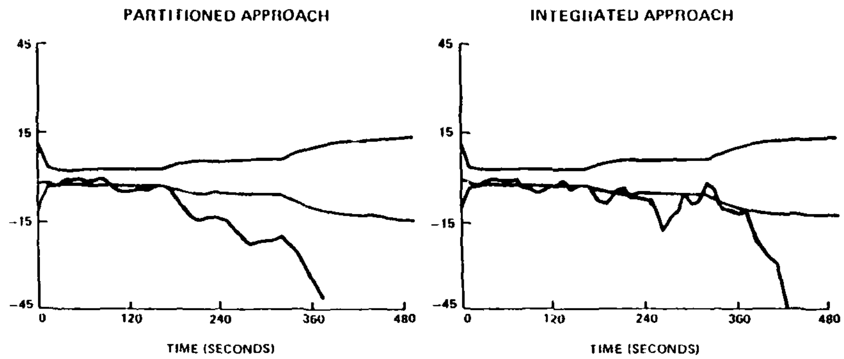


Figure 4. Performance Comparison, Partitioned and Integrated Approaches, for Maximum Jamming Test

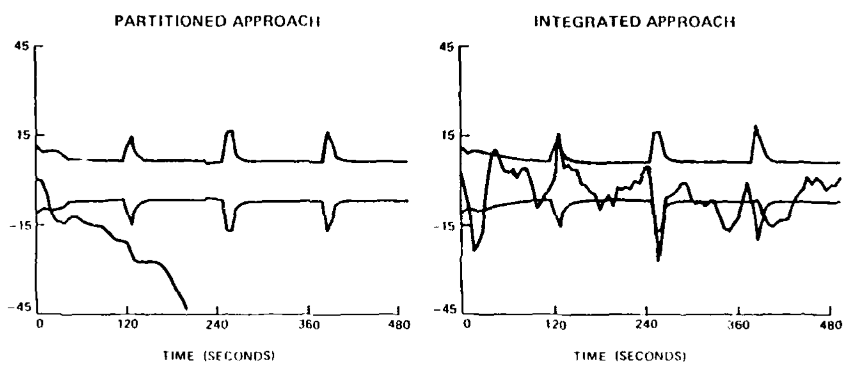


Figure 5. Performance Comparison, Partitioned and Integrated Approaches, for Maximum Jamming and Dynamics Test

AIDING AND INTEGRATION OF A GPS RECEIVER

Lt. Cdr. Per W. Nieuwejaar
Naval Materiel Command
Norway

SUMMARY

The need for receiver aiding due to contingencies such as satellite outages, jamming, and terrain masking is described; alternative aiding methods are also described, such as:

- o Integration GPS/INS or Attitude Heading Reference System (AHRS)
- o Integration GPS and baro-altimeter
- o Integration GPS and atomic clock
- o Use of altitude hold or mean sea level information
- o Integration of GPS/Transit/LORAN C and/or Omega.

1. INTRODUCTION

The GPS system is usually described as an all-weather, global, day and night navigation system. Unfortunately, situations occur where performance is degraded. There will be times when only a limited number of satellites are available. The GPS receiver can also use aiding to improve its operation. It can use time aiding to reduce time to first fix (TTFF) and to perform direct P-code acquisition. In addition, the receiver can use velocity and position aiding to improve its signal processing and to maintain lock under jamming and/or high dynamics. A GPS receiver is also well suited for integration with other navigation systems such as INS.

2. AREAS OF DEGRADED PERFORMANCE

There will be situations wherein the GPS user cannot achieve high accuracy navigation using GPS, because of poor satellite geometry, reduced pseudo range accuracy output from one or more satellites, satellite failure, satellite shadowing, or the limited number of satellites available during constellation buildup. Consequently, GPS will have variable performance, depending on user location and time, due to satellite geometry and availability. In addition, each individual satellite will perform differently and provide a different level of pseudo range accuracy.

2.1 Poor Satellite Geometry

With the planned constellation of 18 satellites and three active spares, there will be short periods during the day, in some locations, when the position dilution of precision (PDOP) will be high due to satellite geometry. The term, "area of degraded performance," is used to describe an area where the PDOP is greater than six during a 24 hour period. These areas are shown in Figure 1. Most GPS users will have sufficient navigation accuracy using GPS as a stand alone system even if the PDOP is greater than six. Many GPS users can navigate safely with navigation accuracy of hundreds of meters (ships in the open ocean, aircraft enroute over oceans, etc.) and can, therefore, use GPS even with a PDOP value of 50 to 100 or more. The relationship between PDOP, pseudo-range accuracy, and position accuracy is shown in Figure 2.

The PDOP can exceed six in some locations for up to 1 hr per day, but most areas of degraded performance last less than 0.5 hr per day (Figure 1). A PDOP greater than 10 will occur for 37 min per day in a few locations, but most areas will only last between 15 and 30 min (Figure 3). A PDOP greater than 50 will only occur for less than 4 min a day in any area of degraded performance (Figure 4).

2.3 Reduced Ranging Signal Accuracy Output from One or More Satellites

The range error due to control- and user segment error should be on the order of 3 to 4 m. Control-segment problems or satellite failures could result in a larger range error for one or more satellite signals. This, combined with the geometry factor (PDOP), will increase the user's position error. The impact of increased pseudo range error depends on geometry, error size, and on how many and which satellites are used in the GPS receiver navigation solution. An illustration of the impact, if all satellites experience the same degradation due to a control segment error, is given in Figure 2.

2.4 Satellite Failure

If one or more of the 21 satellites fails, there will be longer periods of high PDOP over broader areas, and less than four satellites in view. How long these time periods will be, in what locations, etc., depends on which satellites fail.

2.5 Satellite Shadowing

The satellite signal can be shadowed from the receiver antenna because of terrain shadowing or shadowing due to host vehicle (HV) maneuvering

2.6 Satellite Shadowing

The buildup of the satellite constellation will take 2 to 3 years. In the interim there will be a limited number of satellites available for navigation, often with unfavorable geometry.

3. AIDING INPUTS

The following inputs can be used to aid a GPS receiver in acquiring or tracking on GPS satellite signals:

- Position
- Velocity
- Acceleration
- Time
- Roll
- Pitch
- Heading
- Altitude
- HV Fixed Lever Arm

Aiding is required where host vehicle dynamics, jamming, or lack of satellites can preclude a GPS receiver from generating a valid PVT solution. The role of aiding in improving the receiver operation during a period of bad geometry or lack of satellite coverage is shown in Figure 5.

3.1 Position Aiding

When a GPS receiver is first initialized for operation, its approximate position is required to minimize satellite acquisition time. The accuracy requirement for the U.S. DoD program for position information is ±100 km of actual GPS receiver location.

3.2 Velocity Aiding

Velocity aiding is normally supplied to a GPS receiver during initialization, similar to the position aiding information just described. The aiding requirement for the velocity aiding data during initialization is +100 m/s of actual GPS receiver velocity to ensure that satellite acquisition time is within specification. During normal receiver operation, velocity is used as an input to the receiver Kalman filter. The inputs are used for measurement preprocessing, as well as for receiver aiding. Preprocessing provides estimates of pseudo-range, delta-range, and pseudo-range acceleration for use by the receiver processor. Velocity aiding of the code tracking is required during state three (code tracking only) operations.

3.3 Acceleration Aiding

Acceleration aiding is normally not provided to a GPS receiver per se. Instead, velocity aiding data is supplied at a sufficiently high data rate (8 Hz or more) such that the time rate of velocity change (i.e., acceleration) is available for receiver aiding.

3.4 Time Aiding

Time aiding can be used during the initialization process, similar to position and velocity data, as previously described. The time accuracy requirement is +20 sec relative to UTC. This is to ensure that satellite acquisition time is within the specification. Time aiding, if sufficiently accurate, can also be used during the satellite acquisition process. This is to enable a direct P-code acquisition without first acquiring the C/A code.

This type of time aiding is relevant to host vehicles such as submarines, where minimum exposure time of the GPS antenna on the ocean surface is of prime importance. An atomic time standard is necessary to enable direct P code acquisition. An atomic time standard may also be used to aid a GPS receiver when tracking less than four satellites. During normal receiver operation, four pseudo ranges are required to solve for the unknown quantities of position (X, Y, Z) and receiver clock bias C_b. An atomic standard can be substituted for C_b and the unknown values of X, Y, and Z can still be calculated when tracking only three satellites.

3.5 Attitude Aiding

Attitude aiding in the form of roll, pitch, and heading is used, together with lever arm corrections, by a GPS receiver to propagate the navigation solution to a position in the host vehicle other than at the GPS receiver antenna. Normally, an aided GPS receiver will propagate its navigation solution to the center of the aiding source, i. ., to the gimbal center of an INS. Attitude aiding is also used to transform position and velocity aiding parameters into earth centered earth fixed (ECEF) coordinates, if necessary. A GPS receiver usually does all internal calculations in ECEF before carrying out any coordinate transformations on any data. On ships, attitude aiding is used to compensate for antenna motion and, together with water speed and gyro compass heading information, to do relative course and speed calculations.

3.6 Altitude Aiding

3.6.1 Altitude Aiding for an Airborne Receiver

An airborne GPS receiver can use baro-altimeter aiding data as a Kalman filter measurement. Long term altimeter errors are calibrated during periods of four-satellite operations. Subsequently, when less than four satellites are being tracked, the calibrated baro-altimeter data is used as a known Z value in the four unknowns of X, Y, Z, and C_b. Baro-altimeter calibration values are deweighted over a 1000-sec time interval, a rather large time period when aircraft operations are considered. In simulations of high-dynamic platforms, when tracking three satellites augmented with baro-altimeter aiding, a gradual loss of vertical position accuracy has been predicted. Errors of up to 60 m in vertical position information are likely to occur after 500 sec of operation with three satellites and baro-altimeter aiding.

3.6.2 Altitude Aiding for a Shipborne Receiver

A GPS receiver used by surface ships and submarines close to t   urface implies that the receiver knows its altitude and, therefore, can use three   satellites and altitude to obtain a navigation solution. This assumption is only partially true, because the altitude information based on "the receiver is on the ocean surface" is not sufficient for the receiver to use as altitude input to its navigation solution.

3.6.2.1 GPS Receiver Computed Altitude

The GPS receiver calculates its position in ECEF coordinates and converts the position (X, Y, and heights above the geoid) into the datum specified by the operator (e.g., ED 50, WGS 84). This is the most accurate altitude information normally available to the receiver and use of four satellites with good geometry will provide the most accurate navigation solution. The receiver can, therefore, be described as using the horizontal part of a 3-D navigation solution during normal 2-D operation.

3.6.2.2 Use of Mean Sea Level Look-Up Table

Most GPS receivers will have a look-up table for mean sea level (MSL) that contains data for a global grid with algorithms for interpolation. The MSL is normally within 0.5 to 1.0 m of the geoid to which the WGS 84 datum is related. If the receiver knows its height above MSL, the conversion to an altitude expressed in ECEF-coordinates can be done. Most receivers operate with a 10- x 10-deg MSL look-up table and MSL values for positions within a 10- x 10-deg square must be found by interpolation. Such a look-up table is not accurate enough to provide good altitude aiding to a shipborne receiver. If the MSL look-up table is expanded to a 1- x 1-deg grid, the maximum altitude error will be on the order of ±4.5 m with an RMS error of 0.6 m, according to the U.S. Defense Mapping Agency (DMA). The use of three satellites and altitude information from the MSL look-up table should be as accurate as using four satellites with good geometry. External altitude is the same as placing one satellite in the center of the earth; using the PDOP and VDOP will therefore be better than PDOP and vertical dilution of precision (VDOP) based on four satellites. The MSL look-up table gives MSL altitude information only and does not take into account the daily variances due to tide, air pressure change, etc. A tide table can be included in the receiver data base and local/daily variances can be corrected by the receiver when it navigates with four satellites with good geometry. The receiver will then have very accurate altitude aiding available when the receiver does not have four satellites with good geometry in view due to the satellite constellation and/or shadowing.

3.6.2.2.1 The Advantage of Using a 1 x 1-deg MSL Look-Up Table

The major advantage of using a 1 x 1-deg MSL look-up table instead of altitude hold is that it can be used by a receiver that starts receiving, having no altitude information in its memory and only three satellites in view. In this case, the receiver does not have to wait for four satellites to be visible to obtain altitude information. This feature is useful for a submarine that does not have sufficient time aiding (no clock or only an uncalibrated atomic clock). Another user might be a ship close to an obstruction that limits the number of visible satellites when the receiver is turned on before departure and no altitude information is stored in the receiver.

3.6.2.2.2 The Disadvantage of Using a 1- x 1-deg MSL Look-Up Table

The disadvantage with the 1- x 1-deg MSL look-up table is that it requires approximately 0.5 Mb additional memory in the receiver. The MSL data can be stored and computed outside the receiver if the necessary memory is not available in the receiver and only the altitude value to the receiver is available.

3.7 Host Vehicle Fixed Lever Arm

The GPS receiver uses lever arm corrections to offset the generated navigation solution from the nominal GPS antenna location to a HV reference point. The lever arm corrections are only used within the GPS receiver when an external attitude aiding source, such as an INS, is available. The navigation solution should then be propagated to the center of the aiding source, as the HV reference point, to simplify computations. A lever arm vector is provided to the GPS receiver as a vector between the GPS antenna and the host vehicle reference point.

If attitude aiding is removed from the GPS receiver, the navigation solution automatically reverts back to the GPS antenna location. More than one set of lever arm corrections may be stored in the GPS receiver. This is useful for installations having more than one INS aiding source or, in the case of big ships, where position and/or velocity information for different locations on board the ship may be of interest. However, only one attitude aiding source should be used by the GPS receiver at any one time. Hence, the propagated navigation solution will only incorporate the one set of lever arm corrections applicable to the particular aiding source that is providing aiding data to the GPS receiver. Should the aiding source be changed, the lever arm corrections will change automatically within the GPS receiver.

3.8 Combination of Satellites and External Aiding

When only four GPS satellites are in view, the GPS receiver uses either the four pseudo range measurements available, regardless of geometry, or use external aiding instead of one or more satellites. Navigation accuracy using the four available satellites may be degraded because of poor satellite geometry and/or high pseudo-range errors. If the receiver has an external atomic clock, a calibrated baro-altimeter, or (if a seaborne receiver) accurate receiver height above MSL available, the receiver should be able to choose between two options:

- o Use the four GPS satellites available, based on PDOP and User Range Accuracy (URA).
- o Use the number of satellites and external aiding that give the best PDOP and URA.

The selection could be done automatically by the GPS receiver or as an operator interaction with the GPS receiver.

4. POSSIBLE GPS INTEGRATIONS

4.1 Introduction

It is possible to operate a GPS receiver in several configurations, according to the aiding data available to the receiver. Implementations can range from a simple, stand-alone GPS receiver to one that is coupled with an INS, an altimeter, and a time-aiding source. The integration level will largely depend on particular mission requirements for navigation data accuracy and availability. Therefore, navigation system managers need to clearly examine the mission requirements before deciding on the GPS integration level necessary to adequately support military operations.

A single GPS receiver may be capable of operating in several distinct modes, depending on the aiding data available. The measurements data that could be incorporated into a GPS Kalman filter and the type of information it is possible to provide in the state vector output are shown in Figure 6. For example, a stand alone GPS receiver implementation would have only pseudo range and delta-range measurements input to the Kalman filter, with outputs of $\underline{p}$, $\underline{v}$, $\underline{a}$, and C_b. At the other extreme, it would be possible to incorporate all measurements listed in Figure 6 through use of an INS, air data computer, and atomic time standard, and to provide all outputs listed for the state vector. The following paragraphs describe some of the more common GPS integrations that are now being used by the military.

The examples given cover a stand-alone GPS receiver, a GPS/INS integration, and a GPS/doppler integration. Baro altimeter aiding is a (relatively simple) option for each of the integrations described.

4.2 GPS Nominal Solution

A GPS receiver, which has broad application, has an internal Kalman filter that can incorporate satellite measurements and external aiding data into the navigation solution. Depending on the quality of aiding data available to the GPS receiver, the Kalman filter will adapt to the appropriate configuration.

4.3 GPS Stand Alone

A GPS receiver can be operated in a stand alone configuration with no position, velocity, acceleration, or attitude aiding available to it. Altitude aiding may be available, but is unnecessary. When operating as a stand alone navigator, the GPS receiver is said to be operating in the position velocity acceleration (PVA) mode. There are 12 states in the Kalman filter when operating in the PVA mode: 3 position corrections, 3 velocity corrections, 3 acceleration corrections, 2 clock corrections, and 1 baro altimeter bias.

Note: ECEF coordinates are used for all position, velocity, and acceleration quantities internal to the GPS receiver. External aiding data is converted to ECEF format before incorporation into the Kalman filter. Normally, a GPS receiver operating in the PVA mode needs to have valid pseudo range and delta range measurements to four satellites to produce a valid navigation solution. Should the number of satellites drop to three, a valid navigation solution is still possible provided altitude aiding is available. If altitude aiding is not available, the GPS receiver will assume the last valid altitude solution and continue to generate new horizontal solutions. Naturally, the accuracy of the horizontal solution will depend on how accurately the GPS receiver is located with respect to the assumed altitude.

4.4 GPS/INS Mode

One of the most common integrations of a GPS receiver in military aircraft is in combination with an INS. The INS provides position, velocity, attitude, and attitude rate data to the GPS receiver. Baro-altitude aiding may also be available, but is not necessary in the INS mode. Attitude aiding is used in conjunction with lever arm corrections to project the GPS navigation solution to the INS gimbal center.

Attitude data is also used in sequential tracking GPS receivers to transform velocity aiding data into ECEF coordinates if necessary. There are 12 states in the Kalman filter in the INS mode: 3 position errors, 3 velocity errors, 3 platform tilt errors, 1 clock corrections, and 1 baro-altimeter bias. Using a GPS receiver under high dynamics and in areas of high jamming threat virtually dictates the use of an INS as an aiding source to ensure adequate GPS performance. Velocity aiding of the code tracking loops ensures continued generation of a valid GPS navigation solution even under high dynamic situations (5 + G). Additionally, velocity aiding of the GPS receiver tracking loops improves the receivers antijam capability. The GPS correction vector for the INS mode may be used to reset INS outputs to remove position, velocity, and tilt errors within the INS. However, INS errors such as accelerometer biases and gyro bias corrections are not estimated by the GPS receiver. If these additional correction terms are required for a particular integration, these correction terms would need to be calculated in a mission computer Kalman filter located external to the GPS receiver and the INS.

4.5 GPS/DRS Mode

It is possible to provide velocity- and attitude-aiding data of lesser quality than an INS to a GPS receiver. This aiding is commonly referred to as the Dead Reckoning System (DRS) mode. Examples of the DRS mode would be use of doppler radar velocities combined with attitude information from an Altitude Heading Reference System (AHRS). Alternate velocity sources, such as from a central air data computer (CADC), could also be used in the DRS mode. Baro-altitude aiding may be available, but is not necessary for the DRS mode. There are nine states in the Kalman filter in the DRS mode: 3 position errors, 3 velocity errors, 2 clock corrections and 1 baro-altimeter bias. The velocity aiding provided in the DRS mode improves code loop tracking performance during state three operations, thus extending the level of jamming that the GPS receiver can sustain under jamming conditions.

4.6 Integration of a GPS Receiver and an Atomic Clock

A GPS receiver should be capable of receiving and providing precise time via a dedicated interface. Precise time input to a GPS receiver can be used for: reducing TTFF, direct P-code acquisition, and use of a clock instead of one satellite in the navigation solution.

The precise time output can be used for several applications, but one important application is to calibrate an external atomic clock with GPS-derived UTC time, which in turn can feed precise time to the GPS receiver later, if necessary.

4.6.1 Reduced Time to First Fix

Time to first fix (TTFF) is defined in U.S. Government GPS program specification SS-US-200 as ".... the elapsed time from the initial demand on a set that has been turned on for longer than 5 minutes to the subsequent display/output of present position and time ...". The TTFF can be reduced somewhat by using a precise time input for initialization of the GPS receiver.

4.6.2 Direct P-code Acquisition

A P-code GPS receiver should always attempt to acquire P code directly before attempting a C/A-code acquisition with handover to P-code tracking. This allows for shorter TTFF for direct P-code acquisition than going via C/A-code acquisition. It also enables use of L_2 only if L_1 is not available, and it avoids the risk of using a "spoofed" C/A-code on L_1. The specified time available for direct P-code acquisition limits the allowable initial offset of the GPS receiver's P-code generator from the incoming P code signal from the satellite.

4.6.3 Aiding of the GPS Receiver Using an Internal or External Time Reference

The problem of limited satellite visibility and/or bad satellite geometry, as described previously, has several solutions. One possible solution is to use an accurate clock to "coast through" the periods when less than four satellites are available. The GPS receiver needs four satellite pseudo-ranges to determine the GPS receiver position (X, Y, and Z) and the GPS receiver clock offset C_b. If only three satellites are available, the GPS receiver can assume that its time reference is correct and treat the three available satellite range measurements as actual ranges instead of as pseudo-ranges. There are two possible ways of using the receiver clock as GPS system time:

- o Use the receiver clock or an external clock without any corrections to its previously calculated bias and drift.
- o Include the internal or external clock frequency and phase offset and drift rate in the GPS receiver Kalman filter and apply corrections to the clock continuously.

If the GPS receiver is shipborne or stationary with known height, then only three satellites are necessary for normal operation. In this case, only two satellites are necessary when an atomic clock is available.

4.6.3.1 Use of an Uncorrected Clock

If the GPS receiver clock or an external clock is used to maintain GPS time during degraded performance and the clock frequency, phase bias, and drift have not been calculated when the GPS receiver starts to navigate with three satellites, the position accuracy will deteriorate rapidly depending on the quality of the clock and the disturbances that the clock experiences (e.g., accelerations and shock). The GPS receiver clock that is normally used has a drift rate of 10^{-10} sec/sec. Thus, position error will grow to 100 m (95 percent) in a few seconds. An uncalibrated atomic clock with a stability of 10^{-13} sec/sec will keep the GPS receiver within a 100-m (95 percent) position error for several minutes.

4.6.3.2 Use of a Kalman Filter Model for Clock Error

The clock used for the receiver time reference can be monitored by the receiver's Kalman filter such that the clock frequency, phase bias, and drift are known when the receiver starts to navigate with only three satellites. This also applies for a two-satellites-only situation if receiver altitude also is known. The corrections for the clock frequency, phase bias, and drift can then be used to maintain accurate GPS system time. What accuracy can be maintained and for how long a time depends on the error models and the disturbances of the clock during the outage period. Examples of such disturbances are: temperature changes, pressure changes, crystal aging, accelerations, and vibrations. The method of using a clock instead of a satellite is not recommended as a permanent solution, but rather to help the GPS receiver operate during short periods when only a limited set of satellites is available.

4.7 Integrated GPS/AHRS

An AHRS is similar to an INS except that only host vehicle attitude, attitude rates, heading, and acceleration information is available (no position). The cost of an AHRS and the quality of the output information is lower than that of a conventional INS. However, the GPS receiver can provide three-dimensional, bounded velocity information to the AHRS that will improve the AHRS output information (e.g., roll, pitch, heading, and attitude rates). An AHRS can be used for short-term fill-in of velocity information if GPS receiver outputs are lost (Figure 7.) Potentially, AHRS velocity could also aid the GPS receiver tracking loops during short periods of jamming or high dynamics. The integrated GPS/AHRS may become a direct replacement for a high-quality INS. The estimated cost of an integrated GPS/AHRS is less than one-half that of a stand-alone INS. Additionally, yearly maintenance costs are expected to be one-third that of an INS.

4.8 Hybrid Systems

Many military applications require continuous navigation information under a variety of stringent operating conditions; unfortunately, GPS alone cannot always fulfill these stringent requirements. Under these circumstances, a hybrid navigation system that combines the outputs of two or more different navigation systems may be used.

Such a hybrid system makes use of the more desirable features of each navigation system to provide continuous navigation data at the highest accuracy level possible, commensurate with the prevailing circumstances. The examples that follow indicate how GPS may be combined with other navigation systems for specific military applications.

