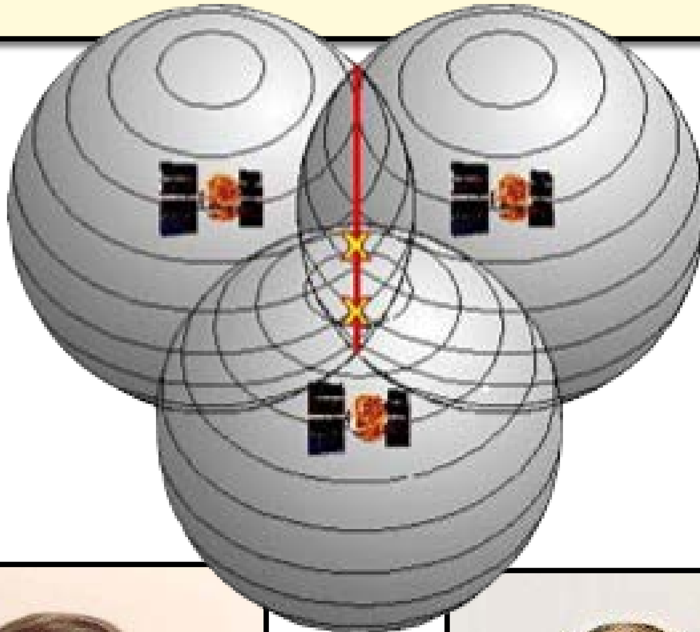
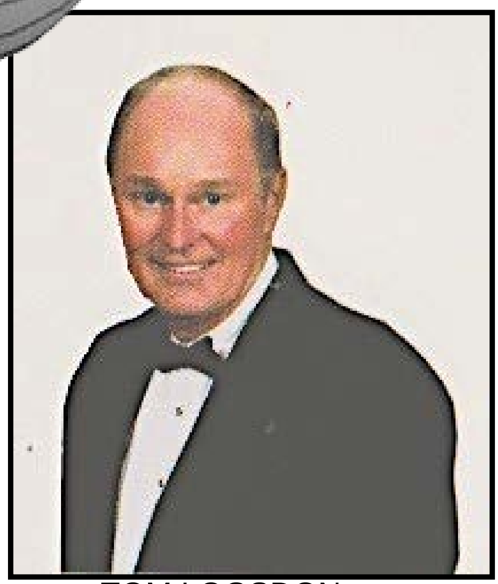
THE GLOBAL POSITIONING SYSTEM — A NATIONAL TREASURE



BOB NELSON



TOM LOGSDON

The Global Positioning System - A National Resource

Note Tom Logsdon is a respected expert on GPS and other satellites systems who teaches several courses for ATIcourses including GPS & Other Radionavigation Systems, Fundamentals of Orbital & Launch Mechanics, Integrated Navigation Systems, and Introduction to Space.

Written by Robert A. Nelson and Updated by Tom Logsdon (May 2013)

The Global Positioning System (GPS) was originally designed jointly by the U.S. Navy and the U.S. Air Force to permit the determination of position and time for military troops and guided missiles. However, GPS has also become the basis for position and time measurement by scientific laboratories and a wide spectrum of applications in a multi-billion dollar commercial industry. Roughly three billion GPS receivers have been sold to delighted consumers throughout the world. Thirty-one GPS satellites are currently broadcasting navigation signals from their high-altitude vantage points in space.

EARLY METHODS OF NAVIGATION

The shape and size of the earth has been known from the time of antiquity. The fact that the earth is a sphere was well known to educated people as long ago as the fourth century BC. In his book On the Heavens, Aristotle gave two scientifically correct arguments. First, the shadow of the earth projected on the moon during a lunar eclipse appears to be curved. Second, the elevations of stars change as one travels north or south, while certain stars visible in Egypt cannot be seen at all from Greece.

The actual radius of the earth was determined within one percent by Eratosthenes in about 230 BC. He knew that the sun was directly overhead at noon on the summer solstice in Syene (Aswan, Egypt), since on that day it illuminated the water of a deep well. At the same time, he measured the length of the shadow cast by a column on the grounds of the library at Alexandria, which was nearly due north. The distance between Alexandria and Syene had been well established by professional runners and camel caravans. Thus Eratosthenes was able to compute the earth's radius from the difference in latitude that he inferred from his measurement. In terms of modem units of length, he arrived at the figure of about 6400 km. By comparison, the actual mean radius is 6371 km (the earth is not precisely spherical, as the polar radius is 21 km less than the equatorial radius of 6378 km). The ability to determine one's position on the earth was the next major problem to be addressed. In the second century, AD the Greek astronomer Claudius Ptolemy prepared a geographical atlas, in which he estimated the latitude and longitude of principal cities of the Mediterranean world. Ptolemy is most famous, however, for his geocentric theory of planetary motion, which was the basis for astronomical catalogs until

Nicholas Copernicus published his heliocentric theory in 1543.

