Chapter 18. Relativity enters technology - the Global Positioning System

This chapter is based on the articles by Neil Ashby [226 – 229].

The US-based Global Positioning System (GPS) is by today one of a few Global Navigation Satellite Systems (GNSS) that are working or are being developed and prepared for operation.

The other ones are the Russian GLONASS (Global Navigation Satellite System), the European Galileo and the Chinese Beidou.

The GPS was the first (launched in 1977), and by now is the longest-working and most widely used, therefore this overview will be focussed on it. It was originally intended for military use only, but is accessible to civilian users since 1989.

The relativistic effects are present in all GNSS.

[226] N. Ashby, Relativistic effects in the Global Positioning System. In: *Gravitation and relativity at the turn of the millenium. Proceedings of the 15th International Conference on General Relativity and Gravitation in Pune, India, 1997*. Edited by N. Dadhich and J. V. Narlikar. Inter-University Centre for Astronomy and Astrophysics, Pune 1998, p. 231.
[227] N. Ashby, *Relativity in the palm of your hand*. **Mercury** 25, 23 (1996).
[228] N. Ashby, Relativity in GNSS. Chapter 24 in *Springer handbook of spacetime*, Springer 2014, pp. 509 - 525.
[229] N. Ashby, GNSS and other applications of general relativity. In: *General Relativity: the most beautiful of theories*. Edited by C. Rovelli, de Gruyter, Studies in Mathematical Physics Vol. 28, 2015.
Consultation by Professor Ashby and the permission to use his figures are gratefully acknowledged. He kindly sent to me Fig. 18.3, while Figs. 18.4 and 18.5 were copied from Ref. [226].

18. 1. Purpose and setup

The GPS provides space and time coordinates of portable receivers on the Earth surface and above it.

It consists of a constellation of satellites that carry atomic clocks, several ground-based monitoring stations, and a Master Control Station in Colorado Springs.

Information between the satellites and the ground stations is transmitted by electromagnetic waves at two frequencies [226].

The orbital planes of the GPS satellites are inclined at 55° to the equatorial plane.
At this angle the perturbations of the satellite paths caused by the Earth's quadrupole moment have zero time-average.

The basic system includes 6 such orbits, regularly arranged with respect to the equator, and 4 satellites on each orbit.

In addition, there are several spare satellites and clocks.

This distribution of satellites in space ensures that, in principle, at least 4 satellites are almost always in the field of view from every point on Earth's surface.

Each satellite carries a cesium clock of accuracy 5×10^{-14} = 4 nanoseconds per 24 hours.

The accuracy in time of 1 ns implies the accuracy in space of 10 cm.

The relativistic effects are much larger than this, as will be seen from the following.

Accuracies of about 1 mm are technically feasible by today, but are accessible only for US military use.

Each satellite sends a stream of circularly polarised electromagnetic waves that carries information about its time-coordinate and position in space.

The signals are received by a network of monitoring stations on the ground and forwarded to the Master Control Station.

There, the state of the whole constellation is analysed and predicted for the next few hours.

Based on the predictions, the clocks in the satellites are periodically corrected.

The users carry locators that determine the user's time, position and velocity using the information transmitted in the signal.

Until 2000, civilians were shielded from the highest precision: the two carrier frequencies were intentionally fluctuated (high-precision positioning required access to undisturbed signals).

From 2000 on, this limitation does not exist, and new signals are being added to increase the precision even further. See Ref. [230] for the currently available menu of signals.

Without relativistic corrections, the GPS would be useless - see the end of this chapter.

This system could be used for experimental tests of relativity, but its precision is not better than that of dedicated experiments.

[230] https://www.gps.gov/systems/gps/modernization/civilsignals/

18.2. The principle of position determination

Let $\vec{r}_i$ be the positions of 4 clocks, i = 1, 2, 3, 4, given at times t_i.

Let $\vec{r}$ and t be the unknown position and time of the receiver; all the times and positions are referred to a local inertial frame.

Consider the signals sent from $\vec{r}_i$ at t_i that are simultaneously received at ($\vec{r}$, t).

Since the velocity of light c is constant, we have

$$c^2 (t - t_i)^2 = \left| \vec{r} - \vec{r}_i \right|^2 \qquad (18.1)$$

These are 4 equations, to be solved for the 4 unknowns ($\vec{r}$, t).

But, in consequence of high precision of the system, the reference frame with respect to which the times and positions are calculated must be carefully defined - see the next sections.

The timing signals are places in the wave trains where the phase of the electromagnetic wave changes sign. At those events, momentarily $F_{\mu\nu} = 0$, and this is independent of the reference system.

The following reference systems will be used in the calculations:

ECEF = Earth-centered, Earth-fixed

This system rotates with the Earth.

ECI = Earth-centered inertial

This system, like the previous one, has its origin at the centre of the Earth, but does not rotate.

The meter and the second of the SI are defined in ECEF.

ECEF = Earth-centered, Earth-fixed
ECI = Earth-centered inertial

$$c^2\left(t-t_i\right)^2=\left|\vec{r}-\vec{r_i}\right|^2 \qquad (18.1)$$

18.3. The reference frames and the Sagnac effect

Equation (18.1) applies in a local inertial frame, while navigators on the Earth use the noninertial ECEF frame. Transformations between them are needed in using the GPS.

