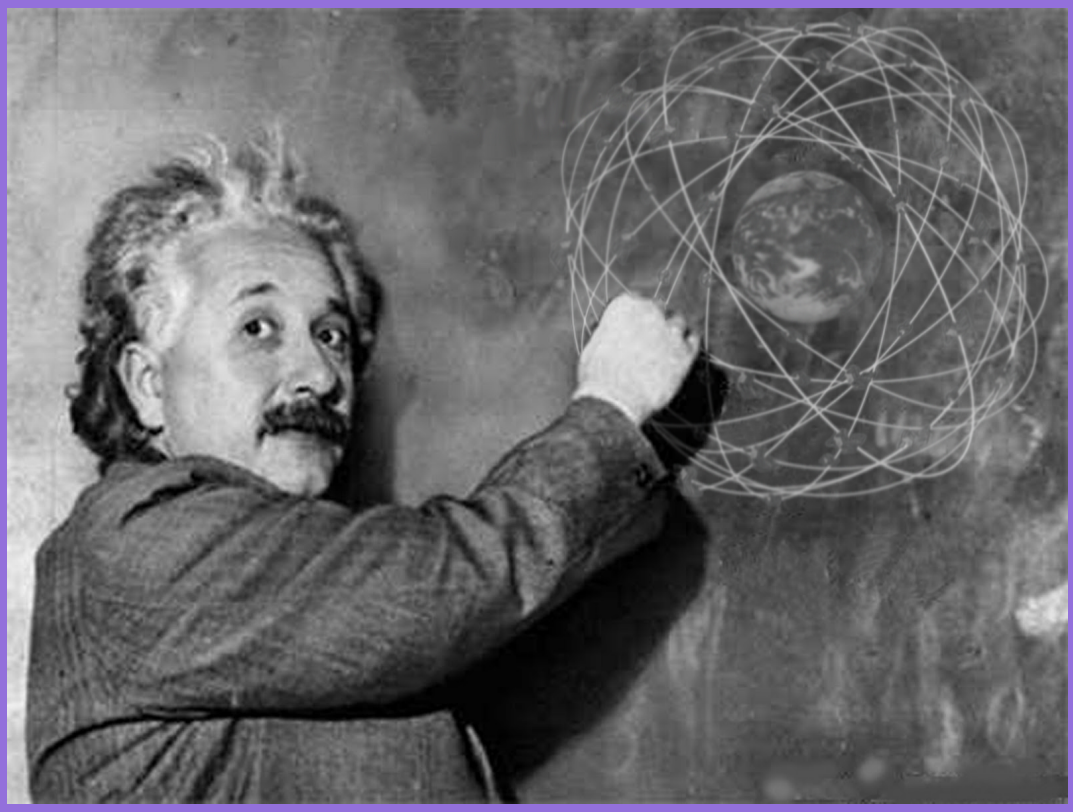
DOI: 10.2420/ACT-BOK-AF
ISSN: 2309-1940

Acta Futura

Issue 7

Relativistic Positioning Systems and their Scientific Applications

Guest Editors: Andreja Gomboc, Martin Horvat and Uroš Kostić



Advanced Concepts Team

http://www.esa.int/act

European Space Agency
Agence spatiale européenne

Publication: Acta Futura, Issue 7 (2013)

Editor-in-chief: Duncan James Barker

Associate editors: Dario Izzo
Jose M. Llorens Montolio
Pacôme Delva
Francesco Biscani
Camilla Pandolfi
Guido de Croon
Luís F. Simões
Daniel Hennes

Published and distributed by: Advanced Concepts Team
ESTEC, PPC-PF
2201 AZ, Noordwijk
The Netherlands
`www.esa.int/gsp/ACT/publications/ActaFutura`
Fax: +31 71 565 8018

Cover image by S. Carloni & A. Gomboc from `http://rgnss.fmf.uni-lj.si/workshop`.

ISSN: 2309-1940
D.O.I.:10.2420/ACT-BOK-AF

Typeset with XeTeX

Workshop on

Relativistic Positioning Systems and their Scientific Applications

19-21 September 2012
Brdo near Kranj, Slovenia

Organised by the Faculty of Mathematics and Physics, University of Ljubljana (UL), ESA (Advanced Concepts Team) and Centre of Excellence SPACE-SI.

Organising Committee:

Andreja Gomboc (UL, CE SPACE-SI)
Martin Horvat (UL)
Uroš Kostić (UL)

Scientific Committee:

Bertram Arbesser-Rastburg (ESA)
Sante Carloni (ESA)
Pacôme Delva (SYRTE, UPMC, Paris Obs.)
Clovis Jacinto de Matos (ESA)
Rune Floberghagen (ESA)
Uroš Kostić (UL)
Leopold Summerer (ESA)

Contents

Foreword

In the last decade we have seen an enormous growth in satellite navigation systems and the utilization thereof. GPS and GLONASS are widely used for a wide range of applications and new systems like Galileo and Beidou are being set up to widen the possibilities. All these systems have in common, that they are based on Newtonian physics. Relativistic effects (sometimes called "Newtonian defects") are treated as deviations that need to be corrected for. In parallel, in the last decade, a number of physicists have worked on theoretical concepts of relativistic positioning and relativistic reference frames.

In a unique collaboration between the ESA Advanced Concepts Team and the Faculty of Mathematics and Physics of the University of Ljubljana with the support of the Slovenian Centre of Excellence for Space Sciences and Technologies this workshop was organised in the picturesque environment of Brdo, Slovenia. The main objectives was to bring physicists and engineers together to discuss the latest theoretical and experimental findings as well as potential implementations.

The programme was set up with a view to maximising the interaction between the participants by allowing sufficient time for discussion after every presentation and at the end of every session. Five well recognized specialists were invited to give plenary lectures on key concepts and the evolution of understanding since Karl Schwarzschild and Hermann Minkowski.

The workshop was also addressing potential applications from fundamental physics to relativistic autonomous navigation as well as enabling technologies such as optical clocks and inter-satellite links.

In the concluding discussion there was strong agreement that the workshop has achieved its objectives and that further similar events should be organised in the future.

Bertram Arbesser-Rastburg
(Scientific Organizing Comittee)

Current Global Navigational Satellite Systems rely on global reference frames which are fixed to the Earth via ground stations so their precision and stability in time are limited by our knowledge of the Earth's dynamics (e.g. plate tectonic motions, tidal effects on the Earth's crust and variations of the Earth's rotation rate). These limitations could be avoided by giving to a constellation of satellites the possibility of constituting by itself a primary and autonomous positioning system, without any a priori realization of a terrestrial reference frame. Such a system leads to numerous advantages, e.g. it is a primary reference frame which is not tied to the Earth, the relativistic effects are already included in the definition of the positioning system, so there is no need to synchronize the clocks, the reference frame is very precise and stable, a better understanding of the principles of positioning systems, the new coordinates defined are measurable directly and as such open new possibilities in experimental physics and astronomy, the system can be used for extra-terrestrial navigation with the use of pulsars as clocks, etc.

The workshop "Relativistic Positioning Systems and their Scientific Applications" was co-organised by the University of Ljubljana, the Center of Excellence SPACE-SI and the European Space Agency (ESA). It was held at Brdo, Slovenia from September 19th to 21st, 2012. It aimed at bringing together people working on different aspects of relativistic positioning systems, and giving them the possibility to interact with each other and exchange ideas and knowledge in a direct way. During the three days of the workshop, we had the opportunity to hear the contributions covering most of the above topics. In these proceedings we try to encompass all the concepts which were presented and discussed on this occasion. The positive responses of the participants show that the area of Relativistic Positioning System research is certainly interesting and lively one, and we hope that it will continue to attract interest and people to work in further development of Relativistic Positioning Systems.

Andreja Gomboc, Martin Horvat and Uroš Kostić
(Guest Editors)

Acta Futura 7 (2013) 11-28
DOI: 10.2420/AF07.2013.11

Acta Futura

Autonomous Spacecraft Navigation With Pulsars

Werner Becker,*
Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße, 85748 Garching, Germany
Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany
Mike G. Bernhardt,
Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße, 85748 Garching, Germany
and Axel Jessner
Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

Abstract. An external reference system suitable for deep space navigation can be defined by fast spinning and strongly magnetized neutron stars, called pulsars. Their beamed periodic signals have timing stabilities comparable to atomic clocks and provide characteristic temporal signatures that can be used as natural navigation beacons, quite similar to the use of GPS satellites for navigation on Earth. By comparing pulse arrival times measured on-board a spacecraft with predicted pulse arrivals at a reference location, the spacecraft position can be determined autonomously and with high accuracy everywhere in the solar system and beyond. The unique properties of pulsars make clear already today that such a navigation system will have its application in future astronautics. In this paper we describe the basic principle of spacecraft navigation using pulsars and report on the current development status of this novel technology.

1 Introduction

Today, the standard method of navigation for interplanetary spacecraft is a combined use of radio data, obtained by tracking stations on Earth, and optical data from an on-board camera during encounters with solar system bodies. Radio measurements taken by ground stations provide very accurate information on the distance and the radial velocity of the spacecraft with typical random errors of about 1 m and 0.1 mm/s, respectively [43]. The components of position and velocity perpendicular to the Earth-spacecraft line, however, are subject to much larger errors due to the limited angular resolution of the radio antennas. Interferometric methods can improve the angular resolution to about 25 nrad, corresponding to an uncertainty in the spacecraft position of about 4 km per astronomical unit (AU) of distance between Earth and spacecraft [37]. With increasing distance from Earth, the position error increases as well, e.g., reaching a level of uncertainty of the order of $\pm$200 km at the orbit of Pluto and $\pm$500 km at the distance of Voyager 1. Nevertheless, this technique has been used successfully to send space probes to all planets in the solar system and to study asteroids and comets at close range. However, it might be necessary for future missions to overcome the disadvantages of this method, namely the dependency on ground-based control and maintenance, the increasing position and velocity un-

*Corresponding author. E-mail: wbecker@mpe.mpg.de

certainty with increasing distance from Earth as well as the large propagation delay and weakening of the signals at large distances. It is therefore desirable to automate the procedures of orbit determination and orbit control in order to support autonomous space missions.

Possible implementations of autonomous navigation systems were already discussed in the early days of space flight [4]. In principle, the orbit of a spacecraft can be determined by measuring angles between solar system bodies and astronomical objects; e.g., the angles between the Sun and two distant stars and a third angle between the Sun and a planet. However, because of the limited angular resolution of on-board star trackers and sun sensors, this method yields spacecraft positions with uncertainties that accumulate typically to several thousand kilometers. Alternatively, the navigation fix can be established by observing multiple solar system bodies: It is possible to triangulate the spacecraft position from images of asteroids taken against a background field of distant stars. This method was realized and flight-tested on NASA's Deep-Space-1 mission between October 1998 and December 2001. The Autonomous Optical Navigation (AutoNav) system onboard Deep Space 1 provided the spacecraft orbit with 1σ errors of ± 250 km and ± 0.2 m/s, respectively [48]. Although AutoNav was operating within its validation requirements, the resulting errors were relatively large compared to ground-based navigation.

In the 1980s, scientists at NRL (United States Naval Research Laboratory) proposed to fly a demonstration experiment called the Unconventional Stellar Aspect (USA) experiment [57]. Launched in 1999 on the Advanced Research and Global Observation Satellite (ARGOS), this experiment demonstrated a method of position determination based on stellar occultation by the Earth's limb as measured in X-rays. This technique, though, is limited to satellites in low Earth orbit.

An alternative and very appealing approach to autonomous spacecraft navigation is based on pulsar timing. The idea of using these celestial sources as a natural aid to navigation goes back to the 1970s when Downs [29] investigated the idea of using pulsating radio sources for interplanetary navigation. Downs analyzed a method of position determination by comparing pulse arrival times at the spacecraft with those at a reference location. Within the limitations of technology and pulsar data available at that time (a set of only 27 radio pulsars were considered), Downs showed that spacecraft position errors on the order of 1500 km could be obtained after 24 hours of signal integration. A possible improvement in precision by a factor of 10 was estimated if better (high-gain) radio antennas were available for the observations.

Chester & Butman [22] adopted this idea and proposed to use X-ray pulsars, of which about one dozen were known at the time, instead of radio pulsars. They estimated that 24 hours of data collection from a small on-board X-ray detector with $0.1\,\mathrm{m}^2$ collecting area would yield a three-dimensional position accurate to about 150 km. Their analysis, though, was not based on simulations or actual pulsar timing analyses; neither did it take into account the technological requirements or weight and power constraints for implementing such a navigation system.

These early studies on pulsar-based navigation estimated relatively large position and velocity errors so that this method was not considered to be an applicable alternative to the standard navigation schemes. However, pulsar astronomy has improved considerably over the last 30 years since these early proposals. Meanwhile, pulsars have been detected across the electromagnetic spectrum and their emission properties have been studied in great detail (cf. [8] for a collection of comprehensive reviews on pulsar research). Along with the recent advances in detector and telescope technology this motivates a general reconsideration of the feasibility and performance of pulsar-based navigation systems. The present paper reports on our latest results and ongoing projects in this field of research. Its structure is as follows: After summarizing the most relevant facts on pulsars and discussing which pulsars are best suited for navigation purposes in § 2, we briefly describe the principles of pulsar-based navigation in § 3. Pulsars emit broadband electromagnetic radiation which allows an optimization for the best suited waveband according to the highest number of bits per telescope collecting area, power consumption, navigator weight and compactness. A possible antenna type and size of a navigator which detects pulsar signals at 21 cm is described in § 4. In § 5 we discuss the possibility of using X-ray signals from pulsars for navigation. The recent developments of low-mass X-ray mirrors and active-pixel detectors, briefly summarized in § 6, makes it very appealing to use this energy band for pulsar-based navigation.

 DOI: 10.2420/AF07.2013.11

2 The Various Types of Pulsars and their Relevance for Navigation

Stars are stable as long as the outward-directed thermal pressure, caused by nuclear fusion processes in the central region of the star, and the inward-directed gravitational pressure are in equilibrium. The outcome of stellar evolution, though, depends solely on the mass of the progenitor star. A star like our sun develops into a white dwarf. Stars above $\approx 8\,M_\odot$ undergo a gravitational collapse once their nuclear fuel is depleted. Very massive stars of more than about $30\,M_\odot$ end up as black holes and stars in the intermediate mass range of about 8 to $30\,M_\odot$ form neutron stars. It is assumed that a neutron star is the result of a supernova explosion, during which the bulk of its progenitor star is expelled into the interstellar medium. The remaining stellar core collapses under its own weight to become a very compact object, primarily composed of neutrons – a neutron star. With a mass of typically $1.4\,M_\odot$, compressed into a sphere of only 10 km in radius, they are quasi gigantic atomic nuclei in the universe. Because of their unique properties they are studied intensively by physicists of various disciplines since their discovery in 1967 [36].

Fast spinning and strongly magnetized neutron stars are observable as pulsars if their spin axis and magnetic field axis are not aligned. Having co-rotating magnetic fields of $B_\perp \approx 10^9$–10^{13} G and spin periods down to milliseconds they radiate broadband electromagnetic radiation along narrow emission cones. If this radiation cone crosses the observer's line of sight a pulse of intensity is recorded in the observing device (cf. Figure 1).

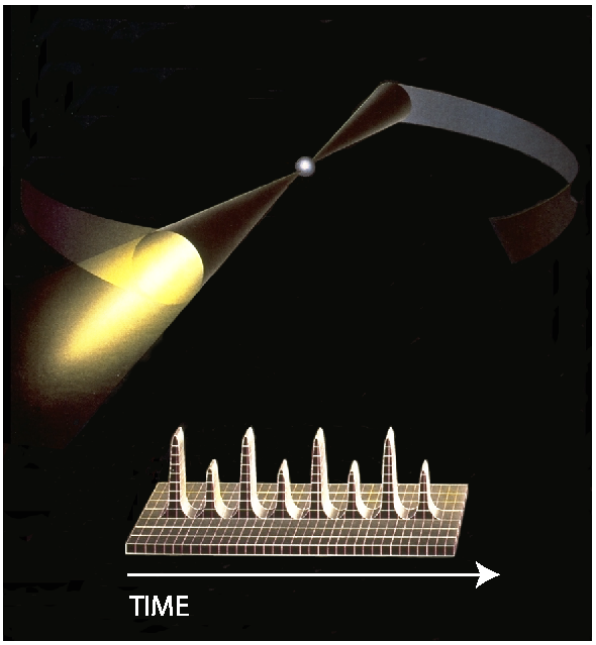


Figure 1. *Artist's impression of a rotation-powered pulsar. The neutron star appears as a pulsating source of radiation if the rotating emission beam crosses the observer's line of sight. Averaging these periodic pulses of intensity over many rotation cycles results in a stable pulse profile. Because of the timing stability of most pulsars, the arrival time of pulses can be predicted with very high precision, which is an essential requirement for a navigation system based on pulsar observations.*

The name *Pulsar* refers to this property. They have been discovered by their radio signals [36]. In source catalogs their common abbreviation is therefore *PSR* which stands for Pulsating Source of Radio, although they have also been detected in other bands of the electromagnetic spectrum meanwhile. Three different classes of pulsars can be distinguished according to the energy source of their electromagnetic radiation. As we will see, only one class is suitable for spacecraft navigation:

- **Accretion-powered pulsars** are close binary systems in which a neutron star is accreting matter from a companion star, thereby gaining energy and angular momentum. There are no radio waves emitted from the accretion process, but these systems are bright in X-rays. The observed X-ray pulses are due to the changing viewing angle of a million degree hot spot on the surface of the neutron star. These hot spots are heated by in-spiraling matter from an accretion disk. The accretion disk and the accretion column itself can also be sources of X-rays. The spin behavior of accretion-powered pulsars can be very complicated and complex. They often show an unpredictable evolution of rotation period, with erratic changes between spin-up and spin-down as well as X-ray burst activities [34]. Although accretion-powered pulsars are usually bright X-ray sources, and thus would give only mild constraints on the sensitivity requirements of a pulsar-based navigation system, their unsteady and non-coherent timing behavior disqualifies them as reference sources for navigation.

- **Magnetars** are isolated neutron stars with exceptionally high magnetic dipole fields of up to 10^{15} G.

All magnetars are found to have rotation periods in the range of about 5 to 10 seconds. RXTE and other X-ray observatories have detected super-strong X-ray bursts with underlying pulsed emission from these objects. According to the magnetar model of Duncan & Thompson [30], their steady X-ray emission is powered by the decay of the ultra-strong magnetic field. This model also explains the X-ray burst activity observed from these objects, but there are also alternative theories, which relate these bursts to a residual fall-back disk [53, 52]. However, their long-term timing behavior is virtually unknown, which invalidates these sources for the use in a pulsar-based spacecraft navigation system.

Concerning their application for navigation, the only pulsar class that really qualifies is that of rotation-powered ones.

- **Rotation-powered pulsars** radiate broadband electromagnetic radiation (from radio to optical, X- and gamma-rays) at the expense of their rotational energy, i.e., the pulsar spins down as rotational energy is radiated away by its co-rotating magnetic field. The amount of energy that is stored in the rotation of the star can be estimated as follows: A neutron star with a radius of $R = 10$ km and a mass of $M = 1.4\ M_\odot$ has a moment of inertia $I \approx (2/5)\, M\, R^2 \approx 10^{45}$ g cm^2. The rotational energy of such a star is $E_{rot} = 2\pi^2\, I\, P^{-2}$. Taking the pulsar in the Crab nebula with $P \approx 33$ ms as an example, its rotational Energy is $E_{rot} \approx 2 \times 10^{49}$ erg, which is comparable with the energy released by thermonuclear burning of our sun in hundred million years. The spin period of a rotation-powered pulsar increases with time due to a braking torque exerted on the pulsar by its magneto-dipole radiation. For the Crab pulsar, the observed period derivative is $\dot{P} = 4.2 \times 10^{-13}$ s s^{-1}, which implies a decrease in rotational energy of $\dot{E}_{rot} = -4\pi^2 I\, \dot{P}\, P^{-3} \approx 4.5 \times 10^{38}$ erg s^{-1}. It has been found, though, that the spin-down energy is not distributed homogeneously over the electromagnetic spectrum. In fact, only a fraction of about $(10^{-7} - 10^{-5})\, \dot{E}_{rot}$ is observed in the radio band whereas it is roughly $(10^{-4} - 10^{-3})\, \dot{E}_{rot}$ in the X-ray band and $(10^{-2} - 10^{-1})\, \dot{E}_{rot}$ in the gamma-ray band [9].

  There are two types of rotation-powered pulsars:

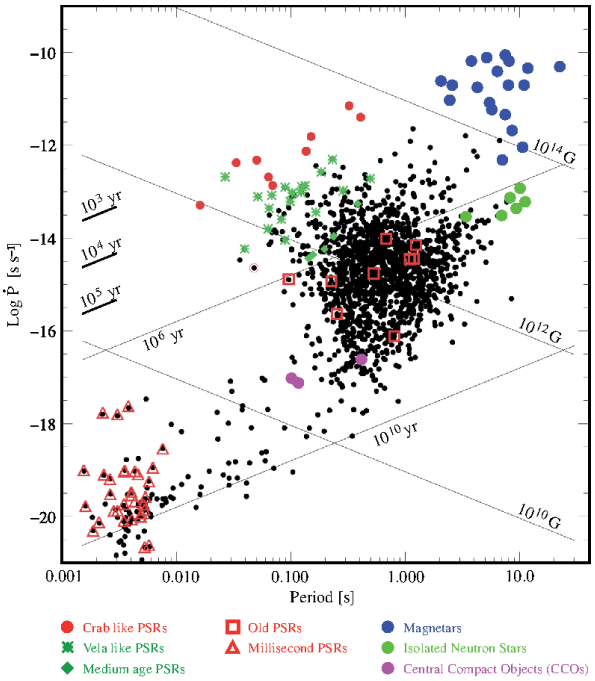


Figure 2. *The P-Ṗ diagram; distribution of rotation-powered pulsars according to their spin parameters. X-ray detected pulsars are indicated by colored symbols. The straight lines correspond to constant ages* $\tau = P/(2\dot{P})$ *and magnetic field strengths* $B_\perp = 3.2 \times 10^{19} (P\dot{P})^{1/2}$ *as deduced within the framework of the magnetic braking model.*

  (1) Field pulsars have periods between tens of milliseconds to several seconds and constitute more than 90% of the total pulsar population. (2) About 10% of the known pulsars are so-called millisecond pulsars, which are defined to have periods below 20 milliseconds. They are much older than normal pulsars, posses weaker magnetic fields and, therefore, relatively low spin-down rates. Accordingly, they exhibit very high timing stabilities, which are comparable to atomic clocks [51, 45]. This property of millisecond pulsars is of major importance for their use in a pulsar-based navigation system. Figure 2 clearly shows that these two types of pulsars belong to distinct populations. Most likely, they are connected by an evolutionary process: It is assumed that millisecond pulsars are born as normal pulsars in a close binary system, but their rotation accelerates as they pass through a phase of accretion in which mass and angular momentum are transferred from the evolving companion star to

DOI: 10.2420/AF07.2013.11

the pulsar (e.g., [19]). However, the fact that millisecond pulsars are often found in binary systems does not affect their suitability for spacecraft navigation as the binary motion can easily be accounted for in pulsar timing [20]. Millisecond pulsars – also referred to as recycled pulsars – were discovered by Backer et al. [2] and studied extensively in the radio band by, e.g., Kramer et al. [38]. Pulsed X-ray emission from millisecond pulsars was discovered by Becker & Trümper [12] using ROSAT. However, only XMM-Newton and Chandra had the sensitivity to study their X-ray emission properties in the $0.5 - 10$ keV band in greater detail. The quality of data from millisecond pulsars available in the X-ray data archives, though, is still very inhomogeneous. While from several of them high quality spectral, temporal and spatial information is available, many others, especially those located in globular clusters, are detected with just a handful of events, not allowing, e.g., to constrain their timing and spectral properties in greater detail. From those millisecond pulsars detected with a high signal-to-noise ratio strong evidence is found for a dichotomy of their X-ray emission properties. Millisecond pulsars having a spin-down energy of $\dot{E} \geqslant 10^{35}$ erg/s (e.g., PSR J0218+4232, B1821−24 and B1937+21) show X-ray emission dominated by non-thermal radiation processes. Their pulse profiles show narrow peaks and pulsed fractions close to 100% (cf. Figure 3). Common for these pulsars is that they show relatively hard X-ray emission, making it possible to study some of them even with RXTE. For example, emission from B1821−24 in the globular cluster M28 is detected by RXTE up to ≈ 20 keV, albeit with limited photon statistics. For the remaining millisecond pulsars the X-ray emission is found to be much softer, and pulse profiles are more sinusoidal. Their typical fraction of pulsed X-ray photons is between 30 and 60%.

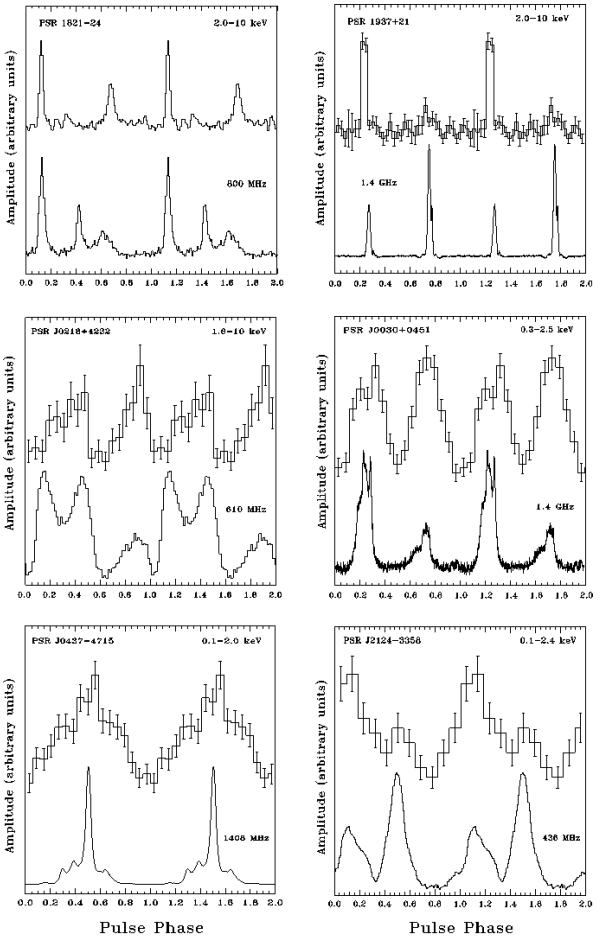


Figure 3. *X-ray and radio pulse profiles for the six brightest millisecond pulsars. Two full pulse cycles are shown for clarity. From Becker [9].*

Some rotation-powered pulsars have shown glitches in their spin-down behavior, i.e., abrupt increases of rotation frequency, often followed by an exponential relaxation toward the pre-glitch frequency [31, 58]. This is often observed in young pulsars but very rarely in old and millisecond pulsars. Nevertheless, the glitch behavior of pulsars should be taken into account by a pulsar-based navigation system.

Today, about 2200 rotation-powered pulsars are known [44]. About 150 have been detected in the X-ray band [9], and approximately 1/3 of them are millisecond pulsars. In the past $30 - 40$ years many of them have been regularly timed with high precision especially in radio observations. Consequently, their ephemerides (RA, DEC, P, $\dot{P}$, binary orbit parameters, pulse arrival time and absolute pulse phase for a given epoch, pulsar proper motion etc.) are known with very high accuracy. Indeed, pulsar timing has reached the 10^{-15} fractional level, which is comparable with the accuracy of atomic clocks. This is an essential requirement for using these celestial objects as navigation beacons, as it enables one to predict the pulse arrival time of a pulsar for any location in the solar system and beyond.

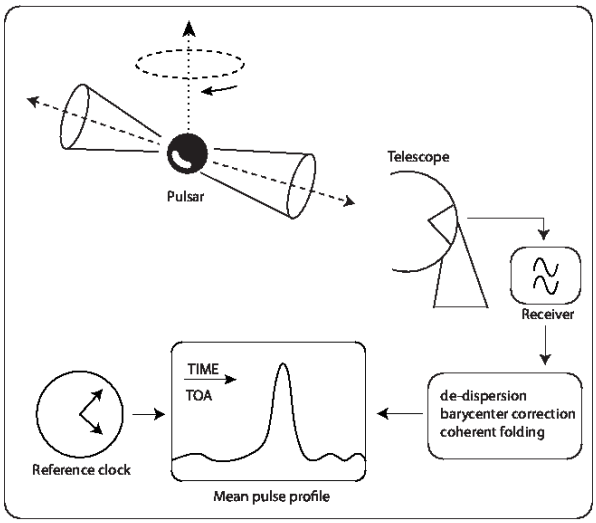


FIGURE 4. *Typical pulsar detection chain. The pulsar beams sweep across the radio antenna. Radio signals are recorded and analyzed in order to produce a mean pulse profile. The data processing comprises a removal of dispersion effects caused by the interstellar medium ("de-dispersion"), correction for the position and proper motion of the observatory ("barycenter correction") and coherent folding of many pulses. The time of arrival (TOA) of the pulse peak is measured against a reference clock.*

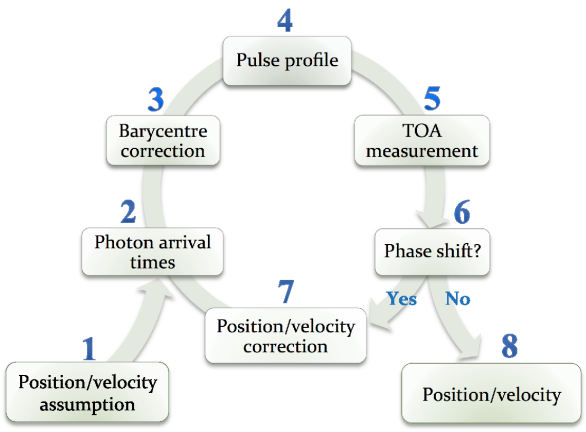


FIGURE 5. *Iterative determination of position and velocity by a pulsar-based navigation system.*

3 Principles of Pulsar-Based Navigation

The concept of using pulsars as navigational aids is based on measurements of pulse arrival times and comparison with predicted arrival times at a given epoch and reference location. A typical chain for detecting, e.g., radio signals from a rotation-powered pulsar is shown in Figure 4.

An important step in this measurement is the barycenter correction of the observed photon arrival times. The pulsar ephemerides along with the position and velocity of the observer are parameters of this correction. Using a spacecraft position that deviates from the true position during the observation results in a phase shift of the pulse peak (or equivalently in a difference in the pulse arrival time). Therefore, the position and velocity of the spacecraft can be adjusted in an iterative process until the pulse arrival time matches with the expected one. The corresponding iteration chain is shown in Figure 5.

An initial assumption of position and velocity is given by the planned orbit parameters of the spacecraft (1). The iteration starts with a pulsar observation, during which the arrival times of individual photons are recorded (2). The photon arrival times have to be corrected for the proper motion of the spacecraft by transforming the arrival times (3) to an inertial reference location; e.g., the solar system barycenter (SSB). This correction requires knowledge of the (assumed or deduced) spacecraft position and velocity as input parameters. The barycenter corrected photon arrival times allow then the construction of a pulse profile or pulse phase histogram (4) representing the temporal emission characteristics and timing signature of the pulsar. This pulse profile, which is continuously improving in significance during an observation, is permanently correlated with a pulse profile template in order to increase the accuracy of the absolute pulse-phase measurement (5), or equivalently, pulse arrival time (TOA). From the pulsar ephemeris that includes the information of the absolute pulse phase for a given epoch, the phase difference $\Delta\phi$ between the measured and predicted pulse phase can be determined (cf. Figure 6).

In this scheme, a phase shift (6) with respect to the absolute pulse phase corresponds to a range difference $\Delta x = cP(\Delta\phi + n)$ along the line of sight toward the observed pulsar. Here, c is the speed of light, P the pulse period, $\Delta\phi$ the phase shift and $n = 0, \pm1, \pm2, \ldots$ an integer that takes into account the periodicity of the observed pulses. If the phase shift is non-zero, the position and velocity of the spacecraft needs to be corrected accordingly and the next iteration step is taken (7). If the phase shift is zero, or falls below a certain threshold, the position and velocity used during the barycenter correction was correct (8) and corresponds to the actual orbit of the spacecraft.

 DOI: 10.2420/AF07.2013.11

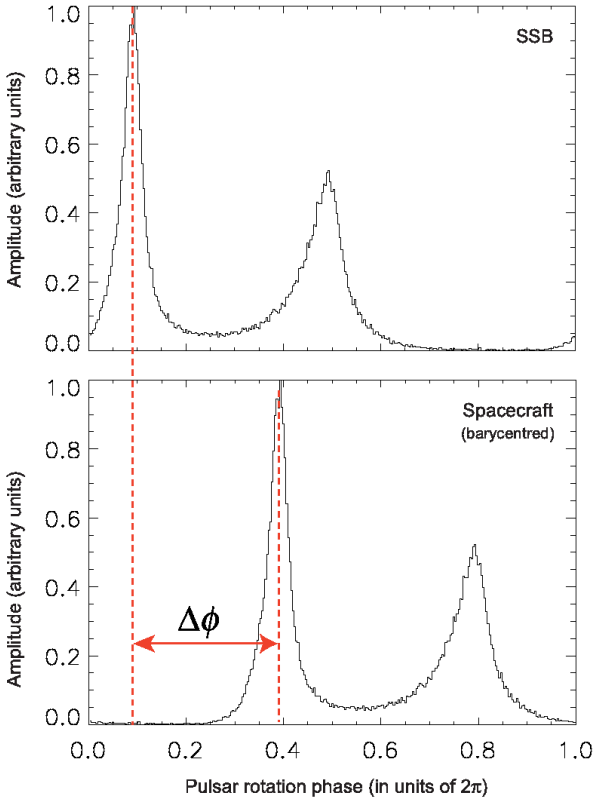


Figure 6. *Measuring the phase difference between the expected and measured pulse peak at an inertial reference location; e.g., the solar system barycenter (SSB). The top profile shows the main pulse peak location as expected at the SSB. The bottom profile is the one which has been measured at the spacecraft and transformed to the SSB by* assuming *the spacecraft position and velocity during the observation. If the position and velocity assumption was wrong, a phase shift* $\Delta\phi$ *is observed.*

A three-dimensional position fix can be derived from observations of at least three different pulsars (cf. Figure 7). If on-board clock calibration is necessary, the observation of a fourth pulsar is required.

Since the position of the spacecraft is deduced from the phase (or pulse arrival time) of a periodic signal, ambiguous solutions may occur. This problem can be solved by constraining the domain of possible solutions to a finite volume around an initial position assumption [18, 17], or by observing additional pulsars as illustrated in Figure 7.

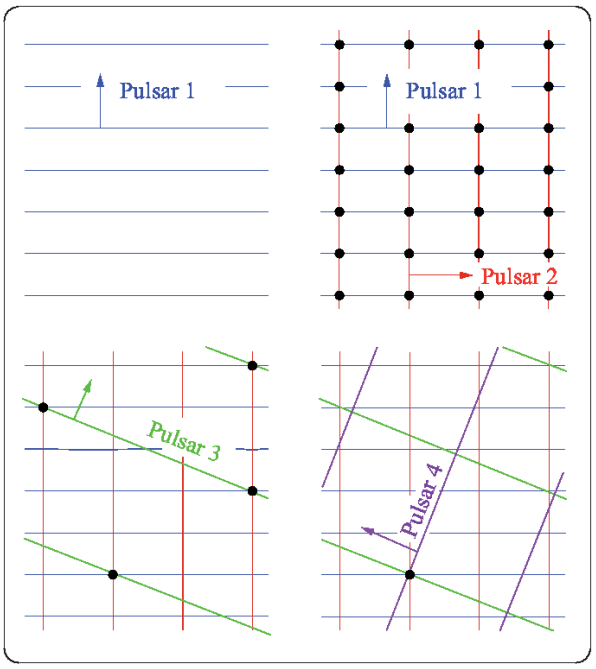


Figure 7. *Solving the ambiguity problem by observing four pulsars (drawn in two dimensions). The arrows point along the pulsar's lines-of-sight. Straight lines represent planes of constant pulse phase; black dots indicate intersections of planes.*

4 Radio Antenna for Pulsar-Based Navigation

Pulsars emit broadband electromagnetic radiation. Therefore, the observing device of a pulsar-based navigator can be optimized for the waveband according to the highest number of bits per telescope collecting area, power consumption, navigator weight, cost and compactness. Especially for the question of navigator compactness it is important to estimate what size a radio antenna would have to have in order to detect the emission from pulsars in a reasonable integration time. In order to estimate this we assume pulsar parameters that are typical for millisecond pulsars. As the radio flux from pulsars shows a $\nu^{-1.5}$ dependence, observations at lower frequencies seem to be preferred, but scintillation and scattering effects are stronger at lower frequency. For a navigation system operating in the radio band of the electromagnetic spectrum, the L-band at 21 cm might therefore be best suited.

For a pulsar detection we require a signal-to-noise ratio of $S/N = 10$, a minimum integration time of $t_{int} = 3600\,s$ and assume a frequency bandwidth of

$\Delta\nu = 100\,\mathrm{MHz}$. For the receiver noise temperature we take $T_{rec} = 100\,\mathrm{K}$. A lower temperature would require active, e.g., cryogenic cooling, which would increase cost, weight, and power consumption of a navigator. Furthermore, active cooling would severely limit the lifetime of the navigator due to consumables like helium. For the sky temperature we take $T_{sky} = 5\,\mathrm{K}$ and for the telescope efficiency $\epsilon = 0.5$. If A_{ant} is the geometrical antenna area, the effective antenna area computes as $A_{eff} = \epsilon\, A_{ant}$. For the period of the pulsar we assume $P = 10\,\mathrm{ms}$ and for the pulse width $W = 2\,\mathrm{ms}$. For the average flux density we adopt $\Delta S = 10\,\mathrm{mJy}$. Using the canonical sensitivity equation [42] which corresponds to the radiometer equation applied to pulsar observations

$$\Delta S_{min} = \frac{2k}{\epsilon A_{ant}} \frac{(T_{rec} + T_{sky})}{\sqrt{n_p\, t_{int}\, \Delta\nu}} \sqrt{\frac{W}{P - W}}$$

and converting it to the geometrical antenna area, including the S/N requirement, we get:

$$A_{ant} = \frac{S}{N} \frac{2k}{\epsilon \Delta S} \frac{(T_{rec} + T_{sky})}{\sqrt{2\, t_{int}\, \Delta\nu}} \sqrt{\frac{W}{P - W}}$$

Here we have also assumed that both polarizations are averaged ($n_p = 2$). From this we compute an antenna area of $A_{ant} \approx 342\,\mathrm{m}^2$ for the parameters specified above. For a parabolic antenna it would mean a radius of about $\sqrt{A_{ant}/\pi} \approx 10.5\,\mathrm{m}$. Increasing the integration time to 4 hours, we find an area of $\approx 171\,\mathrm{m}^2$, which corresponds to a radius of $\approx 7.3\,\mathrm{m}$. For comparison, the radius of the communication antenna used on Cassini and Voyager is 2 m. It may depend on the satellite platform what size and antenna weight is acceptable, but a parabolic antenna does not seem to be very practical for navigation purposes. A navigator would have to observe several pulsars, either at the same time or in series. The pulsars must be located in different sky regions in order to get an accurate navigation result in the x, y and z direction. This, however, means that one either rotates the parabolic antenna or, alternatively, the whole satellite to get the pulsar signals into the antenna focus. Rotating the satellite, though, would mean that the communication antenna will not point to earth any more, which is undesirable. On earth satellites and space missions to the inner solar system, power is usually generated by solar panels. Rotating the satellite would then also mean to bring the solar panels out of optimal alignment with the sun, which is another counter argument for using a parabolic antenna, not to mention the effects of shadowing of the solar panels.

It thus seems more reasonable to use dipole-array antennas for the pulsar observations. Single dipole antennas organized, e.g., as antenna patches could be used to build a larger phased array antenna. Such a phased array would still be large and heavy, though. Depending on the frequency, 10^4–10^5 single patches are required. There have been no phased-array antennas of that size been build for use in space so far, although smaller prototypes exist [26]. From them one may estimate the weight of such antenna arrays. Assuming an antenna thickness of 1 cm and an averaged density of the antenna material of $0.1\,\mathrm{g/cm}^3$ still yields an antenna weight of 170 kg for the $170\,\mathrm{m}^2$ patched antenna array. The signals from the single dipole antennas have to be correlated in phase to each other, which means that all patches have to be connected to each other by a wired mesh and phase correlators. If this phase correlation and the real-time coherent correction for the pulse broadening by interstellar dispersion is done by software, it requires a computer with a Terra-flop GPU of about 500 W power consumption. A clear advantage of a phased antenna array would be that it allows to observe different pulsars located in different sky regions at the same time. That means of course that such an antenna can be smaller if the same S/N ratio is to be achieved for a number $N_{sources}$ of sources within a given time. With a single-dish antenna one would have to increase its diameter by $N_{sources}^{1/4}$ if these had to be observed within the same given time interval.

5 Using X-ray Signals from Pulsars for Spacecraft Navigation

The increasing sensitivity of the X-ray observatories ROSAT, RXTE, XMM-Newton and Chandra allowed for the first time to explore in detail the X-ray emission properties of a larger sample of rotation-powered pulsars. The discovery of pulsed X-ray emission from millisecond pulsars [12], the determination of the X-ray efficiency of rotation-powered pulsars [13] as well as discoveries of X-ray emission from various pulsars (e.g., [14, 10, 11]) and their detailed spatial, spectral and timing studies are just a few of many accomplishments worth mentioning in this context (e.g., [46, 39, 55, 28, 8, 35]). With these new results at hand it was only natural to start looking at their applicability to, e.g., spacecraft navigation based on X-ray data from pulsars. The prospects of this application are of even further interest considering that low-mass X-ray mirrors, which are an

 DOI: 10.2420/AF07.2013.11

important requirement for a realistic implementation of such a navigation system, have been developed for future X-ray observatories.

Given the observational and systematical limitations mentioned above it was a question of general interest, which we found not considered with sufficient gravity in the literature, whether it would be feasible to navigate a spaceship on arbitrary orbits by observing X-ray pulsars.

In order to address the feasibility question we first determined the accuracy that can be achieved by a pulsar-based navigation system in view of the still limited information we have today on pulse profiles and absolute pulse phases in the X-ray band. To overcome the limitations introduced by improper fitting functions, undefined pivot points of pulse peaks and phase shifts by an unmodeled energy dependence in the pulsed signal, we constructed pulse profile templates for all pulsars for which pulsed X-ray emission is detected. Where supported by photon statistics, templates were constructed for various energy ranges. These energy ranges were chosen in order to optimize the S/N ratio of the pulsed signal while sampling as much as possible of the energy dependence of the X-ray pulses.

In the literature various authors applied different analysis methods and often used different definitions for pulsed fraction and pulse peak pivot points. Reanalyzing all data from X-ray pulsars available in the public XMM-Newton, Chandra and RXTE data archives was therefore a requirement to reduce systematic uncertainties that would have been introduced otherwise. The result is a database containing the energy dependent X-ray pulse profiles, templates and relevant timing and spectral properties of all X-ray pulsars that have been detected so far [47, 21].

According to the harmonic content of an X-ray pulse the templates were obtained by fitting the observed pulse profiles by series of Gaussian and sinusoidal functions. The database further includes information on the local environment of a pulsar, i.e., whether it is surrounded by a plerion, supernova remnant or whether it is located in a crowded sky region like a globular cluster. The latter has a severe impact on the detectability of the X-ray pulses as it reduces the S/N ratio of the pulsed emission by the DC emission from background sources. This in turn is important for the selection of the optimal pulsars that emit pulses, e.g., in the hard band (above $\approx 3\,\mathrm{keV}$) in order to blend away the softer emission from a supernova remnant or plerion.

The pulse profile templates allow us to measure pulse arrival times with high accuracy even for sparse photon statistics by using a least-square fit of an adequately adjusted template. The error of pulse-arrival-time measurements is dominated by the systematic uncertainty that comes with the limited temporal resolution of the observed pulse profiles used to construct the templates. The statistical error in fitting a measured profile by a template was found to be much smaller in all cases. Assuming that the temporal resolution of the detector is not the limiting factor, the temporal resolution of a pulse profile is given by the widths of the phase bins used to represent the observed X-ray pulse. The bin width, or the number of phase bins applied, is a compromise between maximizing the S/N ratio per phase bin while sampling as much of the harmonic content as possible. Denoting the Fourier-power of the i-th harmonic by R_i and taking m as the optimal number of harmonics deduced from the H-test [27], an exact expression for the optimal number of phase bins is given by $M = 2.36(\sum_{i=1}^{m} i^2 R_i^2)^{1/3}$ [14]. This formula compromises between information lost due to binning (i.e., zero bin width to get all information), and the effect of fluctuations due to finite statistics per bin (i.e., bin width as large as possible to reduce the statistical error per bin). The total error (bias plus variance) is minimized at a bin width of $1/M$. We applied this in the reanalysis of pulse profiles in our database. Pulsed fractions were computed by applying a bootstrap method [14], which again leads to results that are not biased by the observers "taste" on where to assume the DC level in a profile.

The minimal systematic phase uncertainty for the pulse profile templates in our database is of the order of 0.001 [18]. This uncertainty multiplied by the rotation period P of the pulsar yields the uncertainty in pulse arrival time due to the limited information we have on the exact X-ray pulse profile. Multiplying this in turn by the speed of light yields the spacecraft's position error along the line of sight to the pulsar. It is evident by the linear dependence on P that millisecond pulsars are better suited for navigation than those with larger rotation periods.

The precision of a pulsar-based navigation system thus strongly depends on the choice of pulsars and the accuracy of pulse arrival measurements, which is subject to the quality of the available templates, accuracy of the on-board clock and clock calibration. As mentioned above, in order to obtain three-dimensional position information, timing of at least three different pulsars has to be performed. The spatial arrangement of these pulsars is another parameter of the achievable ac-

curacy. Our simulations show that the systematic error of position determination can be reduced significantly by choosing a pulsar triple that is optimal in the sense that the pulsars are nearly perpendicular to each other. Since a pulsar might be obscured by the sun or a planet and, therefore, its availability for navigation depends on the current position of the spacecraft, the optimal pulsar triple has to be selected from a ranking of possible pulsar combinations. The following Table 1 represents the ranking of pulsar combinations, which according to our analysis provide the highest position accuracy via pulsar navigation [18].

Rank	Pulsar 3-Combination		
1	B1937+21	B1821−24	J0030+0451
2	B1937+21	B1821−24	J1023+0038
3	B1821−24	J0030+0451	J0437−4715
4	B1937+21	J1023+0038	J0218+4232
5	B1821−24	J1023+0038	J0437−4715
6	B1937+21	J0030+0451	J0218+4232
7	B1937+21	B1821−24	J0437−4715
8	B1937+21	J0218+4232	J0437−4715
9	B1821−24	J0218+4232	J0437−4715
10	J1023+0038	J0218+4232	J0437−4715

Table 1. *Ranking of pulsar 3-combinations according to the position accuracy achievable when using them in a navigation system based onf X-ray pulsars . All listed sources are solitary millisecond pulsars except J0437−4715, which is in a binary.*

For the pulsars ranked highest in Table 1 we found position errors of about 5 km as a lower limit (cf. Figure 8). The ranking is independent from a specific spacecraft orbit, but was obtained under the assumption that the navigation system is capable of measuring pulse profiles with the same level of detail and accuracy as the ones used in the simulation. Indeed, this is a severe limitation as those pulse profiles where obtained by powerful X-ray observatories like XMM-Newton and Chandra. It is unlikely, due to weight constraints and power limitations, that a navigation system will have similar capacities in terms of collecting power, temporal resolution and angular resolution.

An improved accuracy can be achieved by means of pulse profile templates of better quality. This, in turn, calls for deeper pulsar observations by XMM-Newton or Chandra as long as these observatories are still available for the scientific community. It would be a valuable task worthwhile the observing time, especially as it is unclear what missions will follow these great observatories and whether they will provide detectors with

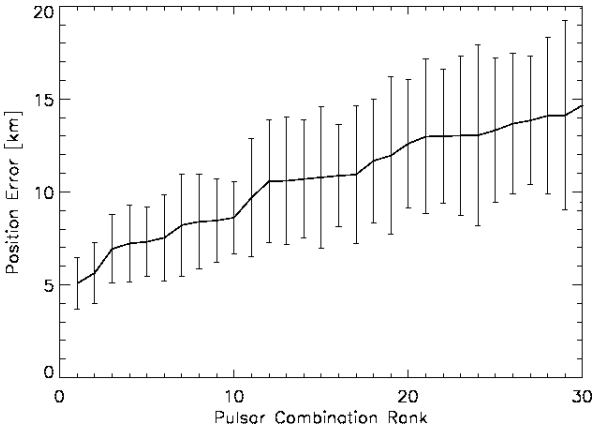


Figure 8. *Spacecraft position error as a function of possible pulsar 3-combinations. The diagram shows the mean position errors and standard deviations for the best 30 combinations. From Bernhardt et al. [18].*

sufficient temporal resolution and on-board clock accuracy.

We extended our simulations in order to constrain the technological parameters of possible navigation systems based on X-ray pulsars. The result of this ongoing project will be a high-level design of a pulsar navigator that accounts for boundary conditions (e.g., weight, cost, complexity and power consumption) set by the requirements of a specific spacecraft and mission design.

The chart shown in Figure 9 illustrates the work logic of our current simulations. Navigator boundary conditions (1) and spacecraft orbit (2) are predefined and constrain the technology parameters (3) of the navigator's X-ray detector, mirror system and on-board electronics. Examples of parameters that will be analyzed in the simulations are detector technology, temporal and energy resolution, on-board-clock accuracy and stability, mirror technology along with collecting power, angular and spatial resolution, focal length, field of view – just to mention the most important ones.

Given their X-ray emission properties some pulsars may not be detectable by the navigator because of its limited angular resolution and sensitivity. The properties of the detector and mirror system along with the trajectory of the virtual navigator thus determine the set of available pulsars (5), from which a suitable selection has to be made (6) according to a predefined ranking of pulsar triples. Using an X-ray-sky simulator (7), which was developed to simulate observations of the future X-ray observatories eROSITA and IXO and which we have

 DOI: 10.2420/AF07.2013.11

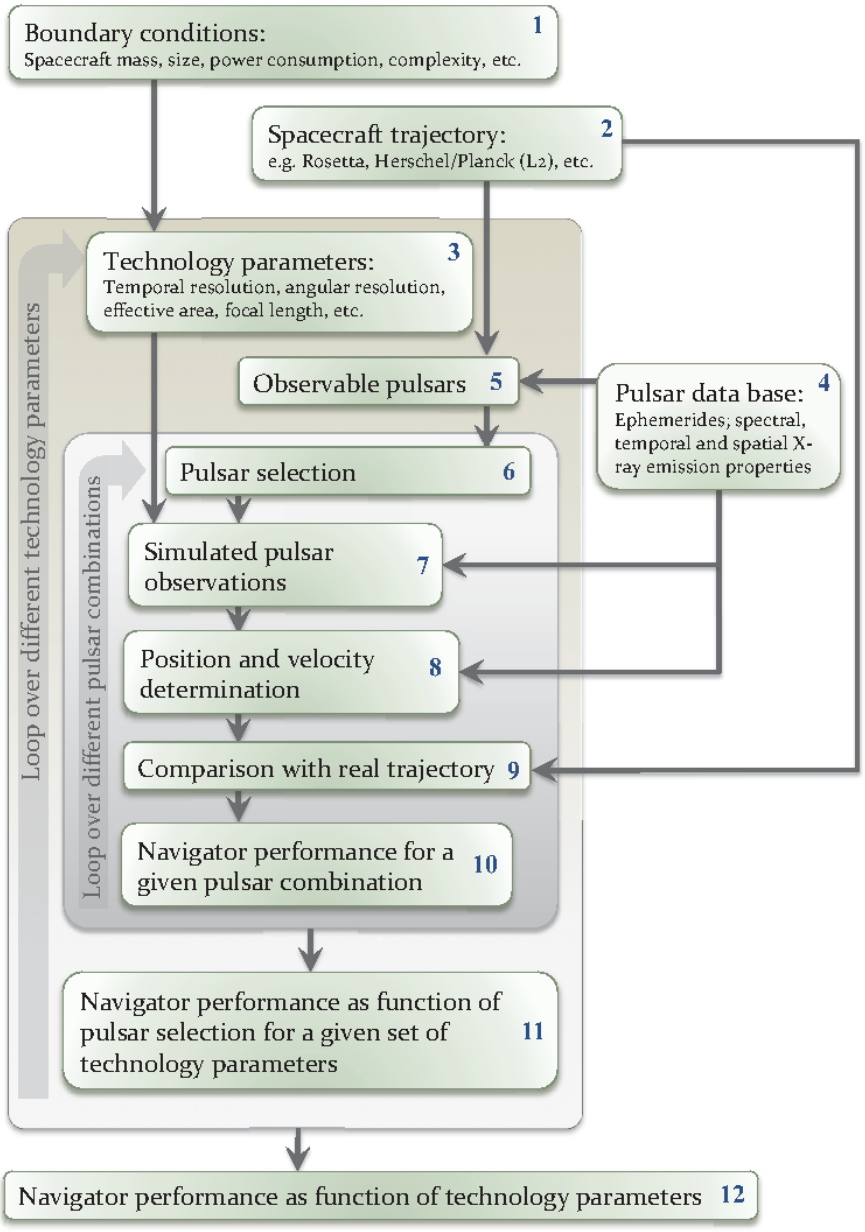


Figure 9. *Work flow and logic of the simulations performed for a technology requirement study and demonstrator high-level design of a pulsar-based navigator.*

modified to include the temporal emission properties of X-ray pulsars, we will be able to create X-ray datasets in FITS-format with temporal, spatial and energy information for the pulsars observed by our virtual navigator. The simulated event files have the same standard FITS-format as those of XMM-Newton and/or Chandra, so that standard software can be used to analyze these data. An autonomous data reduction will then perform the data analysis and TOA measurement in order to obtain the position and velocity of the virtual observer (8). Correlating the result with the input trajectory (9) yields the accuracy of the simulated measurements for a given spacecraft orbit (10), and hence the overall performance of the navigator as a function of the specified detector and mirror system (11,12).

