Astron. Astrophys. 332, 1133–1141 (1998)

ASTRONOMY AND ASTROPHYSICS

General relativistic satellite astrometry

I. A non-perturbative approach to data reduction

F. de Felice[1,⋆], **M.G. Lattanzi**[2,⋆⋆], **A. Vecchiato**[1], **and P.L. Bernacca**[3]

[1] Istituto di Fisica "G. Galiei", Via Marzolo 1, I-35100 Padova, Italy
[2] Osservatorio Astronomico di Torino, Strada Osservatorio 20, I-10025 Pino Torinese, Italy
[3] Astronomy Department, Università di Padova, Vicolo dell'Osservatorio, I-35100 Padova, Italy

Received 28 August 1997 / Accepted 27 November 1997

Abstract. A general relativistic scenario is utilized to build a non-perturbative model, in Schwarzschild metric, for the representation of observed angles among star pairs. This model is then applied to an end-to-end simulation of the GAIA satellite, a concept for global astrometry within the *2000+ scientific program* of the European Space Agency. GAIA is expected to measure positions, parallaxes, and annual proper motions to better than 20 μarcsec for more than 50 million stars brighter than V$\simeq$16 mag. This first attempt at modeling global astrometric data within the framework of general relativity considers a *static* sphere, namely, only Schwarzschild *azimuth* and *colatitude* are estimated, locating stars on the Schwarzschild sphere centered at the Sun. The results show that measurements of large arcs among stars, each measurement good to ~ 100 μarcsec (as expected for V$\simeq$ 16 mag stars), repeated over an observing period of only one year can be modeled to yield $\sim$ 20 μarcsec errors on the estimated relativistic parameters.

Although it remains to be established if the non-perturbative approach can be extended to a more realistic observing scenario (including the oblate and rotating Sun, and the other planets of the solar system), these results provide strong evidence that the lessons learned with Hipparcos apply to the 10^{-5} arcsec regime of the GAIA mission.

Key words: relativity – methods: data analysis – space vehicles – astrometry

1. Introduction

GAIA (Global Astrometric Interferometer for Astrophysics), is a concept for a global astrometric satellite originally proposed by Lindegren and Perryman (1996 and references therein) as a possible interferometric Cornerstone-class mission within the *2000+ program* of scientific satellites of the European Space Agency.

Send offprint requests to: lattanzi@to.astro.it
⋆ INFN, sezione di Padova
⋆⋆ Also: Space Telescope Sc. Institute

This scanning satellite is designed to chart all ~ 50 million stars expected on the celestial sphere down to the operational limiting magnitude of V $\simeq$ 16 mag. The goal is an accuracy of $\sim$ 20 μarcsec on positions and parallaxes (better for brighter stars), and 20 μarcsec per year for proper motions. These values represent an improvement of a factor 100 over the astrometric accuracy of Hipparcos, the highly succesful global astrometry mission considered the predecessor of the GAIA concept. Therefore, GAIA will extend the 10%-error distance horizon by the same factor to ~ 10 kpc, or a factor 10^6 in volume.

A detailed list of outstanding research topics which could be dramatically improved upon, or tackled, with GAIA is discussed in Lindegren and Perryman (1996), and in ensamble of articles presented for the Scientific Aspects parallel session of the first GAIA meeting held in Cambridge in 1995 (Perryman and van Leeuwen 1995). To the science addressed at that meeting we like to add, or call attention to, few selected items. Like the existence of a stellar warp in the disk of our galaxy; a topic that Hipparcos can just begin to address (Smart and Lattanzi 1996), and only the GAIA accuracy will be able to fully characterize its kinematic consequences including small precession rates of ~ 0.5 mas/year, or less. Or like the importance that sub-mas positional accuracies will have in the registration of optical and radio images for a better understanding of energy production and energy transport mechanisms at work in, *e.g.*, Seyfert galaxies (Capetti *et al.* 1997).

Also, recent calculations based on simulated observations support earlier indications (Casertano *et al.*, 1996) that GAIA could discover Jupiter-like planets (orbiting 1-$M_\odot$ parent stars) with a detection probability $\geq 50\%$ up to distances of 100 pc and beyond (Lattanzi *et al.*, 1997). Furthermore, besides its fundamental role in unraveling some of the most intriguing problems of modern astrophysics, GAIA will also be instrumental in fundamental physics. GAIA could be the first in-orbit experiment to try *astrometric detection* of gravitational waves (Fakir 1995; see also Fakir 1994 and Makarov 1995). Although fascinating, gravitational wave astrometry is not on solid footing as yet, and further studies must be devoted to both a rigorous formulation of the effect and its actual observability within the context of the GAIA concept.

Finally, closer to the subject of this article is gravitational light bending. The possibility of accurately measuring the parameter γ of the Parametrized Post-Newtonian (PPN) formulation of gravitational theories is of utmost importance in fundamental physics. The more accurate the measurement, the more decisive it will be in differentiating amongst different metric theories of gravity. Some estimations based on a typical formula for light bending due to a non-rotating Sun (see *e.g.* Misner *et al.* 1973) indicate that GAIA could provide a precision of 1 part in 10^6 or better in the determination of this parameter. It is interesting to note that this is the level of accuracy needed to test theories that claim to provide a route to quantization of gravity (*e.g.* Damour and Nordtvedt 1993).

This clearly suggests that global astrometry with GAIA should be formulated in a fully general relativistic framework. Correction terms of the order $M/R \sim v^2/c^2$ had already been included in Hipparcos (then subtracted from the observations), but the expected increase of observational accuracy with GAIA would sensibly increase the order of higher approximations.

Recent attempts at modeling high-precision astrometric observations in a relativistic framework have been discussed by Soffel (1989, see also Epstein & Shapiro 1980), Brumberg (1991), and Klioner and Kopejkin (1992). Soffel has concentrated on a PPN formulation (which would allow a comparison among different theories of gravity[1]), while Brumberg uses a first post-Newtonian formulation of general relativity. The work of Klioner and Kopejkin, which aimed at the modelization of the $\sim 1 - \mu arcsec$ observations of the proposed astrometric satellite POINTS (Reasemberg et al. 1988), follows Brumberg's approach extending the analysis to the post-post-Newtonian effects of the spherical Sun, and to the corrections due to its oblateness and angular momentum.

The aim of the following pages is to present a non-perturbative general relativistic approach to global astrometry, confining our attention to the case of Schwarzschild space-time geometry.