4.8.1 Hybrid GPS/INS

For military GPS applications that must operate under high dynamics or in areas of heavy jamming, a good hybrid navigation solution is obtained by combining the long term, high accuracy of a GPS receiver with the short term, high accuracy of INS. Present GPS/INS applications involve the use of an external mission computer that estimates the INS error states through use of a system Kalman filter. In such a system, there may be three separate Kalman filters involved, resident in the GPS receiver, the INS, and mission computer. In hybrid GPS/INS applications it is only necessary to have a single Kalman filter to process the GPS and INS data. Then, the sole Kalman filter processes GPS pseudo-range and delta-range measurements along with the INS outputs of specific force, acceleration, etc. The Kalman filter produces a system navigation solution based on the GPS and INS measurements that generates error correction terms for propagating accurate INS data. A hybrid GPS/INS offers several advantages. First, a hybrid can incorporate a single pseudo-range measurement into the filter to bound the INS information in a single dimension during periods of poor satellite coverage. Thus, greater GPS redundance is available when compared with a normal GPS PVT solution, which normally requires a minimum of three satellites to produce a valid result. A hybrid GPS/INS also eliminates possible integration problems that could arise when cascaded Kalman filters are implemented in integrations having separate GPS, INS, and mission computer components. A hybrid GPS/INS may also offer considerable cost savings to the military when compared with the separate purchase of navigation sensors that would then have to be integrated. A hybrid GPS/INS is selfcontained and will provide a continuous output of navigation data, based on the best measurement data available at the time. A hybrid system also permits the use of a much cheaper, medium-to-low-quality INS without experiencing appreciable degradation in navigation accuracy over short periods of GPS non-availability. This can lead to significant cost savings in the purchase of hybrid GPS/INS equipment.

An alternate method of implementing a hybrid GPS/INS navigation system is to keep the GPS solution separate from the INS solution, and to use the GPS-generated PVT data directly in the INS Kalman filter for error correction. Such an approach provides good system redundance; if the INS portion of the system fails, a valid GPS solution would still be available to the operator. System architecture in this approach maintains adequate separation of GPS and INS functions such that a single point failure will not mean the loss of the total system. Such separation extends to individual power supplies for each of the functional modules within the hybrid system.

4.8.2 Hybrid GPS/Transit/Omega

An interim solution to current limited worldwide GPS satellite coverage is to have a hybrid system using the best data available at the time from each of GPS, Transit, and Omega. Navigation system measurements from all three systems can be used to generate an optimum system solution. Naturally, if four GPS satellites are available and PDOP is within limits, the Transit and Omega measurement data is not required.

In view of the launch problems the U.S. Government has experienced, it will be some time before a full GPS satellite constellation is operational. Thus, a hybrid system such as described here can provide many years of useful navigation data. Some manufacturers are producing hybrid GPS equipment and some companies are also offering retrofit GPS kits for installation in existing Transit/Omega receivers. This hybrid option offers the benefit of allowing current users of Transit and/or Omega to evaluate GPS with respect to normal operational usage.

4.8.3 Hybrid GPS/LORAN-C

Another temporary solution to the current limited GPS satellite coverage is to combine a GPS receiver with a LORAN-C receiver. When GPS satellites are in view, the position information can be used to calibrate the LORAN-C receiver for daily and local effects. When GPS satellites are no longer in view, the calibrated LORAN-C receiver is used as a stand-alone system. When GPS satellites become visible again, the LORAN-C data can be used to initialize the GPS receiver and, therefore, to reduce acquisition time. A combination of LORAN-C and GPS data can also be used to produce a position solution.

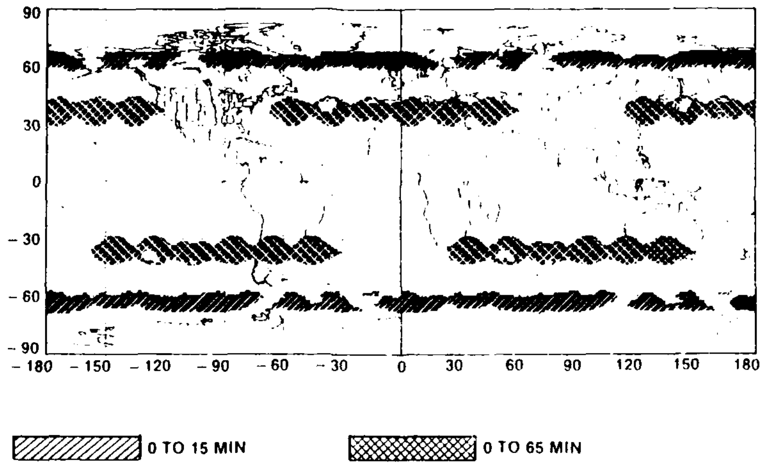


Figure 1. Areas of Degraded Performance
(mask angle - 5°)

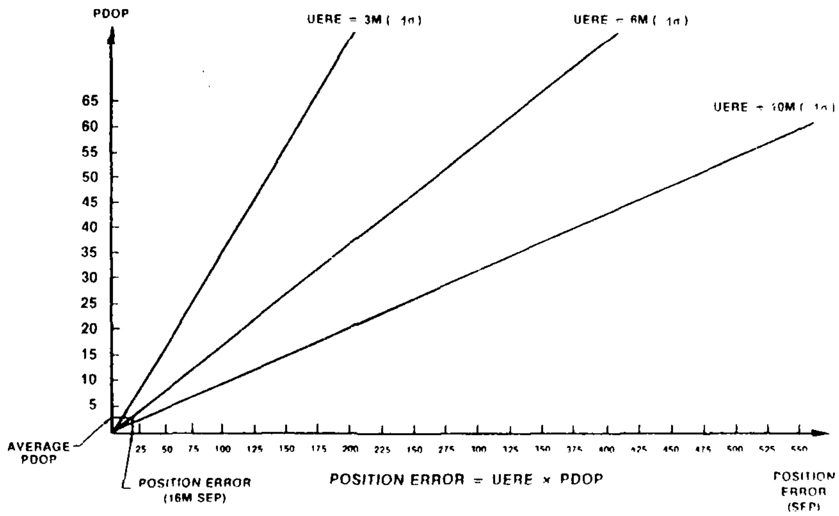


Figure 2. PDOP, Range Accuracy and Position Accuracy Relations

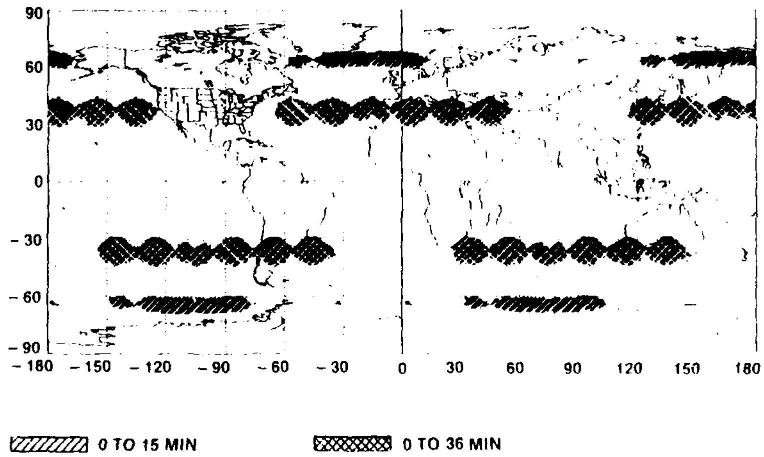


Figure 3. Areas of Degraded Performance where PDOP > 10
(mask angle = 5°)

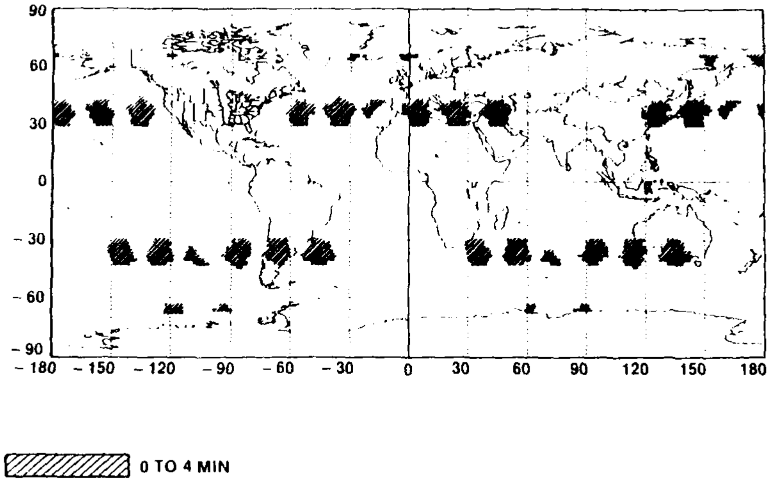


Figure 4. Areas of Degraded Performance where PDOP ≥ 50
(mask angle = 5°)

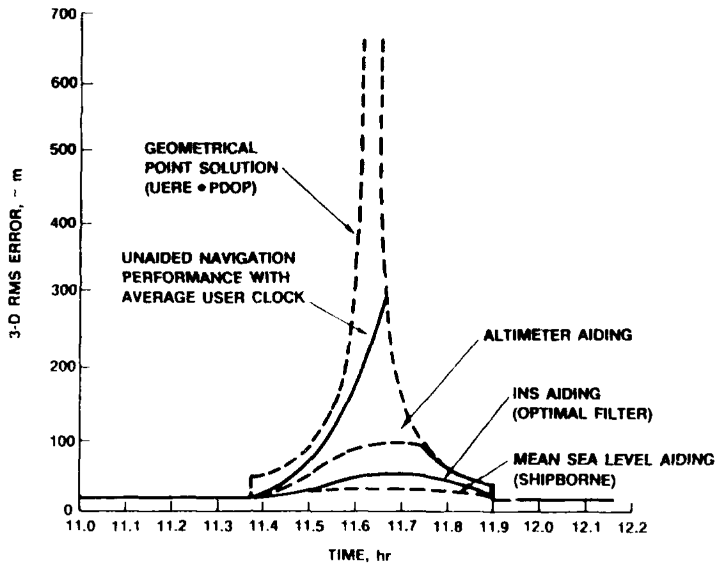


Figure 5. GPS Receiver Performance during a Period of Degraded Performance when Aided

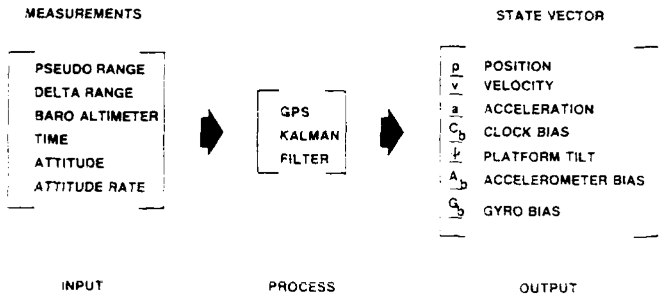


Figure 6. GPS Measurements and Corresponding States

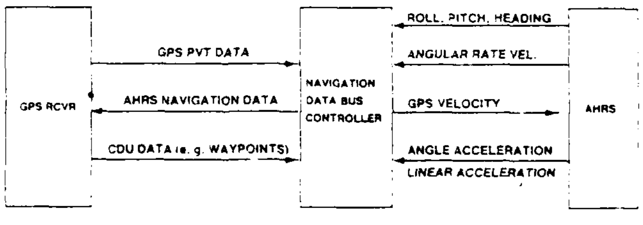


Figure 7. GPS/AHRS Integration

CIVIL AND MILITARY APPLICATIONS OF GPS

HANS J. KUNZE, Lt Cmdr
USAF, Space Division
P.O. Box 92960
Los Angeles, California, USA 90009

SUMMARY

The objectives of this paper are to give an overview of military and civil applications of the Navstar Global Positioning System (GPS), to discuss the resulting benefits for the users, and to present a planned system development that will provide the user community with the GPS information, such as overall system status or satellite data, required for their mission.

The GPS was originally designed by the U.S. Department of Defense (DoD) to satisfy the operational need for a global, highly accurate, and survivable navigation system providing position, velocity, and time. Civilian use of GPS was fostered by the International Maritime Satellite (INMARSAT) Act of 1978, which limits proliferation of federally funded radio navigation systems, and by Senate Resolution 69 and House Resolution 190 (Sep/Oct 1983), which resolved that the GPS should be expedited for the civil sector.

Factor that led the DoD to introduce the GPS into their Armed Forces included its ability to significantly enhance mission performance, coordination of force elements of multiple services/nations, and, ultimately, weapon delivery accuracy. During all phases of the GPS program, tests and evaluations were performed; some of the test results are presented in this paper.

Since the decision was made to make the GPS signal available to the civil user community, a broad range of applications developed. The use of GPS is not only applied to the traditional areas of navigation, i.e. air, marine, and land, but also to the expanding area of space navigation. The system is becoming increasingly important in other applications as well, such as timing and frequency, surveying, and science. This illustrates the diverse range of applications, hence only a brief overview can be given, focusing on a few specific examples.

With the number of military and civil users growing, the need for GPS information is steadily increasing. The DoD has identified a requirement to establish a system for the processing, display, and distribution of GPS status and capability information. Based on requirements from respective user groups, it is planned to distribute GPS data to military and civil users. In August 1987, on request from the DoD, the U.S. Department of Transportation (DoT) agreed to be the focal point for the civil user community and to assume principal oversight and management responsibility for a civil GPS data distribution system.

1. INTRODUCTION

The GPS was developed to satisfy the military operational need for a precise navigation system providing global, highly accurate, position and velocity data in three dimensions and time, and to obviate proliferation of the navigation systems presently required in performance of various missions.

The system was designed such that it eliminates the shortcomings of existing navigation systems, e.g., short range, line-of-sight, atmospheric influences, reliability and survivability, degradation of accuracy over time (Inertial Navigation System - INS), etc. Hence, the major benefits of GPS are design inherent, i.e., a highly survivable, space-based system providing continuous global coverage, a common reference grid, all weather and passive operation, jamming resistance of signal, and an unlimited number of users. These advantages potentially allow GPS users to eliminate the cost for logistic support of ground based radio navigation systems, except for U.S. funding of the GPS control segment. Integration of the GPS and a self-contained system, such as an INS, could improve the jamming resistance and the navigation accuracy of this integrated system and achieve substantial cost savings (up to 50 percent) compared to a high quality INS.

In the Federal Radio-Navigation Plan (FRP), which delineates policies and plans for federally-provided radio navigation systems, the U.S. Government has identified the need to consolidate and reduce the number of radio navigation systems operated by either the U.S. Department of Defense or the U.S. Department of Transportation. The 1986 FRP states that the DoD plans to phase out military use of LORAN-C and OMEGA by 1994 and VOR/DME and Land-Based TACAN by 1997, to replace ILS with MLS, and to discontinue operation of Transit by 1996 (replacing it with GPS). The DoT calls for a transition period of 15 years to discontinue federal operation of VOR/DME, or OMEGA after a decision has been made based on some basic ground rules: Phase-out would be keyed to resolution of the GPS accuracy, coverage, integrity, and financial issues after GPS is fully operational; the GPS would have to meet civil air, marine, and land needs currently met by existing systems, including technical attributes; the GPS civil-user equipment would have to be available at prices economically acceptable to users of existing systems. However, the operation of overseas LORAN-C will be discontinued by 1994, because the U.S. Coast Guard operates these systems based on military (U.S. Navy) requirements.

2. MILITARY APPLICATIONS OF GPS

2.1 Force Enhancements

For the military, the principal influence on selection of a particular navigation sensor or mix of navigation sensors is operational need. Effective execution of most military operations is directly dependent upon the precision of the utilized time and/or navigation information, because these operations encompass functions such as enroute navigation, target acquisition, and weapon delivery. The GPS has the potential to fulfill all military navigational requirements, except for precision approach and harbor entry, and will eventually replace all existing radio navigation systems. The GPS can replace or supplement self-contained systems, e.g., AHRS, INS. The decision will have to be made on a case by case basis to trade off, for example, accuracy and redundancy requirements versus cost. Force effectiveness (i.e., effective execution of missions) is improved greatly by the highly accurate position, velocity, and time information provided by the GPS, while the GPS is present; but it is not degraded when the GPS signal is lost. The GPS acts as a force multiplier, because, in addition to the individual missions, combined operations with force elements of other services/nations can be executed with a flexibility and precision not possible before. The main factors contributing to these increased capabilities are the common reference grid and the precise time provided by the GPS.

Another very important feature of the GPS is that the system supports precise and standard positioning service (PPS and SPS). Access to the PPS (with its navigation accuracy of 16-m SEP) will be restricted primarily to military users, by encrypting the PPS signal. The SPS will be made continuously available to all users, and will provide 100-m, 2-drms navigation accuracy, or better. However, the navigation signal can be selectively corrupted to such a degree that the SPS signal is no longer useful, and, thus, a potential adversary can not employ GPS to his own benefit.

2.1.1 Command, Control, Communications and Intelligence (C^3I)

The designation C^3I refers to a set of functions relating to the collection, processing, and dissemination of battlefield information, such as precise location of friendly and hostile units, together with decision making and dissemination of those decisions. The objective of C^3I is to deploy/employ own force in a timely and effective manner.

NATO developed a concept called follow on force attack, which requires C^3I systems capable of all weather surveillance moving and fixed target acquisition and processing and fusing of information in near real-time, to provide the means to locate an enemy follow on force as early as possible, determine its route of advance, and direct an attack against it in order to cause delay, disruption, and attrition. This section will focus on the potential benefits of GPS for these required C^3I systems.

There are basically three elements to multiple sensor data fusion: to transform the target location given in onboard coordinates to precise geodetic coordinates that may be converted directly into target assignments at the local level; to correlate target reports from multiple sensors; and to reduce target location uncertainty for planning and resource allocation on the theatre level. A study has shown that the GPS, integrated with accurate digital surveillance sensors (electro-optical/infrared sensors or synthetic aperture radar (SAR) systems), makes the new concept of direct target location possible. The benefits are that direct target location, based on GPS, could provide target locations in precise geodetic coordinates in near-real-time with an accuracy of better than 30 m (SEP). With any other navigation reference, except for a ground based system, the accuracy would degrade by at least a factor of three.

The correlation of target reports from multiple sources/sensors (fusion score), as a function of target density, is presented in Figure 1. With the GPS, near perfect fusion scores are achieved, i.e., the target is definitely identified, while the data fusion process breaks down at relatively low target densities with unaided inertial navigation.

The accuracy of target locations can be enhanced by data fusion to as good as 10 m (SEP) with the GPS.

The study concludes that "target location accuracy in a multiple sensor environment is significantly better with the GPS than with any other navigational system studied," and that "GPS could significantly simplify the surveillance process by providing a worldwide, common reference system with relatively minimal logistical support requirement."

The United States is currently developing the Joint Surveillance Tactical Airborne Radar System (JSTARS), which will utilize GPS and will be fielded around 1995. The target acquisition function of JSTARS is based on the integration of GPS, with SAR and digital maps, and the onboard antijamming communication system (the Joint Tactical Information Distribution System), utilizes the precise time output of the GPS receiver to calibrate the onboard clock and, therefore, maintain a stable reference time to be used for frequency hopping techniques (for example).

2.1.2 Tactical Operations

The applications and benefits of a precise navigation system like the GPS cut across all phases of tactical operations. Especially for combined operations of various elements, such as close air support, maritime patrol, air intercept, and amphibious operations, the precise time for synchronization of missions and the common reference grid multiply the benefits the GPS provides for each element. Therefore, and for interoperability considerations, it would be of significant value to supply all services (Army, Air Force, and Navy) throughout the NATO countries with the benefits of the GPS.

Tactical operations can be broken up into three areas: air, naval, and land applications. The potential impact of the GPS on various tactical missions in these areas will be qualitatively analyzed. The functions that must be performed to accomplish these missions are identified, and those whose performance is dependent upon the accuracy of position data are marked for potential applications of GPS. Some field test results from the concept validation phase (Phase I) and full-scale development phase (Phase II) will be presented in section 2.2.

2.1.2.1 Air Operation Applications

Among the various tactical air missions impacted by the GPS are counter air, air interdiction, close air support (CAS), reconnaissance, search and rescue, all weather air drop and transport, and maritime patrol. Effectiveness of the different functions encompassed by these missions will be enhanced by the GPS, such as launch and recovery, enroute navigation, rendezvous, target acquisition, and weapon delivery (see Table 1).

Table 1.

[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	•	•	•		•	•	•	•	•	•	•	•	•	•
[illegible]	•	•	•		•	•	•	•	•	•	•	•	•	•
[illegible]	•	•	•		•	•	•	•	•	•	•	•	•	•
[illegible]	•	•	•		•	•	•	•	•	•	•	•	•	•
[illegible]	•	•		•	•	•	•	•		•	•	•	•	•
[illegible]	•	•	•		•	•	•	•	•	•	•	•	•	•
[illegible]	•	•	•		•	•	•	•					•	•
[illegible]	•	•			•	•	•	•					•	•
[illegible]	•	•												•
[illegible]	•	•			•	•	•	•	•				•	•
[illegible]	•	•		•		•								•
[illegible]	•	•	•		•	•	•	•					•	•
[illegible]	•	•		•		•			•	•	•	•		•
[illegible]	•	•		•		•			•					•
[illegible]	•	•			•	•	•	•	•					•
[illegible]	•	•	•	•			•	•	•				•	•
[illegible]	•	•	•	•		•			•					•
[illegible]	•	•				•	•	•	•				•	•

Launch and Recovery

The GPS can provide INS inflight alignment, which allows fast response takeoff ("scramble"), for air defense, search and rescue, or attack alert; the GPS reduces dependence on ground control/equipment, such as TACAN and VOR/DME; and recovery operations would be unaffected by combat damage to ground navigation aids, because the GPS allows nonprecision type approaches. This makes worldwide bare base operations possible. The requirements of precision approaches for bare-base operations could be met with mobile equipment, possibly a combination of MLS and differential GPS. Overall, the GPS could improve ground survivability of the forces substantially.

Enroute Navigation

Enhanced navigation accuracy (see Fig. 2) provides the flexibility to change routing in flight and allows the planning of optimum routes with respect to, e.g., terrain and defense circumvention. This is most important for missions such as air interdiction and reconnaissance where terrain following flight with aircraft and remotely piloted vehicles is performed. A study showed that integrating the GPS with a digital map (height accuracy approximately 8 m) could make passive terrain following flight with a ground clearance of 50 m possible. Digital maps with this kind of accuracy are currently not available; however, by utilizing the GPS for photo mapping (for example) and establishing a network of coordinated points, digital data bases and, hence, digital maps with the required accuracy could be generated.

Rendezvous

The principal factors in performing a rendezvous (e.g., forward air controller in CAS operations, search and rescue, in-flight refueling) are synchronized time and a common reference grid. The GPS provides these factors and it allows rendezvous under all weather conditions and passively or, at least, with only short periods of radar and/or radio emission. The rendezvous points could be fixed or, as in the case of in-flight refueling, moving. The GPS would decrease the risk, and speed up the execution, of rendezvous operations; hence, the number of aerial refueling tankers required could be reduced (for example).

Target Acquisition

The probability of acquiring a target depends on a number of inputs: flight profile, terrain, target location uncertainty, aircraft navigation system uncertainty, etc. The graph in Figure 3 illustrates the probability of acquiring a small target, such as a SAM site, during a single pass by an aircraft approaching the target at a 600-ft altitude and equipped with a high resolution, forward looking infrared sensor. The target location is assumed to be precisely known and the probability of acquisition is shown as a function of aircraft navigation CEP. Thus, assuring that the aircraft is headed directly toward the reported target location, the probability of acquiring the target improves as the aircraft's navigation system more precisely reports the aircraft's actual position. This example shows clearly that the GPS increases probability of target acquisition, reduces aircraft exposure (single pass), and increases target approach flexibility. These characteristics are very important for missions such as CAS, air interdiction, and maritime patrol.

Weapon Delivery

Weapon delivery accuracy directly depends on the navigation accuracy of the weapon platform. Increased accuracy and the common reference grid of the GPS reduce the probability of hitting friendly forces in CAS operations, allow accurate delivery of air minefields, and provide precise coordinate bombing capability, which improves adverse weather weapon delivery (see Fig. 4).