6 X-ray Detector and Mirror Technology for Pulsar-Based Navigation

The design of an X-ray telescope suitable for navigation by X-ray pulsars will be a compromise between angular resolution, collecting area and weight of the system. The currently operating X-ray observatories XMM-Newton and Chandra have huge collecting areas of 0.43 m^2 and 0.08 m^2 (at 1 keV), respectively [33, and references therein] and, in the case of Chandra, attain very good angular resolution of less than 1 arcsecond. Their focusing optics and support structures, however, are very heavy. To use their mirror technology would be a showstopper for a navigation system.

In recent years, ESA and NASA have put tremendous effort into the development of low-mass X-ray mirrors, which can be used as basic technology for future large X-ray observatories and small planetary exploration missions. Table 2 summarizes the angular resolution and mass of X-ray mirrors used for XMM-Newton and Chandra as well as developed for future X-ray missions. The light-weighted mirrors are of special interest for an X-ray pulsar-based navigator.

	Angular resolution	Mass per effective area (at 1 keV)
Chandra	0.5″	18 500 kg/m^2
XMM-Newton	14″	2300 kg/m^2
Silicon Pore Optics	5″	200 kg/m^2
Glass Micropore Optics	30″	25 kg/m^2

Table 2. *Comparison of current and future X-ray-mirror optics. From Bavdaz et al. [5].*

A typical high-resolution X-ray telescope uses focusing optics based on the Wolter-I design [56]. The incoming X-ray photons are reflected under small angles of incidence in order not to be absorbed and are focused by double reflection off a parabolic and then a hyperbolic surface. This geometry allows for nesting several concentric mirror shells into each other in order to enlarge the collecting area and thereby improve the signal-to-noise ratio. A novel approach to X-ray optics is the use of pore structures in a Wolter-I configuration [6, 5, 16]. X-ray photons that enter a pore are focused by reflections on the walls inside the pore. In contrast to traditional X-ray optics with separate mirror shells that are mounted to a support structure, pore optics form a monolithic, self-supporting structure that is lightweight, but also very stiff and contains many reflecting surfaces in a compact assembly. Two different types of pore optics have been developed, based on silicon and glass.

- **Silicon Pore Optics** [24, 25, 1] use commercially available and mass produced silicon wafers (Figure 10a) from the semiconductor industry. These wafers have a surface roughness that is sufficiently low to meet the requirements of X-ray optics. A chemo-mechanical treatment of a wafer results in a very thin membrane with a highly polished surface on one side and thin ribs of very accurate height on the other side. Several of these ribbed plates are elastically bent to the geometry of a Wolter-I system, stacked together to form the pore structure and finally integrated into mirror modules (Figure 10a). Silicon Pore Optics are intended to be used on large X-ray observatories that require a small mass per collecting area (on the order of 200 kg/m^2) and angular resolution of about 5 arcseconds or better.

- **Glass Micropore Optics** [15, 23, 54] are made from polished glass blocks that are surrounded by a cladding glass with a lower melting point. In order to obtain the high surface quality required for X-ray optics, the blocks are stretched into small fibers, thereby reducing the surface roughness. Several of these fibers can be assembled and fused into multi-fiber bundles. Etching away the glass fiber cores leads to the desired micropore structure, in which the remaining cladding glass forms the pore walls (Figure 10b). The Wolter-I geometry is reproduced by thermally slumping separate multi-fiber

 DOI: 10.2420/AF07.2013.11

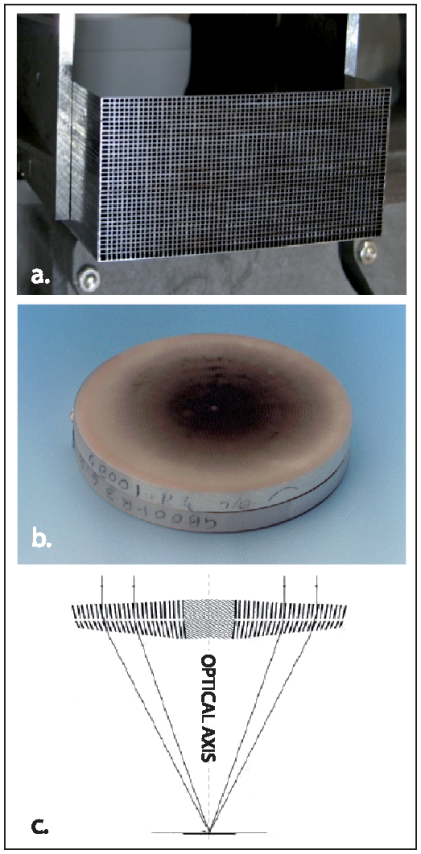


Figure 10. *Silicon pore optics (a) and glass micropore optics (b) represent novel developments for light-weighted X-ray mirrors of the next generation of X-ray observatories. Both mirror types will be used in Wolter-I configuration (c) to focus X-rays in a double reflection. Images from Bavdaz et al. [6, 7, 5].*

plates. Glass Micropore Optics are even lighter than Silicon Pore Optics, but achieve a moderate angular resolution of about 30 arcseconds. They are especially interesting for small planetary exploration missions, but also for X-ray timing missions that require large collecting areas. The first implementation of Glass Micropore Optics on a flight program will be in the Mercury Imaging X-ray Spectrometer on the ESA/JAXA mission Bepi-Colombo, planned to launch in 2014 [32].

Today's detector technology, as in use on XMM-Newton and Chandra, is not seen to provide a detector design that is useful for a navigation system based on X-ray pulsars. Readout noise, limited imaging capability in timing-mode and out-of-time events invalidate CCD-based X-ray detectors for application as X-ray-pulsar navigator. Detectors like those on RXTE that need gas for operation are also not suitable, given the limited live time due to consumables. However, there are novel and promising detector developments performed in semiconductor labs for the use in the next generation of X-ray observatories. Two challenging examples, which are of potential interest for navigation, are:

- **Silicon Drift Detectors (SDDs)** have only limited imaging capability but provide an energy resolution and are capable of managing high counting rates of more than 2×10^6 cts/s. A detector based on this technology was proposed for the High Time Resolution Spectrometer on IXO [3]. The detector technology itself has a high technical readiness. SDD-modules developed in the Semiconductor Lab of the Max Planck Society are working already in the APXS (Alpha Particle X-ray Spectrometer) on-board NASA's Mars Exploration Rovers Spirit and Opportunity and on the comet lander ROSETTA [41]. Detectors based on an SDD technology could be of use, e.g., in designing a navigator for a very specific orbit, for which it is sufficient to navigate according to the signals of pulsars that emit their pulses in the hard X-ray band mostly so that the missing imaging capability does not cause any restrictions on the S/N ratio of the pulsed emission.
- **Active Pixel Sensors (APS)** are an alternative and perhaps more flexible technology (cf. Figure 11), which was the proposed technology for the Wilde Field Imager on IXO [50] and the Low-Energy Detector (LED) on Simbol-X [40]. This detector provides images in the energy band 0.1–25 keV, simultaneously with spectrally and time resolved photon counting. The device, which is under development in the MPE Semiconductor Laboratory, consists of an array of DEPFET (Depleted p-channel FET) active pixels, which are integrated onto a common silicon bulk. The DEPFET concept unifies the functionalities of both sensor and amplifier in one device. It has a signal charge storage capability and is read out demand. The DEPFET is used as unit cell of Active Pixel Sensors (APSs) with a scalable pixel size from 50 µm to several mm and a column-parallel row-by-row readout with a short signal processing time of $\leqslant$ 4 µsec per row. As the pixels are individually addressable the DEPFET APS offers flexible readout

strategies from standard full-frame mode to user-defined window mode.

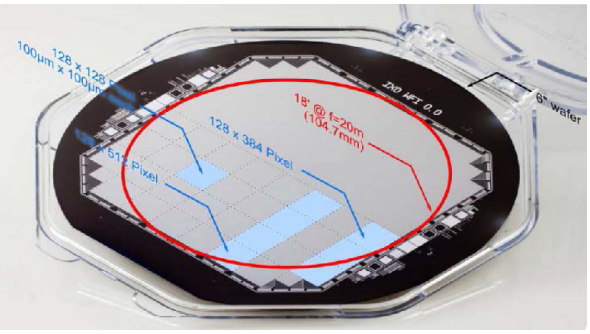


FIGURE 11. *Mechanical sample of an Active Pixel (here 6-inch wafer-scale) detector. Plotted over one hemisphere is the logical layout of the detector. It consists of roughly* 1024×1024 *pixels of* $100 \times 100\ \mu m^2$ *size. From Lechner et al. [41].*

The typical weight and power consumption of these detectors can be estimated from the prototypes proposed for IXO and Simbol-X. The IXO Wide Field Imager with its 17 arcmin field of view and 1024×1024 pixel design had an energy consumption of $\leqslant 22$ W. The Low Energy Detector on Simbol-X had 128×128 pixels and an energy consumption of $\leqslant 8$ W. Power consumption including electronics, filter wheel and temperature control was ca. 250 W. The mass of the focal plane, including shielding and thermal interface, was about 15 kg but could be reduced in a more specific design of an X-ray-pulsar navigator.

7 Concluding Remarks

The knowledge of how to use stars, planets and stellar constellations for navigation was fundamental for mankind in discovering new continents and subduing living space in ancient times. It is fascinating to see how history repeats itself in that a special population of stars may play again a fundamental role in the future of mankind by providing a reference for navigating their spaceships through the Universe (cf. Figure 12).

In the paper we have shown that autonomous spacecraft navigating with pulsars is feasible when using either phased-array radio antennas of at least $150\ m^2$ antenna area or compact light-weighted X-ray telescopes and detectors, which are currently developed for the next generation of X-ray observatories.

Using the X-ray signals from millisecond pulsars we estimated that navigation would be possible with an accuracy of ± 5 km in the solar system and beyond. The error is dominated by the inaccuracy of the pulse profiles templates that were used for the pulse peak fittings and pulse-TOA measurements. As those are known with much higher accuracy in the radio band, it is possible to increase the accuracy of pulsar navigation down to the meter scale by using radio signals from pulsars for navigation.

The disadvantage of radio observations in a navigator, though, is the large size and mass of the phased-antenna array. As we saw in § 4, the antenna area is inversely proportional to the square root of the integration time; i.e., the same signal quality can be obtained with a reduced antenna size by increasing the observation time. However, the observing time is limited by the Allen variance of the receiving system and, therefore, cannot become arbitrarily large. In addition, irradiation from the on-board electronics requires an efficient electromagnetic shielding to prevent signal feedback. This shielding will further increase the navigator weight in addition to the weight of the antenna.

The optimal choice of the observing band depends on the boundary conditions given by a specific mission. What power consumption and what navigator weight might be allowed for may determine the choice for a specific wave band.

In general, however, it is clear already today that this navigation technique will find its applications in future astronautics. The technique behind it is very simple and straightforward, and pulsars are available everywhere in the Galaxy. Today $\approx$ 2200 pulsars are known. With the next generation of radio observatories, like the SKA, it is expected to detect signals from about 20 000 to 30 000 pulsars [49].

Finally, pulsar-based navigation systems can operate autonomously. This is one of their most important advantages, and is interesting also for current space technologies; e.g., as augmentation of existing GPS/Galileo satellites. Future applications of this autonomous navigation technique might be on planetary exploration missions and on manned missions to Mars or beyond.

Acknowledgments

WB acknowledges discussion with David Champion (MPIfR), Horst Baier and Ulrich Walter (TUM). MGB acknowledges support from and participation in

 DOI: 10.2420/AF07.2013.11

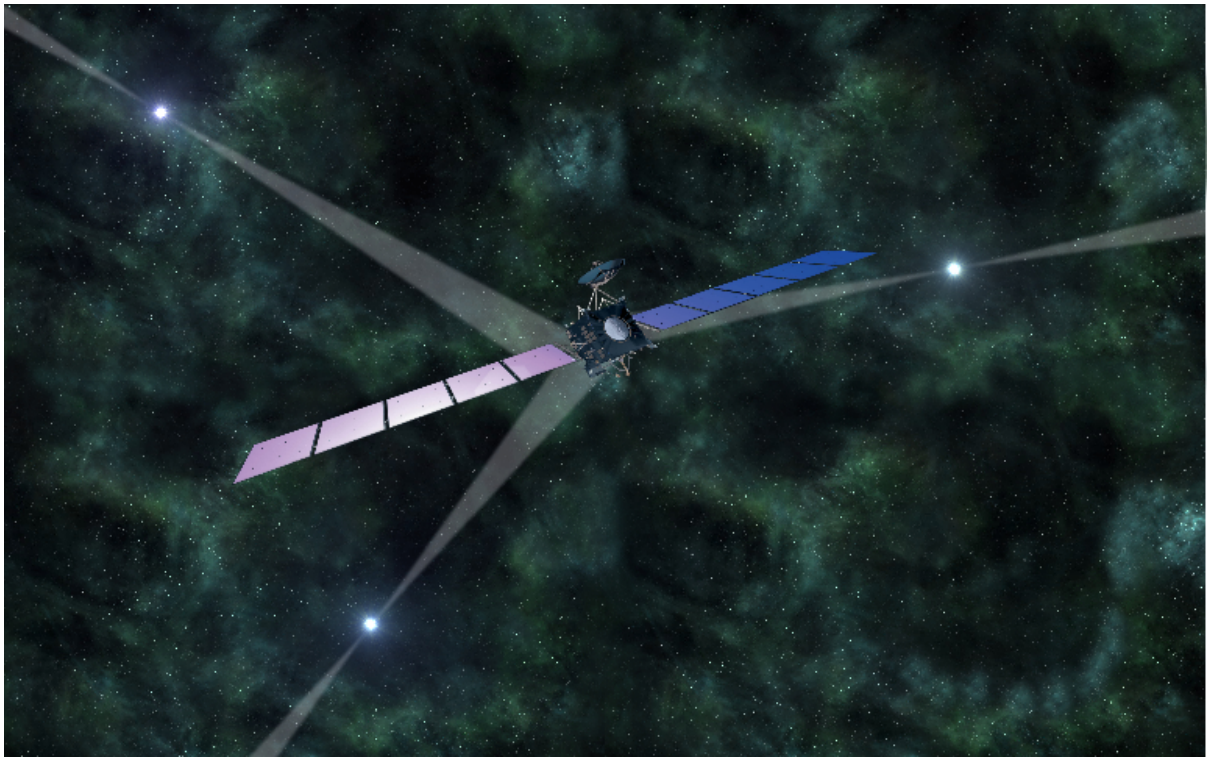


Figure 12. *Artist's impression of Rosetta, if it navigated in deep space using pulsar signals. The characteristic time signatures of pulsars are used as natural navigation beacons to determine the position and velocity of the spacecraft.*

the International Max Planck Research School on Astrophysics at the Ludwig Maximilians University of Munich, Germany.

References

[1] M. D. Ackermann, M. J. Collon, R. Guenther, R. Partapsing, G. Vacanti, E.-J. Buis, M. Krumrey, P. Müller, M. W. Beijersbergen, M. Bavdaz, and K. Wallace. Performance prediction and measurement of Silicon Pore Optics. In *Proc. SPIE*, volume 7437, pages 74371N1–10, 2009.

[2] D. C. Backer, S. R. Kulkarni, C. Heiles, M. M. Davis, and W. M. Goss. A millisecond pulsar. *Nature*, 300:615–618, 1982.

[3] D. Barret *et al.* The High Time Resolution Spectrometer (HTRS) aboard the International X-ray Observatory (IXO). In *SPIE Conference Series*, volume 7732 of *SPIE Conference Series*, 2010.

[4] R. H. Battin. *Astronautical Guidance*. McGraw-Hill, New York, USA, 1964.

[5] M. Bavdaz, M. Collon, M. Beijersbergen, K. Wallace, and E. Wille. X-Ray Pore Optics Technologies and Their Application in Space Telescopes. *X-Ray Optics and Instrumentation*, 2010:1–15, 2010.

[6] M. Bavdaz, A. J. Peacock, V. Lehmann, M. W. Beijersbergen, and S. Kraft. X-ray Optics: new technologies at ESA. In *Proc. SPIE*, volume 4851, pages 421–432, 2003.

[7] M. Bavdaz, A. J. Peacock, E. Tomaselli, M. W. Beijersbergen, M. Collon, S.-O. Flyckt, R. Fairbend, and J.-P. Boutot. Progress at ESA on high-energy optics technologies. In O. Citterio and S. L. O'Dell, editor, *SPIE Conference Series*, volume 5168 of *SPIE Conference Series*, pages 136–147, 2004.

[8] W. Becker, editor. *Neutron Stars and Pulsars*, volume 357 of *Astrophysics and Space Science Library*. Springer, Berlin, Germany, 2009.

[9] W. Becker. X-Ray Emission from Pulsars and Neutron Stars. In W. Becker, editor, *Neutron Stars and Pulsars*, volume 357 of *Astrophysics and Space Science Library*, pages 91–140. Springer, Berlin, Germany, 2009.

[10] W. Becker, A. Jessner, M. Kramer, V. Testa, and C. Howaldt. A Multiwavelength Study of PSR B0628−28: The First Overluminous Rotation-powered Pulsar? *ApJ*, 633:367–376, 2005.

[11] W. Becker, M. Kramer, A. Jessner, R. E. Taam, J. J. Jia, K. S. Cheng, R. Mignani, A. Pellizzoni, A. de Luca, A. Słowikowska, and P. A. Caraveo. A Multiwavelength Study of the Pulsar PSR B1929+10 and Its X-Ray Trail. *ApJ*, 645:1421–1435, 2006.

[12] W. Becker and J. Trümper. Detection of pulsed X-rays from the binary millisecond pulsar J0437−4715. *Nature*, 365:528–530, 1993.

[13] W. Becker and J. Trümper. The X-ray luminosity of rotation-powered neutron stars. *A&A*, 326:682–691, 1997.

[14] W. Becker and J. Trümper. The X-ray emission properties of millisecond pulsars. *A&A*, 341:803–817, 1999.

[15] M. Beijersbergen, M. Bavdaz, E. J. Buis, and D. H. Lumb. Micro-pore X-ray optics developments and application to an X-ray timing mission. In *Proc. SPIE*, volume 5488, pages 468–474, 2004.

[16] M. Beijersbergen, S. Kraft, M. Bavdaz, D. Lumb, R. Guenther, M. Collon, A. Mieremet, R. Fairbend, and A. Peacock. Development of x-ray pore optics: novel high-resolution silicon millipore optics for XEUS and ultralow mass glass micropore optics for imaging and timing. In *Proc. SPIE*, volume 5539, pages 104–115, 2004.

[17] M. G. Bernhardt, W. Becker, T. Prinz, F. M. Breithuth, and U. Walter. Autonomous Spacecraft Navigation Based on Pulsar Timing Information. In *2nd International Conference on Space Technology*, pages 1–4, 2011.

[18] M. G. Bernhardt, T. Prinz, W. Becker, and U. Walter. Timing X-ray Pulsars with Application to Spacecraft Navigation. In *High Time Resolution Astrophysics IV, PoS(HTRA-IV)050*, pages 1–5, 2010.

[19] D. Bhattacharya and E. P. J. van den Heuvel. Formation and evolution of binary and millisecond radio pulsars. *Phys. Rep.*, 203:1–124, 1991.

[20] R. Blandford and S. A. Teukolsky. Arrival-time analysis for a pulsar in a binary system. *ApJ*, 205:580–591, 1976.

[21] F. M. Breithuth. Erstellung einer Pulsardatenbank und ihre Anwendung für die Simulation eines pulsarbasierten Navigationssystems. Master's thesis, Ludwig-Maximilians-Universität München, 2012. (In German).

[22] T. J. Chester and S. A. Butman. Navigation Using X-Ray Pulsars. NASA Tech. Rep. 81N27129, Jet Propulsion Laboratory, Pasadena, CA, USA, 1981.

[23] M. J. Collon, M. W. Beijersbergen, K. Wallace, M. Bavdaz, R. Fairbend, J. Séguy, E. Schyns, M. Krumrey, and M. Freyberg. X-ray imaging glass micro-pore optics. In *Proc. SPIE*, volume 6688, pages 668812/1–13, 2007.

[24] M. J. Collon, R. Guenther, M. Ackermann, R. Partapsing, C. Kelly, M. W. Beijersbergen, M. Bavdaz, K. Wallace, M. Olde Riekerink, P. Mueller, and M. Krumrey. Stacking of Silicon Pore Optics for IXO. In *Proc. SPIE*, volume 7437, pages 74371A1–7, 2009.

[25] M. J. Collon, R. Günther, M. Ackermann, R. Partapsing, G. Vacanti, M. W. Beijersbergen, M. Bavdaz, E. Wille, K. Wallace, M. Olde Riekerink, B. Lansdorp, L. de Vrede, C. van Baren, P. Müller, M. Krumrey, and M. Freyberg. Silicon Pore X-ray Optics for IXO. In *Proc. SPIE*, volume 7732, pages 77321F1–9, 2010.

[26] L. Datashvili, H. Baier, T. Kuhn, H. Langer, S. Apenberg, and B. Wei. A Large Deployable Space Array Antenna: Technology and Functionality Demonstrator. In *Proc. of ESA Antenna Workshop*, 2011.

 DOI: 10.2420/AF07.2013.11

[27] O. C. De Jager, B. C. Raubenheimer, and J. W. H. Swanepoel. A poweful test for weak periodic signals with unknown light curve shape in sparse data. *A&A*, 221:180–190, 1989.

[28] A. De Luca, P. A. Caraveo, S. Mereghetti, M. Negroni, and G. F. Bignami. On the Polar Caps of the Three Musketeers. *ApJ*, 623:1051–1069, 2005.

[29] G. S. Downs. Interplanetary Navigation Using Pulsating Radio Sources. NASA Tech. Rep. 74N34150 (JPL Tech. Rep. 32-1594), Jet Propulsion Laboratory, Pasadena, CA, USA, 1974.

[30] R. C. Duncan and C. Thompson. Formation of very strongly magnetized neutron stars – Implications for gamma-ray bursts. *ApJ*, 392:L9–L13, 1992.

[31] C. M. Espinoza, A. G. Lyne, B. W. Stappers, and M. Kramer. A study of 315 glitches in the rotation of 102 pulsars. *MNRAS*, 414:1679–1704, 2011.

[32] G. W. Fraser *et al.* The mercury imaging X-ray spectrometer (MIXS) on bepicolombo. *Planet. Space Sci.*, 58(1-2):79–95, 2010.

[33] P. Friedrich. Wolter Optics. In J. E. Trümper and G. Hasinger, editors, *The Universe in X-Rays*, Astronomy and Astrophysics Library, pages 41–50. Springer, Berlin, Germany, 2008.

[34] P. Ghosh. *Rotation and Accretion Powered Pulsars*. World Scientific Publishing, New Jersey, USA, 2007.

[35] W. Hermsen *et al.* Synchronous X-ray and Radio Mode Switches: A Rapid Global Transformation of the Pulsar Magnetosphere. *Science*, 339:436–, 2013.

[36] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins. Observation of a Rapidly Pulsating Radio Source. *Nature*, 217:709–713, 1968.

[37] N. James, R. Abello, M. Lanucara, M. Mercolino, and R. Maddè. Implementation of an ESA delta-DOR capability. *Acta Astronautica*, 64(11-12):1041–1049, 2009.

[38] M. Kramer, K. M. Xilouris, D. R. Lorimer, O. Doroshenko, A. Jessner, R. Wielebinski, A. Wolszczan, and F. Camilo. The Characteristics of Millisecond Pulsar Emission. I. Spectra, Pulse Shapes, and the Beaming Fraction. *ApJ*, 501:270, 1998.

[39] L. Kuiper, W. Hermsen, F. Verbunt, S. Ord, I. Stairs, and A. Lyne. High-Resolution Spatial and Timing Observations of Millisecond Pulsar PSR J0218+4232 with Chandra. *ApJ*, 577:917–922, 2002.

[40] P. Lechner. The Simbol-X Low Energy Detector. In J. Rodriguez and P. Ferrando, editor, *American Institute of Physics Conference Series*, volume 1126 of *American Institute of Physics Conference Series*, pages 21–24, 2009.

[41] P. Lechner, C. Amoros, D. Barret, P. Bodin, M. Boutelier, R. Eckhardt, C. Fiorini, E. Kendziorra, K. Lacombe, A. Niculae, B. Pouilloux, R. Pons, D. Rambaud, L. Ravera, C. Schmid, H. Soltau, L. Strüder, C. Tenzer, and J. Wilms. The silicon drift detector for the IXO high-time resolution spectrometer. In *SPIE Conference Series*, volume 7742 of *SPIE Conference Series*, 2010.

[42] D. R. Lorimer and M. Kramer. *Handbook of Pulsar Astronomy*, volume 4 of *Cambridge observing handbooks for research astronomers*. Cambridge University Press, Cambridge, UK, 2005.

[43] R. Maddè, T. Morley, R. Abelló, M. Lanucara, M. Mercolino, G. Sessler, and J. de Vicente. Delta-DOR – a new technique for ESA's Deep Space Navigation. *ESA Bulletin*, 128:68–74, 2006.

[44] R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs. The Australia Telescope National Facility Pulsar Catalogue. *AJ*, 129:1993–2006, 2005.

[45] D. N. Matsakis, J. H. Taylor, and T. M. Eubanks. A statistic for describing pulsar and clock stabilities. *A&A*, 326:924–928, 1997.

[46] G. G. Pavlov, V. E. Zavlin, D. Sanwal, V. Burwitz, and G. P. Garmire. The X-Ray Spectrum of the Vela Pulsar Resolved with the Chandra X-Ray Observatory. *ApJ*, 552:L129–L133, May 2001.

[47] T. Prinz. Zeitanalyse der Röntgenemission von Pulsaren und deren spezielle Anwendung für die Navigation von Raumfahrzeugen. Diploma thesis, Ludwig-Maximilians-Universität München, 2010. (In German).

[48] J. E. Riedel, S. Bhaskaran, S. Desai, D. Han, B. Kennedy, G. W. Null, S. P. Synnott, T. C. Wang, R. A. Werner, and E. B. Zamani. Autonomous Optical Navigation (AutoNav) DS1 Technology Validation Report. Deep Space 1 technology validation reports (Rep. A01-26126 06-12), Jet Propulsion Laboratory, Pasadena, CA, USA, 2000.

[49] R. Smits, M. Kramer, B. Stappers, D. R. Lorimer, J. Cordes, and A. Faulkner. Pulsar searches and timing with the square kilometre array. *A&A*, 493:1161–1170, 2009.

[50] L. Strüder *et al*. The wide-field imager for IXO: status and future activities. In *SPIE Conference Series*, volume 7732 of *SPIE Conference Series*, 2010.

[51] J. H. Taylor, Jr. Millisecond pulsars: Nature's most stable clocks. *IEEE Proceedings*, 79:1054–1062, 1991.

[52] J. E. Trümper, K. Dennerl, N. D. Kylafis, Ü. Ertan, and A. Zezas. An Accretion Model for the Anomalous X-Ray Pulsar 4U 0142+61. *ApJ*, 764:49, 2013.

[53] J. E. Trümper, A. Zezas, Ü. Ertan, and N. D. Kylafis. The energy spectrum of anomalous X-ray pulsars and soft gamma-ray repeaters. *A&A*, 518:A46+, 2010.

[54] K. Wallace, M. J. Collon, M. W. Beijersbergen, S. Oemrawsingh, M. Bavdaz, and E. Schyns. Breadboard micro-pore optic development for x-ray imaging. In *Proc. SPIE*, volume 6688, pages 66881C1–10, 2007.

[55] M. C. Weisskopf, S. L. O'Dell, F. Paerels, R. F. Elsner, W. Becker, A. F. Tennant, and D. A. Swartz. Chandra Phase-Resolved X-Ray Spectroscopy of the Crab Pulsar. *ApJ*, 601:1050–1057, 2004.

[56] H. Wolter. Spiegelsysteme streifenden Einfalls als abbildende Optiken für Röntgenstrahlen. *Ann. Phys.*, 445:94–114, 1952. (In German).

[57] K. S. Wood. Navigation studies utilizing the NRL-801 experiment and the ARGOS satellite. In *Proc. SPIE*, volume 1940, pages 105–116, 1993.

[58] M. Yu, R. N. Manchester, G. Hobbs, S. Johnston, V. M. Kaspi, M. Keith, A. G. Lyne, G. J. Qiao, V. Ravi, J. M. Sarkissian, R. Shannon, and R. X. Xu. Detection of 107 glitches in 36 southern pulsars. *MNRAS*, 429:688–724, 2013.

 DOI: 10.2420/AF07.2013.11

Acta Futura 7 (2013) 29-34
DOI: 10.2420/AF07.2013.29

Acta Futura

Restricted Post-Newtonian two-body problem with spin and its applications

Francesco Biscani*, and Sante Carloni†
ESA – Advanced Concepts Team, European Space Research Technology Center (ESTEC), Keplerlaan 1, Postbus 299, 2200 AG Noordwijk The Netherlands

Abstract. We analyze the relativistic restricted two-body problem with spin employing a Lie series perturbation theory. Taking into full account the complex interplay between the various relativistic effects, we provide a new explicit solution of the averaged equations of motion in terms of elliptic functions. This result is then applied to the case to gravity Probe B recovering the standard results. The possible impact on the development of relativistic positioning systems is briefly discussed.

1 Introduction

One of the most important applications of General Relativity is the understanding of the effect that the curvature of the spacetime has on the motion of celestial objects. The reason behind such statement can be found in the accuracy with which data on the motions of celestial bodies is nowadays acquired. Future space missions hold the promise to ameliorate even further this accuracy. It is natural, then, to strengthen the efforts in the understanding of Post-Newtonian mechanics and to approach the resolution of the many unresolved problems in this field of research. Recently the authors of this contribution used Lie series perturbation theory to perform a reanalysis of the reduced to body problem at 1PN approximation [5] finding an exact solution. The purpose of this paper is to review this result, and discuss the possible future application of these result to Relativistic Positioning Systems.

*E-mail: bluescarni@gmail.com
†Corresponding author. E-mail: sante.carloni@esa.int

2 Hamiltonian formulation

The starting point of our derivation is the well-known 1PN Hamiltonian of the reduced ($m_2 \gg m_1$ and $|\boldsymbol{J}_2| \gg |\boldsymbol{J}_1|$) two-body problem with spin, which, after reduction to the centre-of-mass coordinate system, reads [2, 3, 6]

$$\mathcal{H} = \mathcal{H}_N + \epsilon \mathcal{H}_1. \tag{1}$$

Here $\epsilon = 1/c^2$ is chosen as the "smallness parameter" of our perturbation theory and

$$\mathcal{H}_N = \frac{1}{2}\frac{\boldsymbol{J}_1^2}{I_1} + \frac{1}{2}\frac{\boldsymbol{J}_2^2}{I_2} + \frac{\boldsymbol{p}^2}{2\mu} - \frac{\mathcal{G}M\mu}{r} \tag{2}$$

is the Newtonian Hamiltonian (representing the unperturbed problem). $\mathcal{H}_1$ is defined as

$$\mathcal{H}_1 = \mathcal{H}_{PN} + \mathcal{H}_{SO} + \mathcal{H}_{SS}, \tag{3}$$

where

$$\begin{aligned}
\mathcal{H}_N &= \frac{1}{2}\frac{\boldsymbol{J}_1^2}{I_1} + \frac{\boldsymbol{p}_1^2}{2m_1} - \frac{\mathcal{G}m_1m_2}{r},\\
\mathcal{H}_{PN} &= m_1\left(-\frac{1}{8}\frac{\boldsymbol{p}_1^4}{m_1^4} - \frac{3}{2}\frac{\mathcal{G}m_2}{r}\frac{\boldsymbol{p}_1^2}{m_1^2} + \frac{\mathcal{G}^2m_2^2}{2r^2}\right),\\
\mathcal{H}_{SO} &= \frac{2\mathcal{G}}{r^3}\left(\frac{3}{4}\frac{m_2}{m_1}\boldsymbol{J}_1 + \boldsymbol{J}_2\right)\cdot(\boldsymbol{r}\times\boldsymbol{p}_1),\\
\mathcal{H}_{SS} &= \frac{\mathcal{G}}{r^3}\left[3\left(\boldsymbol{J}_1\cdot\boldsymbol{n}\right)\left(\boldsymbol{J}_2\cdot\boldsymbol{n}\right) - \boldsymbol{J}_1\cdot\boldsymbol{J}_2\right],
\end{aligned} \tag{4}$$

where $\boldsymbol{J}_2$ is now considered as a constant of motion that can be dropped from $\mathcal{H}_N$. In our approximation the origin of the coordinate system coincides with the body 2. Without loss of generality, we can orient the reference system in such a way that the constant $\boldsymbol{J}_2$ is aligned to the positive z axis, so that

$$\boldsymbol{J}_2 = (0, 0, J_2), \tag{5}$$

with $J_2 = |\boldsymbol{J}_2|$.

Let us now express the Hamiltonian in a coordinate system of action-angle variables [1] for the unperturbed problem. We choose the Delaunay elements for the orbital variable whereas the rotational degrees of freedom will be represented by the Serret-Andoyer (SA) variables [10]. Both sets of variables are introduced formally as canonical transformations[1].

The Delaunay arguments (L, G, H, l, g, h) can be introduced via the following standard relations [14]:

$$\begin{aligned}
L &= \sqrt{\mathcal{G}m_2 a}, \qquad & l &= M,\\
G &= L\sqrt{1-e^2}, & g &= \omega,\\
H &= G\cos i, & h &= \Omega.
\end{aligned} \tag{6}$$

Here a, e, i, M, ω and Ω are the classical Keplerian orbital elements describing the trajectory of the secondary body m_1 around the primary m_2. The Keplerian elements are in turn related to the cartesian orbital momentum $\boldsymbol{p}_1$ and position $\boldsymbol{r}$ via well-known relations (e.g., see [15]). In eqs. (6), (L, G, H) play the role of generalised momenta, (l, g, h) are the generalised coordinates.

[1]Note that the post-Newtonian canonical Hamiltonian momenta will differ from the Newtonian ones by terms of order $1/c^2$. This discrepancy will carry over to any subsequent canonical transformation, including the introduction of Delaunay and SA elements. The detailed connection between Newtonian and post-Newtonian Delaunay orbital elements can be found in[17] and [11]. However, in the present work we are concerned with the secular variations of orbital and rotational elements, for which the discrepancy discussed above is of little consequence.

The Serret-Andoyer (SA) [10] variables describe the rotational motion of a rigid body in terms of orientation angles and rotational angular momentum. In our specific case, these variables are obtained in terms of $\boldsymbol{J}_1$, so that the generalized momenta are

$$\begin{aligned}
\tilde{G} &= |\boldsymbol{J}_1|/m_1\\
\tilde{H} &= \tilde{G}\cos I
\end{aligned} \tag{7}$$

where $\cos I = J_{1,z}/|\boldsymbol{J}_1|$ ($J_{1,z}$ is the z component of the rotational angular momentum in the centre-of-mass reference system) and, of the conjugate coordinates $(\tilde{g}, \tilde{h})$, $\tilde{h}$ is nodal angle of spin $\boldsymbol{J}_1$, the angle $\tilde{g}$ conjugated to the $\tilde{G}$. Note that $\tilde{g}$ does not appear in the expression of the Hamiltonian and $|\boldsymbol{J}_1|$ is thus conserved.

We can now express the Hamiltonians (4) in terms of Delaunay and SA arguments, dividing the Hamiltonian and the generalised momenta by m_1 (extended canonical transformation).

The resulting expressions for $\mathcal{H}_N$ and $\mathcal{H}_1$ in terms of the momenta $(L, G, H, \tilde{G}, \tilde{H})$ and coordinates $(l, g, h, \tilde{g}, \tilde{h})$ are given in [5]. With this expression we can analyse the Hamiltonian using the Lie series method.

3 Lie Series analysis

The Lie series perturbative methodology [12, 7] aims at simplifying the Hamiltonian of the problem via a quasi-identity canonical transformation of coordinates depending on a generating function χ to be determined. The Lie series transformation reads

$$\mathcal{H}' = \mathcal{S}_\chi^\epsilon\mathcal{H} = \sum_{n=0}^{\infty}\frac{\epsilon^n}{n!}\mathcal{L}_\chi^n\mathcal{H}, \tag{8}$$

where $\mathcal{L}_\chi^n$ is the Lie derivative of n-th order with generator χ, $\mathcal{H}'$ is the transformed Hamiltonian, and $\mathcal{H}$ is the original Hamiltonian in which the new momenta and coordinates have been formally substituted. At the first order in ϵ the Lie derivative degenerates to a Poisson bracket, and the transformation becomes

$$\mathcal{H}' = \mathcal{H}_N + \epsilon\underbrace{\left(\{\mathcal{H}_N, \chi\} + \mathcal{H}_1\right)}_{\mathcal{K}} + O\left(\epsilon^2\right). \tag{9}$$

We need then to solve the homological equation [1]

$$\{\mathcal{H}_N, \chi\} + \mathcal{H}_1 = \mathcal{K}, \tag{10}$$

 DOI: 10.2420/AF07.2013.29

where χ and $\mathcal{K}$ (the new perturbed Hamiltonian resulting from the transformation of coordinates) have to be determined with the goal of obtaining some form of simplification in $\mathcal{K}$. Since the unperturbed Hamiltonian depends only on the two actions L and $\tilde{G}$, we have

$$\{\mathcal{H}_N, \chi\} = -\frac{\partial \mathcal{H}_N}{\partial L}\frac{\partial \chi}{\partial l} = -\frac{\mathcal{G}^2 m_2^2}{L^3}\frac{\partial \chi}{\partial l}, \qquad (11)$$

where the partial derivative of χ with respect to $\tilde{g}$ (the coordinate conjugated to $\tilde{G}$) can be set to zero as $\tilde{G}$ is a constant of motion. The homological equation (10) then reads

$$\chi = \int \frac{L^3}{\mathcal{G}^2 m_2^2}\left(\mathcal{H}_1 - \mathcal{K}\right) dl. \qquad (12)$$

The technical aspects of the solution of this integral, including the choice of $\mathcal{K}$, are detailed in [5]. Here we will limit ourselves to the following considerations:

- the integral in eq. (12) essentially represents an averaging over the mean motion l. Consequently, the new momenta and coordinates generated by the Lie series transformation are the *mean* counterparts of the original momenta and coordinates;
- the functional form of Hamiltonians in terms of Delaunay and SA variables guarantees that the averaging procedure removes at the same time both l and g from the averaged Hamiltonian;
- the integration is performed in closed form, that is, without resorting to Fourier-Taylor expansions in terms of mean anomaly and eccentricity [8, 16]. The results are thus valid also for highly-eccentric orbits.

The computations involved in the averaging procedure have been carried out with the Piranha computer algebra system [4]. As usual when operating with Lie series transformations, from now on we will refer to the mean momenta and coordinates $\left(L', G', H', \tilde{G}', \tilde{H}'\right)$ and $\left(l', g', h', \tilde{g}', \tilde{h}'\right)$ with their original names $\left(L, G, H, \tilde{G}, \tilde{H}\right)$ and $\left(l, g, h, \tilde{g}, \tilde{h}\right)$, in order to simplify the notation.

After having determined χ from eq. (12), the averaged Hamiltonian generated by eq. (10) reads, in terms of mean elements,

$$\mathcal{H}' = \mathcal{H}_N + \epsilon\left[\mathcal{E}_0 + \mathcal{E}_1\cos\left(\tilde{h} - h\right)\right], \qquad (13)$$

where $\mathcal{E}_0$ and $\mathcal{E}_1$ are given in [5].

Applying the transformation

$$\tilde{H}_* = H + \tilde{H}, \qquad (14)$$

$$h_* = h - \tilde{h}, \qquad (15)$$

we can further reduce the degrees of freedom so that the final averaged Hamiltonian reads

$$\mathcal{H}' = \mathcal{H}_N + \epsilon\left(\mathcal{F}_0 + \mathcal{F}_1\cos h_*\right), \qquad (16)$$

where $\mathcal{F}_0$ and $\mathcal{F}_1$ are functions of the mean momenta only,

$$\begin{aligned}\mathcal{F}_0 = {} & \frac{1}{2}\frac{J_2\mathcal{G}^4\tilde{H}_* {m_2}^3}{G^3L^3} + \frac{3}{2}\frac{H^3J_2\mathcal{G}^4{m_2}^3}{G^5L^3} \\ & + \frac{15}{8}\frac{\mathcal{G}^4{m_2}^4}{L^4} + \frac{3}{2}\frac{HJ_2\mathcal{G}^4{m_2}^3}{G^3L^3} \\ & - \frac{3}{2}\frac{H^2J_2\mathcal{G}^4\tilde{H}_*{m_2}^3}{G^5L^3} - \frac{3}{2}\frac{H^2\mathcal{G}^4{m_2}^4}{G^3L^3} \\ & + \frac{3}{2}\frac{H\mathcal{G}^4\tilde{H}_*{m_2}^4}{G^3L^3} - 3\frac{\mathcal{G}^4{m_2}^4}{GL^3}, \qquad (17)\end{aligned}$$

$$\begin{aligned}\mathcal{F}_1 = {} & -\frac{3}{2}\frac{G_{xy}HJ_2\mathcal{G}^4\tilde{G}_{xy*}{m_2}^3}{G^5L^3} \\ & + \frac{3}{2}\frac{G_{xy}\mathcal{G}^4\tilde{G}_{xy*}{m_2}^4}{G^3L^3}, \qquad (18)\end{aligned}$$

and $\tilde{G}_{xy*}$ is expressed in terms of the new mean momentum $\tilde{H}_*$:

$$\tilde{G}_{xy*} = \sqrt{\tilde{G}^2 - \left(\tilde{H}_* - H\right)^2}. \qquad (19)$$

The one degree-of-freedom averaged Hamiltonian (16) is expressed in terms of the mean momenta $\left(L, G, H, \tilde{G}, \tilde{H}_*\right)$ and the conjugate mean coordinate h_*. Therefore the form equations of motion (see [5] for their detailed form) implies the conservation of all mean momenta apart from H. In particular, the conservation of the mean momentum $\tilde{H}_* = H + \tilde{H}$ corresponds to the conservation of the z component of the total mean angular momentum of the system.

Using the Hamiltonian (16) it is possible to describe as subcases all the classical relativistic effect on the motion of the reduced two body problem like, for example, the Lens Thirring effect of the geodesic effect. A detailed derivation of these effect can be found in [5].

4 An exact solution for $H(t)$

We now derive an analytical solution for the time variation of H. The availability of a closed-form solution for

$H(t)$ allows to obtain immediately the time evolution of $\cos h_*$ via inversion of the Hamiltonian (16)[2]. With $H(t)$ and $\cos h_*(t)$ it is then possible in principle to integrate the equations of motion for the remaining coordinates. Additionally, the analytical expression will allow us to calculate exactly the period of the oscillatory motion of $H(t)$ and to make quantitative predictions about the behaviour of real physical systems in §5.

Recalling the form of the Hamiltonian (16) and combining it with the equation for H,

$$\frac{dH}{dt} = \epsilon \mathcal{F}_1 \sin h_*, \tag{20}$$

we have

$$\frac{dH}{dt} = \pm\sqrt{\epsilon^2\mathcal{F}_1^2 - (\mathcal{H}' - \mathcal{H}_N - \epsilon\mathcal{F}_0)^2}, \tag{21}$$

where $\mathcal{H}'$ is the Hamiltonian constant (whose value is computed by substituting the initial values of the canonical variables in eq. (16)), and with the understanding that the plus sign is now to be taken when $\epsilon\mathcal{F}_1$ and $\sin h_*$ have the same sign. The polynomial under square root of (21) can be proven to be quartic in H. Therefore we can write it as [18]

$$\begin{aligned} f_4(H) &= \epsilon^2\mathcal{F}_1^2 - (\mathcal{H}' - \mathcal{H}_N - \epsilon\mathcal{F}_0)^2 \\ &= a_4 + 4a_3H + 6a_2H^2 + 4a_1H^3 \\ &\quad + a_0H^4, \end{aligned} \tag{22}$$

and rewrite eq. (21) as

$$\int_{H_0}^{H} \pm\frac{dx}{\sqrt{f_4(x)}} = \int_{t_0}^{t} d\tau, \tag{23}$$

where H_0 is the initial value of H and t_0 the initial time. The coefficients of $f_4(H)$ are reproduced in full form in [5]. The left-hand side of this equation is, apart from the sign ambiguity, an elliptic integral in standard form and can be solved to give (see [18], §20.6)

$$\begin{aligned} H(t) = H_0 + \Bigg(&2\left[\wp(t) - \frac{1}{24}f_4''(H_0)\right]^2 \\ &-\frac{1}{48}f_4(H_0)\,f_4^{iv}(H_0)\Bigg)^{-1} \\ &\cdot\Bigg\{\frac{1}{2}f_4'(H_0)\left[\wp(t) - \frac{1}{24}f_4''(H_0)\right] \\ &+\frac{1}{24}f_4(H_0)\,f_4'''(H_0) \\ &\pm\sqrt{f_4(H_0)}\wp'(t)\Bigg\}, \end{aligned} \tag{24}$$

where $\wp(z) \equiv \wp(z; g_2, g_3)$ is a Weierstrass elliptic function defined in terms of the invariants

$$g_2 = a_0a_4 - 4a_1a_3 + 3a_2^2, \tag{25}$$

$$g_3 = a_0a_2a_4 + 2a_1a_2a_3 - a_2^3 - a_0a_3^2 - a_1^2a_4, \tag{26}$$

the derivatives of f_4 are intended with respect to the polynomial variable, the $\pm$ sign is chosen based on the initial signs of $\sin h_*$ and $\mathcal{F}_1$, and we set the initial time $t_0 = 0$ for convenience.

The Weierstrass function is always a real valued function and it can be periodic or non-periodic depending on the values of g [5]. If $g_2 > 0$, $g_3 > 0$ and $\Delta = 0$ $\wp(t)$ simplifies to a simple periodic function whose period depends on the value of the parameters g. An interesting application of this limit is the case of the geodetic precession. In this case the coefficients a_0 and a_1 of the quartic polynomial $f_4(H)$ are both zero, and thus

$$g_2 = 3a_2^2, \tag{27}$$

$$g_3 = -a_2^3, \tag{28}$$

with

$$\begin{aligned} a_2 = &-\frac{1}{2}\epsilon\frac{m_2^4\mathcal{G}^4\mathcal{H}'}{G^3L^3} + \frac{15}{16}\epsilon^2\frac{m_2^8\mathcal{G}^8}{G^3L^7} \\ &+\frac{1}{4}\epsilon\frac{\tilde{G}^2m_2^4\mathcal{G}^4\mathcal{I}_1}{G^3L^3} - \frac{15}{8}\epsilon^2\frac{m_2^8\mathcal{G}^8}{G^4L^6} \\ &-\frac{3}{8}\epsilon^2\frac{\tilde{G}^2m_2^8\mathcal{G}^8}{G^6L^6} - \frac{1}{4}\epsilon\frac{m_2^6\mathcal{G}^6}{G^3L^5}, \end{aligned}$$

and the modular discriminant Δ is also null. Since, by substitution of $\mathcal{H}'$, one obtains

$$\begin{aligned} a_2 = -\frac{3}{8}\epsilon^2\frac{m_2^8\mathcal{G}^8}{L^6G^6}\big[&G^2 - 2H_0^2 + 2H_0\tilde{H}_* \\ &+\tilde{G}^2 + 2G_{xy,0}\tilde{G}_{xy,0}\cos h_{*,0}\big] > 0, \end{aligned} \tag{29}$$

[2] This, of course, is not possible if singular equilibrium points are present or if we are dealing with the indeterminate forms arising when the nodal angles are undefined.

DOI: 10.2420/AF07.2013.29

(the zero subscript represents initial values), because the quantity in the square brackets is the square of the magnitude M of the total mean angular momentum vector. As consequence, g_3 is always positive and the behaviour of $H(t)$ is restricted to be purely periodic and the angular velocity of the periodic motion of $H(t)$ is

$$T = \frac{3}{2}\epsilon\frac{m_2^4 g^4}{L^3 G^3}M \tag{30}$$

which is in agreement with the precessional angular velocity calculated for the geodetic effect.

5 Application to physical systems

At this point we can use the new solutions to deduce the behaviour of real physical systems. It is clear that, because of our initial assumptions, we can consider highly idealised cases and the aim will be only to highlight the role that relativistic effects play in the long term. The most interesting of these effects is the fact that the spin-orbit and spin-spin couplings induce periodic oscillations of the mean orbital plane and, at the same time, of the mean spin vector. Such effects are typically small for the orbital plane, but, because of the conservation of the z component of the total angular momentum $\tilde{H}_*$, they correspond to non-negligible oscillations in the orientation of the mean spin.

In [5] some interesting concrete examples were considered using the *mpmath* multiprecision library [13]. Here we summarize briefly the case of Gravity Probe B. Before starting it is worth to to point out that the perturbative treatment outlined in the previous sections produces results in terms of the components of the mean angular momentum vectors with respect to a fixed (non-rotating) centre-of-mass reference system (rather than in terms of obliquity and relative orientations). Therefore, in the geometrical interpretations of our results, we will always be referring to *absolute* (as opposed to *relative*) orientation angles.

5.1 Earth-orbiting gyroscope: Gravity probe B

Let us consider the gravitomagnetic effects on a gyroscope in low-orbit on board of a spacecraft around the Earth. The parameters of the system are taken from the experimental setup of the Gravity Probe B mission [9]: the orbit is polar ($i = 90.007°$) with a semi-major axis of 7027 km and low eccentricity ($e = 0.0014$). The gyroscope consists of a rapidly rotating quartz sphere (38 mm diameter) whose spin axis is lying on the Earth's equatorial plane (i.e., the spin plane is also "polar"). The spin vector of the gyroscope and the orbital angular momentum vector of the spacecraft, both lying on the equatorial plane, are perpendicular to each other. Translated into mean Delaunay and Serret-Andoyer parameters, this initial geometric configuration implies $H \sim 0$, $\tilde{H}_* \sim 0$, $G_{xy} \sim G$, $\tilde{G}_{xy*} \sim \tilde{G}$ and $h_* = \pm\pi/2$ (with the sign depending on the values of h and $\tilde{h}$ – note that in [9] the configuration shown in Fig. 1 implies $h_* = -\pi/2$). The substitution of these values in the general formula for $H(t)$ yields a period of roughly 195 ka. The solution includes the expected frame-dragging precession, which, after substitution of the appropriate numerical values, amounts to circa $0.04''$ per year [9]. The effect corresponds to a movement of the mean spin axis of the gyroscope along the equatorial plane in the direction of the rotation of the Earth. The other reproduced relativistic motion is the geodetic effect, which is a drift of the mean spin axis on the orbital plane in the direction of the orbital motion. Substitution of the numerical values in this formula yields the expected value of circa $\pm 6.6''$ per year.

6 Conclusions

From the consideration above it is clear that Lie series perturbation theory can lead to the discovery of interesting aspects of Post-Newtonian gravitation which were so far unknown or poorly understood. The role that these results have in terms of the development of Relativistic Positioning Systems lays basically in the fact that RPS in its present formulation requires a characterization of the motion of the satellites in the spacetime generated by Earth. Therefore the possibility to refine our knowledge of the Post-Newtonian dynamics translates directly in an improvement in the accuracy of RPS. It is worth to stress, however, that the work describe above refers explicitly to the problem of two bodies, while a RPS requires at least five (Earth, or other large body and four satellites). Although it seems unrealistic to obtain an exact solution in this complicated case there is no fundamental limitation in Lie series perturbation theory to its treatment. We feel therefore that Lie perturbations can be ascribed as a key tool in the development of a "realistic" RPS.

References

[1] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 2nd edition, May 1989.

[2] B. M. Barker and R. F. O'Connell. Derivation of the equations of motion of a gyroscope from the quantum theory of gravitation. *Physical Review D*, 2(8):1428, 1970.

[3] B. M. Barker and R. F. O'Connell. The gravitational interaction: Spin, rotation, and quantum effects - a review. *General Relativity and Gravitation*, 11(2):149–175, Oct. 1979.

[4] F. Biscani. *Design and implementation of a modern algebraic manipulator for Celestial Mechanics*. Ph.D. dissertation, University of Padua, 2008.

[5] F. Biscani and S. Carloni. A first-order secular theory for the post-newtonian two-body problem with spin – i: The restricted case. *MNRAS (January 21, 2013) 428 (3): 2295-2310*, 03 2013.

[6] T. Damour. Coalescence of two spinning black holes: An effective one-body approach. *Physical Review D*, 64(12), Nov. 2001.