The signature of the adopted metric is $+2$. Also, greek indices take the values 0,1,2, and 3, while latin indices the values 1,2, and 3.

2. A non-perturbative relativistic model

We develop a fully non-perturbative relativistic model based on the following assumptions:

i) the curvature of the space-time is only due to a spherical, non-rotating Sun. Therefore, one can use plain Schwarzschild metric in polar coordinates, i.e.,

$$\mathrm{d}s^2 = -(1 - r_g/r)\,\mathrm{d}t^2 + (1 - r_g/r)^{-1}\,\mathrm{d}r^2 + r^2\,\mathrm{d}\omega^2,$$

where the colatitude $\theta \in [0, \pi]$, the azimuth $\phi \in [0, 2\pi[$, and the 'polar distance' $r \in]0, \infty[$. Also, r_g, the Schwarzschild radius, equals $2 \times \frac{G\,M_\odot}{c^2}$, where $M_\odot$ is the solar mass, and $\mathrm{d}\omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2$.
The choice of the $\{r, \theta, \phi\}$ coordinates presupposes that of three coordinate axes. These can be defined, for example, in such a way as to satisfy the requirements of the new ICRF frame introduced by the IAU (McCarthy 1996, and references therein).

ii) the observer is a GAIA-like triple interferometer (Lindegren & Perryman 1996) placed in a spatially circular orbit around the Sun; the radius of this orbit is an adjustable parameter that, in our simulation, has the value of the mean radius of the Earth orbit $R_\oplus$ (see next section).

The first hypothesis also implies that we do not take into account the effects of the other bodies of the solar system and of the quadrupole component of the gravitational field of the Sun. The second hypothesis fixes the observer's trajectory to an equatorial spatially circular geodesic whose 4-velocity is

$$u^\alpha = \mathrm{e}^\psi(\delta^\alpha_0 + \Omega\delta^\alpha_\phi). \tag{1}$$

Here, δ^i_j is the *Kronecker* delta (=1 for $i = j$, and =0 elsewhere), and

$$\mathrm{e}^\psi = \left(1 - \frac{2M}{r} - \Omega^2 r^2\right)^{-1/2} \tag{2}$$

is the normalization factor;

$$\Omega = \pm\left(\frac{M}{r^3}\right)^{1/2} \tag{3}$$

is the keplerian angular velocity in geometrized units[2], m^{-1}, for an observer at infinity. The choice of such an observer implies that every photon which reaches it must admit a colatitude $\theta_o = \pi/2$.

It should be stressed here that, although the origin of the Schwarzschild coordinate system which we adopt does not coincide with the IAU recommended framework (McCarthy 1996), it is not in contrast with the latter. For, our model only considers the gravitational field of the Sun, therefore its center is also the barycenter of the solar system.

A typical observable in astrometry is the angle between two stars, and space-borne global astrometry, in particular, utilizes angles several tens of degrees wide. The cosine of such an angle is given by the following formula (see, e.g., Brumberg 1991)

$$\cos\psi_{12} = \frac{h_{\alpha\beta}k_1^\alpha k_2^\beta}{(h_{\iota\pi}k_1^\iota k_1^\pi)^{1/2}(h_{\rho\sigma}k_2^\rho k_2^\sigma)^{1/2}}, \tag{4}$$

where k_1^α and k_2^β are the components of the tangents to the null geodesics of the photons emitted by the two stars respectively, and $h_{\alpha\beta} = g_{\alpha\beta} + u_\alpha u_\beta$ is a tensor operator which projects in

[1] Only the γ and β parameters of the Eddington-Robertson formulation of the PPN theory are considered but, as stated in Will (1974, 1981), only fully conservative theories (those with all but γ and β parameters put to zero), appear to be surviving.

[2] Geometrized units are such that the velocity of light and the newtonian constant of gravitation are both equal to unity. In these units the mass of the Sun $M \equiv \frac{GM_\odot}{c^2} \simeq 1,476.64$ m.

the rest frame of the given observer[3]. The components of a null geodesic are (see, e.g., de Felice and Clarke 1990),

$$k \equiv \begin{cases} k^0 = \left(1-\dfrac{2M}{r}\right)^{-1} E \\ k^r = \pm\left[E^2 - \dfrac{L^2}{r^2}\left(1-\dfrac{2M}{r}\right)\right]^{1/2} \\ k^\theta = \pm\dfrac{1}{r^2}\left[L^2 - \dfrac{l^2}{\sin^2\theta}\right]^{1/2} \\ k^\phi = \dfrac{1}{r^2\sin^2\theta} l, \end{cases} \tag{5}$$

in which E is the energy in units of c^2, and L and l are respectively the total and the azimuthal angular momenta of the photon in units of length. From Eqs. (1) and (5), Eq. (4) leads to

$$\cos\psi_{12} = \frac{\left(1-\dfrac{3M}{r_o}\right) F(\lambda_1,\Lambda_1,\lambda_2,\Lambda_2)}{(1-\Omega\lambda_1)(1-\Omega\lambda_2)}, \tag{6}$$

where

$$\begin{aligned} F(\lambda_1,\Lambda_1,\lambda_2,\Lambda_2) = {} & \left[-\left(1-\frac{2M}{r_o}\right)^{-1} + \left(1-\frac{3M}{r_o}\right)^{-1}\right] \\ & -\mathrm{sgn}(\Omega)\left(\frac{M}{r_o^3}\right)^{1/2}\left(1-\frac{3M}{r_o}\right)^{-1}(\lambda_1+\lambda_2) \\ & +\mathrm{sgn}(k_1^r)\mathrm{sgn}(k_2^r)\left(1-\frac{2M}{r_o}\right)^{-1}\times \\ & \times\left[1-\frac{\Lambda_1^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2}\left[1-\frac{\Lambda_2^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2} \\ & +\mathrm{sgn}(k_1^\theta)\mathrm{sgn}(k_2^\theta)\frac{(\Lambda_1^2-\lambda_1^2)^{1/2}(\Lambda_2^2-\lambda_2^2)^{1/2}}{r_o^2} \\ & +\frac{1}{r_o^2}\left[1+\left(1-\frac{3M}{r_o}\right)^{-1}\frac{M}{r_o}\right]\lambda_1\lambda_2. \end{aligned} \tag{7}$$

and Λ, λ are the ratios L/E and l/E respectively. In these formulas r_o is the distance of the observer from the Sun and Ω, as already stated, the angular velocity of the observer.

