Modeling the Global Positioning System Signal Propagation Through the Ionosphere

S. Bassiri
Sharif University of Technology, Tehran, Iran

G. A. Hajj
Tracking Systems and Applications Section

Based on realistic modeling of the electron density of the ionosphere and using a dipole moment approximation for the Earth's magnetic field, one is able to estimate the effect of the ionosphere on the Global Positioning System (GPS) signal for a ground user. The lowest order effect, which is on the order of 0.1–100 m of group delay, is subtracted out by forming a linear combination of the dual frequencies of the GPS signal. One is left with second- and third-order effects that are estimated typically to be ~0–2 cm and ~0–2 mm at zenith, respectively, depending on the geographical location, the time of day, the time of year, the solar cycle, and the relative geometry of the magnetic field and the line of sight. Given the total electron content along a line of sight, the authors derive an approximation to the second-order term which is accurate to ~90 percent within the magnetic dipole moment model; this approximation can be used to reduce the second-order term to the millimeter level, thus potentially improving precise positioning in space and on the ground. The induced group delay, or phase advance, due to second- and third-order effects is examined for two ground receivers located at equatorial and mid-latitude regions tracking several GPS satellites.

I. Introduction

The Global Positioning System (GPS) consists of 24 satellites, evenly distributed in 6 orbital planes around the globe, at an altitude of about 20,200 km. Precise positioning of the GPS satellites, as well as ground and space users, is now reaching a few parts in 10^9 [1–6]. In addition, the GPS has been heavily utilized in a host of geodetic and other applications. These include seismic tectonic motions [7–9], Earth orientation studies [10,11], gravimetry [12], atmospheric water vapor calibration [13,14], and ionospheric monitoring [15]. Precise positioning and other GPS-based applications, however, require a very good understanding of all effects on the GPS signal as it propagates through the Earth's atmosphere, so that all effects can be solved for or modeled.

The GPS transmits two right-hand circularly polarized (RCP) signals at L-band frequencies: L1 at 1574.42 MHz

and L2 at 1227.6 MHz, which correspond to wavelengths of 19.0 cm and 24.4 cm, respectively. These are modulated by a pseudorandom precision code (P-code) at a frequency of 10.23 MHz [16]. [The additional lower frequency course acquisition (C/A) modulation is not of concern here.] A single measurement for a given transmitter and receiver pair will consist of four observables that will be denoted here by L_1, L_2 for the accumulated carrier phase measurements at the two frequencies and P_1, P_2 for the corresponding P-code pseudorange. In addition to the geometric range delay, the signals will experience delays, or phase advances, due to the presence of the ionosphere and neutral atmosphere.

The delay due to the neutral atmosphere is the same for all observables; its effect is on the order of 2 m and can be solved for to better than a centimeter [13,14]. However, due to the dispersive nature of the ionosphere, the group delay caused by it (or phase advance) is frequency dependent, and is on the order of 0.1–100 m, depending on the time of day, the time of year, and the solar cycle. If the ionospheric effect on signal delay (or advance) is expanded in powers of inverse frequencies, then the lowest order term $(1/f^2)$, by far the most dominant, can be solved for and subtracted out by virtue of the dual frequencies of the GPS. Remaining higher order terms are on the order of submillimeters to several centimeters, which remain embedded in the signal and contribute to range and accumulated phase errors. While the first-order term depends simply on the total electron content (TEC), namely the integrated electron density inside a columnar cylinder of unit area between the transmitter and the receiver, higher order terms depend on the coupling between the Earth's magnetic field and the electron density everywhere along the line of sight. In order to estimate the higher order effects on the GPS observables, the authors modeled the ionosphere by a sum of Chapman layers and the Earth's magnetic field by that of a dipole moment. Such a model will make it possible to estimate higher order ionospheric effects at different geographical locations on the ground as well as their sensitivity to the electron density distribution. It will be demonstrated that knowledge of the TEC can be used to calibrate most of the second-order effect and reduce P-code and phase measurement errors to a few millimeters.

Due to the inhomogeneity of the propagation medium, the GPS signal does not travel along a perfectly straight line. Moreover, since the medium is dispersive, the two frequencies will take two slightly different paths. By applying the empirical formula given by Brunner and Gu [17] on the ionospheric model used below, the residual range error between the dual-frequency corrected range and the true range, due to bending alone, is estimated to be ~4 mm at a 10-deg elevation angle and less than a millimeter for elevations above 30 deg. The bending effect will be ignored in the following analysis; the two signals will be assumed to travel along the same straight line.

A more elaborate modeling of higher order ionospheric effects, where bending is taken into account, has been considered by Brunner and Gu [17]; see also [18]. In their paper, the international geomagnetic reference fields (IGRF) and a Chapman profile of the ionosphere were used to estimate the residual range error. They also proposed an improved linear combination that corrects for the second- and third-order terms, as well as for bending. Their improved linear combination requires knowledge of N_m and h_m, the electron density peak and its altitude, respectively. In this article, the second- and third-order terms are considered separately. Here the authors estimate that the second-order term is dominant most of the time over the third-order and the curvature terms. A method of modeling the second-order effect based on a thin shell model of the ionosphere and a dipole magnetic field is suggested. The second- and third-order errors are considered at different geographical locations while tracking different satellites. It is demonstrated that knowledge of the TEC alone can be used to reduce the higher order effects to a few millimeters.

