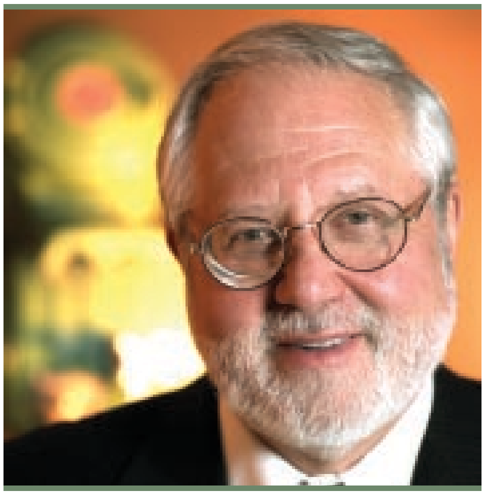
Equipping GPS Satellites with Accelerometers and Satellite-to-Satellite Observables

Proceedings of the 2002 ION National Technical Meeting, San Diego, CA, January 28-30, 2002

Michael E. Ash



ABSTRACT

Millimeter-level accuracy applications of terrestrial Global Positioning System (GPS) phase tracking can be done more robustly with global rather than regional accuracy if the GPS satellites are equipped with (1) accelerometers to measure rather than model nongravitational accelerations, (2) satellite-to-satellite phase tracking as well as ranging cross links to work around inadequately modeled atmospheric propagation effects, and (3) gyroscopes for better satellite attitude control. More complete general relativity modeling is also needed in the orbit fitting and site coordinate estimation process, for which it could be advantageous to employ a combination of Kalman filtering on the orbital motions and maximum likelihood estimation on the site coordinates, accelerometer biases, and other parameters.

INTRODUCTION

Need for New GPS Satellite Instrumentation

The data from geophysics-grade GPS phase tracking receivers are currently used to determine receiving site coordinates with accuracies approaching the millimeter precision of the observables, where the ground observing site coordinates, the GPS orbit initial conditions, and other parameters are simultaneously estimated to best fit the all-in-view phase tracking data.[1]

Among the extra parameters estimated are those in a model of the GPS nongravitational acceleration (~12 nano-g) due to radiation pressure and outgassing. Nongravitational acceleration model inadequacy is a limitation on the site coordinate and orbit fit accuracy obtainable, which problem this paper proposes to circumvent by equipping the GPS satellites with a 3-axis low-g accelerometer (or possibly three single-axis accelerometers), and by the use of satellite-to-satellite phase tracking and ranging observables.

Gyroscopes combined with existing earth and sun sensor data are required to determine the satellite attitude, and hence the position of the transmitting antenna phase center to millimeter accuracy, and to correct for centripetal and angular acceleration effects on the strapdown accelerometers.

Inadequacy of the neutral atmosphere model is another limitation on the site coordinate and orbit fit accuracies obtainable using only ground-based observations. However, use of satellite-to-satellite phase tracking and ranging observables works around this problem. GPS observations from a network of ground sites can then estimate neutral atmosphere characteristics for processing ground-based observations and for input to, e.g., a Navier-Stokes weather prediction model of the atmosphere.

General Relativity Effects

The average earth general relativity effect on GPS clock rate has always been taken into account in the GPS architecture. In addition, the instantaneous earth and sun general relativity effects on GPS clock rate and on orbital motion and radio signal propagation must be taken into account to achieve global millimeter-level orbit and site coordinate determination, and light-time iterations should be done in the solar system barycenter frame, as is done in processing centimeter accuracy lunar laser observations.

Improved Estimation Procedure

The use of maximum likelihood system identification is advocated in the orbit fitting and parameter estimation process, in which an extended Kalman filter is run on the satellite position and velocity states to take account of noise and unmodeled effects in the dynamics, and a maximum likelihood estimator is run on the orbit initial conditions, site coordinates, atmosphere model parameters, accelerometer biases, clock biases, and other parameters.

Applications of Millimeter-Accuracy Satellite Orbit Fits

Applications include:

1. More accurate, robust, and global ground site coordinate determination for, e.g., earthquake prediction and secular change in height monitoring.
2. Real-time determination of neutral atmosphere water vapor content for input to Navier-Stokes weather prediction models.
3. Very accurate airborne gravimetry for, e.g., discovery of mineral deposits or mid-ocean gravity mapping away from differential GPS reference sites.
4. Millimeter accuracy low-altitude satellite orbit determination using GPS observables for, e.g., improved estimates of earth gravity harmonic coefficients and improved interpretation of radar altimeter measurements.
5. Length-of-day and earth wobble monitoring at finer detail between more expensive Very Long Baseline Interferometry (VLBI) observations of stellar sources. The GPS approach is lower cost because the data are available for free in Internet depositories from other applications.[1]
6. Autonomous GPS satellite operation.

All the above applications are being done currently (except the last),[1] but could be done with greater robustness and accuracy with the instrumentation proposed in this paper, especially for global-scale applications.

SATELLITE ACCELEROMETERS

Satellite Nongravitational Acceleration

The radiation pressure and outgassing accelerations on a GPS satellite are of the order 12 nano-g, although somewhat larger than this near the start of mission life when outgassing is larger.[2] The air drag acceleration on a 120-km altitude satellite that could be as large as 500 µg is reduced to 0.1 µg at a 400-km altitude.[3] Thus, the accelerometer provided on low-altitude satellites that observe GPS satellites would have a different dynamic range from the accelerometer provided on the GPS satellites.

The most accurate use of the GPS satellites would be with a low-level accelerometer when the satellites are not thrusting. Occasionally, orbit trim thrusts are done at mg-levels or higher. A higher-level accelerometer for measuring these accelerations could be useful for predicting through these events.

Low-Level Accelerometer Technology

A technology that has been flown in space to measure µg-level nongravitational accelerations with nano-g resolution utilizes a 3-degree-of-freedom electrostatically-supported proof mass.[4] The electrostatic force required to keep the proof mass centered in the accelerometer case is the measure of nongravitational acceleration, where the movement of the proof mass relative to the case is detected electrostatically.

The scale factor of this macro-sized device can be increased by more than an order of magnitude by using a uranium or lead proof mass. The scale factor can be increased further by decreasing the area of the electrostatic pads. In this manner, the 12-nano-g nongravitational acceleration on the GPS satellites can be read with several decimal places of accuracy.

Optical and other technologies for reading out the proof mass position relative to the case and applying restoring forces might yield a better performing low-g accelerometer.