The observation Eq. (6) does not contain explicitly the relevant astrometric unknowns, i.e., the coordinates of the stars; they are hidden in the constants of motion. Integration of Eqs. (5) from the points of origin at the stars (r_i,θ_i,ϕ_i) to the endpoint $(r_o,\frac{\pi}{2},\phi_o)$ at the observer, yields the observation equation in explicit form. It is in general

$$\frac{\mathrm{d}r}{\dot r} = \frac{\mathrm{d}\theta}{\dot\theta}, \qquad \frac{\mathrm{d}\theta}{\dot\theta} = \frac{\mathrm{d}\phi}{\dot\phi}. \tag{8}$$

If we put $k^r = \dot r$, $k^\theta = \dot\theta$ and $k^\phi = \dot\phi$, and assume that all stars have distances $r_i = \infty$ (as in a *static* simulation, i.e., no parallax and proper motion effects are considered), then

$$\int_\infty^{r_o} \frac{\mathrm{sgn}(k^r)\mathrm{d}r}{r^2\sqrt{1-\dfrac{\Lambda^2}{r^2}\left(1-\dfrac{2M}{r}\right)}} = \int_{\theta_i}^{\pi/2} \frac{\mathrm{sgn}(k^\theta)\sin\theta\,\mathrm{d}\theta}{\sqrt{\Lambda^2\sin^2\theta-\lambda^2}} \tag{9}$$

$$\int_{\theta_i}^{\pi/2} \frac{\mathrm{sgn}(k^\theta)\lambda\,\mathrm{d}\theta}{\sin\theta\sqrt{\Lambda^2\sin^2\theta-\lambda^2}} = \int_{\phi_i}^{\phi_o}\mathrm{d}\phi. \tag{10}$$

Integration of the left-hand side of Eq. (9) is in general not possible in terms of elementary functions (it is an elliptic integral).If we express the constant of motion Λ in terms of the minimum distance from the Sun (r_c) reachable by the photon, namely where $k^r = 0$, we obtain

$$\Lambda^2 = r_c^2\left(1-\frac{2M}{r_c}\right)^{-1}, \tag{11}$$

then we can distinguish between the cases $r_c > R_\odot$ and $r_c \le R_\odot$. When $r_c > R_\odot$ the relations between the constants of motion and the Schwarzschild coordinates, as a result of the integration of Eqs. (9) and (10), have the general form

$$\left.\begin{aligned} f(r_c) &= g(\theta_i,\phi_i;\phi_o) \\ \frac{\lambda^2}{\Lambda^2} &= h(\theta_i,\phi_i;\phi_o) \\ \Lambda &= \Lambda(r_c) \end{aligned}\right\} \tag{12}$$

On the other hand, if $r_c \le R_\odot$, the previous relations become

$$\left.\begin{aligned} f(\Lambda) &= g(\theta_i,\phi_i;\phi_o) \\ \frac{\lambda^2}{\Lambda^2} &= h(\theta_i,\phi_i;\phi_o) \end{aligned}\right\}. \tag{13}$$

All the quantities in these expressions are intended in geometrized units. The explicit form of Eqs. (12) and (13) is given in Appendix A.

The lengthy formulas in the Appendices are indeed essential for a clear and correct treatment of the parameters involved in the data reduction process. Besides target coordinates, several other parameters will have to be considered in a more realistic configuration. For example, the orbital velocity of the satellite should be known to 1 cm/sec to allow for proper treatment of the velocity aberration at the 10 μarcsec level; also, the same orbit of the satellite will not be circular. Since general relativistic effects are likely to induce changes in the relevant parameters of magnitude similar to those caused by changes in the orbital setting, it is in our opinion very important to provide fully general relativistic expressions, even for the simple Schwarzschild model, for reference and comparison with more complex simulations.

[3] $g_{\alpha\beta}$ is the metric tensor corresponding to Schwarzschild's solution of the field equations

2.1. Validation of the formulas

The simulation crucially depends on the relations between the constants of motion and the Schwarzschild coordinates of the stars, therefore it is important to check whether they carry the information we expect. Therefore, we tested them for the effects of aberration and gravitational light-bending.

For this validation, we decided to calculate the direction of incoming photons, instead of the angle between two directions. First, we need to define a reference frame for the given observer, i.e., a *tetrad* $\{\lambda_{\hat{\alpha}}\}$ compatible with Eq. (1). We adopted a phase-locked tetrad with the θ-axis orthogonal to the orbital plane, the r-axis pointing along the Sun-observer direction, and the third (ϕ) axis tangent to the geodesic of the observer. The components of this tetrad are (de Felice and Usseglio-Tomasset 1993)

$$\left.\begin{aligned}
\lambda_{\hat{0}} &= \mathrm{e}^{\psi}(\partial_0 + \Omega\partial_\phi)\\
\lambda_{\hat{r}} &= \left(1-\frac{2M}{r}\right)^{1/2}\partial_r\\
\lambda_{\hat{\theta}} &= \frac{1}{r}\partial_\theta\\
\lambda_{\hat{\phi}} &= \frac{\mathrm{e}^{\psi}\Omega r\sin\theta}{\left(1-\frac{2M}{r}\right)^{1/2}}\partial_0 + \frac{\mathrm{e}^{\psi}\left(1-\frac{2M}{r}\right)^{1/2}}{r\sin\theta}\partial_\phi
\end{aligned}\right\} \qquad (14)$$

Then, application of Eqs. (5) through (14) provide the direction cosines of any photon reaching the observer as

$$\cos\Theta_{(\hat{a},k)} = \frac{h_{\rho\sigma}\lambda^{\rho}_{\hat{a}}k^{\sigma}}{(h_{\iota\pi}k^{\iota}k^{\pi})^{1/2}}, \qquad (15)$$

where $\hat{a}$ refers to one of the spatial unit 4-vectors in Eq. (14).

We considered selected directions of given ϕ and θ increasingly near the line joining the Sun to the observer (with the closest approach roughly corresponding to photons grazing the solar limb). These same directions are observed from two different frames, one situated diametrically opposite to the other relative to the Sun. Direction cosines are then calculated for both observers and the results subtracted. By comparing measurements from points 180° out of phase on the plane of the observer's geodesic, we maximize the relative displacements due to the relativistic effects.

