Chapter 1

THE SAGNAC EFFECT IN THE GLOBAL POSITIONING SYSTEM

Neil Ashby
Department of Physics, UCB 390.
University of Colorado.
Boulder, CO 80309-0390 USA.
n_ashby@mobek.colorado.edu

Abstract In the Global Positioning System (GPS) the reference frame used for navigation is an earth-centered, earth-fixed rotating frame, the WGS-84 frame. The time reference is defined in an underlying earth-centered locally inertial frame, freely falling with the earth but non-rotating, with a time unit determined by atomic clocks at rest on earth's rotating geoid. Therefore GPS receivers must apply significant Sagnac or Sagnac-like corrections, depending on how information is processed by the receiver. These corrections can be described either from the point of view of the local inertial frame, in which light travels with uniform speed c in all directions, or from the point of view of an earth-centered rotating frame, in which the Sagnac effect is described by terms in the fundamental scalar invariant that couple space and time. Such corrections are very important for comparing time standards world-wide.

1. Introduction

The purpose of the Global Positioning System (GPS) is accurate navigation on or near earth's surface. GPS also provides an accurate world-wide clock synchronization and timing system. Most GPS users are interested in knowing their position on earth; the developers of GPS have therefore adopted an Earth-Centered, Earth-Fixed (ECEF) rotating reference frame as the basis for navigation. Specifically, in the WGS-84(873) frame, the model earth rotates about a fixed axis with a defined rotation rate, $\omega_E = 7.2921151247 \times 10^{-5}$ rad s^{-1}.[1],[2]

In an inertial frame, a network of self-consistently synchronized clocks can be established either by transmission of electromagnetic signals that propagate

G. Rizzi and M.L. Ruggiero (eds.), Relativity in Rotating Frames, 11-28.

with the universal constant speed c (this is called *Einstein synchronization*), or by slow transport of portable atomic clocks. On the other hand it is well-known[3] that in a rotating reference frame, the Sagnac effect prevents a network of self-consistently synchronized clocks from being established by such processes. This is a significant issue in using timing signals to determine position in the GPS. The Sagnac effect can amount to hundreds of nanoseconds; a timing error of one nanosecond can lead to a navigational error of 30 cm.

To account for the Sagnac effect, a hypothetical non-rotating reference frame is introduced. Time in this so-called Earth-Centered Inertial (ECI) Frame is adopted as the basis for GPS time; this is discussed in Section 2. Of course the earth's mass encompasses the origin of the ECI frame and has significant gravitational effects. To an extremely good approximation in the GPS, however, gravitational effects can be simply added to other effects arising from special relativity. In this article gravitational effects will not be considered. Even time dilation, which is an effect of second order in the small parameter v/c, where v is the velocity of some clock, will be neglected. I shall confine this discussion to effects which are of first order (linear) in velocities. The Sagnac effect is such an effect.

A description of the GPS system, of the signal structure, and the navigation message, needed to understand how navigation calculations are performed, is given in Section 3. In comparing synchronization processes in the ECI frame with those in the ECEF frame, taking into account relativity principles, it becomes evident that the Sagnac effect is a manifestation of the relativity of simultaneity. Observers in the rotating ECEF frame using Einstein synchronization will not agree that clocks in the ECI frame are synchronized, due to the relative motion. In fact observers in the rotating frame cannot even globally synchronize their own clocks, due to the rotation. This is discussed in Section 4. Section 5 discusses Sagnac corrections that are necessary when comparing remote clocks on earth by observations of GPS satellites in common-view. Section 6 introduces the GPS navigation equations and discusses synchronization processes from the point of view of the rotating ECEF frame. Section 7 develops implications of the fact that GPS navigation messages provide satellite ephemerides in the ECEF frame.

2. Local Inertial Frames

Einstein's Principle of Equivalence allows one to discuss frames of reference which are freely falling in the gravitational fields of external bodies. Sufficiently near the origin of such a freely falling frame, the laws of physics are the same as they are in an inertial frame; in particular electromagnetic waves propagate with uniform speed c in all directions when measured with standard rods and atomic clocks. Such freely falling frames are called *locally inertial* frames. For the GPS, it is very useful to introduce such a frame that is non-rotating, with its origin fixed at earth's center, and which falls freely along with the earth in the gravitational fields of the other solar system bodies. This is called an Earth-Centered Inertial (ECI) frame.

Clocks in the GPS are synchronized in the ECI frame, in which self - consistency can be achieved. Thus imagine the underlying ECI frame, unattached to the spin of the earth, but with its origin at the center of the earth. In this non-rotating frame a fictitious set of standard clocks is introduced, available anywhere, all of them synchronized by the Einstein synchronization procedure and running at agreed rates such that synchronization is maintained. These clocks read the coordinate time t. Next one introduces the rotating earth with a set of standard clocks distributed around upon it, possibly roving around. One applies to each of the standard clocks a set of corrections based on the known positions and motions of the clocks. This generates a "coordinate clock time" in the earth-fixed, rotating system. This time is such that at each instant the coordinate clock agrees with a fictitious atomic clock at rest in the local inertial frame, whose position coincides with the earth-based standard clock at that instant. Thus coordinate time is equivalent to time which would be measured by standard clocks at rest in the local inertial frame. [4]

In the ECEF frame used in the GPS, the unit of time is the SI second as realized by the clock ensemble of the U. S. Naval Observatory, and the unit of length is the SI meter. In summary, the reference frame for navigation is the rotating WGS-84 frame, but clocks are synchronized in the underlying hypothetical ECI frame with a unit of time defined by clocks (essentially on the geoid) and a unit of length determined by the defined value of the speed of light, $c = 299792458$ m/s.

