Om globale navigasjonssatellittsystemer og relativitet

Jon Glenn Gjevestad og Øyvind Grøn

Vitenskapelig bedømt (refereed) artikkel

Jon Glenn Gjevestad and Øyvind Grøn: On the Global Navigation Satellite Systems and Relativity

KART OG PLAN, Vol. 73, pp. 32–39, POB 5003, NO-1432 Ås, ISSN 0047-3278

GNSS, or more precisely GNSS-2, is an abbreviation for Global Navigation Satellite Systems – Second generation, and serves as a generic name for the class of modern global satellite based radio navigation systems. GNSS-2 consists mainly of the four major Global Navigation Satellite Systems known as: GPS (U.S.), GLONASS (Russia), Galileo (EU) and Bei-Dou-2 (China). All these global radio navigation systems are based on the same navigation principle, i.e. utilizing ultra-stable clocks in satellites to determine the user position by independent measurements of the transit time of electromagnetic signals transmitted from satellites in orbit, so-called Radio Navigation Satellite Services (RNSS).

The typical performance of these global radio navigation systems is to provide absolute positioning to an observer on the surface of the Earth within the precision of 5-10 meter. However, this precision can be improved utilizing state of the art processing techniques such as Precise Point Positioning (PPP), currently demonstrating absolute positioning of 5-10 centimeters utilizing only one receiver. To achieve this astonishing precision in terms of absolute position, the rate of time as measured on the clock in the satellite must be known to better than a few nanoseconds. Since the satellites are constantly moving with respect to the observer and are also located at highly different gravitational potentials, effects predicted by both the Special- and General theories of Relativity must be considered in order to achieve the desired accuracy in the observed transit times.

These systems are in fact one of the very few man made systems, outside of particle accelerators, that experience significant relativistic effects.

Key words: GNSS, General Relativity

Jon Glenn Gjevestad, Professor, Department of Mathematical Sciences and Technology, Norwegian University of Life Sciences, P O Box 5003, NO-1432 Ås, Norway. E-Mail: jon.glenn.gjevestad@umb.no

Øyvind Grøn: Professor, Faculty of Technology Art and Design, Oslo and Akershus University College of Applied Sciences. Email: Oyvind.Gron@hioa.no

1. Introduction

Within the class of GNSS-2 navigation systems GPS is, through its 43 year history, the first operational and most widely used navigation system and will throughout this article serve as an example to illustrate and quantify some of the most significant relativistic effects experienced in a modern satellite based radio navigation system.

GPS is normally divided into three principal segments: Space segment, Control Segment and User Segment. The Space Segment consists of a nominal constellation of 27 satellites carrying atomic clocks. The nominal constellation was increased from 24 satellites after a reposition procedure known as "Expandable 24" that was completed in June 2011. As a result of this repositioning procedure the GPS constellation has attained a more optimal geometry, maximizing the worldwide coverage. The Control Segment consists of a number of ground based monitoring stations evenly distributed around the Earth close to the Equator. These monitor stations continuously monitor the satellites, distributing this information to the Master Control Station located in Colorado Springs, USA. Here the satellite constellation is analyzed and the satellite ephemerides and satellite clock behavior are predicted for several hours ahead. This information is then uploaded to the satellites for transmission back to the users. The User Segment consists of all users who utilize the transmitted signal from the satellites to determine their position, velocity and time using their local Quartz oscillator.

The GPS clocks are mainly Cesium atomic clocks operating by counting hyperfine transitions of Cesium atoms at a very stable frequency. The precise number of such transitions of a Cesium atom is now being adopted by international agreement as the definition of one atomic second (6). It is the orbiting Cesium clocks that serve as the backbone of the precision demonstrated by GPS.

2. Pseudorange

The principle observable of the GPS is the pseudorange which is basically the distance between a satellite and the receiver on Earth. In other words it is the difference between the reception time, according to the local clock, and the transmission time, according to GPS time, multiplied by the speed of light.

The signals from the satellites can be thought of as continuous timing signals arriving at the receiver from the satellite clocks. Measuring the one-way speed of the satellite signal requires two clocks – one at each end of the path (2).

Measuring this time difference requires synchronization of these clocks, which will be discussed in section 5.