CELESTIAL NAVIGATION

Historically, methods of navigation over the earth's surface have involved the angular measurement of star positions to determine latitude. The latitude of one's position is equal to the elevation of the pole star. The position of the pole star on the celestial sphere is only temporary, however, due to precession of the earth's axis of rotation through a circle of radius 23.5 over a period of 26,000 years. At the time of Julius Caesar, there was no star sufficiently close to the north celestial pole to be called a pole star. In 13,000 years, the star Vega will be near the pole. It is perhaps not a coincidence that mariners did not venture far from visible land until the era of Christopher Columbus, when true north could be determined using the star we now call Polaris. Even then the star's diurnal rotation caused an apparent variation of the compass needle. Polaris in 1492 described a radius of about 3.5 degrees about the celestial pole, compared to today. At sea, however, Columbus and his contemporaries depended primarily on the mariner's compass and dead reckoning.

The determination of longitude was much more difficult. Longitude is obtained astronomically from the difference between the observed time of a celestial event, such as an eclipse, and the corresponding time tabulated for a reference location. For each hour of difference in time, the difference in longitude is 15 degrees.

NAVIGATION AT SEA

Columbus himself attempted to estimate his longitude on his fourth voyage to the New World by observing the time of a lunar eclipse as seen from the harbor of Santa Gloria in Jamaica on February 29, 1504. In his distinguished biography Admiral of the Ocean Sea, Samuel Eliot Morrison states that Columbus measured the duration of the eclipse with an hour-glass and determined his position as nine hours and fifteen minutes west of Cadiz, Spain, according to the predicted eclipse time in an almanac he carried aboard his ship. Over the preceding year, while his ship was marooned in the harbor, Columbus had determined the latitude of Santa Gloria by numerous observations of the pole star. He made out his latitude to be 18 degrees, which was in error by less than half a degree and was one of the best recorded observations of latitude in the early sixteenth century, but his estimated longitude was off by some 38 degrees.

Columbus also made legendary use of this eclipse by threatening the natives with the disfavor of God, as indicated by a portent from Heaven, if they did not bring desperately needed provisions to his men. When the eclipse arrived as predicted, the natives pleaded for the Admiral's intervention, promising to furnish all the food that was needed.

New knowledge of the universe was revealed by Galileo Galilei in his book *The Starry Messenger*.

This book, published in Venice in 1610, reported the telescopic discoveries of hundreds of new stars, the craters on the moon, the phases of Venus, the rings of Saturn, sunspots, and the four inner satellites of Jupiter. Galileo suggested using the eclipses of Jupiter's satellites as a celestial clock for the practical determination of longitude, but the calculation of an accurate ephemeris and the difficulty of observing the satellites from the deck of a rolling ship prevented use of this method at sea. Nevertheless, James Bradley, the third Astronomer Royal of England, successfully applied the technique in 1726 to determine the longitudes of Lisbon and New York with considerable accuracy.

The inability to measure longitude at sea had the potential of catastrophic consequences for sailing vessels exploring the new world, carrying cargo, and conquering new territories. Shipwrecks were common. On October 22, 1707 a fleet of twenty-one ships under the command of Admiral Sir Cloudsley Shovel was returning to England after an unsuccessful military attack on Toulon in the Mediterranean. As the fleet approached the English Channel in dense fog, the flagship and three others foundered on the coastal rocks and nearly two thousand men perished.

Stunned by this unprecedented toss, the British government in 1714 offered a prize of 20,000 British Pounds for a method to determine longitude at sea within a half a degree. The scientific establishment believed that the solution would be obtained from observations of the moon.