[7] A. Deprit. Canonical transformations depending on a small parameter. *Celestial Mechanics*, 1(1):12–30, Mar. 1969.

[8] A. Deprit. Delaunay normalisations. *Celestial Mechanics*, 26:9–21, Jan. 1982.

[9] C. Everitt et al. Gravity probe b: Final results of a space experiment to test general relativity. *Physical Review Letters*, 106(22):221101, May 2011.

[10] P. Gurfil, A. Elipe, W. Tangren, and M. Efroimsky. The Serret-Andoyer formalism in rigid-body dynamics: I. symmetries and perturbations. *Regular and Chaotic Dynamics*, 12(4):389–425, Aug. 2007.

[11] J. Heimberger, M. Soffel, and H. Ruder. Relativistic effects in the motion of artificial satellites - the oblateness of the central body II. *Celestial Mechanics and Dynamical Astronomy*, 47:205–217, 1990.

[12] G. Hori. Theory of general perturbation with unspecified canonical variable. *Publications of the Astronomical Society of Japan*, 18, 1966.

[13] F. Johansson et al. *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.17)*, February 2011.

[14] A. Morbidelli. *Modern Celestial Mechanics: Dynamics in the Solar System*. CRC Press, 1st edition, July 2002.

[15] C. D. Murray and S. F. Dermott. *Solar System Dynamics*. Cambridge University Press, Feb. 2000.

[16] J. Palacián. Closed-form normalizations of perturbed two-body problems. *Chaos, Solitons & Fractals*, 13(4):853–874, 2002.

[17] D. L. Richardson and T. J. Kelly. Two-body motion in the post-Newtonian approximation. *Celestial Mechanics*, 43:193–210, 1988.

[18] E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, fourth edition, 1927. Reprinted 1990.

 DOI: 10.2420/AF07.2013.29

Acta Futura 7 (2013) 35-47
DOI: 10.2420/AF07.2013.35

Acta Futura

Relativistic positioning systems: perspectives and prospects

Bartolomé Coll*
Departament d'Astronomia i Astrofísica, Universitat de València, 46100 Burjassot, València, Spain

Abstract. Relativistic positioning systems are interesting *technical objects* for applications around the Earth and in the Solar system. But above all else, they are basic *scientific objects* allowing developing relativity from its own concepts. Some past and future features of relativistic positioning systems, with special attention to the developments that they suggest for an *epistemic relativity* (relativistic experimental approach to physics), are analyzed. This includes *relativistic stereometry*, which, together with relativistic positioning systems, allows to introduce the general relativistic notion of (finite) *laboratory* (space-time region able to perform experiments of finite size).

1 Introduction

Relativistic positioning systems were born as *scientific objects*.[1] But many people consider them as *technical objects*,[2] a sort of classical positioning systems directly modeled in relativity. Anyway, their handle is not easy, neither as scientific objects nor as technical ones. This is why to meet all of us together and share methods and ideas is an unavoidable step to progress in their development. I want to congratulate the Advanced Concepts Team of the ESA and the Faculty of Mathematics and Physics of Ljubljana for this initiative. This paper is the text of my talk in this meeting.[3]

I believe that most people are better interested on relativistic positioning systems as technical objects. But, since their origin as scientific ones, *relativistic positioning systems are paradigmatic objects able to transform relativity in a truly experimental branch of physics*. Relativistic positioning systems are the first component in the construction of *relativistic laboratories* of finite size.

It seems evident that to be aware of this role may impel us, whatever be our interest, technical or scientific, to solve problems set out by relativistic positioning systems and in finding scientific applications of them.

The purpose of this lecture is twofold: to present my perspective about some of the concepts related to relativistic positioning systems, and to prospect for the first ingredients able to make relativity an experimental approach to physics.

2 Perspectives

Why relativistic positioning systems have taken so long to appear? From my personal experience, the answer is very clear: because the obstructions caused by some prejudices. We shall begin with a brief account of those physical prejudices that, for me, have affected and retarded the natural development of relativistic positioning systems.

*Corresponding author. E-mail: bartolome.coll@uv.es

[1] A 'scientific object' means here an object whose knowledge is interesting by itself, independently of its practical utility.

[2] A 'technical object' means here an object whose knowledge is interesting for practical applications, to control our environment.

[3] Workshop *Relativistic Positioning Systems and their Scientific Applications*, Brdo (Slovenia), 19th to 21st September 2012.

2.1 The genesis of relativistic positioning systems

The idea of relativistic positioning systems appeared almost simultaneously and in all likelihood independently, in Bahder [4], Coll [6] and Rovelli [14]. Every one of these three authors arrived to this idea by very deeply different ways. It is very interesting, and strongly striking, that, in a so short period of time and without apparent precedents, three very similar ideas appeared from three so different ways.

So different ways limit myself to comment only about *my* genesis of the concept.[4] In this genesis, a little number of papers appears as landmarks or precursors for the ideas presented in [6]:

* *Light coordinates in Relativity* (1985) [5],
* *Symmetric frames on Lorentzian spaces* (1991) [8],
* *199 causal classes of space-time frames* (1992) [9],

They have contributed to the weakening[5] of at least one of the following prejudices:

a: a physical frame must involve necessarily an a priori definition of space-like synchronization,

b: no frame of four real null vectors exists in relativity,

c: coordinate systems have no physical meaning.

Prejudice **a** is a mixture of a feeling-based prejudice and an error-based one. On one hand, it is related to the old feeling, current many decades ago in metrology, that a standard of distance must be matter-based and not clock-based, and is a remainder of the feeling that an extended instantaneous space is physically meaningful. On the other hand, it comes from a confusion between the notions of a physical system and of a social or conventional one. It is clear that for our social conventions around the Earth an a priori synchronization is very convenient but not at all necessary.[6] The absence of a strict symmetry of the gravitational field and of the Earth surface implies the *non existence* of an a priori physical synchronization. At the, at present, uncertainty, our conventional synchronization is only possible for our conventional International Atomic Time, not for the proper physical time of every event around the Earth.[7,8]

Prejudice **b** is an error-based prejudice, due to the 'saturation' of the concept of "null" tetrads, produced by the abundance of works in the well-known Newman-Penrose formalism [12]. The error consists in applying unconsciously, to *any* set of four null vectors, the orthogonality condition imposed by Newman and Penrose to their null tetrads. It is perhaps the weaker one of the above three prejudices, and the easier to dilute, but it has been almost 'universal' among relativistic physicists, whatever their renown, up the the last decade.

Prejudice **c** is also an error-based one, due to an incorrect statement of the principle of general covariance. This principle states that *the laws of physics are invariant by the choice of coordinate systems*, and it is an extension of the principle of dimensional invariance, that states the invariance of the laws with respect to the particular units used to obtain them. But curiously, during dozens and dozens of years, this statement has slid to the incorrect form *the laws of physics are independent of the choice of coordinate systems*,[9] generating the prejudice in question, meanwhile the similar statement for the dimensional invariance generated a deep research to improve the definition and construction of physical units.[10]

On the basis of different combinations of the above prejudices, the publication of papers [8] and [9] was strongly retarded[11], paper [5] was forbidden[12] and the research work on this subject, criticized by many colleagues, was underestimate.

For me, this simple sample of the effects of prejudices already explains in part why relativistic positioning systems have taken so long to appear. But in general the damages that prejudices of referees and colleagues pro-

[4] During many years, the germ of relativistic positioning systems covered a corner of my private garden of thoughts for week-ends and holidays. But every flower sprouted in it, every idea, I showed it to my friends Joan Ferrando, Juan Antonio Morales, Albert Tarantola (†2009) and José María Pozo, who watered it carefully. This is the meaning of '*my* genesis'.

[5] I would like to say 'removal', but recent discussions with colleagues show that it is not, unfortunately, the case.

[6] The *local Solar time* everywhere on the Earth is an example of a non a priori synchronization.

[7] Like the Geoid, a physical synchronization on the Earth for a physical time can only be the a posteriori result of continuous careful measures.

[8] Relaxing the space-like condition on a synchronization, i.e. reducing it to the locus of equal time events, a relativistic positioning system does not defines *one* a priori synchronization, but *four* equivalent ones.

[9] For an object ω, its 'independence' of a set $\mathcal{C}$ of objects $\mathbf{c}$ means that its conception, definition, construction and use may be made in *absence* of $\mathcal{C}$, meanwhile its invariance of $\mathcal{C}$ means that for its conception or its definition or its construction or its use, objects of $\mathcal{C}$ are needed, but that their effect on ω are *independent* of the particular objects $\mathbf{c}$, $\mathbf{c}'$, etc. of $\mathcal{C}$ taken for its elaboration or use.

[10] The NIST (National Institute of Standards and Technology), for example, is a good example.

[11] Almost three years for [9].

[12] A hierarchical superior of my research department forbade the submission for publication of an English version of the paper and the continuation of the research on this subject.

 DOI: 10.2420/AF07.2013.35

duce is stronger.[13] To help young researches to avoid them, let me remember that :

> The main obstructions to innovative research are the prejudices. The own prejudices for its conception. Those of the peers for its diffusion.

Prejudices do not belong to the past.[14] The following example concerns a fashionable one.

2.2 An example of an extended prejudice

In 2011, the OPERA experiment between the CERN (Geneva) and the LNGS (Gran Sasso) mistakenly reported neutrinos appearing to travel faster than light [1].

Some scientists expressed their doubts about this result, like Hawking,[15] or their incredulity, like De Rujula,[16] but, as reported in many different media, almost all of them, including the directors of the experiment, believed that, if the result were true, relativity theory would be refuted or, at least, deeply damaged. We can thus state, as a general belief among scientists concerned by the subject, that:

> Neutrinos traveling faster than the velocity of light c between CERN and LNGS are inconsistent with standard relativity theory.

Is this belief correct or is it a prejudice? Let us see the answer in three steps.

Let us begin remembering what is a local theory. A *local theory* is a theory whose general statements and equations are local, i.e. valid in such small space-time[17] regions that any physical quantity not mentioned in the statement or not appearing in the equations has to be supposed as constant. Mathematically a physical theory is local if its formulation is infinitesimal (relates physical fields and their space-time variations at every event).

Now, relativity theory,

- by the concepts used in its construction,
- by the principles on which it is founded,
- by the domain of influence of its equations,
- by the tensor character with which it represents the physical quantities
- by the concept itself of space-time that it proposes,
- and because it gives no phenomenological theory for the construction of its current (energy tensor), but supposes it can be obtained by means of classical balances,

is a local theory in all its constituents.

Consequently, as all others general statements of the theory, the one that says that *the velocity of light* $c = 299792458\ m/s$ *is a physical limit*, or any other equivalent version, is a local statement. And, being local, the velocity cannot but be an *instantaneous velocity*, i.e. measured in a so small time interval that any physical quantity not implied by the concept of velocity is constant. It cannot be, in general, a *mean velocity*. The crucial point is that:

* In Newtonian theory, where time and space are absolute, if the instantaneous velocity of a particle remains constantly *lesser* than a value v during a finite interval, the mean velocity in this interval will also be lesser than v.
* But in relativity, where time and space are different at different events, a particle whose instantaneous velocity remains constantly lesser than a value v, whatever it be, during a finite interval, may have a mean velocity lesser, *equal* or *greater* than v.

This last fact is easy to see in very simple cases, as it is the one of Fig 1, representing an accelerated observer submitted to an acceleration g who, at a proper instant τ_1, sends a light signal to a mirror situated at a proper distance d, and receives it at a proper instant τ_2. Because accelerated clocks slow down with respect to inertial ones, and that the distance d to the mirror is greater than that of the inertial observer that cross at τ_1 and τ_2, the mean velocity, $v_m \equiv d/\Delta\tau$, of the light with respect to the accelerated observer cannot but be greater than c. The precise amount is given by:

$$v_m = c\frac{\frac{gd}{c^2}}{\ln(\frac{gd}{c^2}+1)} \approx \frac{c}{1-\frac{1}{2}\frac{gd}{c^2}}.$$

Note that this is a two-way measure that involves only *one* clock, *one* space and, in it, *one* distance. A one-way measure involves in general, not only *two* clocks, *two*

[13] Prejudices expend time and money, and demoralize their victims.

[14] There exists no panacea to get rid of them. To remove them is an individual inner process for which, if the intention is necessary, it is also frequently insufficient. Fortunately, the prejudices cited in the text belong to the class of those that disappear when one is very careful with the analysis of the conditions under which their assertion is true.

[15] "It is premature to comment on this. Further experiments and clarifications are needed" said him in [10].

[16] "Flabbergasting", said him in [13].

[17] I think that we, physicists, have not yet sufficiently 'symbiosed' the concepts of space and time in practice, so that we have not yet merit the moral right to write 'space-time' without its hyphen of compound word'.

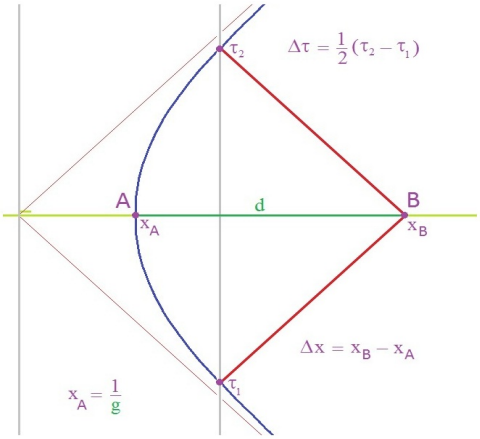


Figure 1. *An observer A with acceleration* g *sends at* τ_1 *a signal to a mirror B situated at a distance* d*, and receives it at* τ_2*. For him, the mean velocity of light is greater than* c.

spaces and *two* distances,[18] but above all, it involves their precise correspondence between them.[19]

In the OPERA experiment, the mean velocity of neutrinos, traveling a region of a non constant gravitational field, is measured between two events, the CERN and the LNGS, of different gravitational acceleration (different clock rates). This situation has been modeled in a relativistic gravitational space-time by Allès [2] and Lüst and Petropoulos [11], and in both cases they have found mean velocities greater than c for the OPERA configuration, although of many orders of magnitude lesser than the experimental obtained value.

The error of the Opera experiment has not been, as generally believed, to obtain a velocity greater than c for neutrinos, because relativity foresee it, but the simple quantitative one of obtaining an inappropriate numerical value.

Let me emphasize this point:

> *In relativity, mean velocities of particles may be lesser, equal or greater than the instantaneous velocity of light.*

[18]Take into account that, in relativity, between two observers A and B, the spatial distance from A to B is the same than that from B to A only if they are locally near, otherwise they are generically different.

[19]In the case of the one-way OPERA experiment, the symmetries (staticity and spherical symmetry) allow to reduce this correspondence to a synchronization between the two clocks.

2.3 Relativistic and classical positioning systems

Many people consider relativistic positioning systems as Newtonian or classical positioning systems directly worked out with relativity. This is not correct. Although intimately related, they are very different objects. Let us see it.

The aim of a relativistic positioning system is:

- to allow any user[20] to know its location in a well defined four-dimensional physical coordinate system,
- to provide the user with its proper time and proper distance (space-time metric),
- to characterize its space-time trajectory dynamically (proper acceleration) and gravitationally (gravimetry).

A relativistic positioning system around the Earth, or RGNSS (Relativistic Global Navigation Satellite System), wants thus to characterize the physics of the space-time region between the constellation of satellites and the Earth surface. Fig 2 represents the gravitational field of this extension in a intuitive form.

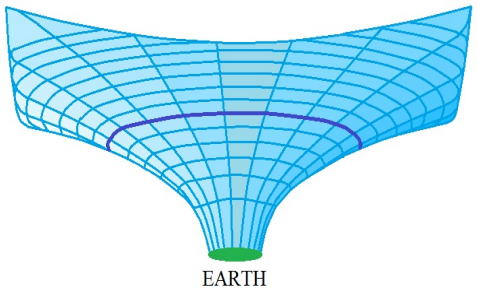


Figure 2. *The aim of a RGNSS is to locate users and to provide them with their proper time and proper units of distance so as to characterize physically the region between the Earth and the satellite constellation.*

On the other hand, the aim of a classical positioning system is:

- to allow any user to know its position with respect to a specific chart of the Earth surface, and its time with respect to a time scale based on the International Atomic Time (TAI).

[20]The word *user* here denotes any person or device able to receive the pertinent emitted data from the relativistic positioning system and to extract from them the corresponding information. For short, we shall refer to this user as *it*.

 DOI: 10.2420/AF07.2013.35

In the positioning systems around the Earth, or GNSS (Global Navigation Satellite Systems), the specific charts of the Earth surface in use are the World Geodetic System (WGS84) or the International Terrestrial Reference Frame (ITRF), differing by less than ten centimeters in their last determinations. The TAI scale, partially physical, social and political, is a weighted average from many national laboratory clocks, and represents a sort of mean proper time on a mean sea surface level. Moreover, its extension all over the space-time region between the Earth surface and the satellite constellation has undoubtedly many practical and social advantages.

But this extension at any altitude of the TAI scale and of the Earth surface chart is tantamount to a Newtonization of the space-time region between the Earth surface and the satellite constellation. Fig 3 represents intuitively this situation.

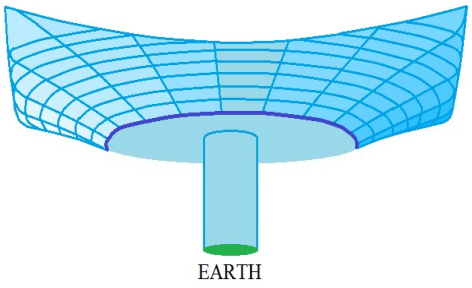


Figure 3. *The extension of the Earth surface chart and of the TAI to all the space up to the satellite constellation, constitute a Newtonization of this region.*

Note that other Newtonizations are possible, as for example the one obtained by the extension to all the region of an averaged time at the satellite constellation level.

Although for a good smooth running of both systems, RGNSS and GNSS, the *same* information is needed, this information is obtained, interpreted and used very differently. Thus, the physical timing of proper clocks of the satellites constitutes the basic data for relativistic positioning systems, meanwhile this timing is used in classical positioning systems to construct the TAI timing at the satellite level, a pure conventional timing at this level. In this sense, the relativistic corrections for the GPS (for example, in Ashby [3]) are used "subtractively" on the physical clocks in motion not to improve their precision but, on the contrary, to better simulate their Newtonian (absolute) behavior.

This situation may be described in short by saying that relativistic positioning systems are *physical* systems meanwhile classical positioning systems are *conventional* ones. To improve both of them, I believe better to first improve the physical systems without regard to the conventional ones, and then to use these results to improve the conventional systems. Anyway, we must be conscientious with what we are doing: using relativity to better 'Newtonize' a GNSS or using relativity to construct a positioning system.

2.4 The main relativistic positioning systems

Let us remember some known concepts about location systems. A *location system* is a physical realization of a coordinate system (for an epistemic definition without reference to mathematics see, for example [7]).

Two important classes of location systems are the *reference systems* and the *positioning systems*. The goal of reference systems is to situate the events of a domain with respect to a given observer (generally located at the origin), meanwhile the goal of positioning systems is to indicate its own position to every event of the domain. In Newtonian theory, as far as the velocity of light is supposed infinite, both goals are exchangeable in a sole location system. But in relativity this is no longer possible and it is impossible to construct a positioning system starting from a reference system, but one can always (and very easily) construct a reference system starting from a positioning system.

Relativistic positioning systems

Positioning systems are *immediate*, what means that every event of their domain may know its proper coordinates without delay (in fact, it is this property that defines them). Here they are also supposed *generic* and (gravity) *free*, guaranteeing their existence in any generic space-time and their construction without the previous knowledge of the gravitational field respectively. It follows from these and the above properties that, whenever possible, it is a positioning system, and not a reference system, that has the most interest to be constructed. From now on, we consider only relativistic positioning systems. Denote by $\mathcal{P}$ the set of all of them.

Auto-locating positioning systems

An important subclass of positioning systems are the *auto-locating positioning systems*, that broadcast their proper time but also the proper time that they receive

from their neighboring satellites. Let τ^{IJ}, $I \neq J$, be the proper time of the satellite clock J received by the satellite I at its proper time instant τ^I. Then, the set of sixteen data $\{\tau^I, \tau^{IJ}\}$ received by an observer contains, of course, the emission coordinates τ^I, $(I = 1, \cdots, 4)$ of this observer but also the coordinates $\{\tau^I, \tau^{IJ}\}$ of every satellite I in the emission coordinate system $\{\tau^I\}$.[21] Denoting by $\mathcal{L}$ the set of all auto-locating positioning systems, we have $\mathcal{L} \subset \mathcal{P}$.

Autonomous positioning systems

Auto-locating systems allow any user to draw univocally the world-lines of the satellites in the emission coordinate system that they broadcast. But the user still does not know how to draw these world-lines in the space time in which it is living.

For a user to be able to do this, the coordinate data $\{\tau^I, \tau^{IJ}\}$ broadcast by the auto-locating system has to be completed with:

* dynamical data of the satellites (acceleration, gradiometry),
* observational data from them (e.g. position of reference quasars or pulsars) and
* gravitational knowledge of the coordinate region (theoretical, experimental or mixed).

The set of this information is called the *autonomous data*. Auto-locating systems broadcasting autonomous data are called *autonomous positioning systems*. Denoting by $\mathcal{A}$ the set of all of them, we have $\mathcal{A} \subset \mathcal{L} \subset \mathcal{P}$.

Generic positioning systems, those in the difference $\mathcal{P} \setminus \mathcal{L}$, have the interest of having shown that relativistic positioning systems generic, free and immediate exist. But, above all, they have the advantage of being easier to study than the auto-locating systems of $\mathcal{L}$. Nevertheless we have seen on one hand that, whenever possible, there are them, and not reference systems, that have the most interest to be constructed and, on the other hand, we have seen that because the absence of autonomous data they need to be referred to a reference system. That is to say: *generic positioning systems are incoherently incomplete or insufficient*. Generic auto-locating systems, those in $\mathcal{L} \setminus \mathcal{A}$, also inherit the above incoherent incompleteness.

[21] The world-lines of the satellites do not belong to the emission coordinate domain of the positioning system, but to its border. Nevertheless, although they are not differentiable along the world-lines, the emission coordinates are well defined on them.

Thus, non-autonomous positioning systems, those in $\mathcal{P} \setminus \mathcal{A}$, appear as intermediate hybrids between relativistic reference systems and autonomous positioning systems.

Autonomous positioning systems are the best location systems. They are the challenge. They were proposed, for the first time, in [6].

3 Prospects

I do not approve the way relativity has been developed during its century of existence.

Relativity is a physical theory of the gravitational field, but it is also a physical theory of the space-time. And it is well established that the relativistic descriptions of both objects, gravitational field and space-time, improve their corresponding homologues in Newtonian theory. For this reason, I think that:

* as a dynamic physical theory, relativity must provide more experiments than simple experimental verifications from time to time, as it is the practice today,
* as an improved theory of the space-time, *any* physical experiment, whatever it be, *ought* to be qualitatively *described in the framework of relativity*, regardless of its quantitative evaluation, for which in many cases Newtonian calculations could suffice,
* as an improved theory of the gravitational field, relativity *ought* to propose experiences and methods of measurement of general gravitational fields (four-dimensional metric), which, up to now, are conspicuous by their absence.

In short, *relativity needs to develop a proper experimental approach to the physical world*. And I believe that we already have the conceptual basic ingredients for this development.

Now, for this purpose, we need to make more precise the idea of a relativistic experimental approach.

3.1 Epistemic relativity

In relativity, a good deal of scientific works analyze physical and geometrical properties of the space-time, but

- don't integrate the physicist as a part of it,

and

 DOI: 10.2420/AF07.2013.35

- forget implicitly that:
  - information is energy,
  - neither the density of energy, nor its velocity of propagation can be infinite in relativity.

Many of these properties of the space-time may be analyzed by a *geometer* on his desk, but to be known by an *experimental physicist* would require the qualities of an *omniscient god !*

For these reasons, we characterize these scientific works as belonging to *ontic relativity*.[22]

Of course, ontic relativity is absolutely *necessary* for physics. The conceptual evaluation of many physical situations in order to be able to conceive physical experiments belongs to ontic relativity. But ontic relativity is also manifestly *insufficient* for a relativistic experimental approach to physics.

At the opposite side, the works in relativity that:

- integrate the physicist as an element of the problem considered,
- concern physical quantities that the physicist can know or measure and
- take into account explicitly what information, when and where, the physicist is able to know,

will be considered as characterizing *epistemic relativity*.[23]

The main objective of epistemic relativity is to provide the physicist with the knowledge and protocols necessary to make relativistic gravimetry in its (a priori unknown) space-time environment.

This is the first and unavoidable step to develop experimental relativity as the natural scientific approach to our physical world.

3.2 Relativistic stereometry

We know that in the space-time, the adequacy between a mathematical model and the physical system that it describes needs of a univocal correspondence between them. Thus, because, in the differentiable manifold of the mathematical model, points are identified by their coordinates, we need to know how to construct a *location system*, that is to say, how to label the events of the physical space-time. But we know also that the best location systems (those which are generic, free and immediate) are the relativistic positioning systems. Consequently, it becomes evident that relativistic positioning systems are the first ingredient of epistemic relativity.

What other else do we generically need in epistemic relativity?

A finite laboratory

In fact, what we need is to be able to consider the space-time region of physical interest as a laboratory. The question is then: what is a laboratory (of finite dimension) in relativity?

A simple reflexion shows that, in fact, and regardless of the specificity of its measurement devices, any laboratory, has to provide us with:

* a precise location of the significant parts of the physical system in question, and
* a precise description of its pertinent intrinsic physical properties.

Similarly to the precise location, which is obtained by means of a system of four clocks (relativistic positioning system), the precise description of the intrinsic properties of a system has to be obtained by means of a system of four observers. Such a system of four observers is called a *stereometric system*. Thus,

> A finite *laboratory* in relativity is a space-time region endowed with
>
> * a relativistic positioning system and
> * a relativistic stereometric system.

Relativistic stereometric systems

In physics, the word 'observer' is rather polysemic. To what notion of observer are attached the relativistic stereometric systems?

Here an *observer* is a 4π-wide hypergon *eye*[24] able to record and to analyze its input. It is a *local* device,[25] defined at every space-time event by its unit velocity, that projects the past light cone of the event onto its *celestial sphere*.[26]

[22] From Greek '*ontos*', 'being', with the meaning of 'what it is' as opposed to 'how it is seen'.

[23] From Greek '*episteme*', 'knowledge', with the meaning of 'how we obtain it'.

[24] Also called 4π-steradian fish eye.

[25] Physically a device is *local* if it takes up such a small space-time region that all physical fields in it may be considered as constants. Mathematically it means that the device needs only of an infinitesimal region around a space-time event to be defined.

[26] The celestial sphere of an observer at a space-time event is the quotient of its three-dimensional space by the set of all the past null directions converging at this event.

Although they are not very abundant, there exist interesting papers on relativistic vision but they differ very much in strategy, starting hypothesis and definition (frequently implicit) of 'eye'. And, unfortunately, almost none of them emphasize the invariants (intrinsic properties) of the configuration that they analyse, a crucial fact for us.[27]

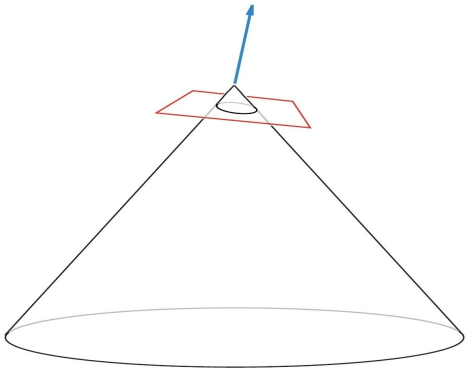


Figure 4. *Three-dimensional representation of a* 4π*-wide hypergon eye as a local device that projects the past-like cone of the event on its celestial sphere.*

A very interesting feature here is that *relativistic stereometric systems are the causal duals of positioning systems.*

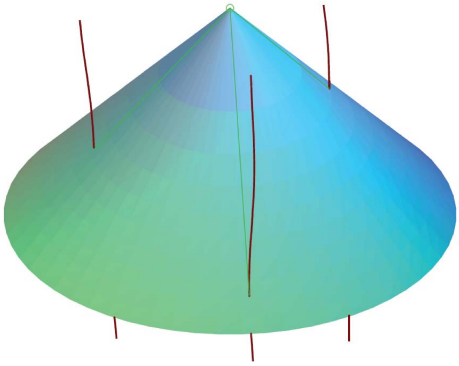


Figure 5. *Three-dimensional representation of a relativistic positioning system.*

They are causal duals operationally: relativistic positioning systems are *passive* for the user meanwhile relativistic stereometric systems are *active*.[28] But they are also causal duals conceptually, as space-time objects, because they differ simply by a timelike inversion (see Figs 5 and 6). It is then clear that many of the properties of one of these systems may be transformed, by simple change of time orientation, in properties of the other system.

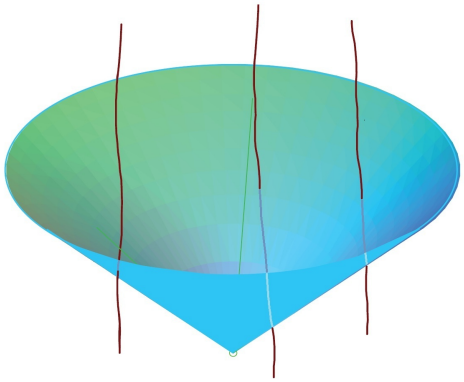


Figure 6. *Three-dimensional representation of a relativistic stereometric system.*

The aim of relativistic stereometry is to obtain the intrinsic properties of physical systems starting from their relative properties seen by four observers.

In relativity, because the space-time objects are *histories*,[29] the intrinsic properties of a system related to its *form* must involve, besides the field of proper distances between its neighboring points, the field of proper times of its local elements. A main set of intrinsic properties are the *visual* ones, obtained by adding to the intrinsic geometric properties, the color field of the local elements.

Thus a first basic problem of relativistic stereometry is to obtain, for every local element of a physical system, its proper color and the proper distances to its neighboring elements starting from the corresponding relative elements observed by the four observers of the relativistic stereometric system.

[27] It would be stimulating to analyze and classify all this material, and to select those result attached to hypergon eyes.

[28] A relativistic stereometric system is also a *location system* for *active* users, i.e. for those users endowed with an instant-identifier, as for example a clock broadcasting its time. The times of reception, by the four observers of the stereometric system, of the signal of an instant of the user, constitute the *reception coordinates* of the user at that instant. Such a reception system, in addition to be active, is also not immediate: it is not a relativistic positioning system.

[29] Namely, they are the histories of the spatial objects of Newtonian theory.

DOI: 10.2420/AF07.2013.35

First theorems in relativistic stereometry

The intrinsic properties of a system are those which are observer-invariant. In relativistic stereometry we have thus to solve, starting from four observer-dependent perspectives, an inverse problem.

In order to show how it works, we shall consider the simplest physical system in the simplest stereometric operational frame: a colored point particle in a two-dimensional Minkowski space-time. In spite of its easy framework, we shall see that the solution to this stereometric inverse problem is interesting enough and constitutes a good example of epistemic relativity.

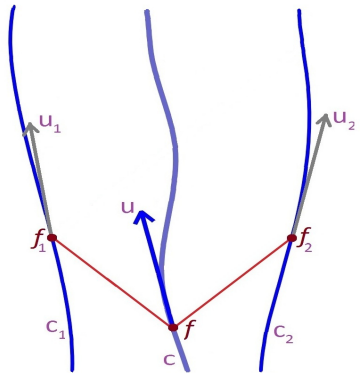


FIGURE 7. *A two-dimensional relativistic stereometric system obtaining the proper frequency of a particle.*

Thus, in Minkowski two-dimensional space-time, let C be the world-line of a colored material point of a physical system, of proper frequency f. Let f_1 and f_2 be the Doppler frequencies received respectively by the observers C_1 and C_2 of the relativistic stereometric system from an instant of C[30] and let ν_{12} be the relative velocity of these observers at the instants of reception of the signals f_1 and f_2. Then we have:

Theorem 1.- *In terms of the received frequencies* f_1 *and* f_2 *and of the relative velocity* ν_{12} *of the system at the reception instants, the proper frequency* f *of the colored point* C *is given by:*

$$f^2 = f_1 f_2 \sqrt{\frac{1+\nu_{12}}{1-\nu_{12}}}$$

Note that, if the colored point C is transported by one of the observers ($f = f_i$ for some $i = 1, 2$), the above expression reduces to the standard one for the Doppler shift.

In addition to the proper frequency, the Doppler frequencies f_1 and f_2 also allow to work out the relative velocities of the material point:

Theorem 2.- *The relative velocities* ν_1 *and* ν_2 *of the material point* C *with respect to the observers* C_1 *and* C_2 *of the relativistic stereometric system at the instants of reception of the signals* f_1 *and* f_2 *are given by:*

$$\nu_1 = \frac{f_2\sqrt{1+\nu_{12}} - f_1\sqrt{1-\nu_{12}}}{f_2\sqrt{1+\nu_{12}} + f_1\sqrt{1-\nu_{12}}}$$

$$\nu_2 = \frac{f_1\sqrt{1+\nu_{12}} - f_2\sqrt{1-\nu_{12}}}{f_1\sqrt{1+\nu_{12}} + f_2\sqrt{1-\nu_{12}}}$$

Note that the results in both theorems depend not only of the measured Doppler frequencies, but also of the relative velocity ν_{12} of the observes of the stereometric system at their instants of reception of these frequencies, a quantity that seems not obvious how to be measured. The question is thus: are these theorems epistemic ?

In fact, they *are not* epistemic. Moreover: by themselves they *cannot be* epistemic. The simple reason is that, without additional specifications, these two theorems do not fulfill any of the above three conditions characterizing epistemic relativity.

To fulfill these conditions, we must complete the above results with the information about:

- what physicist we have choose to make the experiment,
- when and where it[31] is able to be informed of the quantities needed to answer the problem,
- how can it know or measure these quantities.

This information is chosen here as follows:

* the simplest choice of physicist is to take it as one of the observers of the relativistic stereometric system, say C_2, as shown in Fig 8,
* then it will be able to be informed of all the quantities needed to answer our problem at the instant τ_{12} of reception of the pertinent information coming from the observer C_1,

[30] There is no matter here what instant-identifier is used: a clock associated to the point, measuring any non necessarily proper time, a flash reflected or carried by the point or any other pertinent device.

[31] The world *physicist* here denotes any person or device able to receive the pertinent information from the relativistic stereometric system, to record and to analyze it and to perform the computations needed for the problem in question. For short, we shall refer to this physicist as *it*.

* at that instant τ_{12} this physicist is informed of the quantity f_1, it already knows the quantity f_2, measured and recorded by it at τ^2, and it may know the quantity ν_{12} by computation.

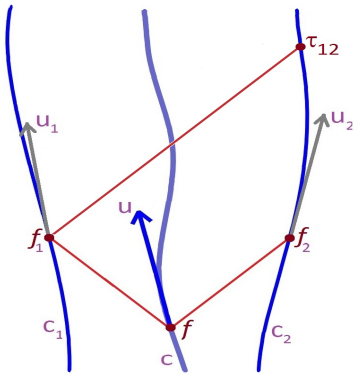


FIGURE 8. *Here, C_2 has been chosen as the physicist of the epistemic problem and, from the instant τ_{12} on, it may know all the quantities of the problem in question.*

The computation of the relative velocity ν_{12} may be made, for example, from the knowledge of the world-lines of the observers of the relativistic stereometric system. In the case where these observers are geodesic, its expression is very simple. If, for short, we call an epistemic theorem an *epistem*, from the above two theorems we have:

Epistem 0.- *In terms of the frequency of the proper time of the observer C_1 with respect to the proper time of the observer C_2, φ_{12}, the relative velocity ν_{12} between the observers of a geodesic relativistic stereometric system is given by:*

$$\nu_{12} = \frac{1-\varphi_{12}^2}{1+\varphi_{12}^2} \, .$$

Now, for this geodesic case, the above two theorems become respectively:

Epistem 1.- *In terms of the received frequencies f_1 and f_2 of a colored point C and of the frequency of proper times φ_{12} of a geodesic stereometric system, the proper frequency f of the colored point C is given by:*

$$f^2 = \frac{f_1 f_2}{\varphi_{12}} \, .$$

Epistem 2.- *The relative velocities ν_1 and ν_2 of the material point C with respect to the observers C_1 and C_2 of a geodesic stereometric system at the instants of reception of the signals f_1 and f_2 are given by:*

$$\nu_1 = \frac{f_2 - f_1\varphi_{12}}{f_2 + f_1\varphi_{12}} \, ,$$

$$\nu_2 = \frac{f_1 - f_2\varphi_{12}}{f_1 + f_2\varphi_{12}} \, .$$

In spite of the very simple context in which they are obtained, these results are interesting because they show the essentials of epistemic relativity. Of course, because the celestial sphere of an observer in a bidimensional space-time reduces to two opposite points, the problem of determining the proper distances to its neighboring elements of a material point cannot be considered in this dimension. The need of extending this work to three or four dimensions is evident.

3.3 What about the mathematics of relativity?

The development of Riemannian geometry at the beginning of the 20th century, the already existing classical theory of deformations and (the belief in) the local character of the fundamental physical laws are at the basis of the mathematical frame of the general theory of relativity, namely the Lorentzian differential geometry.

Insufficiency of differential geometry in relativity

But, in its present form, in front of these historical and conceptual justifications, there exist practical (epistemic!) insufficiencies of differential geometry in relativity. Initially, differential geometry appears appropriate for the description of locally deformed or curved regions. But classical fields are of infinite range, so that little local perturbations of a physical system do not remain confined, but spread indefinitely. This propagation of local little perturbations cannot be neglected, not only because of it physically meaningful character, but because it do not take place subtly, but at the velocity of light, an antroposcopic[32] velocity.[33]

These general facts do not diminish the unavoidable character of differential geometry, its inevitability in the formulation of relativity theory, but show its *insufficiency*. More particularly, in the study of positioning or

[32]Visible to the naked eye and important enough for human activities.

[33]Think, simply, in the important antroposcopic effect of striking a match in the darkness.

 DOI: 10.2420/AF07.2013.35

stereometric systems, or simply in the study of any epistemic problem of finite extension, this insufficiency is dramatic. We are all suffering of this situation.[34]

Finite-differential geometry

I think that it is an urgent task in relativity, for all of us, to try to construct a *finite-differential geometry.*

The purpose of finite-differential geometry is to introduce interchangeable finite versions of the basic ingredients of differential geometry, namely:

* metric g,
* connection Γ,
* curvature Riem.

The distance function $D(x,y)$, or its half-square $\Omega(x,y)$, the Ruse-Synge's *world-function*,

$$\Omega(x,y)=\frac{1}{2}D^2(x,y)\ ,$$

are already finite versions of the metric g. As it is known:

$$\Omega(x,y)=\frac{1}{2}\int_0^1 g\left(\frac{d\gamma}{d\lambda},\frac{d\gamma}{d\lambda}\right)d\lambda\ ,$$

$\gamma(\lambda)$ being the geodesic joining x and y, and their fundamental equations are:

$$g^{\alpha\beta}\partial_\alpha\Omega\partial_\beta\Omega=2\Omega\ ,\quad g^{ab}\partial_a\Omega\partial_b\Omega=2\Omega\ ,$$

where Greek and Latin indices are related to the first and second arguments of $\Omega(x,y)$ respectively.

Distance spaces are well known, but their link with differential geometry has not been yet sufficiently explored.

Let us think, in a given space-time, on a positioning system complemented with a number of additional clocks. This over-determined system will generate an over-determined set of data able to select a distance function with some uncertainty. Well, in spite of its interest, this problem is open for space-time distances.

But not all distant functions come from metric tensors. Thus the first problem to be solved for any proposed[35] distance function is if it is really the (geodesic) distance function of a metric. The constraints for this to be the case will constitute an important tool to improve uncertainties and to delimit parameter values.

I solved this problem some years ago. In order to express its solution, it is convenient to introduce some algebraic functions of the first and second partial derivatives of the symmetric bifunction $D(x,y)\equiv D$ proposed as distance function. Remember that at this level we have no metric at all, and that all subscripts in D denote partial derivatives. Define V^α_{abc} as the following function of second order derivatives of D:

$$V^\alpha_{\ell mn}\equiv\varepsilon^{\alpha\lambda\mu\nu}D_{\ell\lambda}D_{m\mu}D_{n\nu}\ ,$$

$V^{a\alpha}$ as the following combination of first and second ones:

$$V^{a\alpha}\equiv\varepsilon^{a\ell mn}\varepsilon^{\alpha\lambda\mu\nu}D_\ell D_\lambda D_{m\mu}D_{n\nu}\ .$$

and V^α as the quantity:

$$V^\alpha\equiv V^\alpha_{\ell mn}x^\ell y^m z^n\ ,$$

where x^ℓ, y^m, y^n are arbitrary independent directions. Introduce the two scalars

$$\Phi\equiv D_\lambda V^\lambda\ ,\Psi\equiv\varepsilon^{r\ell mn}V^\rho_{\ell mn}D_rD_\rho\ ,$$

and form the two quantities:

$$D^\alpha\equiv\frac{V^\alpha}{\Phi}\ ,\ D^{a\alpha}\equiv 3\frac{V^{a\alpha}}{\Psi}\ .$$

Then, we have:

Theorem 3.- (Structure theorem for distance functions) *The necessary and sufficient condition for a distance function* $D(x,y)$ *to be the geodesic distance function of a metric, is that its derivatives verify the identity:*

$$\begin{aligned}&D_{abc\rho}D^\rho+D_{(ab|\rho|}D_{c)m\sigma}D^{m\rho}D^\sigma\\&-D_{(ab|\rho|}D_{c)}D_{mn\sigma}D^{m\rho}D^nD^\sigma=0\ ,\end{aligned}$$

where the subscripts denote partial derivatives and D^a *and* $D^{a\alpha}$ *are the quantities just defined.*[36]

Note that, in all the above expressions Latin and Greek indices can be exchanged because the symmetric character of the distant function proposed. Another point to be noted, very little known, is that a function may be independent of some parameters meanwhile its algebraic expression depends unavoidably of them, if

[34] Besides this insufficiency, one could add the absolute lack of covariant methods of perturbations and approximations. After beautiful discourses about the importance of the role of the geometrization of physics by relativity, the most simple approximate calculation or deformation of a metric is made, without embarrassment, with non covariant analytical methods devoid of geometrical meanings. It is clear that the usual mathematical methods in relativity are not well adapted to relativity. But this subject will not be considered here.

[35] Or experimentally obtained with some uncertainty.

[36] Parentheses denote sum of the concerned terms for the circular permutations of the indices a, b, c.

these parameters are not scalars. It is here the case of D^α whose expression, from the definition of V^α, depends of the arbitrary vectors x^ℓ, y^m, z^n but that, as a function, it does not dependent of them: the partial derivatives of D^α with respect to any of these vectors vanish.

Once we know that a proposed distance function is truly[37] a geodesic distant function of some metric, the second problem to be solved is to obtain that metric. If, for example, as it may correspond to a natural experimental protocol, the proposed distance function $D(x, y)$ has been obtained from two local groups of points separated by non local distances, very probably the standard method of obtaining the metric by taking the limit when a point of one local group, say y, reaches a point of the other, say x, may have no sense, neither physical nor mathematical. For this reason, one needs to obtain the metric by means of a *finite* method. I did that some years ago, and the result is:

Theorem 4.- (Metric of a distance function) *In terms of the derivatives of the distance function* D *of a metric, the contravariant components* $g^{\alpha\beta}$ *of that metric at the point* x *are given by:*

$$g^{\alpha\beta} = D^\alpha D^\beta + D^{a\alpha} D^{b\beta} D_{ab\gamma} D^\gamma \ .$$

Note that the right hand side of this equation is a combination of a symmetric bifunction $D(x, y)$ and its partial derivatives, which, in general is other bifunction, meanwhile the left hand side is a function of the sole variable x. There is no contradiction: there are the conditions of theorem 3 that guarantee the downfall of the variable y in the right hand side.

This expression is very well adapted for the computation of the metric from an approximate distance function by means of finite difference methods.

The finite analogue of a connection remains a completely open problem. Perhaps this problem is avoidable, but it is not avoidable the quest for a finite version of curvature, because curvature is directly related to the energetic content of physical fields. With my friend Albert Tarantola (1949-2009), in the lustrum 2001-2005, we associated, to every four elements of a space-time, a finite object that seems to be a finite definition of curvature but, unfortunately, we were not able to prove that it is so.

The above results remain, for our needs, elemental. We must still develop them, plan a phenomenology of distance functions, and learn to ask them the analog finite questions that we are asking to a differential geometric structure.

[37] At the admissible uncertainties.

Acknowledgments This work has been supported by the Spanish ministries of "Ciencia e Innovación and "Economía y Competitividad", MICINN-FEDER projects FIS2009-07705 and FIS2012-33582.

References

[1] T. Adam et al. Measurement of the neutrino velocity with the OPERA detector in the CNGS beam. arXiv: 1109.4897v1 [hep-ex] 22 Sep 2011, 2011.

[2] B. Allès. Relativity accommodates superluminal mean velocities. *Phys. Rev. D*, 85:047501, 2012. arXiv:1111.0805v1 [hep-ph] 3 Nov 2011.

[3] N. Ashby. Relativity in the global positioning system. *Living Rev. Relativity*, 6(3):1–45, 2003. http://www.livingreviews.org/Articles/-Volume6/2003-1ashby/.

[4] T. B. Bahder. Navigation in curved space-time. *Am. J. Phys.*, 69(3):315–321, 2001.

[5] B. Coll. Light coordinates in relativity. In *Proc. ERE-1985, Trobades Científiques de la Mediterrània, Menorca, Spain, Ed. Ser. Pub. Univ. Barcelona*, pages 29–38, 1986. http://coll.cc (English version).

[6] B. Coll. Elements for a theory of relativistic coordinate systems. Formal and physical aspects. In *Proc. ERE-2000, Reference Frames and Gravitomagnetism, Valladolid, Spain, World Scientific*, pages 53–65, 2001. http://coll.cc.

[7] B. Coll. Relativistic positioning systems. arXiv:0601110v2 [gr-qc] 1 Feb 2006, 2006.

[8] B. Coll and J. A. Morales. Symmetric frames on Lorentzian spaces. *J. Math. Phys.*, 32(9):2450–2455, 1991.

[9] B. Coll and J. A. Morales. 199 causal classes of space-time frames. *Int. Jour. of Theo. Phys.*, 31(6):1045–1062, 1992.

[10] N. Collins. Science world in shock after CERN light speed claim. *The Telegraph, 23 Sep*, 2011.

 DOI: 10.2420/AF07.2013.35

[11] D. Lüst and M. Petropoulos. Comment on superluminality in general relativity. *Class. Quantum Grav.*, 29:085013, 2012. arXiv:1110.0813v1 [gr-qc] 4 Oct 2011.

[12] E. Newman and R. Penrose. An approach to gravitational radiation by a method of spin coefficients. *Journal of Mathemathical Physics*, 3(3):566–578, 1962.

[13] D. Overbay. Tiny neutrinos may have broken cosmic speed limit. *The New York Times, 22 Sep*, 2011.

[14] C. Rovelli. GPS observables in general relativity. *Phys. Rev. D*, 66:044017, 2002. arXiv: gr-qc/0110003 v1, 1 Oct 2001.

Acta Futura 7 (2013) 49-55
DOI: 10.2420/AF07.2013.49

Acta Futura

From emission to inertial coordinates: observational rule and inertial splitting

Bartolomé Coll,* Joan Josep Ferrando,† and Juan Antonio Morales-Lladosa‡
Departament d'Astronomia i Astrofísica, Universitat de València, 46100 Burjassot, València, Spain

Abstract. This contribution concerns relativistic positioning systems in Minkowski space-time. Attention is focused on the space-time *location problem* and the underlying *bifurcation problem* (where, for a given set of standard data, two locations are possible). The bifurcation problem for a user in inertial coordinates is solved extending the standard set of emission data with an observational rule. The covariant quantities (scalar, vectors and bivector) that give the solution to the location problem are expressed in terms of their time-like and space-like components with respect to the inertial reference observer.

1 Introduction

In Global Navigation Satellite Systems (GNSS), the location problem consists in allowing the users to determine its position in a specific terrestrial frame from the data they receive from four or more satellites of the system.

Formally, the location problem reduces to solving an algebraic system of non-linear equations (cf. Eq. (18) below). Each equation expresses that c times the time-of-flight of the electromagnetic signal is the range from the user to the satellite that broadcasts the signal, with c being the speed of light. From now on we take units so that c is equal to one.

One can solve the equations by an iterative procedure (starting from an approximate user position) or using some well adapted numerical algorithms, for instance those by Bancroft [2] and Krause [8], that are still in use in GPS. The non-uniqueness of the GPS solution, was pointed out by Schmidt [12] and studied by Abel and Chaffee [1, 3] and by Grafarend and Shan [7] who called it the *bifurcation problem*.

In current practical situations in present-day GNSS, the bifurcation problem may be solved by hand. If a user stays near the Earth's surface the right solution is the nearest to the Earth radius. However, if the user is near a satellite, or in extended GNSS (satellites broadcasting in all directions), or in general relativistic positioning systems, the bifurcation problem cannot be so easily avoided. Bifurcation situations will be present when a GPS satellite is located with respect to four Galileo emitters (see [10] and the contribution by D. Sáez and N. Puchades in this workshop [11]).

In order to monitor and to better understand the geometry of the configuration of the emitters and the real solution of the location problem in relativistic positioning, it is desirable to solve it analytically. Therefore we are looking for a general, exact, closed and covariant formula giving the solution.

Here, we are going to consider such formula for rel-

*Corresponding author. E-mail: bartolome.coll@uv.es
†joan.ferrando@uv.es
‡antonio.morales@uv.es

ativistic positioning systems in the Minkowski space-time, focusing the attention on:

(i) the *location problem*: to determine the inertial coordinates of a user from a given set of positioning data,

(ii) the *bifurcation problem*: to choose the *true* solution of the location problem where it admits two formal solutions,

(iii) the *space and time splitting* with respect to inertial coordinates of the quantities providing the user location.

Notation.- The main sign conventions and notation adopted in this paper are:

(i) g is the Minkowski space-time metric, with signature $(-,+,+,+)$.

(ii) η is the metric volume element of g, defined by $\eta_{\alpha\beta\gamma\delta} = -\sqrt{-\det g}\,\epsilon_{\alpha\beta\gamma\delta}$, where $\epsilon_{\alpha\beta\gamma\delta}$ stands for the Levi-Civita permutation symbol, $\epsilon_{0123} = 1$. The Hodge dual operator associated to η is denoted by an asterisk $*$. For instance, in index notation, if x, y, z are space-time vectors, one has

$$[*(x \wedge y \wedge z)]_\alpha = \eta_{\alpha\beta\gamma\delta} x^\beta y^\gamma z^\delta. \qquad (1)$$

where $\wedge$ stands for the *wedge or exterior product* (antisymmetrized tensorial product of antisymmetric tensors).

(iii) $i()$ denotes the interior or contracted product, that is, if x is a vector and T a covariant 2-tensor, one has $[i(x)T]_\nu = x^\mu T_{\mu\nu}$.

(iv) For a given inertial observer of unit velocity u, $u^2 \equiv g(u,u) = -1$, any vector x splits as:

$$x = x^0 u + \vec{x} \qquad (2)$$

where $x^0 = -x \cdot u \equiv -g(x,u)$ and $\vec{x} \in E_\perp$ are the time-like and space-like components of x relative to u, respectively, $E_\perp$ denoting the three-space orthogonal to u.

(v) For vectors $\vec{x}, \vec{y} \in E_\perp$, the vector or cross product is expressed as

$$\vec{x} \times \vec{y} = *(u \wedge \vec{x} \wedge \vec{y}), \qquad (3)$$

and, if $\vec{z} \in E_\perp$, the scalar triple product

$$(\vec{x} \times \vec{y}) \cdot \vec{z} \equiv (\vec{x}, \vec{y}, \vec{z}) \qquad (4)$$

is then given by

$$(\vec{x}, \vec{y}, \vec{z})\, u = *(\vec{x} \wedge \vec{y} \wedge \vec{z}). \qquad (5)$$

2 The location problem

Basically, a *relativistic positioning system* (RPS) is a set of four clocks A ($A = 1,2,3,4$), of world-lines $\gamma_A(\tau^A)$, broadcasting their respective proper times τ^A by means of electromagnetic signals. The set $\mathcal{R}$ of events reached by the broadcast signals is called the emission region of the RPS. The four proper times $\{\tau^A\}$ received at an event P of $\mathcal{R}$ are its *emission coordinates*.

A. Standard emission data set E.- Let $\{x^\alpha\}$ be any given specific coordinate system covering $\mathcal{R}$. Formally, the location problem consists in finding the user position in the coordinate system $\{x^\alpha\}$ as a function of its emission coordinates $\{\tau^A\}$ and a set of suitable data.

The set of the emitter world-lines referred to the coordinates $\{x^\alpha\}$, and the values of the emission coordinates received by a user, $E \equiv \{\gamma_A(\tau^A), \{\tau^A\}\}$, is called the *standard emission data set*. This set E is suitable data for the location problem only in a part of the emission region $\mathcal{R}$, called the central region (see below). Out of this region, an extended set of data including observational information is required.

In Minkowski space-time, the location problem is solved by finding the coordinate transformation, $x^\alpha = \kappa^\alpha(\tau^A)$, from emission $\{\tau^A\}$ to inertial $\{x^\alpha\}$ coordinates. The result may be expressed by means of a general, exact, closed and covariant formula (see [4, 5] and Eq. (10) below).