It must be emphasized here that the comparison of measurements accomplished by two different observers requires that the results be reduced to a homogeneous, Euclidean-like, coordinate system, which we can refer to as the ecliptic system. Therefore, after calculating coordinates for a given observer at ϕ_o, these are transformed, in the euclidean sense, by applying a rotation of $\phi_o - \phi_{ori}$ around the axis perpendicular to the orbital plane; here ϕ_{ori} is the azimuthal reference direction, which we can set to zero. A static Schwarzschild sphere allows a further simplification, i.e., that of taking the Sun as the origin of the local reference frame. Such non-relativistic comparison procedure is however consistent with the relativistic calculations as we reduced the problem to that of comparing measurements taken by two observers whose two origins coincide and that are both inertial at any given instant. The difference between the two reference systems is in their different velocity and, therefore, the two are related by a Lorentz transformation. This is automatically taken care of by the use of a suitable tetrad like that in Eq. (14).

Table 1. Validation results for $\theta_i = \pi/2$. Here, aberration and gravitational light bending both affect directions on the orbital plane only. Note that the effect of aberration, obtained by subtraction of the values in third column from those in second, has the expected value of $-40''.985$. For reference, we recall that the Sun's apparent radius is $\sim 0.^{\circ}27$ for an observer on Earth.

ϕ_i (°)	Δ (")	Δ_0 (")
0.2769	−39.303	1.682
0.2707	−39.264	1.721
0.2667	−39.238	1.747
359.7333	−42.731	−1.747
359.7293	−42.705	−1.721
359.7231	−42.667	−1.682

Table 2. Validation results for $\phi_i = 0$. Here $\Delta\lambda$ and $\Delta\beta$ represent the resulting effects in ("longitude") and off ("latitude") the orbital plane, respectively. Also, the contribution due to aberration is only in longitude, while gravitational light bending splits in two components. As expected, aberration effects vanish for $\Omega = 0$.

	Δ			Δ_0		
θ_i (°)	$\Delta\lambda$ (")	$\Delta\beta$ (")	$\lvert\Delta\rvert$ (")	$\Delta\lambda$ (")	$\Delta\beta$ (")	$\lvert\Delta\rvert$ (")
89.723	−40.985	1.682	41.020	0	1.682	1.682
89.729	−40.985	1.721	41.022	0	1.721	1.721
89.733	−40.985	1.747	41.023	0	1.747	1.747
90.267	−40.985	−1.747	41.023	0	−1.747	1.747
90.271	−40.985	−1.721	41.022	0	−1.721	1.721
90.277	−40.985	−1.682	41.020	0	−1.682	1.682

Tests were run for the cases $\Omega = 0$ and $\Omega = \sqrt{M/r_o^3}$, so to separate the aberration effects from those due to light deflection. The results are reported in Tables 1 and 2, where Δ refers to total "deflection" (i.e., to the combination of aberration and gravitational light deflection), and Δ_0 is the deflection derived for "static" observers ($\Omega = 0$), and it is therefore the deflection due solely to light bending.

3. End-to-end simulation

3.1. Observation equations

The first step in the development of an end-to-end simulation is to write Eq. (4) in an explicit form relating the measurements to the unknown quantities. In the case of a static simulation, the measured arc between a generic star pair, ψ_{12}, is related to the Schwarzschild coordinates of the two stars involved by the general expression $cos\psi_{12} = \eta(\theta_1, \phi_1, \theta_2, \phi_2)$.

We then need to adopt a statistical tool to *estimate* the Schwarzschild coordinates of the stars from the measured arcs.

For this, we decided to utilize linear least-squares which, as observation equations, take the linearized version of $cos\psi_{12}$, i.e.

$$-\sin\psi_{12}\,\delta\psi_{12} = A_1\,\delta\theta_1 + B_1\,\delta\phi_1 + A_2\,\delta\theta_2 + B_2\,\delta\phi_2, \quad (16)$$

where $\delta\psi_{12}$ is the difference between the observed, highly accurate, angle and some a priori knowledge of it derived from approximate values of the Schwarzschild coordinates and, therefore, of the constants of motion of the associated photons. Correspondingly, the small differences $\delta\phi$ and $\delta\theta$ are the corrections to the approximate (catalog) values one wishes to deduce from the least-squares estimation procedure (see next section). The coefficients $A_1 = \frac{\partial\eta}{\partial\theta_1}$, $B_1 = \frac{\partial\eta}{\partial\phi_1}$, *etc.* can be derived analitically [4] and are calculated, likewise $sin\psi_{12}$, from the catalog values of the coordinates.

Differentiation of Eq. (4) with respect to λ_1, λ_2, Λ_1, and Λ_2 leads to

$$-\sin\psi_{12}\,\delta\psi_{12} = a_1\delta\lambda_1 + b_1\delta\Lambda_1 + a_2\delta\lambda_2 + b_2\delta\Lambda_2, \quad (17)$$

where

$$a_1 = \left(1-\frac{3M}{r_o}\right)(1-\Omega\lambda_1)^{-1}(1-\Omega\lambda_2)^{-1}\times \times\left[\frac{\partial F}{\partial\lambda_1} + F\,\Omega\,(1-c\Omega\lambda_1)^{-1}\right], \quad (18)$$

$$b_1 = \left(1-\frac{3M}{r_o}\right)\left[(1-\Omega\lambda_1)(1-\Omega\lambda_2)\right]^{-1}\frac{\partial F}{\partial\Lambda_1}, \quad (19)$$

$$a_2 = \left(1-\frac{3M}{r_o}\right)(1-\Omega\lambda_1)^{-1}(1-\Omega\lambda_2)^{-1}\times \times\left[\frac{\partial F}{\partial\lambda_2} + F\,\Omega\,(1-c\Omega\lambda_2)^{-1}\right], \quad (20)$$

$$b_2 = \left(1-\frac{3M}{r_o}\right)\left[(1-\Omega\lambda_1)(1-\Omega\lambda_2)\right]^{-1}\frac{\partial F}{\partial\Lambda_2}, \quad (21)$$