Ignore for a moment the gravitational field of the Earth. In an inertial frame, in cylindrical coordinates, the Minkowski metric is

$$-ds^2 = -(c\,dt)^2 + dr^2 + r^2\varphi^2 + dz^2. \qquad (18.2)$$

The transformation equations to the ECEF frame with the coordinates (t', r', φ', z') are

$$(t, r, z) = (t', r', z'), \qquad \varphi = \varphi' + \omega_E t', \qquad (18.3)$$

where $\omega_E = 7.292115 \times 10^{-5}$ rad/s is the angular velocity of rotation of the Earth.

The Minkowski metric in the ECEF frame is thus

$$-ds^2 = -(1 - \omega_E^2 r'^2/c^2)(c\,dt')^2 + 2\omega_E r'^2 d\varphi'\,dt' + dr'^2 + r'^2 d\varphi'^2 + dz'^2. \qquad (18.4)$$

ECEF = Earth-centered, Earth-fixed
ECI = Earth-centered inertial

$$- ds^2 = - (1 - \omega_E^2 r'^2/c^2)(c\, dt')^2 + 2\omega_E r'^2 d\varphi' dt' + dr'^2 + r'^2 d\varphi'^2 + dz'^2. \quad (18.4)$$

So, in a rotating frame, the time coordinate t' is not the proper time of the comoving observer (which is $d\tau = \sqrt{1-(\omega_E r'/c)^2}\, dt'$), but the proper time of the ECI observer.

Suppose, a set of observers on the rotating Earth use the Einstein synchronisation method (by sending light rays between them and registering the emission/detection times).

Since $\omega_E r'/c \approx 1.5\times10^{-11}$, $(\omega_E r'/c)^2 = 2.25\times10^{-22} \ll \omega_E r'/c$ and can be neglected. The equation of the light signal is then

$$0 = - ds^2 = - (c\, dt')^2 + 2(\omega_E/c) r'^2 d\varphi' (c dt') + d\sigma'^2 \quad (18.5)$$

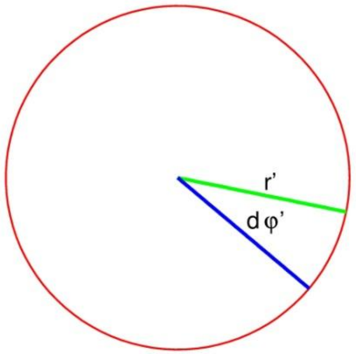
where

$$d\sigma'^2 = dr'^2 + r'^2 d\varphi'^2 + dz'^2. \quad (18.6)$$

The solution of (18.5) up to $\omega_E r'/c$ is

$$c\, dt' = d\sigma' + \omega_E r'^2 d\varphi'/c = d\sigma' + 2\omega_E dA'_z/c, \quad (18.7)$$

$c\,dt' = d\sigma' + \omega_E r'^2 d\varphi'/c = d\sigma' + 2\omega_E dA'_z/c$,   (18.7)



The surface area of the wedge is $dA'_z = (1/2)r'^2\, d\varphi$. This is the area swept by the radius-vector of the head of the light ray, projected on the equatorial plane.

→ Thus, the light pulse traverses a path $\mathcal{P}$ in the time

$$T = \int_{\mathcal{P}} \mathrm{d}t' = \int_{\mathcal{P}} \frac{1}{c}\, \mathrm{d}\sigma' + 2\frac{\omega_E}{c^2} \int_{\mathcal{P}} \mathrm{d}A'_z \qquad (18.8)$$

The first term applies in an inertial frame, the second one is the correction due to rotation.

Now suppose that we try to synchronise clocks on the equator by sending a light ray eastward, all around the Earth. Then the path $\mathcal{P}$ is the whole equator, and

$2\,\omega_E/c^2 = 1.6227 \times 10^{-21}$ s/m^2,

$r' = a_1 = 6\,378\,137$ m is the equatorial radius of the Earth,

$2\omega E \int_{\mathcal{P}} dA'_z/c^2 = 207.4$ ns is the additional time needed for the light ray to catch up with its point of origin.

$$2\omega E\int_P dA'_z/c^2 = 207.4\ \text{ns}$$

This quantity is the error in synchronisation caused by Earth's rotation; it is called the ***Sagnac effect*** [232 – 234].

*

Georges Sagnac was an opponent of special relativity, and he interpreted his effect as a proof of existence of the aether.

The correct explanation of this effect was first given by Paul Langevin [231].

*

For a ray sent around the equator in the westward direction, the error would have the same magnitude, but opposite sign - the ray would meet its point of origin by 207.4 ns earlier than if the Earth were not rotating.

[231] P. Langevin, *C. R. Acad. Sci. Paris* **173**, 831 (1921).
[232] G. Sagnac, L’éther lumineux d'emontré par l’éffet du vent relatif d'éther dans un interféromètre en rotation uniforme. *C. R. Acad. Sci. Paris* **157**, 708 (1913).
[233] G. Sagnac, Sur la preuve de la réalité de l’éther lumineux par l'expérience de l'interférographe tournant. *C. R. Acad. Sci. Paris* **157**, 1410 (1913).
Refs. [232,233] cited after:
[234] J. Frauendiener, Notes on the Sagnac effect in General Relativity. arXiv:1808.07914.