II. Earth's Ionosphere

The Earth's ionosphere extends from an altitude of about 80 to 1000 km. It is a macroscopically neutral ionized gas consisting principally of free electrons, ions, and neutral atoms or molecules. Ions in that region are 2000 to 60,000 times more massive than electrons. Thus, at the frequencies used for radio communication, the range of movement of an ion caused by the electric field of a radio wave is smaller than that of an electron by about the same factor. This implies that the ions can, for most purposes, be ignored [19].

The electron density profile exhibits several distinct regions (E, F1, and F2) as a result of the competing processes of particle production, loss, and transport. The maximum electron densities (10^{12} to 10^{13} m^{-3}) are observed at the F2 peak; the peak altitude ranges from 250 to 350 km at mid-latitudes and from 350 to 500 km at equatorial latitudes. The F1 region, which is present during the day but absent at night, has a peak near the 200-km altitude and is 3–5 times smaller than that of F2. The E peak density is about one order of magnitude smaller than the F2

peak and is typically located at the 100- to 120-km altitude. During daytime there is also a D region below the E region, with a peak at the 80-km altitude [20].

III. Propagation of Electromagnetic Waves in the Ionosphere

When a magnetostatic field $\vec{B}_0$ is applied to a plasma, the plasma becomes anisotropic for the propagation of electromagnetic waves. That is, the scalar dielectric constant of the plasma is transformed into a tensor. To study the propagation and polarization properties of a plane monochromatic wave in a magnetically biased homogeneous lossless plasma, the plasma is regarded as a continuous medium whose conductivity is zero, whose permeability is equal to that of a vacuum, and whose dielectric constant is a tensor. By solving the Helmholtz wave equation subject to proper constitutive relations, one can obtain the expressions for the fields and for the index of refraction. The index of refraction, n, for the Earth's ionosphere is given by the Appleton-Hartree formula [21], as follows:

$$n_{\pm}^2 = 1 - \frac{2X(1-X)}{2(1-X) - Y_{\perp}^2 \pm \sqrt{Y_{\perp}^4 + 4(1-X)^2 Y_{||}^2}} \quad (1)$$

where

$$X = \left(\frac{f_p}{f}\right)^2 = \frac{(Ne^2/4\pi^2\varepsilon_o m)}{f^2} \quad (2)$$

$$Y_{\perp} = Y \sin\ \theta_B;\ Y_{||} = Y \cos\ \theta_B \quad (3)$$

$$Y = \left(\frac{f_g}{f}\right) = \frac{(|e|B_0/2\pi m)}{f} \quad (4)$$

N is the number density of electrons; e and m are the electron charge and mass, respectively; ε_o is the permittivity of the free space; f_p, f_g, and f are the plasma, gyro, and carrier frequencies, respectively; and θ_B is the angle between the Earth's magnetic field, $\vec{B}_0$, and the direction of propagation of the wavefront, $\vec{k}$. By definition, $\vec{Y} = e\vec{B}_0/2\pi f m$, and since e is negative, $\vec{Y}$ is antiparallel to $\vec{B}_0$. The plasma frequency is the natural frequency of oscillation for a slab of neutral plasma with density N after the electrons have been displaced from the ions and are allowed to move freely. The gyro frequency is the natural frequency at which free electrons circle around the magnetic field lines. For the Earth's ionosphere, with $N = 10^{12}$ electrons/m^3, the plasma frequency $f_p \approx 8.9$ MHz. The gyro frequency for an electron in the Earth's magnetic field (2×10^{-5} tesla) is $f_g \approx 0.59$ MHz.

The plus and minus signs of Eq. (1) correspond to the ordinary and extraordinary wave modes of propagation, respectively. In general, these two waves are elliptically polarized with left and right senses of rotation, respectively. As a result of different phase velocities of the two waves, the total wave (the sum of ordinary and extraordinary waves) undergoes Faraday rotation as it passes through the ionosphere. When the carrier frequency is large, as compared with plasma and gyro frequencies, the principal modes of propagation are dominantly circularly polarized. This is the case for the GPS carrier frequencies.

Assuming that $Y \ll 2|\cos\theta_B|(1-X)/\sin^2\theta_B$, the index of refraction can be expanded in inverse powers of frequency. For the GPS carrier frequencies, one has $(f_p/f) = 5.65\times10^{-3}$ and 7.25×10^{-3}, as well as $(f_g/f) = 3.75\times10^{-4}$ and 4.81×10^{-4} for L1 and L2, respectively. Therefore, the stated assumption is valid for GPS frequencies up to a value of $\theta_B \approx 89$ deg. The expansion of Eq. (1) up to the fourth inverse powers of frequency gives

$$n_{\pm} = 1 - \frac{1}{2}X \pm \frac{1}{2}XY|\cos\theta_B| - \frac{1}{4}X\left[\frac{1}{2}X + Y^2(1+\cos^2\theta_B)\right] \quad (5)$$

The second, third, and fourth terms on the right-hand side of Eq. (5) are proportional to the inverse square, inverse cube, and inverse quartic powers of frequency, respectively. The two values of n refer to the ordinary (+) and extraordinary (-) waves. At this point it should be noted from Eq. (5) that the index of refraction is smaller than unity, which corresponds to a phase velocity greater than the speed of light (phase advance). The group refractive index, on the other hand, given by $n^{\text{group}} = n + f(dn/df)$, can be written as

$$n_{\pm}^{\text{group}} = 1 + \frac{1}{2}X \mp XY|\cos\theta_B| + \frac{3}{4}X\left[\frac{1}{2}X + Y^2(1+\cos^2\theta_B)\right] \quad (6)$$