Other examples of GPS benefits are initialization and midcourse guidance of tactical standoff missiles that are used for interdiction missions. Comparison between the GPS and map correlation techniques (for example) shows that the high accuracy of the GPS improves mission effectiveness, with respect to target destruction, by approximately 30 percent. Additionally, the GPS increases operational flexibility, because the missile is not constrained to a predetermined flight corridor and the GPS has worldwide applicability without preparation of source data and maps even over terrain unsuitable for correlation guidance (e.g., over water). An example is the Navy's Standoff Land Missile (SLAM) program. The SLAM is a derivative of the Harpoon missile and will utilize the GPS to satisfy the Navy's need for a precision strike, land attack, air-launched, tactical weapon.

Better weapon delivery reduces the number of sorties and the amount of munitions required to successfully complete a mission. Standoff ranges can be increased and time over target area (exposure to enemy threat) can be reduced and, therefore, aircraft survivability improves.

2.1.2.2 Naval Operational Applications

Naval operations include some air operations similar to those discussed above, as well as mine warfare, surface warfare, amphibious operations, and subsurface warfare. These air operations will not be discussed here any further, although it is recognized that they are very important to certain aspects of naval operations. The emphasis will be placed on additional GPS benefits for naval operations (see Table 2).

Table 2.

GRID-LOCK / LONG RANGE TARGETING / COORDINATED OPS / COVERT RENDEZVOUS / MINE FIELD OPS

NAVAL OPERATIONS \ FUNCTION	Navigation	Own force location	Target location	OTH-T	ASW SOW	Power projection	Vectoring & handoff	Search ops	Aircraft/sub ASW	Task group defense	Off board sensing	Station keeping	Force assembly	Carrier landings	Aerial refueling	Covert landings	Placement	Sweeping	Penetration
BARRIER OPS	•	•	•	•	•		•	•	•		•	•		•			•		
SLBM DEFENSE	•	•	•	•	•		•	•			•	•		•			•		
SURVEILLANCE & RECONN	•	•	•				•	•			•	•			•		•	•	•
OCEAN AREA OPS	•	•	•	•	•		•	•	•		•	•		•	•		•	•	•
NAVAL FORCE PROTECTION	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•				
SHIPPING PROTECTION	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•				
OFFSHORE RESOURCE PROT	•	•	•	•	•		•	•	•		•				•		•		
OPERATIONAL LOGISTICS	•	•						•					•		•	•		•	•
MINE WARFARE	•	•	•								•	•			•		•	•	•
AMPHIBIOUS OPS	•	•	•	•	•	•				•		•	•			•	•	•	•
INSHORE WARFARE	•	•	•	•		•	•	•		•	•	•	•	•	•	•	•	•	•
STRIKES ON LAND TARGETS	•	•	•	•		•						•	•		•	•			

Gridlock

One of the most intriguing aspects of GPS with respect to tactical naval missions, is the common reference grid all ships and boats could utilize, especially since many of them report own position and enemy target locations via data link to their naval headquarters. The good navigation accuracy of each unit improves significantly the data correlation process and, therefore, the overall status picture. The result is that gridlock will be achieved with GPS precision and, hence, subsequent operations can be conducted more effectively.

Some examples for improved navigation, own force, and target location are: convoys under EMCON, at night, or in adverse weather could maintain their station with no added risk; submarines could minimize exposure to enemy threat, because GPS provides updates of the INS, aided with an atomic clock, within seconds after reception of GPS signals; over-the-horizon-targeting (OTHT); and target assignments for naval units without the need to utilize on-board sensors.

Long-Range Targeting

GPS can benefit surface warfare by, e.g., providing precise gridlock between a ship equipped with a stand-off antiship missile and its helicopter employed as a stand-off sensor for OTHT operations, and, thus, increasing kill probability and survivability. The same benefits apply for antisubmarine stand-off weapons (ASW SOW), where the second stage of the missile is a torpedo that performs the terminal phase.

Coordinated Operations

Achieving precise gridlock is also the key to more effective coordinated ASW operations. Patrol aircraft, helicopter, ships and submarines would report their own positions and the line-of-bearing to the target in a common reference grid. This would allow precise target location, assignment to a unit, and, finally, delivery of ASW weapons.

Other possible scenarios include: the coordination of search and rescue operations; amphibious assault, supported by CAS; and surface ship shore bombardment.

Covert Rendezvous

The precise navigation information available to each naval unit and the subsequent achievable gridlock precision makes the assembly of amphibious assault forces under EMCON, at night, and in adverse weather possible. Another example would be the rendezvous of special forces with their "base", such as a submarine.

Mine Field Operations

Mine warfare is of vital interest to the NATO nations, especially to the European nations. The utilization of GPS can result in significant advantages against the majority of threat vessels that must pass through the limited number of straits and intra-ocean passages. Mining operations must be accurately planned and charted to protect friendly vessels. Channels through minefields could be unmarked and tightened up, if GPS receivers are available to these vessels. Minesweeping and other mine countermeasures also depend on accurate position data regarding one's own mines, minefields, and swept areas. The use of GPS on-board receivers by ships and aircraft conducting mining and mine countermeasure operations will greatly enhance the speed, precision, and safety of such missions.

2.1.2.3 Land Operations Applications

While precision of navigation is not always a requirement for land operations, ease and speed of obtaining this information certainly is (see Table 3). Some of the aspects that are important to these applications, such as midcourse guidance of missiles or surveillance systems, have been discussed before.

Table 3.

GROUND FORCE SYSTEMS \ GROUND FORCE FUNCTIONS		NAVIGATION	OWN FORCE LOCATION	TARGET LOCATION	TARGET ACQUISITION	FIRE CONTROL	MIDCOURSE GUIDANCE	TERMINAL GUIDANCE & TARGET DESTRUCTION	RECOVERY
DIRECT FIRE WEAPONS	TROOPS WITH LIGHT GUNS	•	•	•					
	TANKS	•	•	•		•			
INDIRECT FIRE WEAPONS	MORTARS	•	•	•		•			
	ARTILLERY	•	•	•		•			
	ROCKET LAUNCHERS	•	•	•		•			
	S/S MISSILES	•	•		•	•	•	•	
	MINE LAYERS	•	•	•		•			
AIR DEFENCE WEAPONS	ANTI-AIRCRAFT GUNS	•	•	•	•				
	SAMS	•	•	•	•	•	•	•	
SURVEILLANCE SYSTEMS	RADARS & C^3			•					
	LT OBS A/C & HELICOPTERS	•		•	•				•
	RPVS	•		•	•		•	•	•

Direct Fire Weapons

In this case position accuracy does not directly impact kill effectiveness. However, troops and tanks have to navigate as well, and the more this function is automated the less it detracts from the primary mission.

Currently DARPA is sponsoring a development called Small Unit Navigation System (DARPA/SUNS), which will eventually lead to a handheld GPS receiver of the size of a cigarette package ("Virginia Slim"). This would enable ground forces to equip their troops with GPS receivers down to the squadron leader level.

The utilization of GPS can help to accomplish the tasks of locating enemy forces and accurately determine the forward edge of the battle area, thus making CAS and artillery support more effective. The forward air controller in CAS operations will be able to perform his mission more effectively, because all support units will be reporting in a common grid, therefore he will not be burdened with coordinate convergence.

Ranger/special forces in particular will benefit from using GPS receivers, due to their unique mission requirements. A good indication for this is, that the U.S. is developing a special GPS receiver called Parachute Offset Navigation System for high-altitude high-open operations.

Indirect Fire Weapons

Accurate positioning and target location is of primary importance to indirect fire systems such as artillery, rocket launchers, mortars, and surface-to-surface missiles, since it directly effects the probability of hitting the target. Fire positions can be surveyed with a GPS receiver within minutes and without preparation, which translates into more flexibility with respect to position changes. Targets identified through a C^3I system can be acquired much faster and the first round is fired already for effect. A comparison of GPS versus map positioning showed that GPS enhances artillery effectiveness by a factor of 3.

Air Defense Weapons

The potential of using GPS is very limited where intercept or fire control radars directly view both the intercept weapon and its target. However, where the target has to be handed over from a tracking radar to the intercept weapon system such as an aircraft the use of GPS equipment will improve the accuracy of intercept, thus the probability of target kill.

2.1.3 Strategical Operations

The previous paragraph discussed how the GPS improves the effectiveness of various tactical mission functions from launch and recovery to weapon delivery. Strategic weapons can be delivered by aircraft, naval craft, or mobile/fixed land-based launcher; hence, all previously described GPS benefits to the particular weapon platform apply and will not be discussed again.

The principal impact of the GPS on strategic missions is the enhancement of weapon delivery accuracy, especially for sea and air launched cruise missiles (SLCM/ALCM). Improved navigation accuracy of the weapon platforms translates directly into better initialization, thereby reducing landfall trajectory dispersions of SLCMs/ALCMs launched at sea and, thus, the number of warheads required for mission success. Even greater standoff mission performance benefits could be realized if the GPS were also used for midcourse guidance, e.g., integrated with a low-cost INS and digital data base maps. The gained accuracy could enable the use of conventional weapons for many missions previously assigned to nuclear warheads.

The major benefits for mobile land-based missile launches are that the GPS can provide rapid site position survey and precise initialization data for the missile navigation system. This allows the missile to be in a relatively "dormant" state, since ground alignment may be reduced or eliminated, and still have a rapid response time.

The GPS provides the potential for significant improvements in mission flexibility because it makes mission modification, such as route changes or retargeting during missions, technically supportable.

2.2 Operational Demonstration Results

The GPS program is split up into three phases: The concept validation phase (Phase I), the full-scale development phase (Phase II), and the production phase (Phase III). A program overview is presented in Figure 5, showing the test and evaluation time frames with respect to the program Phases I through III and the overall system development.

The Phase I DT&E overall objectives were to validate the concept and design, and to demonstrate the military value. This led to 22 major field test objectives (see Table 4), some of which will be discussed in this section.

Table 4.

NAVIGATION ACCURACY	THREAT PERFORMANCE
POSITION VELOCITY EFFECTS OF DYNAMICS	JAMMING RESISTANCE SELECTING AVAILABILITY
DEMONSTRATIONS OF MILITARY VALUE	ENVIRONMENTAL EFFECTS
PRECISION WEAPON DELIVERY LANDING APPROACH RENDEZVOUS PHOTOMAPPING	PROPELLER AND ROTOR MODULATION FOILAGE ATTENUATION MULTIPATH REJECTION IONOSPHERIC AND TROPOSPHERIC CORRECTION
MAP-OF-EARTH OPERATIONS STATIC POSITIONING COMBINED OPERATIONS CROSS-COUNTRY SHIPBOARD OPERATIONS	SYSTEMS CHARACTERISTICS SATELLITE CLOCK AND EPHEMERIS ACCURACY ACQUISITION AND RE-ACQUISITION TIME TIME TRANSFER SIGNAL LEVELS AND SIGNAL STRUCTURE

Over 600 missions were accomplished comprising of 11 different types of host venicles (jeep, UH-1H helicopter, F-4 aircraft, etc.), and 9 types of user equipment. The Phase I test results demonstrated that the predicted navigation accuracy of 16 m (SEP) for position and 0.1 m/s for velocity could be provided (see Figure 6). Dependent upon test vehicle environment (high, medium, or low dynamic), it proved best to utilize a continuous receiver, which tracks four satellites simultaneously, for high dynamics; a two-channel sequential receiver for medium dynamics; and a one-channel sequential receiver for low dynamics or stationary applications, mainly because of cost versus performance trade-offs.

The conclusions reached with respect to the demonstration of the military value of GPS were: a) it is a force multiplier, because of its proven accuracy, b) force effectiveness is greatly improved while GPS is present and not degraded when it is lost, c) user survivability is enhanced because the system is passive, and d) it is not a replacement, but a highly survivable supplement to self-contained systems.

For example, 25 landing approaches with three different test vehicles were flown. The decision height waypoint and glide-slope course/angle entered into the GPS receiver, and by executing the steering commands of the pilot steering display the aircraft was taken down to the decision height waypoint without visual reference to the ground. Results indicated that horizontal deviations of a GPS approach were well within the ILS beam at 200 ft decision height, while vertical ILS limitations were exceeded by about one-third of the approaches. This GPS system inherent error in calculating vertical position (altitude) could be eliminated by utilizing differential GPS. The results demonstrate that GPS has the required accuracy to guide an aircraft through a nonprecision approach and that, with the implementation of differential GPS, for example, even precision approaches are possible.

Another example is precision weapon delivery. The mission for the F-4 aircraft, was to perform coordinate (or blind) homing using low-drag MK 82 bombs with level and toss delivery methods. About 100 bombs have been dropped with results indicating that GPS may cause a revolution in coordinate bombing accuracies.

In Phase II development and initial operational test and evaluation (DT&E and IOT&E) were executed. The main objective of the DT&E phase, was to demonstrate the utility of the GPS prototype user equipment and the purpose of the IOT&E phase was to evaluate the operational effectiveness and suitability of GPS user equipment. Nine types of vehicles were used for IOT&E, i.e., the USAF F-16 fighter and B-52D bomber, the U.S. Navy A-6E attack aircraft, P-3B submarine hunter, carrier and submarine, the U.S. Army UH-60 attack helicopter, M-60 tank and manpack/vehicular.

The test results demonstrated again that GPS provides accurate position/navigation information as specified, and that it has the potential to enhance combat capability (e.g., precise, passive and low level coordinate bombing (see Figure 7), where properly integrated. The IOT&E showed clearly the importance of proper integration, because integration problems prevented realizing the full potential for certain host vehicles. Other important factors of the IOT&E were reliability and maintainability tests, which resulted in a number of product improvement requirements for Phase III.

2.3 Range Applications

In the summer of 1981 a tri-service committee was formed to evaluate the potential of a GPS-based range-tracking system. Twenty two tri-service ranges were surveyed, and after 1.5 years it was concluded that GPS range equipment satisfies 95 percent of the requirements. It was unanimously agreed that the application of GPS to range tracking has cost and technical advantages over existing range tracking techniques (see Fig. 8).

In January 1983 a Range Applications Joint Program Office (RAJPO) was established with the purpose of developing a family of GPS equipment tailored to the needs of the range tracking community. Several components, such as low and high dynamic instrumentation sets (LDIS/HDIS), a pod with HDIS, inertial reference unit, and data link, have been developed for this GPS receiver-based tracking system (see Fig. 9). The RAJPO chose this approach because this system satisfies the unique requirements of the ranges for time, space, and position information (TSPI). The advantages are that TSPI is available for use by the vehicle; only a narrow band data link (downlink) is necessary for telemetering data to the master station and, thus, the number of vehicles that can be tracked simultaneously is essentially unlimited; and TSPI can be recorded on board when the vehicle is not in the master station line of sight. Clearly, these advantages out weigh the disadvantages for this range tracking application, e.g., loss of signal track due to extremely high vehicle dynamics (jerk) (i.e., a portion of the TSPI may be unrecoverable).

The U.S. Navy took a different approach for their tracking system for the Trident Missile Program. Due to their stringent size and weight, as well as time to first fix and (TTFF) accuracy requirements (range safety), a GPS translator-based tracking system was developed (see Fig. 10). The advantages are that a translator is a much less complex device than a GPS receiver, which reduces the cost, size, and weight of this component up to an order of magnitude; it shifts the computational capability to a ground base, thus allowing this capability to increase, and enabling TTFF of typically less than 5 sec for range safety applications; it allows mission replay; and it gives the system inherent differential GPS accuracy. The requirements of the Trident Mission Program are such that the existing disadvantages are not a factor; these disadvantages are: only a limited number of targets can be tracked simultaneously because of translator retransmission bandwidth requirements, TSPI is not directly available for use by the vehicle, and recording on-board the vehicle is typically impossible--making a line-of-sight relationship to the master station mandatory.

For the same reasons, the GPS translator-based tracking system was chosen for the exoatmospheric reentry vehicle intercepter and the space-based intercepter programs, both part of the Strategic Defense Initiative (SDI) Program. The purpose of these programs is to develop equipment to track high-dynamic kinetic energy weapons and their targets in order to test and evaluate these new weapons.

3. CIVIL APPLICATIONS OF GPS

Guidelines for civilian use of GPS were prepared by the International Maritime Satellite (INMARSAT) Act of 1978, which governs proliferation of federally funded radio navigation systems, and, following the KAL-007 disaster by Senate Resolution 69 the (26 September 1983) and House Resolution 190 (19 October 1983), which resolved that the GPS is to be expedited for use in the civilian section. The FRP reflects these congressional actions, and states that the standard positioning signal (SPS) accuracy will be 100 m (2 drms) in order to meet all published civil navigation requirements other than precision aircraft landing and maritime harbor entry, and that the SPS will be made available free of charge, to civil, commercial, and other users on an international basis at the highest level of accuracy consistent with U.S. national security interests. Limited civil use of the precise positioning service (PPS) may be authorized, if it is in the national interests of the United States; if adequate security can be provided; and if no other source of accuracy is available.

Even with the reduced accuracy of the SPS, potential civil applications of GPS are practically unlimited because of reduced cost of operation through efficient navigation and improved management of resources in a multitude of applications. The larger number of users will bring the cost of GPS receivers down to the point that recreational boaters, or even backpackers, could easily afford it.

3.1 Navigation

The GPS will be used for all of the traditional air, sea, and land applications, as well as for the still developing area of navigation in space. However, it will have a greatly expanded role compared to existing systems due to its high accuracy and its continuous, global coverage.

In the area of air navigation, the GPS has the potential to be certified by the FAA for all phases of flight as the sole means of navigation, except for precision approach, pending resolution of issues related to coverage and integrity. A possible solution to certification of the GPS for precision approaches could be a differential GPS. The current DoD plan to increase the satellite constellation from 18 operational satellites (plus three active spares) to a full 24-operational-satellite constellation by 1995 should resolve the above mentioned issues of coverage and integrity. The GPS could have a major impact on the Air Traffic Control (ATC) System with respect to, among other things, routing and collision avoidance functions. The present fixed-route structure can cause congestion during heavy traffic periods or in adverse weather. During the enroute/ terminal phase, GPS accuracy would allow implementation of a feature such as "dynamic" airways allocation and/or reduction of current separation requirements, and the eradication of congestion. Another benefit would be the capability for overwater ATC. Both benefits would result in reduced flight time and fuel consumption and, hence, reduced cost of operation, as well as increased safety of flight. Several approaches are avail able for utilizing the GPS in a collision avoidance function. One such approach would be to assign each aircraft a time slot in a time division multiple access (TDMA) bus. Each aircraft would then automatically report his position and velocity in his time slot and monitor the bus for other aircraft in the immediate area. The TDMA technique requires accurate synchronization and, therefore, an accurate time reference--something the GPS delivers as a by-product (see section 3.2).

Some of the applications of the GPS in maritime navigation are: worldwide navigation capability would reduce port-to-port transit times with corresponding economical advantages; safety of navigation enhancement in coastal waters, because aids-to-navigation would be positioned with higher accuracy and repeatability and, especially in adverse weather, ships in two way shipways could hold their course more accurately; and search-and-rescue operations. Another application of the GPS could be, development of an integrated GPS and radar collision avoidance system.

Beyond the traditional applications in land navigation the GPS will even bring satellite navigation to private automobiles. Auto manufacturers have plans to incorporate the GPS into their vehicles probably to be introduced in the early 1990s. Automated vehicle monitoring/locating (AVM/AVL) for commercial truck and railroad companies, as well as police, fire, and emergency vehicles in another major area of GPS application. The AVM/AVL capability enables fleet operators and Government agencies, for example, to keep track of vehicles carrying valuable or dangerous cargo, and it allows commercial users to optimize their fleet management, hence reducing the time and cost involved in deploying their vehicles. A development effort sponsored by DARPA has the goal of reducing the GPS receiver to a handheld unit the size of a cigarette box ("Virginia Slim"). It is quite possible, that hikers, for instance, could use such a GPS receiver to cross remote areas safely.

Although the GPS was developed with air, land, and marine user requirements as the main drivers, studies and experiments have shown that the system will provide excellent results for on-board navigation in spacecraft operating in low earth orbits (LEO), i.e., lower than 1000 nmi. The benefits of onboard orbit determination are that it could reduce the cost associated with previously required ground-based tracking stations, and that it would allow a more flexible approach to monitoring and controlling orbital operations. This could become a big factor in the future because, with the advent of the space station, a large increase in the volume of space traffic in low earth orbit is anticipated. The Landsat-4 and -5 satellites (nominal altitude: 705 km) carry the GPS receiver as a onboard navigation tool. Experiments conducted with these satellites compared the GPS receiver derived ephemerides with definitive ephemerides, which were independently derived from ground station tracking data. Results indicated that errors were consistently less than 50 m position error and 6 cm/sec velocity error, utilizing the P code and during periods of good GPS satellite vehicle visibility.

The GPS performance for orbit determination and rendezvous in LEO was evaluated in a study by analysis and simulation. The results indicated that the GPS can be used effectively for both purposes. It was demonstrated in the study that two nearby, cooperative GPS users can obtain relative positioning accuracies of 1.8 m for P code, and 20 m for course acquisition (C/A) code (1-sigma) in each axis, which is superior to each user's absolute accuracy. It was also shown that information provided by an orbiting differential GPS reference station can be used to improve the absolute navigation performance of other GPS user systems in the vicinity, e.g., from 35 to 14 m for a typical C/A code user.

NASA is currently evaluating the feasibility of integrating GPS user equipment into the Space Shuttle's navigation system, to optimize docking maneuvers with the planned space station, by utilizing the relative positioning technique.

3.2 Precise Time

The two main civil applications of precise time provided by the GPS are time transfer and synchronization of communication equipment/systems.

The traditional method of distributing time to remote stations is to physically transport a well calibrated atomic clock to the distant location and then to synchronize the local reference clock to this portable clock. The GPS benefits are that it allows time transfer without transportation of equipment, it is faster, and it can provide even more accurate clock calibration. The U.S. Naval Observatory determines the difference between GPS time and universal time coordinated (UTC) and transmits this difference to the GPS master control station (MCS) for all satellites. The MCS models each satellite clock difference from GPS time, as well as GPS time to UTC time, and then directs the satellites to transmit this information to the user in the navigation message.

There are basically three methods by which the GPS can be utilized for time transfer: nonsimultaneous, simultaneous (or "common view"), and direct measurements; all three methods require that the satellite have an onboard clock and transmit a time-marked signal. The direct measurement method is used quite often because it requires no coordination between stations. Any station can make measurements from the satellites and obtain the difference between its own clock and UTC (for example) whenever a GPS satellite is in view. The accuracy could be as good as 36 ns assuming that the measurements were smoothed, station position is known accurately, and selective availability (SA) is not switched on. The common-view technique or simultaneous measurement method requires that the satellite is in view of both stations at the same time. The advantages are that this method reduces significantly the dependency on satellite clock stability and therefore could reduce the effect of SA. Experiments performed by the National Bureau of Standards (NBS) and U.S. Naval Observatory (USNO) with the Physikalisch Technische Bundesanstalt (PTB) (W-Germany), Tokyo Astronomical Observatory (TAO) and Radio Research Lab (RRL) (Japan), and Paris Observatory (PO) (France) over a period of three years utilized GPS and the common-view technique. The experiments were carried out by utilizing one satellite with a ground track that covered all participating institutions, i.e. NBS, PTB, TAO, and NBS, or USNO, PO, RRL, and USNO. The beginning and end of the measurements series occurred at the same institution (NBS or USNO), so that, theoretically, this time transfer should have zero error. The results showed GPS time transfer accuracies ranging from 5 to 100 ns with respect to UTC, but they were mostly on the order of tens of nanoseconds. Time transfer by nonsimultaneous measurements requires only that the satellite be in an orbit visible to each station, but the stations may be separated by any distance and the satellite does not have to be in view of both stations simultaneously. The disadvantage of this method is that the measurement comparison accuracy depends heavily on the predictability of satellite clock and position.

Another application of GPS precise time is synchronization of communication equipment/systems. An example is the time division multiple access (TDMA) bus. The TDMA technique is used by telephone companies to allocate multiple users to the same bus and it requires typical transmission synchronization accuracies of approximately 1 to 10 μs. A GPS time transfer receiver could provide an accuracy of 0.1 μs with respect to UTC. However, there is really no need for communication system time to be referenced to UTC; hence, the relative accuracy of GPS time transfer receivers to internally communications systems time could be up to one order of magnitude better.