Sources of Accelerometer Noise and Bias Instability

If the accelerometer proof mass is centered in a cavity in a vacuum, the only Brownian mechanical noise is due to noise in the application of the restoring force and in other forces (see below). The readout of the force applied is also a source of noise, whereas noise in the detection of proof mass position is probably not relevant for a low-bandwidth centering control loop.

The bias in the readout of the force applied could vary with time due to instabilities in the electronic components. Variation in other forces on the proof mass is also a source of bias instability, where these other forces are due to the unsymmetric gravitational attraction from the parts of the spacecraft, magnetic forces, and electric forces.

In order to eliminate magnetic forces, the proof mass should be made of nonmagnetic material, and the accelerometer should be shielded from magnetic fields. Proof mass charge accumulation in the radiation environment of space could lead to varying bias electric forces on the proof mass. Changes in the unsymmetric gravitational attraction of the parts of the satellite on the proof mass as fuel is expended can be modeled, or they can be minimized by having a torroidal fuel tank around the center of mass of the satellite where the accelerometer is located. A bladder symmetrically expanding around the circumference of the torus eliminates fuel sloshing.

Note that a 1-kg mass at a 10-cm distance applies an acceleration of 0.68 nano-g to the proof mass.

Accelerometer Noise Requirement

After appropriate high-rate sampling and digital filtering to reduce force application readout noise, let the one-sided acceleration power spectral density (PSD) of the accelerometer have a white noise floor of A $(m/s^2)^2$/Hz at, say, 1 Hz and below. If flicker noise or random walk due to bias instability raised their heads above the white noise floor below the half-orbital revolution frequency of 4.6×10^{-5} Hz, then separate accelerometer biases could be solved for each orbital revolution, or perhaps a trend in bias estimated. Less precise solutions would result if accelerometer biases had to be estimated at fractional revolution intervals.

The velocity random walk resulting from this one-sided PSD noise level is $\sqrt{(A/2)}$ m/s/$\sqrt{s}$. The standard deviation of the velocity noise grows as the square root of time, and the standard deviation of the position is one half this multiplied by time to the 3/2 power. For ±half an orbital period of 6 h, the position standard deviation due to white noise in acceleration is $1.12 \times 10^6\sqrt{A}$ m. Setting this equal to 1 mm yields the requirement

$$\sqrt{A} = 0.091 \text{ nano-g}/\sqrt{\text{Hz}} \quad (1)$$

An acceleration readout precision approaching 0.01 nano-g is desirable.

Accelerometer Calibration and Compensation

The accelerometer proof mass can be dithered at a frequency incommensurate with disturbing acceleration frequencies to calibrate the forcer scale factor relative to the displacement readout scale factor, where the latter is known in terms of the dimensions of the proof mass cavity.

A constant bias over an orbit-fitting interval (ideally at least one 12-h orbital period) can be estimated when doing the orbit fit, as can a bias trend over several orbit periods and misalignments relative to the satellite frame if they were not calibrated adequately before launch.

Any varying gravitational attraction on the accelerometer proof mass from the parts of the satellite (such as due to fuel expenditure or sloshing) has to be modeled and compensated, where the constant part of the satellite gravitational attraction can be absorbed into the bias calibration.

SATELLITE GYROSCOPES

Satellite Attitude Control

Satellite attitude is controlled using momentum wheels and sun and earth sensors. Occasionally, momentum has to be dumped using earth magnetic field torquing.

If gyroscopes were added to the sensor repertoire, the control could be smoother and tighter, especially when the satellite goes through earth shadow.

Attitudinal Motion Effects on a Strapdown Accelerometer

The GPS satellite rotates once every 12 h in order to keep its antenna pointed at the earth (inertial angular velocity $\omega = 1.454 \times 10^{-4}$ rad/s). If the accelerometer were at a distance r = 1 cm from the center of mass of the satellite, its centripetal acceleration $r\omega^2$ would be 0.02 nano-g, which should be corrected in the strapdown accelerometer output before it is used in integrating the satellite equations of motion.

Let satellite attitude angle deviation be $\Delta\theta = A \sin 2\pi ft$, where A is the amplitude and f the frequency of control loop limit cycling. The maximum angular velocity and acceleration effects on the linear accelerometer output due to this limit cycling are $4r(A\pi f)^2$ and $4rA(\pi f)^2$, respectively.

Controlling with sun and earth sensors might have 0.1-deg amplitude, 0.1-Hz limit cycling, whereas controlling with gyroscopes and sun and earth sensors might have 10-μrad amplitude, 10-Hz limit cycling. The maximum angular velocity and acceleration effects of this limit cycling on a linear accelerometer at a distance r = 1 cm from the satellite center of mass are:

lin. acl. due to	nano-g	
	w/out gyros	with gyros
ang. vel.	1.23	0.40
ang. acl.	703	40300

The centripetal acceleration (due to the effect of angular velocity) rectifies and must be compensated for in the accelerometer output, whereas the effect of oscillatory angular acceleration would tend to average zero. High-frequency jitter from thruster firing and momentum wheel ball bearings would make the angular acceleration effect worse, although shock isolation would help.

The low-g accelerometer could be on a gyro-stabilized platform to minimize angular acceleration effects, but there would still be a lever arm effect due to offset from the satellite center of mass. At a distance of r = 10 cm, a 0.1-deg attitude angle motion would cause 0.17-mm motion of the accelerometer proof mass relative to the case.

Thus, a solution to attitude control jitter causing large strapdown linear accelerations (relative to the 12 nano-g dc acceleration) is to have adequate sway space for the accelerometer proof mass relative to the case. The force rebalance proof mass control loop can have low bandwidth, just as long as it keeps the proof mass from hitting the case. The sway space could be larger than a millimeter if the position of the satellite is taken to be the offset of the proof mass from the center of the accelerometer case, plus the result of the double integration of the gravitational acceleration added to the proof mass restoring force measure of nongravitational acceleration.

The low-g accelerometer should be as close to the satellite center of mass as practicable, with the proof-mass sway space and restoring servo-loop performance being tradeoffs in the design. The accelerometer could be on 3-axis guide rails to vernier adjust the location of the accelerometer if it is detected that satellite center of mass has shifted during the mission lifetime.

Antenna Attitude Accuracy Requirement

For an antenna at a distance of 1 m from the satellite center of mass, a 1-mm displacement of the antenna phase center is caused by a 0.057-deg attitude angle error. The attitude either has to be controlled to a fraction of this accuracy for what could even be a larger lever arm, or the angle error has to be telemetered to the ground.