The German cartographer Tobias Mayer, aided by new mathematical methods developed by Leonard Euler, offered improved tables of the moon in 1757. The recorded position of the moon at a given time as seen from a reference meridian could be compared with its position at the local time to determine the angular position west or east. Just as the astronomical method appeared to achieve realization, the British craftsman John Harrison provided a different solution through his invention of the marine chronometer. The story of Harrison's clock has been recounted in Dava Sobel's popular book, *Longitude.*

Both methods were tested by sea trials. The lunar tables permitted the determination of longitude within four minutes of arc, but with Harrison's chronometer the precision was only one minute of arc. Ultimately, portions of the prize money were awarded to Mayer's widow, Euler, and Harrison. In the twentieth century, with the development of radio transmitters, another class of navigation aids was created using terrestrial radio beacons, including Loran and Omega. Finally, the technology of artificial satellites made possible navigation and position determination using line of sight signals involving the measurement of Doppler shift or phase difference.

GLOBAL POSITIONING SYSTEM

The success of Transit stimulated both the U.S. Navy and the U.S. Air Force to investigate more advanced versions of a space-based navigation system with enhanced capabilities. Recognizing the need for a combined effort, the Deputy Secretary of Defense established a

Joint Program Office in 1973. The NAVSTAR Global Positioning System (GPS) was thus created.

In contrast to Transit, GPS provides continuous coverage. Also, rather than Doppler shift, satellite range is determined from phase difference. There are two types of observables. One is pseudorange, which is the offset between a pseudorandom noise (PRN) coded signal from the satellite and a replica code generated in the user's receiver, multiplied by the speed of light. The other is accumulated delta range (ADR), which is a measure of carrier phase.

THE NAVSTAR GPS CONSTELLATION

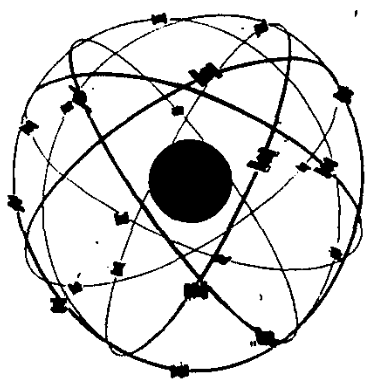


Figure 1. The original GPS constellation reached operational status in 1995. It consisted of 24 GPS satellites arranged in six orbital rings 10,898 nautical miles above the Earth. Each of the rings was tipped 55 degrees with respect to the equator. More than three billion satisfied users now benefit from the GPS signals streaming down from space.

The determination of position may be described as the process of triangulation using the measured range between the user and *four* or more satellites. The ranges are inferred from the time of propagation of the satellite signals. Four satellites are required to determine the three coordinates of position and time. The time is involved in the correction to the receiver clock and is ultimately eliminated from the measurement of position.

High precision is made possible through the use of atomic clocks carried on-board the satellites. Each satellite has two cesium clocks and two rubidium clocks, which maintain time with a precision one part in ten trillionth in over a few hours, or better than 1O nanoseconds. In terms of the distance traversed by an electromagnetic signal at the speed of light, each nanosecond corresponds to about 30 centimeters. Thus the precision of GPS clocks permits a real time measurement of distance to within a few meters. With post processed carrier phase measurements, a precision of a few centimeters can be achieved today.

The design of the GPS constellation had the fundamental requirement that at least four satellites must be visible at all times from any point on earth. The tradeoffs included visibility, the need to pass over the ground control stations in the United States, cost, and sparing efficiency. The orbital configuration approved in 1973 was a total of 24 satellites, consisting of 8 satellites plus one spare in each of three equally spaced orbital planes. The orbital radius was 26,562 km, corresponding to a period of revolution of 12 sidereal hours, with repeating ground traces. Each satellite arrived over a given point four minutes earlier each day. A common orbital inclination of 63° was selected to maximize the on-orbit payload mass with The operational system, as presently deployed, consists of 21 primary satellites and 3 on-orbit spares, comprising *four* satellites in each of six orbital planes. Each orbital plane is inclined at 55° with respect to the equator. This constellation improves on the "18 plus 3" satellite constellation by more fully integrating the three active spares.

There have been several generations of GPS satellites. The Block I satellites, built by Rockwell International, were launched between 1978 and 1985. They consisted of eleven prototype satellites, including one launch failure, that validated the system concept. The ten successful satellites had an average lifetime of 8.76 years.