3. The GPS

The Global Positioning System can be described in terms of three principal "segments:" a Space Segment, a Control Segment, and a User Segment. The Space Segment consists essentially of 24 satellites carrying atomic clocks. (Spare satellites and spare clocks in satellites exist.) There are four satellites in each of six orbital planes inclined at 55° with respect to earth's equatorial plane, distributed so that from any point on the earth, four or more satellites are almost always above the local horizon. Tied to the clocks are navigation and timing signals that will be discussed below.

The Control Segment is comprised of a number of ground-based monitoring stations which continually gather information from the satellites. These data are sent to a Master Control Station in Colorado Springs, CO, which analyzes the constellation and projects the satellite ephemerides and clock behavior forward for the next few hours. This information is then uploaded into the satellites for retransmission to users.

The User Segment consists of all users who, by receiving signals transmitted from the satellites, are able to determine their position, velocity, and the time on their local clocks.

The timing signals transmitted from each satellite are right circularly polarized. A carrier signal of frequency 1.542 MHz is modulated with a series of phase reversals; these phase reversals carry information bits from the transmitter to the receiver. Such phase reversals are conceptually important because the phase of an electromagnetic wave is a relativistic scalar. The phase reversals correspond to physical points in spacetime at which - for all observers - the electric and magnetic fields vanish.

The navigation message contained in these bit streams include values of parameters from which the receiver can compute the satellite's position *in the rotating ECEF frame*, as a function of time of transmission. Also the GPS time on the satellite clock is indicated by a particular phase reversal in the sequence. A receiver distinguishes the signal from a particular satellite by comparing the bit streams, that are unique to each satellite, with bit streams generated by electronic circuitry within the receiver.

Additional information contained in the messages includes an almanac for the entire satellite constellation, information about satellite vehicle health, and information from which Universal Coordinated Time as maintained by the U. S. Naval Observatory–UTC(USNO)–can be determined.

The GPS is a navigation and timing system that is operated by the United States Department of Defense (DoD), and therefore has a number of aspects to it which are classified. Several organizations monitor GPS signals independently and provide services from which satellite ephemerides and clock behavior can be obtained. Accuracies in the neighborhood of 5-10 cm are not unusual. Carrier phase measurements of the transmitted signals are commonly done to better than a millimeter.

For purposes of the remainder of this article, I shall think of a signal from a GPS satellite as containing within itself information about the position and time of a transmission "event". The position is specified in the rotating ECEF frame. GPS time is time in an underlying local inertial frame. The signal propagates with speed c in a straight line in the ECI frame to the receiver, where it is decoded and its arrival time t_R is compared to the time of transmission t_T. The receiver can then form the so-called pseudoranges

$$\rho = c(t_R - t_T). \tag{1.1}$$

A receiver continually forms such pseudoranges for each satellite being observed. A signal can be imagined abstractly as propagating with speed c from transmitter to receiver in a straight line in the ECI frame, with position and

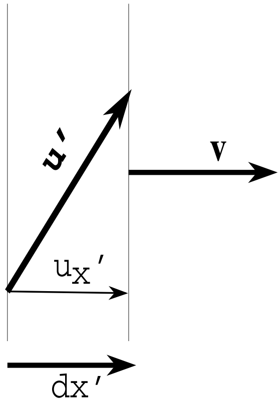

Figure 1.1. Synchronization by transmission of a signal

time of the transmission event “known" by the receiver. Possible clock biases in the receiver prevent the GPS time of the reception event from being known a priori.

4. Relativity of Simultaneity

To establish the connection between the Sagnac effect and the relativity of simultaneity, consider an observer moving with velocity $\mathbf{v}$ in the x direction relative to an inertial frame such as the ECI frame. To be specific, one can imagine measurements of unprimed quantities such as $\mathbf{v}$ and signal velocity $\mathbf{u}$ to be performed in the ECI frame, while primed quantities such as $\mathbf{u}'$ are measured in the rest frame of the moving observer. Referring to Figure 1.1, let a signal be travelling with speed components (u'_x, u'_y) (measured in the moving observer's frame). The vertical lines represent planes at x' and $x'+dx'$. The signal travels a distance dx' in the x direction and the moving observer desires to use this signal to transfer time from clocks in the plane at x' to clocks in the plane at $x' + dx'$. Here I am neglecting higher-order terms in the velocity so $dx = dx'$, there being no appreciable Lorentz contraction. Let the components of signal speed in the ECI frame be (u_x, u_y). The well-known Lorentz transformations for speed include the expression

$$u_x = \frac{u'_x + v}{1 + \frac{u'_x v}{c^2}} \, . \tag{1.2}$$

The terms in the denominator of this expression arise from the time-component of the ordinary Lorentz transformation. In particular the second term in the denominator arises from the relativity of simultaneity, a consequence of the constancy of the speed of light. We wish to compare the propagation time of this signal, measured by the moving observer, with the propagation time measured in the ECI frame. The analysis is performed in the ECI frame.