The GPS umbra lasts 53 min, and umbra plus penumbra is double that. The gyroscopes have to maintain attitude through umbra and part of penumbra without help from the sun sensor, so that 0.01-deg/h gyroscopes are required. There are a number of reliable space-qualified gyroscopes with this performance, but lesser cost 1-deg/h gyroscopes cannot be used.

Gyroscope Calibration

Away from shadow, the gyroscope biases and scale factors can be calibrated simultaneously with controlling the satellite attitude with the gyroscopes and the sun and earth sensors.

NEWTONIAN AND GENERAL RELATIVITY ORBIT AND TIME MODELS

Newtonian Gravity and Force Model

Let $\mathbf{x} = (x^1, x^2, x^3)$ be the vector position of a GPS satellite relative to the center of mass of the earth in an inertially oriented coordinate frame (such as referred to the mean equinox and equator of J2000.0), and let $\mathbf{x}_j$ be the vector position of object j (sun, moon, planets) relative to the center of mass of the earth.

The gravitational potential at the satellite is

$$U = -\frac{GM}{|\mathbf{x}|} + H - \sum_j \frac{GM_j}{|\mathbf{x}_j - \mathbf{x}|} + C \qquad (2)$$

where G is the gravitational constant, M is the mass of the earth with GM = 3.986013 × 10^5 km^3/s^2, M_j is the mass of object j, and H is the potential due to higher harmonics in the earth's gravitational field (largest being J_2). Lunar gravity harmonics can also be included if necessary.

The constant C is specified in Section "Relation Between Terrestrial Atomic Time and Coordinate Time" to make coordinate time and sea level terrestrial atomic clocks run at the same average rate. It has no effect on the equations of motion, but does affect the general relativity equations for clock rate and time delay.

Subtracting the Newtonian equations of motion of the earth from those of the satellite yields equations for the satellite position relative to the earth (i = 1, 2, 3):

$$\frac{d^2x^i}{dt^2} = -\frac{GMx^i}{|\mathbf{x}|^3} - \frac{\partial H}{\partial x^i} + A^i + R^i + \sum_j GM_j\left[\frac{x_j^i - x^i}{|\mathbf{x}_j - \mathbf{x}|^3} - \frac{x_j^i}{|\mathbf{x}_j|^3}\right] \tag{3}$$

where A^i is the measured nongravitational acceleration, R^i is the general relativity effect given below, and where the last perturbing body term is a tidal effect. The GPS satellite motion equations are numerically integrated relative to the earth rather than relative to the solar system barycenter for numerical reasons.

General Relativity Metric Tensor

Let = (x^0, x^1, x^2, x^3) be a general relativity coordinate system with origin at the center of mass of the earth, where x^0 = t is coordinate time. The role of the gravitational potential is played by the symmetric metric tensor $ds^2 = g_{\mu\nu}dx^\mu dx^\nu$ (written using the Einstein summation convention), whose components satisfy the Einstein field equations (second-order hyperbolic partial differential equations that imply the existence of gravitational radiation). Test particles and electromagnetic waves follow geodesics relative to this metric.

The Newtonian approximation for the metric for which the geodesic equations are the Newtonian equations of motion is[5]

$$ds^2 = -\left(1+\frac{2U}{c^2}\right)c^2dt^2 + [(dx^1)^2 + (dx^2)^2 + (dx^3)^2] \tag{4}$$

where c ≈ 3 × 10^5 km/s is the velocity of light.

General Relativity Motion Correction

Inserting Eq. (4) for the metric tensor into the Einstein field equations and utilizing a harmonic coordinate condition yields the post-Newtonian approximation to the metric tensor. The equations for a geodesic in the post-Newtonian approximation yield the Newtonian equations of satellite motion (3) plus the following earth-attraction term in harmonic coordinates to second order in $|\mathbf{v}|/c$:[5]

$$R^i = \frac{GM}{|\mathbf{x}|^3}\left\{x^i\left[\frac{4\alpha}{|\mathbf{x}|} - \frac{|\mathbf{v}|^2}{c^2}\right] + \frac{4v^i(\mathbf{v}\cdot\mathbf{x})}{c^2}\right\} \tag{5}$$

where $\mathbf{v}$ = (v^1, v^2, v^3) is the velocity (coordinate time derivative of $\mathbf{x}$) and where α is the gravitational radius of the earth:

$$\alpha = \frac{GM}{c^2} = 4.435 \text{ mm} \tag{6}$$

The resulting advance of the GPS orbit perigee is[5]

$$\Delta\phi = \frac{6\pi\alpha}{a(1-e^2)} \approx 3.15 \text{ nano-rad/rev} \tag{7}$$

where a ≈ 2.656 × 10^4 km is the semimajor axis and e the eccentricity of the GPS 12-h-period orbit. This angle advance is 83.6 mm/revolution along track, which is important for satellite-to-satellite observables, and is equivalent to the angle advance caused by a –8.8-mm change in a. Hence, the general relativity motion term is also important for millimeter-accuracy ground-based observables.

Formally, the general relativity effect of the sun on a GPS satellite orbit in earth-centered coordinates is the general relativity effect on the satellite minus this effect on the earth, with earth-sun cross interaction terms in the post-Newtonian calculations. It turns out that the secular general relativity effect of the sun on the earth satellite orbit is the same as if the satellite orbit were regarded as a gyroscope with "spin axis" along the earth orbit normal, with the "spin axis" precessing about the pole of the ecliptic due to the gyroscope orbiting the sun.[6]

As discovered by de Sitter, the result is a 2-arcsec/century (0.27 nano-rad/day) precession of the lunar orbit plane around the earth relative to the ecliptic plane (plane of the earth's orbit around the sun). The de Sitter effect has been detected in 30 years of centimeter-accuracy lunar-laser corner-reflector data.[7] The de Sitter effect could be a somewhat larger observable effect on GPS orbits because their 54-deg equatorial orbit inclination implies ecliptic inclinations between 31 deg and 77 deg, whereas the lunar orbit ecliptic inclination is only 5 deg.