The Block II and Block IIA satellites were also built by Rockwell International. Block II consists of nine satellites launched between 1989 and 1990. Block IIA, deployed between 1990 and 1997, consists of 19 satellites with several! navigation enhancements. In April 1995, GPS was declared fully operational with a constellation of 24 operational spacecraft and a completed ground segment. The 28 Block II/IIA satellites have exceeded their specified mission duration of 6 years and are expected to have an average lifetime of more than 1O years. Block IIR comprises 20 replacement satellites that incorporate autonomous navigation based on cross-link ranging. These satellites are being manufactured by Lockheed Martín. The first launch in 1997 resulted in a launch failure. The first IIR satellite to reach orbit was also launched in 1997. The second GPS IIR satellite was successfully launched aboard a Delta 2 rocket on October 7, 1999. One to four more launches are anticipated over the next year. The fourth generation of satellites is the Block II follow-on (Block IIF). This program includes the procurement of 33 satellites and the operation and support of a new GPS operational control segment. The Block IIF program was awarded to Rockwell (now a part of Boeing). Further details may be found in a special issue of the Proceedings of the IEEE for January, 1999.

CONTROL SEGMENT

The Master Control Station for GPS is located at Schriever Air Force Base in Colorado Springs, CO. The MCS maintains the satellite constellation and performs the station keeping and attitude control maneuvers. It also determines the orbit and clock parameters with a Kalman filter using measurements from five monitor stations distributed around the world. The orbit error is about 1.5 meters.

GPS orbits are derived independently by various scientific organizations using carrier phase and post-processing. The state of the art is exemplified by the work of the International GPS Service (IGS), which produces orbits with an accuracy of approximately 3 centimeters within two weeks. The system time reference is managed by the U.S. Naval Observatory in Washington, DC. GPS time is measured from Saturday/Sunday midnight at the beginning of the week. The GPS time scale is a composite "paper clock" that is synchronized to keep step with Coordinated Universal Time (UTC) and International Atomic Time (TAI). However, UTC differs from TAI by an integral number of leap seconds to maintain correspondence with the rotation of the earth, whereas GPS time does not include leap seconds. The origin of GPS time is midnight on January 5/6, 1980 (UTC). At present, TAI is ahead of UTC by 32 seconds, TAI is ahead of GPS by 19 seconds, and GPS is ahead of UTC by 13 seconds.

Only 1,024 weeks were allotted from the origin before the system time is reset to zero because 1O bits are allocated for the calendar function (1,024 is the tenth power of 2). Thus the first GPS rollover occurred at midnight on August 21, 1999. The next GPS rollover will take place May 25, 2019.

SIGNAL STRUCTURE

The satellite position at any time is computed in the user's receiver from the navigation message that is contained in a 50 bit per second data stream. The orbit is represented for each one hour period by a set of 15 Keplerian orbital elements, with harmonic coefficients arising from perturbations, and is updated every four hours.

This data stream is modulated by each of two code division multiple access, or spread spectrum, pseudorandom noise (PRN) codes: the coarse/acquisition C/A code (sometimes called the clear/access code) and the precision P code. The P code can be encrypted to produce a secure signal called the Y code. This feature is known as the Anti-Spoof (AS) mode, which is intended to defeat deception jamming by adversaries. The C/A code is used for satellite acquisition and for position determination by civil receivers. The P(Y) code is used by military and other authorized receivers. The C/A code is a Gold code of register size 10, which has a sequence length of 1023 chips and a chipping rate of 1.023 MHz and thus repeats itself every 1 millisecond. (The term "chip" is used instead of "bit' to indicate that the PRN code contains no information.) The P code is a long code of length 2.3547 x 1014 chips with a chipping rate of 10 times the C/A code of 10.23 MHz. At this rate the P code has a period of 38.058 weeks, but it is truncated on a weekly basis so that 38 segments are available for the constellation. Each satellite uses a different member of the C/A Gold code family and a different one-week segment of the P code sequence.

The GPS satellites transmit signals at two carrier frequencies: the L1 component with a center frequency of 1575.42 MHz, and the L2 component with a center frequency of 1227.60 MHz. These frequencies are derived from the master clock frequency of 10.23 MHz, with L1 = 154 x 10.23 MHz and L2 = 120 x 10.23 MHz. The L1 frequency transmits both the P code and the C/A code, while the L2 frequency transmits only the P code. The second P code frequency permits a dual-frequency measurement of the ionospheric group delay. The P-code receiver has- a two-sigma root-mean-square horizontal position error of about 5 meters.

The single frequency C/A code user must model the ionospheric delay with less accuracy. In addition, the C/A code is intentionally degraded by a technique called Selective Availability (SA}, which introduces errors of 50 to 100 meters by dithering the satellite clock data.