The group delay of a signal passing through the ionosphere, relative to vacuum as a reference, can be rewritten as

$$\tau_{\pm}^{\text{group}} = \frac{1}{c}\int (n_{\pm}^{\text{group}} \cos\alpha - 1)dl \qquad (7)$$

where dl is an element of length along the line of sight, c is the velocity of light in a vacuum, and α is the angle between the wave normal and the ray direction. This angle has significance in anisotropic media, where the direction of the wave normal is, in general, different from the direction of energy propagation. Angle α can be found from the following relation: $\tan\alpha = (1/n)\partial n/\partial\theta_B$. By using Eq. (5) and the definition of α, it is easy to show that for the GPS carrier frequencies $\cos\alpha$ is essentially unity. By using Eqs. (5)–(7), the GPS observables can be written [ignoring the left-hand circularly polarized (LCP) component of the GPS signal, which has <0.35 percent and <2.5 percent of the total power, for L1 and L2, respectively] as

$$P_1 = \rho + \frac{q}{f_1^2} + \frac{s}{f_1^3} + \frac{r}{f_1^4} \qquad (8a)$$

$$P_2 = \rho + \frac{q}{f_2^2} + \frac{s}{f_2^3} + \frac{r}{f_2^4} \qquad (8b)$$

$$L_1 = \rho + n_1\lambda_1 - \frac{q}{f_1^2} - \frac{1}{2}\frac{s}{f_1^3} - \frac{1}{3}\frac{r}{f_1^4} \qquad (9a)$$

$$L_2 = \rho + n_2\lambda_2 - \frac{q}{f_2^2} - \frac{1}{2}\frac{s}{f_2^3} - \frac{1}{3}\frac{r}{f_2^4} \qquad (9b)$$

where

$$q = \frac{1}{2}\int f_p^2 dl = 40.3\int N dl = 40.3\ TEC \qquad (10)$$

$$s = \int f_g f_p^2 |\cos\ \theta_B|\, dl = 7527c\int N B_0 |\cos\ \theta_B|\, dl \qquad (11)$$

$$r = 2437\int N^2 dl + 4.74\times 10^{22}\int N B_0^2 (1+\cos^2\theta_B)\, dl \qquad (12)$$

TEC is the total electron content along the line of sight, and λ is the operating wavelength. In Eqs. (8) and (9), ρ corresponds to the geometrical distance plus all the nondispersive terms that are common to both frequencies, such as clocks, transmitter and receiver delays, and the neutral atmospheric delay. In Eq. (9), $n_1\lambda_1$ and $n_2\lambda_2$ correspond to unknown integer numbers of cycles that are constants for a given transmitter and receiver pair over a continuous tracking period. In addition to the terms shown on the right-hand side of Eqs. (8) and (9), there are terms due to multipath, thermal noise, phase center variations, and a transmitter and receiver relative geometry dependent term; however, these are not the subject of this study, and are omitted from Eqs. (8) and (9).

IV. Ionospheric Layers and Geomagnetic Field Models

To proceed with the computation of the higher order delays, one has to assume models for the electron density, N, and the Earth's magnetic field, B_0. For the electron density distribution, the Chapman layer model is chosen. This model is derived by assuming a homogeneous composition for air at a constant temperature. The curvature of the Earth is neglected, and it is assumed that the atmosphere is horizontally stratified and the scale height H_s is independent of height. As the solar radiation travels downward through the atmosphere, it is absorbed and hence ionization is produced. The rate of electron production is a function of height above mean sea level h and the sun's zenith angle χ, which is the angle between the ray from the sun and the zenith. From considerations of the production of electrons by photoionization and their removal by recombination, the following formula for the electron density distribution can be obtained [22]:

$$N = N_{\max}\exp\left[\frac{1}{2}(1 - z - e^{-z}\sec\chi)\right] \qquad (13)$$

where $N_{\max}$ is the maximum value of the electron density at an altitude of $h_{\max}$ and $z = (h - h_{\max})/H_s$. When χ is near 90 deg, as near sunrise and sunset, the plane Earth approximation fails. To correct for this, $\sec\chi$ in Eq. (13) is replaced with the grazing incidence function $\mathrm{Ch}(x, \chi)$. This function, which applies accurately only to a spherically symmetric atmosphere with H_s independent of height, can be expressed as

$$\mathrm{Ch}(x,\chi) = \left(\frac{1}{2}\pi x\sin\chi\right)^{1/2} e^{1/2x\cos^2\chi} \times\left[1 \pm \mathrm{erf}\left(\frac{1}{2}x\cos^2\chi\right)^{1/2}\right] \qquad (14)$$

where $x = (\mathrm{R_E} + h)/H_s$, $\mathrm{R_E}$ is the Earth's radius, and erf(.) is the error function. The plus (minus) sign refers

to $\chi > 90$ deg ($\chi < 90$ deg). Figure 1 is a plot of the electron density distribution versus height for two different solar zenith angles $\chi = 0$ deg and $\chi = 64$ deg. In obtaining this distribution, three different Chapman layers were added together so that the distribution can resemble the ionospheric F_2, F_1, and E layers: the E layer with a maximum at 110 km, the F1 layer with a maximum at 210 km, and the F2 layer with a maximum at 350 km. This figure is representative of a daytime profile typical of a year near sunspot maximum. The D layer, which is normally present during the daytime, is not included. During nighttime, the F1 layer disappears and the electron density for a given height is about 10–100 times smaller than that of daytime. In a solar minimum, the same features (D, E, F1, and F2 layers) are preserved with the electron density scaled down roughly by a factor of 10.