$- ds^2 = -(1 - \omega_E^2 r'^2/c^2)(c\,dt')^2 + 2\omega_E r'^2 d\varphi'\,dt' + dr^2 + r^2 d\varphi^2 + dz'^2.$ (18.4)

$$T = \int_{\mathcal{P}} \mathrm{d}t' = \int_{\mathcal{P}} \frac{1}{c}\,\mathrm{d}\sigma' + 2\frac{\omega_E}{c^2}\int_{\mathcal{P}} \mathrm{d}A'_z \qquad (18.8)$$

One might think that the error would be avoided by moving a portable reference clock along the equator, with negligible velocity.

This does not help. The proper time of the clock would be, from (18.4)

$$\mathrm{d}\tau^2 = (\mathrm{d}s/c)^2 = \mathrm{d}t'^2\left[1 - \left(\frac{\omega_E r'}{c}\right)^2 - 2\frac{\omega_E r'^2}{c^2}\frac{\mathrm{d}\varphi'}{\mathrm{d}t'} - \frac{1}{c^2}\left(\frac{\mathrm{d}\sigma'}{\mathrm{d}t'}\right)^2\right] \qquad (18.9)$$

Assuming slow motion, $(1/c)\,d\sigma'/dt \ll 1$, and $(\omega_E r'/c)^2 \ll \omega_E r'/c$ again, we obtain
$d\tau \approx dt' - \omega_E r'^2 d\varphi'/c^2$,
so for the travel time all around the equator eastward we again obtain the same formula (18.8), and the same synchronisation error.

The conclusion is:

In synchronising clocks all over the Earth, constant reference must be made to the synchronisation in the underlying inertial frame, ECI. This is the only self-consistent frame.

$$g_{00} = 1 + \frac{2\phi}{c^2} + O\left(\frac{v^3}{c^3}\right), \qquad g_{0I} = O\left(\frac{v^2}{c^2}\right), \qquad g_{IJ} = -\delta_{IJ} + O\left(\frac{v}{c}\right) \tag{10.37}$$

18.4. Earth's gravitation and the SI time units

We assume the following:

1. The mass distribution in the Earth is static,

2. The origin of the local inertial frame is at the Earth's center of mass.

Then the linearised solution of the Einstein equations that describes the Earth's gravitational field is, in agreement with (10.37):

$$-\mathrm{d}s^2 = -\left(1 + \frac{2V}{c^2}\right)(c\mathrm{d}t)^2 + \left(1 - \frac{2V}{c^2}\right)\left(\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2\right) \tag{18.10}$$

where

$$V = -\frac{GM}{r}\left[1 - J_2\left(\frac{a_1}{r}\right)^2 P_2(\cos\vartheta)\right] \tag{18.11}$$

is the Newtonian gravitational potential of the Earth (up to terms of order r^{-3}), M is the mass of the Earth, a_1 is, as before, the equatorial radius of the Earth, and (t, r, z) = (t', r', z'), $\varphi = \varphi' + \omega_E t'$, (18.3) $$-\mathrm{d}s^2 = -\left(1+\frac{2V}{c^2}\right)(c\mathrm{d}t)^2 + \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2\right) \quad (18.10)$$

$$V = -\frac{GM}{r}\left[1 - J_2\left(\frac{a_1}{r}\right)^2 P_2(\cos\vartheta)\right] \quad (18.11)$$

$GM = 3.98600418 \times 10^{14}$ (m^3/s^2),

$J_2 = 1.0826300 \times 10^{-3}$ is the quadrupole moment of the Earth,

$P_2 (\cos \vartheta) = (1/2) (3 \cos^2\vartheta - 1)$ is the second Legendre polynomial; ϑ in this formula is measured down from the rotation axis.

We transform the metric (18.10) to the rotating ECEF frame by eq. (18.3), and neglect terms $(\omega r'/c)^n$ where $n > 2$. The result is

$$\begin{aligned} -\mathrm{d}s^2 &= -\left(1+\frac{2\Phi}{c^2}\right)(c\mathrm{d}t')^2 + 2\frac{\omega_E}{c}\, r'^2\sin^2\vartheta' c\mathrm{d}t'\mathrm{d}\varphi' \\ &+ \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r'^2 + r'^2\mathrm{d}\vartheta'^2 + r'^2\sin^2\vartheta'\mathrm{d}\varphi'^2\right) \end{aligned} \quad (18.12)$$

where

$$\Phi = V - (1/2)(\omega_E r' \sin\vartheta')^2 \quad (18.13)$$

is the ***effective potential*** - the sum of the gravitational and centrifugal potentials. In (18.12), t' is the proper time of observers at rest at infinity.

$$V = -\frac{GM}{r}\left[1 - J_2\left(\frac{a_1}{r}\right)^2 P_2(\cos\vartheta)\right] \qquad (18.11)$$

$$\Phi = V - (1/2)\,(\omega_E r' \sin\vartheta')^2 \qquad (18.13)$$

The SI units of time are defined by atomic clocks at rest in the ECEF system at the mean sea level.

Because of rotational flattening of the Earth these clocks are at different distances from the Earth's centre.

So, for precise synchronisation, we need the equation of the idealised Earth's surface (called ***geoid***) and the values of Φ all over it.