If the moving observer moves a distance vdt in time dt, then the total distance travelled by the signal in the x-direction is $u_x dt$, which is comprised of two contributions: the distance dx, plus the distance vdt required to catch up to the plane at $x' + dx'$. Thus

$$u_x dt = dx + vdt, \tag{1.3}$$

and therefore the time required is

$$dt = \frac{dx}{u_x - v}\,. \tag{1.4}$$

But from the expression for the Lorentz transformation of speed, keeping only terms of linear order in v,

$$u_x - v \approx \frac{u'_x}{1 + \frac{u'_x v}{c^2}}\,. \tag{1.5}$$

and therefore

$$dt = \frac{dx}{u'_x} + \frac{vdx}{c^2}\,. \tag{1.6}$$

The first term in this result is just the time required, in the moving frame, for the signal to travel from the x' plane to the $x' + dx'$ plane. If the moving observer ignores the motion relative to the ECI frame, this would be the time used to synchronize clocks in the $x' + dx'$ plane to clocks in the x' plane. The second term is the additional time required to *synchronize the clocks in the ECI frame*. Note that in this second term, the value of u'_x has cancelled out, so that the value of the signal speed is irrelevant. The signal could be a light signal travelling in a fiber of index of refraction n, or it could even be an acoustic signal. The signal speed could even be variable, the last term would not be affected.

Consider for example an optical fiber loop of length L and index of refraction n which by means of a system of pulleys is made to move with speed v around in a closed circuit, relative to an inertial frame. The circuit itself could be of any shape, such as a figure 8 or an oval. In such a case it is not useful to speak of rotation, although Eq. 1.6 applies to the rotational case as well. Eq. 1.6 applies to each infinitesimal segment of the moving loop, since one can imagine a sequence of moving reference frames each of which is instantaneously at rest with respect to the moving fiber loop and in which Eq. 1.6 is valid. If a signal travels around the loop in a direction parallel to the velocity, then from Eq. 1.6, the total time required for the signal to make one circuit is

$$\Delta t_+ = \oint dt = \oint \frac{dx}{u'_x} + \frac{vL}{c^2}\,, \tag{1.7}$$

and the time required for the signal to make one circuit in the direction opposite to the velocity is

$$\Delta t_- = \oint \frac{dx}{u'_x} - \frac{vL}{c^2}\,, \tag{1.8}$$

The difference is

$$\Delta t = \Delta t_+ - \Delta t_- = \frac{2vL}{c^2}\,, \tag{1.9}$$

and for two counterpropagating monochromatic beams this can be converted into an observable interference fringe shift. If the beams are recombined in the ECI frame where they have angular frequency ω, then the phase difference will be

$$\Delta\phi = \omega \Delta t\,. \tag{1.10}$$

The Sagnac effect in a moving fiber loop is independent of the fiber's index of refraction or of the shape of the loop. This has been confirmed in recent experiments.[5]

For example for electrons of energy $E = \hbar\omega$, the phase difference will be

$$\Delta\phi = \frac{2EvL}{\hbar c^2}\,. \tag{1.11}$$

Interference experiments with electrons have been reported in reference [6], which also has a comprehensive discussion of the many different points of view of the Sagnac effect that can be taken.

In the GPS, a decision was made to synchronize GPS clocks in the ECI reference frame. The above discussion demonstrates that observers on earth, in the ECEF frame, must apply a "Sagnac" correction (the second term in Eq. 1.6) to their synchronization processes in order to synchronize their clocks to GPS time.

The correction can be generalized slightly by noting that the distance dx is in the same direction as the relative velocity $\mathbf{v}$. If $d\mathbf{r}$ is the vector increment of path in the direction of signal propagation, then the Sagnac correction term can be written

$$dt_{Sagnac} = \frac{\mathbf{v}\cdot d\mathbf{r}}{c^2}\,. \tag{1.12}$$

For applications in the GPS, it is useful to describe this correction term another way, in terms of accounting for motion of the receiver during propagation of signals from transmitters to receivers. Henceforth only signals propagating with speed c will be considered. This assumption also applies to measurements made locally by the moving observer in the ECEF frame, since at each instant the measurements of distance and time intervals are the same as they would be in an inertial frame which instantaneously coincides with the observer in the ECEF frame and which moves with the instantaneous velocity $\mathbf{v}$ of the ECEF observer. In Eq. 1.3, the velocity v is present to account for the fact that the signal must catch up to the position at $x' + dx'$ which is moving with velocity $\mathbf{v}$, and to first order in the small quantity v/c leads directly to the Sagnac correction term in Eqs. 1.3 and 1.12. The Sagnac correction can thus be interpreted as an effect which arises in the ECEF frame when one accounts for motion of the receiver during propagation of the electromagnetic signal with speed c.