3.4 Survey

Although GPS is not operational yet, the existing satellite constellation has been used for a variety of air, land, and ocean survey applications, such as oil rig positioning, aerial photomapping, coastal water bottom surveying, and establishment of a network of coordinated points. The main advantages of utilizing the GPS for surveying are: the achievable accuracy is on the order of 1-mm instrument error and 0.1 to 1 ppm (parts per million) variable measurement error between the measurement points; three-dimensional surveying capability; measurement points do not have to be intervisible, high productivity results in low cost of operation. Two techniques (i.e., point positioning with one receiver and relative positioning with a mobile receiver relative to a distant "reference station") are utilized for GPS surveying. The Texas Department of Transportation is currently working to establish a network that will consist of seven reference stations and numerous mobile receivers, because it is estimated that the cost savings will have paid for the GPS equipment in 1 to 2 years.

The achievable position accuracy depends on a number of factors, i.e., point or relative positioning technique used, accuracy with which tropospheric and ionospheric effects can be modeled, and the accuracy of the satellite ephemerides.

3.5 Science/Research

The GPS has been studied extensively by the research and science community, and it is already utilized in applications, such as geophysics (tracking of earth crustal movement, development of better ionospheric models), physics (measurements of earth gravity field), and determining the frequency stability of a neutron star or pulsar.

NASA is currently developing a GPS-based geodetic system for geophysical applications in selected tectonic regions, with a relative position accuracy of 1 to 3 cm on baselines of 1000 km. Recent experiments obtained relative position accuracies of about 3 to 5 cm. The Air Force Geophysics Laboratory has requested proposals for the development of a worldwide system of GPS monitor stations for the purpose of taking continuous ionospheric measurements. This will allow the generation of an extensive ionospheric model and, hence, increased prediction capability for the propagation of HF radio waves. The National Bureau of Standards (NBS) is utilizing GPS time transfer receivers to perform frequency stability measurement of a neutron star (pulsar), by tying NBS-UTC to the remote observation site time reference with the common view technique. The frequency stability achieved is on the order of a few parts times 10^{-14}.

4. GPS INFORMATION DISTRIBUTION

With GPS approaching the fully operational status, more and more sophisticated uses for the system are being established and the need for GPS information for military and civil users is steadily increasing. The DoD has identified the requirement for an Operational Status and Capability (OPSCAP) Reporting and Management System (ORMS). The ORMS will be used to assess the status of the GPS, provide simulation capability for resource management optimization and report operational status information at the mission, segment, subsystem, and component level. This status information will be distributed to a multitude of military and civil users based on established requirements and security considerations.

4.1 OPSCAP Reporting and Management System

The OPSCAP Reporting and Management System (ORMS) is planned to be a single, integrated system that will mainly utilize data extracted from the operational control segment data files, but will also accept inputs from other data sources.

The three main requirements guiding the design and development are to provide "up-channel" and "out-channel" reporting capability, and simulation capability for GPS resource management optimization. The "up-channel" reporting capability will allow the AF Space Command, as operator of the ORMS, to provide information about the GPS mission capability up the chain of command to the U.S. Space Command, the manager of the ORMS, and then to the Joint Chiefs of Staff. The "out-channel" reporting capability addresses a multitude of military and civil users and will provide information about the past, current, or future status of GPS. In order to perform these two functions, the ORMS has to monitor, record, and predict operational status information at the mission, segment (satellite, control), subsystem (master control station, communication links, ground antennae, monitor stations, individual satellite vehicles), and component level.

The OPSCAP information sent to the U.S. Space Command could be that the GPS is: operating/operating with restrictions/not operating, reported as a color status (green/ yellow/red), with amplifying information as may be necessary to adequately explain the cause, determine the impact, and examine response options. Status information distributed to military and civil users will probably be limited to satellite vehicle (configuration and performance data, such as orbit and clock information, user range error, global navigation/time transfer coverage. It is planned to distribute data to several intermediate data nodes, such as U.S. commands in Europe, the Pacific, and Atlantic; NATO (military); and other military users of the GPS, as well as to the DoD/DoT Notices to Airmen (NOTAM) facility and Civil GPS Information Center (CGIC). These intermediate data nodes would act as the focal point for their respective user community, e.g., CGIC for the civil GPS community.

The third function within the ORMS, the simulation capability will model the existing GPS resources (such as Satellite vehicles,and will determine the impact of subsystem or component failures on GPS mission performance and provide response options to such failures. It will allow assessment of the GPS status retroactively in combination with an archival data base and it will have the capability to predict future mission capability based on information such as SV launches, scheduled maintenance, component failure models, etc.

4.2 Civil GPS Information Center

Civil use of the GPS was prepared by several congressional actions, e.g., the International Maritime Satellite Act of 1978 and House and Senate Resolutions in September/October 1983. A brief overview of the broad range of civil GPS applications was given in Section 3, and it becomes quite clear that this diverse group of users, ranging from scientists to recreational boaters, will have very different needs for GPS information.

Hence, in order to evaluate how to satisfy anticipated civil user requirements, the USAF decided to sponsor a study (January 1987 to February 1988) to examine the feasibility of establishing an information service that could serve as a source of GPS information and the point of contact for civil users of GPS. The planned information service would receive and archive unclassified GPS information from the ORMS and other data sources, and distribute this information to the users via recorded message, electronic bulletin board, magnetic tape/disc, personal interaction, and publications.

In August 1987 the DoT agreed to a request from the DoD to assume the principal oversight and management responsibilities for the Civil GPS Service (CGS). The planned three-level management structure for the CGS as drafted by the DoT (October 1987) is shown in Figure 11.

The Administrative Executive (i.e., Secretary of Transportation, or designee), will provide high level, broad policy guidance and will be the signatory of agreements, e.g., between U.S. and foreign Governments. The CGS Program Office will interface with the U.S. Government and user/industry advisory boards and be the focal point for the Civil GPS User Community. The civil PPS office will administer the program under which qualified users will be able to obtain the precise positioning service (PPS) of the GPS. The CGIC Project Office will manage the design and development of the CGIC and monitor and evaluate the performance of the system.

The type of data to be distributed from the CGIC to the user community, utilizing an electronic bulletin board, are short-delay data (within minutes), e.g., SV orbit adjust, GPS SV time steering; intermediate data (within hours), e.g., navigation message as broadcasted or in engineering units; and long-delay data (within days), e.g., precise satellite ephemerides.

Currently, the DoD and DoT are working to survey the GPS user community for inputs to consolidate requirements for CGIC data types and services.

CONCLUSIONS

The GPS Navstar provides the means to significantly reduce the proliferation of radio navigation systems, and, thus, the cost associated with maintaining and utilizing navigation systems.

The military benefits of the GPS lie in the fact that it will not only enhance the overall effectiveness of weapon systems, but will multiply these benefits for coordinated operations through the utilization of a common reference grid. Some argue that relying on the GPS alone for navigation and communication (for example) jeopardizes the independent survivability of these capabilities. This would be true, if the GPS were used as standalone equipment. However, GPS equipment, integrated with a high-quality quartz oscillator (instead of an atomic clock) for communications and a low-cost INS for navigation, should satisfy most existing accuracy, reliability, and redundancy requirements for navigation and communications at a lower price.

Overall, the GPS will increase the safety of navigation, reduce transit time with corresponding economical advantages, and it will provide new capabilities for non-navigational users with applications such as surveying, timing and frequency, and science/research.

DEFINITIONS

PRECISE POSITIONING SERVICE (PPS) - The PPS provides the highest level of accuracy (16 m SEP). Its signal will be encrypted and will be made available initially to U. S. and selected allied military users. Limited civil use of the PPS may be authorized if it can be demonstrated that such use is in the national interest of the U. S., adequate security protection can be provided, and comparable accuracy cannot be obtained from another source.

STANDARD POSITIONING SERVICE (SPS) - Th SPS provides a lower level of accuracy. The U. S. DoD intends that the SPS signal will be broadcast in the clear and will be available for use by any properly equipped user. The SPS will be made available to civil, commercial, and other users on an international basis at the highest level of accuracy consistent with the U. S. national security interest (to date: 100 m, 2 drms).

CIRCULAR ERROR PROBABLE (CEP) - In a circular normal distribution (the magnitudes of the two one-dimensional input errors are equal and the angle of cut is 90 deg), circular error probably is the radius of the circle containing 50 percent of the individual measurements being made, or the radius of the circle inside of which there is a 50 percent probability of being located.

DISTANCE ROOT MEAN SQUARE (drms) - The root-mean-square value of the distances from the true location point of the position fixes in a collection of measurements. As used in this document, 2 drms is the radius of a circle that contains at least 95 percent of all possible fixes that can be obtained with a system at any one place. Actually, the percentage of fixes contained within 2 drms varies between approximately 95.5 and 98.2 percent, depending on the degree of ellipticity of the error distribution.

SPHERICAL ERROR PROBABLY (SEP) - The radius of a sphere within which there is a 50 percent probability of locating a point of being located. SEP is the three-dimensional analogue of CEP.

REFERENCES

1. AGARD, The Impact of Global Positioning System on Guidance and Controls Systems Design of Military Aircraft, Volume I, (1 September 1979), AGARD-AR-147, Vol. 1, France (U).

2. The Analytic Sciences Corp. (TASC), GPS-Aided Passive Terrain Following, (1 June 1985), TR-4185-4-1, Final Report, USAF, Space Division, Los Angeles, California (U).

3. The Analytic Sciences Corp. (TASC), GPS Benefits to Multiple-Sensor Data Fusion, (21 September 1984), TR-4185-3-1, USAF, Space Division, Los Angeles, California (U).

4. P. Axelrad and J. Kelley, "Near Earth Orbit Determination and Rendezvous Navigation Using GPS", (November 1986), IEEE PLANS, IEEE, Position, Location, and Navigation Symposium, New York, New York, pp 184-191 (U).

5. R. Braff, "GPS as a Sole Means for Civil Air Navigation", (November 1984), IEEE PLANS, New York, New York, pp 306-311 (U).

6. G.O. Burnham and R.H.J. Vos, "The Synergism of GPS on Electronic Terrain Map Systems", (November 1984), IEEE PLANS, New York, New York, pp 321-326 (U).

7. R.A. Eastwood, "Early Applications of Civil GPS", (November 1984), IEEE PLANS, New York, New York, pp 180-190 (U).

8. R.L. French, "The Evolving Roles of Vehicular Navigation", (June 1987), Institute of Navigation, Summer Meeting (Annual) Proceedings, Washington, D.C., pp 17-24 (U).

9. G.J. Geier, et al, "Design of an Integrated Navigation System for Robotic Vehicle Applications", June 1987 Proceedings, Institute of Navigation, Washington, D.C., pp 65-69(U).

10. D.W. Henderson, et al, "GPS As a Force Multiplier to Enhance Force Structure", (March 1980), AF University Air Power Symposium, Air War College, Maxwell Air Force Base, Alabama (U).

11. H. Heuberger, "Performance of the GPS Package on LANDSAT-5", (November 1984), IEEE PLANS, New York, New York, pp 147-152 (U).

12. Joachim Heyden, "Forewarned is Forearmed", NATO's Sixteen Nations, Vol. 32, No. 5, (August 1987), Mönch Publishing Group, West Germany, pp 39-41 (U).

13. C.E. Hoefener, "DoD Plans for GPS on Test and Training Ranges", Defense Systems Review, (February 1984), Interstate Electronics Corp., Anaheim, California, pp 23-26 (U).

14. C.E. Hoefener and G.S. Dankner, "Improved GPS Accuracy for TSPI Obtained Through Post-Flight Analysis", (January 1987), Institute of Navigation, Interstate Electronics Corp., Anaheim, California, pp 159-160 (U).

15. C.E. Hoefener and D. Beech, "Test and Evaluation Techniques for High-Dynamic Kinetic Energy Weapons", (October 1987), International Telemetering Conference, Interstate Electronics Corp., Anaheim, California (U).

16. C.E. Hoefener and J.B. McConnell, "A Universal Range Timing System Using GPS", (October 1986), International Telemetering Conference, Interstate Electronics Corp., Anaheim, California (U).

17. L. Jacobsen, L. Wells, and J. Grossman, "GPS Instrumentation for the Trident Weapon System", (November 1984), IEEE Position, Location, and Navigation Symposium, Interstate Electronics Corp., Anaheim, California, pp 272-276 (U).

18. K.M. Joseph, et al, "NAVSTAR GPS Receiver for Satellite Applications", (October 1980), AGARD 31st Guidance and Control Panel Symposium, London (U).

19. W.G. Melbourne, "GPS-Based Geodesy in California, Mexico and the Caribbean", (November 1986), IEEE PLANS Position, Location, and Navigation Symposium, New York, New York, pp 219-229 (U).

20. V.H. Prushan, "Radio Navigation - The Commercial Market and User Viewpoints", Institute of Navigation (January 1987), Winter Meeting (Annual) Proceedings, Washington, D.C., pp 146-149 (U).

21. E.A. Swinney, "Applications of GPS in the Commercial World", (November 1986), IEEE PLANS, Position, Location, and Navigation Symposium, New York, New York, pp 525-528 (U).

22. Jan Peter Tjardts, "Satelliten - Navigations System GPS Navstar Für Weltweite See-, Luft-, und Land-Anwendung", (March 1981), Luftfahrt International, Vol 3, West Germany, pp. 94, 97-98 (U).

23. U.S. DoD and DoT, 1986 Federal Radio Navigation Plan, (1987), DoD-4650.4, DoT-TSC-RSPA-87-3, National Technical Information Services, Springfield, Virginia (U).

24. M. Weiss, "Position Location with the NBS/GPS Time Transfer System" (November 1984), IEEE PLANS New York, New York, pp 175-179 (U).

25. L. Wells, "Field Test Results on the Use of Translated GPS for Trident I", (November 1986), IEEE Position, Location, and Navigation Symposium, Interstate Electronics Corp., Anaheim, California (U).

26. L. Wells, "Translated GPS Real-Time Tracking", (November 1983), IEEE National Telesystems Conference Record, Interstate Electronics Corp., Anaheim, California (U).

27. T.P. Yunck, "Strategies for Sub-Decimeter Satellite Tracking with GPS", (November 1986), IEEE PLANS, IEEE Position, Location, and Navigation Symposium, Proceedings New York, New York, pp 122-128 (U).

28. NAVSTAR GPS User Equipment and Applications, NATO Team, Air Force Space Division (June 1987).

ABBREVIATIONS

AHRS	Attitude Heading Reference System
ALCM	air-launched cruise missile
ASW	antisubmarine warfare
ATC	air traffic control
AVL	automated vehicle locating
AVM	automated vehicle monitoring
C^3I	command, control, communications, and intelligence
CAS	close air support
CEP	circular error probable
CGIC	Civil GPS Information Center
CGS	Civil GPS Service
DARPA	Defense Advanced Research Projects Agency
DT&E	developmental test and evaluation
DoD	Department of Defense
DoT	Department of Transportations
drms	distance root mean square
ECM	electronic countermeasures
EMCON	emission control
FAA	Federal Aviation Administration
ILS	Instrument Landing System

INMARSAT	International Maritime Satellite
INS	Inertial Navigation System
IOT&E	initial operational test and evaluation
JSTARS	Joint Surveillance Tactical Airborne Radar System
MLS	Microwave Landing System
NBS	National Bureau of Standards
OPSCAP	operational status and capability
ORMS	OPSCAP Reporting and Management System
OTHT	over-the-horizon targeting
PO	Paris Observatory
PPS	precise positioning service
PTB	Physikalisch Technische Bundesanstalt
RRL	Radio Research Laboratory
rms	root mean square
SA	selective availability
SAM	surface-to-air missile
SAR	synthetic aperture radar
SEP	spherical error probable
SLAM	standoff land attack missile
SLCM	sea-launched cruise missile
SPS	standard positioning service
TAO	Tokyo Astronomical Observatory
TDMA	time division multiple access
TSPI	time, space, and position information
UE	user equipment
USNO	U.S. Naval Observatory
UTC	universal time coordinated