The geoid is defined as the surface of constant effective potential, thus from (18.11) and (18.13)

$$-\frac{2GM}{r'}\left[1 - J_2\left(\frac{a_1}{r'}\right)^2 \frac{3\cos^2\vartheta' - 1}{2}\right] - (\omega_E r' \sin\vartheta')^2 \stackrel{\text{def}}{=} 2\frac{\Phi_0}{c^2} = \text{constant} \qquad (18.14)$$

Models of the geoid that include higher multipoles are known.

But the sum of the first 20 multipoles gives a correction to Φ_0/c^2 of the order of 10^{-16}, so it is negligible.

$$-\mathrm{d}s^2 = -\left(1+\frac{2\Phi}{c^2}\right)(c\mathrm{d}t')^2 + 2\frac{\omega_E}{c}\, r'^2 \sin^2\vartheta' c\mathrm{d}t'\mathrm{d}\varphi' + \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r'^2 + r'^2\mathrm{d}\vartheta'^2 + r'^2\sin^2\vartheta'\mathrm{d}\varphi'^2\right) \quad (18.12)$$

$$-\frac{2GM}{r'}\left[1 - J_2\left(\frac{a_1}{r'}\right)^2 \frac{3\cos^2\vartheta' - 1}{2}\right] - (\omega_E r'\sin\vartheta')^2 \stackrel{\text{def}}{=} 2\frac{\Phi_0}{c^2} = \text{constant} \quad (18.14)$$

To find the value of Φ_0/c^2, we take it on the equator, where $\vartheta' = \pi/2$ and $r' = a_1$:

$\Phi_0/c^2 = - 6.9693 \times 10^{-10}$. (18.15)

It is seen from (18.12) that τ - the proper time of a clock at rest on the geoid, is related to the inertial time at infinity t' by

$d\tau = ds/c = dt' (1 + \Phi_0/c^2)$. (18.16)

Thus, the clocks on the Earth *run slow* compared to the (fictional, idealised) clocks at rest at infinity (because $\Phi_0 < 0$).

The rate of the two frequencies is $\approx -7 \times 10^{-10}$ ≈ 10 000 times the fractional frequency stability of the best cesium clocks.

$$-\mathrm{d}s^2 = -\left(1+\frac{2\Phi}{c^2}\right)(c\mathrm{d}t')^2 + 2\frac{\omega_E}{c}\, r'^2 \sin^2\vartheta' c\mathrm{d}t'\mathrm{d}\varphi' + \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r'^2 + r'^2\mathrm{d}\vartheta'^2 + r'^2\sin^2\vartheta'\mathrm{d}\varphi'^2\right) \qquad (18.12)$$

Since Φ_0 is constant, all atomic clocks at the sea level beat at the same rate.

→ The time coordinate on the Earth may be rescaled by a constant factor so that it coincides with the proper time of the standard clocks

$$t'' = (1 + \Phi_0/c^2)\, t'. \qquad (18.17)$$

Up to terms of order $1/c^2$, (18.17) changes the metric (18.12) as follows:

$$-\mathrm{d}s^2 = -\left(1+2\frac{\Phi-\Phi_0}{c^2}\right)(c\mathrm{d}t'')^2 + 2\frac{\omega_E}{c}\, r''^2 \sin^2\vartheta' \mathrm{d}\varphi' c\mathrm{d}t'' + \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r'^2 + r'^2\mathrm{d}\vartheta'^2 + r'^2\sin^2\vartheta'\mathrm{d}\varphi'^2\right). \qquad (18.18)$$

On the geoid, where $\Phi = \Phi_0$, the time coordinate t'' is the proper time.

We will drop the primes on t'' from now on.

The ECI metric results from (18.18) when $\omega_E = 0$ and $\Phi = V$.

$$-ds^2 = -\left(1+\frac{2V}{c^2}\right)(cdt)^2 + \left(1-\frac{2V}{c^2}\right)\left(dr^2 + r^2 d\vartheta^2 + r^2\sin^2\vartheta d\varphi^2\right) \qquad (18.10)$$

$t'' = (1 + \Phi_0/c^2)\, t'$. (18.17)

18.5. The realisation of the coordinate time

Now apply the transformation (18.17) to the metric (18.10):

$$-ds^2 = -\left(1+2\frac{V-\Phi_0}{c^2}\right)(cdt)^2 + \left(1-\frac{2V}{c^2}\right)\left(dr^2 + r^2 d\vartheta^2 + r^2\sin^2\vartheta d\varphi^2\right) \qquad (18.19)$$

This metric is in the ECI reference system.

Clocks moving with respect to it will display the time-dilation resulting from their motion, and the gravitational frequency shifts connected with different values of V at their locations.

But the time-coordinate in (18.19) is the well-defined time on the geoid, so it can be used as a universal basis for synchronisation in the whole GP system.

$$-\mathrm{d}s^2 = -\left(1+2\frac{V-\Phi_0}{c^2}\right)(c\mathrm{d}t)^2 + \left(1-\frac{2V}{c^2}\right)\left(\mathrm{d}r^2 + r^2\mathrm{d}\vartheta^2 + r^2\sin^2\vartheta\mathrm{d}\varphi^2\right) \quad (18.19)$$

To take into account the time-dilation and gravitational frequency shifts, let us factor out $(cdt)^2$ from (18.19):

$$\mathrm{d}s^2 = \left[1+2\frac{V-\Phi_0}{c^2} - \left(1-\frac{2V}{c^2}\right)\frac{v^2}{c^2}\right](c\mathrm{d}t)^2 \quad (18.20)$$

where

$$v^2 \stackrel{\text{def}}{=} \left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2 + r^2\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}t}\right)^2 + r^2\sin^2\vartheta\left(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\right)^2 \quad (18.21)$$

is the velocity of the clock in the ECI system. We can neglect the second V/c^2 in (18.20) since it is multiplied by $(v/c)^2$.