5. Time Transfer with the GPS

In the GPS navigation is accomplished by means of signals from four or more satellites, whose arrival times are measured at the location of the receiver. I now consider one such signal in space, transmitted from satellite position $\mathbf{r}_T$ at GPS time t_T. Let the receiver position at GPS time t_T be $\mathbf{r}_R$, and let the receiver have velocity $\mathbf{v}$ in the ECI frame. Let the signal (considered abstractly as a pulse) arrive at the receiver at time t_R. During the time interval $\Delta t = t_R - t_T$, the displacement of the receiver is $\mathbf{v}\Delta t$. Since the signal travels with speed c, the constancy of the speed of light c implies that

$$c^2(\Delta t)^2 = (\mathbf{r}_R + \mathbf{v}\Delta t - \mathbf{r}_T)^2 \,. \tag{1.13}$$

To simplify the equation, I define

$$\mathbf{R} = \mathbf{r}_R - \mathbf{r}_T \,. \tag{1.14}$$

Then to leading order in v,

$$c^2(\Delta t)^2 = (\mathbf{R} + \mathbf{v}\Delta t)^2 \approx R^2 + 2\mathbf{v}\cdot\mathbf{R}\Delta t \,. \tag{1.15}$$

Taking the square root of both sides of Eq. (1.15) and again expanding to leading order in v gives

$$c\Delta t = R + \frac{\mathbf{v}\cdot\mathbf{R}\Delta t}{R} \,. \tag{1.16}$$

This equation can be solved approximately for Δt to give

$$\Delta t = \frac{R}{c} + \frac{\mathbf{v}\cdot\mathbf{R}}{c^2} \,. \tag{1.17}$$

The second term in Eq. 1.17 is the Sagnac correction term, which arises when one accounts for motion of the receiver while the signal propagates from transmitter to receiver. This is illustrated in Figure 1.2.

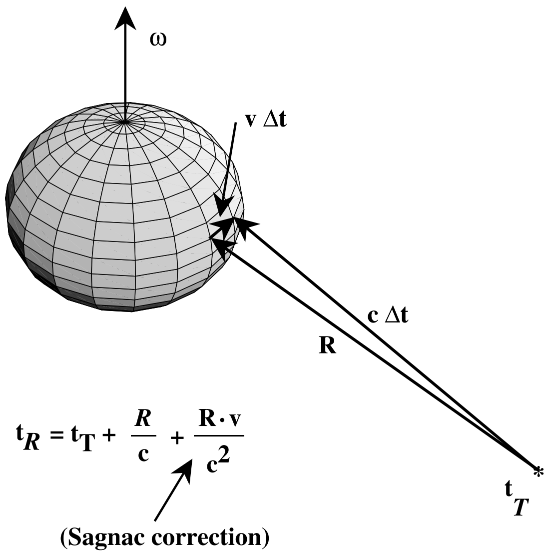

Figure 1.2. Sagnac correction arising from motion of the ECEF observer.

Suppose that the receiver is fixed to the surface of the earth, at a well-surveyed location so that the receiver position $\mathbf{r}_R$ is well known at all times. The velocity of the receiver will be just that due to rotation of the earth with angular velocity $\omega_{\mathbf{E}}$, so

$$\mathbf{v} = \omega_{\mathbf{E}} \times \mathbf{r}_R \,, \tag{1.18}$$

We take $\mathbf{r}_R$ to be the vector from earth's center to the receiver position. Then the Sagnac correction term can be rewritten as

$$\Delta t_{Sagnac} = \frac{\omega_{\mathbf{E}} \times \mathbf{r}_R \cdot \mathbf{R}}{c^2} = \frac{2\omega_{\mathbf{E}}}{c^2} \cdot \left(\frac{1}{2} \mathbf{r}_R \times \mathbf{R} \right) . \tag{1.19}$$

The quantity $2\omega_E/c^2$ has the value

$$\frac{2\omega_E}{c^2} = 1.6227 \times 10^{-21}\ \mathrm{s/m^2} = 1.6227 \times 10^{-6}\ \mathrm{ns/km^2} \,. \tag{1.20}$$

The last factor in Eq. 1.19 can be interpreted as a vector area $\mathbf{A}$:

$$\mathbf{A} = \frac{1}{2} \mathbf{r}_R \times \mathbf{R} \,. \tag{1.21}$$

The only component of $\mathbf{A}$ which contributes to the Sagnac correction is along earth's angular velocity vector $\omega_{\mathbf{E}}$, because of the dot product that appears in the expression. This component is the projection of the area onto a plane normal to earth's angular velocity vector. This leads to a simple description of the Sagnac correction: Δt_{Sagnac} is $2\omega_E/c^2$ time the area swept out by the electromagnetic pulse as it travels from the GPS transmitter to the receiver, projected onto earth's equatorial plane. This is depicted in Figure 1.3, in which the receiver is on earth's surface at the tip of the path vector $\mathbf{R}$.