Relation Between Terrestrial Atomic Time and Coordinate Time

The ratio of a time interval δt shown by an atomic clock to the coordinate time interval Dt is in the Newtonian approximation

$$\frac{\delta t}{Dt} = \left(-\frac{g_{\mu\nu}}{c^2}\frac{dx^\mu}{dt}\frac{dx^\nu}{dt}\right)^{1/2} = \left(1+\frac{2U}{c^2}-\frac{|\mathbf{v}|^2}{c^2}\right)^{1/2} \tag{8}$$

Take the earth gravitational potential at a clock at sea level on the earth to be $-GM/|\mathbf{x}| + C$ and the velocity to be $|\mathbf{x}|\omega_e$cos(latitude), where ω_e is the earth rotation angular velocity. Then the sea level clock rate is

$$1 - 6.9655 \times 10^{-10} + \frac{C}{c^2} + \text{sun, moon effects} \tag{9}$$

relative to coordinate time (CT) rate, where the clock effect of the larger magnitude gravitational potential going toward the pole because of the decrease in radius to the center of the earth is largely counteracted by the clock effect of the decrease in velocity due to the earth rotation (and also earth J_2 effect). At 10^3 m above sea level, the above -6.9655×10^{-10} correction factor is reduced to -6.9644×10^{-10}.

The sun and moon effect on clock rate at the center of the earth are (per Eq. (12) in the next section).

	Average	Eccentricity
Sun	-1.48055×10^{-8}	$\pm 1.654 \times 10^{-10}$
Moon	-2.13×10^{-13}	$\pm 5.8 \times 10^{-15}$

The motion of the earth relative to earth-moon barycenter in the sun's gravitational potential causes up to a $\pm 4.6 \times 10^{-13}$ monthly variation in terrestrial clock rate, whereas the earth rotation through the solar potential causes up to a $\pm 6.3 \times 10^{-13}$ daily variation in terrestrial clock rate.

Choose the constant

$$\frac{C}{c^2} = +1.48055 \times 10^{-8} + 6.9655 \times 10^{-10} \tag{10}$$

so that CT and sea level terrestrial atomic clocks run at the same average rate.

International Atomic Time (TAI) is defined as the average of the time shown by atomic clocks at national time services, which are mostly near sea level, where the International System (SI) second is 9,192,631,770 cycles of a certain cesium hyperfine transition and where

$$CT = TAI + 32.184\ s + \text{sun e effect} \tag{11}$$

The integrated sun eccentricity effect is a yearly ±0.83-ms variation between CT and TAI, which affects interpolation of ephemerides as functions of CT. The motion of the earth relative to the earth-moon barycenter in the sun's gravitational field causes a monthly ±0.2-μs variation between CT and TAI.

The Coordinated Universal Time (UTC) broadcast by the national time services differs from TAI by the leap seconds made every 6 months or a year to keep UTC within 0.9 s of UT1 defined by the rotation of the earth.

General Relativity Clock Rate Correction on Satellite

At true anomaly ψ (angle since perigee) in a GPS orbit due to the earth gravity potential

$$\frac{\delta t}{Dt} \approx 1 + \frac{C}{c^2} - \frac{2\alpha}{a\left(1-e^2\right)}\left(\frac{3}{4} + \frac{e^2}{4} + e\cos\psi\right) + \text{sun, moon} \tag{12}$$

If e = 0.01, there is a ±1.3% variation over an orbit in the general relativity clock rate correction of -2.5047×10^{-10} relative to coordinate time, plus the sun, moon, and C effects, which on average are the same as for a terrestrial clock. Hence

$$\frac{\text{average GPS clock } \delta t}{\text{average sea level clock } \delta t} = 1 + 4.4608 \times 10^{-10} \tag{13}$$

Hence, the GPS clocks are deliberately run faster by the above factor by having fewer cesium cycles in a second in order to keep them in synchronism with sea level clocks. In processing phase tracking data, the variation in GPS clock rate from perigee to apogee should be taken into account, because for e = 0.01, there are 197 fewer cycles of 1.57-GHz GPS frequency data out of 1.7×10^{13} cycles in a 3-h pass around apogee than around perigee.

The variation in the sun gravitational potential effect on clock rate around a GPS orbit is $\pm 2.6 \times 10^{-12}$, and hence must also be taken into account along with the effect of earth orbit eccentricity relative to the sun that affects terrestrial clocks ($\pm 1.654 \times 10^{-10}$ over a year). The variation in the moon gravitational potential effect on clock rate around a GPS orbit is $\pm 2 \times 10^{-14}$.

GROUND-BASED AND SATELLITE-BASED OBSERVABLES

Radio Signal Travel Time

Let $\mathbf{x}_o(t_{rec})$ be the vector from the center of the earth to a GPS receiver on the earth or a satellite at receive time t_{rec}, and let $\mathbf{x}(t_{rec}-\Delta t)$ be the corresponding vector to a GPS satellite at send time $t_{rec} - \Delta t$. Let $\mathbf{x}_e$, $\mathbf{v}_e$ be the vector position and velocity of the earth relative to the solar system barycenter (well known from interplanetary radar and optical observations and spacecraft radio tracking observations).

The one-way travel time delay Δt (determined by iteration in the barycentric coordinate system from interpolating the satellite and earth ephemerides) is

$$\Delta t = |[x(t_{rec}-\Delta t) + x_e(t_{rec}-\Delta t)] - [x_o(t_{rec}) + x_e(t_{rec})]|/c + \Delta t_{ion} + \Delta t_{atm} + \Delta t_{rel} \tag{14}$$

where the ionosphere correction Δt_{ion} is determined from dual frequency measurements, the neutral atmosphere correction Δt_{atm} contains water vapor and other model parameters to be estimated, and the general relativity effect Δt_{rel} is given in Section "General Relativity Radio Propagation Correction."

The difference between doing the light time iteration in the geocentric frame and the solar system barycenter frame is a factor of up to $|\mathbf{v}_e|/c \approx 10^{-4}$ in the light time or range, varying with the GPS-to-receiver vector direction relative to the $\mathbf{v}_e$ direction. This is a large effect, but would be counteracted by the GPS semi-major axis orbit-fit estimate differing when doing the light-time iteration in the geocentric or barycenter frame (essentially due to the Lorentz transformation shortening of measurement rods).

If the various general relativity corrections are included, which are easier to derive in the solar system barycenter frame, it would seem that the light-time iteration should be done in the solar system barycenter frame, as is done in processing lunar laser observations.[7] Numerically integrating the satellite motion equations relative to the center of the earth is done for numerical reasons, where the equations that are integrated are the formal difference (not a Lorentz transformation) of the equations of motion of the GPS satellite and earth in the barycentric frame.

Group Delay Observable

The GPS one-way range group delay observable Δt_{group} is determined by cross-correlation techniques applied to the phase-encoded signal. Its theoretical value is given by Eq. (14) with all the correction terms being positive.