[the derivatives of $F(\lambda_1,\Lambda_1,\lambda_2,\Lambda_2)$ are given in Appendix B.] We can now express $\delta\lambda$ and $\delta\Lambda$ in terms of $\delta\theta$ and $\delta\phi$ using, for $r_c > R_\odot$, Eq. (12); it results

$$\delta\Lambda = \left[\left(\frac{\partial f}{\partial r_c}\right)^{-1}\frac{\partial\Lambda}{\partial r_c}\frac{\partial g}{\partial\theta_i}\right]\delta\theta_i + \left[\left(\frac{\partial f}{\partial r_c}\right)^{-1}\frac{\partial\Lambda}{\partial r_c}\frac{\partial g}{\partial\phi_i}\right]\delta\phi_i = \alpha_\Lambda\delta\theta_i + \beta_\Lambda\delta\phi_i, \quad (22)$$

$$\delta\lambda = \left[\frac{\Lambda^2}{2\lambda}\frac{\partial h}{\partial\theta_i} + \frac{\lambda}{\Lambda}\left(\frac{\partial f}{\partial r_c}\right)^{-1}\frac{\partial\Lambda}{\partial r_c}\frac{\partial g}{\partial\theta_i}\right]\delta\theta_i + \left[\frac{\Lambda^2}{2\lambda}\frac{\partial h}{\partial\phi_i} + \frac{\lambda}{\Lambda}\left(\frac{\partial f}{\partial r_c}\right)^{-1}\frac{\partial\Lambda}{\partial r_c}\frac{\partial g}{\partial\phi_i}\right]\delta\phi_i = \alpha_\lambda\delta\theta_i + \beta_\lambda\delta\phi_i. \quad (23)$$

The derivatives inside the square brackets on the right-hand-side of Eqs. (22) and (23) are provided in Appendix B. Finally, substitution of these expressions in Eq. (17) yields Eq. (16) in explicit form, as it is

$$\begin{aligned} A_1 &= a_1\alpha_{\lambda_1} + b_1\alpha_{\Lambda_1} \\ B_1 &= a_1\beta_{\lambda_1} + b_1\beta_{\Lambda_1} \\ A_2 &= a_2\alpha_{\lambda_2} + b_2\alpha_{\Lambda_2} \\ B_2 &= a_2\beta_{\lambda_2} + b_2\beta_{\Lambda_2}. \end{aligned}$$

The derivation of the case with $r_c < R_\odot$ follows a similar scheme with Eqs. (13) replacing Eqs. (12).

3.2. *Simulating relativistic observations*

The simulation code is an adaptation of that used in Lattanzi et al. (1991; see also Galligani et al. 1989) for the assessment of the astrometric accuracy of the sphere reconstruction in the Hipparcos mission; the most relevant changes concerned the calculations of the relativistic quantities, both observations and coefficients of the linearized observation equations [Eqs. (18) through (23)].

We generate lists of 2000 stars randomly distributed on the (not directly observable) *Schwarzschild sky*, i.e., their positions on the Schwarzschild sphere are uniquely specified using azimuth (ϕ) and colatitude (θ) only. Perspective effects (parallax) and effects due to relative motions throughout the Galaxy (proper motions) were not considered in this first attempt at modeling a relativistic sphere (*static simulation*). Also, stars were all assigned the same survey limiting magnitude V=16. The star positions derived from the random number generators are named *true* positions. From these, approximate locations, called *catalog positions*, are generated by perturbing the true values of 2 mas. Such a *catalog* simulates our best knowledge of the star locations before the satellite measurements. Catalog values are used to compute the coefficients of the linearized equations and the approximate (catalog) values of the angles among star pairs observed by the satellite needed to calculate $sin\psi$ and $\delta\psi$ on the left-hand side of Eq. (16).

The satellite is made to sweep the sky according to a given scanning law quite similar to that successfully implemented on Hipparcos; the spin axis precesses around the Sun at a rate of $\sim$ 6.4 $revolution/year$ and with a constant angle of 43°. Table 3 lists some of the parameters of the reference mission used in this simulation.

Stars that at any given time are 'seen' within a strip 1.6° wide along the great circle being scanned are considered observable and are given further consideration; a great circle (one revolution) is completed in about 2.1 $hours$.

From the true coordinates of the stars in the strip we derive the constants of motion of the emitted photons and from these, utilizing Eq. (4), the *true* angles between each possible

[4] We stress that such linearization of Eq. (4) is not an approximation imposed on our derivations; it is just a necessity of the least-squares estimation procedure utilized

Table 3. Most relevant parameters of the GAIA simulation. Notice that *single-measurement error* refers to the error of one measured angle between two generic stars as observed during one revolution around the spin axis.

parameter	*numerical value*	*comment*
orbital radius (r_o)	1.496×10^{11} m	same as Earth's orbital radius ($R_\oplus$)
precession angle	43°	same as solar aspect angle
satellite spin period	128 min	
angles between interferometers	54° (I1-I2), $78^\circ.5$ (I2-I3)	
coherent FOV of each interferometer	1.6°	
mission duration	1 year	static simulation
No. of simulated stars	2000	
No. of unknowns	4000	only Schwarzschild coordinates estimated
Catalog error	2 mas	error of initial values in linearized observation eq.s
Single-measurement error	100 $\mu arcsec$	as expected for V$\sim$ 16 mag stars

distinct pair that can be formed with the visible stars. Of these angles, only those satisfying the relations $\psi_{ij} = 54^\circ \pm 1.6^\circ$ or $\psi_{ij} = 78.5^\circ \pm 1.6^\circ$ are retained as *observed*. These observability conditions are the result of current expectations for the GAIA payload. The optical configuration of the present baseline mission (Lindegren and Perryman 1996) shows three stacked identical interferometers (I1, I2, and I3) each with a large Field-of-View (FOV) and pointing toward widely separated directions at angles $\sim 60^\circ$ (I1-I2) and $\sim 80^\circ$ (I2-I3) (Table 3).

Finally, for this investigation it is sufficient to say that the GAIA detection system is capable of measuring those wide angles with an error of $\sim 100\ \mu arcsec$ when *both* stars are of magnitude V=16. This is the error added to the true angles (which passed the observability conditions) for the generation of the *observed* angles to be used in the observation equations.

The simulation of one year of uninterrupted observations yielded 78050 observed arcs which were then used for the estimation of 4000 (two coordinates for each of the 2000 simulated stars) stellar unknowns. Each star has, on average, $\sim$39 connections; however, the scanning law favours stars at colatidudes near 43° and 137° with more than twice as many observations as the stars near the "equator" ($\theta \sim 90^\circ$).

4. Results and discussion

The reconstruction of the Schwarzschild sphere (a concept analogous to the unit sphere in Hipparcos) from measurements of arcs among stars proceeds by setting up the appropriate system of linear observation equations to be solved with a least-squares method. Each row of this system is a relation like Eq. (16) linear in the four unknowns $\delta\phi_i$, $\delta\theta_i$ and $\delta\phi_j$, $\delta\theta_j$ of stars i and j. There are $\sim$ 78,000 of such equations available (see previous section) for 4,000 unknowns; therefore, the system is, as expected, overdetermined with quite a comfortable Q (equations to unknowns) ratio of $\simeq$20 [5].