In the early 1980s clocks in remotely situated timing laboratories were being compared by using GPS satellites in "common view", that is when one GPS satellite is observed at the same time by more than one timing laboratory. In one experiment[7] signals from GPS satellites were utilized in simultaneous common view between three pairs of earth timing centers to accomplish a circumnavigation of the globe. The centers were the National Bureau of Standards (now the National Institute of Standards and Technology) in Boulder, Colorado; Physikalisch-Technische Bundesanstalt in Braunschweig, West Germany; and Tokyo Astronomical Observatory. A typical geometrical configuration of ground stations and satellites, with the corresponding projected areas, is illustrated in Figure 1.4. The size of the Sagnac effect calculated varies from about 240 ns to 350 ns depending on the location of the satellites at a particular moment. Sufficient data were collected to perform 90 independent circumnavigations. As Figure 1.4 shows, when a satellite is eastward of one timing center and westward of another, one of the Sagnac corrections is positive and the other is negative, so when computing the difference of times between the two terrestrial clocks, the Sagnac corrections actually add up in a positive sense.

The mean value of the residuals over 90 days of observation was 5 ns, less than 2 percent of the magnitude of the calculated total Sagnac correction. A significant part of these residuals can be attributed to random noise processes in the clocks.

Sagnac corrections of the form of Eq. 1.19 are routinely used in comparisons between distant time standards laboratories on earth.

6. GPS Navigation Equations and the ECEF Frame

The navigation problem in GPS is to determine the position of the receiver in the ECEF reference frame. A by-product of this process is the accurate determination of GPS time at the receiver. In general neither the position nor the time is known, so the assumptions used in previous sections regarding the Sagnac effect are of little use. The principles of position determination and time transfer in the GPS can be very simply stated. Let there be four synchronized atomic clocks which transmit sharply defined pulses from the positions $\mathbf{r}_j$ at times t_j, with $j = 1, 2, 3, 4$ an index labelling the different transmission events. Suppose that these four signals are received at position $\mathbf{r}$ at one and the same instant t. This is called "time-tagging at the receiver", meaning that observations of the various signals are made simultaneously at the receiver at

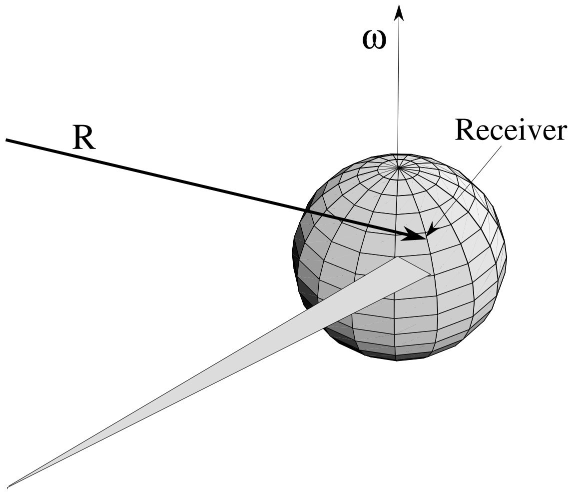

Figure 1.3. Sagnac correction arising from motion of the ECEF observer.

time t. Then from the principle of the constancy of the speed of light,

$$c^2(t - t_j)^2 = |\mathbf{r} - \mathbf{r}_j|^2\,, \quad j = 1, 2, 3, 4. \tag{1.22}$$

These four equations can be solved for the unknown space-time coordinates of the reception event, $(t, \mathbf{r})$. The solution will provide the position of the receiver at the time of the simultaneous reception events, t. No knowledge of the receiver velocity is needed. The Sagnac effect becomes irrelevant. At most one can say that because the solution gives the final position and time of the reception event, the Sagnac effect has been automatically accounted for.

However there are complications from the fact that the navigation equations, Eqs. 1.22, are valid in the ECI frame, whereas users almost always want to know their position in the ECEF frame. For discussions of relativity, the particular choice of ECEF frame is immaterial. Also, the fact the the earth truly rotates about an axis slightly different from the WGS-84 axis, with a variable rotation rate, has little consequence for relativity and I shall not go into this here.

It should be emphasized strongly that the transmitted navigation messages provide the user only with a function from which the satellite position can be calculated *in the ECEF* as a function of the transmission time. Usually the satellite transmission times t_j are unequal, so the coordinate system in which the satellite positions are specified changes orientation from one measurement

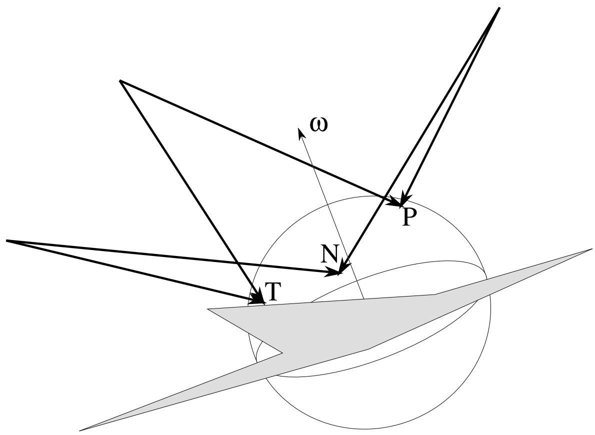

Figure 1.4. Common-view signals from three satellites provide an Around-the-World Sagnac experiment.

to the next. Therefore to implement Eqs. (1.22), the receiver must generally perform a different rotation for each measurement made, into some common inertial frame, so that Eqs. (1.22) apply. After solving the propagation delay equations, a final rotation must then be performed into the ECEF to determine the receiver's position. This can become exceedingly complicated and confusing. I shall discuss this in a later section.