Phase Tracking Observable

The GPS phase tracking observable is the number of cycles of the carrier signal that have occurred between receiver site times t_j and t_{j+1}, where $t_{j+1} - t_j$ is typically between 1 and 30 s for data collected continuously for all-in-view satellites. Sophisticated signal processing makes the count independent of the phase encoding (180-deg phase changes at zero crossings). For an L-band transmit frequency f = 1227.6 or 1575.42 MHz, the phase tracking count observable, which the receiver determines with fractional count resolution, is

$$\phi(t_j, t_{j+1}) = f\,[t_{j+1} - t_j + \Delta t_{phase,j+1} - \Delta t_{phase,j}] \qquad (15)$$

where the transmit frequency f has to be adjusted for the variation in satellite clock rate (see Section "General Relativity Clock Rate Correction on Satellite").

Millimeter parameter estimation accuracy implies some small number w times this accuracy in the ground-based L-band phase tracking observable, so that the phase of the GPS signal is being measured with w × 1.5-deg accuracy (3w ps timing accuracy). The GPS-III satellites will use lower frequency UHF cross links, for which satellite-based phase tracking measurement accuracy could be worse, but the degradation could be counteracted by a larger signal-to-noise ratio.

The theoretical value of the phase delay Δt_{phase} is given by Eq. (14) with the Δt_{atm} and Δt_{rel} corrections being positive and the Δt_{ion} correction being negative.

General Relativity Radio Propagation Correction

The general relativity one-way group or phase delay effect (first derived by Irwin Shapiro,[8] separate from the bending effect) due to the gravitational field of the earth in harmonic coordinates is[9]

$$\Delta t_{rel} = \frac{2\alpha}{c} \log_e \left[\frac{|x| + |x_o| + |x - x_o|}{|x| + |x_o| - |x - x_o|} \right] \qquad (16)$$

where the formula would be changed slightly by the constant C in the gravitational potential, but C is ignored here in order to see the magnitude of the effect using previously-derived formulas.

For a terrestrial observing site, the zenith $c\Delta t_{rel}$ = 12.65 mm and the horizon $c\Delta t_{rel}$ = 18.67 mm for a 6-mm variation. From one GPS satellite to another, the largest $c\Delta t_{rel}$ = 37.34 mm and the smallest $c\Delta t_{rel}$ = 9.74 mm for a 27.6-mm variation.

Applying Eq. (16) to the gravitational field of the sun by replacing α by α_s = 1.4766 km, $\mathbf{x}$ by $\mathbf{x} - \mathbf{x}_s$ with $|\mathbf{x}_s| = 1.496 \times 10^8$ km, etc., yields a correction for a terrestrial receiver between 398 mm and 509 mm (total variation 111 mm) and a correction for a GPS satellite receiver between 524 mm and 1017 mm (total variation 493 mm). The lunar general relativity radio propagation effect is less than 0.05 mm.

Raw and Difference Observables

Great success in estimating regional millimeter-level observing site coordinates has been obtained by fitting to doubly-differenced geophysics phase tracking all-in-view observables, where the difference of different receivers observing the same satellite are taken, and then the difference of different satellites.[10] There is common-mode rejection of many error effects, such as receiver and satellite clock errors, neutral atmosphere effects, etc. Similar results are obtained using nondifferenced raw observables, with receiver and satellite clock biases being estimated for each observing pass.[1]

For estimating neutral atmosphere parameters for input to numerical weather prediction models and for using the fitted GPS orbits in such applications as airborne gravimetry, the raw group delay and phase tracking ground-based and satellite-based all-in-view observables should be used.

Number of Observables

If each of 24 GPS-III satellites observed all the other GPS-III satellites continuously when not occulted by the earth, there could be 24 × 23 = 552 phase tracking and group delay data streams telemetered to the earth, plus the accelerometer, gyro, sun and earth sensor, and other housekeeping data streams for each satellite. The number of phase tracking data streams would be cut in half if a satellite did not observe a satellite that was observing itself. The phase delay count interval could range between 1 and 30 s, for a total of 1.5 to 47 million phase tracking observables a day, plus an equal number of group delay observables.

To these can be added the all-in-view phase tracking and group delay observables from hundreds or thousands of ground stations, yielding tens or hundreds of millions of observables per day.

Processing these data with the FIMLOF estimation technique described in the next section will strain the capabilities of ground-based computers, even as they get faster and faster. A more straightforward Kalman filter or least squares estimator could be used by a satellite-based radiation-hard computer insofar as meter rather than millimeter accuracy is required for GPS autonomous operation if contact with ground control stations were lost.

FITTING TO DATA

FIMLOF Overview

It is advocated that an extended Kalman filter be run on the time-varying satellite position and velocity and a maximum likelihood estimator be run on the parameters and satellite initial conditions.[11],[12] This approach, which is used in adaptive control,[13] has been called maximum likelihood system identification, or Full Information Maximum Likelihood Optimal Filtering (FIMLOF). It allows noise in the satellite dynamics caused by unmodeled effects and nongravitational acceleration measurement noise, and at the same time, uses an optimal maximum likelihood estimator on the constant parameters such as site coordinates, accelerometer biases, satellite initial osculating elliptic orbital elements, etc.

Note that the site coordinates are constant during the time span of an orbit fit after compensation for the ±30-cm half-daily deformation due to the lunar-solar solid earth tide and ocean tide loading. The movement of the site coordinates from one orbit fit interval to the next, or a detected shift during an orbit fit interval, is what is of interest for possible earthquake prediction.

Given nominal values of the satellite initial osculating elliptic orbital elements at time t_0 and other parameters such as accelerometer biases, earth gravitational potential harmonics, observing site coordinates, atmosphere parameters, clock biases, etc., the satellite equations of motion are numerically integrated along with the equations for the partial derivatives of the motion with respect to the parameters (except for parameters that do not affect satellite motion such as site coordinates). Earth gravitational harmonic coefficient partial derivatives are calculated only if low-altitude satellite observations are included.

An extended Kalman filter is run on the satellite motions given the observables $Z(t_k)$ at times t_k ($k = 1, \ldots, N$), where the plant noise covariance in the Kalman filter equations is sized to encompass any unmodeled effects and nongravitational acceleration measurement noise. Using the partial derivatives of the observables with respect to the initial osculating elliptic orbital elements and other parameters, a maximum likelihood adjustment is made to these quantities. The process is repeated until convergence is obtained.