Neglecting also other terms of the order $(v/c)^4$, the proper time of a moving clock is related to its time-coordinate t by

$$\mathrm{d}s/c \stackrel{\text{def}}{=} \mathrm{d}\tau = \left(1+\frac{V-\Phi_0}{c^2} - \frac{v^2}{2c^2}\right)\mathrm{d}t \quad (18.22)$$

Solving this for t, again only up to terms of order $(v/c)^2$, we obtain

$$\int_{\text{path}} \mathrm{d}t = \int_{\text{path}} \left(1-\frac{V-\Phi_0}{c^2} + \frac{v^2}{2c^2}\right)\mathrm{d}\tau \quad (18.23)$$

$$\int\limits_{\text{path}} \mathrm{d}t = \int\limits_{\text{path}} \left(1 - \frac{V - \Phi_0}{c^2} + \frac{v^2}{2c^2}\right) \mathrm{d}\tau \qquad (18.23)$$

This is the set of corrections to be applied in synchronising the GPS clocks.

The last term in (18.23) is often called the ***transverse*** or second order ***Doppler effect***.

$$\int_{\text{path}} \mathrm{d}t = \int_{\text{path}} \left(1 - \frac{V - \Phi_0}{c^2} + \frac{v^2}{2c^2}\right) \mathrm{d}\tau \qquad (18.23)$$

18.6. Selected corrections of the orbits of the GPS satellites

18.6.1. Corrections for gravity and velocity

Φ_0 includes the Earth's quadrupole moment, but the quadrupole contribution in Φ_0/c^2 is

$GMJ_2/(2c^2a_1) \approx 3.76 \times 10^{-13}$.

At the satellites' positions this would be $\approx 10^{-14}$, and we will neglect it for a while.

We also assume that the GPS satellites follow Keplerian orbits.

This is not the case for low orbits, but the GPS orbits have radii of ≈ 25 000 km, and for them the assumption is fulfilled with sufficient accuracy.

With the quadrupole term neglected, the satellites move in the potential $V = - GM/r$.

Let a be the semimajor axis of the orbit, and e its eccentricity.

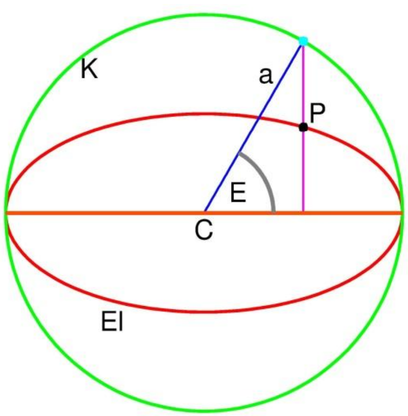
The orbits are then ellipses, and their equation is

$$r = a\,\frac{1-e^2}{1-e\cos\varphi}, \qquad e = \sqrt{1-(b/a)^2} \tag{18.24}$$

where φ is the azimuthal coordinate of the satellite in the polar coordinates whose origin coincides with the center of the Earth and with the focus of the ellipse (this coordinate is also called *true anomaly*).

Another useful description of the orbit is via the *eccentric anomaly* :

r = a (1 – e cos E). (18.25)



Definition of the eccentric anomaly E for the point P on the ellipse El with semimajor axis a.

The circle K has radius a and is tangent to El at both ends of its longer axis.

Imagine the ellipse being stretched in the vertical direction by the ratio a/b, where b is its seminor axis.

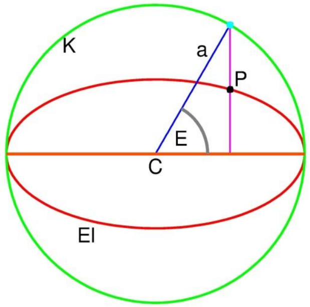
After this, El will coincide with K.

E is the azimuthal coordinate of the point on K where P will land after the stretching.

$$r = a\,\frac{1-e^2}{1-e\cos\varphi} \qquad (18.24) \qquad J^2 = GMm^2a(1-e^2), \qquad mr^2 d\varphi/dt = J, \qquad v^2 = (dr/dt)^2 + r^2(d\varphi/dt)^2$$

The relation between the two anomalies is

$$\cos\varphi = \frac{\cos E - e}{1 - e\cos E} \qquad (18.26)$$



The dependence of E on the Newtonian time is

$$E - e\sin E = \sqrt{\frac{GM}{a^3}}\;(t - t_p) \qquad (18.27)$$

where t_p is the time of passage of the satellite through the perigee E = 0.