The purpose of the present discussion is to examine first-order relativistic effects from the point of view of the ECEF frame. Consider the simplest instance of a transformation from an inertial frame, in which the space-time is Minkowskian, to a rotating frame of reference. Thus ignoring gravitational potentials, the metric in an inertial frame in cylindrical coordinates is

$$-ds^2 = -(c\,dt)^2 + dr^2 + r^2 d\phi^2 + dz^2\,, \tag{1.23}$$

and the transformation to a coordinate system $\{t', r', \phi', z'\}$ rotating at the uniform angular rate ω_E is

$$t = t', \quad r = r', \quad \phi = \phi' + \omega_E t', \quad z = z'\,. \tag{1.24}$$

This results in the following well-known metric (Langevin metric) in the rotating frame:

$$-ds^2 = -\left(1 - \frac{\omega_E^2 r'^2}{c^2}\right)(cdt')^2 + 2\omega_E r'^2 d\phi' dt' + (d\sigma')^2\,, \tag{1.25}$$

where the abbreviated expression $(d\sigma')^2 = (dr')^2 + (r'd\phi')^2 + (dz')^2$ for the square of the coordinate distance has been used.

The time transformation $t = t'$ in Eqs. (1.24) is a result of the convention to determine time t' in the rotating frame in terms of time in the underlying ECI frame.

Now consider a process in which observers in the rotating frame attempt to use Einstein synchronization (that is, the principle of the constancy of the speed of light) to establish a network of synchronized clocks. Light travels along a null worldline so I may set $ds^2 = 0$ in Eq. (1.25). Also, it is sufficient for this discussion to keep only terms of first order in the small parameter $\omega_E r'/c$. Then

$$(cdt')^2 - \frac{2\omega_E r'^2 d\phi'(cdt')}{c} - (d\sigma')^2 = 0\,, \tag{1.26}$$

and solving for (cdt'),

$$cdt' = d\sigma' + \frac{\omega_E r'^2 d\phi'}{c}\,. \tag{1.27}$$

The quantity $r'^2 d\phi'/2$ is just the infinitesimal area dA'_z in the rotating coordinate system swept out by a vector from the rotation axis to the light pulse, and projected onto a plane parallel to the equatorial plane. Thus the total time required for light to traverse some path is

$$\int_{\rm path} dt' = \int_{\rm path} \frac{d\sigma'}{c} + \frac{2\omega_E}{c^2}\int_{\rm path} dA'_z. \qquad [\,\text{light}\,] \tag{1.28}$$

Observers fixed on the earth, who were unaware of earth rotation, would use just $\int d\sigma'/c$ for synchronizing their clock network. Observers at rest in the underlying inertial frame would say that this leads to significant path-dependent inconsistencies, which are proportional to the projected area encompassed by the path. Consider for example a synchronization process which follows earth's equator in the eastwards direction. For earth, $2\omega_E/c^2 = 1.6227 \times 10^{-21}$ s/m^2 and the equatorial radius is $a_1 = 6,378,137$ m, so the area is $\pi a_1^2 = 1.27802 \times 10^{14}$ m^2 . Thus the last term in Eq. (1.28) is

$$\frac{2\omega_E}{c^2}\int_{\rm path} dA'_z = 207.4 \text{ ns}. \tag{1.29}$$

From the underlying inertial frame, this can be regarded as the additional travel time required by light to catch up to the moving reference point. Simple-minded use of Einstein synchronization in the rotating frame gives only $\int d\sigma'/c$, and thus leads to a significant error. Traversing the equator once eastward, the last clock in the synchronization path would lag the first clock by 207.4 ns.

Traversing the equator once westward, the last clock in the synchronization path would lead the first clock by 207.4 ns.

In an inertial frame a portable clock can be used to disseminate time. The clock must be moved so slowly that changes in the moving clock's rate due to time dilation, relative to a reference clock at rest on earth's surface, are extremely small. On the other hand, observers in a rotating frame who attempt this find that the proper time elapsed on the portable clock is affected by earth's rotation rate. Factoring $(dt')^2$ out of the right side of Eq. (1.25), the proper time increment $d\tau$ on the moving clock is given by

$$(d\tau)^2 = (ds/c)^2 = dt'^2 \left[1 - \left(\frac{\omega_E r'}{c} \right)^2 - \frac{2\omega_E r'^2 d\phi'}{c^2 dt'} - \left(\frac{d\sigma'}{cdt'} \right)^2 \right] . \tag{1.30}$$

For a slowly moving clock $(d\sigma'/cdt')^2 << 1$ so the last term in brackets in Eq. (1.30) can be neglected. Also, keeping only first order terms in the small quantity $\omega_E r'/c$,

$$d\tau = dt' - \frac{\omega_E r'^2 d\phi'}{c^2} \tag{1.31}$$

which leads to

$$\int_{\text{path}} dt' = \int_{\text{path}} d\tau + \frac{2\omega_e}{c^2} \int_{\text{path}} dA'_z. \qquad [\,\text{portable clock}\,] \tag{1.32}$$

This should be compared with Eq. (1.28). Path-dependent discrepancies in the rotating frame are thus inescapable whether one uses light or portable clocks to disseminate time, while synchronization in the underlying inertial frame using either process is self-consistent.