The satellite motion equations can be numerically integrated ahead in time for prediction purposes with zero plant noise and projections of the nongravitational acceleration, which probably repeats to a great extent from one orbit to the next.

Partial Derivatives of Satellite Motion

Given initial osculating elliptic orbit elements, the initial Cartesian position and velocity are calculated as initial conditions for numerically integrating equations of motion (3) with R^i being the general relativity term (5) plus the sun de Sitter effect. The equations for the partial derivatives with respect to parameters β_ℓ are also numerically integrated:

$$\begin{aligned} \frac{d^2 \partial x^i / \partial\beta_\ell}{dt^2} &= -\frac{GM}{|\mathbf{x}|^3}\left[\frac{\partial x^i}{\partial\beta_\ell} - \frac{3x^i}{|\mathbf{x}|^2}\sum_{k=1}^{3} x^k \frac{\partial x^k}{\partial\beta_\ell}\right] \\ &\quad - \frac{\partial}{\partial\beta_\ell}\frac{\partial H}{\partial x^i} + \frac{\partial A^i}{\partial\beta_\ell} + \frac{\partial R^i}{\partial\beta_\ell} \\ &\quad + \frac{\partial}{\partial\beta_\ell}\sum_j GM_j\left[\frac{x_j^i - x^i}{|\mathbf{x}_j - \mathbf{x}|^3} - \frac{x_j^i}{|\mathbf{x}_j|^3}\right], \quad i = 1,2,3 \end{aligned} \tag{17}$$

with zero initial conditions at time t_0, unless β_ℓ is an initial osculating elliptic orbital element, in which case, the initial conditions are the partial derivatives of the initial position and velocity with respect to the orbital element. Estimating initial osculating elliptic orbital elements instead of initial position and velocity is a standard approach in celestial mechanics, which separates the partial derivative with respect to semimajor axis a that grows with time from the other initial condition partial derivatives that do not grow with time.

The equations for the second-order partial derivatives $\partial^2 x^i/\partial elem_0 \partial\beta_\ell$ with respect to initial osculating elliptic orbital elements $elem_0$ and the parameters β_ℓ are also numerically integrated with zero initial conditions, except that the initial conditions are nonzero for β_ℓ one of the $elem_0$.

Extended Kalman Filter State Equation

Let $x(t_k)$ denote the position and velocity vector of dimension 6M at time t_k of the M satellites, as determined by interpolation from the numerically integrated ephemeris files (one file for each satellite) given values for the initial conditions and parameters. For the GPS satellite constellation, M is nominally 24, but it could be more or less, depending on the health of the individual satellites and whether spare satellites are activated. If other satellites observing GPS were included in the orbit fit in addition to the GPS satellites, M would be larger than 24.

Let $y(t_k)$ be the linear Kalman filter correction to $x(t_k)$, so that the observable vector $Z(t_k)$ at time t_k is a function $h(x(t_k) + y(t_k))$. Take the initial state of the linearized Kalman filter state vector to be $y_0 = y(t_0) = 0$ with an a priori covariance $P_0 = P(t_0)$ defined as a reasonable measure of the uncertainty in $x(t_0)$, in order to start off the Kalman filter for $y(t)$. Note that $x(t_0)$ will be maximum likelihood adjusted after running the Kalman filter, since the initial osculating elliptic orbital elements are so adjusted.

The propagation of the state vector y from time t_{k-1} to time $t_k \geq t_{k-1}$ is given by the state equation

$$y(t_k) = \Phi(t_k\,;\beta)y(t_{k-1}) + L\xi(t_{k-1}) \tag{18}$$

where ξ is a zero mean white plant noise vector with covariance $E\{\xi(t_j)\xi(t_k)^T\} = Q\delta_{jk}$. The state transition matrix Φ as function of the parameters $\beta=(\beta_1, \ldots, \beta n)$ is

$$\Phi(t_k) = \left[\frac{\partial x(t_k)}{\partial x(t_{k-1})}\right] = \left[\frac{\partial x(t_k)}{\partial\, elem_0}\right]\left[\frac{\partial x(t_{k-1})}{\partial\, elem_0}\right]^{-1} \tag{19}$$

where the $6M \times 6M$ matrix Φ is zero away from 6×6 blocks down the diagonal (hence only 6×6 submatrices need to be inverted), with the blocks being determined by numerically integrating the equations for the partial derivatives of satellite position and velocity with respect to initial osculating elliptic orbital elements $elem_0$.

The partial derivatives $\partial\Phi/\partial\beta_\ell$ can also be calculated using the second-order partial derivatives described at the end of the last section.

Processing Observations

Take the Kalman filter observable vector $z(t_k)$ to be the actual observable $Z(t_k)$ minus the theoretical value of the observable $h(x(t_k))$ evaluated for the nominal satellite motion state $x(t_k)$. Then, the linearized Kalman filter observables equation is

$$z(t_k) = H(t_k\,;\beta)y(t_k) + \theta(t_k) \tag{20}$$

where θ is a zero mean white measurement noise vector with covariance $E\{\theta(t_j)\theta(t_k)T\} = R\delta_{jk}$, and $H(t_k) = \partial h / \partial x(t_k)$.

Given values for the parameters β (including initial osculating elliptic orbital elements) and the initial state $y(t_0) = 0$ and its covariance P_0, running a Kalman filter yields the propagated state $\hat{y}(t_k \mid t_{k-1}) = \Phi\hat{y}(t_{k-1})$, its covariance

$$P(t_k \mid t_{k-1}) = \Phi\, P(t_k)\, \Phi^T + L\, Q\, L^T \tag{21}$$

the zero mean pre-update residual innovations sequence and its covariance

$$r(t_k) = z(t_k) - H\,\hat{y}(t_k \mid t_{k-1}) \tag{22}$$

$$S(t_k) = E\{r(t_k)\, r(t_k)^T\} = H\, P(t_k \mid t_{k-1})\, H^T + R \tag{23}$$

and the updated state and its covariance

$$\hat{y}(t_k) = \hat{y}(t_k \mid t_{k-1}) + K(t_k)\, r(t_k) \tag{24}$$

$$P(t_k) = E\{\hat{y}(t_k)\,\hat{y}(t_k)^T\} = [I - K(t_k)\,H]P(t_k \mid t_{k-1}) \quad (25)$$

where I is the identity matrix and K is the Kalman filter gain matrix[14]

$$K(t_k) = P(t_k \mid t_{k-1})\,H^T\,[H\,P(t_k \mid t_{k-1})\,H^T + R]^{-1} \quad (26)$$

As the Kalman filter is run on the states y, the partial derivatives of the various Kalman filter matrices are also calculated to allow a maximum likelihood adjustment to the parameters β upon completion of the Kalman filter.