3. Relativistic time effects

According to the general physical interpretation of a timelike spacetime interval in the general theory of relativity, the proper time interval, $d\tau$, measured on a clock moving with a velocity having components $v^i = dx^i / dt$ (we use Latin letters for spatial indices and Greek letters for spacetime indices) in a coordinate system $\left(x^0, x^1, x^2, x^3\right)$, where $x^0 = ct$ and t is the coordinate time, is [1]

$$d\tau = \left(-\frac{g_{\mu\nu} dx^\mu dx^\nu}{c^2} \right)^{1/2} \quad , \quad \mu, \nu = 0,1,2,3 \, , \tag{1}$$

where $g_{\mu\nu}$ are the components of the metric tensor. This may be written

$$d\tau = \left(-g_{00} - 2g_{i0}\frac{v^i}{c} - \frac{v^2}{c^2} \right)^{1/2} dt \quad , \quad i = 1,2,3 \; , \tag{2}$$

where $v = \left(g_{ij} v^i v^j\right)^{1/2}$ is the coordinate velocity of the clock.
Due to the rotation of the Earth there is Kerr spacetime outside the Earth. Then

$$\begin{gathered} g_{00} = -\frac{1 - \frac{R_S}{r} + \frac{a^2}{r^2}\cos^2\theta}{1 + \frac{a^2}{r^2}\cos^2\theta} \; , \quad g_{0i} = \frac{\frac{R_S}{r}\frac{a}{r} r \sin^2\theta}{1 + \frac{a^2}{r^2}\cos^2\theta} \; , \quad g_{rr} = \frac{1 + \frac{a^2}{r^2}\cos^2\theta}{1 - \frac{R_S}{r} + \frac{a^2}{r^2}} \; , \\ g_{\theta\theta} = \frac{\left(1 + \frac{a}{r^2}^2\right)^2 - \left(1 - \frac{R_S}{r} + \frac{a^2}{r^2}\right)\frac{a^2}{r^2}\sin^2\theta}{1 + \frac{a^2}{r^2}\cos^2\theta} r^2 \sin^2\theta \; , \quad g_{\phi\phi} = \left(1 + \frac{a^2}{r^2}\cos^2\theta\right) r^2 \; . \end{gathered} \tag{3}$$

Here $R_S = 2GM / c^2 = 0.01m$ is the Schwarzschild radius of the Earth, and a is the length corresponding to the angular momentum per unit mass of the Earth, $a = J / Mc$, where J is the angular momentum of the Earth. Inserting the angular velocity, mass and radius of the Earth gives $a = 2.0m$. The radius of the Earth is $r_\oplus = 6.4 \cdot 10^6 m$. Hence, at the surface of the Earth $R_S / r_\oplus = 1.6 \cdot 10^{-9}$ and $a / r_\oplus = 3.2 \cdot 10^{-7}$. Let us compare the result of using the Kerr metric versus the simpler, but not so accurate Schwarzschild metric which has $a = 0$ when we calculate the relativistic time effects for the GPS-system.

In order to find the magnitude of the relativistic effects it is sufficient to consider a satellite moving along a circular path with radius r_s in the equatorial plane. Then $\theta = \pi / 2$ and $v^r = v^\theta = 0$ and eq.(2) reduces to

$$d\tau = \left(1 - \frac{R_S}{r_s} - 2\frac{R_S}{r_s}\frac{a}{r_s}\frac{v^\phi}{c} - \left(\frac{v^\phi}{c}\right)^2\right)^{1/2} dt \ , \quad v^\phi = r\frac{d\phi}{dt}. \tag{4}$$

The coordinate clocks of the Kerr metric are all synchronized to show the same time as a standard clock at rest infinitely far from the mass distribution. Hence the coordinate clocks proceed at the same rate independent of their position. The term R_S / r_s represents the gravitational time dilation. The larger r_s is, i.e. the higher up the satellite is, the faster will the satellite clock tick. The next term which is proportional to the product of the angular momentum of the mass distribution and the velocity of the satellite, may be called the Kerr-term. The last term is the usual velocity dependent time dilation which comes from the special theory of relativity.