Eqs. 1.28 and 1.32 can be reinterpreted as a means of realizing coordinate time $t' = t$ in the rotating frame, if after performing a synchronization process appropriate corrections of the form +$2\omega_E \int_{path} dA'_z/c^2$ are applied. It is remarkable how many different ways this can be viewed. The different ways discussed so far in this article include the fact that from the inertial frame it appears that the reference clock from which the synchronization process starts is moving, requiring light to traverse a different path than it appears to traverse in the rotating frame. The Sagnac effect can also be regarded as arising from the relativity of simultaneity in a Lorentz transformation to a sequence of local inertial frames co-moving with points on the rotating earth, or as the difference between proper times of a slowly moving portable clock and a Master reference clock fixed on earth's surface.

This was recognized in the early 1980s by the Consultative Committee for the Definition of the Second and the International Radio Consultative Committee who formally adopted procedures incorporating such corrections for the comparison of time standards located far apart on earth's surface. For the GPS it means that synchronization of the entire system of ground-based and orbiting atomic clocks is performed in the local inertial frame, or ECI coordinate system.

7. Sagnac-like effects due to rotation of the ECEF frame

By design, the ephemerides (positions) of the GPS satellites are broadcast in such a way that the receiver can compute their positions at the instant of transmission *in the rotating WGS-84 reference frame*. For time-tagging at the receiver, the propagation delays from different satellites can vary from about 67 ms to 86 ms. During this approximately 19 ms transmission time variation, the ECEF reference frame can rotate more than a microradian and the positions of the satellites due to this rotation alone can vary by over 30 meters while the satellites move in inertial space by as much as 60 meters. If this is not carefully accounted for, unacceptable navigation errors can occur.

It would lead to serious error to assert Eqs. 1.22 were valid in the ECEF frame. What the receiver must do is rotate the positions of each of the satellites, that have been computed in the rotating frame, into some chosen ECI frame. Then Eqs. 1.22 are valid and can be solved in the ECI frame. The resulting position found in the ECI frame is finally rotated into the WGS-84 frame and used for navigation.

To illustrate that these rotations give rise to Sagnac-like effects, suppose the chosen ECI frame instantaneously coincides with the WGS-84 frame at the instant of arrival of the earliest of the four signals. I denote the GPS time of arrival of this particular signal by t_1, and the position of this particular satellite at this time as $\mathbf{r}_1$. Let the time intervals between the arrival of this signal and the other three signals be denoted by

$$\Delta t_i = t_i - t_1\,, \quad i = 1, 2, 3, 4 \tag{1.33}$$

where for simplicity I have taken $\Delta t_1 = 0$. During the time interval Δt_i the ECEF frame has rotated the amount $\omega_E \Delta t_i$. An active rotation of the satellite position $\mathbf{r}_i(ECEF)$ by the amount $+\omega_E \Delta t_i$ is necessary in order to express the position of satellite i in the inertial frame in which the position $\mathbf{r}_1$ is expressed. This rotation operation can be expressed as

$$\mathbf{r}_i(ECI) = \mathbf{r}_i(ECEF) + \boldsymbol{\omega}_E \times \mathbf{r}_i(ECEF)\Delta t_i\,. \tag{1.34}$$

The navigation equations then become

$$c^2(t - t_i)^2 = |\mathbf{r} - \mathbf{r}_i(ECEF) - \boldsymbol{\omega}_E \times \mathbf{r}_i(ECEF)\Delta t_i|^2 \tag{1.35}$$

and if I put $\Delta t = t - t_1$ (no subscript on t) and $\mathbf{R}_i = \mathbf{r} - \mathbf{r}_i(ECEF)$ I obtain

$$c^2(\Delta t - \Delta t_i)^2 = |\mathbf{R}_i - \boldsymbol{\omega}_E \times \mathbf{r}_i(ECEF)\Delta t_i|^2 \tag{1.36}$$

Eqs. 1.36 have within them the four unknowns $(\Delta t, \mathbf{r})$. The position solution for $\mathbf{r}$ will be in the ECI frame chosen for computation. After finding this position, the result must then be rotated into the ECEF frame for navigation. Since the ECEF frame rotates an amount $\omega_E \Delta t$ during the time interval Δt, the final solution for the position in the ECEF frame will be

$$\mathbf{r}(ECEF) = \mathbf{r} - \omega_E \times \mathbf{r}\Delta t\,. \tag{1.37}$$

The size of the correction term in this last equation can easily be estimated, since $\Delta t \approx .015$ s and $r \approx 6.4 \times 10^6$ m. A typical value will be about 9 meters. Eq. 1.36 can be solved approximately for Δt by expanding the square on the right side, keeping only linear terms in ω_E, and then taking a square root, similar to the approximations made in deriving Eq. 1.17. The result is

$$\Delta t = \Delta t_i + \frac{R_i}{c} + \frac{\omega_E \times \mathbf{r}_i(ECEF) \cdot \mathbf{R}_i}{cR_i}\Delta t_i\,. \tag{1.38}$$

The last term in the above equation is a Sagnac-like correction. I can estimate its magnitude by substituting in an approximate expression for Δt_i:

$$\Delta t_i \approx \frac{R_i}{c} - \frac{R_1}{c} \tag{1.39}$$

So the correction term becomes, after interchanging dot and cross products,

$$\frac{\omega_E \cdot \mathbf{r}_i(ECEF) \times \mathbf{R}_i}{c^2}\left(1 - R_1/R_i\right)\,. \tag{1.40}$$

Is this really a Sagnac correction? It is linear in the rotational velocity, the coefficient can be interpreted in terms of an area, and it is relativistic (there is a factor $1/c^2$).