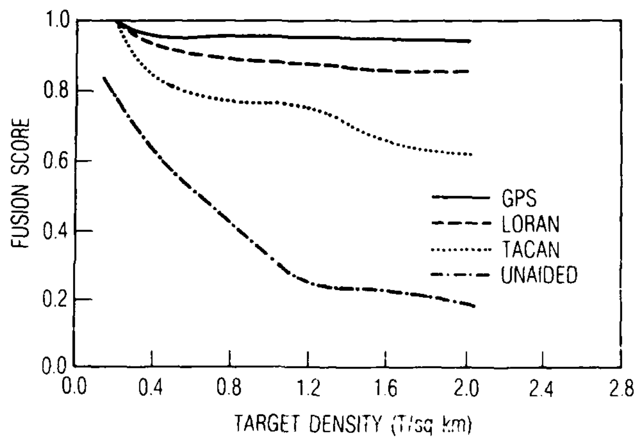


Figure 1

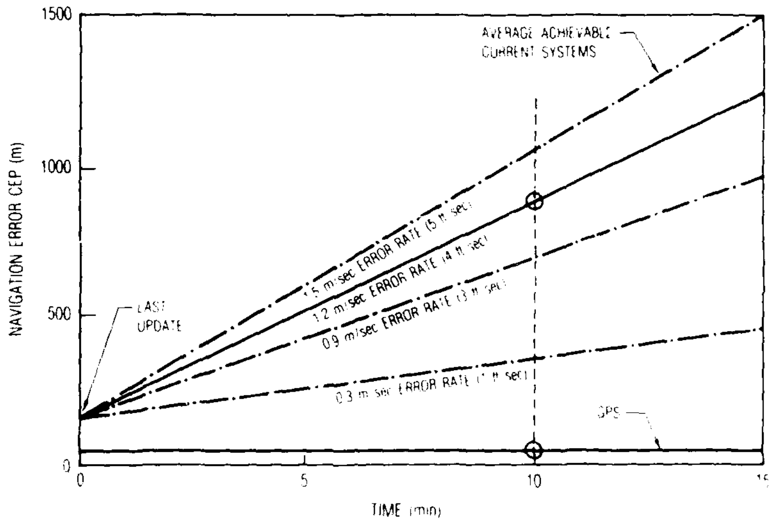


Figure 2

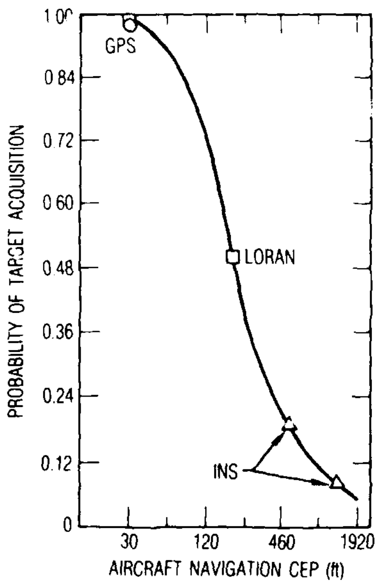


Figure 3

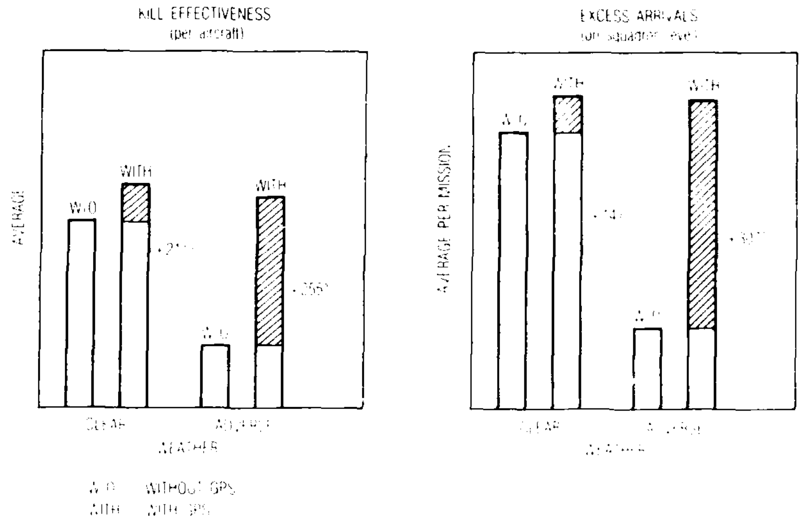


Figure 4

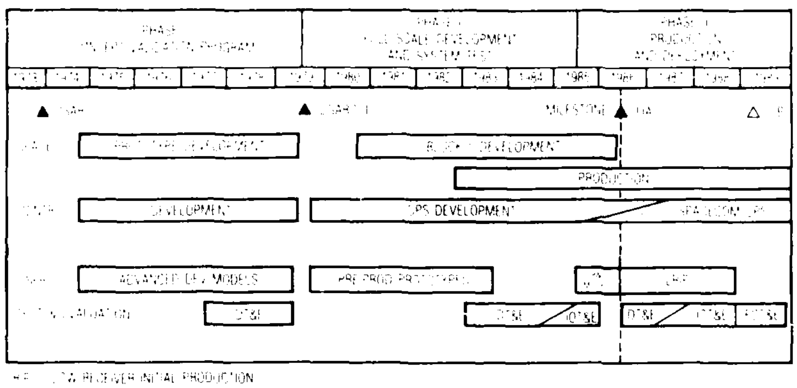


Figure 5

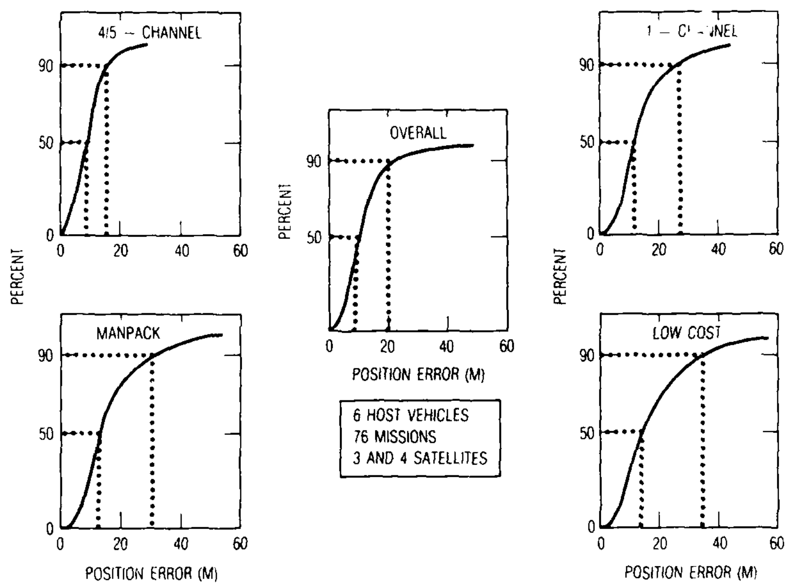


Figure 6

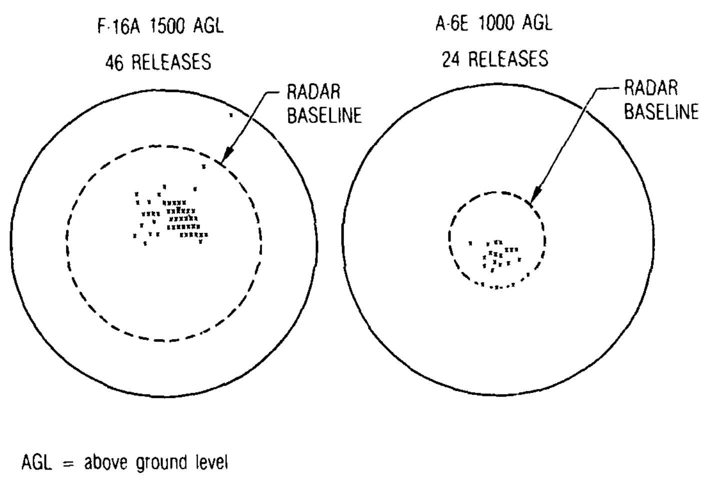


Figure 7

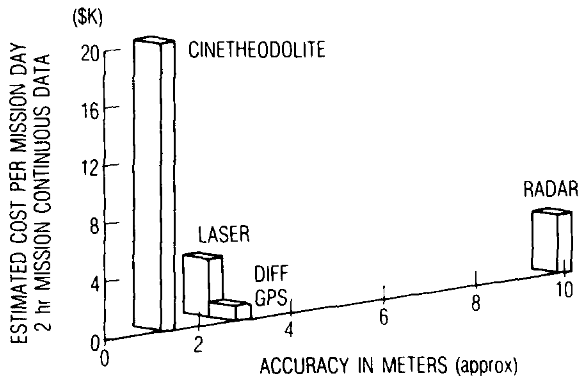


Figure 8

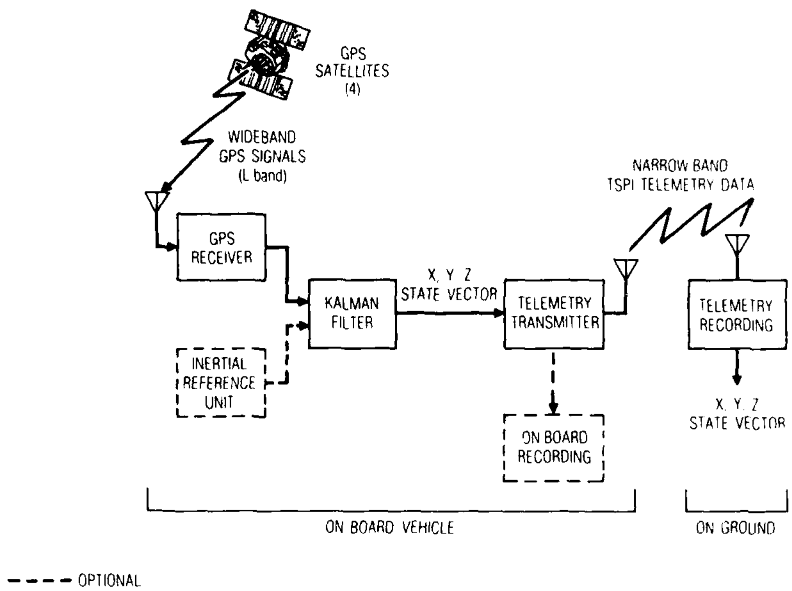


Figure 9

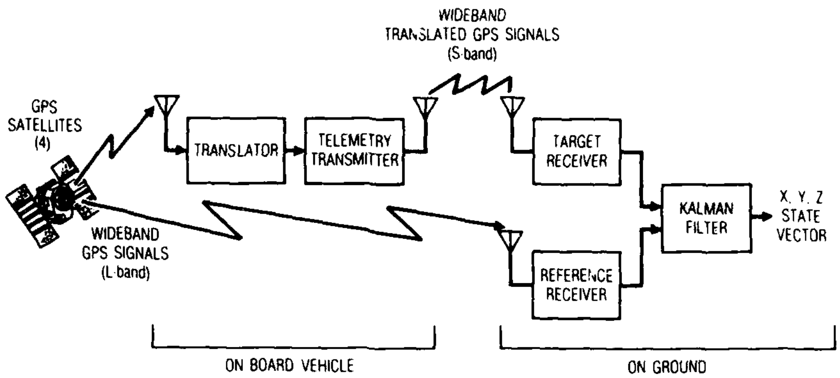


Figure 10

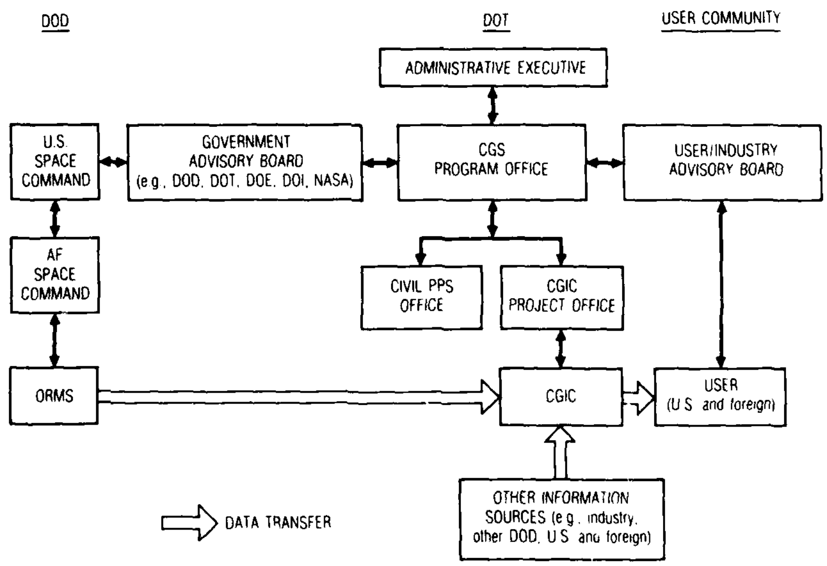


Figure 11

DIFFERENTIAL OPERATION OF NAVSTAR GPS FOR ENHANCED ACCURACY

Robert P. Denaro
Rudolph M. Kalafus

TAU Corporation
Los Gatos, California

INTRODUCTION

The Navstar Global Positioning System has demonstrated impressive absolute accuracy in tests and operational use. Most applications will benefit from the increased accuracy afforded by GPS over conventional systems. However, many applications require even better accuracy than GPS can provide in the standalone mode.

Differential GPS is a practical solution in many situations, providing few-meter level accuracy under dynamic conditions. It has been noted that the majority of error sources are related to the satellites and the propagation environment. With the satellites at approximately 20,000 kilometer altitude, the relative geometry between two close terrestrial sites is very similar, so the range errors tend to be highly correlated in a local geographic area. Conclusive data are not yet available, but the "local" area appears to be as large as 300-600 kilometer before significant spatial decorrelation of the range errors occurs. This area more than adequately covers many applications, such as offshore oil survey, harbor control and navigation, aircraft landing approach, and local military missions.

The subject of differential GPS is fundamentally a simple concept, but implementation details can be more complex, and handling of all cases takes special care to achieve continuously reliable results. This paper covers the concept and objectives of differential GPS, and presents a methodology for implementation. Example test results are presented which validate the potential of differential GPS.

DIFFERENTIAL GPS CONCEPT

Differential GPS is a concept that eliminates some of the common, bias errors experienced by conventional GPS. Differential GPS derives its potential from the fact that the measurement errors are highly correlated between different users (as well as highly autocorrelated). By employing a second GPS receiver with comparison to truth, slowly varying, correlated errors can be isolated and eliminated. In addition, depending on the relative rates, Selective Availability (SA), intentional degradation of the C/A signal, may be eliminated by differential GPS as well [1]. Measurement errors are also highly correlated between satellites for any particular user, but such common errors are removed by the conventional GPS solution as they are indistinguishable from user clock bias, hence corrupt only that estimate.

In differential GPS, a receiver reference station is located in the local area where greater accuracy is desired, as illustrated in Figure 1. The correlated errors that a receiver experiences (such as satellite ephemeris errors) should be common to all users in a relatively close geographical area. If the reference station can obtain a reliable estimate of its actual error and transmit that to dynamic users, the dynamic users may be able to compensate for a large portion of their errors.

DIFFERENTIAL SYSTEM ARCHITECTURE

A typical high-level architecture of a differential GPS system is illustrated in Figure 2. At the reference station, a (preferably) multichannel GPS receiver tracks all satellites in view. The station computes the differential corrections in some manner that optimizes the tradeoff of response to true dynamics in the errors and filter of noise that is uncorrelated to the errors experienced by the remote users. Before output data message formatting, it is important to perform some type of integrity management to ensure that the transmitted corrections are valid. The resultant valid corrections for each satellite in view are formatted in a standardized protocol and modulated onto the broadcast radio signal.

At the user end, the GPS receiver tracks (typically) the best four satellites, either in parallel or sequentially. The data link receiver decodes the differential corrections. The position processor applies the corrections to the pseudorange measurements from the receiver, matching satellite IDs, and computes a solution from the corrected pseudorange measurements (as well as uncorrected deltaranges, as appropriate). The improved positions are sent to a display or other interface.

DIFFERENTIAL GPS ACCURACY ENHANCEMENT POTENTIAL

The accuracy enhancement potential with differential GPS is evident in the diagram of Figure 3, drawn approximately to scale. The GPS satellites are in roughly circular orbits at an orbital altitude of about 3 earth radii. The major error sources for GPS are indicated on the diagram. Clearly the space segment errors of ephemeris and clock drift errors are nearly identical for the two users. It is also a good assumption that the ionospheric delay, and to a lesser extent, the tropospheric delay will be very similar. Only the receiver unique errors such as multipath and receiver noise are likely to be uncorrelated between receivers.

IMPLEMENTATION TYPES

There are several options for the manner in which differential GPS is applied. The variations are based on the type of information transmitted, the type of datalink used and the use of a static reference station. Table 1 identifies the various types and implementations of differential GPS.

Table 1. Types and Implementations of Differential GPS

Type of Correction Data	• Range Corrections • Solution Corrections • Translated Signal • Kinematic Phase
Data Link Options	• Uplink • Downlink • Pseudolite • Post-Mission
Reference Systems	• Multiple Reference • Integrity Monitoring • Dynamic Relative Navigation

The information transmitted in a differential GPS installation can vary widely. The most common is to transmit corrections to the range measurements, so that the mobile user needs to simply difference the received measurement corrections with his independent range (deltarange) measurements. This requires intercept of the range measurements inside the receiver prior to navigation solution processing, or output of the raw range measurements for external processing.

The alternative is to transmit corrections in the orthogonal solution frame, such as longitude, latitude, altitude, and apply them to the navigation solution in the same frame. This is somewhat simpler from the receiver standpoint, but requires complexity in the datalink structure. The problem is that the navigation solution coordinate errors are dependent on which satellites are used by the mobile unit. Since, in general, the static reference station has no knowledge which satellites might be in use by the mobile unit, it must compute and transmit navigation corrections for all possible combinations of satellites, which may number in the hundreds. In the case of measurement corrections, the total quantity of corrections is limited by the number of satellites in view, usually 6 to 10. Due to this advantage of the range corrections technique, the Radio Technical Commission for Marine Services (RTCM) Special Committee (SC) 104 has recommended the standard use of range corrections [2].

Another option for the type of information transmitted gives rise to the translator method [3]. Here, the mobile unit (or reference station) carries a GPS signal translator which does not track the signal but merely translates it to a communication frequency which is retransmitted to the ground. On the ground it is received by a GPS receiver with its front end modified to receive this translated frequency, and formulates a GPS solution which is now the solution of the remote mobile unit (this is possible since the time delay of the additional broadcast is identical for all satellite signals, so only the clock bias calculation is affected). A separate, conventional ground-based GPS receiver provides the differential correction computation. If used in the opposite direction, with the translator on the ground, the result is a real-time differential solution on the vehicle.

The fourth option in information transmission is the use of dynamic phase tracking [4]. This technique establishes relative phase position of the GPS carrier including the integer ambiguity, then attempts to maintain absolute carrier lock during movement. This is actually a mobile version of the geodetic survey technique using the phase observable of GPS. The geodetic survey technique is, in itself, a differential technique in that two sites are required, one unknown and one reference pre-surveyed location. It is not discussed in this paper due to its unique characteristics and static application.

Dynamic phase tracking is a potential technique for specially controlled dynamic applications. Typically the carrier tracking loop bandwidth is opened quite wider than normal to tolerate higher dynamics and to reduce the possibility of loss of carrier lock. This technique is extremely accurate, less than 1 meter, but suffers from the necessity to maintain lock continuously or reinitialize and start over. Tracking more than the minimum number of satellites can be used to recover phase ambiguity on a single lost signal in some cases.

Data link options present more alternatives. The major choice is whether the differential corrections are uplinked to the mobile unit for a real-time differential solution, or whether the mobile standalone solution is downlinked to the reference station and the differential solution is calculated only on the ground. The latter case is typical of test range applications where the role of differential GPS is for precise vehicle surveillance [?]. Note that these transmission options can be operated with either range or solution corrections without the usual drawbacks inherent in the uplink mode.

An additional alternative in data links that has some interesting features is the use of the so-called "pseudolite," a GPS signal and code generator that broadcasts the differential corrections over an L-band signal that is modulated identically to the satellite signals, except that the message includes the differential correction information for the actual satellites being tracked. The attractive feature of this configuration is that a separate data link receiver is not required for the differential correction information. Also, if one ties the pseudolite into the GPS satellite constellation by proper time synchronization, then the additional signal can actually serve as another ranging signal. This has advantages in outage periods or during periods of poor satellite geometry, especially in the vertical component [5]. However, one must have authority to broadcast on L-band, and there is a danger of interference with the real satellite signals since the dynamic range of the signal strength of a pseudolite is likely to be far greater than a satellite. The RTCM SC-104 has also devised a time division multiple access (TDMA) format to minimize such self interference [6].

If real-time differential accuracy is not required (that is, real-time standalone accuracy will suffice), then post-mission processing is an option that eliminates the need for a datalink altogether. For many operations this configuration will suffice.

The reference station of the differential system can take on several forms. The simplest is a single, correction computing unit that outputs its corrections to the appropriate data link. If multiple, geographically dispersed reference stations are used in an integrated manner, the possibility exists for detection of and correction for the effects of spatial decorrelation of the observed range errors from any particular satellite.

Since differential operation often applies to cases where the expense or safety of operations is critical, many reference station concepts employ some sort of integrity monitoring and management, assuring that only valid corrections are transmitted. This is possible because of the static, precisely known location of the reference station.

Carried to the extreme, the reference station can be designed as primarily an integrity monitoring station for GPS, alerting users when satellite signals are outside of some specified threshold in accuracy due to some reason. This is of particular interest to countries outside of the U.S. who are dependent on the U.S. DOD for the effective operation of the satellite network for reliable GPS operation. Integrity monitors, while they cannot afford any degree of control of the system, can at least provide effective monitoring, fault detection and isolation to achieve operational reliability.

Finally, there is a method of differential GPS that does not require a static reference station; relative GPS. Any two receivers, whether in motion or not, will be precise in their position with respect to each other with differential accuracy. The difference in this case is that neither receiver knows absolutely where it is to anything greater than the standalone capability of GPS. However, for mobile sea ranges or certain relative measurement applications, relative accuracy is all that is required.

EFFORTS IN DIFFERENTIAL GPS

The applications of differential GPS have expanded greatly since its beginnings in the early 1980s. Several programs have been conducted in various governments as well as commercially. While it would be prohibitive to describe all efforts in this paper, some of the more prominent efforts are summarized here.

In 1985, the U.S. Department of Transportation Research and Special Projects Agency (RSPA) and U.S. Coast Guard initiated a program to investigate the use of differential GPS for Harbor Approach and Harbor operations. The program was administered technically by the DOT Transportation Systems Center with the assistance of the Coast Guard Research and Development Center. The program included development of standards for differential GPS broadcast (transitioned to the RTCM SC-104 [2]), research into GPS and differential GPS concepts through university grants, development of a prototype differential GPS reference station, development of a prototype Mid-Frequency Marine Radiobeacon data link, and field testing [7].

The objectives of this development were to demonstrate the validity of the RTCM differential message standard, develop and implement concepts for maximizing the accuracy of the differential corrections, develop and test methods to insure the integrity of these corrections, and to demonstrate of cost-effective hardware architecture capable of meeting station requirements [8].

The U.S. National Oceanic and Atmospheric Administration (NOAA) is conducting a program to investigate the use of differential GPS to support its hydrographic and bathymetric survey operations. The National Ocean Service (NOS), a component of NOAA, is investigating new technologies to improve the efficiency of NOAA vessels conducting hydrographic surveys, oceanographic and fisheries research while simultaneously using less manpower and requiring a lower capital investment [9].

The availability of differential GPS as a worldwide marine positioning and navigation system will have a dramatic effect upon the NOAA hydrographic fleet. It will no longer be necessary to establish extensive networks of shore-based radio positioning systems and to determine geodetic positions. Using differential GPS, vessels will be able to operate on a 24 hour basis without the need to transit to calibration sites or await favorable propagation conditions. The benefits to be derived from the availability of GPS for scientific research are less quantifiable, but does make available a new research tool at a potentially less capital investment.

To meet NOS accuracy requirements for positioning, differential GPS using the C/A-code is being explored as a prime candidate. An accuracy of 5 meters, 2 drms, would satisfy all of the NOAA fleet's requirements, while an accuracy of 10 meters would satisfy nearly 90% of these requirements. The GPS P-code is also a potential candidate but the inherent security problems make it unwieldy for the majority of the NOAA fleet. Recent tests of a prototype system verified this accuracy potential in at sea trials [9].

The NASA Ames Research Center conducted a research program to evaluate differential GPS concepts for civil helicopter navigation [10]. The civil helicopter community will probably be an early user of GPS because of the unique mission operations in areas where precise navigation aids are not available. Many of these applications will have accuracy requirements that are very demanding, beyond that of conventional GPS. Such applications include remote area search and rescue, offshore oil platform approach, remote area precision landing, and other precise navigation operations.

Ames Research Center initiated a program to investigate the use of the C/A-code in a differential mode to provide a precision-approach capability for civil operations. As a first step, several alternative mechanizations of differential GPS were studied to determine the preferred configuration for supporting precision approaches into remote areas [11]. The study phase was followed by static laboratory tests using a single-channel, sequential GPS receiver. This receiver was installed in a NASA SH-3G helicopter and operated during a series of flight tests designed to assess its performance in several types of helicopter operations. The installation included study of accelerometer and altimeter aiding sensors to improve the vertical axis performance. Test results were encouraging [12].

A major test range application of differential GPS is being developed by a joint program in the U.S. DOD, headed by the Air Force at Eglin AFB, Florida [13]. This program is developing a new class of GPS P-code and C/A-code receivers that fit inside of a standard AIM-9 aircraft instrumentation pod (internal diameter approximately 14 cm). Called the GPS Range Applications Program, the system will be installed on U.S. DOD test ranges worldwide. Smaller, less extensive C/A-code variations of this system are in development in the commercial sector for application to smaller ranges, as well.

An interesting application of differential GPS is the NASA Jet Propulsion Laboratory's Topographic Experiment, or TOPEX. TOPEX will use differential GPS as the reference for precise ocean height measurement from the orbiting TOPEX satellite [14]. The goal is submeter accuracy, using a radar altimeter from the orbital altitude of approximately 1300 km.

The NASA Johnson Space Center has conducted a study of the use of differential GPS for local traffic control and docking on the Space Station. This is actually the relative mode of differential GPS discussed earlier, since the reference station is aboard the Space Station.

Canada has been very active in the study of differential GPS, and has pioneered the use of dynamic phase and is also studying the use of integrity monitoring [15]. This program augments their extensive work in geodetic survey using GPS. In the U.S., the Kennedy Space Center conducted some airborne tests of kinematic phase processing with success [16].

Commercial operations in Norway have implemented differential GPS for North Sea oil exploration. Results have been favorable in this region of the world [17,18]. The U.S. Gulf of Mexico is active in the use of differential GPS for oil exploration (currently seismic survey), and commercial services are now available for differential users [19]. The oil exploration field is an early operational user of differential GPS since economic value can still be achieved with the limited visibility available today.

Clearly the subject of differential GPS has become very popular in recent years, with major programs underway at a number of locations. Commercial use of differential GPS is expected to grow rapidly as the satellites become available.

DIFFERENTIAL GPS SYSTEM DESIGN

Although the implementation types of differential GPS can take many forms, the work of the RTCM has provided some amount of standardization for the conventional range correction update method. Since this is currently the most frequent application of differential GPS, it is described in this section as a typical example of how the system and algorithms are implemented.

The differential GPS system consists of the reference station, the data link, and the mobile units. The following discussion describes the features of the systems designed and implemented by TAU Corporation for the U.S. Coast Guard, NOAA, the U.S. Air Force and commercial users as representative of operational differential GPS systems.

REFERENCE STATION DESIGN

A Differential GPS reference station performs the overall functions of:

- Tracking of all GPS satellites in view
- Computation of real-time differential GPS corrections for each satellite
- Monitoring of all factors influencing the integrity of the differential corrections
- Decision on when to initiate, interrupt, or terminate correction broadcast on each satellite
- Preparation and communication of RTCM format differential correction message for broadcast
- Operator interface for testing, monitoring, data analysis, message formulation, and special message communication.
- Data storage and archival

These functions are depicted in Figure 4. In addition, for a P-code differential GPS reference station, the use of dual frequency ionospheric calculation and possibly tropospheric measurement and modeling is optional. In general, the added noise inherent in processing the second frequency for ionospheric correction is less desirable than assuming a common delay at both ends and leaving the effect in the corrections. Similarly, local measurement of the refractivity and resultant tropospheric delay is less precise than the same assumption, assuming that corrections are made for altitude differences of the mobile units for airborne applications. However, there may be cases where the local conditions contain sufficient gradients, or the user-to-reference separation distances are sufficiently large that such pre-calculation and removal is called for.

To accomplish the function of all-in-view tracking, the station controls and monitors a continuously tracking, multi-channel GPS receiver. A sequencing receiver can be used at the reference station, but there will be some loss of accuracy due to the interrupted carrier tracking and gaps in code measurement. An atomic frequency standard may be used in lieu of the reference receiver's quartz oscillator for increased stability and reliability. Increased noise in the reference station processing is more serious than in the GPS standalone case because of the additive effect of errors when combined with the mobile solution in the correction differencing process. The reference station maintains control of the reference receiver for satellite channel assignment and moding.

Computation of real-time differential corrections is a three-step process involving pseudorange error estimation, clock bias correction, and linear correction prediction (since the broadcast parameters are range error and rate of change of range error). Both receiver pseudorange (code) and deltarange (carrier) information are processed to compute the ultimate broadcast quantities of pseudorange error and rate-of-change of pseudorange error. However, the use of carrier must be handled properly since ionospheric delay changes and multipath affect the code and carrier in different ways (group delay, phase advance phenomenon). The pseudorange error estimation process employs some type of filtering for noise reduction and best response to measurement error dynamics. If a real-time Kalman filter is used, by-products of the Kalman filter processing are also useful for integrity monitoring.

This last feature, integrity monitoring, is a particularly significant function of an operational reference station. Users of differential GPS are often characterized by two factors which mandate reliability of the differential corrections: safety and cost of operations, either or both of which may be jeopardized by broadcast of erroneous corrections. Therefore, the integrity monitoring feature may be an important function of the differential GPS reference station.

The broadcast message for differential corrections and other station information, which may be the RTCM standard data message format for differential GPS corrections, is constructed continuously. Order and frequency of message types is established manually by the station operator (or defaulted) depending on user needs and datalink capacity. The message is sent and buffered at a communication port for interface with the data link controller.

Although reference stations are generally designed for autonomous operation (including automatic power-out recovery), an optional operator CRT interface is desirable for testing or monitoring. Data storage and archival are provided for post-operation anomaly and event review, data analysis, special testing, or demonstration.

An example of pseudorange error estimation accomplished by a Kalman filtering process is illustrated in Figure 5. Using the broadcast ephemeris and the known (survey) location of the reference station antenna phase center, the pseudorange error measurement, z_i, formulated as the error between the pseudorange measurement, PR_i, and the calculated range:

$$z_i = PR_i - |S_i - x|$$

where

S_i = i^{th} satellite position calculated from ephemeris data
x = Surveyed Reference Station location

Integrity management is the term used to describe the set of functions and checks performed to assure that the differential corrections in fact improve the conventional navigation solution. The various functions performed by integrity management may include:

- Detection of user clock jumps
- Detection of unusually high rate pseudorange errors
- Compensation for changes in measurement noise
- Avoidance of excessive low-elevation tropospheric and ionospheric delay uncertainty
- Detection of unusually high ionospheric delay rates

Typical techniques for the integrity management function are statistical residuals tests, data editing, reasonableness checks, and threshold tests to detect these anomalies. Response to any integrity anomaly always includes an operator message and event log, and may trigger temporary cessation of message broadcast. Of course, these methods cannot detect the errors due to spatial decorrelation, for which interstation links would be necessary.

DATALINKS AND MESSAGE FORMULATION

Any type of datalink can be used for differential GPS, depending on the operational constraints. The approximate data transition rate for effective measurement correction is 150 - 250 bits per second. As low as 50 bits per second may suffice for usual conditions (use of the pseudolite would be limited to this rate, for example). Other than that, the usual restrictions of line-of-sight, power, security and interference determine datalink selection.