Next, one must model the Earth's magnetic field. A first approximation to the geomagnetic field near the surface of the Earth is an Earth-centered dipole with its axis tilted to intersect the Earth at 78.5 deg N latitude, 291.0 deg E longitude, which corresponds to the geomagnetic north pole; and at 78.5 deg S latitude, 111.0 deg E longitude, which corresponds to the geomagnetic south pole [20] (see Fig. 2).

At this point one must distinguish between two reference frames with a common origin at the Earth's center. The geodetic frame is Earth-fixed and is given by $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$, where $\hat{\mathbf{z}}$ is along the Earth's spin axis, and $\hat{\mathbf{x}}$ is pointing toward 0 deg longitude. The geomagnetic frame, on the other hand, is obtained by first rotating the geodetic frame by an angle $\beta = 291$ deg around its $\hat{\mathbf{z}}$ axis, and then applying a second rotation by an angle $\delta = 11.5$ deg around the new $\hat{\mathbf{y}}_m$ axis (Fig. 3). This geomagnetic frame is denoted by $\hat{\mathbf{x}}_m, \hat{\mathbf{y}}_m, \hat{\mathbf{z}}_m$ and is constructed so that $\hat{\mathbf{z}}_m$ is along the magnetic dipole. A vector transformation from the geodetic to the geomagnetic frame is given by

$$\vec{V}_m = \begin{pmatrix} \cos\delta\cos\beta & \cos\delta\sin\beta & -\sin\delta \\ -\sin\beta & \cos\beta & 0 \\ \sin\delta\cos\beta & \sin\delta\sin\beta & \cos\delta \end{pmatrix} \vec{V} \qquad (15)$$

At a point on the Earth's surface, local geodetic east, north, and vertical are denoted by $\hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}}$, and geomagnetic east, north, and vertical are denoted by $\hat{\mathbf{X}}_m, \hat{\mathbf{Y}}_m, \hat{\mathbf{Z}}_m$ (Fig. 3). The magnetic field vector is given by

$$\vec{B}_0 = B_g \left(\frac{\mathrm{R_E}}{r_m}\right)^3 \sin\theta_m \hat{\mathbf{Y}}_m - 2B_g \left(\frac{\mathrm{R_E}}{r_m}\right)^3 \cos\theta_m \hat{\mathbf{Z}}_m \quad (16)$$

where r_m is the radial distance, and θ_m is the magnetic colatitude. The value B_g is the amplitude of the magnetic field at the Earth's surface at the magnetic equator, and is equal to 3.12×10^{-5} tesla.

V. Analysis

A. First-Order Effect

According to Eqs. (8)–(10), the first-order ionospheric delay can be written as $4.48 \times 10^{-16} \lambda^2 TEC$ (meters). For the GPS L1 and L2 frequencies, respectively, this translates to 16.2 cm and 26.7 cm of group delay (or phase advance) for every one TEC unit (1 TEC unit = 10^{16} electrons/m^2). Daytime and nighttime, as well as solar minimum and maximum ground TEC measurements, vary between 1 and 500 TEC units. Therefore, first-order ionospheric group delay (phase advance) ranges between ~0.2 and 80 m for L1 and ~0.3 and 130 m for L2.

The first-order ionospheric term, which is about three orders of magnitude larger than higher order terms, can be eliminated by using the "ionospheric free" linear combination, which, based on Eq. (8), is given by

$$\left(\frac{f_1^2}{f_1^2 - f_2^2}\right) P_1 - \left(\frac{f_2^2}{f_1^2 - f_2^2}\right) P_2 = \rho - \frac{s}{f_1 f_2 (f_2 + f_1)} - \frac{r}{f_1^2 f_2^2} \qquad (17)$$

As the first-order ionospheric term is eliminated, the dominant ionospheric errors are due to the second- and third-order terms, which are discussed below.

B. Second-Order Effect

The term $B_0 |\cos\theta_B|$ in Eq. (11) represents the absolute value of the component of the $\vec{B}_0$ field along the line of propagation; therefore, it can be replaced by $|\vec{B}_0 \cdot \vec{k}|$, where $(\cdot)$ represents the inner product and $\vec{k}$ is the unit vector in the direction of propagation.

Consider a station with magnetic colatitude and longitude θ_m and ϕ_m, respectively, observing a satellite with elevation E_m and azimuth A_m, where A_m is measured from magnetic north. Then $\vec{k}$ is given by

$$\vec{k} = -(\cos E_m \sin A_m \hat{\mathbf{X}}_m + \cos E_m \cos A_m \hat{\mathbf{Y}}_m + \sin E_m \hat{\mathbf{Z}}_m) \qquad (18)$$

therefore,

$$\left|\mathbf{B}_0 \cdot \vec{k}\right| = B_g \left(\frac{\mathrm{R_E}}{r_m}\right)^3 \left|\sin\theta_m' \cos E_m \cos A_m - 2\cos\theta_m' \sin E_m\right| \quad (19)$$

where θ_m', r_m are the magnetic colatitude and radial distance of a point along the link, respectively. This term, multiplied by the electron density, is the integrand of Eq. (11), where one must think of r_m and θ_m' as varying along the line of integration. While the exact distribution of electron density along the line of sight is needed to calculate the second-order delay term, a useful approximation can be derived by assuming that the ionosphere consists of a very thin layer at altitude H. Then, the corresponding r_m and θ_m' at the intersection point between the line of sight and the ionospheric layer are given (for $E_m > 10$ deg) by

$$r_m = \mathrm{R_E} + H \quad (20a)$$

$$\theta_m' = \theta_m - \frac{H}{\mathrm{R_E}\sin E_m}\cos A_m \cos E_m + O\left(\frac{H^2}{\mathrm{R_E}^2}\right) \quad (20b)$$