Through differential GPS measurements, however, position accuracy can be improved by reducing selective availability and environmental errors. The transmitted signal from a GPS satellite has right hand circular polarization. According to the GPS Interface Control Document, the specified minimum signal strength at an elevation angle of 5 degrees into a linearly polarized receiver antenna with a gain of 3dB (approximately equivalent to a circularly polarized antenna with a gain of O dB) is - 160 dBW for the L1 C/A code, - 163 dBW far the L1 P code, and - 166 dBW for the L2 P code. The L2 signal is transmitted at a lower power level since it is used primarily for the ionospheric delay correction.

PSEUDORANGE

The fundamental measurement in the Global Positioning System is pseudo-
range. The user equipment receives the pseudorandom code from a satellite and, having identified the satellite, generates a replica code. The phase by which the replica code must be shifted in the receiver to maintain maximum correlation with the satellite code, multiplied by the speed of light, is approximately equal to the satellite range. It is called the pseudorange because the measurement must be corrected by a variety of factors to obtain the true range.

The corrections that must be applied include signal propagation delays caused by the ionosphere and the troposphere, the space vehicle clock error, and the user's receiver clock error. The ionosphere correction is obtained either by measurement of dispersion using the two frequencies L1 and L2 or by calculation from a mathematical model, but the tropospheric delay must be calculated since the troposphere is non dispersive. The true geometric distance to each satellite is obtained by applying these corrections to the measured pseudo-
range.

Other error sources and modeling errors continue to be investigated. For example, a recent modification of the Kalman filter has led to improved performance. Studies have also shown that solar radiation pressure models may need revision and there is some new evidence that the earth's magnetic field may contribute to a small orbit period variation in the satellite clock frequencies.

CARRIER PHASE

Carrier phase is used to performance measurements with a precision that greatly exceeds those based on pseudorange. However, a carrier phase measurement must resolve an integral cycle ambiguity, whereas the pseudorange is unambiguous.

The wavelength of the L1 carrier is about 19 centimeters. Thus with a cycle resolution of one percent, a differential measurement at the level of a few millimeters is theoretically possible. This technique has important applications to geodesy and analogous scientific programs.

RELATIVITY

The precision of GPS measurements is so great that it requires the application of Albert Einstein's special and general theories of relativity for the reduction of its measurements. Professor Carroll Alley of the University of Maryland once articulated the significance of this fact at a scientific conference devoted to time measurement in 1979. He said, "I think it is appropriate to realize that the first practical application of **Einstein's ideas in actual** engineering situations are with us in the fact that clocks are now so stable that one must take these small effects into account in a variety of systems that are now undergoing development or are actually in use in comparing time worldwide. It is no longer a matter of scientific interest and scientific application, but it has moved into the realm of engineering necessity."

According to relativity theory, a moving clock appears to run slow with respect to a similar clock that is at rest. This-effect is called "time dilation." In addition, a clock in a weaker gravitational potential appears to run fast in comparison to one that is in a stronger gravitational potential. This gravitational effect is known in general as the "red shift" (only in this case it is actually a "blue shift").

GPS satellites revolve around the earth with a velocity of 3.874 km/s at an altitude of 20,184 km. Thus on account of the its velocity, a satellite clock appears to run slow by 7 microseconds per day when compared to a clock on the earth's surface. But on account of the difference in gravitational potential, the satellite clock appears to run fast by 45 microseconds per day. The net effect is that the clock appears to run fast by 38 microseconds per day. This is an enormous rate difference for an atomic clock with a precision of a few nanoseconds. Thus to compensate for this large secular rate, the clocks are given a rate offset prior to satellite launch of - 4.465 parts in 10 to the tenth power from their nominal frequency of 10.23 MHz so that on average they appear to run at the same rate as a clock on the ground. The actual frequency of the satellite clocks before launch is thus 10.22999999543 MHz. Although the GPS satellite orbits are nominally circular, there is always some residual eccentricity. The eccentricity causes the orbit to be slightly elliptical, and the velocity and altitude vary over one revolution. Thus, although the principal velocity and gravitational effects have been compensated by a rate offset, there remains a slight residual variation that is proportional to the eccentricity. For example, with an orbital eccentricity of 0.02 there is a relativistic sinusoidal variation in the apparent clock time having an amplitude of 46 nanoseconds. This correction must be calculated and taken into account in the GPS receiver.