Using the formulae written at the top, the total energy per unit mass of a satellite is

$$\frac{v^2}{2} - \frac{GM}{r} = -\frac{GM}{2a} \qquad (18.28)$$

so it is constant on the orbit, as expected.

$$\int_{\text{path}} \mathrm{d}t = \int_{\text{path}} \left(1 - \frac{V - \Phi_0}{c^2} + \frac{v^2}{2c^2}\right) \mathrm{d}\tau \qquad (18.23)$$

$$\frac{v^2}{2} - \frac{GM}{r} = -\frac{GM}{2a} \qquad (18.28)$$

After using (18.28) in (18.23) to eliminate v^2, and rearranging the terms we obtain

$$\int_{\text{path}} \mathrm{d}t = \int_{\text{path}} \left[1 + \frac{3GM}{2ac^2} + \frac{\Phi_0}{c^2} - \frac{2GM}{c^2}\left(\frac{1}{a} - \frac{1}{r}\right)\right] \mathrm{d}\tau \qquad (18.29)$$

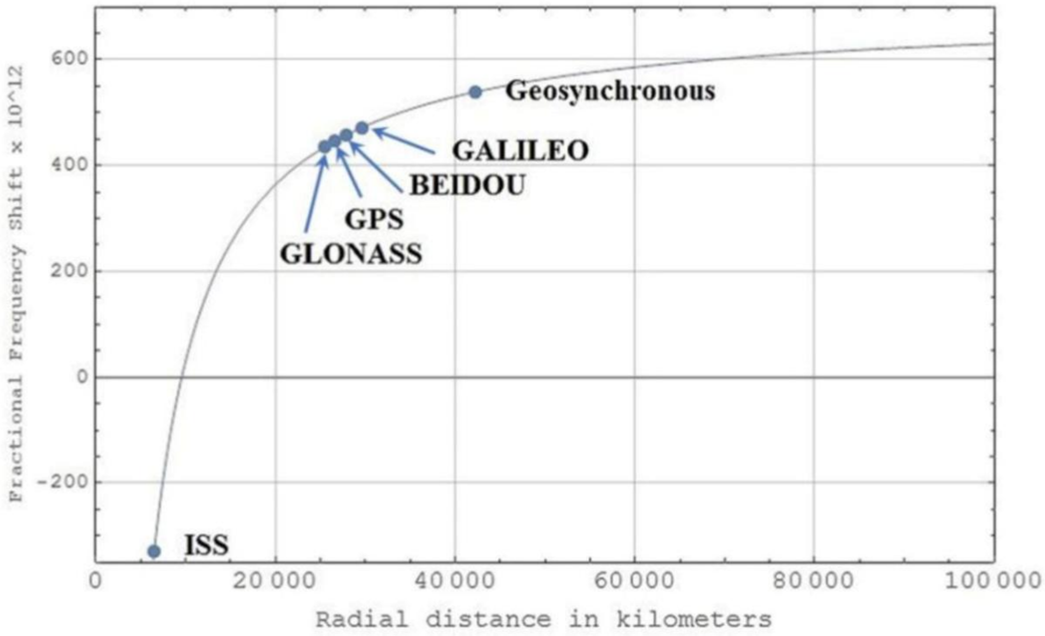
The term in parentheses is zero on circular orbits, while the two constant terms, on a GPS satellite orbit, sum up to

$$\frac{\Delta t_{\text{GPS}}}{\Delta\tau} \stackrel{\text{def}}{=} \frac{3GM}{2ac^2} + \frac{\Phi_0}{c^2} = -4.465 \times 10^{-10} \qquad (18.30)$$

Since $\Delta t_{GPS} < 0$, a clock on a circular orbit with *a* the same as for GPS satellites goes faster than a clock on the Earth.

→ In order that the receivers on the Earth see the right frequency of the clock signal, 10.23 MHz, the clocks on satellites are adjusted before launch to go at the corrected frequency:

$$\nu_{\text{corr}} = \left(1 - 4.465 \times 10^{-10}\right) \times 10.23\ \text{MHz} = 10.22999999543\ \text{MHz} \qquad (18.31)$$



$$\frac{\Delta t_{\mathrm{GPS}}}{\Delta \tau} \stackrel{\mathrm{def}}{=} \frac{3GM}{2ac^2} + \frac{\Phi_0}{c^2} = -4.465 \times 10^{-10} \tag{18.30}$$

Fractional frequency shifts at various distances from the Earth's center. The distance is the *a* in Eq. (18.30).

On low orbits (for example the Space Shuttle orbit) $\Delta\nu < 0$ because *a* in (18.30) is smaller, so $3GM/(2ac^2) > 0$ is large, and makes $\Delta t/\Delta\tau > 0$, so $\Delta\nu/\nu < 0$.

On high orbits $\Delta\nu/\nu > 0$.

The two effects cancel out at $a \approx 9545$ km.

At the launch of the first GPS satellite on 23 June 1977 some engineers did not believe that relativistic effects were meaningful.

But, just in case, a frequency synthesiser was built into the satellite clock system so that corrections could be switched on if relativity turned out to be correct.

The uncorrected system ran for 20 days, during which the frequency measured for the satellite clock was by

442.5×10^{-12}

larger than for the clocks on the ground.

The prediction of relativity was

446.5×10^{-12}.

This was in fact an experiment confirming the combination of two relativistic effects (transverse Doppler and gravitational redshift).