B. Quantities associated to the configuration of the emitters.- The *configuration of the emitters for an event* P is the set of the four events $\{\gamma_A(\tau^A)\}$ of the emitters at the emission times $\{\tau^A\}$ received at P. Let us denote by $x \equiv OP$ the position vector with respect to the origin O of the coordinate system $\{x^\alpha\}$. Consider a user at P that receives the broadcast times $\{\tau^A\}$, and denote by γ_A the position vectors of the emitters at the emission times, $\gamma_A \equiv O\gamma_A(\tau^A)$. Let us choose the fourth emitter as the *reference emitter* and name the other emitters the *referred emitters*, whose relative position vectors with respect to the reference emitter are respectively $e_a = \gamma_a - \gamma_4$ ($a = 1,2,3$). The trajectories followed by the light signals from the emitters $\gamma_A(\tau^A)$ to the reception event P are represented by the vectors $m_A \equiv x - \gamma_A$ that are null, $m_A^2 = 0$, and future pointing, that is, $u \cdot m_A < 0$ for any observer u (see Fig. 1).

The configuration of the emitters has associated the

 DOI: 10.2420/AF07.2013.49

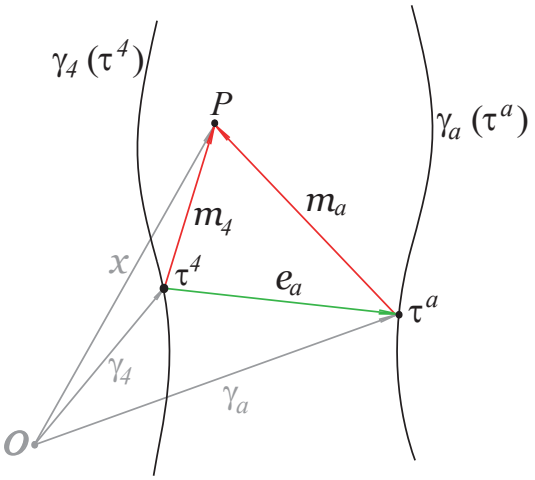


Figure 1. *Configuration of emitters. Taking γ_4 as the reference emitter, the position vector of the event P is m_4 and the relative positions e_α of the referred emitters γ_α are $e_\alpha = \gamma_\alpha - \gamma_4 = m_4 - m_\alpha$ ($\alpha = 1,2,3$). They determine the configuration scalars Ω_α, vector χ, and bivector H defined in the text.*

following quantities: the *configuration scalars*,

$$\Omega_\alpha = \frac{1}{2}\, e_\alpha^2 \quad (\alpha = 1,2,3) \tag{6}$$

which are the world function of the pairs of emitters $\{\gamma_\alpha, \gamma_4\}$, the *configuration vector*,

$$\chi \equiv *(e_1 \wedge e_2 \wedge e_3) \tag{7}$$

is orthogonal to the hyperplane containing the four events, and the *configuration bivector*,

$$H \equiv *(\Omega_1\, e_2 \wedge e_3 + \Omega_2\, e_3 \wedge e_1 + \Omega_3\, e_1 \wedge e_2). \tag{8}$$

All these quantities are computable from the only standard data set E because they are defined from e_α.

Here, we suppose that the emitter configuration is *regular* for the reception event P, that is, that the four emission events $\{\gamma_A(\tau^A)\}$ determine a hyperplane, the *configuration hyperplane* for P. Thus, we assume that $\chi \neq 0$. Non-regular emitter configurations, with $\chi = 0$, can occur in current GNSS as it was considered in [1, 3].

C. Covariant expression of the solution.- A RPS defines a function, $\Theta(P) = \{\tau^A\}$, that maps each event P in $\mathcal{R}$ to its emission coordinates $\{\tau^A\}$. Θ is called the *characteristic emission function* of the RPS and the *emission coordinate region* of the RPS on which Θ is invertible is denoted by $\mathcal{C}$. The *orientation* of a RPS at the event P is the orientation of the emission coordinates at P,

$$\hat{\epsilon} \equiv \mathrm{sgn}[*(d\tau^1 \wedge d\tau^2 \wedge d\tau^3 \wedge d\tau^4)] \tag{9}$$

and it coincides with the sign of the Jacobian determinant of Θ, $\hat{\epsilon}(P) = \mathrm{sgn}\, J_\Theta(P)$.

In the case of flat space-time, the explicit form of the coordinate transformation $x^\alpha(\tau^A)$ is given by (see [4, 5]):

$$x = \gamma_4 + y_* - \lambda\chi, \tag{10}$$

where

$$\lambda = \frac{y_*^2}{(y_* \cdot \chi) + \hat{\epsilon}\sqrt{\Delta}}. \tag{11}$$

and

$$y_* = \frac{1}{\xi \cdot \chi}\, i(\xi) H, \tag{12}$$

ξ being any vector transversal to the configuration, $\xi \cdot \chi \neq 0$, and Δ being the following quadratic invariant of H,

$$\Delta \equiv (y_* \cdot \chi)^2 - y_*^2\chi^2 = -\frac{1}{2} H_{\mu\nu} H^{\mu\nu} \tag{13}$$

that is non-negative, $\Delta \geqslant 0$. Furthermore, it holds that $H_{\mu\nu}(*H)^{\mu\nu} = 0$ and then Δ is the sole essential scalar invariant of the configuration bivector H (see [5]).

Note that y_* and Δ can be computed by using only the standard data set E.

The orientation $\hat{\epsilon}$ involves the unknown x and *is not always* computable from the only data set E. Next, we suppose that the user is able to observe the relative positions of the emitters on its celestial sphere. This observational data will allow the user to determine the orientation $\hat{\epsilon}$ at the reception event, and then, its location. The set E completed with the above observational data is called the *extended emission data set*.

D. Bifurcation problem.- The map Θ is not injective: for any RPS there always exist pairs of different events having the same emission coordinates (bifurcation problem). In fact, we have [4, 5]:

(i) When $\chi^2 \leqslant 0$ there is only one emission solution x, and therefore no bifurcation. To obtain the solution, take $\hat{\epsilon} = \mathrm{sgn}(u \cdot \chi)$ (where u is any future pointing time-like vector).

(ii) When $\chi^2 > 0$ there are two emission solutions. They only differ by their orientation $\hat{\epsilon}$. In this case, the standard data set E is insufficient to solve the bifurcation problem and an extended data set (including observational information) is necessary to determine $\hat{\epsilon}$,

and then to obtain the solution corresponding to the real user.

The region $\mathcal{C}^C \equiv \{x \in \mathcal{C} \,|\, \chi^2 \leqslant 0\}$ is called the *central region* of the RPS. The orientation $\hat{e}$ is constant on $\mathcal{C}^C$, and may be evaluated by considering only the data E. The bifurcation problem always appears in the time-like region $\mathcal{C}_t \equiv \{x \in \mathcal{C} \,|\, \chi^2 > 0\} = \mathcal{C} - \mathcal{C}^C$ and, in order to solve it, we need to calculate $\hat{e}$ at the 'true' space-time location x of the user.

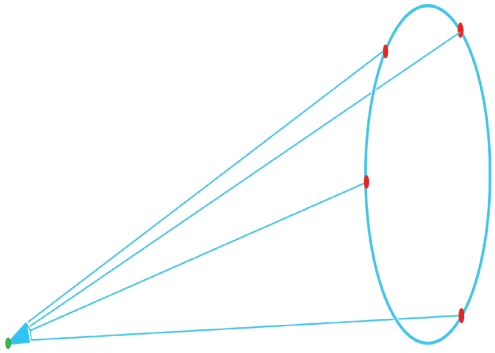


FIGURE 2. *The four emitters are on a circle on the celestial sphere of the user if, and only if, the user is on an event* x *where the Jacobian of* Θ *vanishes,* $J_\Theta(x) = 0$.

E. Observational rule to obtain $\hat{e}$.- It becomes apparent that the Jacobian hypersurface

$$\mathcal{J} \equiv \{x \,|\, J_\Theta(x) = 0\} \tag{14}$$

is of relevant interest in relativistic positioning, according to the following result by Coll and Pozo [6, 9]:

- *$\mathcal{J}$ consists in those events for which any user at them can see the four emitters on a circle on its celestial sphere.*

This result follows from the fact that four null directions on an event are linearly dependent if, and only if, the corresponding space-like directions for an observer at the event lie in a cone (see Fig. 2).

Moreover, the above result (being Lorentz invariant) suggests that at any event of the emission coordinate region $\mathcal{C} = \mathcal{R} - \mathcal{J}$, the orientation $\hat{e}$ could be obtained from the relative positions of the emitters on the celestial sphere of the user at this event. In fact, the following statement was proved elsewhere [5].

Observational rule to determine $\hat{e}$. In the coordinate region $\mathcal{C}$, the orientation $\hat{e}$ of the positioning system may be obtained as follows:

(i) consider the circle on the celestial sphere of the user containing the three referred emitters, $a = 1, 2, 3$,

(ii) turn this circle around its center in the increasing sense $1 \to 2 \to 3$ to orient the visual axis of the user by the rule of the right-hand screw,

(iii) if the fourth emitter $A = 4$ is in the spherical cap pointing out by this oriented axis, then the orientation is $\hat{e} = -1$, otherwise $\hat{e} = +1$.

By applying this observational rule, the users receiving the extended emission data set can determine the orientation $\hat{e}$ and their position in inertial coordinates.

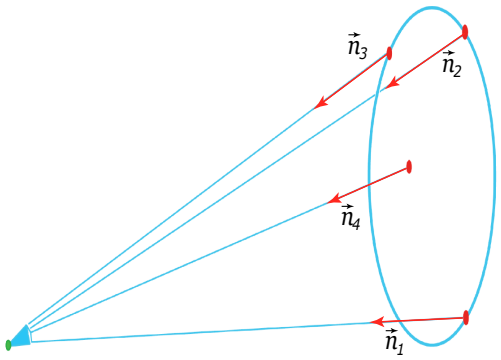


FIGURE 3. *A configuration with the reference emitter in the interior of the circle determined by the referred emitters and* $(\vec{n}_1, \vec{n}_2, \vec{n}_3) < 0$. *Any user at the reception event concludes that the orientation at its location* x *is positive, that is* $\hat{e} = +1$, *or equivalently,* $J_\Theta(x) > 0$.

Denote by $\vec{n}_A$ the unit vectors along the relative directions of propagation of the signals with respect to a user of unit velocity u,

$$m_A = -(u \cdot m_A)(u + \vec{n}_A). \tag{15}$$

It follows that directions $-\vec{n}_1$, $-\vec{n}_2$ and $-\vec{n}_3$ are contained in the cone confined by the circle, and then, if the reference emitter is in the interior of the circle, $-\vec{n}_4$ is in the interior of the cone and $\hat{e}$ is given by $\hat{e} = -\mathrm{sgn}[(\vec{n}_1, \vec{n}_2, \vec{n}_3)]$. Otherwise, $\hat{e} = \mathrm{sgn}[(\vec{n}_1, \vec{n}_2, \vec{n}_3)]$ (see Fig. 3).

3 Splitting of the solution

Consider the inertial observer associated to the specific inertial coordinate system $\{x^\alpha\}$, of unit velocity u, $u^2 = -1$. This section presents the decomposition, with respect to this inertial observer, of the quantities appearing in the transformation (10) from emission to inertial coordinates.

DOI: 10.2420/AF07.2013.49

A. Splitting of the emitter trajectories.- Relatively to u, the position vector of the emitter A is decomposed as

$$\gamma_A = t_A u + \vec{\gamma}_A, \qquad (A = 1,2,3,4) \tag{16}$$

where $t_A \equiv \gamma^0_A$ is the value of the inertial time of the observer u at the event $\gamma_A(\tau^A)$. The vectors m_A are written as

$$m_A \equiv x - \gamma_A = (x^0 - t_A)u + \vec{x} - \vec{\gamma}_A, \tag{17}$$

and because they are null and future pointing, we have

$$(x^0 - t_A)^2 = (\vec{x} - \vec{\gamma}_A)^2, \quad x^0 > t_A, \tag{18}$$

which constitute the relative expression of the 'navigation equations'.

The position vector of the emitter a with respect to the reference emitter is written as

$$e_a = \sigma_a u + \vec{e}_a \qquad (a = 1,2,3) \tag{19}$$

where $\sigma_a = t_a - t_4$ and $\vec{e}_a = \vec{\gamma}_a - \vec{\gamma}_4$. Then, the configuration scalars are expressed as:

$$\Omega_a = \frac{1}{2}(\vec{e}_a^{\,2} - \sigma_a^2). \tag{20}$$

B. Splitting of χ and H.- Substituting (19) in (7) and taking into account (3) and (5) one has

$$\begin{aligned}\chi = (\vec{e}_1, \vec{e}_2, \vec{e}_3)\, u + \sigma_1\, \vec{e}_2 \times \vec{e}_3 \\ + \sigma_2\, \vec{e}_3 \times \vec{e}_1 + \sigma_3\, \vec{e}_1 \times \vec{e}_2\end{aligned} \tag{21}$$

so that, $\chi = \chi^0 u + \vec{\chi}$ with

$$\chi^0 = (\vec{e}_1, \vec{e}_2, \vec{e}_3), \ \vec{\chi} = \frac{1}{2}\epsilon^{abc}\sigma_a\, \vec{e}_b \times \vec{e}_c. \tag{22}$$

Then, we see that:

(i) $|\chi^0|$ is the volume of the parallelepiped defined by the relative positions $\vec{e}_a$ of the referred emitters.

(ii) $\vec{\chi}$ represents a weighted vector-area. The area of the face generated by $\vec{e}_1$ and $\vec{e}_2$ is weighted with a complementary σ_3 factor. The other faces of the parallelepiped are weighted correspondingly.

As we have seen, on $\mathcal{C}^C$ the orientation is constant and is given by $\hat{\epsilon} = \mathrm{sgn}(u \cdot \chi)$ for any future pointing time-like vector u. In particular, if u is an inertial observer,

$$\hat{\epsilon} = -\mathrm{sgn}(\chi^0) = -\mathrm{sgn}[(\vec{e}_1, \vec{e}_2, \vec{e}_3)], \tag{23}$$

so that on the central region of a relativistic positioning system, the orientation $\hat{\epsilon}$ is $+1$ or -1 if $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ is, respectively, negative or positive.

On the other hand, the configuration bivector can be written as

$$H = u \wedge \vec{S} - *(u \wedge \vec{B}), \tag{24}$$

where

$$\vec{S} = -i(u)H, \quad \vec{B} = -i(u) * H \tag{25}$$

are, respectively, the electric-like and magnetic-like parts of H relative to u. They are expressed as:

$$\vec{S} = \Omega_1\, \vec{e}_2 \times \vec{e}_3 + \Omega_2\, \vec{e}_3 \times \vec{e}_1 + \Omega_3\, \vec{e}_1 \times \vec{e}_2 \tag{26}$$

$$\begin{aligned}\vec{B} \;=\; & \Omega_1(-\sigma_2\vec{e}_3 + \sigma_3\vec{e}_2) + \\ & \Omega_2(-\sigma_3\vec{e}_1 + \sigma_1\vec{e}_3) + \\ & \Omega_3(-\sigma_1\vec{e}_2 + \sigma_2\vec{e}_1),\end{aligned} \tag{27}$$

as it follows from (8), (3) and the identity $**F = -F$ for any bivector F. Note that $\vec{S}^2 \geqslant \vec{B}^2$ because $\Delta = \vec{S}^2 - \vec{B}^2 \geqslant 0$. In [5] it was stated that Eq. (10) allows a user to locate itself at the hypersurface $\mathcal{J}$ of vanishing Jacobian, and also, that this region is invariantly characterized by the condition $\Delta = 0$. Then, an inertial observer u will recognize these events by testing where the electric and magnetic parts of H become equimodular.

On the other hand, it always occurs that $\vec{S} \cdot \vec{B} = 0$, because the invariant $H_{\mu\nu}(*H)^{\mu\nu}$ identically vanishes.

C. Splitting of y_* and λ.- From Eq. (12), the relative decomposition of the quantity y_* in (10) is obtained by splitting the vector $i(\xi)H$. To begin with, notice that the transversal vector ξ can be chosen so that its time-like component, ξ^0, is equal to one, that is $\xi = u + \vec{\xi}$. Thus, the transversality condition says that $\chi^0 \neq \vec{\xi} \cdot \vec{\chi}$ and by Eqs. (22) and (4) it is written as

$$(\vec{e}_1, \vec{e}_2, \vec{e}_3) \neq \frac{1}{2}\epsilon^{abc}\sigma_a\, (\vec{\xi}, \vec{e}_b, \vec{e}_c), \tag{28}$$

and then, from (24), we have

$$\begin{aligned} i(\xi)H \;&=\; -\vec{S} - (\vec{\xi} \cdot \vec{S})u + *(\xi \wedge u \wedge \vec{B}) \\ &=\; -(\vec{\xi} \cdot \vec{S})u - \vec{S} - \vec{\xi} \times \vec{B}. \end{aligned} \tag{29}$$

Therefore, the splitting of $y_* = y^0_*\, u + \vec{y}_*$ is given by:

$$y^0_* = -\frac{\vec{\xi} \cdot \vec{S}}{D}, \quad \vec{y}_* = -\frac{\vec{S} + \vec{\xi} \times \vec{B}}{D}, \tag{30}$$

where $\vec{S}$ and $\vec{B}$ are obtained from (26) and (27), and

$$D \equiv \vec{\xi} \cdot \vec{\chi} - (\vec{e}_1, \vec{e}_2, \vec{e}_3) \neq 0. \tag{31}$$

It is worthy to say that we still have a complete freedom in choosing the transversal vector ξ. This freedom may be exploited in order to make the expression of the solution as simple as possible. For example, we can get $\xi = u$ whenever $\chi^0 = (\vec{e}_1, \vec{e}_2, \vec{e}_3) \neq 0$. Then $y^0_* = 0$ and y_* takes the simple form:

$$y_* = \frac{\vec{e}_a^{\,2} - \sigma_a^2}{2\,(\vec{e}_1, \vec{e}_2, \vec{e}_3)}\, \vec{e}_{a+1} \times \vec{e}_{a+2}. \tag{32}$$

Eq. (10) gives the solution x for the location problem: Eqs. (22), (26), (27), (30) and (31) allow to determine $\chi = \{\chi^0, \vec{\chi}\}$, $\Delta = S^2 - B^2$ and $y_* = \{y^0_*, \vec{y}_*\}$, and then λ,

$$\lambda = \frac{-(y^0_*)^2 + \vec{y}_*^{\,2}}{(-y^0_*\chi^0 + \vec{y}_* \cdot \vec{\chi}) + \hat{e}\sqrt{\vec{S}^2 - \vec{B}^2}}. \tag{33}$$

Remember that, in any region of $\mathcal{C}$, the orientation $\hat{e}$ of the RPS may be determined from the observational rule that we have explained, and that, in addition, this orientation can be calculated in the central region from the sole standard data set E.

4 Summary

(i) The standard emission data set E is generically insufficient to locate a user of a positioning system in an inertial system.

(ii) It is the time-like region $\mathcal{C}_t$ ($\chi^2 > 0$) where the *bifurcation problem* cannot be solved from the sole standard data set E.

(iii) The data of the orientation joint to the standard emission data set E solves completely the bifurcation problem, and then the location problem in relativistic positioning.

(iv) The covariant solution of this location problem may be expressed in terms of the space-like and time-like components of the emission data with respect to the reference inertial observer.

Acknowledgments This work has been supported by the Spanish ministries of "Ciencia e Innovación" and "Economía y Competitividad" MICINN-FEDER projects FIS2009-07705 and FIS2012-33582.

References

[1] J. S. Abel and J. W. Chaffee. Existence and uniqueness of GPS solutions. *IEEE Transactions on Aerospace and Electronic Systems*, 27(6):952–956, 1991.

[2] S. Bancroft. An algebraic solution of the GPS equations. *IEEE Transactions on Aerospace and Electronic Systems*, AES-21(7):56–59, 1985.

[3] J. W. Chaffee and J. S. Abel. On the exact solutions of the pseudorange equations. *IEEE Transactions on Aerospace and Electronic Systems*, 30(4):1021–1030, 1994.

[4] B. Coll, J. Ferrando, and J. Morales-Lladosa. Positioning systems in Minkowski spacetime: from emission to inertial coordinates. *Classical Quantum Gravity*, 27:065013, 2010.

[5] B. Coll, J. Ferrando, and J. Morales-Lladosa. Positioning systems in Minkowski spacetime: Bifurcation problem and observational data. *Physical Review D*, 86:084036, 2012.

[6] B. Coll and J. M. Pozo. Constructions in 4D: Algebraic properties and special observers. Lecture delivered in Salamanca, 2005.

[7] E. W. Grafarend and J. Shan. A closed-form solution of nonlinear pseudo-ranging equations (GPS). *Artificial Satellites, Planetary Geodesy No 28*, 31(3):133–147, 1996.

[8] L. O. Krause. A direct solution to GPS type navigation equations. *IEEE Transactions on Aerospace and Electronic Systems*, AES-23(2):225–232, 1987.

[9] J. M. Pozo. Some properties of emission coordinates. *arXiv:gr-qc/0601125*, 2006.

[10] N. Puchades and D. Sáez. Relativistic positioning: four-dimensional numerical approach in Minkowski spacetime. *Astrophysics and Space Science*, 341(2):631, 2012.

[11] D. Sáez and N. Puchades. Relativistic positioning systems: Numerical simulations. In *Proceedings of the on Relativistic Positioning Systems and their Scientific Applications, Brdo near Kranj, Slovenia, 19-21 September 2012*. Acta futura, 2012.

 DOI: 10.2420/AF07.2013.49

[12] R. O. Schmidt. A new approach to geometry of range difference location. *IEEE Transactions on Aerospace and Electronic Systems*, AES-8(6):821–835, 1972.

DOI: 10.2420/AF07.2013.49

Acta Futura 7 (2013) 57-65
DOI: 10.2420/AF07.2013.57

Acta Futura

The concept of Autonomous Basis of Coordinates

Andrej Čadež*
Univ. of Ljubljana, Ljubljana, Slovenia

Abstract. The tremendous development of clock accuracy since the first introduction of the GPS positioning system, together with the ever more accepted relativistic notion of space-time, requires a new look at the function and at the basic theoretical concepts in introducing a system of coordinates. Based on Newton's notion of geometry as being driven by dynamics and Laplace's understanding of dynamics as being able to describe motion, we introduce the covariant notion of an Autonomous Basis of Coordinates, which is tied to a constellation of freely moving satellites in space. Each satellite is assumed to carry its own (ideal) proper time clock and is able to send timing signals (emission coordinates) into space, and receive those signals (emission coordinates) from all other satellites in the constellation. We demonstrate that exchange of timing signals – emission coordinates – between satellites provides the measure of local space-time curvature, and thus gives enough information for building a local quasi Minkowski system of coordinates in which dynamics of free body motion is describable with the precision limited only by the accuracy of timing information exchanged between satellites.

1 Coordinates through time

The notion of coordinates as labels for positions in space (on the Earth) is obviously a very ancient one. Observations of certain relations (distance, direction, shortest path) between positions with different coordinates (names) led to the development of geometry, while the observation of changes in our world gave impetus to begin writing history. Both geometric and historical observations led to the understanding of certain symmetries in space that seemed to be independent of time. The basic axioms that followed from such observations could be stated as follows:

- Earth, planets and the sky make up the permanent Universe.
- There are permanent features in the Universe that can be described by maps in sufficient detail to be recognized by seekers.
- Symmetries of the sky are related to those of the Earth and vice versa

These axioms were sufficient for the ancients to draw maps of the known world and successfully use them in travels and expeditions between distant places. The new understanding and usefulness of coordinates that emerged in the 17^{th} century with Descartes, Galileo, Newton, Laplace, Lagrange ... established coordinates as sets of numbers that could be used in algebraic expressions and allowed the description of symmetries and dynamics in terms of physical causes. With the invention of accurate clocks by Harrison [7], it became possible to connect symmetries of the sky and the Earth in a simply applicable way. By reading positions of stars, it

*Corresponding author. E-mail: andrej.cadez@fmf.uni-lj.si

became possible to find observer's coordinates with unprecedented precision. Until the second half of the 20th century the only limit to position accuracy appeared to be limited by clock accuracy. This ceased to be a limitation, as the accuracy of best available clocks has increased by seven (!) orders of magnitude in the last fifty years. It became thus possible to demonstrate that the Earth is not rotating as uniformly as expected and is not as accurately elliptic or homogeneous as was thought before. Rotation, shape and gravitational potential of the Earth now emerge as parameters limiting position accuracy. The accuracy by which time and position can be measured amply demonstrated the validity of general relativity and necessitated the introduction of relativistic corrections to the Newtonian concept of space. These axioms were sufficient for the ancients to draw maps of the known world and successfully use them in travels and expeditions between distant places. The new understanding and usefulness of coordinates that emerged in the 17th century with Descartes, Galileo, Newton, Laplace, Lagrange ... established coordinates as sets of numbers that could be used in algebraic expressions and allowed the description of symmetries and dynamics in terms of physical causes. With the invention of accurate clocks by Harrison [7], it became possible to connect symmetries of the sky and the Earth in a simply applicable way. By reading positions of stars, it became possible to find observer's coordinates with unprecedented precision. Until the second half of the 20th century the only limit to position accuracy appeared to be limited by clock accuracy. This ceased to be a limitation, as the accuracy of best available clocks has increased by seven (!) orders of magnitude in the last fifty years. It became thus possible to demonstrate that the Earth is not rotating as uniformly as expected and is not as accurately elliptic or homogeneous as was thought before. Rotation, shape and gravitational potential of the Earth now emerge as parameters limiting position accuracy. The accuracy by which time and position can be measured amply demonstrated the validity of general relativity and necessitated the introduction of relativistic corrections to the Newtonian concept of space.

In view of these developments, ESA's Advanced concepts team initiated a study "Mapping the Spacetime Metric with a Global Navigation System". The invitation to the project stipulated a novel perspective on the meaning and definition of a system of coordinates in the 4-dimensional spacetime. The first step was to rewrite the basic code, such as used by the GPS system, into the fully relativistic form based on emission coordinates introduced by [1]. It was demonstrated in [4] that position determination from emission coordinates data is numerically equally efficient yet more accurate than the classical post-newtonian approach. However, the main advantage of the fully relativistic formalism is in its conceptual simplicity and accordance with general structure of spacetime, which, as we now know, must be described by general relativity.

The second part of the project [5] addressed the self consistency of a relativistic system of coordinates defined by means of emission coordinates. Namely, we investigated the accuracy to which the redundant number of emission coordinates from the constellation of satellites gives the same four spacetime coordinates of a given event, especially in light of discussions considering the possibility of intersatellite communication [8].

2 Intersatellite communication and the ABC concept

The introduction of coordinates in the four dimensional spacetime requires a consideration of basic premises on which a system of coordinates is based. We propose that the the system of coordinates, that is most suitable to describe the local spacetime manifold, be defined in such a way, that generators of equations of motion (Lagrangian, Hamiltonian) can be correctly expressed in the chosen system of coordinates, in accordance with Laplace's famous quote: "An intelligence which at a given instant knew all the forces acting in nature and the position of every object in the universe – if endowed with a brain sufficiently vast to make all necessary calculations – could describe with a single formula the motions of the largest astronomical bodies and those of the smallest atoms. To such an intelligence, nothing would be uncertain; the future, like the past, would be an open book."

In this respect the International Celestial Reference Frame (ICRF), which is defined as a quasi inertial reference frame centered at the barycenter of the Solar System, is a kinematic frame, formed on the assumption that space time is flat from here to distant quasars. This assumption appears to be satisfied locally to a high degree of accuracy, but it is not tested dynamically within the definition of the reference frame. This attitude does not satisfy the principle of general covariance, which considers spacetime as a smooth four dimensional manifold with Minkowski local structure, possessing no prior geometry. Spacetime seems to be flat only on grand

 DOI: 10.2420/AF07.2013.57

scales encompassing large volumes [6]. Therefore, we propose to base the definition of spacetime coordinates on the ancient requirement for its functionality, i.e. a system of coordinates should allow accurate (as accurate as theoretically possible) orientation in the part of spacetime covered by its coordinates. The vehicle to apply this principle is Laplace's understanding of dynamics. Thus we propose that the global navigation system (GNS) of coordinates be based solely on dynamics of all the satellites composing the GNS constellation of satellites (GNSS).

The basic idea of ABC is that all satellites of a GNSS move according to laws of motion that follow from a Hamiltonian, common to all satellites. Ideally, satellites move only under the influence of gravity and are thus following geodesics of spacetime. It is expected, that additional forces acting on satellites, such as collisions with micrometeorites, solar radiation pressure, solar wind pressure, etc. are small and can be counterbalanced in a drag-free mode. Since light, which propagates emission coordinates, is tracing spacetime geodesics as well, one can consider sets of spacetime events along light rays connecting world lines of two freely falling GNSS satellites, as two dimensional geodetic manifolds with specific position in the fabric of spacetime. I would like to call **panels** such manifolds. All panels, formed by pairs of all satellites in a GNSS constellation, form a boundary of the spacetime 4-manifold which can be defined as the **interior** of the constellation. We intend to show how panels can be constructed on the basis of emission coordinates exchanged between pairs of satellites and how panels can be combined to form the boundary of the interior. Once the geometry of the boundary of the interior is defined by measurement, Einstein's equations can be used to extend coordinates inward by using multipole expansions as described by [2] and [10].

Tiles and Panels

A panel is a two sheeted geodesic two manifold formed by lightlike geodesics connecting world lines $\mathcal{C}_A$ and $\mathcal{C}_B$ of two freely falling satellites A and B. Lightlike geodesics from A to B form the first sheet, and those from B to A the second sheet of the panel, Fig.1. Panels are usually spatially thin, since their relative thickness is generally proportional to the mutual acceleration of supporting satellites. Since any spacetime is localy Minkowskian, a panel, spanning a sufficiently small region of spacetime can be embedded in a local Minkowski frame. Such a small panel or section of a large panel may be called a **tile**. It is convenient to choose the local frame of a tile S_{AB} in such a way that the two timelike geodesics A and B have opposite velocities as in Fig.1. The world lines $\mathcal{C}_A$ and $\mathcal{C}_B$ can be expressed as follows:

$$\mathcal{C}_A : \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma c(\tau_A + \Delta\tau) \\ \gamma v(\tau_A + \Delta\tau - \tau_0) \\ -y_0 \\ 0 \end{pmatrix} \tag{1}$$

$$\mathcal{C}_B : \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \gamma c(\tau_B - \Delta\tau) \\ -\gamma v(\tau_B - \Delta\tau - \tau_0) \\ y_0 \\ 0 \end{pmatrix}, \tag{2}$$

where $\gamma = \frac{1}{\sqrt{1-v^2/c^2}}$. The world lines $\mathcal{C}_A$ and $\mathcal{C}_B$ define the two timelike edges of the tile and the spacelike geodesic (coeval in the proper frame of the tile) connecting lower ends $\{\mathcal{C}_A(\tau^i+\Delta\tau), \mathcal{C}_B(\tau^i-\Delta\tau)\}$ can be called the base of the tile and $\{\mathcal{C}_A(\tau^f+\Delta\tau), \mathcal{C}_B(\tau^f-\Delta\tau)\}$ the top of the tile. Note the choice of the local Minkowski coordinate system orientation: x axis points in the direction of mutual velocity and the y axis points along the direction of shortest distance between the two satellites and $z = 0$.

Tile parameters velocity v, proper time misalignment $\Delta\tau$, proper time mark of closest approach τ_0 and closest distance y_0 are constants of motion of geodesics supporting the tile, thus we call them mutual constants of motion. They can be determined from exchange of emission coordinates between A and B as follows: Let ${\tau_A}^*(\tau_B)$ be the the emission coordinate (i.e. proper time of emission at A) of A received by B at τ_B and ${\tau_B}^*(\tau_A)$ the emission coordinate of B received by A at τ_A. Using the null property of emission coordinates and equations 1 and 2, one calculates the functions ${\tau_A}^*(\tau_B)$ and ${\tau_B}^*(\tau_A)$. In order to express the result in a simple way, we introduce $v = \tanh(\xi)$ to obtain:

$$\begin{aligned} {\tau_A}^*(\tau_B) &= (\tau_B - \tau_0 - \Delta\tau)\cosh(2\xi) \\ &- \sqrt{(\tau_B - \tau_0 - \Delta\tau)^2 \sinh^2(2\xi) + 4y_0^2} \\ &+ \tau_0 - \Delta\tau \end{aligned} \tag{3}$$

$$\begin{aligned} {\tau_B}^*(\tau_A) &= (\tau_A - \tau_0 + \Delta\tau)\cosh(2\xi) \\ &- \sqrt{(\tau_A - \tau_0 + \Delta\tau)^2 \sinh^2(2\xi) + 4y_0^2} \\ &+ \tau_0 + \Delta\tau \end{aligned} \tag{4}$$

The constant $\Delta\tau$ is readily determined by observing that ${\tau_A}^*(\tau+\Delta\tau)+\Delta\tau = {\tau_B}^*(\tau-\Delta\tau)-\Delta\tau = s(\tau)$, i.e. the

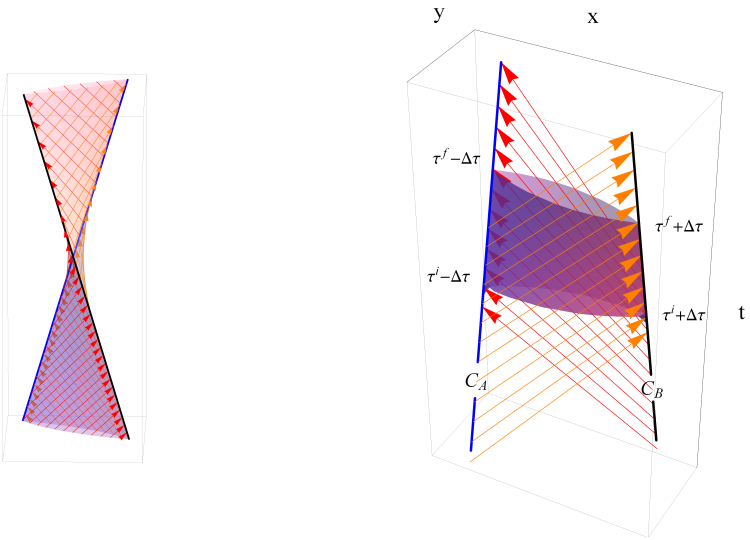


Figure 1. *A panel (left) is a two sheeted geodesic manifold formed by lightlike geodesics connecting two timelike geodesics. A tile (right) is a panel supported by two geodesics moving with opposite velocities with respect to its proper local Minkowski frame. Lightlike geodesics colored in orange form the front sheet of the tile and transport emission coordinate $\tau_A{}^*$ to satellite B, while the lightlike geodesics colored in red transport $\tau_B{}^*$ to A, and form the back sheet of the tile. Timelike sides of a tile are specified by proper time intervals of both satellites and are denoted as $\{\tau^i-\Delta\tau,\tau^f-\Delta\tau\}$ and $\{\tau^i+\Delta\tau,\tau^f+\Delta\tau\}$ respectively. The edges $\mathcal{C}_A(\tau^i-\Delta\tau)$ and $\mathcal{C}_B(\tau^i+\Delta\tau)$ are coeval in the proper frame of the tile, thus $\Delta\tau$ may be considered as the proper time misalignment of the two clocks.*

curves $\tau_A{}^*(\tau)$ and $\tau_B{}^*(\tau)$ are identical hyperbolas, but displaced. The other three constants follow by fitting $s(\tau)$ to a parabola[1] and solving the parabola coefficients for ξ, τ_0 and y_0.

Joining tiles

Each pair of satellites generates panels which possess a definite position in spacetime. As satellites move in time, newly emerging panels tile the spacetime and build up the boundary of the interior. Three satellites support three panels, four satellites support four panels etc. Four panels are sufficient to close the boundary of a four dimensional tube in spacetime, while more than four satellites create a redundant number of panels creating a multi layer boundary of the four dimensional spacetime tube. The question is how to join panels and represent them in a common system of coordinates. Taking into account that, to the best of our knowledge, the spacetime is locally flat, we can imagine breaking panels into tiles, small enough to be treated as in flat spacetime. Tile has a known – measurable – geometry and fits to its neighbors. Fitting tiles into the spacetime tube is just like fitting LEGO blocks according to the instruction book that comes with the set. The only difference between fitting LEGO blocks and tiles is that LEGO blocks are moved by generators of the Galileo group in a flat space, while tiles are moved by generators of the local Poincare group in a curved spacetime.

We first consider a tiling procedure in the locally flat spacetime volume, starting by joining three adjacent tiles P_{12}, P_{23}, P_{32}, generated by orbits $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{C}_3$. Since the base of a tile is coeval, we start by translating the three tiles, so that their bases join and form a triangle. If this was the first triangle formed, one can align the coordinates of the common Minkowski frame so that this triangle is in the $x-y$ plane. At this stage the position of tiles in the common Minkowski frame looks like the one shown in Fig.2. Each geodesic is still represented by edges of two adjacent tiles. In order to join them, one simply needs to apply boosts in directions perpendicular to tile's bases, which by now have been oriented. The procedure is somewhat tedious but straightforward and is helped by taking into account the laws of motion along geodesics in flat spacetime which lead to the following statements: a) the sum of tile velocities must point along the common time axis and b) the total angular momentum of tiles is perpendicular to the $x-y$

[1] Parabola is a good enough approximation of a hyperbola on a small tile.

 DOI: 10.2420/AF07.2013.57

plane. The angular momentum of a tile with respect to the common Minkowski frame is the sum of its spin and orbital angular momentum, which are defined as:

$$s^{\mu}_{T} = \sum_{i=1}^{2} \varepsilon^{\mu}{}_{\nu\lambda\sigma} U^{\nu} \left(\mathcal{C}^{\lambda}_{i} \dot{\mathcal{C}}^{\sigma}_{i} - \mathcal{C}^{\sigma}_{i} \dot{\mathcal{C}}^{\lambda}_{i} \right)$$
$$l^{\mu}_{orb} = 2\varepsilon^{\mu}{}_{0\lambda\sigma} \left(\mathcal{C}^{\lambda}_{T} U^{\sigma} - \mathcal{C}^{\sigma}_{T} U^{\lambda} \right), \qquad (5)$$

where $\mathbf{C}_T(\tau)$ is the worldline (geodesic) of the tile origin with respect to the common Minkowski frame origin and $\mathbf{U}$ is the velocity four vector of $\mathbf{C}_T$. The above construction gives a complete description of motion of satellites #1, #2 and #3 on the time interval considered, with respect to their common Minkowski frame, defined by the base triangle T_{123}. One can repeat the same procedure with tiles supported by satellites #2, #3 and #4 to obtain a triangle T_{234}, which shares the side S_{23} with T_{123}, and so on, until all the four triangle faces of a tetrahedron are obtained. Finally the faces are joined into a tetrahedron oriented in the local Minkowski frame of triangle T_{123}. This completes the joining of all panels supported by the four satellites and coordinatizes a four dimensional structure V_{1234} enclosing a section of the spacetime manifold.

As long as spacetime was believed to be flat, it was expected that the complete spacetime, from here to distant quasars, can be filled with exactly equal tetrahedral structures.

The above procedure to join tiles to build triangles, triangles to build tetrahedra and tetrahedra to build the four-dimensional simplex structures is based on geometry, which requires tiles being translated to fit their bases, and on dynamics, which, in flat space, assures that tiles, supported by timelike geodesics, can be joined along timelike edges, if their total angular momentum points perpendicular to the base triangle. This second condition may not be met in curved space. As a consequence, tiles can in curved space only be matched along two timelike edges and may leave a gap along the remaining edge as shown in Fig.3. This gap is the direct consequence of curvature and can be expressed with the Riemann tensor as

$$\delta U^{\mu} = \frac{1}{2} R^{\mu}{}_{\nu\lambda\sigma} U^{\nu} a^{\lambda} b^{\sigma}, \qquad (6)$$

where $\mathbf{U}$ is the timelike edge of the tile along the gap, which is transported along the bases of the three panels, $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$, and $\delta\mathbf{U}$ is the difference across the gap. In principle one can always make the four volume of the structure V_{1234} sufficiently small to make the gap negligible. So one can continue adding small tetrahedrons, such as V_{1234}, to adjacent sides, just as a good tiler can tile a crooked wall with small square tiles. However, the infinitesimally small misalignments eventually add up and cause straight lines to appear curved.

The ABC proposal

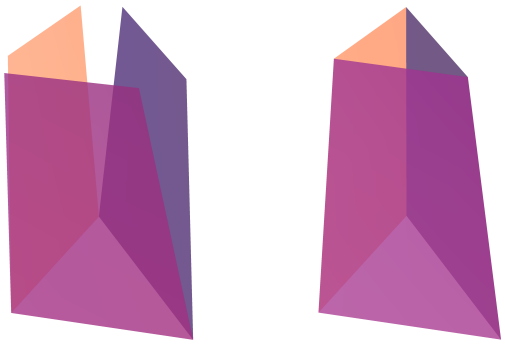
The Autonomous Basis of Coordinates proposal is based on the idea that spacetime, whether flat or curved, can in principle be tiled [9].

Flat spacetime (if it would exist) could be tiled with a periodic structure of equal tiles and would be coordinatized by four numbers counting the number of tiles from the origin in direction of time and in the three spatial directions. The usefulness of such a Cartesian coordinatization comes from our ability to express the laws of motion in such a system of coordinates. For a long time the Cartesian coordinatization was considered as a natural realization of inherent spacetime symmetries, and was thus taken as an axiom needing no testing. Today, this belief can no longer be supported as practical, since tiles and panels supported by GPS satellites can be measured with sufficient precision to actually demonstrate the curvature of spacetime. We propose to anchor the ABC system of coordinates on the tiling of spacetime panels provided by GNSS satellites. Since the spacetime around the Earth, where GNSS satellites are moving, is sufficiently curved to keep satellites indefinitely orbiting the Earth, and it is not practical to increase the density of satellites so that they would define very small tiles, it makes sense to join the small tiles in panels, which can fit the curvature of spacetime for as as long and as far, as one can find agreement between measured and dynamically predicted panel characteristics. As a first step in constructing such panels we start with the Schwarzschild space time, which may conveniently be coordinatized by isotropic coordinates with the metric[2]:

$$ds^2 = -\left(\frac{1 - \frac{M}{2r}}{1 + \frac{M}{2r}} \right)^2 dt^2 + \left(1 + \frac{M}{2r} \right)^4 (dx^2 + dy^2 + dz^2), \qquad (7)$$

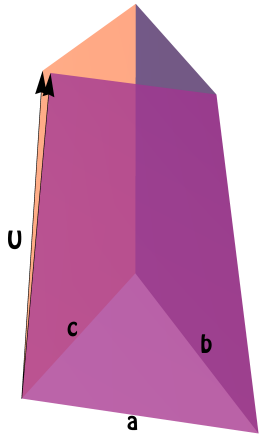
where $r = \sqrt{x^2 + y^2 + z^2}$. A typical panel, supported by two satellites is represented in Fig4. Since the Schwarzschild spacetime does not possess translational

[2]Units $c = G = 1$ are used.



Joining three tiles. First join bases at $t_{initial}$ *and obtain a triangle, whose internal angles can be calculated. In left panel each geodesic is represented in two panels. In order to join the two images of a geodesic into a single line, one boosts each panel with a velocity perpendicular to its lower edge.*

FIGURE 2.



Three tiles do not automatically fit in the tangent Minkowski space of a curved manifold. The vector **U**, *when parallel transported along the base triangle, suffers a geodesic deviation* $\delta\mathbf{U}$ *that shows as a gap between the tips of the two images of* **U**.

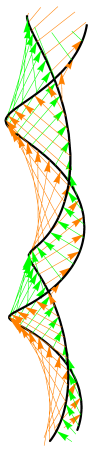
FIGURE 3.

symmetry, panels can not be translated or boosted, but can only be rotated. This fact actually simplifies the procedure of panel fitting, since two panels with a common edge can always be joined along this edge by applying a rotation.

A Schwarzschild panel can be characterized by nine mutual constants of motion belonging to its pair of satellites (A and B). We use the following quantities as mutual constants: semi-major axes a_A, a_B, orbital eccentricities ε_A, ε_B, mutual inclination ι, longitude of line of nodes of A's with respect to B's orbit Ω, longitude of A's periastron with respect to line of nodes ω, proper times of first periastron passage of the two satellites τ_A^{peri}, τ_B^{peri}. We could show, to some surprise, that mutual constants can be determined by each of the satellites solely by fitting the function $\tau_A^*(\tau_B)$ (or $\tau_B^*(\tau_A)$) to theoretically predicted relations, with mutual constants as fitting parameters [5]. The fitting procedure was tested numerically on different sets of simulated data, stretching from 1 to 4 orbital periods in time. Our simulated data imitated data that would be obtained from Galileo GNSS and included timing noise simulations expressed in terms of Allan deviation. In all cases the fitting procedure returned values of mutual constants, which allowed the precision of orbital prediction to be limited essentially by timing accuracy of satellite clocks[3]. A sample of

[3]This statement reflects a general experience with our numerical

 DOI: 10.2420/AF07.2013.57



A panel in the Schwarzschild spacetime displayed in isotropic quasi Cartesian coordinates

Figure 4.

our results showing the accuracy of orbital axis determination and the uncertainty of timing origin as functions of Allan deviation of satellite's clocks is shown in Fig.5.

Thus, clocks, and clocks only, define the measure of panels. In this respect it is worth noting that each panel is measured by two satellites. Therefore, mutual constants obtained by both satellites must be equivalent, which is a constraint defining the meaning of proper time. A disagreement can only arise, in the framework of general relativity, if satellite clocks fail to measure proper time, thus, the comparison of mutual constants offers a rigorous check of clock rates. Atomic clocks, with their fundamentally defined rate, thus provide the fundamental gauge for spacetime geometry.

The spacetime around the Earth is not quite Schwarzschild, since it is perturbed by the Sun, Moon, Earth obliquity, rotation, planets... As a result, Schwarzschild panels do not fit indefinitely in the spacetime of our neighborhood. Sooner or later geodesic deviation, such as described by equation (6), with Riemann curvature replaced by its deviation from the Schwarzschild one, would become apparent. One can imagine to coordinatize our spacetime with sufficiently small Schwarzschild panels, which would eventually show deviations from large Schwarzschild panels. After deviations have built into measurable discrepancies, one can determine deviations of the Riemann tensor from the assumed one, and correct the metric accordingly. In this way one can construct new, better fitting panels. The procedure can be repeated until the smallest measurable mismatch between panels is observed.

The ABC system of coordinates is thus a system of coordinates based on panels that fit the local spacetime geometry to the highest precision measurable.

In order to get an idea of the order of magnitude of expected mismatch between panels we note that the Schwarzschild Riemann tensor components can be expressed in the form:

$$R^{\mu}{}_{\nu\lambda\sigma} = \frac{GM}{c^2 r^5 (1+\frac{M}{2r})^2} P^{\mu}{}_{\nu\lambda\sigma}(x,y,z)\,, \qquad (8)$$

where $P^{\mu}{}_{\nu\lambda\sigma}(x,y,z)$ are quadratic polynomials, for example $P^{0}_{101} = 2x^2 - y^2 - z^2$. Taking $\frac{GM_{\text{Earth}}}{c^2 r^3}$ as the rough magnitude of Riemann tensor components due to Earth mass at the the distance of the Galileo satellites, we obtain $R^{\mu}{}_{\nu\lambda\sigma} \sim 3.5 \times 10^{-25}$ m^{-2}. Three satellites orbiting in a plane in an equilateral triangle formation, form a triangle with the side $a \sim 4 \times 10^4$ km. According to equation (6), three Minkowski tiles would be mismatched after 1 second by approximately $\delta U \sim 3.5\times10^{-25}$ m$^{-2}\frac{1}{2}(4\times10^7$ m$)^2\frac{\sqrt{3}}{2}$ 1 sec 3×10^8 m/sec $\sim$ 0.07 m. In other words $\delta U/U \sim 2.4 \times 10^{-10}$, which is readily measurable. The contribution of the Moon to the local Riemann curvature is about 2.4×10^{-6} that of the Earth, while that of the Sun is half as much the Moon's. Thus, Schwarzschild panels with a time base of about 3 days ($= \frac{1 \text{ sec}}{3.6\times10^{-6}}$) are expected to show a 0.07 m gap. Such a gap represents a $\sim 10^{-15}$th part of the temporal side of the panel, and is about at the limit of what an average a standard cesium atomic clock can measure

experiments performed on a limited set of parameters, and with limited numerical accuracy.

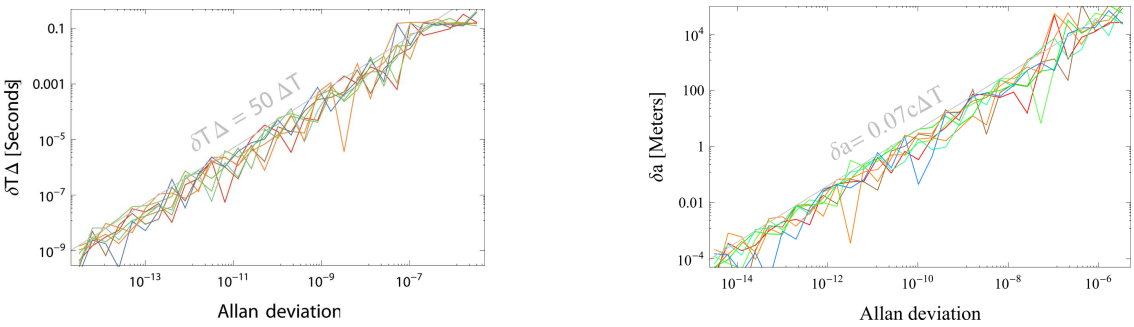


Figure 5. *Time origin and orbital axis uncertainty as functions of clock timing noise.*

today, but a hundred times above the limit of the best clock reported as of today [3].

3 Conclusion

The ABC system of coordinates is a concept based on parceling spacetime into covariant panels, which are constructed with null surfaces, connecting world lines of bodies (satellites), freely falling in the local spacetime geometry. In the ABC system coordinates are defined only by clocks. Therefore, the dynamical position of a body in the part of spacetime, covered by ABC coordinates, is defined with the accuracy of clock reading.

In order to establish an ABC system, two requirements must be fulfilled: 1) GNSS satellites must communicate and exchange emission coordinates, and 2) ABC satellites must move on geodesics of spacetime, i.e. they should by influenced only by local gravity and no other forces (or other forces would have to be accurately known). The second requirement could in principle be met by making satellites drag-free or by measuring their acceleration with sufficient accuracy. Small, random forces such as due to micrometeorite hits are sufficiently rare that the sudden change of orbital parameters of one satellite in the constellation would be readily detected and taken into account by other members of the constellation.

The benefits of establishing an ABC would be numerous and would grow with time as the system of coordinates would increase its accuracy through better understanding of nongravitational forces influencing satellite dynamics, through improved accuracy of clocks and through improved accuracy of intersatellite communication. Sub-millimeter accuracy which appears reachable with today's standard atomic clocks (c.f. Fig.5), could be a formidable tool for studying Earth dynamics including continental drift, tectonic and volcanic activity, magma flow, ocean currents, erosion, etc. Also, the ABC provides a logical and most precise definition of coordinate time by which ABC should be connected to ICRS.

Acknowledgements

I would like to express my deep appreciation of Massimo Calvani's friendship who, as many times before, read my manuscript and helped improve it.

References

[1] B. Coll and J. M. Pozo. Relativistic positioning systems: the emission coordinates. *Class. Quantum Grav.*, 23:7395, 2006.

[2] C. J. S. Clarke and D. W. Sciama. Static gravitational multipoles — the connection between field and source in general relativity. *General Relativity and Gravitation*, 2:331–345, 1971.

[3] C. W. Chow, D. B. Hume, T. Rosenband, and D. J. Weinland. Optical clocks and relativity. *Science*, 329:1630–1633, 2010.

[4] A. Čadež, U. Kostić, and P. Delva. Ariadna final report ID:09/1301: Mapping the spacetime metric with a global navigation satellite system. Technical report, Advanced Concepts Team, European Space Agency, 2010.

[5] A. Čadež, U. Kostić, P. Delva, and S. Carloni. Ariadna final report ID:09/1301: Mapping the spacetime metric with a global navigation satellite system — extension of study: Recovering of

 DOI: 10.2420/AF07.2013.57

orbital constants using inter-satellite links. Technical report, Advanced Concepts Team, European Space Agency, 2011.

[6] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, 1973.

[7] D. Sobel. *Longitude: The True Story of a Lone Genius Who Solved the Greatest Scientific Problem of his Time*. Fourth Estate, 2005.

[8] L. Arona, J. Pulido, F. Soualle, A. Fernández, and M. Sanchez-Nogales. ESA Final Report GNSSPLUS-DMS-TEC-FIR01-11-E-R. Technical report, ESA, 2006.

[9] P. Delva, A. Čadež, U. Kostić, and S. Carloni. A relativistic and autonomous navigation satellite system. *ArXiv e-prints*, 1106(3168), 2011.

[10] K. S. Thorne. Multipole expansion of gravitational radiation. *Reviews of Modern Physics*, 52(4):299–340, 1980.