In a Hipparcos-like formulation, where the equations for the sphere reconstruction would be written in Euclidean (flat) space, angular distances are invariant for any rigid rotation of the coordinate frame. This naturally leads to a rectangular matrix (or *design matrix*, how it is often called) with a rank deficiency of 3. In practice, this means that measuring arcs among stars can "*close*" the sphere but cannot *observe*, i.e. fix, the reference frame.

Something similar can be expected in our relativistic formulation although invariance of the angular measurements is defined differently; because of the role of the *observer*, the transformations which leave the angular measurements unchanged are those which do not alter the relative orientation of the observer and the stars. We did not investigate the issue of rank deficiency of the relativistic equations; however, the occurence of a singular design matrix does not prevent us from obtaining a unique solution satisfying the least-squares system of equations, as discussed below.

The numerical code used for the solution of this system of equations is, a part for minor changes to update it to faster machines, the one described in Galligani et al. (1989) and utilized on simulated as well as real Hipparcos data. The code implements an iterative algorithm which uses the conjugate-gradient method to solve, in the least-squares sense, large and sparse overdetermined systems like ours (the number of non zero elements represents $\simeq$ 0.1% of the total). These methods can take a long time before reaching a solution as the number of iterations goes linearly with the number of unknowns. However, standard techniques exist that can sharply improve upon rate of convergence. In our particular application, the use of a simple rescaling of the columns of the associated rectangular matrix always provided convergence in less than 350 iterations.

This iterative method has a "built-in" way of dealing with rank-deficient matrices. In the presence of rank deficiency the solution returned by the program is that of *minimum norm*, i.e., among the infinite n-dimensional vectors (n is the number of

[5] In a more realistic scenario, relevant subsystems on the satellite (interferometers, detectors, etc.) will deviate from nominal during the measuremets and with the operational lifetime (instrumental degradation), and a certain number of extra unknown parameters will have to be introduced in the expression for the observation equation to model such effects. Then, part of the consistent Q ratio is profitably exploited to estimate the extra degrees of freedom introduced by the real system.

Table 4. Mean (μ) and standard deviation (σ) of the true errors of the least-squares adjustment.

	μ ($\mu arcsec$)	σ ($\mu arcsec$)	*No. of stars*
$\delta\hat{\phi}$	$< 1\ \mu arcsec$	36	1971 (98.6%)
$\delta\hat{\theta}$	$< 1\ \mu arcsec$	21	1986 (99.3%)

columns of the desing matrix) satisfying the system of m equations, that with the smallest modulus is chosen.

Figs. 1 and 2 and Table 4 show the results of a typical sphere adjustment run on a set of simulated data. The individual differences shown in the plots and averaged in Table 4 are the *true* errors defined as $\delta\hat{\phi}_i = \phi_i(true) - \phi_i(adj)$ (and similarly for $\delta\hat{\theta}_i$), where $\phi_i(true)$ is the true Schwarzschild azimuth of star i and $\phi_i(adj)$ is the value derived from the results of the least-squares adjustment; it is

$$\phi_i(adj) = \delta\tilde{\phi}_i + \tilde{\phi}_i,$$

where $\delta\tilde{\phi}_i$ is the least-sqares estimate for star i and $\tilde{\phi}_i$ the corresponding initial, *catalog*, value used in the calculation of the coefficients of the linearized equations.

The third column of Table 4 shows that there were few outliers that were removed before calculating the average errors. These are typically stars with a critically low number of connections.

In the absence of correlations among observations one would anticipate average errors on $\phi(adj)$ and $\theta(adj)$ of the order of $\frac{\sqrt{2}\ 100}{\sqrt{40}} \simeq 23\ \mu arcsec$, where the factor 100 $\mu arcsec$ is the single-measurement error and the value 40 is the typical average number of connections per stars (see previous section). On the other hand, because of the geometry of the arcs observed by GAIA, the ϕ error is expected to be larger than the θ error. From the values in Table 4 we derive the empirical value $\sigma_\phi/\sigma_\theta \simeq 1.7$. This is in quite good agreement with the theoretical prediction $\sigma_\phi/\sigma_\theta \simeq 1.6$ derived by Betti and Sanso' (1985) in similar circumstances but in the context of the Hipparcos mission.

The very small values obtained for μ_ϕ and μ_θ suggest that the reconstructed Schwarzschild sphere defines, on average, a set of spherical coordinates very close to the original, true, set. In a Newtonian framework, we would have said that the reconstructed sphere defines a coordinate system closely related to that of the true positions, thus pointing to the absence of overall residual distortions in the reconstructed sphere.

5. Summary and conclusions

We have shown that, under simplifying assumptions, it is possible to tackle the problem of global astrometric observations in a fully general relativistic framework. Despite the complications of the model, the results of the simulation have validated the expectations that angular observations accurate to 100-μarcsec would result in positions good to 20 μarcsec, an improvement of approximately a factor of 100 over Hipparcos!

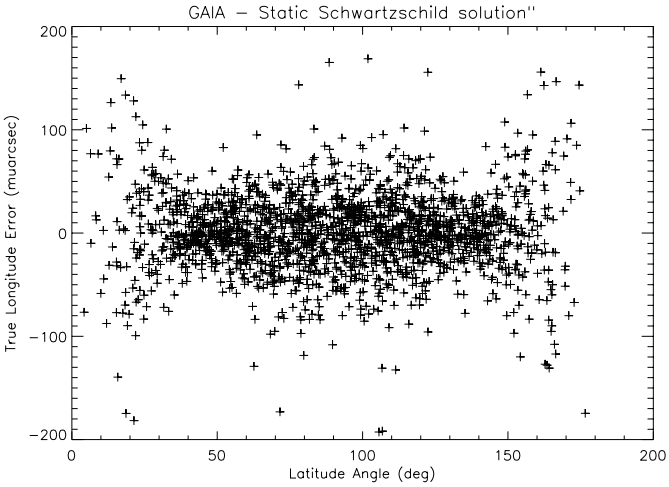

Fig. 1. *true* ϕ errors versus colatitude (θ). Units on the ordinate axis are $\mu arcsec$. Notice the errors become smaller near the colatidudes 43° and 137°. This is where the number of connections (observations) per stars peaks as a result of the particular sky coverage startegy simulated.