Commercial offshore marine operators are studying the use of the already deployed marine radiobeacon transmitters, which broadcast in the 285-335 KHz band. A promising technique has been developed which employs Minimum Shift Keying (MSK) modulation and convolutional coding with interleaving to achieve low sideband, impulse-noise-resistant operation at ranges well in excess of the rated range of the beacon [20]. The technique involves placing the differential GPS transmissions on a subcarrier about 325 Hz above the beacon CW carrier. It has been estimated that the affect on both manual and automatic direction-finders (including airborne units, which sometimes tune in to the marine beacons) is small, causing less than 2 degrees of bearing error. The U.S. Coast Guard has fabricated and tested a prototype radiobeacon data link.

Two examples of a reliable high-frequency system have been devised, one by TAU Corporation of California and one by Sercel of France [9,21]. Both schemes use multiple frequencies and multiple time slots for the redundant message to maximize probability of reception.

In the case of the TAU system, a time-multiplexing modem accepts an RS-232C serial input at about 80 baud from the reference station, and time-multiplexes the data over 7 channels. The data output by the demodulator is generated by a consensus among the seven channels. Each channel on the demodulator has an indicator that lights if the data from that channel is corrupted, so that the quality of the data, as well as the signal, can be observed. Operational tests showed excellent results. The Sercel system was demonstrated at very long ranges with similar success. The only drawback to these systems is that they require access to multiple HF slots.

The RTCM developed a recommendation for a general message format that was patterned after the GPS Navigation Message Format as specified in the system document ICD-GPS-200 [2]. The major difference between the formats is that the differential messages will use a variable length format, whereas the GPS format has fixed length subframes.

The format does not dictate the data rate, but the recommendations are premised on a minimum 50 bits per second, the same as GPS (highly desirable if pseudolites are employed). Each message frame is made up of several 30-bit words, always headed by two standard words which provide a fixed preamble, message type identifier, station identification, timing data, message frame length, and station health information. Allowance is made for 64 message types. Table 2 lists the first 16 message types.

Table 2. Message Types

Type No.	Message Type	No.	Message Type
1	Differential Corrections	9	High Rate Differential Corrections
2	Delta Differential Corrections	10	P-Code Differential Corrections (Reserved)
3	Station Parameters	11	C/A-Code L1,L2 Delta Corrections (Reserved)
4	Surveying (Carrier Phase)	12	Health Message (ASCII String)
5	Constellation Health	13	Undefined
6	Null Frame	14	Undefined
7	Beacon Almanacs	15	Undefined
8	Pseudolite Almanacs	16	Special Message (ASCII String)

Message Type 1 is the primary message: it contains the pseudorange and rate-of-change of pseudorange corrections to be used for navigation. The satellite pseudorange measured by the receiver is corrected as follows, using the Message Type 1 parameters:

$$PR(t) = PR_m(t) + PR_o + (dPR_o/dt)(t-t_o)$$

where $PR_m(t)$ is the measured pseudorange at time t in meters, PR_o is the pseudorange correction at t_o, dPR_o/dt is the rate-of-change of correction, and t_o is the modified Z-count of the second word, applied to that satellite indicated by the satellite ID. Also provided is satellite health indication, which includes an estimate of the differential error in terms of User Differential Range Error (UDRE) as estimated by the reference station. The satellite health is not just an echo of the satellite message-provided data, but is determined by the reference station, based on a comparison of true and measured pseudorange, as well as the quality of the data and the measurement.

Issue of Data (previously called "Age of Data") tells the user whether the reference station is basing the corrections on the most recent satellite data. Most of the time the user and reference station will base the position estimates on the same satellite data. If the user's data is older than that of the reference station, the information contained in the Type 2 Message is used to further correct the measured pseudorange. This enables a user to modify the received corrections when the satellite ephemeris and clock data set with which it is operating is older than the data set being used by the reference station. It will rarely occur, but if the user's data is newer than that of the reference station, the new data should not be used until the reference station is also basing the corrections on the new data.

The other message types provide further details on the station or are used for special purposes such as surveying or using pseudolites.

Message formatting is accomplished for multiple RTCM message types in the typical reference station. Repetition rates vary (as selected by operator) according to mission and user needs. A typical sequence might be:

- Type 1 as often as possible
- Type 2 one per minute
- Alternate types 3 and 5 every two minutes

DIFFERENTIAL GPS ERROR SOURCES

The factors affecting differential GPS accuracy in most applications are:

- Spatial decorrelation of "common" errors
- Temporal decorrelation/correction update rate
- Uncompensated vehicle motion-induced errors
- Reference station or vessel position computational (filtering) errors

GPS error sources have varying sensitivities to navigator-to-reference station separation distance. The generally accepted set of GPS error sources is listed in Table 3 [22]. Their relative error contributions under differential operation conditions are also indicated.

Table 3. Differential GPS Error Sources

	Nominal Error	Differential Error
•	Satellite Ephemeris Error	Small
•	Satellite Clock Drift Error	Very Small
•	Selective Availability (C/A-Code)	Small to Large
•	Ionospheric Propagation Delay Error	Large
•	Tropospheric Propagation Delay Error	Moderate
•	Multipath Error	Moderate to Large
•	Receiver Clock Drift Error	Very Small
•	Receiver Noise	Small

Satellite ephemeris and clock errors have small error effects on the differential user caused by slightly different look angles to each satellite by the ship and reference station. The effect is small due to the high altitude of the GPS satellite orbits which dwarfs any reasonable user-to-reference station separation distance, as illustrated in Figure 3. The decorrelation in this case is caused by the different line-of-sight component of the satellites' three-dimensional orbital error on the two different lines-of-sight. As illustrated in Figure 3, for a separation distance of 300 km, the angular separation of the two lines-of-sight is less than 18 milliradians (1 degree). In total, the differential error due to ephemeris can be expected to cause a magnitude of about .001% of the baseline distance for typical ephemeris errors (e.g., 100 m in-track, 15 m cross-track, and 2 m radial) [23].

Satellite clock drift error is completely observed in the line-of-sight range. However, if differential updates (including range correction and rate of change of range correction) are accomplished infrequently, the residual drift between the clock's actual drift and the linear assumption inherent in the correction rate of change will introduce an error. This, too, should be negligible for updates even as seldom as every few minutes, unless some sort of clock anomaly transient occurs.

Selective availability (S/A) is a potentially large source of differential error. In analysis of typical S/A data, Kalafus [1] has shown that the expected range error and first two derivatives of range error due to S/A should be:

$$E(\bar{r}) = E(\dot{\bar{r}}) = E(\ddot{\bar{r}}) = 0 \qquad \sigma_r = 100 \text{ m}$$
$$\sigma_{\dot{r}} = .14 \text{ m/s}$$
$$\sigma_{\ddot{r}} = .004 \text{ m/s}^2$$

As with other errors, a constant range error due to S/A will not introduce any residual differential error, but range rate or higher derivatives will cause the actual error to deviate from the predicted linear estimate in correction rate of change. Once again, this is a correction update rate issue. For the above data, an error due to S/A of .2 m can be expected for an update interval of 10 sec.

The ionosphere is a major source of error in standalone C/A-code GPS. For differential GPS, the well-behaved temporal and spatial variability of ionospheric delay in mid-latitude regions make it generally well-compensated by differential GPS. However, the ionosphere does have a peculiar personality. Studies by the Air Force Geophysics Laboratory [24] have shown somewhat ill-behaved results at high latitudes and at the equator (the former due in part to the Aurora Borealis, the latter due to the electrojet). Ionospheric delay is caused by an electron layer at about 100 to 1700 km altitude above the earth. Figure 6 illustrates the general shape of the ionosphere, indicating its diurnal variation and its latitudinal variation. Moving satellite lines of sight that intersect high gradients in this ionospheric shell may experience higher dynamic errors than experienced in more benign regions.

The above considerations refer to the temporal correlation of ionospheric errors. The spatial decorrelation of the ionosphere is perhaps a more significant problem for differential GPS. Figure 7 illustrates worldwide isoclines for vertical ionospheric total electron density at a "snapshot" in time [24]. As a rule of thumb, 10^{16} el/m^2 in density is equivalent to about 15 cm in delay. Over a 200 mile baseline, total delays can vary by several meters, although variances are typically less than 1 foot.

The tropospheric delay is caused by a layer of water vapor a lot closer to the surface, from 1 to 20 miles or so. Because of its closeness to the surface and its causative factors, local variations in tropospheric effects are more likely than with the ionosphere. However, for all but the lowest elevation angles, the troposphere is highly modelable and is generally of lesser overall magnitude than the ionospheric delay. For differential considerations, a local weather front can cause reasonably fast variations in the troposphere. Furthermore, spatial variations of a meter or so over 200 mile baselines are possible, though rare.

Multipath error is a potentially serious effect for differential GPS. Much of the problem can be eliminated by good antenna design and placement, but for more casual implementation of differential system, multipath could cause differential errors of a few meters or even preclude reference station operation at certain low elevation/azimuth combinations.

Multipath effects can be seen in the pseudorange measurement as a satellite rises and the reflected signal moves in and out of phase. The effect is alternate positive and negative reinforcement of the primary signal. Figure 8 was plotted from two very diverse sets of data and illustrates this effect. The first plot is from some Navy dockside tests of a GPS antenna mounted on a ship mast with GPS signal reflections occurring off the ocean. Notice the periodic ±2 meter "waves" in the pseudorange measurement error, which have been attributed to multipath. Similarly, the second plot [25] is for an antenna located near some trailers in an Air Force Geophysics Laboratory ionospheric delay test. They exhibit as much as ±4 meter changes.

For moving vehicles, for higher elevation satellites, and for better designed RF absorbant antenna ground planes, multipath is generally not observed other than as additional noise, but care must be taken in design to assure this. The danger, of course, is that any multipath effects are likely to be completely uncorrelated between navigator and reference station.

Navigator receiver clock drift error should generally have a minor effect on differential accuracy. If the reference receiver is a slow-sequencing single channel tracking unit, then some residual error due to clock drift may be induced. However, for the typical case of a multi-channel simultaneously tracking reference receiver, the clock drift error contribution to the differential solution is well sub-meter. The only real threat from the user clock is possible frequency transients in either receiver and acceleration sensitive errors in a navigator receiver. Integrity management techniques can virtually eliminate the former while proper mounting should mitigate the latter.

Receiver noise, caused primarily by measurement noise, is particularly bothersome in differential applications because the reference station and navigation receivers will have uncorrelated noise such that nay noise in the differential corrections will root-sum-square with the ship receiver noise. Navigator receiver noise can be smoothed only outside of the bandwidth of the vehicle dynamics which must be tracked, unless other aiding sensors are used to allow increased isolation and smoothing of measurement noise. Reference receiver noise is smoothable outside of the expected bandwidth of the GPS signal error dynamics, which is a much more satisfying result. However, while the error dynamics can be highly filtered, care must be taken to allow for tracking of anomalous ionospheric transients and, in the case of C/A-code operation, selective availability signal dynamics.

TEST RESULTS

Differential GPS has matured to the point where a number of significant tests have been conducted which demonstrate performance, including a few preliminary kinematic phase tests. Some of the major recent results are described in this section.

In 1984, the Air Force GPS Joint Program Office operating location at the Yuma Proving Ground Test Range conducted a series of differential GPS tests in C/A and P-code, static and dynamic [26]. Final errors were generally in the range of 2-3 meters horizontal, 4-6 meters vertical. C/A-code exhibited about 2 to 3 times the noise of P-code. An interesting result of these tests was a data comparison between Yuma and Edwards AFB, California, a distance of about 400 km. Even at this baseline distance, C/A-code differences were less than 5 meters.

In June of 1988, Sercel S.A. of France conducted C/A-code differential GPS tests [21]. Tests were conducted at ranges of 200, 430, and 680 km. Results were generally in the 3-6 meter range horizontally, with 1-3 meters vertically. An interesting result of these tests was an apparent day-to-day repeatability of the bias errors.

Norway has also been active in differential GPS with their DIFFSTAR system [17]. Their results in Northern Norway yielded horizontal accuracies in the 5-8 meter range over short baselines, and they cite similar results over 200-500 km. An interesting observation by them was that differential noise content was worse than standalone, but, of course, biases were mostly removed.

The NASA-Ames Research Center has been conducting studies of differential GPS for several years. The recent flight tests of an experimental helicopter precision landing system are interesting, incorporating barometric altimeter and vertical accelerometer data to reduce the vertical axis error as much as possible [10]. Their findings indicated that differential GPS substantially reduced bias errors, and that addition of a barometric altimeter reduced the noise on the vertical axis solution by 40%. The vertical accelerometer had only a 10% effect.

The U.S. Coast Guard Tests showed some very dramatic results in marine vessel applications of differential GPS [7]. The Coast Guard used TAU and Magnavox differential equipment and TI, Trimble, and Magnavox receivers. Operating at speeds from 10 to 20 knots, the Coast Guard obtained results of less than 2 meters radial error. They also simultaneously compared standalone P-code tracking with differential C/A-code, and found standalone differential C/A-code to be marked by superior. Figure 9 illustrates the performance from these tests.

The NOAA conducted differential GPS tests on a moving vessel as well in July 1987, using TAU differential equipment and Trimble C/A-code receivers [9]. The tests compared standalone and differential results, and included differential results from two reference stations, one on shore nearby the test area (Elizabeth River near Norfolk, Virginia) and one approximately 480 kilometers to the north in Groton, Connecticut.

The results of these test were very encouraging. Figures 10 and 11 show good periods of data in a radial scatter diagram and cumulative probability distribution, respectively. Table 4 shows the statistical results. Note that results for the Groton data, over the longer baseline, were excellent. However, subsequent analysis showed that satellite clock drift in one of the satellites was the dominant error source for these tests, hence, very little spatial decorrelation is expected.

Table 4. Performance Comparisons

	Standalone	Local Reference	Remote Reference
Case 1			
$\bar{x}/\sigma_x$	27.3/8.8	1.3/2.0	1.6/2.1
$\bar{y}/\sigma_x$	0.7/5.3	0.3/1.5	-0.3/1.6
h_{rms}	29.4	2.9	3.1
Case 2			
$\bar{x}/\sigma_x$	-2.7/2.9	1.8/2.2	0.3/2.3
$\bar{y}/\sigma_x$	-2.5/1.9	0.0/2.2	-2.[illegible]/2.2
h_{rms}	5.0	3.[illegible]	4.2

A more austere configuration of differential GPS was represented by some tests conducted by Trimble Navigation with one of their sequential GPS receivers and TAU's Reference Station [27]. They showed 2-3 meter performance, even with a sequencing receiver, and dramatically illustrated the insensitivity of differential GPS operation to either selection of a new satellite set or reception of a new satellite upload. In both of these latter cases, residual errors were with the overall (2-3 meter) noise content of the differential solution.

Finally, some results of kinematic differential phase tracking are shown. Nortech Surveys of Canada and the Canadian Hydrographic Service conducted ground vehicle tests over a 1.5 km baseline [28]. Results showed 1-3 meter errors. The low noise value was expected due to incorporation of phase data, but the low bias value was attributed to a very low contribution by the ionosphere of only about 1 meter. Also, cesium clocks were used for several of the tests which significantly reduced measurement noise. In addition, the Goddard Space Flight Center did tests on a P-3 aircraft. Using a laser altimeter off a lake surface as a reference, the results in the vertical axis showed decimeter accuracy. An example is shown in Figure 12 [16].

CONCLUSIONS

Differential GPS has emerged as a successful technique for achieving extremely precise navigation with GPS, up to an order of magnitude improvement over conventional operation. Although the basic concept is straightforward, handling of anomalous conditions and insuring reliability through differential operation are design challenges requiring sophisticated approaches. Numerous programs are currently underway, and a few commercial systems have appeared. Initial testing indicates accuracy potentials on the few-meter range, sufficient to support a wide variety of applications beyond that envisioned for standalone GPS.

REFERENCES

1. Kalafus, R., Vilcans, J., and Knable, N., "Differential Operation of Navstar GPS," NAVIGATION, Volume II, The Institute of Navigation, Washington, DC, March 1983.

2. Kalafus, R., Dierendonck, A.J., and Pealer, N.A., "Special Committee 104 Recommendation for Differential GPS Service," NAVIGATION, Volume III, The Institute of Navigation, Washington, DC, June 1986.

3. Blackwell, E.G., "Overview of Differential GPS Methods," *Journal of the Institute of* Navigation, Vol. 32, No. 2, Summer 1985.

4. Remondi, B.W., "Performing Centimeter-Level Surveys in Seconds with GPS Carrier Phase: Initial Results," NAVIGATION, Volume III, The Institute of Navigation, Washington, DC, September 1985.

5. Klein, D., and Parkinson, B.W., "The Use of Pseudo-Satellites for Improving GPS Performance," NAVIGATION, Volume III, The Institute of Navigation, Washington, DC, June 1984.

6. Stansell, T.A., "RTCM SC-104 Recommended Pseudolite Signal Specification," NAVIGATION, Volume III, The Institute of Navigation, Washington, DC, June 1986.

7. Pietraszewski, D., Spalding, J., Viehweg, C., and Luft, L., "Status of United States Coast Guard Sponsored Differential GPS Demonstration System Development," Pr ceedings of the Satellite Division First Technical Meeting, The Institute of Navigation Satellite Division, Colorado Springs, Colorado, September 21-25, 1987.

8. Denaro, R.P., Kalafus, R., and Quill, J., "Differential GPS Reference Station Design and Development," Proceedings of the National Technical Meeting, Institute of Navigation, Anaheim, California, January 1987.

9. Kremer, G.T., Denaro, R.P., Swale, S.L., and Rulon, T., "Test Results of a Differential GPS System for NOAA Offshore Surveys," Proceedings of the National Technical Meeting, The Institute of Navigation, Santa Barbara, California, January 26-29, 1988.

10. Edwards, F.G., Paielli, R.A., and Hegarty, D.M., "Helicopter Terminal Approach Using Differential GPS with Vertical-Axis Enhancement," Proceedings of the Satellite Division First Technical Meeting, The Institute of Navigation Satellite Division, Colorado Springs, Colorado, September 21-25, 1987.

11. Denaro, R.P., "Application of Differential GPS to Civil Helicopter Terminal Guidance," Proceedings of Digital Avionics Systems Conference, December 1984.

12. Edwards, F.G., Hegarty, D.M., Turner, R.N., van Graas, F., and Sharma, S., "Validating the Airborne and Ground Based Components of a Differential GPS System," Proceedings of the National Technical Meeting, The Institute of Navigation, Santa Barbara, California, January 26-29, 1988.

13. Lichten, S.C., Wu, J.T., and Yunck, T.P., "Predicted TOPEX Positioning Accuracy with Differential GPS Techniques," Proceedings of First International Symposium on Precise Positioning with the Global Positioning System, April 1985.

14. Beser, J., and Lyon, M., "Differential GPS in the Range Applications Program," Proceedings of the National Technical Meeting of the Institute of Navigation, January 1986.

15. Delikaraoglon, Demitris, "Report on a Canadian Active Control System Using GPS," Proceedings of the 42nd Annual Meeting of the Institute of Navigation, Seattle, Washington, June 1986.

16. Krabill, W., and Martin, C., "Aircraft Positioning Using GPS Carrier Phase Data," Journal of the Institute of Navigation, Spring 1987.

17. Hervig, K., "DIFFSTAR - A Project Based on Differential GPS in Northern Norway," IEEE PLANS '86, Position Location and Navigation Symposium, Las Vegas, Nevada, November 4-7, 1986.

18. Karkalik, F., Fitzgibbon, K., and Vargus, R., "Differential GPS Tests in the North Sea," Proceedings of the National Technical Meeting, The Institute of Navigation, Santa Barbara, California, January 26-29, 1988.

20. Enge, P.K., Ruane, M.F., and Sheynblatt, L., "Marine Radiobeacons for the Broadcast of Differential GPS Data," IEEE PLANS '86, Position Location and Navigation Symposium, Las Vegas, Nevada, November 4-7, 1986.

21. Nard, G., "Reliable High Accuracy Long Range Real-Time Differential GPS Using a Lightweight High Frequencies Data Link," IEEE PLANS '86, Position Location and Navigation Symposium, Las Vegas, Nevada, November 4-7, 1986.

22. Lachapelle, G., Falkenberg, W. and Casey, M.,"Use of Phase Data for Accurate Differential GPS Kinematic Positioning," IEEE PLANS '86, Position Location and Navigation Symposium, Las Vegas, Nevada, November 4-7, 1986.

22. Denaro, R.P., Loomis, P.V.W., and Yoerger, D., "The Application of Differential GPS to Marine Vessel Dynamic Positioning," Proceedings of the 42nd Annual Meeting of the Institute of Navigation, Seattle, Washington, June 1986.

23. Beser, J., and Parkinson, B., "The Application of Navstar Differential GPS in the Civilian Community," NAVIGATION, Volume II, The Institute of Navigation, Washington, DC, June 1981.

24. Klobuchar, J.A., "Ionospheric Effects on Earth-Space Propagation," AFGL-TR-84-0004, December 1983.

25. Bishop, G.J., and Klobuchar, J.A., "Multipath Effects on the Determination of Absolute Ionospheric Time Delay from GPS Signals," Radio Science, 20:3, pages 388-396, May 1985.

26. Druczynski, L.R., Porter, W.W., Abby, D.G., and Wilson, E.T., "Global Positioning System Differential Navigation Tests at the Yuma Proving Ground," Journal of the Institute of Navigation, Vol. 32, No. 2, Summer 1985.

27. Eschenback, R., and Tiwari, A., "Differential GPS with a Sequencing Receiver," Proceedings of the Satellite Division First Technical Meeting, The Institute of Navigation Satellite Division, Colorado Springs, Colorado, September 21-25, 1987.

28. Kleusberg, A. and Wells, D., "High Precision Differential GPS Navigation," IEEE PLANS '86, Position Location and Navigation Symposium, Las Vegas, Nevada, November 4-7, 1986.