A typical velocity of the GPS-satellites is $v^\phi = 4 \cdot 10^3 m / s$. Hence the magnitudes of the terms inside the parenthesis are

$$\frac{R_S}{r_s} \approx 5 \cdot 10^{-10} \ , \quad 2\frac{R_S}{r_s}\frac{a}{r_s}\frac{v^\phi}{c} \approx 10^{-21} \ , \quad \left(\frac{v^\phi}{c}\right)^2 = 1.7 \cdot 10^{-10}. \tag{5}$$

This shows that the Kerr-term is much smaller than the other terms which give rather small relativistic corrections. Hence the Kerr-term may be neglected. Also we see that the gravitational term and the kinematical term are of the same order of magnitude which is to be expected due to the virial theorem.

If the relativistic effects are neglected, the satellite clocks would advance at the same rate as the clocks on the Earth. Hence, neglecting the Kerr-term the relativistic time-effects on the GPS-satellite clocks are given by

$$\Delta\tau = d\tau(r_s, v) - d\tau(r_s + h, 0)\,, \tag{6}$$

where h is the height of the satellite clocks above the surface of the Earth. Neglecting the Kerr-term and calculating to 1. order in R_S / r and $(v^\phi / c)^2$ leads to

$$\Delta\tau = \Delta\tau_G + \Delta\tau_V \approx \frac{1}{2}\frac{R_S h}{r_\oplus (r_\oplus + h)}\Delta t - \frac{1}{2}\left(\frac{v^\phi}{c}\right)^2 \Delta t. \tag{7}$$

Here $\Delta\tau_G$ is the gravitational effect and $\Delta\tau_V$ the kinematical effect. Inserting numerical values give for a coordinate time interval $\Delta t = 24h \approx 8{,}6 \cdot 10^4 s$ that $\Delta\tau_G = 5.2 \cdot 10^{-5} s$ and $\Delta\tau_V \approx 7 \cdot 10^{-6} s$. Hence due to the gravitational time dilation a satellite clock goes faster than the Earth clock with about 52 microseconds in 24 hours, while it slows down with 7 microseconds due to the kinematical time dilation.

Ashby [1] wrote "Timing errors of one ns will lead to positioning errors of the order of 30 cm." This comes about by multiplying the time difference by the velocity of light. Then an error of $5.2 \cdot 10^{-5} s$ per day due to neglecting the gravitational time effect would correspond to a position error of 16 km per day. It should be noted that this refers to an error of the distance between the satellite emitting a signal and the receiver at the surface of the Earth, i.e. it refers to the error in the pseudorange corresponding to a timing error in the satellite clock.

However the corresponding error in the determination of the position of an object on the surface of the Earth is smaller. In the case of a signal from one satellite it is the projection of $c\Delta\tau$ along the surface of the Earth. If, for example, the signal moves vertically, then there would be a positioning error on the Earth due to a time error on the satellite only because this time error leads to a wrong determination of the position of the satellite at a given satellite time. In this case the positioning error on the Earth is given by multiplying the time error of the satellite clock by the velocity of the satellite, not by the velocity of light. Then the time difference $5.2\cdot 10^{-5}s$ corresponds to a position difference of 20 cm per day. Such a drift would amount to an error of 6 meters in a month, which is still not negligible.

4. Sagnac Effect

In the experiment of Sagnac two light signals were forced by means of mirrors to move along a closed path in opposite directions, and made to interfere at the arrival. The equipment was positioned on a rotating table and Sagnac measured how the interference patterns changed with the angular velocity of the table. This change is an expression of the difference in travelling time for the two light signals. From the point of view of the non-rotating laboratory frame this travel time difference is readily explained as a result of the tangential movement of the equipment during the travel of the light signals.

This proved the absolute character of rotational velocity for Sagnac who performed the experiment in 1913 before Einstein had constructed the general theory of relativity. However according to Einstein's theory an observer on the table may consider herself as at rest and the environment and the rest of the Universe as rotating [3]. Let us see how the dependence of the travelling time of the light upon the angular velocity of the environment is explained according to the general theory of relativity.