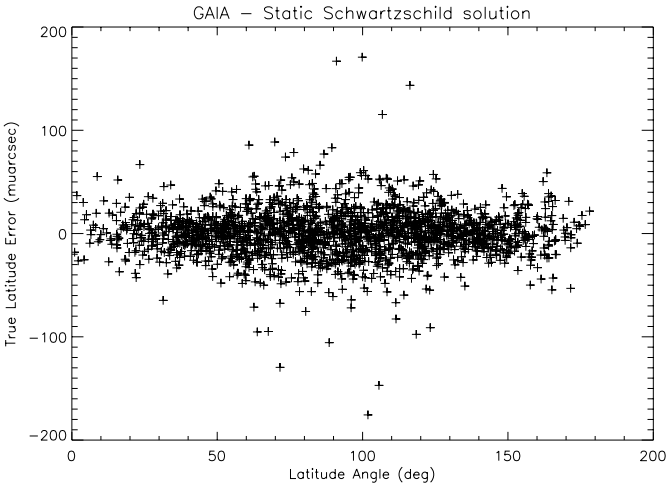

Fig. 2. *true* θ errors versus θ. As for the ϕ errors, the results appear better at $\theta \sim 43°$ and $\sim 137°$.

We anticipate that, if the spherical non-rotating Sun is retained as the only cause for the space-time curvature, fully dynamical equations (i.e., including perspective effects -parallax- and intrinsic motions of the stars -$d\phi_i/dt$ and $d\theta_i/dt$) can be derived in a way similar to that described here. This is the subject of a forthcoming paper.

Acknowledgements. We wish to thank the referee, Dr. S.A. Klioner, and the editor, Prof. J. Lequeux, for several comments and suggestions which improved the manuscript. Work partially supported by the Italian Space Agency (ASI), the italian Ministero della Ricerca Scientifica e Tecnologica (MURST), and by the Gruppo Nazionale per la Fisica Matematica del CNR.

Appendix A: relations among Schwarzschild coordinates and constants of motion

The systems of Eqs. (12) and (13), in the last part of Sect. 2, use the following functions:

$$f(r_c) = \mathrm{flag}_2 \left(1-\frac{2M}{r_c}\right)^{-1/2} \left[\left(1-\frac{M}{r_c}\right)\arcsin\frac{r_c}{r_o} - \frac{M}{r_c}\left(\frac{r_c}{r_o}+2\right)\sqrt{\frac{r_o-r_c}{r_o+r_c}} + \mathrm{flag}_2\frac{2M}{r_c} - \mathrm{flag}_1\pi\left(1-\frac{M}{r_c}\right)\right], \quad \text{(A1)}$$

$$f(\Lambda) = \left(\frac{1}{6r_o^3}-\frac{M}{4r_o^4}\right)\Lambda^3 + \frac{\Lambda}{r_o}, \quad \text{(A2)}$$

$$g(\theta_i,\phi_i;\phi_o) = \mathrm{flag}_1\pi + \mathrm{flag}_2 \arcsin\sqrt{\frac{2-\sin^2\theta_i\{\cos[2(\phi_i-\phi_o)]+1\}}{2}}, \quad \text{(A3)}$$

$$h(\theta_i,\phi_i;\phi_o) = \frac{\sin^2\theta_i\{\cos[2(\phi_i-\phi_o)]-1\}}{\sin^2\theta_i\cos[2(\phi_i-\phi_o)]+\sin^2\theta_i-2}, \quad \text{(A4)}$$

$$\Lambda^2(r_c) = r_c^2\left(1-\frac{2M}{r_c}\right)^{-1}. \quad \text{(A5)}$$

The parameters flag_1 and flag_2 are defined as

$$\mathrm{flag}_1 = \begin{cases} 0 & \text{when } -\frac{\pi}{2} < \tilde{\phi} < \frac{\pi}{2} \\ 1 & \text{when } \tilde{\phi} < -\frac{\pi}{2} \text{ o } \frac{\pi}{2} < \tilde{\phi} \end{cases}$$

and

$$\mathrm{flag}_2 = \begin{cases} +1 & \text{when } -\frac{\pi}{2} < \tilde{\phi} < \frac{\pi}{2} \\ -1 & \text{when } \tilde{\phi} < -\frac{\pi}{2} \text{ o } \frac{\pi}{2} < \tilde{\phi} \end{cases},$$

where

$$\tilde{\phi} = \begin{cases} \phi_i-\phi_o & \text{when } \phi_i-\phi_o < \pi \\ (\phi_i-\phi_o)-2\pi & \text{when } \phi_i-\phi_o > \pi \\ \phi_i-\phi_o & \text{when } \phi_i-\phi_o < 0 \\ (\phi_i-\phi_o)+2\pi & \text{when } \phi_i-\phi_o < -\pi \end{cases}.$$

Appendix B: derivatives for the coefficients of the observation equations

The partial derivatives of $F(\lambda_1,\Lambda_1,\lambda_2,\Lambda_2)$ with respect to each constant of motion used in Eqs. (18) through (21), are

$$\frac{\partial F}{\partial\lambda_1} = -\Omega\left(1-\frac{3M}{r_o}\right)^{-1} - \mathrm{sgn}(k_1^\theta)\mathrm{sgn}(k_2^\theta)\frac{\lambda_1}{r_o^2}\frac{(\Lambda_2-\lambda_2)^{1/2}}{(\Lambda_1-\lambda_1)^{1/2}} + \frac{\lambda_2}{r_o^2}\left[1+\left(1-\frac{3M}{r_o}\right)^{-1}\frac{M}{r_o}\right], \quad \text{(B1)}$$

$$\frac{\partial F}{\partial\Lambda_1} = -\mathrm{sgn}(k_1^r)\mathrm{sgn}(k_2^r)\frac{\left[1-\frac{\Lambda_2^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2}}{\left[1-\frac{\Lambda_1^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2}}\frac{\Lambda_1}{r_o^2} + \mathrm{sgn}(k_1^\theta)\mathrm{sgn}(k_2^\theta)\frac{\Lambda_1}{r_o^2}\frac{(\Lambda_2-\lambda_2)^{1/2}}{(\Lambda_1-\lambda_1)^{1/2}}, \quad \text{(B2)}$$