The displacement of a receiver on the surface of the earth due to the earth's rotation in inertial space during the time of flight of the signal must also be taken into account. This is a third relativistic effect that is due to the universality of the speed of light. The maximum correction occurs when the receiver is on the equator and the satellite is on the horizon. The time of flight of a GPS signal from the satellite to a receiver on the earth is then 86 milliseconds and the correction to the range measurement resulting from the receiver displacement is 133 nanoseconds. An analogous correction must be applied by a receiver on a moving platform, such as an aircraft or another satellite. This effect, as interpreted by an observer in the rotating frame of reference of the earth, is called the Sagnac effect. It is also the basis for a laser ring gyro in an inertial navigation system.

GPS MODERNIZATION

In 1996, a Presidential Decision Directive stated the president would review the issue of Selective Availability in 2000 with the objective of discontinuing selective availability no later than 2006. In addition, both the L1 and L2 GPS signals would be made available to civil users and a new civil 10.23 MHz signal would be authorized. To satisfy the needs of aviation, the third civil frequency, known as L5, would be centered at 1176.45 MHz, in the Aeronautical Radio Navigation Services (ARNS) band, subject to approval at the World Radio Conference in 2000. According to Keith McDonald in an article on GPS modernization published in the September, 1999 GPS World, with selective availability removed, the civil GPS accuracy would be improved to about 1O to 30 meters. With the addition of a second frequency for ionospheric group delay corrections, the civil accuracy would become about 5 to 10 meters. A third frequency would permit the creation of two beat frequencies that would yield one-meter accuracy in real time.

A variety of other enhancements are under consideration, including increased power, the addition of a new military code at the L1 and L2 frequencies, additional ground stations, more frequent uploads, and an increase in the number of satellites. These policy initiatives are driven by the dual needs of maintaining national security while supporting the growing dependence on GPS by commercial industry. When these upgrades would begin to be implemented in the Block IIR and IIF satellites depends on GPS funding.

Besides providing position, GPS is a reference for time with an accuracy of 10 nanoseconds or better. Its broadcast time signals are used for national defense, commercial, and scientific purposes. The precision and universal availability of GPS time has produced a paradigm shift in time measurement and dissemination, with GPS evolving from a secondary source to a fundamental reference in itself.

The international community wants assurance that it can rely on the availability of GPS and continued U.S. support for the system. The Russian Global Navigation Satellite System (GLONASS) has been an alternative, but economic conditions in Russia have threatened its continued viability. Consequently, the European Union is considering the creation! of a navigation system of its own, called Galileo, to avoid relying on the U.S. GPS and Russian GLONASS programs.

The Global Positioning System is a vital national resource. Over the past thirty years it has made the transition from concept to reality, representing today an operational system on which the entire world has become dependent. Both technical improvements and an enlightened national policy will be necessary to ensure its continued growth into the twenty first century.

Dr. Robert A. Nelson, P.E. was president of Satellite Engineering Research Corporation in Bethesda, Maryland, a Lecturer in the Department of Aerospace Engineering at the University of Maryland and Technical Editor of Via Satellite magazine. Dr. Nelson was the instructor for the ATI course Satellite Communications Systems Engineering for more than 20 years. Dr. Nelson passed away in May 2013. He will be remembered and missed for his many contributions to the field of Satellite Engineering.

Based on an article originally published in Via Satellite. Updated on May 28, 2013 Tom Logsdon has lectured extensively and has taught 300 short courses on a variety of technical topics in 31 different countries scattered across six continents. He as written and sold 1.8 million words including 32 nonfiction books. His words; spoken and written, have been translated into a dozen different languages including French, Spanish, Serbo-Croatian, Russian, Latvian, Japanese, and International Sign Language. Tom is an expert on GPS and other navigation satellites who teaches several courses for ATIcourses including GPS & Other Radio navigation Satellites ,Fundamentals of Orbital & Launch Mechanics, Integrated Navigation Systems , and Introduction to Space.

About Applied Technology Institute Courses (ATIcourses or ATI) ATIcourses is a national leader in professional development seminars in the technical areas of space, communications, defense, sonar, radar,engineering, and signal processing. Since 1984, ATIcourses has presented leading-edge technical training to defense and NASA facilities, as well as DOD and aerospace contractors. ATI courses create a clear understanding of the fundamental principles and a working knowledge of current technology and applications. ATI offers customized on-site training at your facility anywhere in the United States, as well as internationally, and over 200 annual public courses in dozens of locations. ATI is proud to have world-class experts instructing courses. Call 410-956- 8805/888-501-2100, or visit them on the web at www.ATIcourses.com.

CONTACT: Jim Jenkins Ph: (888) 501-2100 Ph: (410) 956-8805 .