Acta Futura 7 (2013) 67-78
DOI: 10.2420/AF07.2013.67

Acta Futura

Atomic clocks: new prospects in metrology and geodesy

Pacôme Delva*, and Jérôme Lodewyck
LNE-SYRTE, Observatoire de Paris, CNRS and UPMC ; 61 avenue de l'Observatoire, 75014 Paris, France

Abstract. We present the latest developments in the field of atomic clocks and their applications in metrology and fundamental physics. In the light of recent advents in the accuracy of optical clocks, we present an introduction to the relativistic modelization of frequency transfer and a detailed review of chronometric geodesy.

1 Introduction

Atomic clocks went through tremendous evolutions and ameliorations since their invention in the middle of the twentieth century. The constant amelioration of their accuracy (figure 1) and stability permitted numerous applications in the field of metrology and fundamental physics. For a long time cold atom Caesium fountain clocks remained unchallenged in terms of accuracy and stability. However this is no longer true with the recent development of optical clocks. This new generation of atomic clock opens new possibilities for applications, such as chronometric geodesy, and requires new developments, particularly in the field of frequency transfer. The LNE-SYRTE laboratory (CNRS/LNE/Paris Observatory/UPMC) is involved in many aspects of the development of atomic clocks and their applications.

In section 2 we present the latest developments in the field of atomic clocks: microwave clocks, optical clocks, their relation to international time-scales, means of comparisons and applications. Section 3 is an intro-

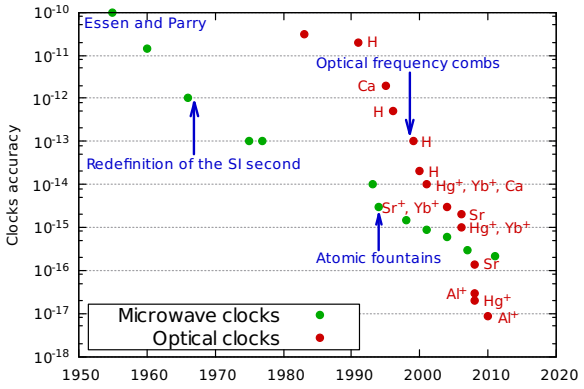


Figure 1. *Accuracy records for microwave and optical clocks. From the first Cs clock by Essen and Parry in the 1950's, an order of magnitude is gained every ten years. The advent of optical frequency combs boosted the performances of optical clocks, and they recently overcome microwave clocks.*

*Corresponding author. E-mail: Pacome.Delva@obspm.fr

duction to relativistic time transfer, the modelization of remote frequency comparisons, which lies at the heart of many applications of atomic clocks, such as the realization of international time-scales and chronometric geodesy. Finally, section 4 is a detailed review of the field of chronometric geodesy, an old idea which could become reality in the near future.

2 Atomic clocks

In 1967, the definition of the SI second was changed from astronomical references to atomic references by setting the frequency of an hyperfine transition in the Cs atom [39]. Since then, the accuracy of atomic clocks has improved by five orders of magnitude, enabling better and better time-keeping. More recently, a new generation of atomic clocks, based on atomic transitions in the optical domain are challenging the well established Cs standard and thus offer opportunities for new applications in fundamental physics and geodesy.

2.1 Microwave clocks

In a microwave atomic frequency standard, a microwave electro-magnetic radiation excites an hyperfine electronic transition in the ground state of an atomic species. Observing the fraction of excited atoms p after this interaction (or transition probability) gives an indicator of the difference between the frequency ν of the microwave radiation and the frequency ν_0 of the hyperfine atomic transition. This frequency difference $\nu - \nu_0$ (or error signal) is fed in a servo-loop that keeps the microwave radiation resonant with the atomic transition. According to Fourier's relation, the frequency resolution that can be achieved after such an interrogation procedure grows as the inverse of the interaction time T, and since consecutive interrogations are uncorrelated, the frequency resolution further improves as the square root of the total integration time τ. Quantitatively, the residual frequency fluctuation of the microwave radiation locked on the atomic resonance are (in dimensionless fractional units, that is to say divided by the microwave frequency):

$$\sigma_y(\tau) = \frac{\xi}{\nu_0 T\sqrt{N}}\sqrt{\frac{T_c}{\tau}}, \qquad (1)$$

where T_c is the cycle time (such that τ/T_c is the number of clock interrogations), and N is the number of simultaneously (and independently) interrogated atoms. ξ is a numerical constant, close to unity, that depends on the physics of the interaction between the radiation and the atoms. This expression is the ultimate *frequency (in)stability* of an atomic clock, also called the *Quantum Projection Noise* (QPN) limit, referring to the quantum nature of the interaction between the radiation and the atoms. It is eventually reached if all other sources of noise in the servo-loop are made negligible.

As seen from eq. (1), an efficient way to improve the clock stability is to increase the interaction time T. The first atomic clocks therefore comprised a long tube in which a thermal beam of Cs atoms is traveling while interacting with the microwave radiation.

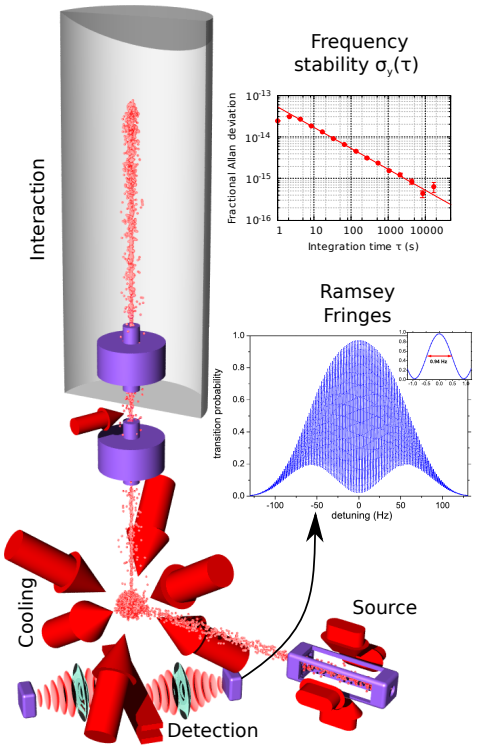


Figure 2. *Atomic fountain clock. About* 10^6 *cold atoms, at a temperature of 1* μK *are launched upwards by pushing laser beams. At the beginning and the end of their trajectory, they interact with a microwave radiation in a resonator and the hyperfine transition probability is measured by an optical detection. When scanning the microwave frequency, the transition probability follows a fringe (Ramsey) pattern much like a double slit interference pattern. During the clock operation, the microwave frequency is locked on the top of the central fringe. The frequency stability, defined by eq. (1), is a few* 10^{-14} *after 1 s, and a statistical resolution down to* 10^{-16} *is reached after a few days of continuous operation [12, 30].*

DOI: 10.2420/AF07.2013.67

The advance in the physics of cold atoms enabled to prepare atoms with a smaller velocity and consequently largely increase the interaction time, thus the frequency stability of atomic standards. In cold atoms clocks, the interrogation time is only limited by the time after which the atoms, exposed to gravity, escape the interrogation zone. In the atomic fountain clock (see fig. 2), a set of cold atoms are launched vertically in the interaction zone and are interrogated during their parabolic flight in the Earth gravitational field.

In micro-gravity, atoms are not subject to the Earth gravitational field and thus can be interrogated during a longer time window T. This will be the case for the Pharao space microwave clock [7].

The reproducibility of an atomic clock stems from the fact that all atoms of a given species are rigorously indistinguishable. For instance, two independent atomic clocks based on caesium will produce the same frequency (9,192,631,770 Hz exactly in the SI unit system), regardless of their manufacturer or their location in space-time.

However, in laboratory conditions, atoms are surrounded by an experimental environment that perturbs their electronic state and thus slightly modify their resonance frequency in a way that depends on experimental conditions. For instance, atoms will interact with DC or AC electro-magnetic perturbations, will be sensitive to collisions between them in a way that depends on the atomic density or to the Doppler effect resulting from their residual motion…Such systematic effects, if not evaluated and corrected for, limit the universality of the atomic standard. Consequently, the accuracy of a clock quantifies the uncertainty on these systematic effects. Currently, the best microwave atomic fountain clocks have a relative accuracy of 2×10^{-16}, which is presumably their ultimate performances given the many technical obstacles to further improve this accuracy.

2.2 SI second and time scales

About 250 clocks worldwide, connected by time and frequency transfer techniques (mainly through GNSS signals) realize a complete architecture that enable the creation of atomic time scales. For this, local time scales physically generated by metrology laboratories are *a posteriori* compared and common time-scales are decided upon. the International Bureau for Weights and Measures (BIPM) in Sèvres (France) is responsible for establishing the International Atomic Time (TAI). First a free atomic time-scale is build, the EAL. However, this time-scale is free-running and the participating clocks do not aim at realizing the SI second. The rate of EAL is measured by comparison with a few number of caesium atomic fountains which aim at realizing the SI second, and TAI is then derived from EAL by applying a rate correction, so that the scale unit of TAI is the SI second as realized on the rotating geoid [11]. It necessary to take into account the atomic fontain clocks frequency shift due to relativity (see section 3). Recently, a change of paradigm occured in the definition of TAI. According to UAI resolutions [36], TAI is a realization of Terrestrial Time (TT), which is defined by applying a constant rate correction to Geocentric Coordinate Time (TCG). This definition has been adopted so that the reference surface of TAI is no longer the geoid, which is not a stable surface (see section 4). Finally, the Coordinated Universal Time (UTC) differs from TAI by an integer number of second in order to follow the irregularities of the Earth rotation.

The publication of such time-scales enables worldwide comparison of Cs fountain clocks [30, 31] and the realization of the SI second.

2.3 Optical clocks

A new generation of atomic clocks have appeared in the last 15 years. These clocks consist in locking an electromagnetic radiation in the optical domain ($\nu = 300$ to 800 THz) to a narrow electronic transition. As seen from eq. (1), increasing the frequency ν_0 of the clock frequency by several orders of magnitude drastically improves the ultimate clock stability, even though the number of interrogated atoms N is usually smaller in optical clocks. Increasing the clock frequency also improves the clock accuracy since most systematic effects (sensitivity to DC electromagnetic fields, cold collisions…) are of the same order of magnitude in frequency units, and thus decrease in relative units. However, two notable exceptions remain, and both have triggered recent research in optical clocks. First, the sensitivity of the clock transition frequency on the ambient black-body radiation is mostly rejected in microwave clocks but not for optical transitions. Thus, the uncertainty on this effect usually only marginally improves when going to optical clocks (with the notable exception of a few atomic species such as Al^+ for which the sensitivity is accidentally small). Therefore, extra care has to be taken to control the temperature of the atoms environment. Second, the Doppler frequency shift $\delta\nu$ due to the residual velocity v of the atoms scales as the

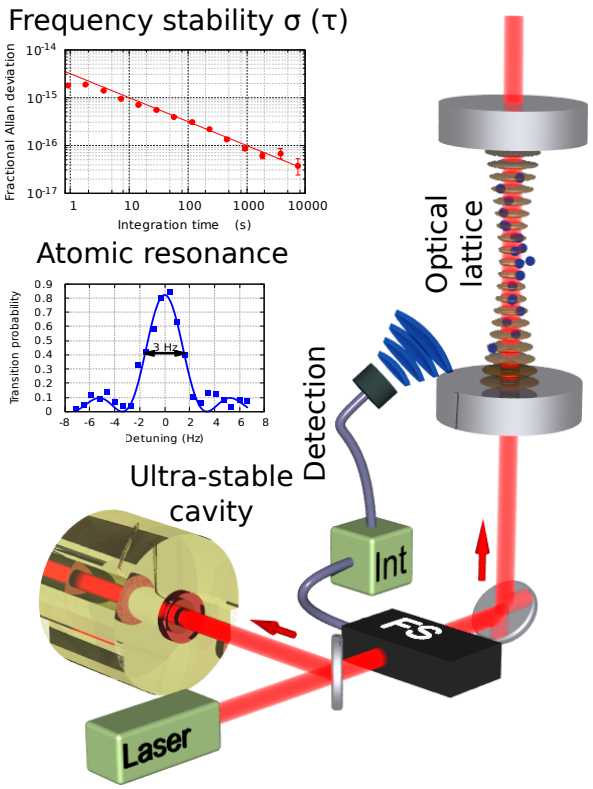


FIGURE 3. *Sr optical lattice clock. The clock laser, prestabilized on an ultra-stable optical cavity, probes an optical transition of* 10^4 *atoms confined in an optical lattice. The excitation fraction (or transition probability) is detected by a fluorescence imaging and a numerical integrator acts on a frequency shifter (FS) to keep the laser on resonance with the atomic transition. The width of the resonance is Fourier-limited at 3 Hz, which, given the clock frequency* $\nu_0 = 429$ *THz, yields a resonance quality factor* $Q = 1.4 \times 10^{14}$*, compared to* 10^{10} *for microwave clocks.*

clock frequency ν_0, such that the relative frequency shift remains constant:

$$\frac{\delta\nu}{\nu_0} = \frac{v}{c} \tag{2}$$

For this reason, the fountain architecture, for which the Doppler effect is one of the limitations, cannot be applicable to optical clocks, and the atoms have to be tightly confined in a trapping potential to cancel their velocity v. To achieve this goal, two different technologies have been developed. First, a single ion is trapped in a RF electric field. These *ion optical clocks* [34, 8, 16, 25] achieve record accuracies at 9×10^{-18}. However, since only a single ion is trapped ($N = 1$, because of the electrostatic repulsion of ions), the stability of theses clocks is limited at the QPN level of 2×10^{-15} at 1 s. Second, a more recent technology involves trapping of a few thousands neutral atoms in a powerful laser standing wave (or optical lattice) by the dipolar force. Due to its power, this trapping potential is highly perturbative, but for a given "*magic*" wavelength of the trapping light [38], the perturbation is equal for the two clock levels, hence cancelled for the clock transition frequency. These *optical lattice clocks* [38, 21, 29, 23, 26] have already reached an accuracy of 1×10^{-16} and are rapidly catching up with ion clocks. Furthermore, the large number of interrogated atoms allowed the demonstration of unprecedented stabilities (a few 1×10^{-16} at 1 s), heading toward their QPN below 1×10^{-17} at 1 s.

2.4 Remote comparison of optical frequencies

The recent breakthrough of performances of optical clocks was permitted by the development of frequency combs [44], which realizes a "ruler" in the frequency domain. These lasers enable the local comparison of different optical frequencies and comparisons between optical and microwave frequencies. However, although optical clocks now largely surpass microwave clocks, a complete architecture has to be established to enable remote comparison of optical frequencies, in order to validate the accuracy of optical clocks, build "optical" time-scales, or enable applications of optical clocks such as geodesic measurements Indeed, the conventional remote frequency comparisons techniques, mainly through GPS links, cannot reach the level of stability and accuracy realized by optical clocks.

 DOI: 10.2420/AF07.2013.67

Fibre links

To overcome the limitations of remote clock comparisons using GNSS signals, comparison techniques using optical fibres are being developed. For this, a stable and accurate frequency signal produced by an optical clock is sent through a fibre optics that links metrology institutes, directly encoded in the phase of the optical carrier. Because of vibrations and temperature fluctuations, the fibre adds a significant phase noise to the signal. This added noise is measured by comparing the signal after a round trip in the fibre to the original signal, and subsequently cancelled. Such a signal can be transported through a dedicated fibre (dark fibre) when available [33], or, more practically along with the internet communication (dark channel) [24]. These comparison techniques are applicable at a continental scale ; such a network will presumably be in operation throughout Europe in the near future. In particular, the REFIMEVE+ project will provide a shared stable optical oscillator between a large number of laboratories in France, with a number of applications beside metrology. This network will be connected to international fibre links (NEAT-FT project[1]) that will enable long distance comparisons between optical frequency standards located in various national metrology laboratories in Europe. The ITOC project[2] aims at collecting these comparisons result to demonstrate high accuracy frequency ratios measurements and geophysical applications.

Space-based links

When considering inter-continental time and frequency comparisons, only satellite link are conceivable. In this aspect, The two-way satellite time and frequency transfer (TWSTFT) involves a satellite that actively relays a frequency signal in a round-trip configuration. Also, the space mission Pharao-ACES [7] that involves an ensemble of clocks on board the International Space Station will comprise a number of ground receiver able to remotely compare optical clocks.

2.5 Towards a new definition of the SI second based on optical clocks

Unlike microwave clocks, for which Cs has been an unquestionable choice over half a century, a large number of atomic species seem to be equally matched as a new frequency standard based on an optical transition. Ion clocks with Hg^+, Al^+, Yb^+, Sr^+, Ca^+ have been demonstrated, as well as optical lattice clocks with Sr, Yb and more prospectively Hg and Mg. As an illustration, four optical transitions have already been approved by the CIPM (International Committee for Weights and Measures) as secondary representation of the SI second. Therefore, although the SI second would already gain in precision with optical clock, a clear consensus has yet to emerge before the current microwave defined SI second can be replaced.

2.6 Applications of optical clocks

The level of accuracy reached by optical clocks opens a new range of applications, through the very precise frequency ratios measurements they enable. In fundamental physics, they enable the tracking of dimension-less fundamental constants such as the fine structure constant α, the electron to proton mass ratio $\mu = m_e/m_p$, or the quantum chromodynamics mass scale m_q/Λ_{QCD}. Because each clock transition frequency have a different dependence on these constants, their variations imply drifts in clock frequency ratios that can be detected by repeated measurements. Currently, combined optical-to-optical and optical-to-microwave clock comparisons put an upper bound on the relative variation of fundamental constants in the 10^{-17}/yr range [34, 13, 21, 22]. Other tests of fundamental physics are also possible with atomic clocks. The local position invariance can be tested by comparing frequency ratios in the course of the earth rotation around the earth [13, 21], and the gravitational red-shift will be tested with clocks during the Pharao-ACES mission.

The TAI time scale is created from Cs standards, but recently, the Rb microwave transition started to contribute. In the near future, secondary representations of the SI second based on optical transitions could contribute to TAI, which would thus benefit from their much improved stability.

Clock frequencies, when compared to a coordinate time-scale, are sensitive to the gravitational potential of the Earth (see section 3). Therefore it is necessary to take into account their relative height difference when building atomic time-scales such as TAI. However, the most accurate optical clocks can resolve a height difference below 10 cm, which is a scale at which the global gravitational potential is unknown. Because of this, optical clock could become a tool to precisely measure this potential. They come as a decisive addition to relative

[1] http://www.ptb.de/emrp/neatft_home.html
[2] http://projects.npl.co.uk/itoc/

and absolute gravimeters which are sensitive to the gravitational field, and to satellite based measurement which lack spatial resolution (see section 4).

3 Relativistic frequency transfer

In distant comparisons of frequency standards, we are face with the problem of curvature of space and relative motion of the clocks. These two effects change locally the flow of proper time with respect to a global coordinate time. In this section we describe how to compare distant clock frequencies by means of an electromagnetic signal, and how the comparison is affected by these effects.

3.1 The Einstein Equivalence Principle

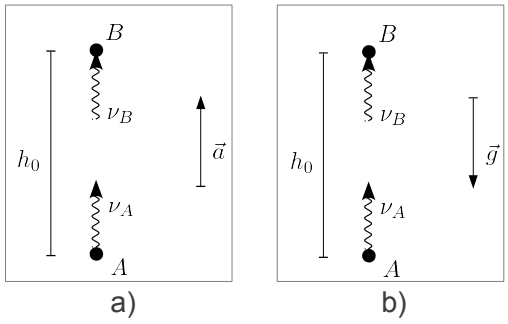


Figure 4. *A photon of frequency ν_A is emitted at point A toward point B, where the measured frequency is ν_B. a) A and B are two points at rest in an accelerated frame, with acceleration $\vec{a}$ in the same direction as the emitted photon. b) A and B are at rest in a non accelerated (locally inertial) frame in presence of a gravitational field such that $\vec{g} = -\vec{a}$.*

Before we treat the general case, let's try to understand in simple terms what is the frequency shift effect. Indeed, it can be seen as a direct consequence of the Einstein Equivalence Principle (EEP), one of the pillars of modern physics [27, 41]. Let's consider a photon emitted at a point A in an accelerated reference system, toward a point B which lies in the direction of the acceleration (see fig.4). We assume that both point are separated by a distance h_0, as measured in the accelerated frame. The photon time of flight is $\delta t = h_0/c$, and the frame velocity during this time increases by $\delta v = a\delta t = ah_0/c$, where a is the magnitude of the frame acceleration $\vec{a}$. The frequency at point B (reception) is then shifted because of Doppler effect, compared to the frequency at point A (emission), by an amount:

$$\frac{\nu_B}{\nu_A} = 1 - \frac{\delta v}{c} = 1 - \frac{ah_0}{c^2} \tag{3}$$

Now, the EEP postulates that a gravitational field $\vec{g}$ is locally equivalent to an acceleration field $\vec{a} = -\vec{g}$. We deduce that in a non accelerated (locally inertial) frame in presence of a gravitational field $\vec{g}$:

$$\frac{\nu_B}{\nu_A} = 1 - \frac{gh_0}{c^2} \tag{4}$$

where $g = |\vec{g}|$, ν_A is the photon frequency at emission (strong gravitational potential) and ν_B is the photon frequency at reception (weak gravitational field). As $\nu_B < \nu_A$, it is usual to say that the frequency at the point of reception is "red-shifted". One can consider it in terms of conservation of energy. Intuitively, the photon that goes from A to B has to "work" to be able to escape the gravitational field, then it looses energy and its frequency decreases by virtue of $E = h\nu$, with h the Planck constant.

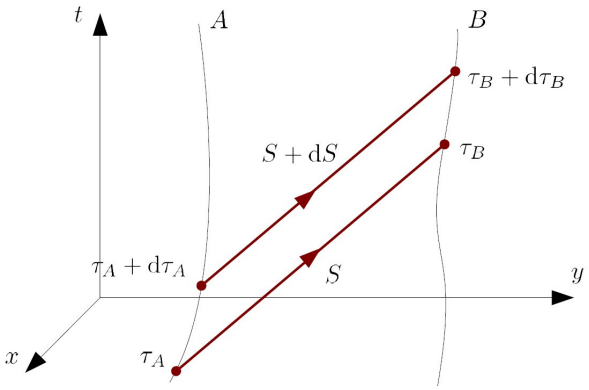


Figure 5. *Two clocks A and B are measuring proper time along their trajectory. One signal with phase S is emitted by A at proper time τ_A, and another one with phase $S + dS$ at time $\tau_A + d\tau_A$. They are received by clock B respectively at time τ_B and $\tau + d\tau_B$.*

3.2 General case

The principle of frequency comparison is to measure the frequency of an electromagnetic signal with the help of the emitting clock, A, and then with the receiving clock, B. We obtain respectively two measurements ν_A and ν_B [3]. Let $S(x^\alpha)$ be the phase of the electromagnetic

[3] However, in general one measures the time of flight of the electromagnetic signal between emission and reception. Then the ratio ν_A/ν_B can be obtain by deriving the time of flight measurements.

DOI: 10.2420/AF07.2013.67

signal emitted by clock A. It can be shown that light rays are contained in hypersurfaces of constant phase. The frequency measured by A/B is:

$$\nu_{A/B} = \frac{1}{2\pi}\frac{dS}{d\tau_{A/B}} \tag{5}$$

where $\tau_{A/B}$ is the proper time along the worldline of clock A/B (see fig.5). We introduce the wave vector $k_\alpha^{A/B} = (\partial_\alpha S)_{A/B}$ to obtain:

$$\nu_{A/B} = \frac{1}{2\pi} k_\alpha^{A/B} u_{A/B}^\alpha \tag{6}$$

where $u_{A/B}^\alpha = dx_{A/B}^\alpha/d\tau$ is the four-velocity of clock A/B. Finally, we obtain a fundamental relation for frequency transfer:

$$\frac{\nu_A}{\nu_B} = \frac{k_\alpha^A u_A^\alpha}{k_\alpha^B u_B^\alpha} \tag{7}$$

This formula does not depend on a particular theory, and then can be used to perform tests of general relativity. Introducing $v^i = dx^i/dt$ and $\hat{k}_i = k_i/k_0$, it is usually written as:

$$\frac{\nu_A}{\nu_B} = \frac{u_A^0}{u_B^0}\frac{k_0^A}{k_0^B}\frac{1+\frac{\hat{k}_i^A v_A^i}{c}}{1+\frac{\hat{k}_i^B v_B^i}{c}} \tag{8}$$

From eq. (5) we deduce that:

$$\frac{\nu_A}{\nu_B} = \frac{d\tau_B}{d\tau_A} = \left(\frac{dt}{d\tau}\right)_A \frac{dt_B}{dt_A}\left(\frac{d\tau}{dt}\right)_B \tag{9}$$

We suppose that space-time is stationary, ie. $\partial_0 g_{\alpha\beta} = 0$. Then it can be shown that k_0 is constant along the light ray, meaning that $k_0^A = k_0^B$. We introduce the time transfer function:

$$\mathcal{T}(x_A^i, x_B^i) = t_B - t_A \tag{10}$$

Deriving the time transfer function with respect to t_A one obtains:

$$\frac{dt_B}{dt_A} = \frac{1+\frac{\partial\mathcal{T}}{\partial x_A^i}v_A^i}{1-\frac{\partial\mathcal{T}}{\partial x_B^i}v_B^i} \tag{11}$$

Inserting eq. (11) in eq. (9), and comparing with eq. (8), we deduce:

$$\hat{k}_i^A = c\frac{\partial\mathcal{T}}{\partial x_A^i}\ ,\ \hat{k}_i^B = -c\frac{\partial\mathcal{T}}{\partial x_B^i} \tag{12}$$

General formula for non-stationary space-times can be found in [14, 20].

As an exemple, let's take the simple time transfer function:

$$\mathcal{T}(x_A^i, x_B^i) = \frac{R_{AB}}{c} + \mathcal{O}\left(\frac{1}{c^3}\right) \tag{13}$$

where $R_{AB} = |R_{AB}^i|$ and $R_{AB}^i = x_B^i - x_A^i$. Then we obtain:

$$\frac{dt_B}{dt_A} = \frac{1+\frac{\vec{N}_{AB}\cdot\vec{v}_A}{c}+\mathcal{O}\left(\frac{1}{c^3}\right)}{1+\frac{\vec{N}_{AB}\cdot\vec{v}_B}{c}+\mathcal{O}\left(\frac{1}{c^3}\right)} \tag{14}$$

where $N_{AB}^i = R_{AB}^i/R_{AB}$. Up to the second order, this term does not depend on the gravitational field but on the relative motion of the two clocks. It is simply the first order Doppler effect.

The two other terms in eq. (9) depend on the relation between proper time and coordinate time. In a metric theory one has $c^2 d\tau = \sqrt{g_{\alpha\beta}dx^\alpha dx^\beta}$. We deduce that:

$$\frac{u_A^0}{u_B^0} = \frac{\left[g_{00}+2g_{0i}\frac{v^i}{c}+g_{ij}\frac{v^i v^j}{c^2}\right]_B^{1/2}}{\left[g_{00}+2g_{0i}\frac{v^i}{c}+g_{ij}\frac{v^i v^j}{c^2}\right]_A^{1/2}} \tag{15}$$

3.3 Application to a static, spherically symmetric body

As an example, we apply this formalism for the case of a parametrized post-Minkowskian approximation of metric theories. This metric is a good approximation of the space-time metric around the Earth, for a class of theories which is larger than general relativity. Choosing spatial isotropic coordinates, we assume that the metric components can be written as:

$$\begin{aligned} g_{00} &= -1 + 2(\alpha+1)\frac{GM}{rc^2} + \mathcal{O}\left(\frac{1}{c^4}\right) \\ g_{0i} &= 0 \\ g_{ij} &= \delta_{ij}\left[1 + 2\gamma\frac{GM}{rc^2} + \mathcal{O}\left(\frac{1}{c^4}\right)\right] \end{aligned} \tag{16}$$

where α and γ are two parameters of the theory (for general relativity $\alpha = 0$ and $\gamma = 1$). Then:

$$\frac{d\tau}{dt} = 1 - (\alpha+1)\frac{GM}{rc^2} - \frac{v^2}{2c^2} + \mathcal{O}\left(\frac{1}{c^4}\right) \tag{17}$$

where $v = \sqrt{\delta_{ij} v^i v^j}$. The time transfer function of such a metric is given in eq. (108) of [20]. Then we calculate:

$$\frac{dt_B}{dt_A} = \frac{q_A + \mathcal{O}\left(\frac{1}{c^5}\right)}{q_B + \mathcal{O}\left(\frac{1}{c^5}\right)} \tag{18}$$

where

$$q_A = 1 - \frac{\vec{N}_{AB}\cdot\vec{v}_A}{c} - \frac{2(\gamma+1)GM}{c^3}\times \frac{R_{AB}\vec{N}_A\cdot\vec{v}_A + (r_A+r_B)\vec{N}_{AB}\cdot\vec{v}_A}{(r_A+r_B)^2 - R_{AB}^2} \tag{19}$$

$$q_B = 1 - \frac{\vec{N}_{AB}\cdot\vec{v}_B}{c} - \frac{2(\gamma+1)GM}{c^3}\times \frac{R_{AB}\vec{N}_A\cdot\vec{v}_B - (r_A+r_B)\vec{N}_{AB}\cdot\vec{v}_B}{(r_A+r_B)^2 - R_{AB}^2} \tag{20}$$

Let assume that both clocks are at rest with respect to the chosen coordinate system, ie. $\vec{v}_A = \vec{0} = \vec{v}_B$ and that $r_A = r_0$ and $r_B = r_0 + \delta r$, where $\delta r \ll r_0$. Then we find:

$$\frac{\nu_A - \nu_B}{\nu_B} = (\alpha+1)\frac{GM}{r_0^2 c^2}\delta r + \mathcal{O}\left(\frac{1}{c^4}\right) \tag{21}$$

The same effect can be calculated with a different, not necessarily symmetric gravitational potential $w(t, x^i)$. The results yields:

$$\frac{\nu_A - \nu_B}{\nu_B} = \frac{w_A - w_B}{c^2} + \mathcal{O}\left(\frac{1}{c^4}\right) \tag{22}$$

This is the classic formula given in textbooks for the "gravitational red-shift". However one should bear in mind that this is valid for clocks at rest with respect to the coordinate system implicitly defined by the space-time metric (one uses usually the Geocentric Celestial Reference System), which is (almost) never the case. Moreover, the separation between a gravitational red-shift and a Doppler effect is specific to the approximation scheme used here. One can read the book by Synge [37] for a different interpretation in terms of relative velocity and Doppler effect only.

We note that the lowest order gravitational term in eq. (21) is a test of the Newtonian limit of metric theories. Indeed, if one wants to recover the Newtonian law of gravitation for $GM/rc^2 \ll 1$ and $v \ll 1$ then it is necessary that $\alpha = 0$. Then this test is more fundamental than a test of general relativity, and can be interpreted as a test of Local Position Invariance (which is a part of the Einstein Equivalence Principle). See [41, 42] for more details on this interpretation, and a review of the experiments that tested the parameter α.

A more realistic case of the space-time metric is treated in article [3], in the context of general relativity: all terms from the Earth gravitational potential are considered up to an accuracy of 5.10^{-17}, specifically in prevision of the ACES mission [7].

4 Chronometric geodesy

Instead of using our knowledge of the Earth gravitational field to predict frequency shifts between distant clocks, one can revert the problem and ask if the measurement of frequency shifts between distant clocks can improve our knowledge of the gravitational field. To do simple orders of magnitude estimates it is good to have in mind some correspondences calculated thanks to eqs. (21) and (22):

$$\begin{aligned} 1 \text{ meter } &\leftrightarrow \frac{\Delta\nu}{\nu} \sim 10^{-16} \\ &\leftrightarrow \Delta w \sim 10 \text{ m}^2.\text{s}^{-2} \end{aligned} \tag{23}$$

From this correspondence, we can already imagine two direct applications of clocks in geodesy: if we are capable to compare clocks to 10^{-16} accuracy, we can determine height differences between clocks with one meter accuracy (levelling), or determine geopotential differences with 10 $\text{m}^2.\text{s}^{-2}$ accuracy.

4.1 A review of chronometric geodesy

The first article to explore seriously this possibility was written in 1983 by Martin Vermeer [40]. The article is named "chronometric levelling". The term "chronometric" seems well suited for qualifying the method of using clocks to determine directly gravitational potential differences, as "chronometry" is the science of the measurement of time. However the term "levelling" seems too restrictive with respect to all the applications one could think of using the results of clock comparisons. Therefore we will use the term "chronometric geodesy" to name the scientific discipline that deals with the measurement and representation of the Earth, including its gravitational field, with the help of atomic clocks. It is sometimes named "clock-based geodesy", or "relativistic geodesy". However this last designation is improper as

 DOI: 10.2420/AF07.2013.67

relativistic geodesy aims at describing all possible techniques (including e.g. gravimetry and gradiometry) in a relativistic framework [18, 28, 35].

The natural arena of chronometric geodesy is the four-dimensional space-time. At the lowest order, there is proportionality between relative frequency shift measurements – corrected from the first order Doppler effect – and (Newtonian) geopotential differences (see eq.(22)). To calculate this relation we have seen that we do not need the theory of general relativity, but only to postulate Local Position Invariance. Therefore, it is perfectly possible to use clock comparison measurements – corrected from the first order Doppler effect – as a direct measurement of geopotential differences in the framework of classical geodesy, if the measurement accuracy does not reach the magnitude of the higher order terms.

Comparisons between two clocks on the ground generally use a third clock in space. For the comparison between a clock on the ground and one in space, the terms of order c^{-3} in eqs. (19)-(20) reach a magnitude of $\sim 10^{-15}$ and $\sim 3 \times 10^{-14}$ for respectively the ground and the space clock, if they are separated radially by 1000 km. Terms of order c^{-4} omitted in eqs. (21)-(22) can reach $\sim 5 \times 10^{-19}$ in relative frequency shift, which corresponds to a height difference of ~ 5 mm and a geopotential difference of $\sim 5 \times 10^{-2}$ $m^2.s^{-2}$. Clocks are far from reaching this accuracy today, but it cannot be excluded for the future.

In his article, Martin Vermeer explores the "possibilities for technical realisation of a system for measuring potential differences over intercontinental distances" using clock comparisons [40]. The two main ingredients are of course accurate clocks and a mean to compare them. He considers hydrogen maser clocks. For the links he considers a 2-way satellite link over a geostationary satellite, or GPS receivers in interferometric mode. He has also to consider a mean to compare the proper frequencies of the different hydrogen maser clocks. However today this can be overcome by comparing Primary Frequency Standards (PFS), which have a well defined proper frequency based on a transition of Caesium 133, used for the definition of the second. However, this problem will rise again if one uses Secondary Frequency Standards which are not based on Caesium atoms. Then the proper frequency ratio between two different kinds of atomic clocks has to be determined locally. This is one of the purpose of the European project "International timescales with optical clocks"[4], where optical clocks based on different atoms will be compared one each other locally, and to the PFS. It is planned also to do a proof-of-principle experiment of chronometric geodesy, by comparing two optical clocks separated by a height difference of around 1 km using an optical fibre link.

In the foreseen applications of chronometric geodesy, Martin Vermeer mentions briefly intercontinental levelling, and the measurement of the true geoid (disentangled from geophysical sea surface effects) [40]. Few authors have seriously considered chronometric geodesy. Following Vermeer idea, Brumberg and Groten [5] demonstrated the possibility of using GPS observations to solve the problem of the determination of geoid heights, however leaving aside the practical feasibility of such a technique. Bondarescu *et al.* [4] discuss the value and future applicability of chronometric geodesy, including direct geoid mapping on continents and joint gravity-geopotential surveying to invert for subsurface density anomalies. They find that a geoid perturbation caused by a 1.5 km radius sphere with 20 per cent density anomaly buried at 2 km depth in the Earth's crust is already detectable by atomic clocks of achievable accuracy. Finally Chou *et al.* demonstrated the potentiality of the new generation of atomic clocks, based on optical transitions, to measure heights with a resolution of around 30 cm [9].

4.2 The chronometric geoid

Arne Bjerhammar in 1985 gives a precise definition of the "relativistic geoid" [1, 2]:

> "The relativistic geoid is the surface where precise clocks run with the same speed and the surface is nearest to mean sea level"

This is an operational definition. Soffel et al. [35] in 1988 translated this definition in the context of post-Newtonian theory. They also introduce a different operational definition of the relativistic geoid, based on gravimetric measurements: a surface orthogonal everywhere to the direction of the plumb-line and closest to mean sea level. He calls the two surfaces obtained with clocks and gravimetric measurements respectively the "u-geoid" and the "a-geoid". He proves that these two surfaces coincide in the case of a stationary metric. In order to distinguish the operational definition of the geoid from its theoretical description, it is less ambiguous to give a name based on the particular technique to

[4] http://projects.npl.co.uk/itoc/

measure it. Relativistic geoid is too vague as Soffel et al. have defined two different ones. The names chosen by Soffel et al. are not particularly explicit, so instead of "u-geoid" and "a-geoid" one can call them chronometric and gravimetric geoid respectively. There can be no confusion with the geoid derived from satellite measurements, as this is a quasi-geoid that do not coincide with the geoid on the continents [15]. Other considerations on the chronometric geoid can be found in [18, 19, 28].

Let two clocks be at rest with respect to the chosen coordinate system ($v^i = 0$) in an arbitrary space-time. From formula (9), (11) and (15) we deduce that:

$$\frac{\nu_A}{\nu_B} = \frac{d\tau_B}{d\tau_A} = \frac{[g_{00}]_B^{1/2}}{[g_{00}]_A^{1/2}} \tag{24}$$

In this case the chronometric geoid is defined by the condition $g_{00} = \text{cst}$.

We notice that the problem of defining a reference surface is closely related to the problem of realizing Terrestrial Time (TT). TT is defined with respect to Geocentric Coordinate Time (TCG) by the relation [36, 32]:

$$\frac{dTT}{dTCG} = 1 - L_G \tag{25}$$

where L_G is a defining constant. This constant has been chosen so that TT coincides with the time given by a clock located on the classical geoid. It could be taken as a formal definition of the chronometric geoid [43]. If so, the chronometric geoid will differ in the future from the classical geoid: a level surface of the geopotential closest to the topographic mean sea level. Indeed, the value of the potential on the geoid, W_0, depends on the global ocean level which changes with time [6]. With a value of $dW_0/dt \sim 10^{-3}$ m^2.s^{-2}.y^{-1}, the difference in relative frequency between the classical and the chronometric geoid would be around 10^{-18} after 100 years (the correspondence is made with the help of relations (23)). However, the rate of change of the global ocean level could change during the next decades. Predictions are highly model dependant [17].

5 Conclusion

We presented recent developments in the field of atomic clocks, as well as an introduction to relativistic frequency transfer and a detailed review of chronometric geodesy. If the control of systematic effects in optical clocks keep their promises, they could become very sensible to the gravitational field, which will ultimately degrade their stability at the surface of the Earth. One solution will be to send very stable and accurate clocks in space, which will become the reference against which the Earth clocks would be compared. Moreover, by sending at least four of these clocks in space, it will be possible to realize a very stable and accurate four-dimensional reference system in space [10].

References

[1] A. Bjerhammar. On a relativistic geodesy. *Bulletin Geodesique*, 59:207–220, 1985.

[2] A. Bjerhammar. Relativistic geodesy. Technical Report NON118 NGS36, NOAA Technical Report, 1986.

[3] L. Blanchet, C. Salomon, P. Teyssandier, and P. Wolf. Relativistic theory for time and frequency transfer to order c^{-3}. *A&A*, 370:320–329, Apr. 2001.

[4] R. Bondarescu, M. Bondarescu, G. Hetényi, L. Boschi, P. Jetzer, and J. Balakrishna. Geophysical applicability of atomic clocks: direct continental geoid mapping. *Geophysical Journal International*, 191:78–82, Oct. 2012.

[5] V. A. Brumberg and E. Groten. On determination of heights by using terrestrial clocks and GPS signals. *Journal of Geodesy*, 76:49–54, Jan. 2002.

[6] M. Burša, S. Kenyon, J. Kouba, Z. Šíma, V. Vatrt, V. Vítek, and M. Vojtíšková. The geopotential value W_0 for specifying the relativistic atomic time scale and a global vertical reference system. *Journal of Geodesy*, 81:103–110, Feb. 2007.

[7] L. Cacciapuoti and C. Salomon. Space clocks and fundamental tests: The ACES experiment. *The European Physical Journal - Special Topics*, 172:57–68, 2009. 10.1140/epjst/e2009-01041-7.

[8] C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland, and T. Rosenband. Frequency comparison of two high-accuracy Al^+ optical clocks. *Phys. Rev. Lett.*, 104:070802, Feb 2010.

[9] C. W. Chou, D. B. Hume, T. Rosenband, and D. J. Wineland. Optical Clocks and Relativity. *Science*, 329:1630–, Sept. 2010.

 DOI: 10.2420/AF07.2013.67

[10] P. Delva, A. Čadež, U. Kostić, and S. Carloni. A relativistic and autonomous navigation satellite system. In *Gravitational Waves and Experimental Gravity*, pages 277–280, 2011. Proceedings of the XLVI$^{\text{th}}$ RENCONTRES DE MORIOND And GPhyS Colloquium.

[11] P. Giacomo. News from the BIPM. *Metrologia*, 17:69–74, Apr. 1981.

[12] J. Guéna, M. Abgrall, D. Rovera, P. Laurent, B. Chupin, M. Lours, G. Santarelli, P. Rosenbusch, M. Tobar, R. Li, K. Gibble, A. Clairon, and S. Bize. Progress in atomic fountains at LNE-SYRTE. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 59(3):391–409, 2012.

[13] J. Guéna, M. Abgrall, D. Rovera, P. Rosenbusch, M. E. Tobar, P. Laurent, A. Clairon, and S. Bize. Improved tests of local position invariance using ^{87}Rb and ^{133}Cs fountains. *Phys. Rev. Lett.*, 109:080801, Aug 2012.

[14] A. Hees, S. Bertone, and C. Le Poncin-Lafitte. Frequency shift up to the 2-PM approximation. In S. Boissier, P. de Laverny, N. Nardetto, R. Samadi, D. Valls-Gabaud, and H. Wozniak, editors, *SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pages 145–148, Dec. 2012.

[15] B. Hofmann-Wellenhof and H. Moritz. *Physical geodesy*. Springer, 2006.

[16] N. Huntemann, M. Okhapkin, B. Lipphardt, S. Weyers, C. Tamm, and E. Peik. High-accuracy optical clock based on the octupole transition in $^{171}Yb^{+}$. *Phys. Rev. Lett.*, 108:090801, Feb 2012.

[17] S. Jevrejeva, J. C. Moore, and A. Grinsted. Sea level projections to AD2500 with a new generation of climate change scenarios. *Global and Planetary Change*, 80:14–20, Jan. 2012.

[18] S. M. Kopeikin. Relativistic Manifestations of gravitational fields in gravimetry and geodesy. *Manuscripta Geodaetica*, 16:301–312, Jan. 1991.

[19] S. M. Kopeikin, M. Efroimsky, and G. Kaplan. *Relativistic celestial mechanics of the solar system*. Wiley-VCH, 2011.

[20] C. Le Poncin-Lafitte, B. Linet, and P. Teyssandier. World function and time transfer: general post-Minkowskian expansions. *Class. Quantum Grav.*, 21:4463–4483, Sept. 2004.

[21] R. Le Targat, L. Lorini, Y. Le Coq, M. Zawada, J. Guéna, M. Abgrall, M. Gurov, P. Rosenbusch, D. G. Rovera, B. Nagórny, R. Gartman, P. G. Westergaard, M. E. Tobar, M. Lours, G. Santarelli, A. Clairon, S. Bize, P. Laurent, P. Lemonde, and J. Lodewyck. Experimental realization of an optical second with strontium lattice clocks. *Nature Communications*, 4:2109, July 2013.

[22] N. Leefer, C. T. M. Weber, A. Cingöz, J. R. Torgerson, and D. Budker. New limits on variation of the fine-structure constant using atomic dysprosium. *ArXiv e-prints*, Apr. 2013.

[23] N. D. Lemke, A. D. Ludlow, Z. W. Barber, T. M. Fortier, S. A. Diddams, Y. Jiang, S. R. Jefferts, T. P. Heavner, T. E. Parker, and C. W. Oates. Spin-1/2 optical lattice clock. *Phys. Rev. Lett.*, 103:063001, Aug 2009.

[24] O. Lopez, A. Haboucha, B. Chanteau, C. Chardonnet, A. Amy-Klein, and G. Santarelli. Ultra-stable long distance optical frequency distribution using the internet fiber network. *Opt. Express*, 20(21):23518–23526, Oct 2012.

[25] A. A. Madej, P. Dubé, Z. Zhou, J. E. Bernard, and M. Gertsvolf. $^{88}Sr^{+}$ 445-THz single-ion reference at the 10^{-17} level via control and cancellation of systematic uncertainties and its measurement against the SI second. *Phys. Rev. Lett.*, 109:203002, Nov 2012.

[26] J. J. McFerran, L. Yi, S. Mejri, S. Di Manno, W. Zhang, J. Guéna, Y. Le Coq, and S. Bize. Neutral atom frequency reference in the deep ultraviolet with fractional uncertainty $= 5.7 \times 10^{-15}$. *Phys. Rev. Lett.*, 108:183004, May 2012.

[27] C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. San Francisco: W.H. Freeman and Co., 1973.

[28] J. Müller, M. Soffel, and S. A. Klioner. Geodesy and relativity. *Journal of Geodesy*, 82:133–145, Mar. 2008.

[29] T. L. Nicholson, M. J. Martin, J. R. Williams, B. J. Bloom, M. Bishof, M. D. Swallows, S. L. Campbell, and J. Ye. Comparison of two independent Sr optical clocks with 1×10^{-17} stability at 10^3 **s**. *Phys. Rev. Lett.*, 109:230801, Dec 2012.

[30] T. E. Parker. Long-term comparison of caesium fountain primary frequency standards. *Metrologia*, 47(1):1, 2010.

[31] G. Petit and G. Panfilo. Comparison of frequency standards used for TAI. *Instrumentation and Measurement, IEEE Transactions on*, 62(6):1550–1555, 2013.

[32] G. Petit and P. Wolf. Relativistic theory for time comparisons: a review. *Metrologia*, 42:138, June 2005.

[33] K. Predehl, G. Grosche, S. M. F. Raupach, S. Droste, O. Terra, J. Alnis, T. Legero, T. W. Hänsch, T. Udem, R. Holzwarth, and H. Schnatz. A 920-kilometer optical fiber link for frequency metrology at the 19th decimal place. *Science*, 336(6080):441–444, 2012.

[34] T. Rosenband, D. B. Hume, P. O. Schmidt, C. W. Chou, A. Brusch, L. Lorini, W. H. Oskay, R. E. Drullinger, T. M. Fortier, J. E. Stalnaker, S. A. Diddams, W. C. Swann, N. R. Newbury, W. M. Itano, D. J. Wineland, and J. C. Bergquist. Frequency ratio of Al^+ and Hg^+ single-ion optical clocks; metrology at the 17th decimal place. *Science*, 319(5871):1808–1812, 2008.

[35] M. Soffel, H. Herold, H. Ruder, and M. Schneider. Relativistic theory of gravimetric measurements and definition of the geoid. *Manuscr. Geod.*, 13:143–146, 1988.

[36] M. Soffel, S. A. Klioner, G. Petit, P. Wolf, S. M. Kopeikin, P. Bretagnon, V. A. Brumberg, N. Capitaine, T. Damour, T. Fukushima, B. Guinot, T.-Y. Huang, L. Lindegren, C. Ma, K. Nordtvedt, J. C. Ries, P. K. Seidelmann, D. Vokrouhlický, C. M. Will, and C. Xu. The IAU 2000 Resolutions for Astrometry, Celestial Mechanics, and Metrology in the Relativistic Framework: Explanatory Supplement. *The Astronomical Journal*, 126:2687–2706, Dec. 2003.

[37] J. L. Synge. *Relativity, The general theory*. North-Holland Publishing Company, Amsterdam, 1960, 1960.

[38] M. Takamoto, F.-L. Hong, R. Higashi, and H. Katori. An optical lattice clock. *Nature*, 435:321–324, 2005.

[39] J. Terrien. News from the international bureau of weights and measures. *Metrologia*, 4(1):41, 1968.

[40] M. Vermeer. Chronometric levelling. *Reports of the Finnish Geodetic Institute*, 83(2):1–7, 1983.

[41] C. M. Will. *Theory and Experiment in Gravitational Physics*. Theory and Experiment in Gravitational Physics, by Clifford M.˜Will, pp.˜396.˜ISBN 0521439736.˜Cambridge, UK: Cambridge University Press, March 1993., 1993.

[42] C. M. Will. The Confrontation between General Relativity and Experiment. *Living Reviews in Relativity*, 9(3), 2006.

[43] P. Wolf and G. Petit. Relativistic theory for clock syntonization and the realization of geocentric coordinate times. *A&A*, 304:653, Dec. 1995.

[44] J. Ye and S. T. Cundiff. *Femtosecond optical frequency comb: Principle, Operation and Applications*. Springer, 2004.

 DOI: 10.2420/AF07.2013.67

Acta Futura 7 (2013) 79-85
DOI: 10.2420/AF07.2013.79

Acta Futura

Relativistic Positioning Systems and Gravitational Perturbations

Andreja Gomboc*,
Faculty of Mathematics and Physics, University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia
Centre of Excellence SPACE-SI, Aškerčeva cesta 12, 1000 Ljubljana, Slovenia
Uroš Kostić, Martin Horvat,
Faculty of Mathematics and Physics, University of Ljubljana, Jadranska ulica 19, 1000 Ljubljana, Slovenia
Sante Carloni,
ESA - Advanced Concepts Team, ESTEC, Keplerlaan 1, Postbus 299, 2200 AG Noordwijk, Netherlands
and Pacôme Delva
LNE-SYRTE, Observatoire de Paris, CNRS and UPMC ; 61 avenue de l'Observatoire, 75014 Paris, France

Abstract. In order to deliver a high accuracy relativistic positioning system, several gravitational perturbations need to be taken into account. We therefore consider a system of satellites, such as the Galileo system, in a space-time described by a background Schwarzschild metric and small gravitational perturbations due to the Earth's rotation, multipoles and tides, and the gravity of the Moon, the Sun, and planets. We present the status of this work currently carried out in the ESA Slovenian PECS project *Relativistic Global Navigation System*, give the explicit expressions for the perturbed metric, and briefly outline further steps.

1 Introduction

Current Global Navigation Satellites Systems (GNSS), such as the Global Positioning System and the European Galileo system, are based on Newtonian concept of absolute time and space. The signals from four satellites are needed for a receiver to determine its position and time via the time difference between the emission and the reception of the signal. However, due to the inertial reference frames and curvature of the space-time in the vicinity of Earth, space and time cannot be considered as absolute. In fact, general relativistic effects are far from being negligible [2]. Current GNSS deal with this problem by adding general relativistic corrections to the level of the accuracy desired. An alternative, and more consistent, approach is to abandon the concept of absolute space and time and describe a GNSS directly in general relativity, i.e. to define a Relativistic Positioning System (RPS) with the so-called emission coordinates [5, 16, 3, 4, 17]. A user of an RPS receives, at a given moment, four signals from four different satellites. It is able to determine the proper time τ of each satellite at the moment of emission of these signals. Then these four proper times $(\tau_1, \tau_2, \tau_3, \tau_4)$ constitute its emission coordinates. By receiving them at subsequent times, the receiver therefore knows its trajectory in the emission coordinates.

There are several advantages of an RPS. Firstly, the emission coordinates are covariant quantities; they are

*Corresponding author. E-mail: andreja.gomboc@fmf.uni-lj.si

independent of the observer (although dependent on the set of satellites chosen and their trajectories). Moreover, if each satellite broadcasts its own and also receives proper times of other satellites, the system of satellites is autonomous and constitutes a primary reference system, with no need to define a terrestrial reference frame. Therefore tracking of satellites with ground stations is necessary only to link an RPS to a terrestrial frame, although this link can also be obtained by placing several receivers at the known terrestrial positions. There is no need to synchronize satellite clocks to a time-scale realized on the ground (like it is done today with GPS time, which is a realization of the TT Terrestrial Time). No relativistic corrections are necessary, as relativity is already included in the definition of the positioning system.

To demonstrate feasibility, stability and accuracy of an RPS, two ESA Ariadna projects were carried out in 2010 and 2011 [6, 7, 8]. In these projects, the authors modelled an RPS in the idealized case of Schwarzschild geometry and found that relativistic description of a satellite constellation with inter-satellite links provides numerically accurate, stable and autonomous system. They called such a system the Autonomous Basis of Coordinates (ABC). By communicating their proper times solely, two satellites can determine their orbits (i.e., their constants of motion). Any additional satellite would serve to increase the system's accuracy.

Outline of the PECS Relativistic GNSS project and its goals

Continuation of this work is presently being carried out in the ESA Slovenian PECS project *Relativistic Global Navigation System* (2011-2014). In this project, we aim to demonstrate that an RPS and the ABC concept are highly accurate and stable also if the space-time is not purely spherically symmetric, but contains small gravitational perturbations due to the Earth's multipoles, tides and rotation and gravitational influences of the Moon, the Sun, Jupiter and Venus. In this work we use the same approach that was used in two Araidna projects (description of satellites' orbit in the ABC, emission coordinates, inter-satellite communication and recovery of their orbits) and add to a Hamiltonian gravitational perturbations.