$$\frac{\partial F}{\partial\lambda_2} = -\Omega\left(1-\frac{3M}{r_o}\right)^{-1} - \mathrm{sgn}(k_1^\theta)\mathrm{sgn}(k_2^\theta)\frac{\lambda_2}{r_o^2}\frac{(\Lambda_1-\lambda_1)^{1/2}}{(\Lambda_2-\lambda_2)^{1/2}} + \frac{\lambda_1}{r_o^2}\left[1+\left(1-\frac{3M}{r_o}\right)^{-1}\frac{M}{r_o}\right], \quad \text{(B3)}$$

$$\frac{\partial F}{\partial\Lambda_2} = -\mathrm{sgn}(k_1^r)\mathrm{sgn}(k_2^r)\frac{\left[1-\frac{\Lambda_1^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2}}{\left[1-\frac{\Lambda_2^2}{r_o^2}\left(1-\frac{2M}{r_o}\right)\right]^{1/2}}\frac{\Lambda_2}{r_o^2} + \mathrm{sgn}(k_1^\theta)\mathrm{sgn}(k_2^\theta)\frac{\Lambda_2}{r_o^2}\frac{(\Lambda_1-\lambda_1)^{1/2}}{(\Lambda_2-\lambda_2)^{1/2}}. \quad \text{(B4)}$$

Additional calculations provide us with the derivatives that appear in Eqs. (22) and (23):

$$\frac{\partial\Lambda}{\partial r_c} = \left(1-\frac{3M}{r_c}\right)\left(1-\frac{2M}{r_c}\right)^{-3/2}, \quad \text{(B5)}$$

$$\frac{\partial f}{\partial r_c} = -\left(1-\frac{2M}{r_c}\right)^{-1}\times\frac{M}{r_c^2}f(r_c) + \mathrm{flag}_2\left(1-\frac{2M}{r_c}\right)^{-1/2}\times \left[\frac{M}{r_c^2}\arcsin\frac{r_c}{r_o} + \left(1-\frac{M}{r_c}\right)\left(1-\frac{r_c^2}{r_o^2}\right)^{-1/2}\frac{1}{r_o} + \frac{M}{r_c^2}\left(\frac{r_c}{r_o}+2\right)\sqrt{\frac{r_o-r_c}{r_o+r_c}} - \frac{M}{r_c}\frac{1}{r_o}\sqrt{\frac{r_o-r_c}{r_o+r_c}} + \frac{M}{r_c}\left(\frac{r_c}{r_o}+2\right)\frac{r_o}{(r_o+r_c)^2}\sqrt{\frac{r_o+r_c}{r_o-r_c}} - \mathrm{flag}_2\frac{2M}{r_c^2} - \mathrm{flag}_1\pi\frac{M}{r_c^2}\right], \quad \text{(B6)}$$

$$\frac{\partial g}{\partial\theta_i} = -\mathrm{flag}_2\sqrt{\frac{\cos[2(\phi_i-\phi_o)]+1}{2-\sin^2\theta_i\{\cos[2(\phi_i-\phi_o)]+1\}}}\times \times\cos\theta_i, \quad \text{(B7)}$$

$$\frac{\partial g}{\partial \phi_i} = \mathrm{flag}_2 \sqrt{\frac{\{\cos[2(\phi_i - \phi_o)] + 1\}^{-1}}{2 - \sin^2\theta_i \{\cos[2(\phi_i - \phi_o)] + 1\}}} \times$$
$$\times \sin\theta_i \sin[2(\phi_i - \phi_o)], \tag{B8}$$

$$\frac{\partial h}{\partial \theta_i} = -\frac{4 \sin\theta_i \cos\theta_i \{\cos[2(\phi_i - \phi_o)] - 1\}}{\{\sin^2\theta_i \cos[2(\phi_i - \phi_o)] + \sin^2\theta_i - 2\}^2}, \tag{B9}$$

$$\frac{\partial h}{\partial \phi_i} = \frac{4 \sin^2\theta_i \cos^2\theta_i \sin[2(\phi_i - \phi_o)]}{\{\sin^2\theta_i \cos[2(\phi_i - \phi_o)] + \sin^2\theta_i - 2\}^2}. \tag{B10}$$

References

Betti, B., and Sanso', F., 1985; The Second FAST Thinkshop, J. Kovalevsky (Ed.), Marseille

Brumberg, V. A., 1991; Essential Relativistic Celestial Mechanics, Adam Hilger

Capetti, A., Macchetto, F.D., and Lattanzi, M.G., 1997; ApJ, 476, L67

Casertano, S., Lattanzi, M.G., Perryman, M.A.C., & Spagna, A., 1996; Ap&SS, 241, 89

Damour, T., Nordtvedt, K., 1993; General Relativity as a Cosmological Attractor of Tensor-Scalar Theories; Phys. Rev. Letters, 70 (15), 2217-2219

de Felice, F., and Usseglio-Tomasset, S. 1993; Gen. Rel. Grav., 28 (2), 179-192

de Felice, F., Clarke, C.J.S., 1990; Relativity on curved manifolds, Cambridge University Press

Epstein, R., Shapiro, I., 1980; Phys. Rev., D22, 2947

Fakir, R., 1993; ApJ, 418, 202-207

Fakir, R., 1994; ApJ, 426, 74-78

Klioner, S. A., Kopejkin, S. M., 1992; AJ, 104, 897-914

Lattanzi, M.G., Bucciarelli, B., and Bernacca, P.L. 1991; ApJS, 73, 481

Lattanzi, M.G., et al., 1997; submitted to MNRAS

Lindegren, L., and Perryman, M.A.C., 1996; A&AS, 116, 579

Makarov, V.V., 1995; *Future Possibilities for Astrometry in Space*, ESA SP-379, 117

McCarthy, D.D., 1996; *IERS Technical Note 21*, Central Bureau of IERS, Observatoire de Paris.

Misner, C. W., Thorne, K. S., Wheeler, J. A., 1973; Gravitation; W. H. Freeman and Company

Perryman, M.A.C., and van Leeuwen, F. (Eds.), 1995; ESA SP-379

Smart, R.L., and Lattanzi, M.G., 1996; A&A, 314, 104

Soffel, M.H., 1989; Relativity in Astrometry, Celestial Mechanics and Geodesy, Springer-Verlag

Will, C. M., 1974; The Theoretical Tools of Experimental Gravitation; in: Experimental Gravitation: Proceedings of Course 56 of the International School of Physics "Enrico Fermi", ed. by B. Bertotti, Academic Press, New York

Will, C. M., 1981; Theory and Experiment in Gravitational Physics; Cambridge University Press