By combining Eqs. (8), (11), and (19), one can approximate the second-order ionospheric group delay (in units of distance) by

$$\text{second order ion. group delay} = 2.61 \times 10^{-18}\lambda^3 \left(\frac{\mathrm{R_E}}{r_m}\right)^3 \times \left|\sin\theta_m' \cos E_m \cos A_m - 2\cos\theta_m' \sin E_m\right| TEC \quad (21)$$

where r_m and θ_m' are given by Eq. (20). Setting H at 300 km and ignoring the factor between the absolute signs, Eq. (21) implies that in the dipole approximation, the second-order ionospheric group delay is on the order of 0.16 mm and 0.33 mm for L1 and L2, respectively, for each TEC unit. The second-order ionospheric phase advance, on the other hand, is one-half of this effect. When forming the ionospheric free linear combination, some cancellation in the second-order term takes place; the residual range error (RRE), which is defined as the difference between the dual-frequency corrected range [left-hand side of Eq. (17)] and the true range, is then on the order of −0.11 mm per TEC unit.

The relations between the magnetic colatitude and longitude, θ_m and ϕ_m, and the geographical colatitude and longitude, θ and ϕ, are given by

$$\cos\theta_m = \sin\delta\cos\beta\sin\theta\cos\phi + \sin\delta\sin\beta\sin\theta\sin\phi + \cos\delta\cos\theta \quad (22)$$

$$\tan\phi_m = \frac{-\sin\beta\sin\theta\cos\phi + \cos\beta\sin\theta\sin\phi}{\cos\delta(\cos\beta\sin\theta\cos\phi + \sin\beta\sin\theta\sin\phi) - \sin\delta\cos\theta} \quad (23)$$

The satellite elevation in local magnetic east–north–vertical coordinates, E_m, is the same as the elevation in local geodetic east–north–vertical coordinates, E. On the other hand, the azimuths in these two coordinates are related through

$$A_m = A + \arccos(\sin\phi\sin\phi_m\cos\delta\cos\beta + \cos\phi\cos\phi_m\cos\beta + \sin\phi\cos\phi_m\sin\beta - \cos\phi\sin\phi_m\cos\delta\sin\beta) \quad (24)$$

Figure 4 shows the absolute value of the RRE due to the second-order term. This is shown for two stations at different longitudes and latitudes, tracking different GPS satellites, as indicated on the figure. These errors are calculated using the exact integral form of Eq. (11) and assuming the Chapman layer distribution of Fig. 1 and the magnetic field of a tilted dipole, as described above. The angle χ in Eqs. (13) and (14) is determined based on the assumption that the $\hat{\mathbf{x}}$ axis (Fig. 3) is pointing toward the sun at 12h UT. The exact calculation, referred to as truth, is compared with an approximation obtained from Eqs. (20)–(24). According to the examples of Fig. 4, the true second-order absolute RRE has an rms value of 1.25 cm, and can be as large as 4 cm at the lowest elevation angle (10 deg). Using the thin-layer model at the 300-km altitude as described above, it is possible to approximate this effect to better than 90 percent on the average. The difference between the truth and the approximation has an average of 0.11 cm and a variance of 0.25 cm. This suggests that a thin-layer model of the ionosphere can be very useful in calibrating the second-order ionospheric effect and therefore improving GPS-user range measurements.

C. Third-Order Effect

Upon examining Eq. (12), one finds that the second term, except during times of very strong magnetic storms, contributes no more than a submillimeter of range error for gigahertz frequencies. Therefore, one must consider the first term, which can be simplified to (in units of meters, kilograms, and seconds)

$$\text{third-order ion. group delay} = 3.0 \times 10^{-31} \lambda^4 \int N^2 dl \tag{25}$$

To get an approximate estimate of the integral of Eq. (25), the authors use the shape parameter η, defined by Brunner and Gu [17] as

$$\eta \equiv \frac{\int N^2 dl}{N_{\max} \int N dl} \tag{26}$$

For a single Chapman layer, η was estimated to be ~0.66 and almost independent of elevation [23,17]. Since this ionospheric profile is dominated by a single layer (F2), the authors believe that the shape parameter η in this case will be close to 0.66. Therefore, one can approximate the integral of Eq. (25) by $0.66 \times N_{\max} \times TEC$. For $N_{\max} = 3.0 \times 10^{12} (e/\mathrm{m}^3)$ and $TEC = 10^{18} (e/\mathrm{m}^2)$ the third-order term is estimated to be ~0.86 mm for L1, ~2.4 mm for L2, and ~–0.66 mm for the RRE. A more exact estimate of the third-order term based on Eq. (12) and the Chapman distribution of Fig. 1 is shown in Fig. 4. In the examples of Fig. 4, the delay ranges between 1 and 4 mm.

VI. Conclusion

The above results are summarized in Table 1, which shows the amount of group delay due to first-, second-, and third-order ionospheric terms in the zenith direction, assuming a zenith $TEC = 10^{18} (e/\mathrm{m}^2)$.