After 20 days, the frequency correction was switched on. In later-built satellites, the correction is built in into the hardware.

r = a (1 – e cos E). (18.25)

$$E - e\sin E = \sqrt{\frac{GM}{a^3}}\ (t - t_p) \qquad (18.27)$$

$$\int_{\text{path}} \mathrm{d}t = \int_{\text{path}} \left[1 + \frac{3GM}{2ac^2} + \frac{\Phi_0}{c^2} - \frac{2GM}{c^2}\left(\frac{1}{a} - \frac{1}{r}\right)\right] \mathrm{d}\tau \qquad (18.29)$$

18.6.2. The eccentricity correction

The last term in (18.29) can be integrated exactly if we observe, from (18.27) and (18.25), that

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \frac{1}{1 - e\cos E}\sqrt{\frac{GM}{a^3}}, \qquad (18.32)$$

$$\frac{1}{r} - \frac{1}{a} = \frac{e\cos E}{a(1 - e\cos E)} \equiv \frac{e}{a}\sqrt{\frac{a^3}{GM}}\ \cos E\frac{\mathrm{d}E}{\mathrm{d}t} \qquad (18.33)$$

This last term in (18.29) contains the factor $1/c^2$, and the difference between dt and dτ is also of order $1/c^2$.

→ If we replace dτ by dt in the integral, the error will be of order $1/c^4$ - negligible at our assumed accuracy. Consequently, using (18.32) we obtain

$$\Delta t = \int \frac{2GM}{c^2}\left(\frac{1}{a} - \frac{1}{r}\right)\mathrm{d}\tau = \frac{2GM}{c^2}\int\left(\frac{1}{a} - \frac{1}{r}\right)\mathrm{d}t = \frac{2\sqrt{GMa}}{c^2}e(\sin E - \sin E_0) \qquad (18.34)$$

The additive constant ~ sin E_0 only resets the origin of time.

$$\Delta t = \int \frac{2GM}{c^2}\left(\frac{1}{a} - \frac{1}{r}\right) \mathrm{d}\tau = \frac{2GM}{c^2}\int \left(\frac{1}{a} - \frac{1}{r}\right) \mathrm{d}t = \frac{2\sqrt{GMa}}{c^2} e(\sin E - \sin E_0) \qquad (18.34)$$

The variable term in (18.34) is

$$\Delta t = +\left(4.4428 \times 10^{-10} \frac{\mathrm{s}}{\sqrt{\mathrm{m}}}\right) e\sqrt{a}\sin E \qquad (18.35)$$

This correction could have been applied in the satellite clocks.

But insufficient computing power available in the 1970s dictated the decision to apply (18.35) in the receivers.

Now it is too late to reverse this decision because of the large investment already made by manufacturers into the devices existing on the market.

This means that each hand-held GPS locator applies this relativistic correction (and the same is true for other GNSS-s).

$$\Delta t = + \left(4.4428 \times 10^{-10} \frac{\mathrm{s}}{\sqrt{\mathrm{m}}}\right) e\sqrt{a} \sin E \qquad (18.35)$$

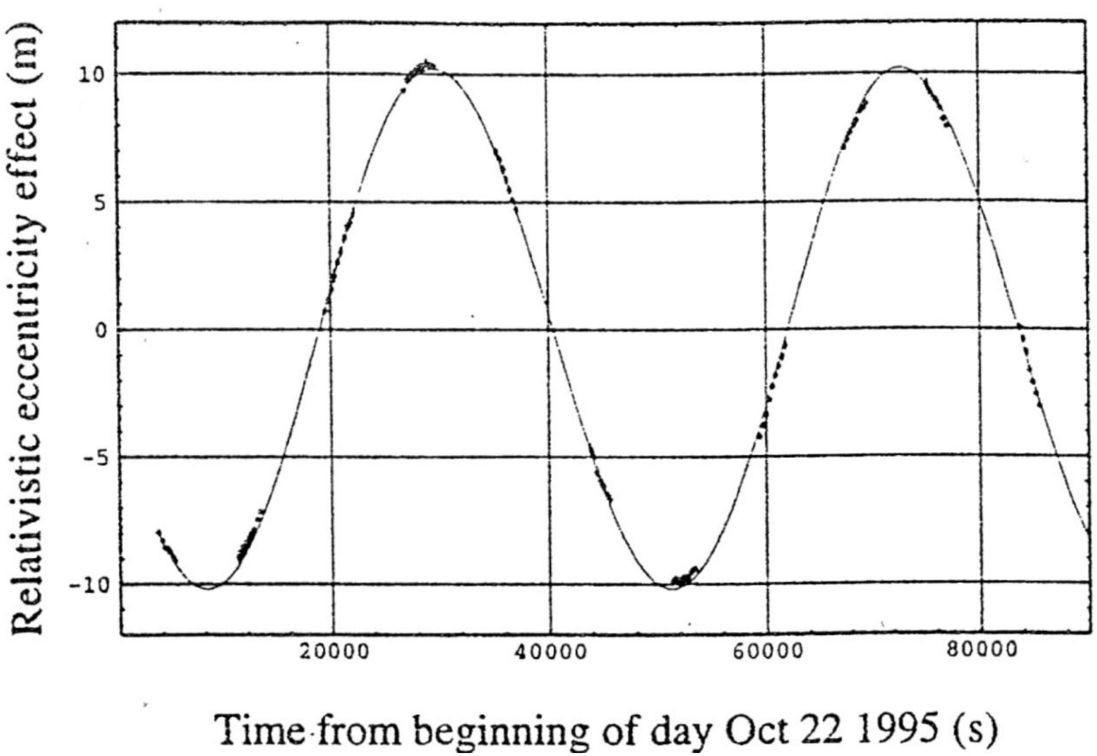
The eccentricity effect (18.35) allows for one more test of relativity.