In the case of time-tagging at the transmitters, signals are chosen for processing which leave the transmitters at some chosen time t_T. Then the broadcast ephemerides will all be calculated by the receiver in one and the same ECEF frame. It would then be natural to choose for application of the navigation equations (Eqs. 1.22) an inertial frame which coincides with this ECEF frame at the instant t_T of GPS time. But then the signals do not arrive simultaneously at the receiver, and the receiver motion during the interval between arrival of the first and last signals must be accounted for.

To illustrate the size of the Sagnac-like effects that occur in this situation, let $\mathbf{r}$ denote the receiver position at transmission time t_T, and let $\mathbf{r}_i$ denote the transmitter position at time t_T. Imagine these positions to be expressed in an inertial frame which coincides instantaneously with the ECEF frame at time t_T. Let t_i denote the arrival time at the receiver, of the signal from the ith satellite. The receiver position at time t_i will be modified by earth rotation and will be

$$\mathbf{r} + \omega_E \times \mathbf{r}(t_i - t_T)\,. \tag{1.41}$$

The navigation equations in this inertial reference frame will be

$$c^2(t_i - t_T)^2 = |\mathbf{r} + \omega_E \times \mathbf{r}(t_i - t_T) - \mathbf{r}_i|^2 \tag{1.42}$$

Because of the similarity of this equation to Eq. 1.13 it is clear that Sagnac-like corrections will enter solution of the equations. The times t_i are however known only to within an added constant, because of a possible error or systematic bias in the receiver's clock. If the arrival times actually measured in the receiver are t_i', then

$$t_i = t_i' + b\,. \tag{1.43}$$

where b is the receiver clock bias then the navigation equations become

$$c^2(t_i' + b - t_T)^2 = |\mathbf{r} + \omega_E \times \mathbf{r}(t_i' + b - t_T) - \mathbf{r}_i|^2 \tag{1.44}$$

and the unknowns are $(b, \mathbf{r})$. Obviously there are many other ways of formulating the problem of accounting for receiver motion. A technical note[8] discusses these issues in more detail, with numerical examples.

8. Summary

In the GPS, the Sagnac effect arises because the primary reference frame of interest for navigation is the rotating Earth-Centered, Earth-Fixed frame, whereas the speed of light is constant in a locally inertial frame, the Earth-Centered Inertial frame. Additional Sagnac-like effects arise because the satellite ephemerides are broadcast in a form allowing the receiver to compute satellite positions in the ECEF frame. In the case of time-tagging of observations at the receiver, it is necessary to rotate the satellite positions into a common ECI reference frame in order apply the principle of the constancy of c. In the rotating frame of reference the effect appears to arise from a Coriolis-like term in the fundamental scalar invariant. Whether synchronization procedures are performed by using electromagnetic signals or slowly moving portable clocks, to leading order the same Sagnac effect arises. The effect is of significant magnitude and must be taken into account for accurate navigation. It is also necessary to apply Sagnac corrections when comparing remote clocks on earth's surface.

References

[1] *World Geodetic System 1984*, (National Imagery and Mapping Agency, 1994) Report No. 8350.2, Third Edition, Amendment 1. NIMA Stock No. DMATR83502WGS84, MSN 7643-01-402-0347.

[2] Malys, S., and Slater, "Maintenance and Enhancement of the World Geodetic System 1984", in *Proceedings of the 7th International Technical Meeting of the Satellite Division of the Institute of Navigation, Sept.*

20-23, Salt Lake City, UT (ION GPS-94), 17-24. The Institute of Navigation (1994).

[3] Landau, L., and E. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon, New York p. 254 (1997).

[4] Ashby, N., and Allan, D.W., "Practical Implications of Relativity for a Global Coordinate Time Scale", *Radio Science* **14**, 649-669 (1979).

[5] Wang, R., Y. Zheng, A. Yao, and D. Langley, "Modified Sagnac experiment for measuring travel-time difference between counter-propagating light beams in a uniformly moving fiber", *Phys. Letts*. **A 312**, pp 7-10, (2003).

[6] Hasselbach, F., and M. Nicklaus, *Phys. Rev.* **A48**, pp 143-151, (1993).

[7] Allan, D. W., M. Weiss, and N. Ashby, "Around-the-World Relativistic Sagnac Experiment", *Science* **228**, pp 69-70 (5 April 1985).

[8] Ashby, N., and M. Weiss, "Global Positioning Receivers and Relativity", NIST Technical Note 1385, U. S. Government Printing Office, Washington, D.C., March (1999).