The velocity dependent time dilation implies that a moving clock goes at a slower rate the faster it moves through space, and in the limit of the velocity of light the clock does not proceed at all. This means that the proper time interval as measured by a clock that moves with the velocity of light vanishes. Therefore the propertime intervals vanish along the world line of a light signal, meaning that in this case $d\tau = 0$ in eq.(1). For a light signal that moves along a circular path in the equatorial plane of the Kerr spacetime this means that eq.(1) reduces to

$$g_{tt}dt^2 + 2g_{t\phi}dtd\phi + g_{\phi\phi}d\phi^2 = 0, \tag{8}$$

or

$$\frac{g_{\phi\phi}}{r^2}\left(v^\phi\right)^2 + 2\frac{g_{t\phi}}{r}v^\phi + g_{tt} = 0. \tag{9}$$

Hence the angular velocity of the light signals are

$$v^\phi = \frac{-g_{t\phi} \pm \sqrt{g_{t\phi}^2 - g_{tt}g_{\phi\phi}}}{g_{\phi\phi}}r. \tag{10}$$

This formula shows a property of light which may come as a surprise for those familiar with the special theory of relativity. This theory is based upon two postulates: 1. The principle of relativity for accelerated motion, and 2. The postulate that the velocity of light is isotropic and independent of the velocity of the emitter. However, the formula (10) shows that with respect to a reference frame with a metric where $g_{t\phi} \neq 0$ the velocity of light is anisotropic.

Consider two light signals moving around a circle in opposite directions. Using eq.(10) we find that the difference of their travel times is

$$\Delta t = 2\pi r\left(\frac{1}{v_+^\phi} + \frac{1}{v_-^\phi}\right) = 4\pi \frac{g_{t\phi}}{|g_{tt}|}. \tag{11}$$

It is natural to call this travel time difference for the *Sagnac effect*.
In order to see as clearly as possible the meaning of the last two formulae we shall first consider the original Sagnac experiment. Then there is an inertial laboratory frame with cylindrical coordinates (T,R,Θ,Z). With these coordinates the line-element of the flat Minkowski spacetime takes the form

$$ds^2 = -c^2dT^2 + dR^2 + R^2 d\Theta^2 + dZ^2. \tag{12}$$

The table rotates with an angular velocity ω. The coordinates co-moving with the rotating table are (t,r,θ,z), and the coordinate transformation is

$$t = T \ , \quad r = R \ , \quad \theta = \Theta - \omega T \ , \quad z = Z. \tag{13}$$

In this coordinate system the line-element has the form

$$ds^2 = -\left(1 - \frac{r^2\omega^2}{c^2}\right)c^2dt^2 + 2r^2\omega dt d\theta + dr^2 + r^2 d\theta^2 + dz^2. \tag{14}$$

Hence the non-vanishing components of the metric tensor are

$$g_{tt} = -\left(c^2 - r^2\omega^2\right) \ , \quad g_{t\theta} = r^2\omega, \quad g_{rr} = 1 \ , \quad g_{\theta\theta} = r^2 \ , \quad g_{zz} = 1. \tag{15}$$

Inserting these into eq.(10) gives

$$v_+^\theta = -r\omega + c \ , \quad v_-^\theta = -r\omega - c. \tag{16}$$

This shows that in the rotating reference frame the velocity of light is greater in the same direction as that of the rotating environments than in the opposite direction. Hence there is a travel time difference as given by eqs. (11) and (15),

$$\Delta t = \gamma^2 \frac{4A\omega}{c^2} \ , \quad \gamma = \left(1 - \frac{r^2\omega^2}{c^2}\right)^{-1/2}, \tag{17}$$

where $A = \pi r^2$ is the area enclosed by the path. The difference in the velocity of light travelling in opposite directions is the general relativistic explanation of the change of the interference pattern with reference to the rest frame of the apparatus. In the general theory of relativity this difference vanishes in a frame in which the environment does not rotate.
The experiment was repeated by Michelson and Gale in 1925 [4] with light path enclosing an area about $2 \cdot 10^5 m$ and with the Earth as the rotating table with an angular velocity $\omega_\oplus = 3 \cdot 10^{-7} rad/s$. In this experiment the light travel time difference is $\Delta t \approx 3 \cdot 10^{-18} s$ which was measured from an analysis of the two light beams with an uncertainty of only 2%.
In the general theory of relativity there is an inertial dragging effect inside a rotating mass distribution. A rough calculation [4] shows that there may be *perfect dragging* in our Universe, meaning that the inertial frames are dragged on together with the cosmic masses. Hence in an inertial frame the cosmic masses have no angular velocity. In this frame the velocity of light is isotropic and the time difference (11) vanishes.