In employing a Chapman distribution and a dipole approximation for the magnetic field, it was possible to estimate the higher order ionospheric effects on range and phase measurements. The second-order error can be several centimeters for range as well as phase during daytime, for a year near sunspot maximum. Moreover, since the magnetic field is fixed to the Earth, and the GPS orbit, as seen from a ground station, repeats itself daily (shifted by ~4 min per day), the diurnal shape of the second-order error is most likely to repeat its overall structure for several days, at least to the extent that the overall electron density distribution remains unchanged. Such daily repeatable errors in range and phase will be mapped directly into orbital and baseline estimation. This study shows that a rough ionospheric model consisting of a thin shell at 300 km, plus a knowledge of the TEC, allows one to calibrate the second-order term to better than 90 percent. This implies reducing the second-order ionospheric error to less than 2 mm on the average and, therefore, potentially improving orbit determination and baseline solutions.

Acknowledgments

The authors wish to thank Sien Wu and Thomas Yunck of JPL for helpful comments and suggestions on this article.

References

[1] T. P. Yunck, W. G. Melbourne, and C. L. Thornton, "GPS-based satellite tracking system for precise positioning," *IEEE Tr. Geosci. and Rem. Sensing*, vol. GE-23, pp. 450–457, July 1985.

[2] S. M. Lichten and J. S. Border, "Strategies for high precision Global Positioning System orbit determination," *J. Geoph. Res.*, vol. 92, no. 10, pp. 12,751–12,762, November 1987.

[3] T. P. Yunck, S. C. Wu, and J. T. Wu, "Precise Near-Earth Navigation With GPS: A Survey of Techniques," *TDA Progress Report 42-91*, vol. July–September 1987, Jet Propulsion Laboratory, Pasadena, California, pp. 29–45, August 15, 1987.

[4] S. C. Wu, T. P. Yunck, and C. L. Thornton, "Reduced Dynamic Technique for Precise Orbit Determination of Low Earth Satellites," *J. Guidance, Control and Dynamics*, vol. 14, no. 1, pp. 24–30, January–February 1991.

[5] S. M. Lichten, "Toward GPS Orbit Accuracy of Tens of Centimeters," *Geoph. Res. Let.*, vol. 17, no. 3, pp. 215–218, March 1990.

[6] S. M. Lichten, "Estimation and filtering for high-precision GPS positioning application," *Manuscripta Geodaetica*, vol. 15, pp. 159–179, April 1990.

[7] C. L. Thornton, J. L. Fanselow, and N. A. Renzetti, "GPS-based geodetic measurement systems," *Space Geodesy and Geodynamics*, A. Anderson and A. Cazenave, eds., New York: Academic Press, 1986.

[8] J. N. Kellogg and T. H. Dixon, "Central and South America GPS geodesy—CASA UNO," *Geoph. Res. Let.*, vol. 17, no. 3, pp. 195–198, March 1990.

[9] J. T. Freymueller and J. N. Kellogg, "The Extended Tracking Network and Indications of Baseline Precision and Accuracy in the North Andes," *Geoph. Res. Let.*, vol. 17, no. 3, pp. 207–210, March 1990.

[10] R. P. Malla and S. C. Wu, "GPS Inferred Geocentric Reference Frame for Satellite Positioning and Navigation," *Bul. Geod.*, vol. 63, pp. 263–279, 1989.

[11] A. P. Freedman, "Measuring Earth Orientation With Global Positioning System," *Bul. Geod.*, vol. 65, pp. 53–65, 1991.

[12] W. I. Bertiger, J. T. Wu, and S. C. Wu, "Gravity Field Improvement Using GPS Data From TOPEX/POSEIDON: A Covariance Analysis," *J. Geoph. Res.*, vol. 97, no. B2, pp. 1965–1971, February 10, 1992.

[13] S. M. Lichten, "Precise Estimation of Tropospheric Path Delays With GPS Techniques," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1–12, February 15, 1990.

[14] D. M. Tralli and S. M. Lichten, "Comparison of Kalman filter estimates of zenith atmospheric path delays using the Global Positioning System and very long baseline interferometry," submitted to *Radio Science*.

[15] D. Coco, "GPS—Satellites of opportunity for ionospheric monitoring," *GPS World*, vol. 2, no. 9, pp. 47–50, October 1991.

[16] J. J. Spilker, "GPS Signal Structure and Performance Characteristics," *Navigation*, vol. 25, pp. 29–54, 1978.

[17] F. K. Brunner and M. Gu, "An Improved Model for the Dual Frequency Ionospheric Correction of GPS Observations," *Manuscripta Geodaetica*, vol. 16, no. 3, pp. 205–214, 1991.

[18] M. Gu and F. K. Brunner, "Theory of the Two Frequency Dispersive Range Correction," *Manuscripta Geodaetica*, vol. 15, pp. 357–361, 1990.

[19] K. G. Budden, *The Propagation of Radio Waves*, New York: Cambridge Press, 1985.

[20] A. S. Jursa, ed., *Handbook of Geophysics and the Space Environment*, Air Force Geophysics Laboratory, National Technical Information Services, Springfield, Virginia, 1985.

[21] C. H. Papas, *Theory of Electromagnetic Wave Propagation*, New York: McGraw-Hill, 1965.

[22] H. Rishbeth and O. K. Garriott, *Introduction to Ionospheric Physics*, New York: Academic Press, 1969.

[23] G. K. Hartmann and R. Leitinger, "Range Errors Due to Ionospheric and Tropospheric Effects for Signal Frequencies Above 100 MHz," *Bul. Geod.*, vol. 58, pp. 109–136, 1984.