It was carried out because an official document distributed in 1997 by the Aerospace Corporation claimed that the correction (18.35) could be neglected for a receiver on the orbit.

This error, if actually implemented in the system, would have catastrophic consequences(*): the next generation of the GPS satellites is able to work independently of ground control for periods of up to 180 days.

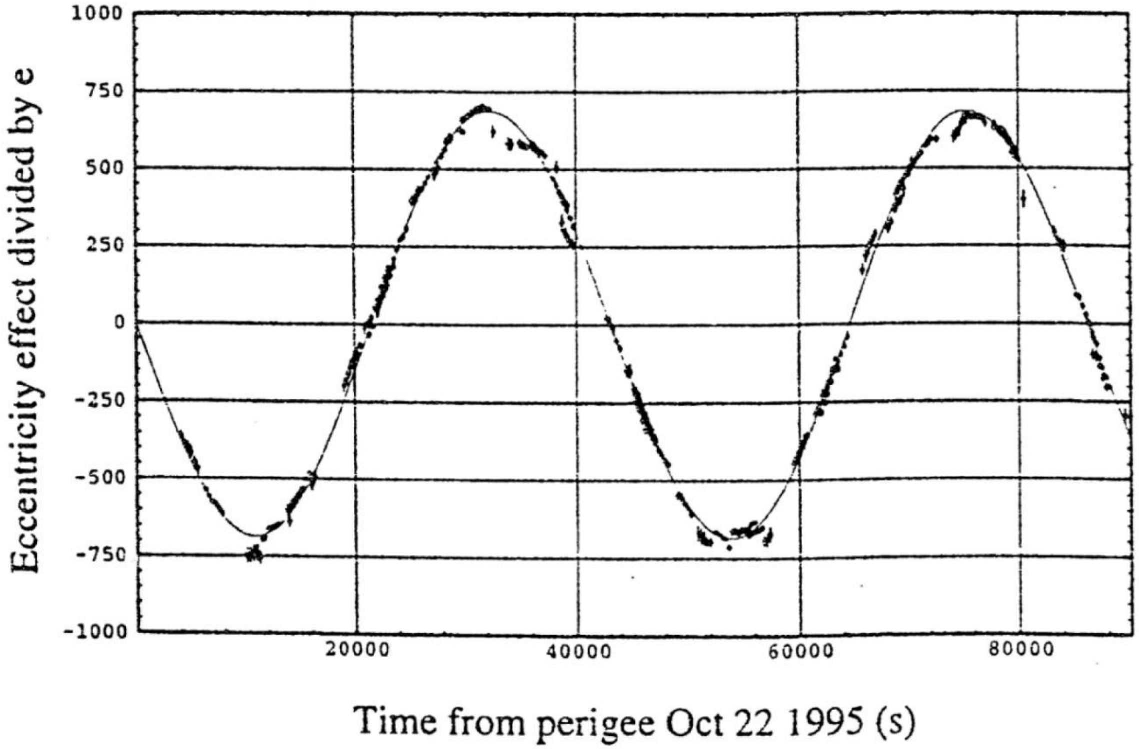
They determine their positions by picking up signals from other GPS satellites.

(*) My private comment: This is an example of what can happen when self-assured but poorly informed individuals pose to be experts and use rhetorical arguments. This is a frequent phenomenon in astrophysics.



The relativistic eccentricity effect 1: Comparison of experimental results with the prediction of the theory for the GPS satellite with the largest e = 0.01486.

The mean difference in positions between theory and experiment is 22 cm, which is 2.2% of the amplitude, 10.2 m.



The relativistic eccentricity effect 2: Δt/*e* measured experimentally vs. the prediction of theory for 5 GPS satellites with the largest values of *e*.

Here the agreement between the theory and experiment was at the level of 2.5%.

18.7. The 9 largest relativistic effects in the GPS

The relativistic effects influence the rate of time flow in the components of the system.

Because of the large value of c, small discrepancies in timing translate into large differences in calculated positions of the receivers.

If the relativistic corrections were neglected, the errors would accumulate over time.

An illustrative measure of the importance of those corrections is the error in determining the position of the receiver if the corrections were neglected for 24 hours.

The table on the next page, based on Ref. [227], shows the 9 largest effects.

[227] N. Ashby, *Relativity in the palm of your hand*. **Mercury** 25, 23 (1996).

The influence of relativistic effects on the precision of location in the GPS
(Adapted from Ref. [227])

Effect	Error after 24 hours
Influence on the clocks in the receivers	
Earth's gravitational field	**18 km**
Oblateness of the Earth	9.7 m
Altitude of the clock (e.g. 10 km)	28 m
Rotation of the Earth (on the equator)	31 m
Velocity of the clock (e.g. 600 km/h)	10 m
Synchronisation on the rotating Earth (Sagnac effect)	up to 62 m
Influence on the clocks in the satellites	
Earth's gravitational field	**4.3 km**
Orbital velocity of the satellite	**2.2 km**
Influence on the propagation of the electromagnetic waves	
Rotation of the Earth	up to 41 m

Take-home message:

Every time when we determine our position using a GPS receiver and the result is correct, we carry out an experiment that confirms the predictions of general relativity.