A similar experiment with light moving around the Earth in opposite directions along a circular path from a satellite with no angular velocity relative to the stars, would give no travel time difference according to Newton's theory. According to the general theory of relativity there is Kerr spacetime outside the Earth. When this is described with Boyer-Lindquist coordinates the metric components are given in eq.(3). This is a rigid coordinate system where the reference particles have no angular velocity relative to the stars.

Consider satellites at rest in the Boyer-Lindquist coordinate system. The Sagnac experiment is now performed with light emitted in opposite direction from such a satellite and then made to interfere when the light beams meet at the satellite again after a trip around the Earth. In this case then the time travel difference, Δt_{KS}, may be called the Kerr-Sagnac effect, and is found by inserting the metric components (3) into eq.(11), which gives

$$\Delta t_{KS} = \frac{4\pi R_S a}{c\left(r_s - R_S\right)} \approx \frac{4\pi R_S a}{c r_s} = 4.2 \cdot 10^{-17} s. \tag{18}$$

If the Sagnac experiment is performed with a GPS-satellite there is an additional time difference due to its motion [1], given by eq.(17),

$$\Delta t_v \approx \frac{4 r_S v^\phi}{c^2} = 414\,\mathrm{ns}\,, \tag{19}$$

which is much larger than the Kerr-Sagnac effect.

5. On the synchronization of the GPS clocks

There is a difficulty in synchronizing the satellite clocks due to the relativity of simultaneity. The satellites are at rest in a rotating frame of reference. As shown in section 3, in a non-rotating frame with center coinciding with the center of the Earth, there is with sufficiently good accuracy the Schwarzschild metric. The coordinate clocks of the Schwarzschild metric are all synchronized to show the same time at a standard clock at rest infinitely far from the mass distribution.

In order to discuss the synchronization of the satellite clocks, we assume that the satellites are moving with a velocity v^ϕ in the equatorial plane of the Earth along circular paths with radius r_s. Consider two satellites A and B with an angular distance $\Delta\phi$ between them, and with A in front. There is a distance $\Delta l = r_s \Delta\phi$ between the satellites. Assume that one tries to Einstein synchronize the satellite clocks around the path for example by emitting light signals backwards and forwards from a point midway between two clocks. Einstein synchronization means that the clocks are set to show the same time when these signals hit them. However, as observed from the non-rotating frame of the Schwarzschild clocks, the signal moving backwards hits the satellite clock B before the forward moving signal hits the clock A. Hence the satellite clock B is ahead of the clock A when compared with the Einstein synchronized coordinate clocks of the non-rotating reference frame.

If this local Einstein synchronization process is performed around the circle it corresponds to emit light signals in opposite directions from an emitter at the opposite sides of the circle relative to the position A of the clock that shall be synchronized. The light is then moving a half circle in opposite directions to the clock A. This gives a travel time difference equal to one half of that in the Sagnac experiment. Hence the difference of time of the locally Einstein synchronized satellite clocks is half of that given by eq. (17). Thus for the satellite clock there is a time discontinuity at A of locally Einstein synchronized clocks, equal to 207 ns. This means that as measured on the globally synchronized Schwarzschild coordinate clocks the satellite clock A is not synchronized with itself. In other words it is impossible to Einstein synchronize globally the satellite clocks which are at rest in a rotating reference frame.

Therefore the satellite clocks of the GPS-system are not locally Einstein synchronized in their rest frame. Instead all the clocks in the GPS-system are synchronized in an Earth-centered inertial (ECI) reference frame [1]. This frame is not rotating with the Earth, neither with the satellites. The Kerr- or Schwarzschild coordinate clocks are globally Einstein synchronized in the ECI-frame. So the GPS-clocks are not showing proper time. If they had done that, the clocks in the satellites would proceed at a faster rate than the clocks at the surface of the Earth. Instead the rate of the satellite clocks is adjusted to account for the increased rate of time due to their height and the decreased rate due to their velocity in the ECI-frame. In this way one obtains the result that all of the clocks ticks at the same rate, like the coordinate clocks of the Schwarzschild metric.