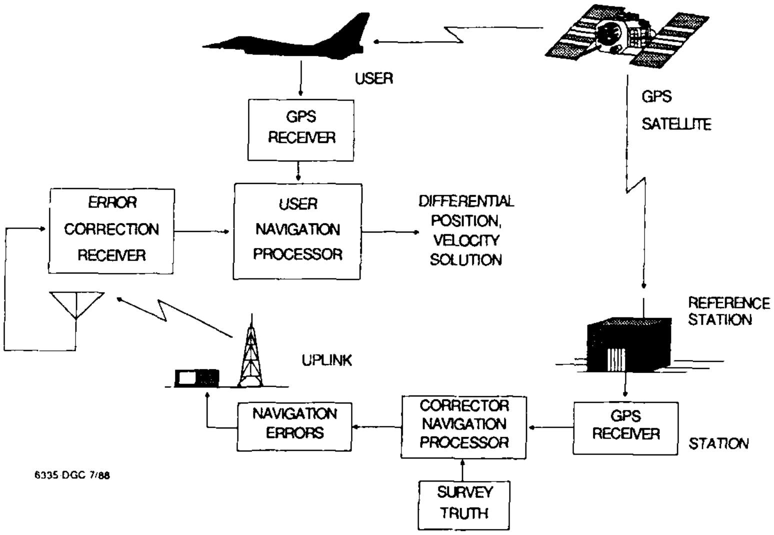


Figure 1. Differential GPS Concept

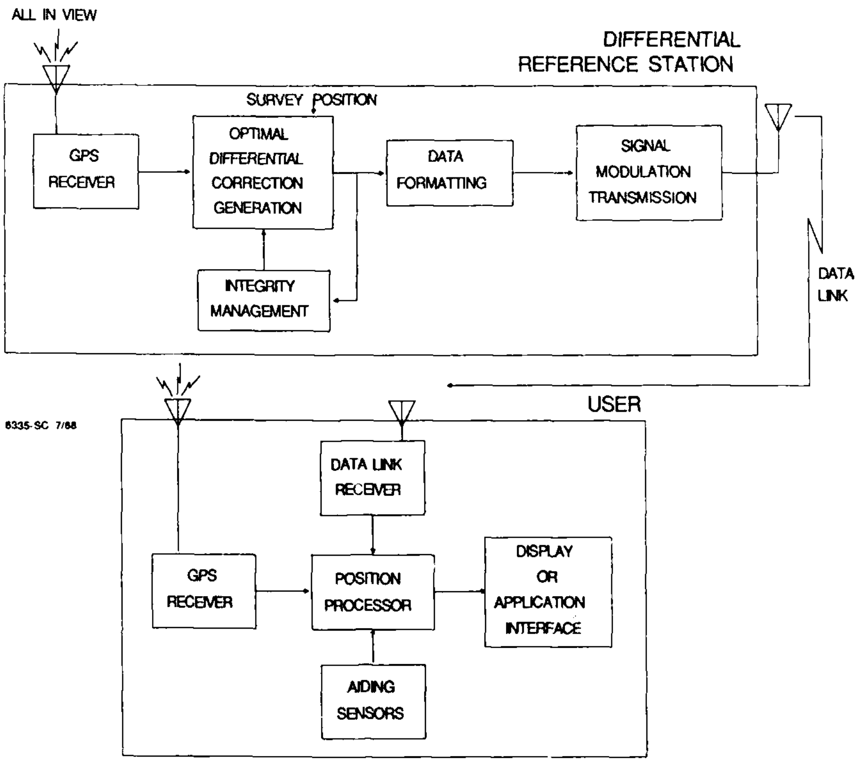


Figure 2. Differential GPS System Architecture

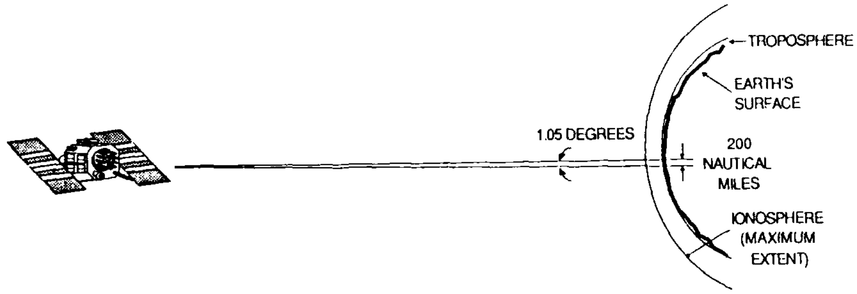


Figure 3. Relative Geometry in Typical Differential GPS Application

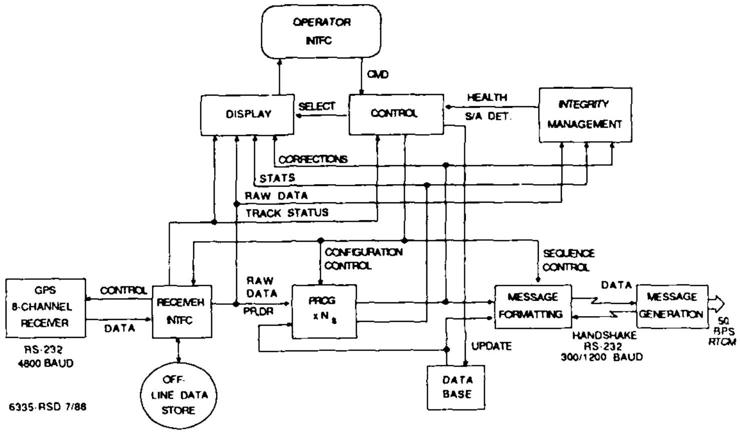


Figure 4. Differential GPS Reference Station Functional Design

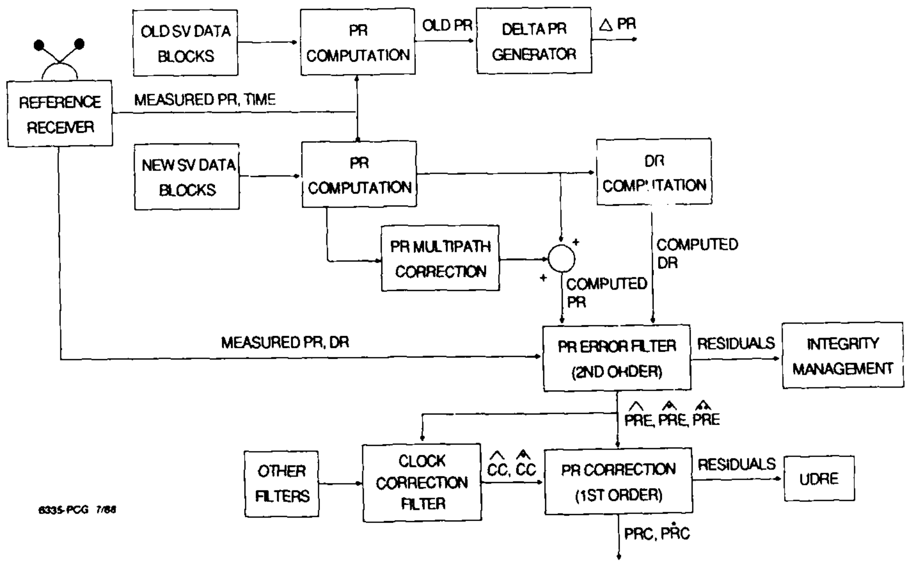


Figure 5. Pseudorange Error Estimation Process in the Reference Station

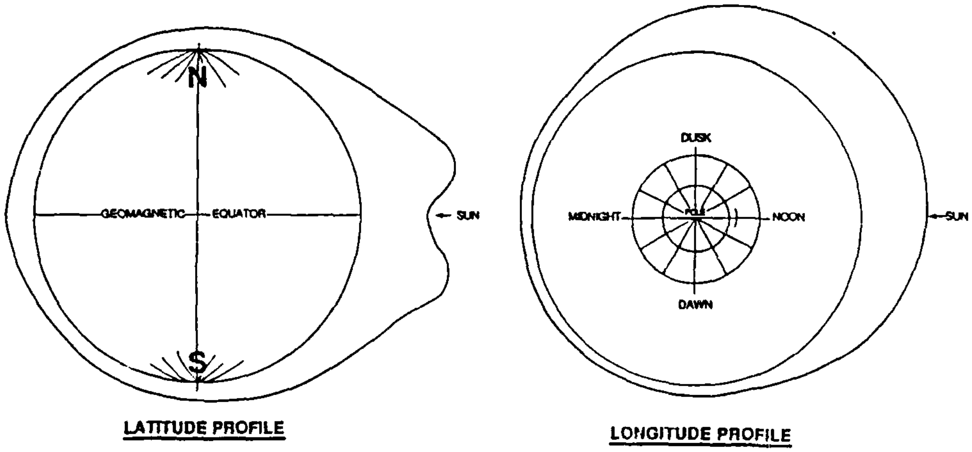
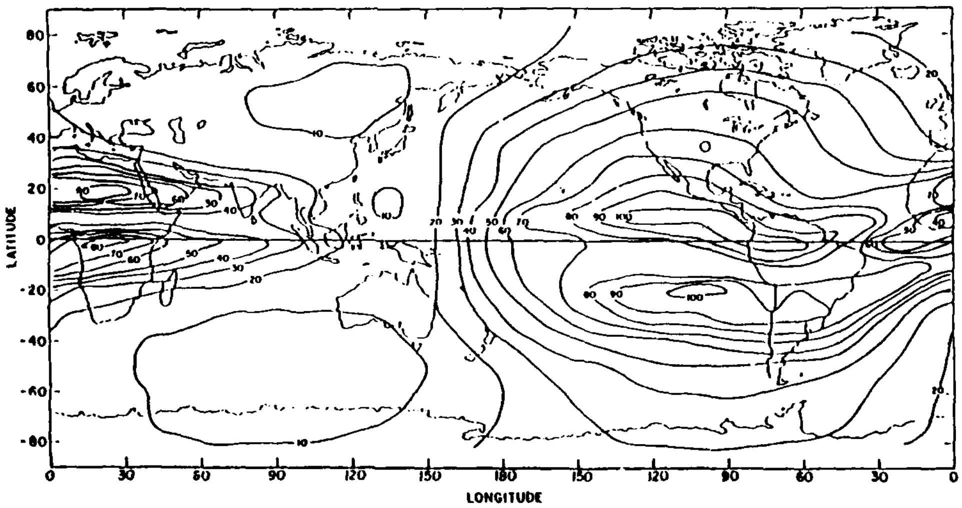


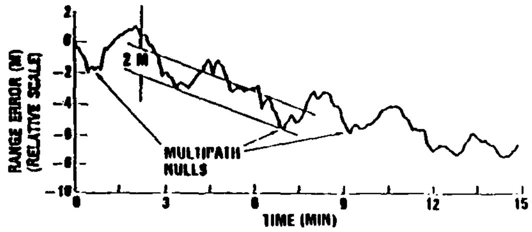
Figure 6. General Ionospheric Layer Shape



CONTOURS OF VERTICAL TEC, IN UNITS OF 10^{16} EL/M^2 COLUMN, FOR 2000 UT, MARCH 1980.

Figure 7. Worldwide Ionospheric Total Electron Density at a Single Point in Time

MULTIPATH CANCELLATION AT SHIPBOARD ANTENNA



Refs.

R. Major, NOSC

G. Bishop, AFGL

MULTIPATH EFFECTS AROUND BUILDINGS

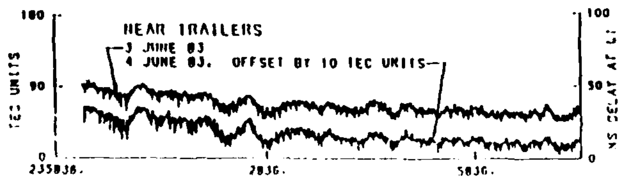
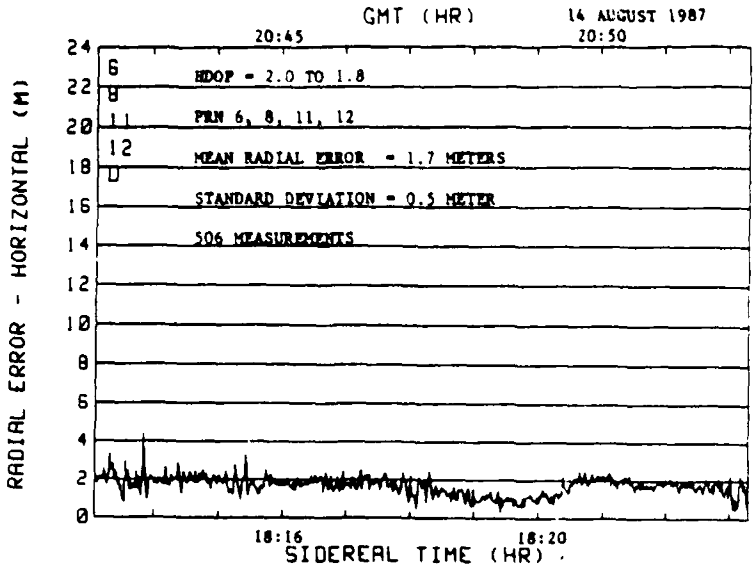


Figure 8. Multipath at Dockside and in a Field



Fi~ure 9. Horizontal Radial Error for U.S. Coast Guard Sea Tests

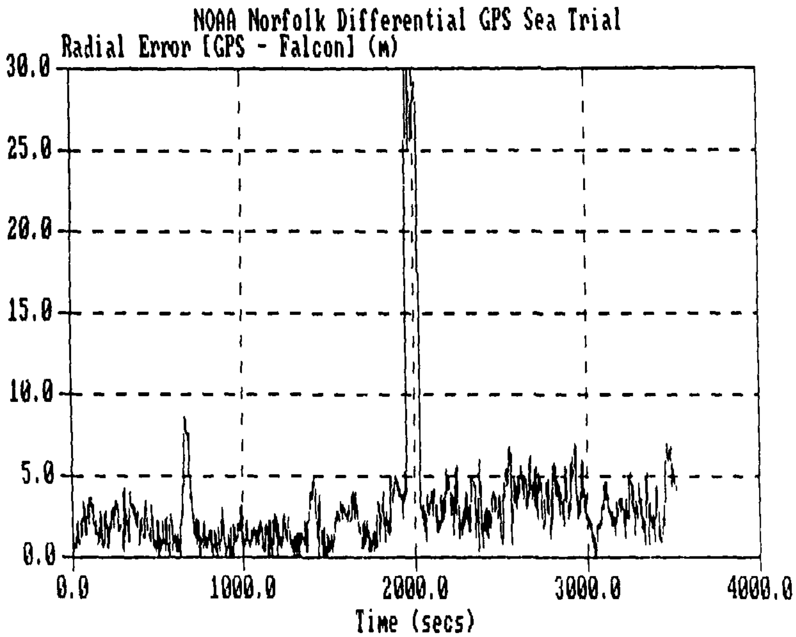


Figure 10. Radial Differential GPS Error During NOAA Sea Test

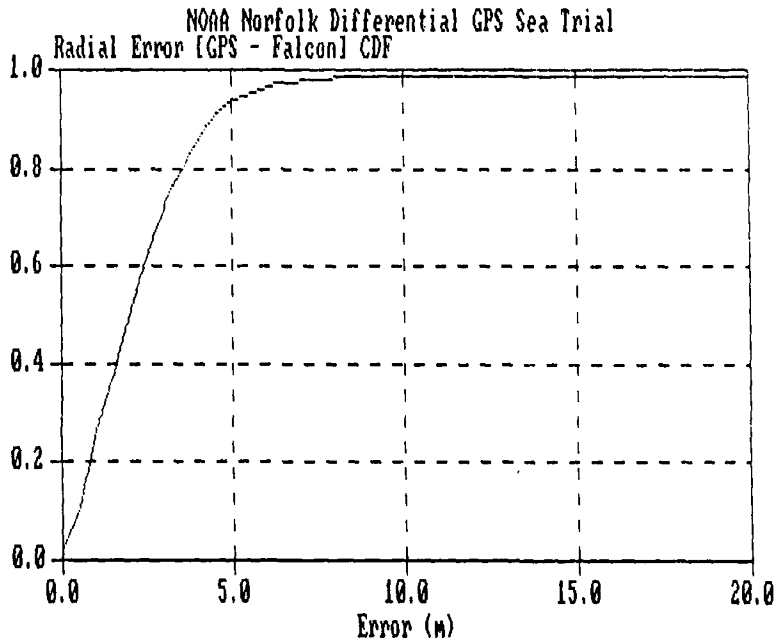


Figure 11. Cumulative Distribution of Radial Errors for NOAA Test

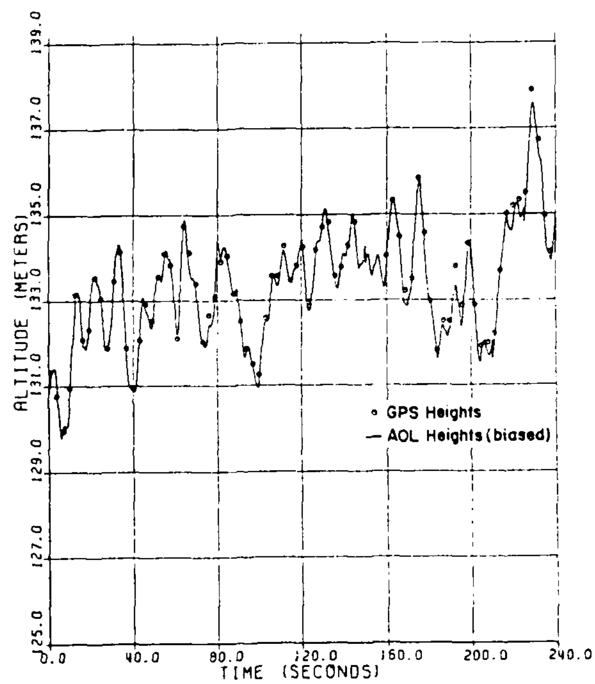


Figure 12. P-3 Kinematic Phase Tracking Performance

GPS NAVIGATION PROCESSING AND KALMAN FILTERING

Robert P. Denaro
Peter V. W. Loomis
TAU Corporation
Los Gatos, California

INTRODUCTION

Measurement processing in the GPS receiver involves two observables from the received signal: code-based pseudorange and carrier-based deltarange or doppler. In a typical receiver, tracking loops are implemented independently for each observable. While code tracking is a digital correlation process, the carrier can be implemented in analog or digital, although most modern receivers employ a digital implementation. The measurement generation process yields a range and range rate to the satellites being tracked.

Conversion of the ranges involves solution of a set of simultaneous quadratic equations, usually accomplished indirectly by least squares or Kalman filtering. The Kalman filter is particularly well-suited to the GPS problem by virtue of its measurement weighting and geometric transformation capability, transforming the range measurements into three orthogonal position coordinates and three velocity coordinates

NAVIGATION MEASUREMENT PROCESSING

Pseudorange Generation

The process of generating a range measurement in the GPS receiver involves precise signal timing. This is accomplished by a correlation process which despreads the received wideband, spread spectrum signal to yield the level of processing gain necessary to track the very low-power GPS signal. The correlation is performed by matching an internally-generated code to the received signal. The resultant time slew (difference from broadcast time) required to achieve correlation establishes the received time relative to the internal oscillator driving the internally-generated code. When differenced with the time of transmission and scaled by the speed of light, the resultant measurement is a "pseudorange." An unknown oscillator offset in the receiver, common to all satellite signals received, is what makes the measurement a pseudorange instead of a pure range. The pseudorange measurement is the primary observable in the GPS signal processing.

A generic block diagram of a GPS receiver is shown in Figure 1. The GPS signal, after reception and intermediate frequency (IF) downconversion, is input to the receiver baseband processing. Many modern receivers convert the final IF stage to a digital signal through analog-to-digital converters, performing the correlation function digitally [1,2]. Early and late replicas of the code generated internally are used for code tracking (delay lock loop) while the prompt version is used for carrier tracking and data demodulation. Some receivers employ extended range detection and tracking which enables generation of a tracking error signal over 1 chip wide (29 meters P-Code, 293 meters C/A-Code) which is useful for maintaining lock under conditions of high error dynamics and narrow tracking loop bandwidths [1,2]. Since phase tracking is commonly performed prior to data demodulation, the Costas form of a phased lock loop is usually used for carrier tracking [3]. The code tracking process yields the pseudorange measurement for navigation processing. The input to the Numerically Controlled Oscillator (NCO) of the carrier tracking loop is typically integrated over a predetermined interval (e.g., 1 second or less) to produce the deltarange measurement for navigation processing.

The code tracking loop is usually mechanized as some form of a delay lock loop (DLL); the error signal can be generated coherently (with the carrier in lock) or noncoherently, depending on the application. When carrier tracking is maintained, many receiver designs aid the code tracking loop with the very precise range rate information derived by carrier tracking. For applications in which an INS is present, INS-derived velocity replaces the carrier aiding signal when carrier tracking cannot be maintained. This simplifies the design of the code loop filter, since the code loop does not have to track host vehicle dynamics.

The code loop correlation process is illustrated in Figure 2. By multiplying the received code by the internally generated code, with a variable offset time, the output of the multiplier is the autocorrelation function of the code itself. The product is essentially zero everywhere but within ±1 code chip interval. Unfortunately, this output is not sufficient for code tracking, because it does not provide information on the sign of the tracking error which is essential for closed loop tracking. Therefore, the delay-lock loop employs an early and late replica of the code through additional multipliers which drive the code NCO to advance or retard the internal code, maintaining code lock. Note that the resultant correlation function now has a "spike" shape between ±1.5 chip widths [4].

The source for the local replica code and carrier timing and other signal processing timing functions is the local oscillator, usually a high quality, temperature-controlled or temperature-compensated quartz oscillator. The oscillator can approach 1 part in 10^9 in stability.

The code loop tracking error is a function of the bandwidth of the tracking loop and the carrier-to-noise ratio of the received signal [5]. Typical code error is on the order of 1-2 meters for most C/A-Code implementations, and about 10% of that for P-Code. Steady-state vehicle dynamics and some signal error sources (such as steady multipath) will produce bias in the code loop error, as well.

Deltarange Generation

The reason the Costas form of the phase lock loop is used for carrier tracking is because the carrier has phase reversals on it from the code. The Costas loop detector is essentially a dual phase-locked loop which ignores the 180-degree phase shifts of the carrier [6].

Carrier tracking is usually performed by a Costas tracking loop, but the more classical Phase Lock Loop (PLL) is also used in some applications. Selection of appropriate bandwidths for the code and carrier tracking loops is, of course, driven by the expected range of dynamics of the host and the anticipated noise environment. Usually, the lowest possible bandwidth which can tolerate the worst-case dynamic transients and maintain signal lock is chosen, to permit generation of the most accurate (i.e., minimum noise content) pseudo and deltarange measurements for navigation processing. However, the stability of the local oscillator can also establish lower limits on the possible tracking loop bandwidth, particularly in very low dynamics, or when the dynamic tracking is provided external to the receiver (e.g., when INS-aided) [7]. The accuracies with which pseudorange and deltarange measurements are derived vary significantly, depending upon the application. Pseudorange errors can vary from the submeter level to several meters, one sigma, while deltarange errors can vary from the subcentimeter level to several centimeters. Generally speaking, carrier tracking will be lost when the signal to noise ratio (C/N_0) falls to 27 dB-Hz (Costas mechanization) or 21 dB-Hz (PLL mechanization). After carrier lock has been lost, code tracking can continue to roughly 12 dB-Hz if INS aiding is present.

In addition to their random walk error characteristics, most quartz oscillators have strong acceleration-induced errors in each of their orthogonal axes. The g-sensitivity of the crystal results in an error in signal phase input to the receiver tracking loops. This is interpreted as vehicle dynamic motion, and may be of sufficient magnitude to cause loss on carrier lock in a narrow bandwidth receiver. Hence, for high acceleration cases, the g-sensitivity of the quartz oscillator may be the primary constraint for minimum bandwidth selection. The result is higher susceptibility to jamming and dynamics-induced noise in the solution [7].

Since vehicle dynamics and other factors can cause problems in maintaining carrier lock compared with the benign case, there is an incentive to incorporate an adaptive tracking loop so that bandwidths can be kept as narrow as possible and maximum antijam performance can be maintained under all conditions. When velocity aiding is available, loop bandwidths can be reduced to minimize interference. Many modern receivers employ such functionality to achieve performance optimization [8,9].

Several types of implementation of carrier-based measurement generation are possible in the receiver, and are used by various manufacturers. The basic observable is the incremental phase of the carrier. This can be continuously observed or sampled, depending on the phase lock technique. A common implementation is to measure the change in equivalent code phase by integrating the carrier phase change over an interval, a deltarange. If the phase is observed directly as a cumulative value, the measurement is an accumulated deltarange (ADR).

Carrier tracking loop errors have similar characteristics to code loop errors, except that, as a percentage of the carrier wavelength versus the code "wavelength" (for C/A-Code, 19 cm vs. 293 m, a ratio of 1/1542), the residual carrier loop tracking errors are very small. For this reason, the carrier-derived deltarange measurement is often used to "smooth" the code measurements. If continuous and contiguous (i.e., no gaps between deltarange measurement intervals) deltarange measurements are available, the effect on smoothing the pseudorange measurements is dramatic. Such carrier smoothing can be accomplished directly in the measurement generation process, or indirectly in the measurement noise specification and process modeling in a Kalman filter.

Navigation Message Decoding

As a final process, the navigation message must be recovered from the remaining carrier signal. The punctual code signal is used for this correlation and the data bit stream is recovered. The data is encoded by phase reversals/non-reversals every 20 C/A-Code lengths, or 50 Hz.

The navigation message is divided into five sequential subframes, each divided into ten words of thirty bits each. Each word has six parity bits, and two words of each subframe are used for timing and message identification. The other eight words in each subframe are used for data communication as follows:

subframe 1: SV Data Block I (clock and ionospheric parameters)
subframe 2: First half of SV Data Block II (Keplerian orbit parameters)
subframe 3: Second half of SV Data Block II (Keplerian orbit parameters)
subframe 4: Message Block
subframe 5: SV Data Block III (constellation health and almanac)

The total navigation message of five subframes takes thirty seconds to transmit. Note that navigation requires subframes 1, 2, and 3 as a minimum to calculate the SV position and clock error. Subframe 4, the message block, is included for future use as a one-way communication device to users. Subframe 5 carries a portion of a large data block, SV Data Block III, which contains coarse orbit, clock, and health information for the whole GPS constellation. Each subframe contains such data for a single satellite; twenty-five subframes are required for the twenty-four satellites in the constellation (one is a dummy subframe of all zero's). To collect a complete almanac data base requires twenty-five navigation message cycles, for a total of 750 seconds. This information is required for knowing which satellites are visible and for quick acquisition of new satellites.