Table 1. Estimated zenith ionospheric group delay due to $1/f^2$, $1/f^3$, and $1/f^4$ terms, for an arbitrary wavelength λ (microwave region), L1 and L2 frequencies as well as the residual range error with dual-frequency calibration. It is assumed that the zenith $TEC = 10^{18}$ (e/m^2). The phase advance can be read from this table by multiplying each number by = 1, –1/2, and –1/3 for the $1/f^2$, $1/f^3$, and $1/f^4$ terms, respectively.

Ionospheric expansion term	λ, MKS[a] units	L1	L2	RRE
$1/f^2$	$4.48 \times 10^{-16}\lambda^2 TEC$	16.2 m	26.7 m	0.0
$1/f^3$	$\approx\ a\ 2.61 \times 10^{-18}\lambda^3 TEC$ $(0 < a < 2)$	~1.6 cm	~3.3 cm	~ –1.1 cm
$1/f^4 (N_{max} = 3.0 \times 10^{12} e/m^2)$	$\approx 2.0 \times 10^{-31}\lambda^4 N_{max} TEC$	~0.86 mm	~2.4 mm	~ –0.66 mm
Calibrated $1/f^3$ based on a thin-layer ionospheric model				~ 1–2 mm

[a]Meters, kilograms, and seconds.

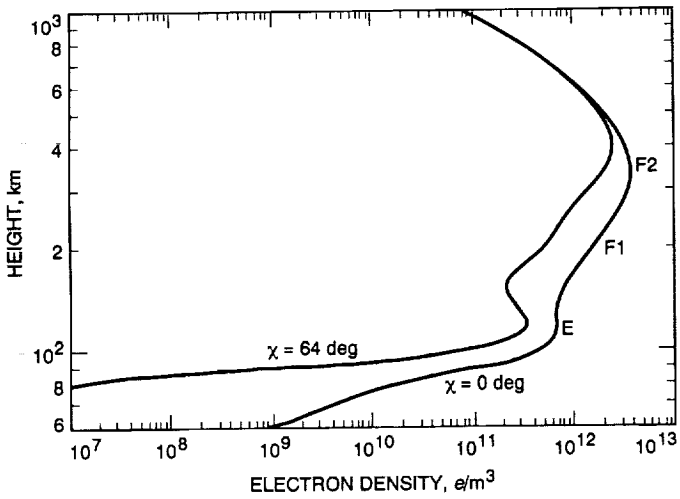

Fig. 1. Ionospheric profile modeled as the sum of three Chapman layers.

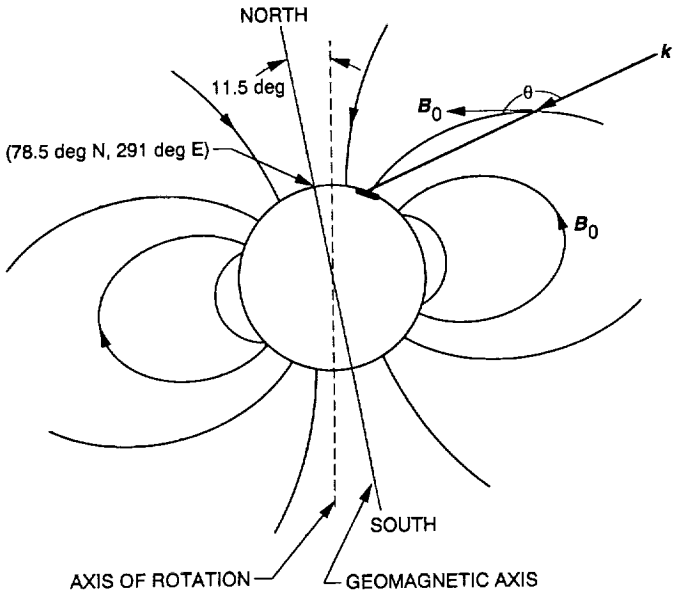

Fig. 2. The Earth's magnetic field modeled as an Earth-centered dipole, aligned along the geomagnetic axis.