6. Factory Offset

GPS satellites are corrected for the relativistic effects due to circular orbit and constant Earth potential prior to launch in what is merely known as the "factory offset". The line-element of the Schwarzschild metric is given as:

$$ds^2 = -\left(1+\frac{2\Phi}{c^2}\right)c^2dt^2 + (1+\frac{2\Phi}{c^2})^{-1}dr^2 + r^2d\theta^2 + r^2\sin^2\theta d\theta. \tag{20}$$

where r, θ, ϕ is the coordinates of the satellite and $\Phi = -GM/r$ the gravitational potential.

First assume a non-rotating reference system with a circular satellite orbit lying in the equatorial plane (i.e. $dr = d\theta = 0$, $\sin(\theta) = 1$). Then the Schwarzschild line-element reduces to:

$$ds^2 = -\left(1+\frac{2\Phi}{c^2}\right)c^2dt^2 + r^2d\phi^2. \tag{21}$$

Since the acceleration is equal to the gradient of the potential we have:

$$\frac{v^2}{r} = \frac{GM}{r} \quad \Rightarrow \quad v^2 = -\Phi. \tag{22}$$

Setting $ds^2 = -c^2d\tau^2$ for the Schwarzschild line-element we can write:

$$-c^2d\tau^2 = -(1+\frac{2\Phi}{c^2})c^2dt - \Phi dt^2 = -(1+\frac{3\Phi}{c^2})c^2dt. \tag{23}$$

By taking the square root we obtain the following relationship:

$$d\tau = (1+\frac{3\Phi}{2c^2})dt. \tag{24}$$

The remaining step is to convert from this time into international atomic time (TAI) realized by atomic clocks at rest on the geoid. Thus we have:

$$d\tau = (1+\frac{\Phi_0}{c^2})dt \quad \Rightarrow \quad dt = (1-\frac{\Phi_0}{c^2})dt', \tag{25}$$

where Φ_0 is the combined gravitational- and rotational potential on the geoid. Inserting (25) into (24) we obtain the final result:

$$d\tau = (1+\frac{3\Phi}{2c^2}-\frac{\Phi_0}{c^2})dt'. \tag{26}$$

It is the two last terms in this equation that contains the relativistic correction which is subtracted from the nominal frequency of the satellite clocks prior to launch.
At the GPS orbital radius of $r = 26562km$, the fractional clock rate offset is as follows:

$$\frac{3\Phi}{2c^2}-\frac{\Phi_0}{c^2} = -2.5045336\cdot 10^{-10} + 6.9692842\cdot 10^{-10} = 4.46475\cdot 10^{-10}. \tag{27}$$

In order for the satellite clock to beat at the frequency of $10.23MHz$ as seen from an observer on Earth, the fundamental frequency needs to be shifted high with $10.23MHz \cdot 4.46475 \cdot 10^{-10} = 0.004567Hz$, and the GPS satellite clocks are correspondingly lowered in frequency to $10.2299999954MHz$ prior to launch in order to compensate for this effect (3).

7. Conclusion

Modern satellite based navigation systems will experience significant relativistic effects, which can be predicted by involving both the Special- and General Theory of Relativity. The reasons for these relativistic effects are mainly due to the large velocity of the GNSS satellites, the difference in gravitational potential between the satellites and the receiver and the effects of Earth rotation.

Due to the fact that GNSS utilizes precision atomic clocks to estimate the position of an object on the surface of the Earth at the cm-level, these relativistic effects will become significant and must be taken into account in order for the systems to perform as wanted.

References

N. Ashby, *Relativity in the Global Positioning System.* Living.Rev.Relativity. **6,** 1–42 (2003).

N. Ashby, NIST Technical Note 1385 (1999)

N. Ashby, *Introduction to Relativistic Effects on the Global Positioning System*, AIAA. (1995)

Ø. Grøn, *Lecture Notes on the General Theory of Relativity.* Lecture Notes in Physics 772. Springer (2009).

Ø. Grøn and S. Hervik, *Einstein's General Theory of Relativity.* Springer (2007).

BIPM, 13th General Conference on Weights and Measures (CGPM) 1967/68, Resolution 1; CR, 103 and Metrologia, 1968, **4**, 43