NAVIGATION EQUATIONS

Overview of Navigation Processing

GPS signals are timed at their arrival at the receiver by the code loop correlation process; the total slew of the bit edge that achieves maximum correlation with the incoming code is the time offset from the local oscillator reference time (therefore an arbitrary time). The time of broadcast is contained in the navigation message, which is decoded in the receiver after correlation, as are the polynomial coefficient corrections to this time based on ground observation of each satellite clock that are uploaded along with the new ephemeris predictions.

Multiplication of the calculated time of transit by the speed of transmission (light) results in a measurement of pseudorange. Corrections may be made to this pseudorange for the assumed, modeled, or measured tropospheric and/or ionospheric delays. In standalone operation of C/A or P-Code receivers, the modeled corrections are generally applied. Two-frequency P-Code receivers will often measure the ionospheric delay directly and apply a smoothed value.

With four pseudorange measurements to four different satellites, the next step is to convert the measurements to an absolute geodetic position. Frequently the selected computation coordinate frame is earth-centered, earth-fixed, although earth-centered inertial and local level are also viable alternatives. A direct solution is possible (although there is a solution ambiguity on the other side of the satellites), but does not provide any particular computational advantages. Instead, an indirect solution is usually implemented with an assumed rough location and iterative updates to that location, essentially an application of Newton's root-finding method. This can be implemented as a least squares solution or a Kalman filter.

Basic GPS Navigation Equations

The basic equation relating the line-of-sight pseudorange measurements, PR_i, i=1, ..., 4 satellites, and the satellite positions, S_{xi}, S_{yi}, S_{zi}, to the orthogonal position coordinates, U_x, U_y, U_z is:

$$PR_i = \sqrt{(S_{xi} - U_x)^2 + (S_{yi} - U_y)^2 + (S_{zi} - U_z)^2} + b_u \qquad (1)$$

where

PR_i =	Pseudorange to the i^{th} satellite (measured in the code correlation process)
S_{xi}, S_{yi}, S_{zi} =	Three coordinates of the position of the i^{th} satellite (known from the decoded navigation message)
U_{xi}, U_{yi}, U_{zi} =	Three coordinates of user position (to be solved)
b_u =	User clock offset from GPS time (to be solved)

With four pseudorange measurements there are four simultaneous quadratic equations in four unknowns, the three coordinates of user position and the user's receiver clock bias. Except in unusual geometry conditions, there exists a solution. In practice, there are many computations to be made before arriving at this equation. For example, the satellite positions are broadcast as orbital parameters (ephemerides) as a function of current time. In all, 24 variables must be computed or solved from the available information [10].

Clock and Atmospheric Delay Corrections

The 50 bit-per-second navigation message contains individual satellite corrections for the satellite frequency standard drifts as well as standard corrections for the ionospheric delay (a diurnal and seasonally variable function). Note that receivers authorized to track the P-Code can directly compute the ionospheric delay from the difference in the observed L1 and L2 signals at 1575.42 MHz and 1227.6 MHz, respectively, as they do not require the standard ionospheric correction parameters. These corrections are generally applied to the pseudorange and deltarange measurements before navigation solution processing.

The satellite clock correction is a polynomial fit to the drift computed prior to the last upload by the Control Segment. The receiver computes the current offset using current time, t, and time of applicability, t_{oc}, and the three polynomial coefficients, a_o, a_1, a_2:

$$\Delta t_{SV} = a_0 + a_1(t-t_{0C}) + a_2(t-t_{0C})^2 \qquad (2)$$

The single frequency user must employ a model of the ionosphere since he has no direct measurement of the ionospheric delay. The general daily profile of the ionosphere is low at night, rising sharply at dawn, peaking shortly afternoon, and decaying to a night-time low after dusk. After considerable statistical study, the ionosphere model profile chosen to be supported by GPS for the user community was a simple half-sine curve, centered at 2:00 p.m. local standard time at the ionospheric pierce point with amplitude and period broadcast in SV Data Block I:

$$\text{iono vertical delay in nanoseconds} = \begin{cases} 5.0 + A\cos\omega(t-14{:}00) & (t \text{ in hours}) \\ 5.0 & (\text{if } \cos\omega(t-14{:}00) < 0.0). \end{cases}$$

Five nanoseconds is the general night-time low level. The constants A and ω, roughly on the order of 10-100 nanoseconds and $(12\ \text{hrs.})^{-1}$ respectively, are broadcast as coefficients for cubic equations in (geomagnetic) latitude to better model the latitude dependence of the earth's ionosphere.

When the signal travels at a slant through the ionosphere the distance traveled through the ionosphere is extended and the ionospheric delay increased by an "obliquity factor." The commonly accepted obliquity factor, as a function of satellite elevation angle, multiplies the vertical ionospheric delay to yield

$$\text{total ionospheric delay} = (1.0 + 16.0(0.53 - E)^3)\ \text{vertical delay}.$$

The dual frequency user can compute his ionospheric delay from the different effect that the ionosphere has on the signal at the different frequencies. The effect on the L1 signal is $(77/60)^2$ times the effect on L2 signal, proportional to the square of the wavelength. The iono-compensated signal is thus

$$LC = \frac{(77)^2\ L1 - (60)^2\ L2}{(77)^2 - (60)^2}$$

This compensation formula applies both to carrier and code.

Tropospheric corrections are generally applied by a model within the receiver that is simply a function of elevation angle. External temperature, pressure and humidity measurements can be used to model the tropospheric effects, but stand-alone receivers generally do not use such inputs. There is no universally accepted tropo model, especially at low elevation angles. An approximate expression is:

$$\text{tropospheric delay} = \operatorname{cosec}\theta_0\ (1.4588 + 0.0029611\ N_S)$$

where θ_0 is the elevation angle and N_S = surface refractivity. Refractivity can vary tremendously (250 - 400), so it must be either measured or an average refractivity value computed as a function of latitude and season must be used.

Note that some applications may not require all of these corrections. Under differential GPS operation, for example, the ionospheric delay corrections and tropospheric delay corrections would typically not be applied. This is because the differential correction data are expected to compensate for these effects better than the models or the dual frequency ionospheric measurement.

GPS NAVIGATION KALMAN FILTERING

Overview of Basics

The Kalman filter is a recursive estimator that produces the minimum variance estimate in a least squares sense. Intuitively, the Kalman filter sorts out information and weighs the relative contributions of measurements compared with its assumed model of the process. In a radionavigation system problem, the measurements are the received signals and the process is represented by the assumed model of how the "vehicle" will maneuver in time. For simplicity in describing the basics in this section, the discussion will assume use of the discrete formulation of the Kalman filter [11].

The concept of the covariance of the state is central to the Kalman filter operation. The covariance is the statistical uncertainty in the state, and is propagated from the initial covariance set in the filter. The time propagation enlarges the covariance as confidence in the model of the process wanes with time, and measurements reduce the covariance. In essence, therefore, after a long absence of measurements, a new measurement will be weighted heavily due to the large amount of relative information it provides as an actual observation, however noisy it may be, about the position of the vehicle which, up to receipt of the measurement, had been assumed pursuing a course according to some prestored model.

The basic "process" is the model of how the state transitions over time, from timestep to timestep in the case of a discrete Kalman filter. The measurement is assumed to be a function "h" of the state. These two processes are represented by the following equations:

$$\underline{x}_K = \Phi_{K-1}\,\underline{x}_{K-1} + \underline{w}_{K-1} \qquad (6)$$

$$\underline{z}_K = \underline{h}(\underline{x}_K) + \underline{v}_K$$

The additive terms $\underline{v}_K$ and $\underline{w}_{K-1}$ are Gaussian white noise terms to account for the uncertainty in the state extrapolation and measurement extrapolation models, respectively. The variance of the state noise is Q_K, and the variance of the measurement noise is R_K.

As indicated earlier, the Kalman filter alternates between extrapolation of the state and covariance in time, and update of these variables with new measurement information. Figure 3 is a simplified diagram of the Kalman filter as it processes new measurements and propagates in time. Measurement incorporation is referred to as "updates," and the extrapolation of the state and propagation of the covariance over the time period between measurements are called "propagation."

The state and covariance are propagated according to equations which assume some dynamic model for the state and apply the same assumption to the covariance. The equations describing this propagation or extrapolation of the state and covariance are:

$$\hat{\underline{x}}_K(-) = \Phi_{K-1}\,\hat{\underline{x}}_{K-1}(+) \qquad (7)$$

$$P_K(-) = \Phi_{K-1}\,P_{K-1}(+)\,\Phi^T_{K-1} + Q_{K-1}$$

In these equations, the state "transitions" from the previous time step, k, to a new time step, k + 1. The transition matrix, Φ, determines this temporal relationship. The process noise term, Q, accounts for the error in the modeling assumptions of Φ. The hat notation, ^, means "best estimate," the state output of the filter. The (-) and (+) notation is used in addition to the k, k+1 notation to distinguish between filter estimates immediately before and after a measurement, respectively. In propagating over a timestep, the state or covariance increment their subscript, k, by 1. The notation (-) refers to the state or covariance immediately prior to measurement incorporation, and the notation (+) refers to the state or covariance at the same point in time as (-), but after the inclusion of the measurement information.

Applying this terminology to the state propagation equation as an example, the best estimate of the state ($\underline{x}$) at the current time (k), prior to measurement incorporation (-), is the state transition matrix (Φ) evaluated at the previous time (k-1) times the best estimate of the state at the previous time step ($\hat{\underline{x}}_{K-1}$) and after the last measurement (+).

The measurement "update" is the incorporation of additional information into the filter at some (usually regular) instants in time. The state immediately after measurement incorporation is generally accepted as the most "optimal" state in the filter since it is punctual and has the most recent information (measurement). This step is where the new measurement is compared with the propagated estimate of the state. The difference between them is incorporated in the new state estimate, after scaling by the Kalman gain:

$$\hat{\underline{x}}_K(+) = \hat{\underline{x}}_K(-) + K_K[\underline{z}_K - \underline{h}(\hat{\underline{x}}_K)] \qquad (8)$$

The Kalman gain is a new calculation performed at each measurement update time, based on the propagated covariance from the previous time, $P_K(-)$; the measurement noise covariance, R_K; and the sensitivity of the measurement to small changes in the state, $H_K = \partial\underline{z}_K/\partial\underline{x}_K$:

$$K_K = H^T_K[H_K P_K(-) H^T_K + R_K]^{-1}. \qquad (9)$$

Although this equation seems complex, a simple example will help develop an intuitive feel for this gain calculation. Assume that the state and measurement are in the same coordinate frame, so H is the identity matrix. Suppressing notation and matrix formulation for simplicity, the gain is:

$$K = \frac{P}{P + R} \qquad (10)$$

R is the measurement noise, so for large uncertainty in the state model (represented in P) compared to the uncertainty in the measurement (R), the gain applied to the new measurement is near unity. In other words, given the large uncertainty in the state, the new measurement is assumed to be a much better estimate of where we are than is the unreliable state extrapolation. Conversely, for large uncertainty in the measurement compared with the state, represented by a condition where R >> P, then the Kalman gain is very small.

After the measurement is incorporated, the final step is to update the covariance for the next state update (after the next propagation to the time of that measurement:

$$P_K(+) = [I - K_K H_K] P_K(-) \qquad (11)$$

Again in this equation, the Kalman gain determines the amount by which the covariance (state uncertainty) has been improved by the new measurement. For a large gain, due to a highly confident measurement, the covariance will diminish reflecting the confidence supplied by that measurement. Of course, if the model of the time transition for the covariance is poor, represented by a large Q in equation 6 earlier, then the covariance will grow substantially from this update as the covariance is propagated prior to the next measurement incorporation.

Application of Kalman Filtering to GPS

To cast the Kalman filter equations in GPS form for the unaided receiver, the state vector must be formulated to completely describe the measurements and its own propagation. Typically, an eight state vector is chosen: position (x, y, z), receiver clock phase error (b_u), velocity (V_x, V_y, V_z), and receiver clock frequency error (f_u).

The measurement equations for GPS are developed as follows:

range vector $\underline{R}_i$ $= (S_{xi}-x,\ S_{yi}-y,\ S_{zi}-z)$

range R_i $= |\underline{R}_i|$

pseudorange measurement PR_i $= R_i - b;$

pseudorange rate measurement $= (V_{xi} - V_x,\ V_{yi} - V_y,\ V_{zi} - V_z) \cdot \underline{R}_i / R_i + f$

where S_{xi}, S_{yi}, S_{zi} and V_{xi}, V_{yi}, V_{zi} are the components of the computed satellite position and velocity, respectively. The matrix H has two different types of rows, one for pseudorange

$$((x-S_{xi})/R_i,\ (y-S_{yi})/R_i,\ (z-S_{zi})/R_i,\ 1,0,0,0,0)$$

and a similar row for pseudorange rate:

$$(0,0,0,0,\ (x-S_{xi})/R_i,\ (y-S_{yi})/R_i,\ (z-S_{zi})/R_i,\ 1).$$

The state transition matrix for the dynamics model can take various forms. Generally receivers will model the vehicle motion as a constant velocity. In this case, the Q matrix represents the expected acceleration deviation from this assumption. Obviously, this is a judgment call and becomes a matter of filter tuning.

The propagation equations are thus:

$$\dot{x} = Vx$$

$$\dot{y} = Vy$$

$$\dot{z} = Vz$$

$$\dot{b} = f$$

$$\dot{V}x = 0$$

$$\dot{V}y = 0$$

$$\dot{V}z = 0$$

$$\dot{f} = 0$$

so the Φ matrix is, for a propagation time interval of Δt:

$$\begin{bmatrix} I & I\,\Delta t \\ 0 & I \end{bmatrix}$$

Process noise covariance for this system is: $\begin{bmatrix} Q_v\,\Delta t^3/3 & Q_v\,\Delta t^2/2 \\ Q_v\,\Delta t^2/2 & Q_v\,\Delta t \end{bmatrix}$

where Q_v [12] is a 4 by 4 symmetric matrix with units (meter^2/seconds^3), describing the variance of the uncertainty of velocity propagation per unit of time.

A modification of this formulation is to include 3 acceleration states in addition to the position and velocity states. Although there is no direct measurement of acceleration in the unaided GPS receiver, these augmented states aid the filter in sorting out non-zero mean errors. Specifically, if these states are included, and the vehicle undergoes constant acceleration, the apparent discrepancy in the velocity data will build up as a bias in the acceleration states, and the resultant filter accuracy will improve. In essence, these states represent an unknown bias error in the states, related to the velocity terms by their first difference, so the filter assumes that any such errors belong in these states. Of course, if the acceleration is not constant, the acceleration states will not perfectly track the error, and in fact the filter will respond more sluggishly to the velocity changes. But for the proper case, an aircraft with constant acceleration turns, for example, the extended state filter will outperform the 8-state filter substantially, especially in the turns.

The final area of consideration in GPS Kalman filter design is the use of adaptive tuning in the filter. Specifically, this refers to setting of the process noise, Q, dynamically depending on vehicle motion. The reason to consider this is that the Q term is, at best, a poor substitute for the mismodeling in the state transition model. The modeling errors are not Gaussian noise, but may be biases in turns as already shown. Therefore, the "right" Q value depends on the vehicle profile. For straight and level flight, a small Q is appropriate. For turns or higher dynamics, Q must be larger. For filter stability reasons, Q must be set to the highest level of uncertainty expected. This means that in straight and level flight, for example, the Q will be overly pessimistic and will force processing of too much noise from the measurements due to larger Kalman gains than needed.

The answer is to adapt Q, by setting it as small as possible, then using some other observation to boost Q when needed. Some schemes tried in GPS receivers including making Q a function of the ratio of the observed measurement residuals with the assumed measurement noise. The only danger here is that if Q is allowed to adapt too quickly, the filter can get into a positive feedback loop and cause instability. This happens when observed noise opens Q which creates more noise, etc. The resolution of this problem is to make Q adaptation very slow so that only longer trend conditions cause a change in Q. In practice, the adaptation may be implemented directly on the covariance rather than the Q term, but the effect is similar.

SUMMARY

The processing of GPS signals is a complex process of dealing with two observables, the range and the doppler of the signal. This provides opportunities for substantial sophistication in processing techniques in the tracking loops, and later in the navigation processing of the Kalman filter or other filter. In total, the combination of techniques is a closely interrelated set of algorithms that provide near-optimum performance.

REFERENCES

1. Cox, Jr., D.B., "Integration of GPS with Inertial Navigation Systems," AGARD Conference Proceedings No. 298, France.

2. Vachenlon, J.F., Huneke, A.C., Kaiser, G.M., Forseth, D.C., and Justice, J.H., "GPS Phase III Multi-Channel User Equipment," The Institute of Navigation, Proceedings of the Satellite Division First Technical Meeting, Colorado Springs, Colorado, September 1987.

3. Divakaruni, S.P., Stensland, R.A., and Brenner, M.B., "Ring Laser Gyro Inertial and GPS Integrated Navigation System for Commercial Aviation," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

4. Spilker, J.J., "GPS Signal Structure and Performance Characteristics," Papers published in Journal of Institute of Navigation, Vol. I, 1980.

5. Hemeseth, N.B., "Performance Enhancement of GPS User Equipment," Papers published in Journal of Institute of Navigation, Vol. I, 1980.

6. Eschenbach, R., and Helkey, R., "Performance/Cost Ratio Optimized for GPS Receiver Design," Microwave Systems News, November 1984.

7. Tafel, Jr., R.W., and Krasnjanski, D., "Rapid Alignment of Aircraft Strapdown Inertial Navigation Systems Using Navstar Global Positioning System (GPS)," AGARD Conference Proceedings No. 298, France.

8. Teasley, S.P., "Flight Test Results of an Integrated GPS and Strapdown Inertial System," IEEE PLANS '86 (Position Location and Navigation Symposium), IEEE AES Society, Las Vegas, Nevada, November 1986.

9. Schwartz, H.M., Sensitivity Analysis of an Integrated Navstar GPS/INS Navigation System to Component Failure," Journal of the Institute of Navigation, Vol. 30, No. 4, Winter 1983-84.

10. Van Dierendonck, A.J., et al, "The GPS Navigation Message," Papers published in Journal of Institute of Navigation, Vol. I, 1980.

11. Gelb, A., Applied Optimal Estimation, The Analytic Sciences Corporation, Technical Staff, The M.I.T. Press, Cambridge, Massachusetts, 1979.

12. Kao, M.H., and Eller, D.H., "Multiconfiguration Kalman Filter Design for High-Performance GPS Navigation," Kalman Filtering: Theory and Application, edited by Harold W. Sorenson, IEEE Press, New York, 1985.

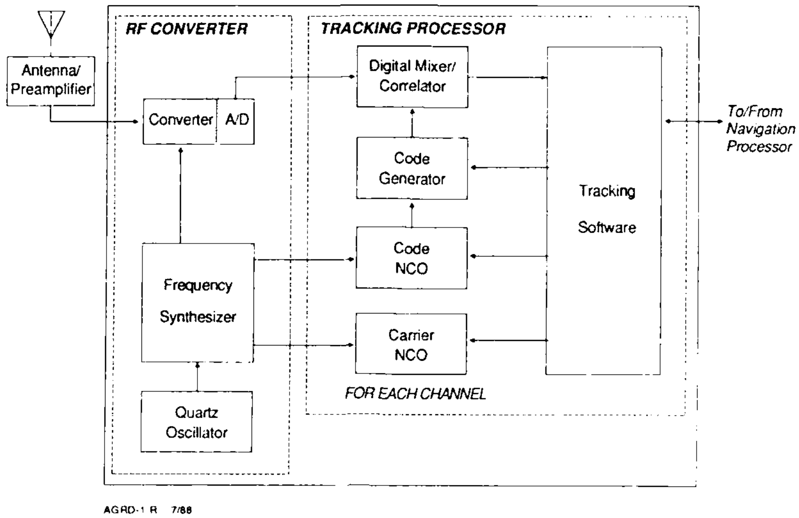


Figure 1. Generic Block Diagram for a GPS Receiver

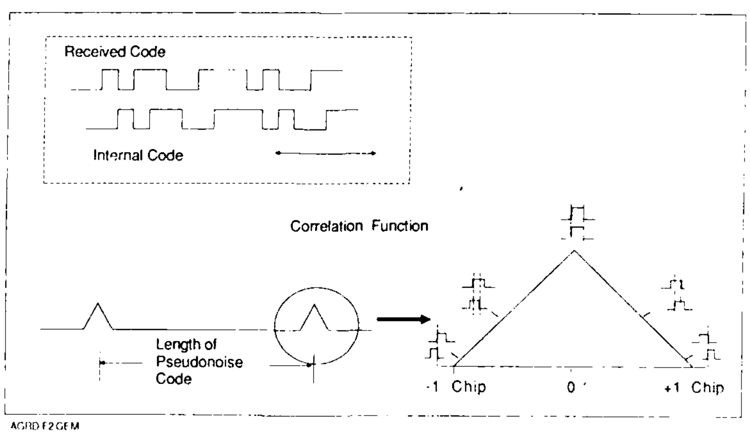
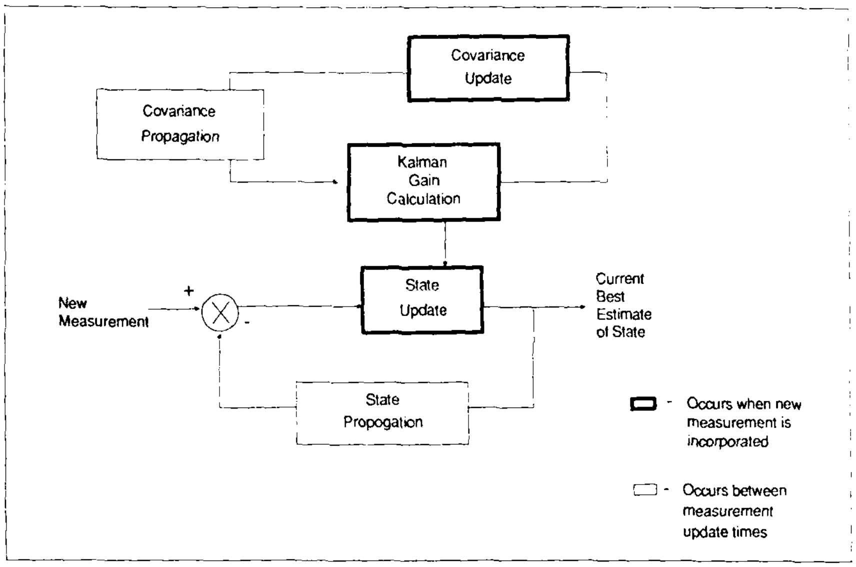


Figure 2 Code Loop Correlation Process



AGRD-F3.CEM

Figure 3. Simplified Diagram of a Kalman Filter

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-LS-161	ISBN 92-835-0477-1	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title

THE NAVSTAR GPS SYSTEM

7. Presented at

8. Author(s)/Editor(s)	9. Date
Various	September 1988

10. Author's/Editor's Address	11. Pages
Various	144

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

Navigation satellites
Position indicators
Navigation

14. Abstract

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

AGARD Lecture Series No.161
Advisory Group for Aerospace Research and Development, NATO
THE NAVSTAR GPS SYSTEM
Published September 1988
144 pages

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

ISBN 92-835-0477-1

AGARD-LS-161

Navigation satellites
Position indicators
Navigation

AGARD Lecture Series No.161
Advisory Group for Aerospace Research and Development, NATO
THE NAVSTAR GPS SYSTEM
Published September 1988
144 pages

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

ISBN 92-835-0477-1

AGARD-LS-161

Navigation satellites
Position indicators
Navigation

AGARD Lecture Series No.161
Advisory Group for Aerospace Research and Development, NATO
THE NAVSTAR GPS SYSTEM
Published September 1988
144 pages

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

ISBN 92-835-0477-1

AGARD-LS-161

Navigation satellites
Position indicators
Navigation

AGARD Lecture Series No.161
Advisory Group for Aerospace Research and Development, NATO
THE NAVSTAR GPS SYSTEM
Published September 1988
144 pages

A comprehensive and up-to-date Lecture Series on the Navstar GPS System is presented. Among the subjects covered are the system concept, navigation principles, system design, user applications and technologies involved. The Lectures are a balanced presentation of theory, application and measured performance.

This Lecture Series, sponsored by the Avionics Panel of AGARD, has been implemented by the Consultant and Exchange Programme.

ISBN 92-835-0477-1

AGARD-LS-161

Navigation satellites
Position indicators
Navigation