In brief: we start modelling of an RPS by writing the perturbed Hamiltonian, corresponding equations of motion and calculating orbits of satellites. Then we define the system based on the ABC with four satellites, simulate their inter-satellite communication and use the emission/reception of their proper times to determine their orbital parameters. Our work is divided in the following steps:

1 Add first order gravitational perturbations to the Schwarzschild metric: find perturbation coefficients describing all known gravitational perturbations, i.e., due to Earth's multipoles, tides and rotation; gravity of the Moon, the Sun, and planets (Venus and Jupiter).

2 Solve the perturbed geodesic equations: use Hamiltonian formalism and perturbation theory to obtain time evolution of zeroth order constants of motion. Simulate satellites' orbits.

3 Find accurate constants of motion: use solely inter-satellite links over many orbital periods to determine satellites' orbital parameters and study the stability and possible degeneracies between them.

4 Refine values of gravitational perturbation coefficients: use additional satellites and residual errors between orbit prediction and orbit determination through inter-satellite communication to improve the accuracy of position determination and to probe the space-time, i.e., measure perturbations. Discuss possible scientific applications.

Before starting detailed calculations it is useful to get an estimation of the order of magnitude of various perturbations on a satellite orbit (as in e.g. [13, Fig. 3.1.]). At GNSS altitudes of about 20.000 km above Earth, the most important gravitational perturbations are due to Earth's multipoles, followed by the gravitational field of the Moon and the Sun. Several orders of magnitude smaller are perturbations due to Solar radiation pressure and the Earth's albedo (not considered in our project) and due to Earth's tides. About one order of magnitude smaller are relativistic effects and the gravitational influence of Jupiter and Venus. Relativistic effects due to Earth's rotation are about an order of magnitude smaller again.

Because our project is still in progress, we will present here only the results of the first part of the project: we will give the expressions for the perturbed metric around Earth.

 DOI: 10.2420/AF07.2013.79

2 Metric perturbations in the Schwarzschild background

We describe the spherically symmetric and time independent background with the Schwarzschild metric $g^{(0)}_{\mu\nu}$ and denote metric perturbations with $h_{\mu\nu}$. Because gravitational perturbations are several orders of magnitude smaller than the Earth's gravitational GM term ($h_{\mu\nu} \ll g^{(0)}_{\mu\nu}$), we use linear perturbation theory and write the perturbed metric as

$$g_{\mu\nu} = g^{(0)}_{\mu\nu} + h_{\mu\nu} \, . \tag{1}$$

Because we are interested in the space-time outside Earth, the perturbed metric must satisfy the Einstein equation for vacuum:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 0 \tag{2}$$

where

$$R_{\mu\nu} = R^{(0)}_{\mu\nu} + \delta R_{\mu\nu} \tag{3}$$

is the Ricci tensor (symbol $^{(0)}$ denotes unperturbed quantities and δ perturbations). The Einstein equation becomes:

$$\begin{aligned} &h_{\alpha}{}^{\alpha}{}_{;\mu\nu} - h_{\mu}{}^{\alpha}{}_{;\nu\alpha} - h_{\nu}{}^{\alpha}{}_{;\mu\alpha} + h_{\mu\nu;}{}^{\alpha}{}_{\alpha} \\ &+ g^{(0)}_{\mu\nu}(h_{\alpha}{}^{\lambda}{}_{;}{}^{\alpha}{}_{\lambda} - h_{\lambda}{}^{\lambda}{}_{;\alpha}{}^{\alpha}) - h_{\mu\nu} R^{(0)} \\ &+ g^{(0)}_{\mu\nu} h_{\lambda\sigma} R^{(0)\lambda\sigma} = 0 \end{aligned} \tag{4}$$

where a semi-colon (;) denotes covariant derivative, calculated with respect to the unperturbed metric $g^{(0)}_{\mu\nu}$.

To find solutions of these equations for vacuum we use the Regge-Wheeler-Zerilli (RWZ) framework [15, 18] . In the RWZ formalism, the metric perturbation $h_{\mu\nu}$ is expanded into a series of independent tensor harmonics, a tensor analog to spherical harmonic functions, labeled by indices n (degree), m (order), and parity: odd or even. By adopting the notation from [14], the general expansion of the metric perturbation $h_{\mu\nu}$ can be written as

$$h_{\mu\nu} = \sum_{n=2}^{\infty} \sum_{m=-n}^{n} (h^{nm}_{\mu\nu})^{(o)} + (h^{nm}_{\mu\nu})^{(e)} \, , \tag{5}$$

where the expansion terms $(h^{nm}_{\mu\nu})^{(o)}$ and $(h^{nm}_{\mu\nu})^{(e)}$ are the odd-parity and the even-parity metric functions (or modes), respectively. We find it most convenient to work in the gauge from [15], where the odd parity metric functions are

$$(h^{nm}_{\mu\nu})^{(o)} = \left[\begin{array}{cc|cc} 0 & 0 & -h_0 \csc\theta \partial_\varphi & h_0 \sin\theta \partial_\theta \\ 0 & 0 & -h_1 \csc\theta \partial_\varphi & h_1 \sin\theta \partial_\theta \\ \hline \star & \star & 0 & 0 \\ \star & \star & 0 & 0 \end{array}\right] Y^m_n \, , \tag{6}$$

and for even parity, the metric functions are

$$(h^{nm}_{\mu\nu})^{(e)} = \left[\begin{array}{cc|cc} H_0\chi & H_1 & 0 & 0 \\ \star & H_2\chi^{-1} & 0 & 0 \\ \hline 0 & 0 & r^2 K & 0 \\ 0 & 0 & 0 & r^2 K \sin^2\theta \end{array}\right] Y^m_n \, , \tag{7}$$

where $\star$ indicates the symmetric part of the tensor, $\chi = 1 - r_s/r$, and $r_s = 2GM/c^2$ is the Schwarzschild radius. Expressions h_i, H_i, and K depend on Schwarzschild coordinates (t, r) and indices (n, m), which are omitted in the expressions for clarity. It is shown in [15, 18] that in vacuum $H_0 = H_2$, therefore, both functions are marked with H.

After inserting (6) and (7) into (4), we obtain a number of partial differential equations. We find and write their solutions for time-independent and time-dependent perturbations, separately.

2.1 Time-independent metric perturbations

Even-parity contributions

In the case of time-independent perturbations of even parity $H_1 = 0$ [15]. It follows that the even metric mode (7) is diagonal:

$$\begin{aligned} &(h^{nm}_{\mu\nu})^{(e)} = \\ &\mathrm{diag}(H\chi, H\chi^{-1}, r^2 K, r^2 K \sin^2\theta) Y^m_n \, . \end{aligned} \tag{8}$$

Inserting it in (4) gives equations for H and K. We find the solution for H to be:

$$H(r) = A_{nm} \frac{P^{(0)}_n\left(\frac{r_s}{r}\right)}{r^n (r - r_s)} + B_{nm} \frac{r^{n+1} R^{(0)}_n\left(\frac{r_s}{r}\right)}{r - r_s} \, , \tag{9}$$

where the functions $P^{(0)}_n$ and $R^{(0)}_n$ are Gaussian hypergeometric functions ${}_2F_1$ [1] (for more details see[11]).

The solution for function K is determined by H (see [11]) and can be written as

$$K(r) = A_{nm} \frac{P^{(1)}_n\left(\frac{r_s}{r}\right)}{r^{n+1}} + B_{nm} r^n R^{(1)}_n\left(\frac{r_s}{r}\right) \, , \tag{10}$$

where functions $P_n^{(1)}$ and $R_n^{(1)}$ can be expressed via $P_n^{(0)}$ and $R_n^{(0)}$ [11].

In our case of satellites in orbit around Earth, time-independent perturbations are due to the Earth's multipoles (for the time being, we neglect Earth's rotation and tides). In order to determine the constants A_{nm} and B_{nm} in (9) and (10), we compare equation (9) with its Newtonian counterpart, i.e., the gravitational potential Φ. For a non-rotating object of mass M the potential Φ can be expanded into a series of multipole contributions (eq. 3.61 in [12]):

$$\Phi = \frac{GM}{r} + \sum_{nm} (M^{\oplus}_{nm} r^{-n-1} + M^{\ominus}_{nm} r^{n}) Y^m_n \,, \tag{11}$$

where $M^{\oplus}_{nm}$ and $M^{\ominus}_{nm}$ are time-independent spherical multipole momenta and notation $\sum_{nm} \equiv \sum_{n=2}^{\infty} \sum_{m=-n}^{n}$ is used. The first term in the sum describes the gravitational potential of the perturbing sources positioned within the radius r, while the second term corresponds to those outside r. Comparing (9) with (11), we notice the same behavior (i.e., the superposition of r^{+n} and r^{-n-1} functional dependence) in the perturbative part of (11) and it is evident that the coefficients A_{nm} and B_{nm} are related to the multipole momenta. The relation between both is found from the weak field approximation

$$\frac{c^2}{2}(1 + g_{00}) \sim \Phi \,. \tag{12}$$

By inserting

$$g_{00} = g^{(0)}_{00} + \sum_{nm} (h^{nm}_{00})^{(e)} = \chi \left(-1 + \sum_{nm} H_{nm} \right) \tag{13}$$

into the above relation together with the Newtonian potential (11), we find that in the weak field limit A_{nm} and B_{nm} are asymptotically related to Newtonian spherical multipole momenta $M^{\oplus}_{nm}$ and $M^{\ominus}_{nm}$ as

$$A_{nm} \sim \frac{2}{c^2} M^{\oplus}_{nm} \quad \text{and} \quad B_{nm} \sim \frac{2}{c^2} M^{\ominus}_{nm} \,. \tag{14}$$

Note that for finite c, $M^{\oplus}_{nm}$ and $M^{\ominus}_{nm}$ only approximate A_{nm} and B_{nm}.

Odd-parity contributions

In case of time-independent perturbations, the odd metric functions $(h^{nm}_{\mu\nu})^{(o)}$ in (6) have $h_1 = 0$ [15] and can be written with a single function h_0 as

$$(h^{nm}_{\mu\nu})^{(o)} = -h_0 \csc\theta \, Y^m_{n\,,\varphi} (\delta_{0,\mu}\delta_{2,\nu} + \delta_{2,\mu}\delta_{0,\nu}) + h_0 \sin\theta \, Y^m_{n\,,\theta} (\delta_{0,\mu}\delta_{3,\nu} + \delta_{3,\mu}\delta_{0,\nu}) \,. \tag{15}$$

The solution for h_0 is:

$$h_0(r) = \alpha_{nm} \frac{P^{(2)}_n\left(\frac{r_s}{r}\right)}{r^n} + \beta_{nm} r^{n+1} R^{(2)}_n \left(\frac{r_s}{r}\right) \,, \tag{16}$$

where functions $P^{(2)}_n$ and $R^{(2)}_n$ are Gaussian hypergeometric functions ${}_2F_1$ (for more details see[11]).

To determine the constants α_{nm} and β_{nm} we note that off-diagonal terms in the metric tensor are associated with frame-dragging effects. In our case frame-dragging effects from 'external' objects (the Sun, the Moon, other planets) are negligible. Consequently, we set $\beta_{nm} = 0$.

We do take into account the frame-dragging due to the Earth's rotation. To determine α_{nm}, we notice that for $n = 1$ and $m = 0$ the corresponding h_0 matches the weak field and slow rotation approximation of the Kerr metric: if $r \gg r_s$ and Earth's angular parameter $a \ll 1$, then for $\alpha_{10} = a r_s \sqrt{4\pi/3}$ it follows

$$h_0(r) = a \frac{r_s}{r} \sqrt{\frac{4\pi}{3}} \,, \tag{17}$$

where we keep only the terms linear in a.

For higher multipoles ($n > 1$), it turns out that their dependence on a is not linear [10]. Therefore, the only multipole we include in the odd-parity metric function is the monopole, i.e., the one belonging to the linear (in a) part of the Kerr effect.

2.2 Time-dependent metric perturbations

Due to the Earth's rotation its multipoles vary periodically. Earth's solid and ocean tides introduce additional time dependency in its multipoles (additional variability with different frequency, phase and varying amplitude, depending on the position of the Moon and the Sun). In addition, the gravitational influence of other celestial bodies introduces time dependent perturbations to the space-time around Earth, because their relative positions with respect to Earth change with time. These

 DOI: 10.2420/AF07.2013.79

perturbations can be expanded in a series of multipoles and treated with the same procedure as the Earth's multipoles.

We therefore consider time dependent metric perturbations for the case of perturbations oscillating slowly with angular velocities, which are smaller or of the same order of magnitude as the angular velocity of Earth. All angular velocities are defined with respect to the Schwarzschild time t.

Even-parity contributions

Even parity modes are connected to the Newtonian gravitational potential Φ, which in the case of time dependent multipoles can be written as:

$$\Phi = \frac{GM}{r} + \sum_{nm} (M^{\oplus}_{nm}(T) r^{-n-1} + M^{\ominus}_{nm}(T) r^{n}) Y^{m}_{n} , \tag{18}$$

where $T = ct$, or alternatively in frequency domain:

$$\Phi = \frac{GM}{r} + \sum_{nm} \int_{-\infty}^{\infty} dk\, e^{ikT} \times \left[\widetilde{M}^{\oplus}_{nm}(k) r^{-n-1} + \widetilde{M}^{\ominus}_{nm}(k) r^{n} \right] Y^{m}_{n} , \tag{19}$$

where k is the wavenumber and $\widetilde{M}^{\oplus}_{nm}$, $\widetilde{M}^{\ominus}_{nm}$ are the Fourier transforms of time dependent multipoles:

$$\widetilde{M}^{u}_{nm}(k) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dT\, e^{-ikT} M^{u}_{nm}(T) , \tag{20}$$

where $u = \oplus, \ominus$.

Each time dependent multipole generates a time-dependent even metric perturbation $(h^{nm}_{\mu\nu})^{(e)}$. Functions H, H_1, and K determining the modes can be expressed with their Fourier transforms as

$$(H(T), H_1(T), K(T)) = \int_{-\infty}^{\infty} dk\, e^{ikT} (\widetilde{H}(k), \widetilde{H}_1(k), \widetilde{K}(k)) \tag{21}$$

Because in our case k is very small ($k \ll 1/r$), we solve differential equations for H, H_1, and K perturbatively in k. We assume that $\widetilde{H}$, $\widetilde{H}_1$, and $\widetilde{K}$ are smooth functions of k and write them as a power series of dimensionless $\kappa = k r_s$:

$$(\widetilde{H}, \widetilde{H}_1, \widetilde{K}) \sim \sum_{i=0}^{\infty} \kappa^{2i} (\widetilde{H}^{(i)}, i\kappa \widetilde{H}^{(i)}_1, \widetilde{K}^{(i)}) , \tag{22}$$

The first terms in the expressions for H and K are already known: $\widetilde{H}^{(0)}$ is given in (9) and $\widetilde{K}^{(0)}$ in (10), where instead of (14) we use:

$$A_{nm} \sim \frac{2}{c^2} \widetilde{M}^{\oplus}_{nm} \quad \text{and} \quad B_{nm} \sim \frac{2}{c^2} \widetilde{M}^{\ominus}_{nm} . \tag{23}$$

For $\tilde{H}_1$ we find that

$$\widetilde{H}^{(0)}_1(r) = A_{nm} \frac{r^{-n+1} P^{(3)}_n\left(\frac{r_s}{r}\right)}{r_s(r - r_s)} + B_{nm} \frac{r^{n+2} R^{(3)}_n\left(\frac{r_s}{r}\right)}{r_s(r - r_s)} , \tag{24}$$

where functions $P^{(3)}_n$ and $R^{(3)}_n$ are given as a series in $\frac{r_s}{r}$ [11].

Because k is very small for all time-dependent perturbations considered, we neglect all higher than leading terms in the expansion (22) and use the following approximations:

$$H(T, r) \approx \int_{-\infty}^{\infty} dk\, e^{ikT} \widetilde{H}^{(0)}(k, r) , \tag{25}$$

$$H_1(T, r) \approx \int_{-\infty}^{\infty} dk\, ikr_s e^{ikT} \widetilde{H}^{(0)}_1(k, r) , \tag{26}$$

$$K(T, r) \approx \int_{-\infty}^{\infty} dk\, e^{ikT} \widetilde{K}^{(0)}(k, r) . \tag{27}$$

A metric perturbation expressed with these functions is accurate up to the linear order in frequency. Since higher order perturbations naturally give rise to contributions with higher orders of frequencies, our approximation of a perturbation is consistently linear, i.e., it is linear in frequencies and in the order of perturbation.

Odd-parity contributions

For odd-parity solutions we use the same notation as in (21) and we find that the asymptotic behaviour of solutions $\tilde{h}_0$ and $\tilde{h}_1$ is not flat:

$$\tilde{h}_1(r) \asymp r \sin(kr + \phi) \tag{28}$$

$$\tilde{h}_0(r) \asymp r \cos(kr + \phi) \tag{29}$$

These solutions are therefore not relevant in our case.

3 Metric around Earth

Finally, we can write the metric perturbation $h_{\mu\nu}$, which in (5) was expressed as a series of normal modes

$(h^{nm}_{\mu\nu})^{(o)}$ and $(h^{nm}_{\mu\nu})^{(e)}$. Based on the positions of the sources of perturbations, these modes can be grouped into two terms:

$$h_{\mu\nu} = h^{\oplus}_{\mu\nu} + h^{\ominus}_{\mu\nu} \,. \tag{30}$$

The term $h^{\oplus}_{\mu\nu}$ represents the Earth's time dependent (exterior) multipoles and the frame-dragging effect of Earth. The former arise from the shape of the Earth, which changes with time due to rotation and tidal forces. For the latter, Kerr effect, there is no non-relativistic counterpart.

The term $h^{\ominus}_{\mu\nu}$ represents the time dependent (interior) multipoles of other celestial bodies. This term arises from the perturbative effect of other planets, the Moon, and the Sun, whose positions relative to Earth's change with time. Their frame-dragging effect is neglected.

To simplify expressions, we introduce the normalized complex multipoles ($u = \oplus, \ominus$):

$$\overline{M}^{u}_{nm} := \frac{2}{c^2} M^{u}_{nm} \,. \tag{31}$$

Metric perturbation due to Earth's multipoles and rotation can be written as

$$\begin{aligned}
[h^{\oplus}_{\mu\nu}] &= \sum_{nm} \overline{M}^{\oplus}_{nm} Y^{m}_{n} \times \\
&\operatorname{diag}\left(\frac{P^{(0)}_n}{r^{n+1}}, \frac{P^{(0)}_n}{r^{n-1}(r-r_s)^2}, \frac{P^{(1)}_n}{r^{n-1}}, \frac{P^{(1)}_n \sin^2\theta}{r^{n-1}} \right) \\
&+ \sum_{nm} \overline{M}^{\oplus}_{nm,T} Y^{m}_{n} \frac{P^{(3)}_n}{r^{n-1}(r-r_s)} (\delta_{\mu,1}\delta_{\nu,0} + \delta_{\nu,1}\delta_{\mu,0}) \\
&- a_{\oplus} \frac{r_s}{r} \sin^2\theta (\delta_{\mu,3}\delta_{\nu,0} + \delta_{\nu,3}\delta_{\mu,0}) \,,
\end{aligned} \tag{32}$$

where Earth's multipoles $\overline{M}^{\oplus}_{nm}$ are functions of time and include rotation, ocean and solid tides; and $\overline{M}^{\oplus}_{nm,T}$ are their time derivatives.

Metric perturbations due to other celestial bodies are

$$\begin{aligned}
[h^{\ominus}_{\mu\nu}] &= \sum_{nm} \overline{M}^{\ominus}_{nm} Y^{m}_{n} \times \\
&\operatorname{diag}\left(r^n R^{(0)}_n, \frac{r^{n+2} R^{(0)}_n}{(r-r_s)^2}, r^{n+2} R^{(1)}_n, r^{n+2} R^{(1)}_n \sin^2\theta \right) \\
&+ \sum_{nm} \overline{M}^{\ominus}_{nm,T} Y^{m}_{n} \frac{r^{n+2} R^{(3)}}{r-r_s} (\delta_{\mu,1}\delta_{\nu,0} + \delta_{\nu,1}\delta_{\mu,0}) \,,
\end{aligned} \tag{33}$$

where $\overline{M}^{\ominus}_{nm}$ are summed multipoles of other celestial bodies (in our case the Sun, the Moon, Jupiter and Venus).

The first order approximations of the metric perturbations given by (32) and (33) are fully determined by multipole momenta $\overline{M}^{\oplus}_{nm}$, $\overline{M}^{\ominus}_{nm}$, Kerr parameter a, and functions $P^{(i)}_n$ and $R^{(i)}_n$ (for details see [11]).

4 Conclusions

The aim of our project is to model the Galileo GNSS in general relativity including all relevant gravitational perturbations and to test to what level can this new approach improve the accuracy and stability of the Galileo GNSS reference frame.

In this contribution, we show how we can include gravitational perturbations with linear perturbation theory on a Schwarzschild background and present explicit expressions (32) and (33) for the perturbative metric.

In the next steps, we will use this metric in the perturbative Hamiltonian formalism to obtain time derivatives of zeroth order constants of motion. With the derivatives known, we will be able to determine the time evolution of these slowly changing constants of motion and apply them to analytical solutions for Schwarzschild geodesics to obtain satellites' orbits in perturbed space-time. Once the orbits are known, we can use them to do the relativistic positioning, as well as build the ABC in the perturbed space-time around Earth, using only inter-link communication between GNSS satellites.

Acknowledgements

AG, MH and UK acknowledge the financial support by ESA Slovenian PECS project *Relativistic Global Navigation System*.

References

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, fifth edition, 1964.

[2] N. Ashby. Relativity in the global positioning system. *Living Reviews in Relativity*, 6(1):43, 2003.

[3] M. Blagojević, J. Garecki, F. W. Hehl, and Y. N. Obukhov. Real null coframes in general relativity and GPS type coordinates. *Phys. Rev. D*, 65(4):044018–+, Feb. 2002.

 DOI: 10.2420/AF07.2013.79

[4] B. Coll. A principal positioning system for the Earth. In N. Capitaine and M. Stavinschi, editors, *Journées 2002 - systèmes de référence spatio-temporels. Astrometry from ground and from space, Bucharest, 25 - 28 September 2002, edited by N. Capitaine and M. Stavinschi, Bucharest: Astronomical Institute of the Romanian Academy, Paris: Observato*, volume 14, pages 34–38, 2003.

[5] B. Coll and J. A. Morales. Symmetric frames on Lorentzian spaces. *Journal of Mathematical Physics*, 32(9):2450, 1991.

[6] A. Čadež, U. Kostić, and P. Delva. Mapping the spacetime metric with a global navigation satellite system, european space agency, the advanced concepts team, ariadna final report (09/1301). Technical report, European Space Agency, 2010.

[7] A. Čadež, U. Kostić, and P. Delva. Mapping the spacetime metric with a global navigation satellite system-extension study: Recovery of orbital constants using intersatellite links, european space agency, the advanced concepts team, ariadna final report (09/1301 ccn). Technical report, European Space Agency, 2011.

[8] P. Delva, U. Kostić, and A. Čadež. Numerical modeling of a Global Navigation Satellite System in a general relativistic framework. *Advances in Space Research*, 47:370–379, Jan. 2011.

[9] V. Gurzadyan, J. Makino, M. J. Rees, G. Meylan, R. Ruffini, and J. A. Wheeler. *Black Holes, Gravitational Waves and Cosmology*. Advances in Astronomy and Astrophysics Series. Cambridge Scientific Publishers Limited, 2005.

[10] J. B. Hartle. Slowly Rotating Relativistic Stars. I. Equations of Structure. *Astrophysical Journal*, 150:1005, Dec. 1967.

[11] M. Horvat, U. Kostić, and A. Gomboc. Perturbing the Schwarzschild metric with Newtonian multipoles. *submitted*, 2013.

[12] J. D. Jackson. *Classical Electrodynamics Third Edition*. Wiley, third edition, Aug. 1998.

[13] O. Montenbruck and E. Gill. *Satellite Orbits: Models, Methods, Applications*. Springer Verlag, 2005.

[14] A. Nagar and L. Rezzolla. Gauge-invariant non-spherical metric perturbations of Schwarzschild black-hole spacetimes. *Classical and Quantum Gravity*, 22(16):R167, 2005.

[15] T. Regge and J. A. Wheeler. Stability of a Schwarzschild Singularity. *Phys. Rev.*, 108:1063–1069, Nov 1957.

[16] C. Rovelli. GPS observables in general relativity. *Phys. Rev. D*, 65(4):044017, Feb. 2002.

[17] A. Tarantola, L. Klimes, J. M. Pozo, and B. Coll. Gravimetry, Relativity, and the Global Navigation Satellite Systems. *ArXiv e-prints*, 2009:35, 5 2009.

[18] F. J. Zerilli. Gravitational Field of a Particle Falling in a Schwarzschild Geometry Analyzed in Tensor Harmonics. *Phys. Rev. D*, 2:2141–2160, Nov 1970. See erratum by Zerilli in Appendix A-7 of [9].

Acta Futura 7 (2013) 87-96
DOI: 10.2420/AF07.2013.87

Acta Futura

Preliminary study for the measurement of the Lense-Thirring effect with the Galileo satellites

Beatriz Moreno Monge, Rolf Koenig*, Grzegorz Michalak,
GFZ German Research Center for Geosciences, Germany

Ignazio Ciufolini,
Dipartimento di Ingegneria dell'Innovazione, Universitá del Salento, Italy

Antonio Paolozzi, and Giampiero Sindoni
Scuola di Ingegneria Aerospaziale and DIAEE, Sapienza Universitá di Roma, Italy

Abstract. The precession of the orbital node of a particle orbiting a rotating mass is known as Lense-Thirring effect (LTE) and is a manifestation of the general relativistic phenomenon of dragging of inertial frames or frame-dragging. The LTE has already been measured by using the node drifts of the LAGEOS satellites and GRACE-based Earth gravity field models with an accuracy of about 10% and will be improved down to a few percent with the recent LARES experiment. The Galileo system will provide 27 new node observables for the LTE estimation and their combination with the LAGEOS and LARES satellites can potentially reduce even more the error due to the mismodeling in Earth's gravity field. However, the accurate determination of the Galileo orbits requires the estimation of many different parameters, which can absorb the LTE on the orbital nodes. Moreover, the accuracy of the Galileo orbits and hence, of their node drifts, is mainly limited by the mismodeling in the Solar Radiation Pressure (SRP). Using simulated data we analyze the effects of the mismodeling in the SRP on the Galileo nodes and propose optimal orbit parameterizations for the measurement of the LTE from the future Galileo observations.

*Corresponding author. E-mail: rolf.koenig@gfz-potsdam.de

1 Introduction

In 1918, Lense and Thirring [9] proved that a particle orbiting around a central body endowed with an angular momentum J experiences a nodal precession $\dot{\Omega}$ according to the expression

$$\dot{\Omega} = \frac{2GJ}{c^2 a^3 (1-e^2)^{3/2}} \tag{1}$$

where a and e are the orbital semimajor axis and eccentricity, G is the gravitational constant and c is the speed of light.

One of the main error sources in the estimation of the Lense-Thirring effect (LTE) by using the nodes of an Earth satellite arises from the uncertainties in the Earth's gravity field model. In particular, the largest errors proceed from uncertainties in the first even zonal harmonics of a spherical expansion of the Earth potential, i. e., J2, J4, J6, ..., and their variations in time.

The combination of two node observables in order to remove the influence of the first even zonal

harmonic, J2, on the node drift was first proposed in [2]. Then, the two laser-ranged LAGEOS satellites and the high-accurate Earth gravity field models based on GRACE observations (f. i., EIGEN-GRACE02S, EIGEN-GRACE03S, JEM03G), have provided for the measurement of the LTE with an accuracy of about 10% [3]. The combination with the new LARES satellite [12] will allow to eliminate the uncertainties due to both, J2 and J4, thus being possible to obtain a measurement of the LTE with an accuracy of the order of 1% [3].

The Galileo system will provide a new node observable from a total of 27 satellites, whose combination with the LAGEOS and LARES satellites will potentially reduce the uncertainty due to the mismodeling in the Earth gravity field. However, there are two issues that greatly impact the LTE measurement with the Galileo satellites. The first of those is that the accurate determination of the Galileo orbits and their node drift requires the estimation of many different parameters, such as initial state vectors, empirical accelerations, clock offsets, station coordinates, etc, which can absorb partially or completely the LTE. The second issue is that the final accuracy of the estimated Galileo orbits is mainly limited by the mismodeling in the Solar Radiation Pressure (SRP), which constitutes the primary error source in the determination of the LTE with the Galileo nodes.

In the present work, the effects of both, the mismodeling in the SRP and the orbit parameterization, on the Galileo node drift determination are analyzed by means of simulated Galileo orbits and observations.

2 Simulation of Galileo orbits and observations

A set of Galileo orbits and observations corresponding to days 1-4 of January, 2008 are simulated and used in a series of different numerical tests. In the simulation, the EPOS-OC software (Earth Parameter and Orbit System– Orbit Computation [16]) is used.

The Galileo orbits are simulated according to the specifications given by ESA [5], a Walker 27/3/1 configuration with a semi-major axis of 29600 km, an inclination of 56 deg, zero eccentricity, a separation between planes equal to 120 deg and a change of mean anomaly for equivalent satellites in neighbouring planes of 13.3 deg. The main physical background models used in the simulations are presented in Table 1. The corresponding datasets of Galileo code and phase observations for a global network of 80 stations are obtained. In the data simulation and in the subsequent orbit recovery, identical background models are used and hence, the errors due to uncertainties in these models are not addressed in this work.

In EPOS-OC the SRP acceleration $\ddot{\mathbf{r}}$ is computed by means of the expression

$$\ddot{\mathbf{r}} = \left(\frac{A}{R}\right)^2 \frac{1}{m} F_{rad} \frac{\mathbf{R}}{R} \tag{2}$$

where A represents the Astronomical Unit (AU) in meters, m is the satellite mass, $\mathbf{R}$ is the heliocentric vector pointing out to the satellite with module R and F_{rad} is the direct pressure force computed by means of a model. For the time being, the SRP models implemented in EPOS-OC are the ROCK4 model [6] for GPS-type satellites and the cannon ball or macro models dedicated to specific non-GPS satellites (f. i., [10]). In our simulations the Galileo satellites were considered to be GPS-Block-II-like satellites and thus, the ROCK4 model was applied. The ROCK4 model is recommended by the IERS conventions [11] for the modeling of the SRP effect. It includes the dimensions and optical properties of the GPS spacecraft surfaces.

Gravity field	EIGEN-6C 12 x 12 [7]
Earth tide	IERS Conv. [14]
Ocean tide	EOT11a [15]
Atmospheric tide	Biancale-Bode [1]
Lunisolar and planetary perturbations	JPL DE421 [8]
Ocean pole tide	Desai [4]
Earth Orientation Parameters	EOP08C04
Nutation and precession	IERS Conv. [14]
Earth albedo	Analytic model by Heurtel

Table 1. *Physical background models used in the simulation of the Galileo orbits and observations*

In order to account for deficiencies in the SRP model, the Eq. (2) is multiplied by a global scaling factor (F0 hereafter) and different parameters can also be added, like global biases in the X, Y and Z directions in a satellite body-fixed reference system. The influence of the Earth's and Moon's shadow is also considered.

 DOI: 10.2420/AF07.2013.87

The creation of a more accurate SRP model for Galileo will require the knowledge of the vehicles characteristics in terms of shape, size, weight and surface optical properties. Then, a macro model could be adopted for the Galileo SRP computation, along with an appropriate attitude model which also accounts for the yaw turns during satellite midnights (shadow turns) and noons. As of today, little information has been published about the characteristics of the Galileo satellites. Some general features taken from [5] are compiled in Table 2.

Bus dimensions	2.7 x 1.1 x 1.2 m^3
Solar array span	13 m
Mass	700 Kg

Table 2. *Galileo satellite features*

3 Optimal parameterizations for the LTE estimation

According to Eq. (1), the nodal precession of the Galileo satellites due to the LTE is $1.7 \cdot 10^{-9}$ deg/d. This node drift holds for all 27 Galileo satellites of the full constellation. Figure 1 shows the comparison of two sets of *simulated* orbits over a period of 26 h generated with and without modeling of the LTE.

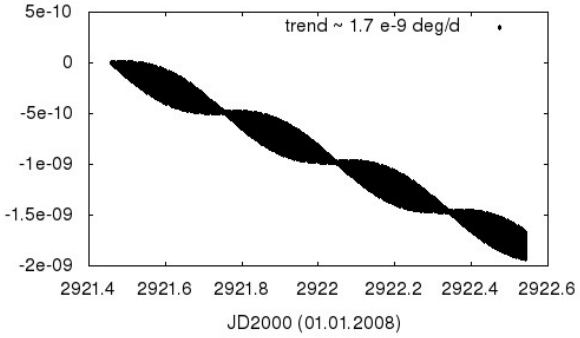


Figure 1. *Orbital node differences (deg) between two sets of simulated Galileo orbits with and without LTE.*

To measure the LTE from Galileo observations the nodal drift due to the LTE must be present in the node positions *estimated* in a precise orbit determination process where the LTE is not modeled. This means that the estimated node positions without LTE modeling must differ from the true node positions and this difference must not be absorbed by the estimated parameters. In order to find out under which circumstances this holds, a set of 26 h of Galileo orbits and noise-free code and phase observations are simulated including the modeling of the LTE. Then, the simulated observations are used to recover the Galileo orbits without modeling the LTE, thereby estimating the following parameters:

- Initial orbital elements for each satellite (position and velocity)
- Empirical accelerations in the along-track and normal directions (4 coefficients per arc)
- SRP F0, Y- and Z-bias
- Earth's albedo scaling factor
- Station coordinates
- Tropospheric delays (10 per satellite-station pair)
- Phase ambiguities
- Clock offsets

The *estimated* (without LTE) and *simulated* (with LTE) node positions are compared and the differences analyzed in order to identify the parameters absorbing the LTE. An optimal parameterization for the LTE measurement must allow to observe the node drift when comparing the estimated and simulated orbits.

It was found that the free estimation of the station coordinates introduces large errors in the node positions of the order of 10^{-8} deg and, as a consequence, the station coordinates must be either fixed or highly constrained, what can be done by imposing a set of No-Net-Translation-Rotation-Scale (NNTRS) conditions on the whole ground network. In addition, the estimation of empirical accelerations in the normal direction to the orbital plane absorbs part of the LTE and consequently, their estimation must be avoided, with a concurrent loss in orbit recovery accuracy. Provided that these two conditions are fulfilled the LT signal is observed in the differences between the estimated and simulated nodes positions, like in Fig. 2 (a). A linear regression of the node position differences provides an estimate of the LTE drift of approximately $1.5 \cdot 10^{-9}$ deg/d, close to the expected value, $1.7 \cdot 10^{-9}$ deg/d, with a standard deviation of $3 \cdot 10^{-12}$ deg/d and a small post-regression RMS of $9 \cdot 10^{-11}$ deg. In conclusion, this

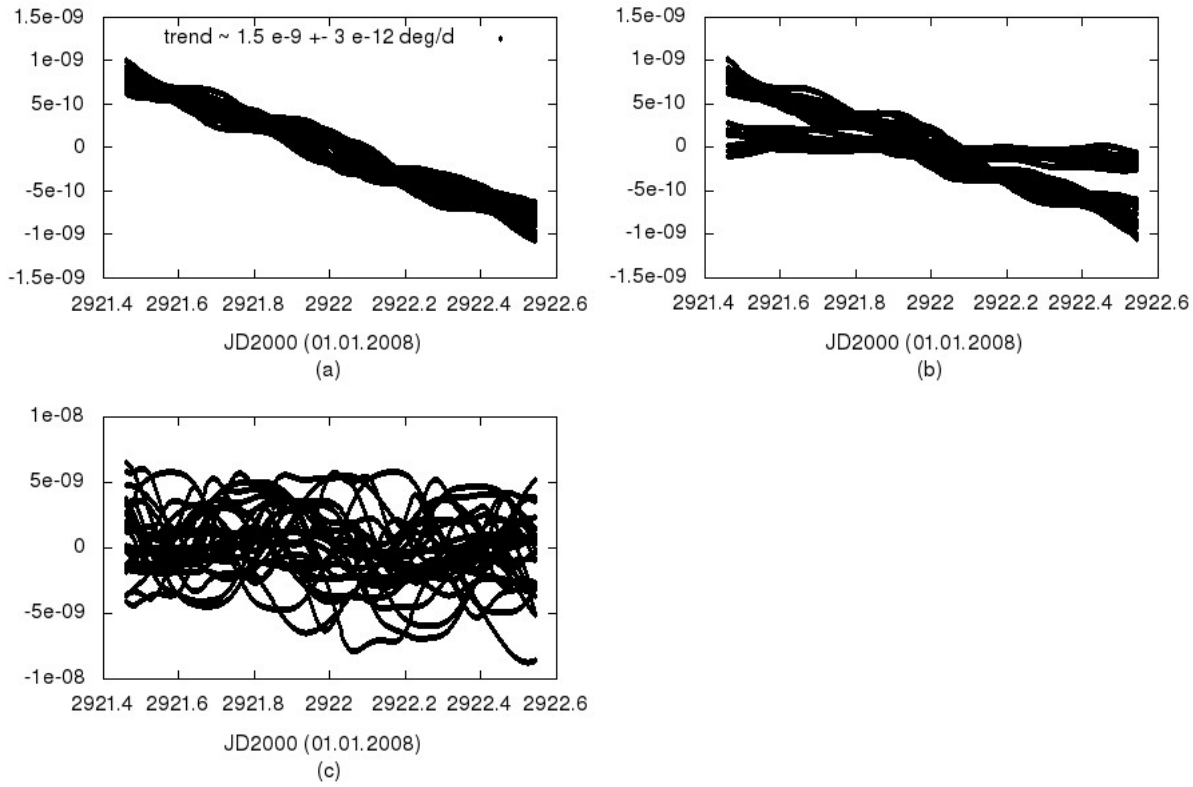


Figure 2. *Orbital node differences (deg) between estimated (without LTE) and simulated (with LTE) Galileo orbits when (a) station coordinates are constrained with NNTRS conditions and the empirical accelerations in the orbit normal direction are fixed (optimal parameterization), (b) the X-bias is additionally estimated and (c) the optimal parameterization is applied to noisy observations.*

parameterization seems to be optimal for the LTE estimation from 1-day Galileo arcs in the sense that the LTE is not absorbed by the estimated parameters.

An additional test has been performed by estimating another parameter of the SRP model, the X-bias, in addition to F0, Y- and Z-bias. The differences in the node positions are presented in Fig. 2 (b). In this case the LT signal has been absorbed for some satellites and hence, the estimation of the SRP X-bias must be avoided. This can be safely done, since the estimation of the X-bias does not introduce a significant improvement in the accuracy of the recovered orbit and therefore is usually dismissed.

The results obtained when applying the optimal parameterization to noisy observations are also analyzed. For this purpose, a Gaussian noise has been introduced to the simulated observations with standard deviations of 50 cm for code and 3 mm for phase ranges. The differences between estimated and simulated node positions are shown in Fig. 2 (c). In this case, the node drifts due to the LTE is hidden behind the large noise. In fact, the RMS of the node differences is about $2.7 \cdot 10^{-9}$ deg, quite larger than the expected node displacement from the LTE for 1 day, i. e., $1.7 \cdot 10^{-9}$ deg/d. Therefore, in order to obtain a more precise estimation of the LTE, longer arcs shall be used, f. i. with 3-day arcs the LTE amounts to $5.1 \cdot 10^{-9}$ deg/d, which is significantly larger than the noise mentioned before.

Thus, for the next tests, 3-day Galileo orbits and noise-free observations are simulated with modeling of the LTE. Then, the Galileo orbits are estimated without modeling the LTE from the observations by applying the optimal parameterization for 1-day arcs. The same number of empirical coefficients are estimated for the 3-day arcs as for the 1-day arcs. The differences between the estimated and simulated node positions are presented in Fig. 3 (a), where it can be observed that the LT signal is absorbed for some satellites and hence, the

 DOI: 10.2420/AF07.2013.87

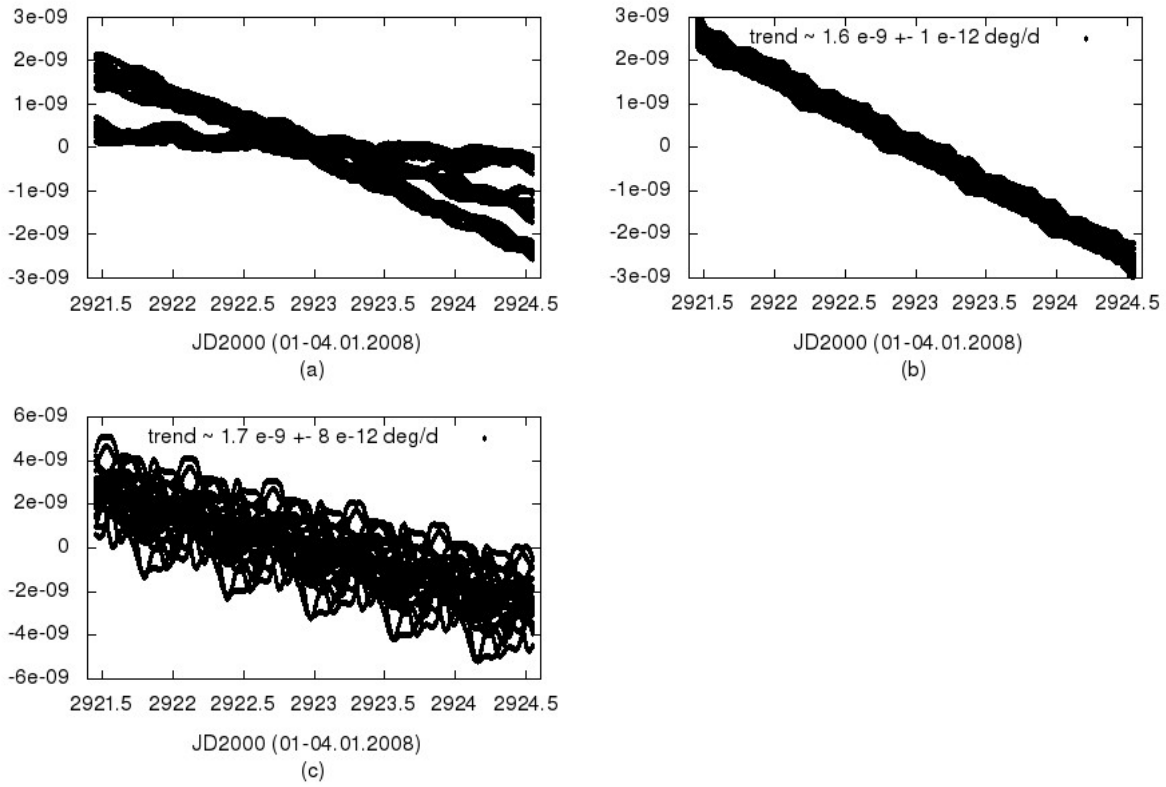


Figure 3. *Orbital node differences (deg) between estimated (without LTE) and simulated (with LTE) 3-day Galileo orbits when (a) the optimal parameterization for 1-day arcs is applied, (b) the Earth albedo is additionally fixed (optimal parameterization for 3-day arcs) and (c) the optimal parameterization for 3-day arcs is applied to noisy observations.*

optimal parameterization for 1-day arcs is not suitable for 3-day arcs.

The parameters absorbing the LTE are identified as the SRP Y-bias and the Earth's albedo scaling factor. Fixing either of them results in a perfect recovery of the LTE, as shown in Fig. 3 (b). Nevertheless, fixing the Y-bias to an incorrect value would introduce large errors in the node positions and has to be avoided. On the contrary, the influence of Earth's albedo on the Galileo orbits is very small and therefore can safely be neglected. Consequently, the optimal parameterization for 3-day arcs coincides with that for 1-day arcs with the exception of the Earth's albedo scaling factor, which must be fixed for the 3-day arcs.

Finally, the optimal parameterization for 3-day arcs is applied to noisy observations, the differences in the node positions are in Fig. 3 (c). The linear regression provides a trend of $1.7 \cdot 10^{-9}$ deg/d, with an RMS of $1.0 \cdot 10^{-9}$ deg, 5 times smaller than the total displacement of the Galileo nodes due to the LTE after 3 days. In conclusion, the 3-day arcs in combination with the optimal parameterization proposed here seem suitable for the precise measurement of the LTE from noisy Galileo observations in absence of other modeling errors. Then, using real Galileo observations and estimated orbits, the LTE will be measured as the difference between the estimated node positions in the overlap of two consecutive Galileo orbital arcs. Assuming a precision of the estimated node positions at the level of the RMS, this is $1.0 \cdot 10^{-9}$ deg for 3-day arcs, the combination of 27 node observables to compute the difference in the overlap yields a precision of $1.0 \cdot 10^{-9} \cdot \sqrt{2}/\sqrt{27} = 0.27 \cdot 10^{-9}$ deg for the LTE estimation. Thus, a precision of about 5% of the LTE would be reached with two 3-day Galileo arcs. Similarly, the average of LTE estimations obtained in a series of n overlaps increases the precision by a factor $1/\sqrt{n}$. Thus, in order to achieve a precision in the LTE estimation of 1%, a minimum of 30 Galileo 3-day arcs need to be

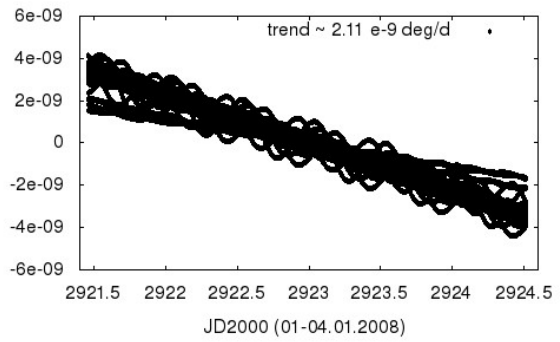


Figure 4. *Orbital node differences (deg) between estimated 3-day GPS orbits with and without modeling the LTE when the optimal parameterization is applied.*

processed.

The optimal parameterization has been verified with real GPS data through the estimation of a 3-day GPS orbit in two different cases: with and without modeling of the LTE. The differences in the node positions between the two sets of estimated orbits are in Fig. 4 and show a drift of the GPS nodes of about $2.1 \cdot 10^{-9}$ deg/d, close to the theoretical value of $2.3 \cdot 10^{-9}$ deg/d. This means that this optimal parameterization allows to observe the LTE in the GPS nodes provided that there are no errors in the background models and the nodes are accurately estimated to the few 10^{-9} deg level. This however is difficult to achieve mainly due to SRP modeling errors.

4 Effects of the SRP on the Galileo nodes

In this section we analyze the effects of the SRP on the Galileo nodes and the errors in the *estimated* nodes due to mismodeling of SRP. For that purpose, a set of Galileo orbits is simulated by including various parameters of the SRP model, then they are compared to a set simulated without SRP model. In first place a set of Galileo orbits with F0 equal to 1 and X-, Y- and Z-biases equal to 0 is tested. The effect is shown in Fig. 5 (a). A periodical displacement of the node is observed with an amplitude of up to $3 \cdot 10^{-5}$ deg and a period equal to a complete revolution of the Galileo satellites. Moreover the SRP introduces a drift of $7 \cdot 10^{-9}$ deg/d, about twice the level of the LTE.

The effects due to the presence of the X-, Y- and Z-bias (all simulated with magnitude 10^{-10} m/s^2) are shown in Figs. 5 (b), (c) and (d) respectively. It can be observed that the X- and Y-bias yield a significant trend of up to $6 \cdot 10^{-8}$ deg/d and up to $1 \cdot 10^{-7}$ deg/d respectively, in both cases depending on the orbital planes. Conversely, the Z-bias produces a periodical displacement of the node with an amplitude of $2.5 \cdot 10^{-8}$ deg and a period of one revolution. The resulting node drift due to the Z-bias is $5 \cdot 10^{-12}$ deg/d only. In summary, the SRP parameters must be handled carefully, since they can introduce large errors in the node thus easily masking the sought for LTE.

As a next step, the error in the *estimated* nodes due to the mismodeling in the SRP is analyzed in two different cases. In the first case 1 day of Galileo orbits and observations are simulated with F0 = 1 and the orbits are recovered from the observations by fixing F0 = 1.2. This represents a deliberate error of 20% of the full SRP model coming from a mismodeling in the satellite area, mass, surface properties, etc. In GPS orbit determination F0 varies periodically ranging from 0.5 to 1.5 or even more and hence, an error of 20% seems realistic. No X-, Y- and Z-biases are considered in this first case and the optimal parameterization for 1-day arcs is used in the recovery. Like in the previous tests the LTE is introduced in the data simulation but it is not modelled in the recovery of the orbits. The differences in the node between estimated and simulated orbits are presented in Fig. 6 (a). A periodical error with a magnitude of up to $1 \cdot 10^{-7}$ deg is clearly observed, the node drift due to the LTE is not visible. As a consequence, the global scaling factor of the SRP model shall not be fixed or highly constrained, but rather estimated. The estimation of F0 absorbs the deficiencies in the SRP model and does not affect the measurement of the LTE, as shown in Section 3.

In a second case, the Galileo orbits and observations are simulated by including Y- and Z-biases in the SRP model and an increasing F0 from the beginning to the end of the arc, according to the values given in Table 3. These values are realistic since they are obtained from the determination of GPS orbits. In the subsequent estimation of the Galileo orbits, F0 is fixed to 1 and the Y- and Z-biases are fixed to 0. The differences between estimated and simulated nodes are shown in Fig. 6 (b). This time, the errors observed in the estimated nodes reach more than $1 \cdot 10^{-7}$ deg mainly due to the mismodeling of the Y- and Z-bias, the LT node drift is not visible. Thus, the Y- and Z-bias of the SRP shall not be fixed but rather estimated since fixing them to incorrect values introduces errors in the node position two orders of magnitude larger than the LTE.

 DOI: 10.2420/AF07.2013.87

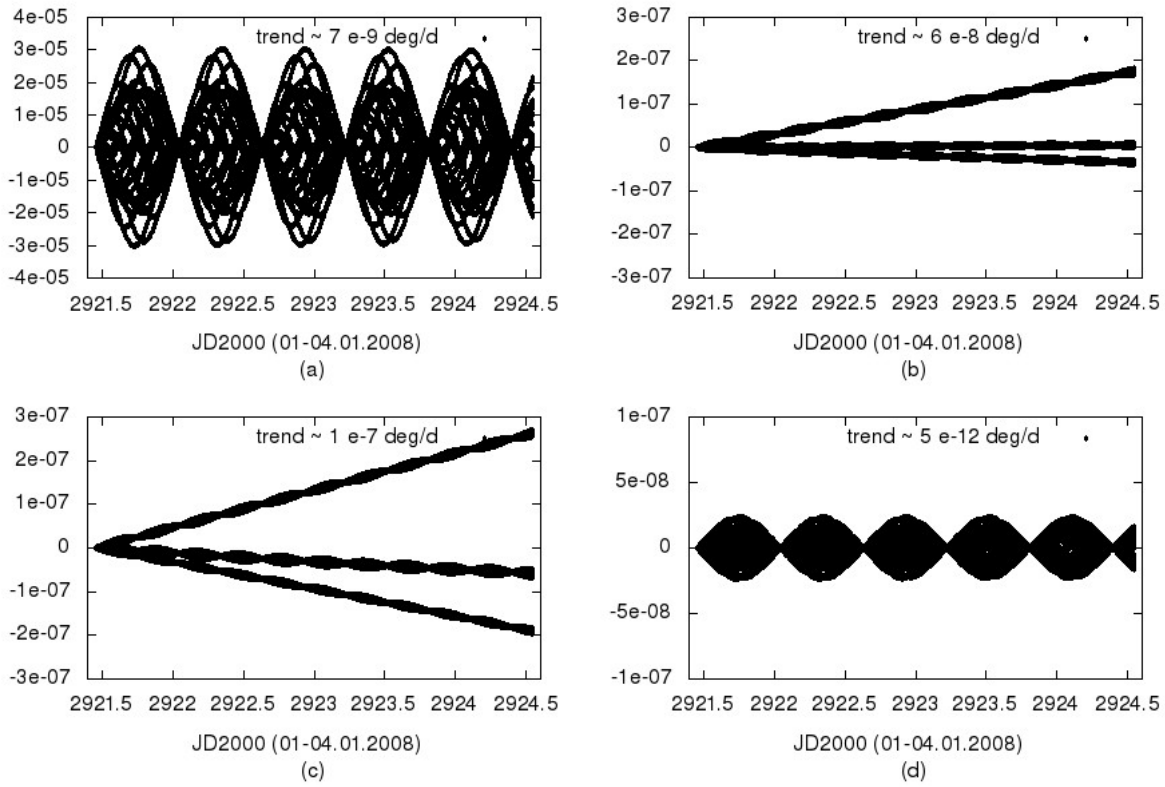


Figure 5. *Effect of the SRP on the Galileo nodes of simulated orbits (deg) when the following parameters are considered: (a) F0 only, (b) X-bias only, (c) Y-bias only and (d) Z-bias only.*

parameter	value
F0	0.999 - 1.001 (total variation = 0.002)
Y-bias	-0.59600E-10
Z-bias	-0.13400E-07

Table 3. *SRP parameters as obtained from GPS orbit determination and used in the simulations*

Finally, to analyze the effects of the SRP depending on the satellite surface properties, a test is performed by using a macro model for the Galileo satellites. A macro model (sometimes also called box-wing model in the literature) takes into account the size and reflection properties of each surface of the satellite. The Galileo attitude model is neglected here. It must be stressed however, that the attitude model is critical for the computation of the SRP effects on the Galileo orbits, since their panels are continuously reoriented to face the Sun. Thus, the results obtained here are a simplification of the real behaviour of the Galileo nodes.