Negative Log Likelihood Function

The joint probability density $p(z^N)$ of the observables $z^N = [z(t_N), z(t_{N-1}), \ldots, z(t_1), y(t_0)]$ through time t_N is

$$p(z^N) = p(z(t_N)|z^{N-1})p(z(t_{N-1})|z^{N-2})\ldots p(z(t_1)|y_0)p(y_0) \quad (27)$$

by Bayes rule, where the Gaussian conditional probability density of $z(t_k)$ given z^{k-1} is

$$p(z(t_k)|\,z^{k-1}) = (2\pi)^{-\rho/2}\det[S(t_k)]^{-1/2} \times \exp[r(t_k)^T\,S(t_k)^{-1}\,r(t_k)] \quad (28)$$

where ρ is the dimension of the observation vector. The negative log likelihood $\zeta = -\ln[p(z^N)]$ of the measurements through time t_N is

$$\zeta(z^N;\beta) = \sum_{k=1}^{N}\zeta(z(t_k)\,|\,z^{k-1};\beta) + \zeta(y(t_0);\beta) \quad (29)$$

where by Eq. (28) with 6M the dimension of the initial state vector $y(t_0)$

$$\zeta(z(t_k)\,|\,z^{k-1};\beta) = \frac{\rho}{2}\ln(2\pi) + \frac{1}{2}\ln\{\det[S(t_k)]\} + \frac{1}{2}r(t_k)^T\,S(t_k)^{-1}\,r(t_k) \quad (30)$$

$$\zeta(y(t_0);\beta) = \frac{6M}{2}\ln(2\pi) + \frac{1}{2}\ln\{\det[P_0]\} + \frac{1}{2}[y(t_0) - y_0]^T\,P_0^{-1}\,[y(t_0) - y_0] \quad (31)$$

The constant terms involving 2π can be ignored in taking partial derivatives, as can all of $\zeta(y(t_0);\beta)$ because of the way in which we have formulated the extended Kalman filter with $y(t_0) = y_0 = 0$ and P_0 not involving parameters to be estimated.

The partial derivatives of the negative log likelihood are

$$\frac{\partial\zeta(z^N;\beta)}{\partial\beta_\ell} = \sum_{k=1}^{N}\frac{\partial\zeta(z(t_k)\,|\,z^{k-1};\beta)}{\partial\beta_\ell} + \frac{\partial\zeta(y(t_0);\beta)}{\partial\beta_\ell} \quad (32)$$

where

$$\frac{\partial\zeta(z(t_k)\,|\,z^{k-1};\beta)}{\partial\beta_\ell} = \frac{1}{2}\mathrm{trace}\left[S(t_k)^{-1}\frac{\partial S(t_k)}{\partial\beta_\ell}\right] + r(t_k;\beta)^T\,S(t_k)^{-1}\frac{\partial r(t_k)}{\partial\beta_\ell} - \frac{1}{2}r(t_k)^T\,S(t_k)^{-1}\frac{\partial S(t_k)}{\partial\beta_\ell}S(t_k)^{-1}\,r(t_k) \quad (33)$$

$$\frac{\partial\zeta(y(t_0);\beta)}{\partial\beta_\ell} = 0 \quad (34)$$

Maximum Likelihood Estimation

Let $\Delta\beta_\ell = \hat{\beta}_\ell - \beta_{\ell o}$ be adjustments from guesses $\beta_{\ell o}$ toward maximum likelihood estimates $\hat{\beta}_\ell$ that minimize the negative log likelihood $\zeta(z^N;\beta)$. A Taylor series expansion of the gradient of $\zeta(z^N;\beta)$ yields

$$\sum_{\ell=1}^{n}A_{m\ell}\,\Delta\beta_\ell = B_m\,, \quad m = 1, \ldots, n \quad (35)$$

$$A_{m\ell} = \left.\frac{\partial^2\zeta(z^N;\beta)}{\partial\beta_m\partial\beta_\ell}\right|_{\beta=(\beta_{1o},\ldots,\beta_{no})} \quad (36)$$

$$B_m = \left.\frac{\partial\zeta(z^N;\beta)}{\partial\beta_m}\right|_{\beta=(\beta_{1o},\ldots,\beta_{no})} \quad (37)$$

The expected value of the matrix A is the Fisher information matrix $\mathbf{I}$. When solving for the adjustments $\Delta\beta_\ell$ from the normal equations (35), the Fisher information approximation is used, where A is replaced by $\mathbf{I}$:

$$A_{m\ell} \approx \mathbf{I}_{m\ell} \equiv E\left[\frac{\partial^2\zeta(z^N;\beta)}{\partial\beta_m\partial\beta_\ell}\right] = E\left[\frac{\partial\zeta(z^N;\beta)}{\partial\beta_m\partial\beta_\ell}\;\frac{\partial\zeta(z^N;\beta)}{\partial\beta_m\partial\beta_\ell}\right] \quad (38)$$

Taking expected values using Eq. (32) yields[11]

$$A_{m\ell} \approx \mathbf{I}_{m\ell} = \sum_{k=1}^{N}\left\{\mathrm{trace}\left[\frac{\partial r(t_k)^T}{\partial\beta_m}\frac{\partial r(t_k)}{\partial\beta_\ell}S(t_k)^{-1}\right] + \frac{1}{2}\mathrm{trace}\left[\frac{\partial S(t_k)}{\partial\beta_m}S(t_k)^{-1}\frac{\partial S(t_k)}{\partial\beta_\ell}S(t_k)^{-1}\right]\right\} \quad (39)$$

The normal equations (35) are solved with the Fisher information approximation (39) to obtain parameter adjustments, and the satellite equations of motion and the equations for the partial derivatives of the motion are numerically integrated with the new values of the initial osculating elliptic orbital elements and other parameters. The extended Kalman filter is rerun and the normal equations for adjustments to the parameters reformed and solved to get further adjustments to the parameters. The iteration continues until convergence is reached. The uncertainty of the parameter estimates has the Cramer-Rao lower bound $\mathbf{I}^{-1}$.[11]

CONCLUSIONS

Satellite Instrumentation

Adding gyroscopes and a low-g 3-axis linear accelerometer as well as satellite-to-satellite observable and communication cross links to the GPS-III satellites would have to be engineered carefully into the satellite design. Cross-link antennas as well as earth-coverage antennas are required, and interference has to be avoided for simultaneous transmission and reception of cross-link signals. It is important that the GPS satellite-based receivers include phase tracking as well as group delay measurements.