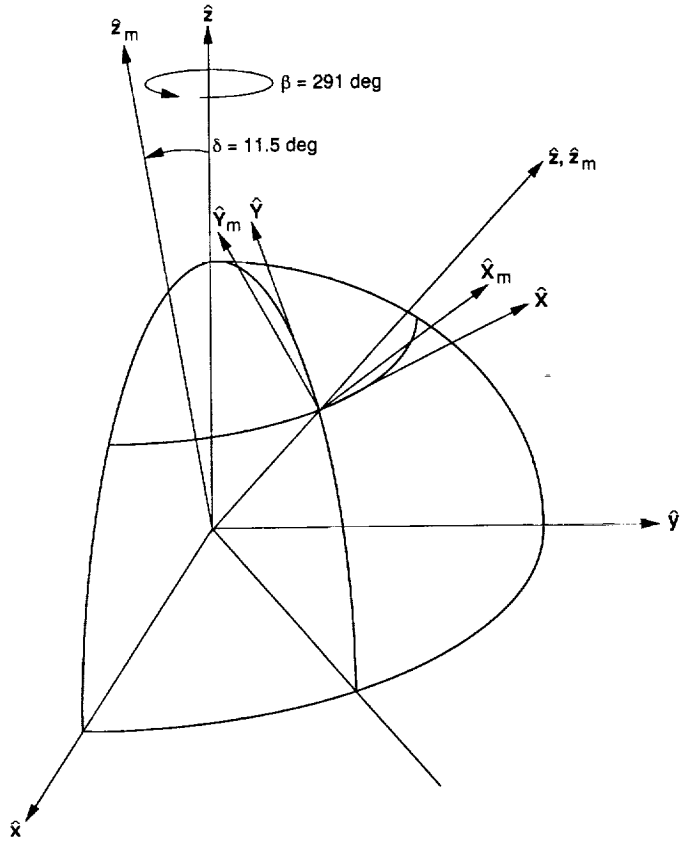

Fig. 3. A graphical illustration of all the frames used in the text. The vectors $\hat{x}$, $\hat{y}$, and $\hat{z}$ correspond to the geodetic frame; the vectors $\hat{x}_m$, $\hat{y}_m$, and $\hat{z}_m$ correspond to the geomagnetic frame; the vectors $\hat{X}_m$, $\hat{Y}_m$, and $\hat{Z}_m$ correspond to geodetic local east, north, and vertical; and the vectors $\hat{X}_m$, $\hat{Y}_m$, and $\hat{Z}_m$ correspond to geomagnetic local east, north, and vertical.

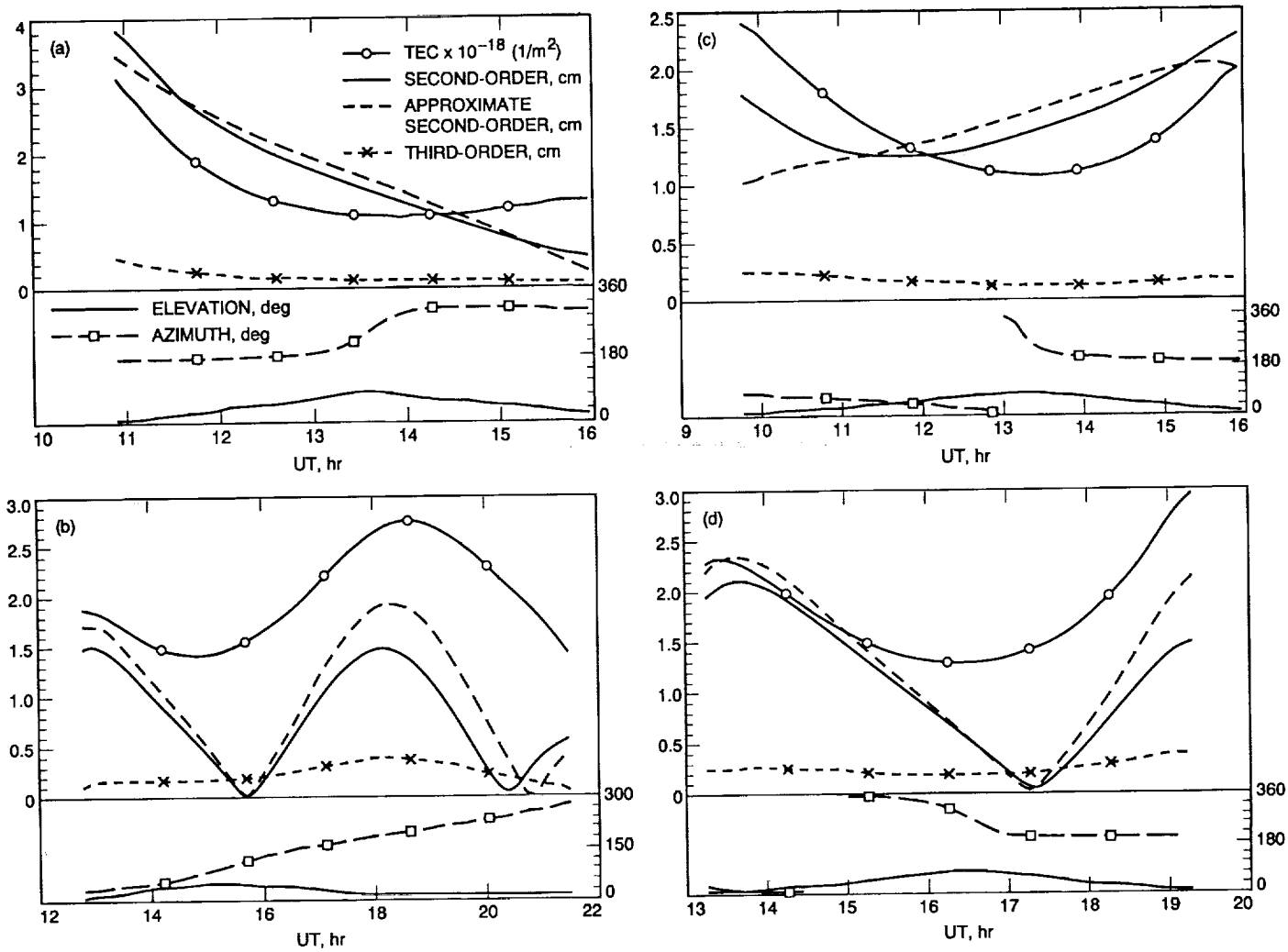

Fig. 4. TEC and absolute second- and third-order ionospheric residual range errors along the line of sight for different GPS–ground receiver pairs. Shown also are the elevation and azimuth of the observed satellite as functions of time: (a) GPS 1, station at 40 deg N latitude, 0 deg longitude; (b) GPS 9, station at 0 deg latitude, 75 deg W longitude; (c) GPS 20, station at 40 deg N latitude, 0 deg longitude; and (d) GPS 16, station at 0 deg latitude, 75 deg W longitude.