According to different illustrations of the Galileo satellites (e. g. [5] or [13]), two different sets of parameters are set up, corresponding to two different satellite coatings, a gold coating and a silver (similar to aluminium) coating. The parameters of the macro models used are summarized in Table 4. It can be noted that the differences of the coefficients due to the choice of the coating amounts to 25% already. This will directly transfer into the SRP modeling and corresponds roughly with the uncertainties of the SRP model used in the above. The effect on the nodes due to the choice between the two different coatings can be observed in Fig. 7. The node differences are tremendous reaching $2.8 \cdot 10^{-7}$ deg peak to peak. In the end however, in Galileo orbits estimation the error due to a wrong choice of the surface coating will be reduced by the estimation of certain SRP parameters.

In conclusion, the measurement of the LTE with the Galileo nodes requires an accurate SRP model for what

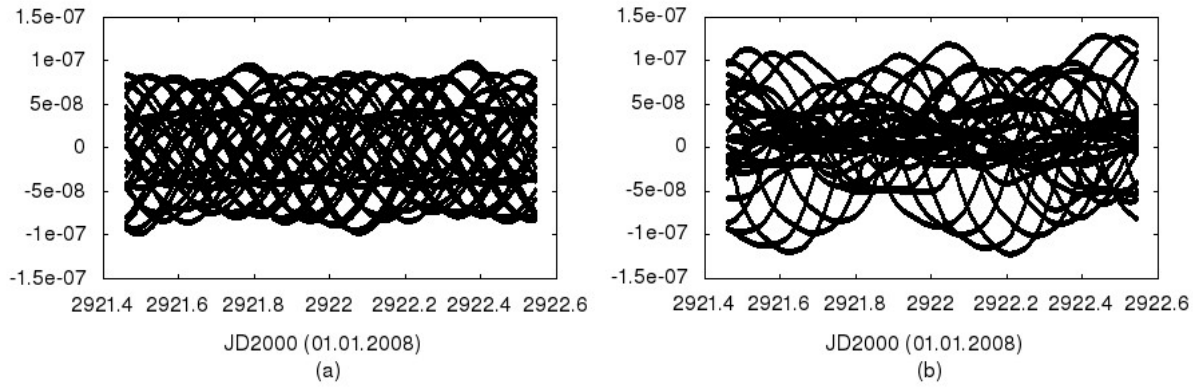


Figure 6. *Orbital node differences (deg) between estimated (without LTE) and simulated (with LTE) Galileo orbits (a) with an error in F0 of 20% and (b) with an error in F0, Y- and Z-bias.*

Surface	Area (m^2)	Refl. coeff. visible			
		gold		silver	
		geom	diff	geom	diff
bus top	1.32	0.14	0.56	0.18	0.72
bus bottom	1.32	0.14	0.56	0.18	0.72
bus left	2.75	0.14	0.56	0.18	0.72
bus right	2.75	0.14	0.56	0.18	0.72
bus front	3.00	0.14	0.56	0.18	0.72
bus back	3.00	0.14	0.56	0.18	0.72
panel left	11.70	0.04	0.16	0.04	0.16
panel right	11.70	0.04	0.16	0.04	0.16

Table 4. *Macro model parameters for the Galileo satellites based on gold and silver coatings.*

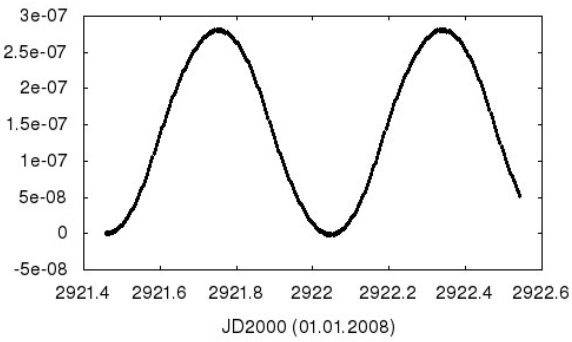


Figure 7. *Orbital node differences (deg) between simulated orbits considering gold or silver/aluminium coatings.*

it is essential to know the Galileo vehicles shape and surface properties. Also some model parameters such as F0, the Y- and Z-biases need to be estimated. The uncertainty arising from the SRP model is unavoidable, however a significant reduction could be achieved only if the Galileo satellites would be equipped with accelerometers, which measure all non-conservative forces acting on the satellites.

5 Discussion and Conclusions

Using simulated Galileo orbits and observations, the effect of the orbit parameterization on the detection of the LTE and the effects of SRP on the Galileo nodes are analyzed.

The best parameterizations for the LTE estimation with 1-day and 3-day Galileo orbits are proposed. It is shown that the best parameterization depends on the arc length. In general, the station coordinates cannot be freely estimated but they must be fixed or highly constrained, f. i. by imposing a set of NNTRS conditions. In addition, the empirical accelerations in the normal direction to the orbital plane and the X-bias of the SRP model must not be estimated, since they absorb the LT signal in the orbital nodes. When processing 1-day arcs, the SRP parameters (F0, Y- and Z-bias) and the Earth's albedo scaling factor can simultaneously be estimated, however with 3-day arcs the Earth albedo must be fixed. These optimal parameterizations allow to estimate the LTE from *noise-free* observations provided that there are no errors in the background models. The possible errors due to the mismodeling of the background models shall be analyzed in future tests.

The 3-day arcs in combination with the optimal parameterization seem suitable for a precise estimation of the LTE from *noisy* Galileo observations, since the noise

 DOI: 10.2420/AF07.2013.87

in the estimated nodes is 5 times smaller than the total displacement of the node due to the LTE after 3 days. In the future this must be confirmed with real Galileo observations. Then, assuming no errors in the background and SRP models, the LTE seems to be estimable from real Galileo data with a precision of 1% by using the nodes of 27 satellites and a minimum of 30 3-day arcs, again assuming no errors in the background models.

The SRP introduces large periodical displacements on the Galileo nodes, four orders of magnitude larger than the LTE. The deviations mainly cancel out after each complete revolution. The resulting node drift is twice the value of the LTE. In addition, the presence of X- and Y-biases produces significant nodal drifts, two orders of magnitude larger than the LTE depending on the orbital plane. The Z-bias introduces small periodical effects and an insignificant node drift. Fixing the SRP parameters (F0, Y- and Z-bias) to incorrect values yields large errors in the estimated nodes. Hence, these parameters shall not be fixed or highly constrained in the estimations, allowing them to absorb the deficiencies in the SRP model. The choice of wrong reflectivity coefficients of the surface of the satellites can give place to deviations of the nodes of up to two orders of magnitude larger than the LTE, which in the end can be reduced by the estimation of certain SRP parameters.

As a consequence, the measurement of the LTE by using the Galileo nodes requires an accurate SRP model, like f. i. a macro model fitted to the shape, size, weight and surface properties of the Galileo satellites. The macro model needs to be accompanied by an appropriate attitude model. The SRP parameters shall be estimated (F0, Y- and Z-bias), allowing them to absorb the deficiencies of the SRP model.

Acknowledgements

The authors wish to thank the European Space Agency for supporting this research under the ESTEC project "Space Tests of General RElativity Using the GAlileo Constellation (REGAL)", contract No. 4000103504/2011/NL/WE.

References

[1] R. Biancale and A. Bode. Mean annual and seasonal atmospheric tide models based on 3-hourly and 6-hourly ECMWF surface pressure data. Scientific Technical Report. STR 06/01, GeoForschungsZentrum, 33. Technical report, GFZ, 2006.

[2] I. Ciufolini. Measurement of the Lense-Thirring drag on high-altitude, laser-ranged artificial satellites. *Phys. Rev. Lett.*, 56:278–281, 1986.

[3] I. Ciufolini, A. Paolozzi, E. C. Pavlis, J. Ries, R. Koenig, R. Matzner, G. Sindoni, and H. Neumayer. Testing gravitational physics with satellite laser ranging. *Eur. Phys. J. Plus*, 126:72, 2011.

[4] S. Desai. Observing the pole tide with satellite altimetry. *J. Geophys. Res.*, 107 (C11):3186, 2002.

[5] ESA. Navigation Website: http://www.esa.int/esaNA/galileo.html, 2012.

[6] H. F. Fliegel, T. E. Gallini, and E. R. Swift. Global Positioning System Radiation Force Model for Geodetic Applications. *J. Geophys. Res.*, 97(B1):559–568, 1992.

[7] C. Foerste, S. Bruinsma, R. Shako, J. C. Marty, F. Flechtner, O. Abrikosov, C. Dahle, J. M. Lemoine, H. Neumayer, R. Biancale, F. Barthelmes, R. Koenig, and G. Balmino. EIGEN-6 A new combined global gravity field model including GOCE data from the collaboration of GFZ-Potsdam and GRGS-Toulouse. In *EGU General Assembly 2011, April 2011, Vienna, Austria.*, 2011.

[8] W. M. Folkner, J. G. Williams, and D. H. Boggs. The Planetary and Lunar ephemeris DE 421 - IPN Progress Report. Technical report, JPL - NASA, 2009.

[9] J. Lense and H. Thirring. *Phys. Z.*, 19, 156, 1918.

[10] J. A. Marshall, P. G. Antreasian, G. W. Rosborough, and B. H. Putney. Modelling radiation forces acting on satellites for precision orbit determination - AAS 91-357, NASA Technical Memorandum 104564. Technical report, NASA, 1992.

[11] D. D. McCarthy. IERS Technical Note 21. IERS Conventions. Technical report, IERS, 1996.

[12] L. Mission. Website: http://www.lares-mission.com/, 2012.

[13] OHB. Website: http://www.ohb-system.de, 2012.

[14] G. Petit and B. Luzum. IERS Conventions. Technical report, IERS, 2010.

[15] R. Savcenko and W. Bosch. EOT11a - a new tide model from Multi-Mission Altimetry. In *OSTST Meeting, October 19-21, San Diego, 2011*, 2011.

[16] S. Zhu, C. Reigber, and R. Koenig. Integrated adjustment of CHAMP, GRACE, and GPS data. *J. Geod*, 78(1-2), 2004.

 DOI: 10.2420/AF07.2013.87

Acta Futura 7 (2013) 97-101
DOI: 10.2420/AF07.2013.97

Acta Futura

GPS observables in general relativity

Carlo Rovelli*
Aix Marseille Université, CNRS, CPT, UMR 7332, 13288 Marseille, France
Université de Toulon, CNRS, CPT, UMR 7332, 83957 La Garde, France

Abstract. I review the motivations and the main ideas that have lead to the introduction of relativistic positioning systems in the form of the so called "GPS" variables. These variables allow us to define a *gauge invariant* metric, and a gauge-invariant notion of length, curvature... in the context of general relativity. I show explicitly how the metric can be measured locally.

1 GPS coordinates

It is common to read that the metric is not an observable in general relativity (GR), because it is not gauge invariant, or, equivalently, in not invariant under coordinate transformations. But this is not always true. When the coordinates are defined in terms of physical degrees of freedom, the quantities that depend on these physical coordinates are gauge-invariant, and observable. The GPS coordinates [11] are particularly simple physical coordinates, that are easy to implement concretely and realistically. In fact, they are *already* implicitly implemented concretely by the global positioning systems, such as the Global Navigation Systems, or by pulsar signal receivers. To define these coordinate theoretically, indeed, it suffices to have four objects in spacetime, and a way to track light signals.

Furthermore, the GPS coordinates of a spacetime point depend only on (the closure of) the past of that point. As a consequence, evolution equations in these coordinates are *local*. A remarkable property that is absent in most other manners to define physical coordinates.

Previous attempts to define a complete class of observables abound in the literature. For instance, in the presence of matter we can localize things with respect to the matter [15, 10, 4]. Other (earlier) attempts to write physical coordinates in GR is to use curvature scalars [3]. The result is however mathematically very intricate and physically extremely unrealistic.

GPS coordinates, introduced in [11], are much simpler: consider a general covariant system including the gravitational field as well as four small bodies. These are taken to have negligible mass; they will be considered as point particles for simplicity, and called "satellites". They follow timelike worldliness, and we assume that a preferred point O_α, $\alpha = 1, 2, 3, 4$ is marked in each of them. Then (there is a region $\mathcal{R}$ of spacetime for which) we can uniquely associate four numbers s^α to each spacetime point p as follows. The past lightcone of p intersect the four geodesics in four points p_α. The numbers s^α are defined as the time lapsed between O_α and p_α along the α world line. We can use the s^α's as physically defined coordinates.

The components $g_{\alpha\beta}(s)$ of the metric tensor in these coordinates are observable quantities. They are invariant under four-dimensional diffeomorphisms (because, of course, these deform the metric as well as the satel-

*E-mail: rovelli@cpt.univ-mrs.fr

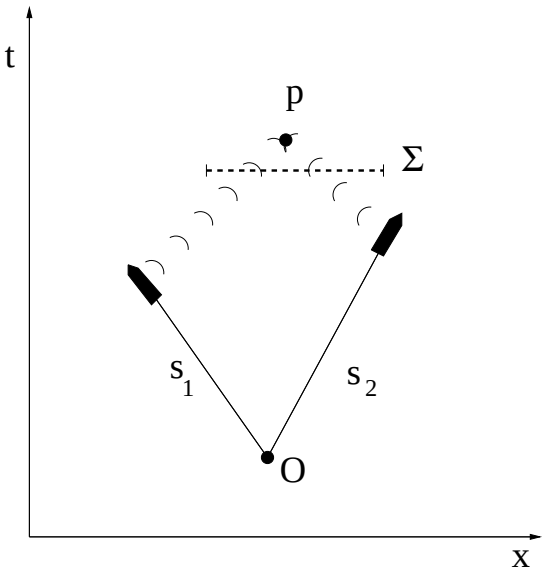


FIGURE 1. s_1 *and* s_2 *are the GPS coordinates of the point* p. Σ *is a Cauchy surface with* p *in its future domain of dependence.*

lites' worldlines). They define a complete set of observables for the region $\mathcal{R}$.

The physical picture is simple, and its realism is transparent. Imagine that the four "satellites" are in fact satellites, each carrying a clock that measures the proper time along its trajectory, starting at O_α. Imagine also that each satellite broadcasts its local time with a radio signal. At the point p, a receiver can receives the four signals and displays the four readings. See Figure 1. These four numbers are precisely the four physical coordinates s^α defined above. Current technology permits to perform these measurements with accuracy well within the relativistic regime [2, 5, 6, 7, 8, 12, 13, 14]. A rod and a clock measuring physical distances between s^α coordinates points, measure the components of the metric tensor in the physical coordinate system. (In the terminology of Ref. [10], the s^α's are *partial* observables, while $g_{\alpha\beta}(s)$ are *complete* observables.)

The physical coordinates s^α have nice geometrical properties. First, they are characterized by

$$g^{\alpha\alpha}(s) = 0, \qquad \alpha = 1, \dots, 4. \tag{1}$$

This is because the constant s^α surface, for a given α, is null, therefore the surface element has null volume, and $g^{\alpha\alpha}$ is the volume element of this surface. Second, the evolution equations for $g_{\alpha\beta}(s)$ are local, as shown in [11]. These evolution equations can be written explicitly using the Arnowitt-Deser-Misner (ADM) variables [1] and Lapse and Shift turn out to be fixed local functions of the three metric.

The coordinates s^α are null coordinates, but we can easily use them to define more common space and time coordinates in a region. For instance, we can define s^μ around some point p by

$$s^\mu = W^\mu_\alpha \, \delta s^\alpha \tag{2}$$

where δs^α is he difference with the respect of the coordinates of p and W^μ_α is a fixed matrix that transforms null direction is space and time directions.

2 Measuring the metric

How to measure the spacetime metric with local measurements? Here I recall that this is possible, and how it can be done.

The setting is the following. Assume that the spacetime region R of interest is described by a metric field $g_{\mu\nu}(x)$ which is not known. The only assumption about this metric is that in the region of interest the maximum curvature scale is given by a value A, which we assume to be know. In other words, the metric is not too "rough", below a given scale. Assume that somewhere in the universe there are four emitters that broadcast a signal, which can be received in the region R. Concretely, this can be a Navigation System-like artificial signal broadcasting the emission proper time, or the periodic signal of a pulsar. Assume that the measuring device for measuring the metric is a swarm of freely falling flying objects equipped with the following devices:

- a clock measuring proper time along the world line of each device,
- a receiving antenna that detects the signals of the four emitters and reads the emitter's proper time. In the case of the pulsar, I assume that onboard clock can resolve time by a greater precision than the millisecond pulsar period, by simply linearly interpolating between pulses.
- a laser ranging technology that allows each flying object to track the distance of its immediate neighbors in the swarm, by sending a light signal to the neighbor, and measuring the proper time between emission and reception of the reflected signal.
- a way to detect the moment at which the radar signal emitted by a neighboring device is reflected.

 DOI: 10.2420/AF07.2013.97

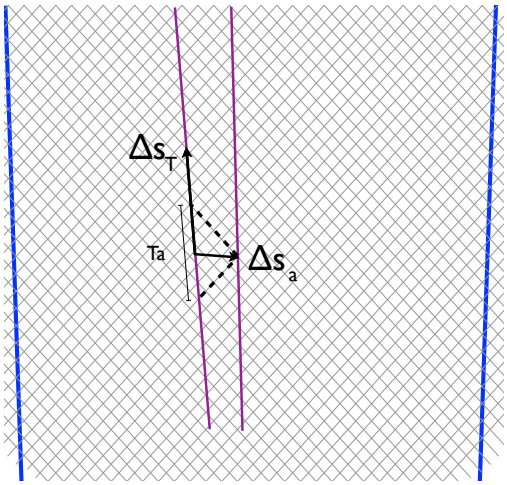


Figure 2. *Local measurements. The blue lines are the external emitters, the grid gives the null coordinates, the purple lines are the free falling objects.*

With these data, the metric can be reconstructed with arbitrary precision, if the average distance between the flying devices and the clock's accuracy are sufficiently small with respect to A.

This is how the reconstruction works. On a given object O, at some given moment, let $s^\alpha = (s^1, s^2, s^2, s^3, s^4)$ be the reading of the four signals emitted by the four external emitters (interpolated is necessary for the pulsars), that is, GPS coordinates. These readings will be used as a concrete coordinatization of the points in R. At some given moment p, let this reading be s^α_p. Wait for a proper time T (much shorter than A) and take the new reading of the four signals. Let this be $s^\alpha_T = s^\alpha_p + \Delta s^\alpha_T$. We can use the basic definition of the metric field, which is

$$ds^2 = g_{\mu\nu}(x)dx^\mu dx^\mu \tag{3}$$

and since we have assumed that T is much shorter than A, we can treat the finite Δs^α_T intervals as infinitesimals, and get a condition for the components $g_{\alpha\beta}(s)$ of the metric in the coordinate system defined by the coordinates s^α:

$$T = g_{\alpha\beta}(s_p)\Delta s^\alpha_T \Delta s^\beta_T. \tag{4}$$

Next, at the same moment s^α_p, a signal was sent from O to the three of the neighboring devices on the swarm (labelled $a = 1, 2, 3$), was received back on O after a time $2T_a$, and the coordinates of the neighboring object at the moment of the reflection is recorded (on the object a, and later communicated) and called $s^\alpha_a = s^\alpha_p + \Delta s^\alpha_a$. Then, as before, this implies

$$cT_a = -g_{\alpha\beta}(s_p)\Delta s^\alpha_a \Delta s^\beta_a. \tag{5}$$

Next, with the same technology, let cT_{ab} be the distance between the object a and the object b at the reception time. Since these distances are much smaller than A, we can consider spacetime locally flat and from the three lengths of the three sides of the triangle formed by the three objects O, a and b, we can compute the angle ϕ_{ab} at of this triangle at O (which is the angle under which a and b at seen by O). This is given by Hero formula

$$T_a T_b \cos\phi_{ab} = T_a^2 + T_b^2 - T_{ab}^2, \tag{6}$$

So we have immediately one more condition on the metric

$$c(T_a^2 + T_b^2 - T_{ab}^2) = -g_{\alpha\beta}(s_p)\Delta s^\alpha_a \Delta s^\beta_b. \tag{7}$$

Finally, let's observe that the surfaces where one coordinate is constant are null surfaces, therefore the norm of their covariant normal vanishes, as this is given by the diagonal matrix elements of the controvariant metric, we have (no summation)

$$0 = g^{\alpha\alpha}(s_p). \tag{8}$$

Now in these equations, all the quantities $T, T_a, T_{ab}, \Delta s^\alpha_T, \Delta s^\alpha_a, s^\alpha_p$ are concretely measurable. The s^α and Δs^α quantities are readings form the four distant emitters (or given by counting pulsate pulses plus interpolation), while the times T, T_a, T_{ab} are directly read by onboard clocks. Therefore the equations can be solved for $g_{\alpha\beta}(s_p)$ at the given point s^α_p. This can be done on a grid of points. If the grid is fine enough, the metric can be reconstructed with arbitrary precision. Notice that the individual components $g_{\alpha\beta}(s)$ of the metric in the coordinates s^α are gauge-invariant quantities in general relativity: they are uniquely determined physically by the matter and the gravitational field. Therefore they observable quantities, in the sense of gauge theories. Here I have shown how they can be concretely measured.

Where is the curvature? Within each cell defined by the grid, there cannot be curvature, by assumption, of course: if in reality there is curvature, it is averaged away by the measuring procedure. But at the immediately higher scale, curvature is immediately detected. To see this, imagine for a moment that we focus for simplicity on spacial curvature, which is easier to have intuition

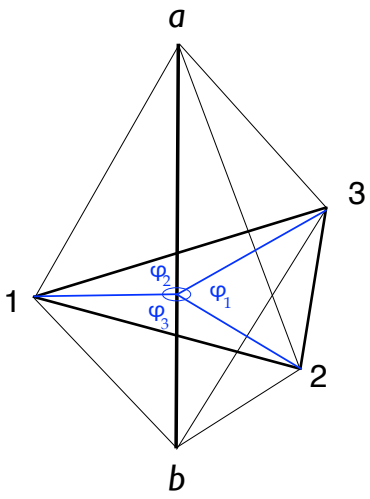


FIGURE 3. *Curvature is given by the difference between* 2π *and* $\varphi_1 + \varphi_2 + \varphi_3$.

about. Consider two objects, say a and b. And a ring of three other objects labelled as $1, 2, 3$ forming a triangle that surrounds the line between a and b. The four devices $a, b, 1, 2$ form a tetrahedron. In a tetrahedron, the dihedral angle φ_{ab} at the side (a, b) can be computed explicitly from the lengths of the sides. It is given by the formula

$$\cos\varphi_{ab} = \frac{\cos\phi_{a12} - \cos\phi_{ab1}\cos\phi_{ab2}}{\sin\phi_{ab1}\sin\phi_{ab2}}, \quad (9)$$

where ϕ_{abc} is the angle at a in the abc triangle. Let φ_1 be the dihedral angle at the side ab of the tetrahedron $a, b, 2, 3$ and so on cyclically. Then clearly if space is flat the three dihedral angles must sum up to 2π. If they don't, this means that there is curvature at the side ab. Therefore, following Regge, we can define the deficit angle

$$\delta_{ab} = 2\pi - \sum_{i,1}^{3} \varphi_i. \quad (10)$$

The deficit angle is the discrete version of the curvature. (It is given by the components of the Riemann tensor that rotate in the plane normal to the side, and integrated in a plane normal to the side: $\delta_{ab} \sim R^{\alpha}_{\beta\gamma\theta}\Delta s_{1\alpha}\Delta s_2^{\beta}\Delta s_1^{\gamma}\Delta s_2^{\theta}$) The dihedral angle is an explicit function of the measured quantities. Hence these measures detect curvature systematically. The same construction (which I do not give here for simplicity) works in four dimensions.

In conclusion, given GPS coordinates and a swarm of objects equipped with clocks and capable of tracking the relative distances from their immediate neighbors via laser tracking, the spacetime geometry of a region can be fully reconstructed down to the scale of the average objects' distance.

It is may be useful to recall here the distinction between partial and complete observables [10]. A partial observable is a quantity to which a measuring procedure can be associated. A complete observable is an observable quantity that can predicted by the theory, or, equivalently, whose knowledge provides information on the state of the system. The GPS coordinates are partial observables: we can associate them a measuring procedure (this is what has been done in this paper), but we can of course not "predict" s[1]. The complete observables, or true observables, are the quantities $g_{\mu\nu}(s)$, for any given value of the coordinates s^{μ}. These quantities are diffeomorphism invariant, are uniquely determined by the initial data and in a canonical formulation are represented by functions on the phase space that commute with all constraints.

These observables are straightforward generalization of Einstein's "point coincidences"[1]. In a sense, they are *precisely* Einstein's point coincidences. Einstein's "material points" are just replaced by photons (light pulses): the spacetime point s^{α} is characterized as the meeting point of four photons designated by the fact of carrying the radio signals s^{α}.

References

[1] R. L. Arnowitt, S. Deser, and C. W. Misner. *Gravitation: an introduction to current research*, chapter The dynamics of general relativity, pages 227–265. John Wiley & Sons, 1962.

[2] T. B. Bahder. Fermi coordinates of an observer moving in a circle in Minkowski space: Apparent behavior of clocks. Technical Report ARL-TR-2211, Army research laboratory, 1998.

[3] P. G. Bergmann. Observables in general relativity. *Rev. Modern Phys.*, 33:510–514, 1961.

[1] "All our space-time verifications invariably amount to a determination of space-time coincidences. If, for example, events consisted merely in the motion of material points, then ultimately nothing would be observable but the meeting of two or more of these points" [9].

 DOI: 10.2420/AF07.2013.97

[4] J. D. Brown and D. Marolf. Relativistic material reference systems. *Physical Review D*, 53(4):1835, 1996.

[5] F. De Felice, M. Lattanzi, A. Vecchiato, and P. Bernacca. General relativistic satellite astrometry. i. a non-perturbative approach to data reduction. *Astronomy and Astrophysics*, 332:1133–1141, 1998.

[6] B. Guinot. Application of general relativity to metrology. *Metrologia*, 34(3):261, 1997.

[7] B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins. *Global Positioning System. Theory and practice.* Wien (Austria): Springer, 1993.

[8] E. D. Kaplan and C. J. Hegarty. *Understanding GPS: principles and applications*. Artech House Publishers, 2006.

[9] W. Perrett and G. Jeffery. *The principle of relativity*. Dover, New York, 1952.

[10] C. Rovelli. What is observable in classical and quantum gravity? *Classical and Quantum Gravity*, 8(2):297, 1991.

[11] C. Rovelli. Gps observables in general relativity. *Phys. Rev. D*, 65:044017, Jan 2002.

[12] J. Spilker. *GPS Theory and Applications, Vol. I, Progress in Astronautics and Aeronautics*, chapter Tropospheric Effects on GPS, pages 517–546. Amererica Institute of Aeronautics, 1996.

[13] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr. *Interferometry and synthesis in radio astronomy*. Wiley-Vch, 2008.

[14] P. Visser. Gravity field determination with GOCE and GRACE. *Advances in Space Research*, 23(4):771–776, 1999.

[15] L. Witten. *Gravitation: an introduction to current research*, volume 1. New York: Wiley, 1962.

Acta Futura 7 (2013) 103-110
DOI: 10.2420/AF07.2013.103

Acta Futura

Relativistic Positioning Systems: Numerical Simulations

Diego Sáez*, and Neus Puchades
Departamento de Astronomía y Astrofísica, Universidad de Valencia, 46100 Burjassot, Valencia, Spain

Abstract. The motion of satellite constellations similar to GPS and Galileo is numerically simulated and, then, the region where bifurcation (double positioning) occurs is appropriately represented. In the cases of double positioning, the true location may be found using additional information (angles or times). The zone where the Jacobian, J, of the transformation from inertial to emission coordinates vanishes is also represented and interpreted. It is shown that the uncertainties in the satellite world lines produce positioning errors, which depend on the value of |J|. The smaller this quantity the greater the expected positioning errors. Among all the available 4-tuples of satellites, the most appropriate one –for a given location– should minimize positioning errors (large enough |J| values) avoiding bifurcation. Our study is particularly important to locate objects which are far away from Earth, e.g., satellites.

1 General considerations

Global Navigation Satellite Systems (GNSSs) are satellite constellations broadcasting signals, which may be used to find the position of a receiver (user) [7]. GNSSs are mainly used to locate receivers on Earth surface. In this case, users receiving signals from four or more satellites may be always located with admissible positioning errors. However, if the user to be located is far away from Earth, two main problems may arise, the first one is the existence of two possible positions for the object (bifurcation), some previous considerations about this problem may be found in [1, 3, 10]. The second problem is related to the positioning errors due to uncertainties in the world lines of the satellites. These errors are proved to be too big for some user positions close to points where the Jacobian J –of the transformation from inertial to emission coordinates– is too small. In this paper, we find and represent the regions where the above problems arise. Only numerical calculations may be used to find these regions in the case of realistic GNSSs. Moreover, their representation is an additional problem, which have been solved by choosing appropriate sections –of the 4D emission region– and by using suitable methods previously developed in other research fields (see below). The mentioned regions have been only studied in the interior of a big sphere having a radius of 10^5 km, which is centred in a point E of the Earth surface (see Fig. 1).

Quantities G, $M_\oplus$, t, and τ stand for the gravitation constant, the Earth mass, the coordinate time, and the proper time, respectively. Greek (Latin) indexes run from 0 to 3 (1 to 3). Quantities $\eta_{\alpha\beta}$ are the covariant components of the Minkowski metric tensor. Our signature is (+,+,+,–). The unit of distance is assumed to be the Kilometre, and the time unit is chosen in such a way that the speed of light is $c = 1$. The index A numerates the four satellites necessary for relativistic positioning.

GPS and GALILEO satellite constellations are simulated [14, 15]. Satellite trajectories are assumed to be

*Corresponding author. E-mail: Diego.Saez@uv.es

circumferences in the Schwarzschild space-time created by an ideal spherically symmetric Earth. A first order approximation in $GM_\oplus/R$ is sufficient for our purposes. The angular velocity is $\Omega = (GM_\oplus/R^3)^{1/2}$, and coordinate and proper times are related as follows: $\gamma = \frac{dt}{d\tau} = (1 - \frac{3GM_\oplus}{R})^{-1/2}$. Angles θ and ϕ fix the orbital plane (see [15]), and the angle $\alpha_A(\tau) = \alpha_{A0} - \Omega\gamma\tau$ localizes the satellite on its trajectory. This simple model is good enough as a background configuration. Deviations with respect to the background satellite world lines will be necessary to develop our study about positioning accuracy (see below).

Other known world lines (no circumferences) of Schwarzschild space-time might be easily implemented in the code, but the new background satellite configurations would lead to qualitatively comparable numerical results; at least, for the problems considered in this paper.

The angle $\alpha_A(\tau)$ may be calculated for every τ, and the two angles fixing the orbital plane (θ and ϕ) are constant. From these three angles and the proper times τ^A, the inertial coordinates of the four satellites, x^α_A, may be easily found to first order in $GM_\oplus/R$ [15]. This means that the world lines of the background satellites [functions $y^\alpha = x^\alpha_A(\tau^A)$] are known for every satellite A. Hence, given the emission coordinates $(\tau^1, \tau^2, \tau^3, \tau^4)$ of a receiver, the inertial coordinates $x^\alpha_A \equiv (x_A, y_A, z_A, t_A)$ of the four satellites –at emission times– may be easily calculated. The knowledge of the satellite world lines is necessary for positioning; namely, to find the inertial coordinates from the emission ones [4].

The satellite world lines are also necessary to go from the inertial coordinates of an user to its emission ones. This transformation is now considered under the assumption that photons move in the Minkowski space-time, whose metric has the covariant components $\eta_{\alpha\beta}$. This approach is good enough for us. Since photons follow null geodesics from emission to reception, the following algebraic equations must be satisfied:

$$\eta_{\alpha\beta}[x^\alpha - x^\alpha_A(\tau^A)][x^\beta - x^\beta_A(\tau^A)] = 0\,. \qquad (1)$$

These four equations must be numerically solved to get the four emission coordinates τ^A. The four proper times are the unknowns in the system (1), which may be easily solved by using the well known Newton-Raphson method [13]. Since the satellite world lines are known, functions $x^\alpha_A(\tau^A)$ may be calculated for any set of proper times τ^A, thus, the left hand side of Eqs. (1) can be computed and, consequently, the Newton-Raphson method may be applied. A code has been designed to implement this method. It requires multiple precision. Appropriate tests have been performed [15].

Moreover, given four emission coordinates τ^A, Eqs. (1) could be numerically solved to get the unknowns x^α, that is to say, the inertial coordinates (positioning); however, this numerical method is not used. It is better the use of a certain analytical formula giving x^α in terms of τ^A, which was derived in [4]. The analytical formula is preferable because of the following reasons: (i) the numerical method based on Eqs. (1) is more time consuming and, (ii) the analytical formulation of the problem allows us a systematic and clear discussion of bifurcation, and also a study of the positioning errors close to points of vanishing Jacobian.

The analytical formula [4] has been described in various papers [4, 7, 5, 15], and numerically applied in [14, 15]. This formula involves function χ^2, which is the modulus of the configuration vector, and a discriminant Δ. The definitions of both χ^2 and Δ may be found in [4, 7, 6]. It is very important that these two quantities may be calculated by using only the emission coordinates τ^A. From the analytical formula giving the inertial coordinates in terms of the emission ones, and taking into account some basic relations of Minkowski space-time, the following propositions have been previously proved [4, 5, 7]:

(a) For $\chi^2 \leqslant 0$, there is only a positioning (past-like) solution.

(b) For $\chi^2 > 0$ there are two positioning solutions; namely, there are two sets of inertial coordinates (two physical real receivers) associated to the same emission coordinates τ^A.

(c) The Jacobian J of the transformation giving the emission coordinates in terms of the inertial ones vanishes if and only if the discriminant Δ vanishes.

(d) The Jacobian J may only vanish if $\chi^2 > 0$; namely, in the region of double positioning (bifurcation).

(e) The Jacobian J may only vanish if the lines of sight –at emission times– of the four satellites belong to the same cone (with vertex in the user).

These conclusions are basic for the numerical estimations and discussions presented below. In particular, after calculating χ^2 and Δ from the emission coordinates, propositions (a) and (b) allow us to find the regions where bifurcation takes place, whereas the zones with vanishing Jacobian (infinite positioning errors) may be found by using proposition (c).

 DOI: 10.2420/AF07.2013.103

2 Emission region

For every set of four satellites, the so-called 4D emission region [4, 7, 15] is studied by considering 3D sections $t = constant$. Each of these sections is covered by points according to the method described in Fig. 1. Let us here give some additional details. Healpix package –first used in cosmic microwave background researches [11]– is used to define 3072 directions. A segment with one of its ends at E and having length $L_{max} = 10^5$ km is associated to each direction. A great enough number of points are uniformly distributed along every direction. The inertial coordinates of the chosen points are known by construction; then, the Newton-Raphson method –implemented in our numerical codes– gives the associated emission coordinates, which allow us the computation of χ^2 and Δ (see section 1). With these quantities, we may find the zeros of both χ^2 and J [proposition (c) of section 1] along any given direction. As it is displayed in Fig. 1, various zeros of χ^2 and J may be found in each direction. They may be distributed in different ways by obeying proposition (d) of section 1.

Once the zeros of χ^2 and Δ have been found for all the Healpix directions (3072), an appropriate method is necessary to display the results. Our method is based on the Healpix pixelisation and the mollwide projection (see [15] for more details). Healpix associates a pixel to each direction. The pixel colour measures the value of some chosen quantity according to the colour bar displayed in our figures, which are mollwide projections of the pixelised sphere.

The structure of the 3D section considered in Fig. 2 ($t = 25$ h after the time origin) is displayed in seven panels. The blue pixels of panel (a) show the directions of the four chosen satellites when they emitted the signals received at point E. Since satellite velocities are much smaller than the speed of light, the satellite positions at emission times are very similar for every point of the 3D section under consideration. In order to understand panels (b)–(h) the reader need to know the quantity associated to every colour bar and the meaning of the grey pixels. In panel (b) [(e)], the colour bar shows the distance from E to the first zero of χ^2 [J], and along the directions corresponding to the grey pixels, function χ^2 [J] does not vanish, at least, up to a distance L_{max} from E. In panel (c) [(f)], the colour bar gives the distance from the first to the second zero of χ^2 [J], and function χ^2 [J] does not vanish two times in the directions of the grey pixels (up to L_{max}) and, finally, in panel (d) [(g)] the colour bar displays distances from the second to the third zero of χ^2 [J], and no a third zero of χ^2 [J] has been found along the directions of the grey pixels (up to L_{max}). In some cases, there are no directions with more than one zero of functions χ^2 (see [15]) and J. In any case, panels (c)–(g) of Fig. 2 show that there are only few directions with two zeros of these functions, and also that directions with three zeros are very scarce. Other 3D sections and other 4-tuples of satellites have been studied with similar results.

From Figs. (1) and (2), it follows that the emission region has the following structure: there are directions without zeros of χ^2 and J. These directions subtend a great solid angle [grey pixels of panels (b) and (e)]. The complementary solid angle corresponds to directions with one or more zeros of χ^2. From point E to the first zero of χ^2 there is no bifurcation, but it appears beyond the first zero. For the small number of directions having a second zero, there is no bifurcation beyond this zero, but it occurs again beyond the third zero (see Fig. 1), which only exists for very scarce directions. We have verified that, according to proposition (d) of section 1, the Jacobian only vanishes in regions with bifurcation.

3 Positioning errors and satellite uncertainties

The background world lines of the satellites are the circumferences of section 1, whose equations have the form $y^\alpha = x^\alpha_A(\tau^A)$. Let us first suppose that the background world lines are exactly followed by the satellites (without uncertainties). Under this assumption, given the inertial coordinates x^α of an user, the background world line equations, Eqs. (1), the Newton-Raphson method, and multiple precision may be used to find the emission coordinates $\tau^1, \tau^2, \tau^3, \tau^4$ with very high accuracy. Finally, the chosen inertial coordinates x^α may be recovered from the emission ones –with very high accuracy– by using the analytical solution derived in [4]. This process is useful to prove that our numerical codes work with high accuracy.

Let us now suppose that there are uncertainties in the satellite world lines, whose equations are $y^\alpha = x^\alpha_A(\tau^A) + \xi^\alpha_A$, where ξ^α_A are deviations with respect to the background world lines due to known or unknown external actions on the satellites. Let us now take the above inertial coordinates x^α, the equations $\eta_{\alpha\beta}[x^\alpha - x^\alpha_A(\tau^A) - \xi^\alpha_A][x^\beta - x^\beta_A(\tau^A) - \xi^\beta_A] = 0$, the

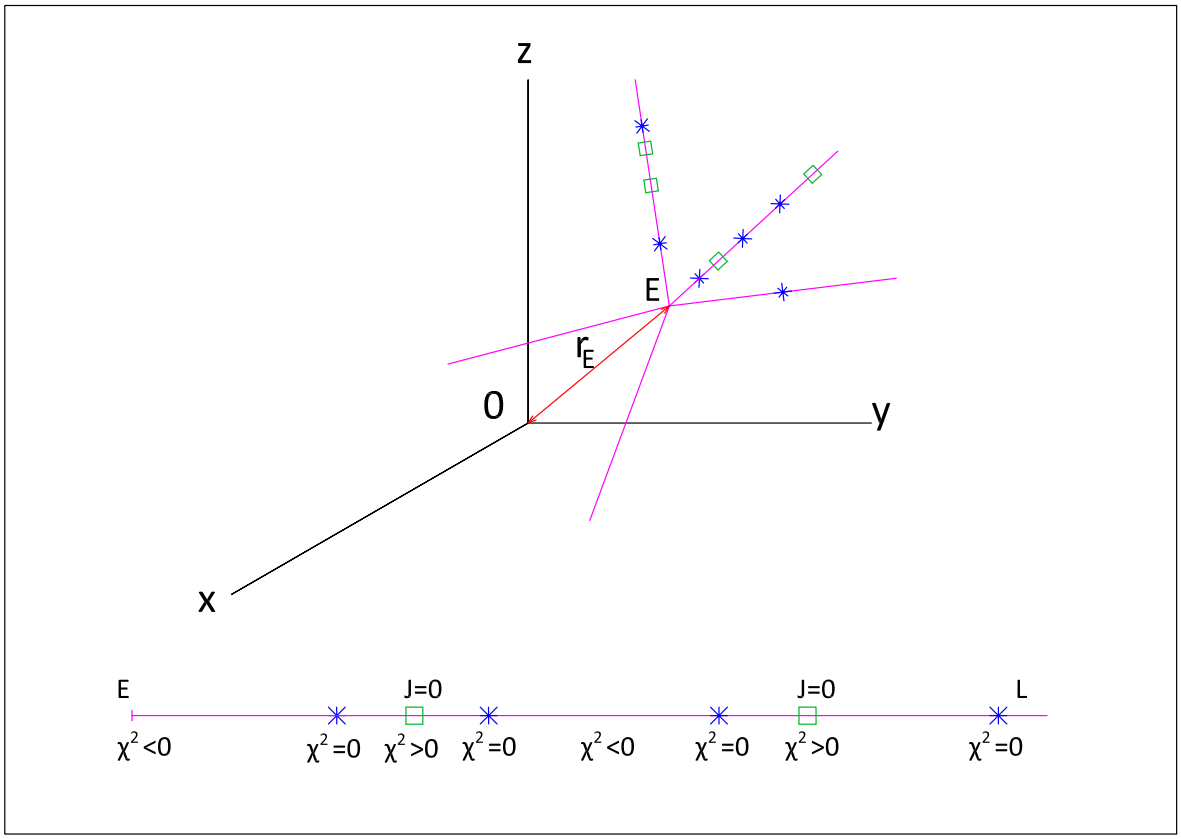


Figure 1. *3D sections (*$t = \mathrm{constant}$*) of the emission region are considered. Point E is an arbitrary centre. Its distance to the origin* O *is the Earth radius.* L *is the distance from E to another point of the section. 3072 segments starting from E and having distinct directions cover each 3D section. Along each direction our study is restricted to* $0 < L < L_{max} = 10^5$ km*. We look for the zeros of* χ^2 *and* J*. The Jacobian only may vanish in the segments where* $\chi^2 > 0$*, which are limited by the first and second or by the third and fourth* χ^2*-zeros.*

Newton-Raphson method, and multiple precision, to get the perturbed emission coordinates $[\tau^1+\Delta(\tau^1), \tau^2+\Delta(\tau^2), \tau^3+\Delta(\tau^3), \tau^4+\Delta(\tau^4)]$. Since the time deviations $\Delta(\tau^A)$ are all small, quantities ξ^α_A may be assumed to be constant in the short interval $[\tau^A, \tau^A + \Delta(\tau^A)]$. Finally, by using the analytical solution mentioned above, new inertial coordinates $x^\alpha + \Delta(x^\alpha)$ may be obtained from the emission coordinates τ^A and the deviations ξ^α_A. Coordinates $x^\alpha + \Delta(x^\alpha)$ are to be compared with the inertial coordinates x^α initially assumed.

Quantity $\Delta_d = [\Delta^2(x^1) + \Delta^2(x^2) + \Delta^2(x^3)]^{1/2}$ is a good estimator of the positioning errors produced by the assumed uncertainties, ξ^α_A, in the satellite motions.

For a certain direction, we have taken an interval of 200 km centred at a zero of function J and, then, quantity Δ_d has been calculated in 200 uniformly distributed points of the chosen interval. In each of these points, the same deviations ξ^α_A have been used to perturb the satellite world lines. The three quantities ξ^i_A have been written in terms of the modulus $\Xi_A = [(\xi^1_A)^2+(\xi^2_A)^2+(\xi^3_A)^2]^{1/2}$ and two angles Θ and Φ (spherical coordinates) and, then, quantities Ξ_A, Θ, Φ, and ξ^4_A have been generated as random uniformly distributed numbers in the intervals $[0, 10^{-3}]$ in km, $[0, \pi]$, $[0, 2\pi]$, and $[0, 10^{-3}]$ in time units (see section 1), respectively. Results are presented in Fig. 3, where we see that our estimator of the positioning errors Δ_d is very large close to the central point where $J = 0$. This fact is important in section 4 (satellite positioning). It is due to the fact that the satellite to be located may cross the region of vanishing J.

The Jacobian is being numerically calculated –in any

 DOI: 10.2420/AF07.2013.103

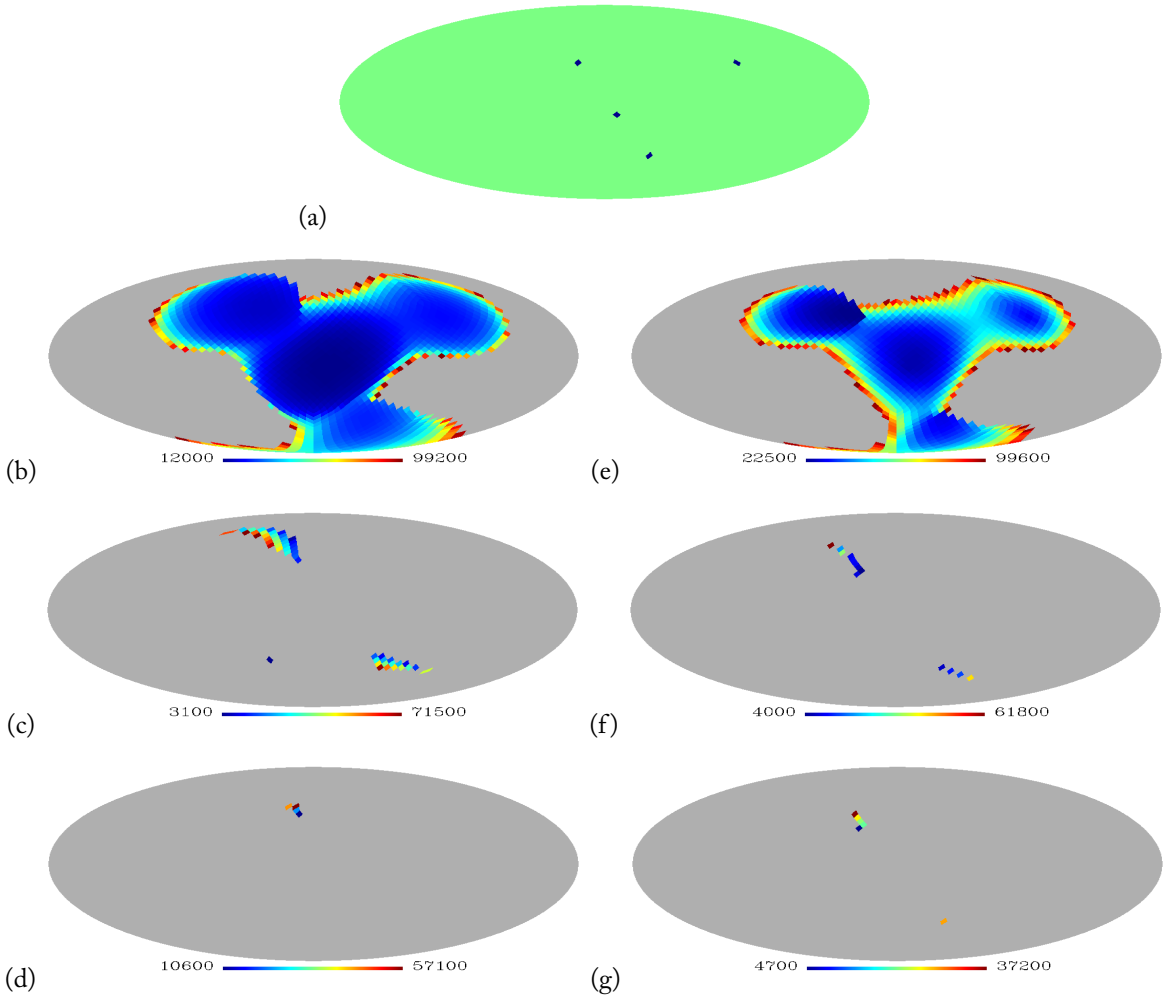


FIGURE 2. *Structure of section* $x^4 = t = 25$ h *for* $L < 10^5$ km

point of the emission region– to perform a relativistic study of the so-called *dilution of precision* [12]; namely, to look for the relation between the geometry of the system satellites-user and the amplitude of the positioning errors. This study must be developed for users on Earth, as well as for users far away from Earth (satellites).

4 Looking for the position of a satellite

In this paper, we are concerned with the location of users which move far away from Earth as, e.g., an user in a satellite. Of course, we use the emission times broadcast by four satellites, which might belong, e.g., to GPS or Galileo GNSSs. Two particular cases are considered. In the first (second) one, the user travels in a Galileo (GPS) satellite and the emitters are four GPS (Galileo) satellites. Thus, the world lines of the user and the emitters are known (see section 1). As it follows from section 2, there is no bifurcation for distances to E smaller than about 10^4 km, which means that GPS and Galileo satellites, which have altitudes of 20200 and 23222 km, may pass from a regions with bifurcation to other region with single positioning or vice versa. The parts of the user circumference where bifurcation occurs are now determined in the two cases. Results of the first (second) case are presented in panel (h) [(i)] of Fig. (4).

7200 equally spaced points are considered on the user world line. In each point, the emission coordinates are

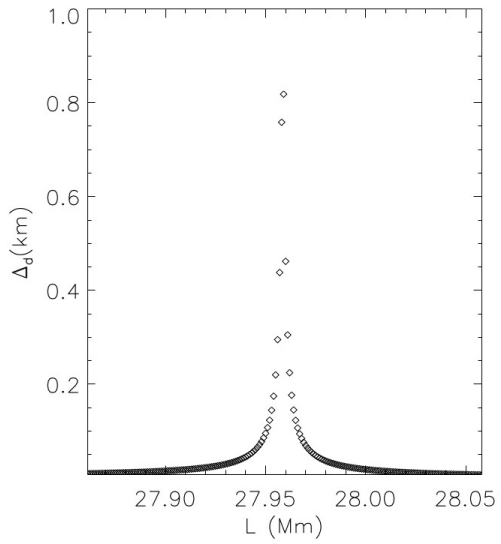


Figure 3. *Estimator* Δ_d *(in* km*) against the distance* L *to* E *(in* $Mm \equiv 10^3$km*) along the chosen direction in a* 200 km *interval centred at a point where* $J = 0$*. Close to this point, quantity* Δ_d *takes on very large values*

calculated (Newton-Rhapson) and, from them, the sign of χ^2 is found, this sign tell us if the point is single valued or it has an associated false position (bifurcation). Single valued points are red and are always located on the user circumference.

Four sets of 1800 points have been selected and –in the case of bifurcation– these sets have been ordered in the sense of growing time (dextrogyre) by using the following sequence of colours: black, fuchsia, dark blue and light blue.

Initial points may be: a single valued red point (represented by a star) or a bifurcation represented by two black stars.

Since GPS and GALILEO satellites have not the same period, the final point has not always the same χ^2 sign as the initial one. In the case of bifurcation one of the points is on the circumference and the other point is an external light blue star.

In the transition from red (single positioning) to any other colour (bifurcation), one of the positions is on the circumference and the other one tends to infinity. The same occurs from any colour (bifurcation) to red (single positioning). Any other colour change is continuous. It is due to the fact that we have decided to change the colour to follow the satellite motion.

In panel (i) of Fig. (4), there are no bifurcation at all (red points). This is an exceptional case. A more frequent situation with zones of bifurcation is given in panel (h), where the asymptotic behaviour at the ends of the bifurcation intervals is displayed.

5 General discussion

In our approach, satellites move in Schwarzschild space-time, so the effect of the Earth gravitational field on the clocks is taken into account, e.g., it has been verified that GPS clocks run more rapid than clocks at rest on Earth by about 38.4 microseconds per day. This prediction agrees with previous ones, which strongly suggests that our methods and codes work.

Since the Earth gravitational field produces a very small effect on photons while they travel from the satellites to the receiver (the covered distance is not large and the gravitational field is weak), photons have been moved in Minkowski space-time.

We are currently moving photons in Schwarzschild, Kerr, and PPN space-times; however, only small corrections arise with respect to the approach assumed here. Previous work on this subject has been performed in various papers [8, 9, 2].

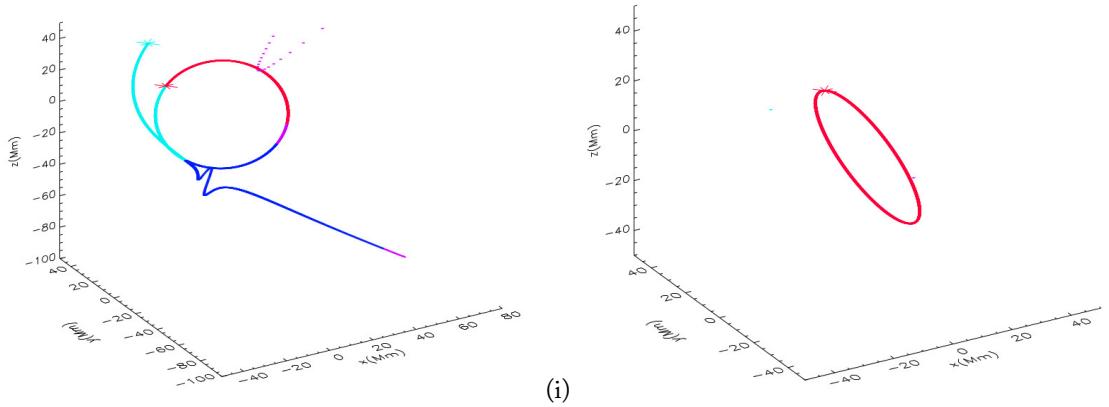
In this paper, the emission coordinates are calculated, from the inertial ones, by using accurate numerical codes based on the Newton-Raphson method. However, the inertial coordinates are obtained, from the emission ones (positioning), by means of the analytical transformation law derived in [4].