The 3-axis accelerometer should be as close to the center of mass of the satellite as possible, with fuel tanks symmetrically expending fuel relative to the center of mass with no sloshing. The satellite attitude should be smoothly and tightly controlled by blending the outputs of the gyroscopes and sun and earth sensors, including when going through shadow.

The gyroscope, accelerometer, sun and earth sensor, and satellite-to-satellite observable data would be telemetered to the ground to be used along with ground-based observations for robust millimeter-accuracy orbit fits and ground site position determination. With satellite-based observables, orbit fits on the satellites are possible (perhaps to meter rather than millimeter accuracy) for autonomous operation independent of ground-control sites, if adequate onboard computation and data storage capacity are provided.

Effects of General Relativity

The orbital motion, radio propagation, and clock rate effects of general relativity have to be modeled carefully to post-Newtonian order, including earth-sun cross coupling. These effects are most easily modeled in the solar system barycenter frame, so that light-time iterations should be done in that frame for the millimeter-level orbit fits.

For more conventional lower-accuracy applications with broadcast ephemerides, either the orbit fits have to be repeated in a geocentric frame or the orbit fit in the barycenter frame has to be Lorenz transformed to the geocentric frame, since the light time iteration is traditionally done in the geocentric frame or not at all, where a large portion of the geocentric light time effect is absorbed in the estimated clock bias when doing receiver triangulation navigation from pseudo-range to four or more satellites.

Estimation Approach and Applications

A mixture of extended Kalman filtering and maximum likelihood estimation should be used in fitting the GPS orbits to data (maximum likelihood system identification or FIMLOF), which, however, requires the calculation of some second-order partial derivatives. If satellite-based accelerometer measurements result in low enough plant noise in the satellite dynamics, then least-squares maximum likelihood estimation could be adequate using only first-order partial derivatives.

Not being corrupted by atmospheric effects, the satellite-based observables determine the relative distances within the GPS constellation very accurately and give some visibility into the GPS constellation's angular orientation because of earth gravity harmonic and lunar-solar perturbations. The ground-based observables better determine the GPS constellation's angular orientation, since they are made from the rotating earth with a known or estimated transformation from earth-fixed coordinates to inertial coordinates referred to the mean equinox and equator of J2000.0, which is the coordinate frame in which the GPS satellite motions are numerically integrated.

Robust millimeter positioning has many applications, some of which are listed in Section "Applications of Millimeter-Accuracy Satellite Orbit Fits," including estimating earth rotation parameters (wobble and UT1 – UTC).

ACKNOWLEDGMENTS

Thanks are due to Dr. Robert King and Prof. Thomas Herring of MIT, to Dr. John Chandler at the Harvard-Smithsonian Center for Astrophysics, and to Dr. Richard Greenspan of Draper Laboratory for several helpful discussions.

REFERENCES

[1] International GPS Service Web Site: http://igscb.jpl.nasa.gov

[2] Fliegel, H., T. Gallini, and E. Swift, "Global Positioning System Radiation Force Model for Geodetic Applications," *Journal of Geophysical Research*, Vol. 97, No. B1, 1992, pp. 559-568.

[3] Champion, K.S.W. and F.A. Marcos, "The Triaxial-Accelerometer System on Atmosphere Explorer," *Radio Science*, Vol. 8, 1973, pp. 297-303.

[4] Blanchard, R.C., M.K. Hendrix, J.C. Fox, D.J. Thomas, and J. Nicholson, "The Orbital Acceleration Research Experiment," Paper 86-1195, AIAA 3rd Space Systems Technology Conference, 1986.

[5] Weinberg, S., *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*, John Wiley & Sons, New York, 1972.

[6] Will, C.M., *Theory and Experiment in Gravitational Physics*, 2nd edition, Cambridge University Press, Cambridge, England, 1993.

[7] Shapiro, I.I., R.D. Reasenberg, J.F. Chandler, and R.W. Babcock, "Measurement of the de Sitter Precession of the Moon: A Relativistic Three-Body Effect," *Phys. Rev. Letters 61*, 1968, pp. 2643-2646.

[8] Shapiro, I.I., G.H. Pettengill, M.E. Ash, M.L. Stone, W.B. Smith, R.P. Ingalls, and R.A. Brockelman, "Fourth Test of General Relativity: Preliminary Results," *Phys. Rev. Letters 20*, 1968, pp. 1265-1269.

[9] Soffel, M.H., *Relativity in Astrometry, Celestial Mechanics and Geodesy,* Springer-Verlag, Berlin, 1989.

[10] Kogan, M.G., G.M. Steblov, R.W. King, T.A. Herring, D.I. Frolov, S.G. Egorov, V.Y. Leven, A. Lerner-Lam, and A. Jones, "Geodetic Constraints on the Rigidity and Relative Motion of Eurasia and North America," *Geophy. Res. Lett. 27*, 2000, pp. 2041-2044.

[11] Sandell, Jr., N. and K. Yared, "Maximum Likelihood Identification of State Space Models for Linear Dynamic Systems," MIT Electronics Systems Laboratory Report ESL-R-814, Cambridge, MA, 1978.

[12] Ash, M.E. and M.E. Skeen, "Maximum Likelihood Estimation of Fractional Brownian Motion Parameters and Markov Noise Parameters," *J. of Guidance, Control, and Dynamics 18*, 1994, pp. 379-382.

[13] Maybeck, P., *Stochastic Models, Estimation, and Control*, Vol. 2, Chap. 10, Academic Press, San Diego, CA, 1982.

[14] Gelb, A. (ed.), *Applied Optimal Estimation*, MIT Press, Cambridge, MA,1974.

Michael E. Ash

Michael E. Ash is a Principal Member of the Technical Staff in the System Integration, Test, and Evaluation Division, where he works on inertial sensor and system modeling, simulation, and testing. He previously worked at the MIT Lincoln Laboratory on an interplanetary radar test of general relativity and on satellite orbit determination. He is Chair of the Accelerometer Committee of the IEEE/Aerospace Electronics System Society (AESS) Gyro and Accelerometer Panel and an Associate Fellow of the American Institute of Aeronautics and Astronautics (AIAA). He received a BS from MIT and a PhD from Princeton, both in Mathematics.