From the emission coordinates and the satellite world lines, one easily finds the number of possible receiver positions. If this number is two, there is bifurcation. In this case, it has been proposed a method (based on angle measurements) to select the true position [7]. Other methods (based on time measurements) are possible (see [15]).

We have proved that small uncertainties in the satellite world lines produce large positioning errors if $J \simeq 0$. A more detailed study of this type of errors is in progress.

The emission region has been studied for a certain 4-tuple of satellites. The zones with bifurcation and those having small values of $|J|$ have been found, and appropriate methods have been used to their representation. We have seen that satellites moving at altitudes greater than about 10^4 km may cross these zones, which leads to problems due to bifurcation and large positioning errors. In a GNSS there are various 4-tuples of satellites

 DOI: 10.2420/AF07.2013.103



(h) (i)

Figure 4. *Left: positioning a Galileo satellite with four GPS emitters. Right: same for a GPS satellite and four Galileo emitters.*

which may be used to find the position of a certain user. Among the possible 4-tuples without bifurcation, we should choose the 4-tuple leading to the greatest value of $|J|$ to minimize positioning errors.

Positioning on Earth surface is always single ($\chi^2 \leqslant 0$) and the Jacobian does not vanish in this case. Hence, our study is particularly relevant in the case of users moving far away from Earth.

Acknowledgments This work has been supported by the Spanish Ministries of *Ciencia e Innovación* and *Economía y Competitividad*, MICINN-FEDER projects FIS2009-07705 and FIS2012-33582.

References

[1] J. S. Abel and J. W. Chaffee. Existence and uniqueness of gps solutions. *IEEE Transactions on Aerospace and Electronic Systems*, 27(6):952–956, 1991.

[2] D. Bunandar, S. Caveny, and R. Matzner. Measuring emission coordinates in a pulsar-based relativistic positioning system. *Phys. Rev. D*, 84:104005–9p, 2011.

[3] J. W. Chaffee and J. S. Abel. On the exact solutions of the pseudorange equations. *IEEE Transactions on Aerospace and Electronic Systems*, 30(4):1021–1030, 1994.

[4] B. Coll, J. Ferrando, and J. Morales-Lladosa. Positioning systems in Minkowski space-time: from emission to inertial coordinates. *Class. Quantum Grav.*, 27:065013–17p, 2010.

[5] B. Coll, J. Ferrando, and J. Morales-Lladosa. From emission to inertial coordinates: an analytical approach. *J. Phys. Conf. Ser.*, 314:012105–4p, 2011.

[6] B. Coll, J. Ferrando, and J. Morales-Lladosa. From inertial to emission coordinates: splitting of the solution relatively to an inertial observer. *submitted to Acta Futura*, 2012.

[7] B. Coll, J. Ferrando, and J. Morales-Lladosa. Positioning systems in Minkowski space-time: Bifurcation problem and observational data. *Phys. Rev. D*, 86:084036–10p, 2012.

[8] A. Čadež, U. Kostić, and P. Delva. Mapping the spacetime metric with a global navigation satellite system. *Advances in Space Research*, Final Ariadna Report 09/1301, Advanced Conceps Team. European Space Agency:1–61, 2010.

[9] P. Delva, U. Kostić, and A. Čadež. Numerical modeling of a global navigation satellite system in a general relativistic framework. *Advances in Space Research*, 47:370–379, 2011.

[10] E. Grafarend and J. Shan. A closed-form solution of the nonlinear pseudo-ranging equations (gps). *ARTIFICIAL SATELLITES, Planetary Geodesy*, 31(3):133–147, 1996.

[11] K. G. E. Hivon and B. Wandelt. Analysis issues for large cmb data sets. *In Proc. of the MPA/ESO*

Conference on Evolution of Large Scale Structure. Ipscam, Enschede, pages 37–42, 1999.

[12] R. Langley. Dilution of precision. *GPS World*, 10(5):52–59, 1999.

[13] W. Press, S. Teukolski, W. Vetterling, and B. Flannery. Numerical recipes in fortran 77: the art of scientific computing. *Cambridge University Press, New York*, pages 355–362, 1999.

[14] N. Puchades and D. Sáez. From emission to inertial coordinates: a numerical approach. *J. Phys. Conf. Ser.*, 314:012107–4p, 2011.

[15] N. Puchades and D. Sáez. Relativistic positioning: four-dimensional numerical approach in Minkowski space-time. *Astrophys. Space Sci.*, 341:631–643, 2012.

 DOI: 10.2420/AF07.2013.103

Acta Futura 7 (2013) 111-124
DOI: 10.2420/AF07.2013.111

Acta Futura

Relativistic space-time positioning: principles and strategies

Angelo Tartaglia*
Department of Applied Science and Technology, Politecnico di Torino, and INFN, Italy

Abstract. Starting from the description of space-time as a curved four-dimensional manifold, null Gaussian coordinates systems as appropriate for relativistic positioning will be discussed. Different approaches and strategies will be reviewed, implementing the null coordinates with both continuous and pulsating electromagnetic signals. In particular, methods based on purely local measurements of proper time intervals between pulses will be expounded and the various possible sources of uncertainty will be analyzed. As sources of pulses both artificial and natural emitters will be considered. The latter will concentrate on either radio- or X ray-emitting pulsars, discussing advantages and drawbacks. As for artificial emitters, various solutions will be presented, from satellites orbiting the Earth to broadcasting devices carried both by spacecrafts and celestial bodies of the solar system. In general the accuracy of the positioning is expected to be limited, besides the instabilities and drift of the sources, by the precision of the local clock, but in any case in long journeys systematic cumulated errors will tend to become dominant. The problem can be kept under control properly using a high level of redundancy in the procedure for the calculation of the coordinates of the receiver and by mixing a number of different and complementary strategies. Finally various possibilities for doing fundamental physics experiments by means of space-time topography techniques will shortly be presented and discussed.

*E-mail: angelo.tartaglia@polito.it

1 Introduction

The problem of positioning is as old as the history of wandering of mankind especially by see. Since the oldest times the problem was tackled looking at the sky and associating the observation with time measuring. Initially time was determined using the rotation of the earth as a clock and one had to wait until the 18th century for the invention of the chronometer to reach an accuracy appropriated to the development of modern technological societies.

In our days the global positioning on earth and close to it is obtained by means of global positioning systems. The first and most used system, starting to be deployed at the end of the '70's of the last century, is named after its acronym GPS. A similar one developed by the former Soviet Union is GLONASS; Europe has started the deployment of its own Galileo system, which at the moment has three satellites in the sky. China is planning to build its global navigation and positioning system, named Bei Dou (North star). India and Japan are also planning to develop national systems, and others are also considering the possibility to do the same. The reason for this vast interest is mainly political, since both GPS and GLONASS are under military control, even when they are used for civilian purposes. In any case all these systems, deployed or under implementation, are such that a continuous control and intervention from the ground is needed.

GPS, which is a sort of an archetype of all current

positioning systems, is like a chimera made of different pieces. Basically, even using a constellation of 31 satellites (in this very moment) distributed in six orbital planes, it is transposing at the global terrestrial scale old techniques of Euclidean geometry together with Newtonian physics. It essentially determines ranges between the observer and a number of satellites (normally six of them) by means of a time of flight determination of signals sent from the orbiting emitters. At the scale of the needed positioning it is immediately clear that special and general relativistic effects cannot be neglected, so they are introduced as corrections to the classical data. Relativity enters the process in order to take into account the behaviour of the orbiting atomic clocks: on one side their pace appears to be slowed down with respect to a similar clock at rest on the surface of the earth because of the orbital speed of the satellite; on the other side the frequency of the orbiting device is increased because of the gravitational blue-shift depending on the hight of the orbit. Finally the orbiting clocks must be synchronous with respect to one another and to the local clock of the user on earth, because of the need to measure times of flight. However the purely kinematical relativistic Sagnac effect produces de-synchronization of each clock with itself at each revolution, so that from earth one has to periodically re-align all clocks.

Besides these complications and the way they are managed, it is also true that GPS is not fit to guide spacecrafts navigating across the solar system. For such navigation other techniques are used, all requiring an almost continuous guidance from earth. The distance to the spacecraft can be determined with a good accuracy (in the order of millimeters) by means of laser or radio ranging from earth, but the transverse positioning is far worse and the accuracy rapidly decays with distance. Usually the spacecrafts for the exploration of the solar system are equipped with limited capacities of self-guidance; for instance they carry pictures of the sky that allow them to an autonomous control of their trim; similarly real or reconstructed images of the final destination enable the spacecraft to autonomously guess its distance and position with respect to the intended arrival. All this is however rather complicated and not easy to manage.

The above drawbacks, despite the enormous strength of political and commercial constraints, have begun to stimulate the search for a more up-to-date approach to positioning. To say the least, we speak today of space-time as a continuous four-dimensional Riemannian manifold with Lorentzian signature. We should try and start from that fact for the building of appropriate methods to be used in order to navigate across space-time; such an approach would include general relativity from scratch and not as a set of "corrections" to be made more or less by hands.

The consequent studies have brought to a better definition of the concepts at the base of positioning, to the introduction of light coordinates and to the development of a number of proposals [14, 10, 6, 9, 7, 2, 17, 8]. Here I shall present a relativistic positioning system that has been implemented to the level of algorithms and simulations [16, 15, 21, 22, 20]. It is based on the local measurement of the "length" (i.e. the proper time interval) of a stretch of the world-line of an observer between the arrivals of subsequent pulses from not less than four independent sources represented by known world-lines in space-time. An idea considered by many authors is the one of using pulsars [12, 1, 19, 18, 11],[16, 15, 21], but other solutions can also be envisaged.

The timing, be it of laser or radio pulses, combined with relativistic positioning can also be of paramount importance for fundamental physics and I will shortly review some possibilities in the final part of the present document.

2 Reference frames

The very idea of finding a position in space-time implies the definition and assumption of a reference system with respect to which the position is defined. There can exist reference frames at various scales according to the peculiar applications one is interested in, however, in the end, some global frame needs be defined within which all other local and partial frames are located. Of course what I am writing implies that a global reference frame *can* indeed exist and uniquely be defined, which issue is not at all trivial when applied to the whole visible universe.

In practice the background reference frame that people commonly use is the one of the "fixed stars". Today by "fixed stars" quasars are meant. Quasars (quasi stellar objects) are, according to the most accepted interpretation, active galactic nuclei; the source of their energy is commonly ascribed to the presence of a massive black hole, but there are various hypotheses concerning the mass to energy conversion mechanism. What matters here, however, is that those bright objects are very far away, from approximately 3 to approximately 13 billion light years. Their distance implies that, at the human

 DOI: 10.2420/AF07.2013.111

time scales, the quasars appear as being fixed in the sky despite any proper motion they might be endowed with. From this fact arises the possibility of having fixed directions pointing along the axes of any Cartesian non-orthogonal reference frame. Thousands of quasars are known; fig. 1 taken from the Sloan Digital Survey shows the distribution in the sky of a few of them. The reciprocal angular positions of the quasars in the sky are determined by the Very Long Baseline Interferometry (VLBI) and are known, at the moment, with an accuracy of the order of 10^{-9} rad.

Of course, in order to have a reference frame, three fixed non-co-planar directions are not enough. An origin also needs be chosen and this can be done arbitrarily, in principle. In practice the best we can do, at the moment, is to choose the origin of our frame located at the barycenter of the solar system. The barycenter of the solar system is indeed moving with respect to the quasars and its motion is not an inertial one because it is constrained by the gravitational field of the galaxy, but we assume that the acceleration due to the galaxy is negligible and consequently we pretend the motion to be inertial.

This essentially is the International Celestial Reference System (ICRS), being the quasars the International Celestial Reference Frame (ICRF).

We may then refer to the International Terrestrial Reference System (ITRS) which can be connected to the ICRS using the Earth Orientation Parameters given by the IERS (International Earth Rotation Service, now International Earth Rotation and Reference Systems Service).

2.1 Space-time

Everything I have written above is OK in three dimensions, however, if we wish to have a fundamental relativistic description, we must refer to space-time as a Riemannian continuum with Lorentzian signature. This means that we need to include time among the coordinates. When defining an origin for our reference system we need to introduce an origin of time as well; it can be arbitrary, of course, but what we really need is to associate a duration standard to our space origin. We consequently imagine to place an atomic clock in the barycenter of the solar system and to use its time as our coordinate time. This assumption is not trivial at all, since we know that, if we compare the readings of two identical clocks located along two different world-lines, we find they can differ from one another because of relative motion of the two clocks and them being placed in different gravitational potential wells.

In fact, considering space-time, we see that it appears locally as a sort of crumpled manifold like in fig. 2.

The warps between the quasars and the local terrestrial observers can be no problem provided they stay stable during our progressive exploration of our world-line, over times of the order of typical human times. This assumption can be reasonable for the path along our past light-cone out of the solar system, but may be questionable in the final portion close to the observer, where the proper motion of the latter in the local gravitational potential well can introduce non-negligible time changes.

When setting up a global reference system in a general relativistic framework, our definitions actually rest on a number of implicit assumptions that are added to the explicit ones. For instance the ICRS implicitly assumes that space-time is asymptotically flat. Treating quasars as fixed (point-like) objects amounts to say that their world-lines are straight and parallel: this implies that "there" space-time is flat. Most likely, however, in the universe there is no asymptotic flatness; at most we may say that in between galaxies, far away from any matter bunch, space-time is *almost* flat.

Putting everything together, we should rather say that the ICRS (or any other analogous reference system) is defined as being "drawn" on a flat Minkowski space-time which coincides with the tangent space-time in the origin of our reference frame. In order to use the proper time of the atomic clock located in the origin as the global coordinate time of our reference implies that we treat it as being in a globally flat environment.

We have no such problem at the moment, but the use of the local tangent space-time would make it not trivial at all to uniquely and understandably transfer position information to another observer a few million light years away.

2.2 Coordinates and geodesics

Looking at the problem of defining efficient reference systems for a Riemannian manifold we need also to decide how to uniquely and smoothly attribute to each event in the manifold a quadruple of numbers i.e. an appropriate coordinate set. This may be done in principle drawing four independent families of curves. The curves of each family do not intersect each other and densely cover the whole manifold. Labeling each curve by a progressive real number, any intersection of four curves from different families identifies an event on the

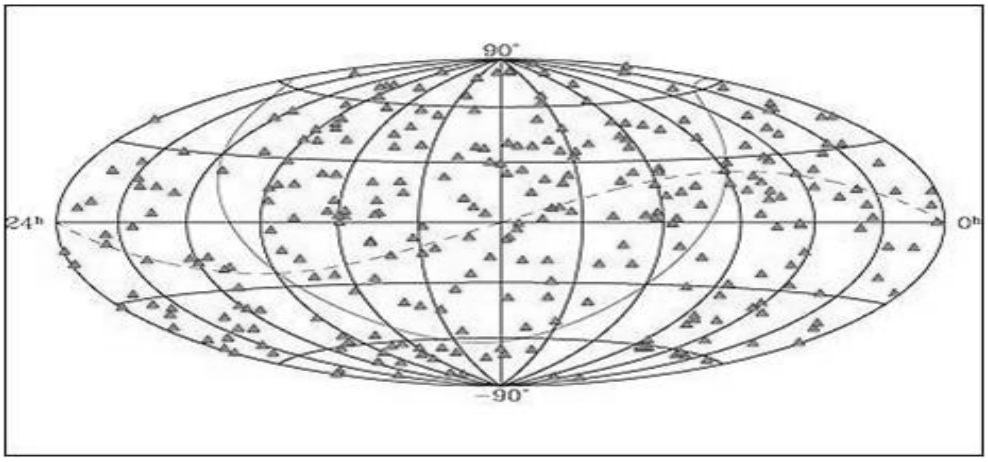


Figure 1. *The distribution in the sky of a few quasars taken from the Sloan Digital Survey*

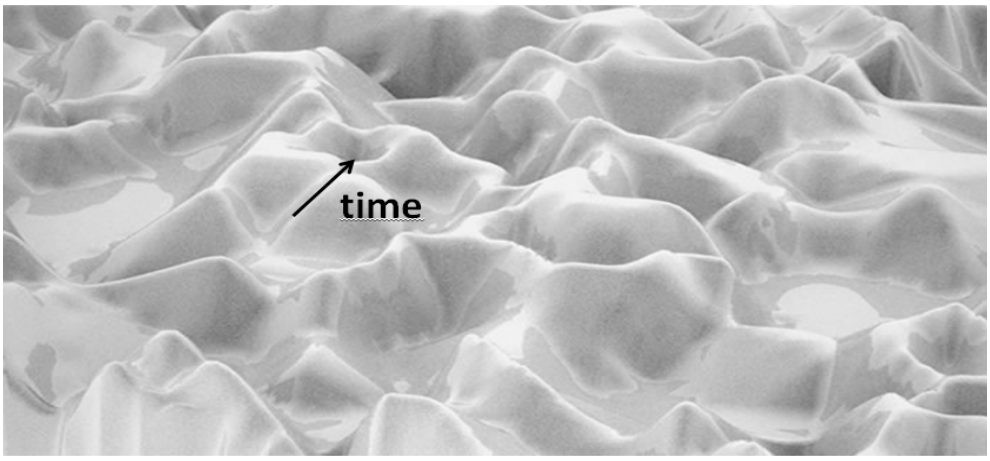


Figure 2. *Artistic three-dimensional view of a "crumpled" space-time*

 DOI: 10.2420/AF07.2013.111

manifold and endows it with four coordinates. What I have described by this process is known as a Gaussian coordinates system. As for the curves to be used they can for instance be geodesics of the manifold. Not considering singularities and defects, geodesics do indeed cover the whole manifold. In the case of space-time and in view of the importance of the null cones a good choice for building a Gaussian coordinate system is to use independent families of null geodesics. Fig. 3 visualizes our choice in a simplified bi-dimensional space-time.

In a flat manifold a whole family of null geodesics may be identified by means of a tangent vector to anyone of them. It would be a null four-vector like the one written in eq. (1).

$$\chi = cT(1, \cos\alpha, \cos\beta, \cos\gamma) \tag{1}$$

It is indeed $\chi^2 = 0$. The space components of (1) are the direction cosines of a three-vector defined with respect to some global reference system of our choice. The cT factor is optional and does not change the null character of χ, but contains an additional information on the period of the signals traveling along the null geodesics of the family identified by χ, provided they are indeed periodic.

Suppose now that we have four independent null four-vectors $\chi_{1,2,3,4}$ and use them as a basis for vectors in the manifold; the position of any event in the same manifold is identified by a "radial" four-vector, expressed as a linear combination of the basis vectors:

$$r = \left(\frac{\tau}{T}\right)_a \chi^a \tag{2}$$

The index a runs from 1 to 4; the periods of the signals are assumed to be different from one another in the quadruple; τ's are named "light coordinates" of the event and the ratios $(\frac{\tau}{T})_a = x_a$ are pure numbers due to the choice for the scale factor in eq. (1).

The number of actual degrees of freedom in our representation deserves a comment. All χ's of the basis are on the light cone of the origin of the reference frame (actually on any light cone if the manifold is globally flat), which means that, as far as we stay on the light cone, only three of the null wave vectors can be mutually independent. Three χ's are enough to localize events on the light cone; we need four of them for time-like or space-like events, i.e. for events out of the light cone.

Rather than using the families of null geodesics, we may adopt a dual vision. Each χ is associated with a null four-dimensional hyperplane obtained by Hodge conjugation. The corresponding four-form ϖ is $\varpi = *\chi$, or explicitly:

$$\varpi_{abc} = \epsilon_{abcd}\chi^d \tag{3}$$

ϵ_{abcd} is the Levi-Civita fully antisymmetric tensor. Now we have four independent families of hyperplanes covering the whole space-time and intersecting each other. The hyperplanes of a family are null and orthogonal to the corresponding χ.

All this is globally true if the manifold is flat; if it is curved it holds locally.

3 Positioning

On the bases laid down in the previous section, we may outline a fully relativistic positioning method. Suppose you have (not less than) four independent sources of electromagnetic signals located at infinity; suppose then that they emit pulses at the rate of $1/T$ per second. The T parameter of formula (1) is now interpreted as the repetition time of the pulses rather than the period of a monochromatic continuous wave. Once this has been specified we may apply the procedure outlined in the previous section. The χ's are associated to the four (or more) sources; we may identify as duals to the χ's four discrete sets of hyperplanes ϖ covering space-time with an egg crate whose spacings along the directions of the basis vectors are given by the T's, when measured along the time axis of the background global reference frame.

The situation is schematically shown in fig. 4.

The world-line of an observer necessarily crosses the walls of successive boxes of the egg crate. If we are able to label each cell of the crate, we are also able to reconstruct the position of the observer in the manifold. The use of pulses implies that, realistically, the walls of the cells are "thick". In practice the hypersurfaces on the graph correspond to "sandwich waves" carrying the pulse. A typical emission diagram of one of the sources will more or less be like the one sketched in fig. 5.

The shape of the pulse is not important as well as it is not the spectral content of it. What matters is its reproducibility and the stability of the repetition time. Considering natural pulses, as the ones coming from pulsars, we find repetition times ranging from several seconds down to a few milliseconds and lasting a fraction of the period. As an example of artificial pulses the highest performance is obtained with lasers: GHz frequencies are possible with pulses as short as $\sim 10^{-15}$ s.

Once pulses are used, we may label them in order, by

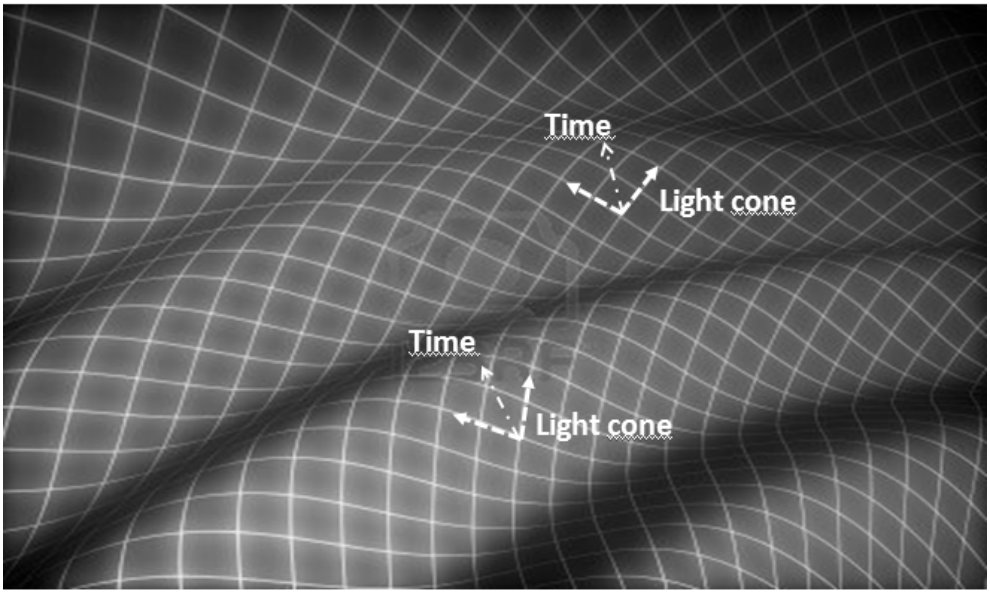


Figure 3. *Two sets of null geodesics covering a bidimensional curved manifold*

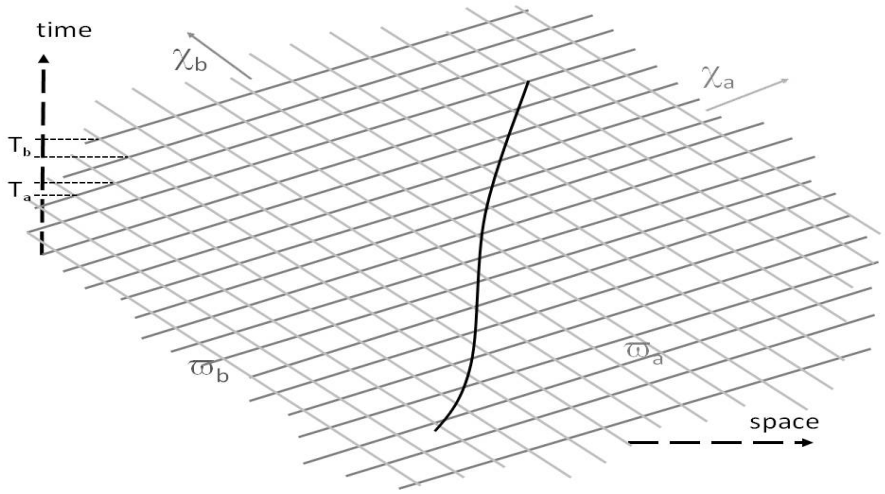


Figure 4. *A bidimensional flat space-time covered by a grid made of null hypersurfaces (actually lines) conjugated to the null vectors* $\chi_{a,b}$. *The wavy line is the world-line of an observer.*

 DOI: 10.2420/AF07.2013.111

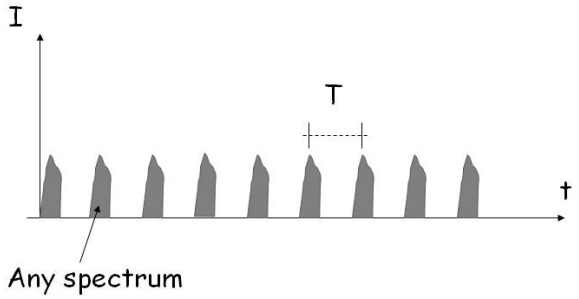


FIGURE 5. *Typical emission sequence of the pulses from a source. Vertically intensities are drawn; the profile of the pulse is not important; times are proper times of the emitter.*

integer numbers, as it can schematically be seen in fig. 6.

The integers can be though of as rough coordinates identifying the cells of the grid. At this level the approximation would be rather poor, being of the order of the size of each cell. If the periods are of the order of milliseconds this corresponds to hundreds of kilometers. Looking at fig. 6 we may however notice that the intersections of a given world-line with the walls of the cells are labeled by a quadruple of numbers, at least one of which is an integer: these numbers are the coordinates of the intersection points. We may write the typical coordinate of a position in the crate as $\xi_a = n_a + x_a$; the n's are the integers, whilst the x's are the fractional parts. If we have a means to determine the x's the localization of an intersection event can be done with an accuracy much better than the hundreds of km I mentioned above. Considering that the intersections coincide with the arrivals of pulses from different sources, the determination of the fractional part of the coordinates is indeed a trivial task, provided the traveler carries a clock, the space-time is flat and the world-line is straight. Once one measures the proper intervals between the arrivals of successive pulses a simple linear algorithm based on elementary four-dimensional flat geometry produces the x's [21]. The corresponding light coordinates are $\tau_a = [(n+x)T]_a$. The accuracy of the result depends on the precision of the clock which is being used in order to measure the proper intervals between pulses and on the stability of the period of the pulses, which in turn tells us what the effective "thickness" of the walls of the cells of our space-time crate is. Just to fix some order of magnitude, let me remark that nowadays to have a portable clock with a 10^{-10}s accuracy is quite easy (much better can be achieved in the lab); on the other side, considering pulsars, we have some, whose period is known and stable down to 10^{-15}s. With these figures the final positioning can be within a few centimeters.

Of course the traveler's motion will not in general be an inertial one and space-time will not be flat, however a short enough stretch of the world-line can always be confused with the tangent straight line to it and a small enough patch of space-time can always be confused with a portion of the local tangent space. In practice we work on the local tangent space and on a linearized portion of the world-line. The acceptability of these assumptions depends on the accuracy required for the positioning and on the constraints posed by the linear algorithm in use. The reconstruction of a piece of the world-line requires the knowledge of at least eight successive arrival times of pulses from the minimal set of independent sources (four) [21]. So, if $\delta\tau$ is the maximum proper

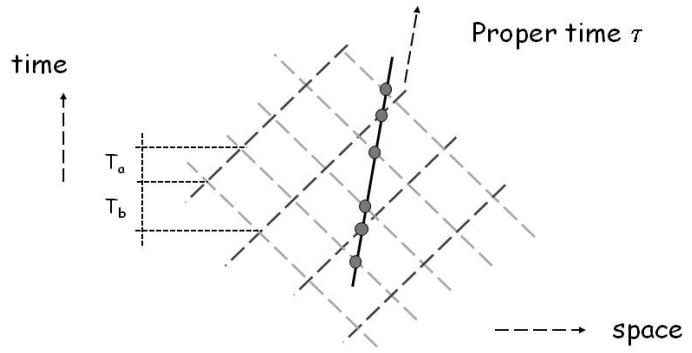


Figure 6. *A straight portion of a world-line is shown in a local flat patch of space-time. The lines of the grid correspond to different pulses labeled by their ordinal integer. The intersections of the world-line with the walls are localized by quadruples (actually pairs in the figure) of real numbers one of which is always an integer.*

time inaccuracy that we decide to be tolerable, the final relative accuracy of the positioning will be:

$$\left|\frac{\delta x}{x}\right| \leqslant 4\left(\frac{1}{\tau_{i,i+4n}} + \frac{\tau_{i,i+1}}{\tau^2_{i,i+4n}}\right)\delta\tau \qquad (4)$$

The index i in eq. (4) labels the order of the arrival events; $\tau_{i,i+4n}$ is the proper time interval between the ith and the $(i+4n)$th arrival, being $n \geqslant 1$ an integer; n should assume the highest value compatible with the straightness hypothesis for the world-line. Of course the number of pitches that can safely be considered depends on the periods T_a of the emitting sources: the shorter are the periods, the bigger is the number of paces that can be used within the linearity assumption.

A pictorial view of what we are doing is as follows. Imagine to embed the real four-dimensional manifold, together with its tangent space at the start event, in a five-dimensional flat manifold; then consider the real world-line of the traveler and project it onto the tangent space. The world-line on the tangent space is what we are piecewise reconstructing by our linear algorithm: in practice we are building a flat chart containing the projection of our space-time trajectory. The time dependence of the a-dimensional coordinates of the projected world-line may of course be written in the form of a power series, as:

$$x_a = u_a\frac{\tau}{T_a} + \frac{1}{2}\alpha_a\frac{\tau^2}{T_a^2} + ... \qquad (5)$$

The coefficients u_a and α_a are proportional to the four-velocity and four-acceleration of the traveler. The individual segments used for the reconstruction are short enough so that the second and further terms of (5) are negligible with respect to the linear one. Going on, after a number of paces, the possible presence of an extrinsic curvature of the projected world-line shows up; we know that locally it is impossible to distinguish a gravitational field from a non-gravitational acceleration so we need additional information for that purpose. In the case of a gravitational field evidenced by the reconstruction process I am describing, we get from the data the gradient of the Newtonian gravitational potential.

In order not to cumulate the distortion introduced by the projection from the real curved manifold to the tangent space at a given event, we need periodically to restart from a further event on the world-line, i.e. to pass to the tangent space at a different event. If the visible curvature of the line on the tangent space as well as the tilt of the successive tangent spaces continues for long in the same sense, the linearization process, as in all

 DOI: 10.2420/AF07.2013.111

similar cases, tends to produce a growing systematic discrepancy with respect to the real world-line, so that periodically one has to have recourse to some independent position fixing means in order to reset the procedure.

3.1 Pulsars

I have already mentioned pulsars as possible natural sources of pulses. This kind of neutron stars are indeed good pulse emitters because of their extreme stability and long duration. As we know, their emission is in the form of a continuous beam. The apparent periodicity is due to the fact that the emission axis (the magnetic axis) does not coincide with the spin axis of the object so that it steadily rotates, together with the whole star, about the direction of the angular momentum. The pulses arise from the periodic illumination of the earth by the rotating beam. The stability is guaranteed by the angular momentum conservation.

The advantages of pulsars are numerous. Their period is extremely stable and is sometimes known with the accuracy of 10^{-15} s; it tends to decay slowly (the relevant times are at least months), but with a very well known trend, determined by the emission of gravitational radiation. Typically the fractional decay rate of the period is in the order of one part in 10^{12} per year. The number of such sources is rather high, so that redundancy in the choice of the sources is not a problem: at present approximately 2000 pulsars are known and their number continues to increase year after year. Being these stars at distances of thousands of light years from the earth, they can be treated as being practically fixed in the sky; in any case their slow apparent motion in the sky is known, so corrections for the position are easily introduced. Just to recall some numbers, the rate of change of a typical angular coordinate α in the position in the sky is

$$\left|\frac{\delta\alpha}{t}\right| \approx 10^{-6}\left(\frac{100\mathrm{pc}}{\mathrm{distance}}\right)\frac{\mathrm{rad}}{\mathrm{year}} \qquad (6)$$

Unfortunately pulsars have also major drawbacks. One is that their distribution in the sky is uneven, since they are mostly concentrated in the galactic plane, which fact brings about the so called "geometric dilution" of the accuracy of the final positioning: sources located on the same side of the observer produce an amplification of the inaccuracy originating in the intrinsic uncertainties. Furthermore individual pulses differ in shape from one another so that some integration time is needed in order to reconstruct a fiducial series of pulses; this fact, also considering the length of the repetition time, can conflict with the linearization of the world-line of the traveler. It should also be mentioned that most pulsars are subject to sudden jumps in the frequency (glitches), caused by matter falling onto the star; these unpredictable changes can be made unoffensive by means of redundancy, i.e. making use of more than four sources at a time.

However the most relevant inconvenience with pulsars is their extreme faintness. In the radio domain their signals can be even 50 dB below the noise at the corresponding frequencies; to overcome this problem big antennas are required (not less than 100 m^2) and convenient integration times accompanied with "folding" techniques must be employed. In principle at least four different sources must be looked at simultaneously and this is not an easy task, especially with huge antennas. The weakness problem has led to consider X-ray- rather than radio-pulsars for positioning. A few hundreds X-ray emitting pulsars are indeed known; their signals are weak too, and can be received only outside the atmosphere, but the background noise is far smaller than the one typical in the radio domain; as for the hardware, X-ray antennas can be much smaller than the typical radio-antennas. Since many galactic X sources emit also at radio frequencies, one can envisage the opportunity to combine both X-ray and radio pulses from one single source for the positioning process.

3.2 Artificial and blended solutions

In principle what can be done using pulsars can as well be done by means of artificial emitters of electromagnetic pulses. Artificial emitters can have far higher intensities than pulsars; the repetition time can easily be in the range of ns or less, thus making the linearization process more reliable. The stability over time of the source is not as good as for pulsars, but this can represent no inconvenience as far as the number of sources is redundant and they are kept under control. A problem is in the sources clearly not being at infinite distance, which implies a more complicated geometry and of course the need for a good knowledge of the world-line of the emitter in the background reference frame.

One could think of building a Solar System reference frame made of pulse emitters laid down on the surface of various celestial bodies whose orbits are well known and reproducible: the earth of course, the moon, Mars, maybe some of the asteroids; even some space station following a well defined, highly stable orbit around the sun or a planet.

A blended solution for self-guided navigation in the solar system could combine some artificial emitters, as quoted above, together with a limited number of pulsars (the most intensely emitting ones).

The fully relativistic method I have described is of course specially fit for space navigation, but it can also be useful in rather limited areas. Think for instance to the accurate mapping of a depopulated region, where traditional topography may be rather expensive. If one puts a limited number of antennas (not less than four in any case) emitting pulses, located in precisely defined positions at the boundary of the region to map, their signals may be used by a moving vehicle to draw a chart of the area within centimeter accuracy. Of course the same can be done using differential GPS, but with the specific features of GPS, mainly the fact that it is under military control and the best performance of the system is not reserved to ordinary civilian applications.

One could also think that the new relativistic method will be implemented in the next generations of global positioning systems, even though the approach tends to be rather conservative there, the reason being that the huge amount of money already spent for developing and deploying the traditional GPS makes present applications based on it cheap, whereas any new solution would initially be more expensive. Probably a gradual transition will happen, triggered by new applications, especially outside the terrestrial environment, and, last but not least, political reasons.

4 Positioning and fundamental physics in space

The method I have being describing for positioning purposes is based on electromagnetic signals and their accurate timing. The same kind of technology can be used for various experiments aimed to the detection of fundamental properties of space-time. It is worth mentioning a few possibilities.

4.1 Intercommunicating swarms of satellites

Consider a swarm of identical satellites (as the ones of the future Galileo system), equipped with pulse emitters and receivers and able to accurately measure the arrival times and to recognize the origin of each pulse (this could be achieved tuning the emitters on different individual frequencies). The information gathered by the whole constellation would allow for space-time geodesy, based on multiple triangulations performed on null triangles. It would be a means to reconstruct the average curvature of the patch of the manifold where the world-lines of the satellites lie. The Galileo satellites will indeed be able to intercommunicate and also the present GPS satellites may communicate with each other even though this possibility (introduced for military reasons) is not actually used at the moment.

4.2 Ring-lasers

Electromagnetic waves can be used as probes for the structure of space-time and in particular the gravito-magnetic part of the gravitational interaction, exploiting the anisotropic propagation of light induced by the chiral symmetry associated with a rotating mass. This possibility is the basis of the proposal to use ring lasers for the measurement of the Lense-Thirring (frame dragging) effect of the earth [3]. If a light beam is obliged, by conveniently located mirrors, to follow a closed path in space, the total time of flight for a loop is different according to the fact that light is moving in the same sense as the rotation of the central mass or in the opposite sense. In fact the difference in the proper (i.e. of the laboratory) times of flight for one turn in co- and counter-rotating sense, is obtained as [3]:

$$\delta\tau = -2\sqrt{g_{00}}\oint \frac{g_{0\phi}}{g_{00}}\,d\phi \qquad (7)$$

Polar coordinates centered on the earth are assumed and the $g_{\mu\nu}$'s are elements of the metric tensor of an axially symmetric stationary space-time. A ring laser converts the time of flight difference in a beat note obtained from the two counter-rotating beams in steady state; the beat note arises from the different equilibrium frequencies of the two beams. The device in practice can measure effective angular velocities, which contain the kinematical effect (classical Sagnac effect), the geodetic (or de Sitter) effect (coupling of the gravito-electric field of the earth with the kinematical rotation of the apparatus) and the gravito-magnetic contribution (Lense-Thirring frame dragging). The latter two terms are 9 orders of magnitude smaller than the classical Sagnac effect when measured on the surface of the planet, so that a very high sensitivity is needed, but contemporary laser technologies are approaching the required accuracy level. Both the de Sitter and Lense-Thirring effects have already been measured in space by a different technique based on the behaviour of mechanical gyroscopes. The most difficult to reveal is the Lense-Thirring drag and it has been measured with an accuracy of 19% by the

 DOI: 10.2420/AF07.2013.111

GP-B experiment [13] and 10% by the laser ranging of the LAGEOS satellites [5]; laser ranging of the orbit is also being used by the LARES experiment launched on February 13th 2012, with the purpose of reaching the 1% accuracy [4]. The newly proposed ring laser GINGER experiment [3] is aimed at reaching a 1% accuracy for the physical terms in a terrestrial laboratory.

Here I would like to mention the possibility of bringing a ring laser experiment in space. One could for instance think of a three-dimensional array of four mirrors, rigidly attached to one another in the shape of a tetrahedron; the whole thing could be in free fall (stable orbit) around the earth. Each face of the tetrahedron would coincide with a triangular ring laser; the signal extracted from each face would give information on the projection of the total rotation vector on the normals to the faces. In the case of a circular equatorial orbit the frequency of the beat note extracted from one of the faces of area S and perimeter length P would be:

$$f_{beat} = 4\frac{S}{\lambda PR}\sqrt{G\frac{M}{R}}\cdot \left[\left(1-\frac{3}{4}\sqrt{G\frac{M}{c^2R}}+4G\frac{M}{c^2R}\right)\hat{n_a}\cdot\hat{n_S} - \left(\frac{1}{2}G\frac{M}{c^2R}+\frac{GJ}{c^4R}\right)\hat{n_\theta}\cdot\hat{n_S}\right] \quad (8)$$

M is the mass of the earth; J is its angular momentum; G is Newton's constant; R is the radius of the orbit; λ is the wavelength of the light of the laser; $\hat{n_S}$, $\hat{n_a}$, $\hat{n_\theta}$ are unit vectors, respectively, perpendicular to the plane of the ring, aligned with the axis of the earth, aligned with the local meridian in the sense of increasing co-latitude (here, in practice, perpendicular to the equatorial plane).

The term depending on the angular momentum of the earth in eq. (8) is the smallest and is eight orders of magnitude below the biggest; an extremely good accuracy is always required, but in free fall one has a far smaller environmental noise than on earth.

4.3 Linear cavities

Another interesting possibility is represented by simple linear resonating cavities as the ones in Fabry-Pérot interferometers. In fact, when describing a simple bounce back from a mirror to the other in four-dimensional space-time, one has a bidimensional graph, like the one shown on fig. 7 where an active region is assumed in the middle of two mirrors.

As it can be seen, the light beams moving back and forth in the cavity delimit a closed contour at each cycle. This fact implies that the effects of the curvature and the chiral symmetry of space-time may be expressed in terms of the Riemann tensor and the contoured area. Considering the electromagnetic tensor $F_{\mu\nu}$, its change after one cycle is given by:

$$\delta F^{\mu\nu} = (R^{\mu}_{\epsilon 0 i}F^{\epsilon\nu} + R^{\nu}_{\epsilon 0 i}F^{\mu\epsilon})\delta S^{0i} \quad (9)$$

Latin indices are used for space-coordinates; $\delta S^{\mu\nu}$ is the antisymmetric area 2-form.

For practical purposes, only the least useful approximation of the Riemann tensor needs be retained down to the order of the angular momentum of the earth. An example of the approximated version of one of the equations (9) is for instance:

$$\delta F^{\theta\phi} \cong \left(\frac{GM^{3/2}}{c^3R^{7/2}} - 3\frac{GJ}{c^3R^4}\right)\frac{\cos\theta}{\sin^2\theta}\frac{l^2}{R}F^{\theta r} \quad (10)$$

$\delta F^{\theta\phi}$ is the change in the radial component of the magnetic field expressed as a function of the East-West component; l is the length of the cavity and R is its radial position with respect to the center of the earth.

The result depends on the orientation of the cavity and builds up with the successive reflections. One could think of combining this effect along an array of mutually perpendicular freely falling cavities.

4.4 An orbital ring cavity

The satellites of the GPS constellation are distributed on 6 different orbital planes so that 5 of them are on the same orbit; the Galileo system, when fully deployed will have 30 satellites on 3 orbital planes so that 10 satellites will share the same orbit. The presence of at least three satellites on the same orbit opens an interesting possibility if they are enabled to communicate with each other. Suppose each satellite is sending laser pulses to the others who are forwarding them along the orbit. In practice we would have a sort of ring laser at orbital scale. If at least one of the satellites is equipped with interferometric devices or can accurately measure the arrival time difference between pulses having completed a clockwise turn with respect to the ones revolving in the counterclockwise sense, the whole system behaves as a gigantic ring laser (or Sagnac interferometer) with a sensitivity measured by the huge scale factor given by the ration between the contoured area and the length of the perimeter of the polygon followed by the light pulses.

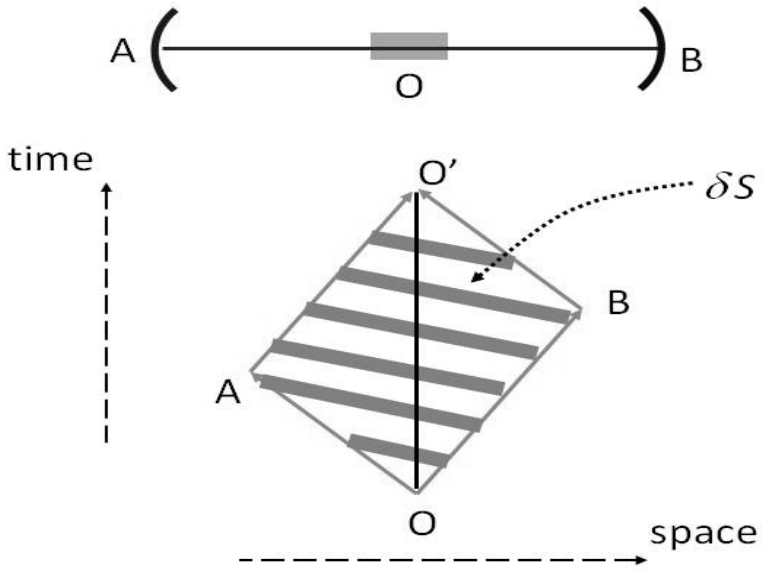


Figure 7. *World-sheet representation of light being reflected back and forth in a linear cavity. A and B are mirrors; O-O' is the world-line of an active region in the middle of the two mirrors; the gray stripes identify a space-time area* δS.

5 Conclusion

Summing up, I have shown how the approaches to positioning and navigation could be implemented in order to become fully relativistic. The idea is in the use of space-time as such as a reference, and in the exploitation of four-dimensional geometry. In practice a generalization is possible of the ordinary three-dimensional topographic techniques, upgrading them to four dimensions and the Lorentzian signature. The method, with the related algorithms, is per se simple and relies on plain proper time measurements made by the traveller needing to localize himself in a given background reference frame. Reducing everything to the essentials, we see that the pattern of the proper arrival times of regular pulses from not less than four independent sources is uniquely related to the position of the receiver in space-time and to its starting event. The relativistic positioning method has positively been tested with simulators. Fig. 8 shows for instance the reconstruction of three days of the absolute motion of the antenna of the radio-telescope at the Parkes observatory,obtained by the simulated timing of four real pulsars [15].

This approach, either based on signals from pulsars (X ray or radio waves emitters) or on artificial sources laid down on the surface of the earth and other bodies of the solar system, will probably raise growing interest little by little as the need for positioning systems freed from the control of any specific power will increase. The same holds with the expansion of navigation within the solar system. I have also given a few examples of the importance of timing measurements together with the use of laser beams or pulses for fundamental physics. It turns out that light is indeed a perfect relativistic probe for testing the structure of space-time.

References

[1] J. S. Alvarez, A. U. Planas, and J. V. Piera. Feasibility study for a spacecraft navigation system relying on pulsar timing information, ARIADNA project 03/4202. Technical report, ESA Advanced Concepts Team, 2004.

[2] D. Bini, A. Geralico, M. L. Ruggiero, and A. Tartaglia. Emission versus Fermi coordinates: applications to relativistic positioning systems. *Classical and Quantum Gravity*, 25(20):205011, 2008.

[3] F. Bosi, G. Cella, A. Di Virgilio, A. Ortolan, A. Porzio, S. Solimeno, M. Cerdonio, J. P. Zendri, M. Allegrini, J. Belfi, N. Beverini,

 DOI: 10.2420/AF07.2013.111

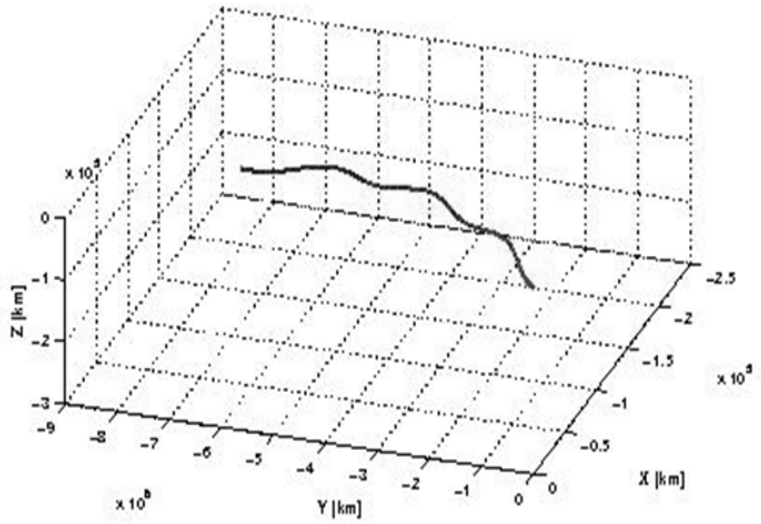


Figure 8. *The picture shows the motion of the Parkes observatory carried by the earth in its rotation and revolution motion, during three days. The reconstruction has been made applying the relativistic positioning method to the simulated arrival times of the signals from four real pulsars. The reconstructed trajectory is superposed to the real one.*

B. Bouhadef, G. Carelli, I. Ferrante, E. Maccioni, R. Passaquieti, F. Stefani, M. L. Ruggiero, A. Tartaglia, K. U. Schreiber, A. Gebauer, and J.-P. R. Wells. Measuring gravitomagnetic effects by a multi-ring-laser gyroscope. *Phys. Rev. D*, 84:122002, Dec 2011.

[4] I. Ciufolini, A. Paolozzi, and C. Paris. Overview of the LARES mission: orbit, error analysis and technological aspects. *Journal of Physics: Conference Series*, 354(1):012002, 2012.

[5] I. Ciufolini, A. Paolozzi, E. Pavlis, J. Ries, R. Koenig, R. Matzner, G. Sindoni, and H. Neumayer. Testing gravitational physics with satellite laser ranging. *The European Physical Journal Plus*, 126(8):1–19, 2011.

[6] B. Coll. Relativistic positioning systems. In *A century of Relativity Physics ERE 2005; XXVIII Spanish Relativity Meeting*, volume 841, pages 277–284. AIP Conf. Proc., 2006.

[7] B. Coll, J. J. Ferrando, and J. A. Morales. Positioning with stationary emitters in a two-dimensional space-time. *Phys. Rev. D*, 74:104003, Nov 2006.

[8] B. Coll, J. J. Ferrando, and J. A. Morales. Four causal classes of Newtonian frames. *Found. of Phys*, 39:1280–1295, 2009.

[9] B. Coll, J. J. Ferrando, and J. A. Morales-Lladosa. Two-dimensional approach to relativistic positioning systems. *Physical Review D*, 73:084017, 2006.

[10] B. Coll and A. Tarantola. A galactic positionning system. In N. C. A. Finkelstein, editor, *Proceedings of the "Journées 2003 Systèmes de Référence Spatio-Temporels"*, pages 333–334. Institute of Applied Astronomy, St.Petersburg, 2004.

[11] B. Coll and A. Tarantola. Using pulsars to define space-time coordinates. arxiv:0905.4121v1. 2009.

[12] G. S. Downs. Interplanetary navigation using pulsating radio sources. In *NASA Technical Reports N74-34150*, pages 1–12. NASA, 1974.

[13] C. W. F. Everitt, D. B. DeBra, B. W. Parkinson, J. P. Turneaure, J. W. Conklin, M. I. Heifetz, G. M. Keiser, A. S. Silbergleit, T. Holmes, J. Kolodziejczak, M. Al-Meshari, J. C. Mester, B. Muhlfelder, V. G. Solomonik, K. Stahl, P. W.

Worden, W. Bencze, S. Buchman, B. Clarke, A. Al-Jadaan, H. Al-Jibreen, J. Li, J. A. Lipa, J. M. Lockhart, B. Al-Suwaidan, M. Taber, and S. Wang. Gravity Probe B: Final results of a space experiment to test general relativity. *Phys. Rev. Lett.*, 106:221101, May 2011.

[14] C. Rovelli. GPS observables in general relativity. *Phys. Rev. D*, 65:044017, Jan 2002.

[15] M. L. Ruggiero, E. Capolongo, and A. Tartaglia. Pulsars as celestial beacons to detect the motion of the earth. *International Journal of Modern Physics D*, 20(06):1025–1038, 2011.

[16] M. L. Ruggiero, E. Capolongo, and A. Tartaglia. *Solar System: Structure, Formation and Exploration*, chapter A Relativistic Positioning System Exploiting Pulsating Sources for Navigation across the Solar System and Beyond, pages 75–95. Space Science, Exploration and Policies. Nova Science Publishers, 2012.

[17] M. L. Ruggiero and A. Tartaglia. Mapping cartesian coordinates into emission coordinates: Some toy models. *International Journal of Modern Physics D*, 17(02):311–326, 2008.

[18] S. I. Sheikh, R. W. Hellings, and R. A. Matzner. High-order pulsar timing for navigation. In *Proceedings of the 63rd Annual Meeting of The Institute of Navigation*, pages 432–443. ION Publications, 2007.

[19] S. I. Sheikh, D. J. Pines, P. S. Ray, K. S. Wood, M. N. Lovellette, and M. T. Wolff. Spacecraft navigation using X-ray pulsars. *Journal of Guidance, Control, and Dynamics*, 29(1):49–63, Jan 2006.

[20] A. Tartaglia. Emission coordinates for the navigation in space. *Acta Astronautica*, 67:539–545, Sept. 2010.

[21] A. Tartaglia, M. L. Ruggiero, and E. Capolongo. A null frame for spacetime positioning by means of pulsating sources. *Advances in Space Research*, 47(4):645 – 653, 2011.

[22] A. Tartaglia, M. L. Ruggiero, and E. Capolongo. A relativistic navigation system for space. *Acta Futura*, 4:33–40, 2011.

 DOI: 10.2420/AF07.2013.111