RELATIVISTIC EFFECTS IN HIPPARCOS DATA

B F Schutz
Department of Applied Mathematics & Astronomy
University College, Cardiff, UK

ABSTRACT

Relativistic effects in the HIPPARCOS data affect both the data analysis and the selection of programme stars. Light-deflection by the Sun will be measureable over a considerable fraction of the celestial sphere, which suggests that a very accurate test of general relativity is possible, and even that direct measurement of the metric is feasible. Selection of distant programme stars enhances the possibility of detecting a black hole by the anomalous deflection it produces in a star's apparent position.

Keywords:Relativistic effects,Light-deflection, General relativity, Black holes

1. INTRODUCTION

The gravitational deflection of light by the Sun was the first prediction of general relativity to be confirmed by a specially-designed observation, and its discovery made Einstein a popular as well as a scientific hero (Ref.1). The effect was in those days only barely measureable: 1."75 for stars near the limb of the Sun. The accuracy planned for the HIPPARCOS satellite is three orders of magnitude better, and the deflection will therefore have a systematic effect on the observations of stars over a considerable fraction of the sky. More than that, the wealth of accurate data the satellite should produce can be used for a variety of purposes from testing general relativity to searching for black holes. Relativistic considerations should be taken into account in two distinct areas, which I will address separately: the design of the software for data reduction, and the selection of programme stars for the observing list.

2. RELATIVISTIC EFFECTS IN THE DATA

It is important to accept that any attempt to 'correct' data for relativistic effects or to use the data to test theoretical predictions depends on a theoretical model of the relativistic effects, and that model has to be chosen from the start. I shall discuss three models -- general relativity, the 'parametrized post-Newtonian' framework, and a general metric framework -- and comment on the usefulness of adopting one or another of them. Much of what I shall say in this section results from very illuminating converstions with M. Cruise of the Mullard Space Sciences Laboratory.

2.1 Predictions of general relativity

The full angle by which a light ray is deflected on passing near a perfectly spherical Sun is

$$\delta = 4GM_{\odot}/c^2 r = 1''.7505\ R_{\odot}/r, \tag{1}$$

where r is the impact parameter of its original trajectory (Fig.1). A star observed near the limb

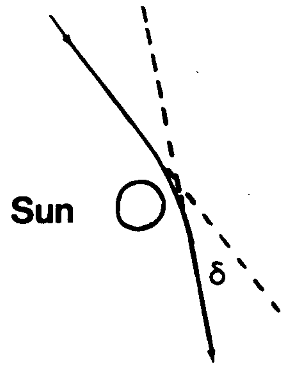

Figure 1. Full deflection defined by asymptotically straight incoming and outgoing paths.

of the Sun suffers the full deflection, but light observed at an angle of 90° from the Sun has undergone only half its deflection, and moreover has an impact parameter r=1 A.U., so that for it δ=4.1 milliarcsec. The general expression (Refs.2-5) for an observer located a distance d from the Sun, and its value for d=1 A.U., are

$$\delta = \frac{2GM_{\odot}}{c^2 d} \cot\frac{\alpha}{2} = 4.07195 \cot\frac{\alpha}{2} \text{ milliarcsec}, \tag{2}$$

where α is the angle between the Sun and the apparent position of the star (Fig.2). The accuracy with which α is known from ground-based observations is of course sufficient to calculate δ, so this correction could be applied rapidly to stellar positions, for example to check the self-consistency of the observations in a single scan.

The Sun is not spherical, of course, and its

Proceedings of an International Colloquium on The Scientific Aspects of the Hipparcos Mission, held at Strasbourg, France, 22-23 February 1982. (ESA SP-177).

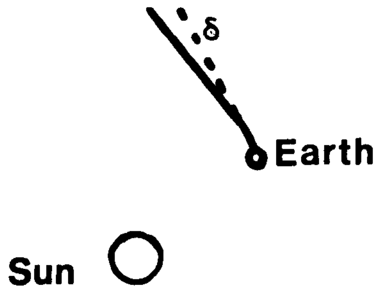

Figure 2. Partial deflection.

quadrupole moment and angular momentum can both in principle affect the deflection. Using the Dicke-Goldenberg (Ref.6) value for the quadrupole moment, which is certainly an upper limit, its effect at the limb of the Sun is 1.5×10^{-5} that of the spherical Sun. Moreover, it falls off as r^{-3}, so it can be neglected everywhere. The angular momentum of the Sun directly affects the deflection through the 'dragging of inertial frames': circular planetary orbits have slightly smaller periods in the co-rotating sense than in the counter-rotating sense, and a light ray is deflected less if it flies past in the co-rotating sense than otherwise. If the Sun has angular momentum $J_\odot$, then this effect is (Ref.7)

$$|\delta| = \frac{2GJ_\odot}{d^2c^3}(1-\cos\alpha)^{-1}$$

in the notation of Eq.2. If the Sun rotates roughly rigidly with its surface angular velocity, then this effect is only of order 10^{-6} arcsec at the limb. But it may be worthwhile incorporating a test for it in the data analysis, first because it is the only direct measure of $J_\odot$ we are likely to get in the near future, and second because it may be possible to dig deeper into the noise for this effect: unlike parallax and the main light-bending of the Sun, it produces an asymmetry of the star's apparent position about the Sun.

Other frame-dragging effects, such as that due to the movement of the Sun about the baycenter of the solar system, are much smaller. The motion of the spacecraft about the Earth induces two relativistic precessional effects (Ref.4). One, an analogue of the atomic Thomas precession, is of magnitude a few microarcsec per orbit for a spacecraft in geostationary orbit. The other comes from the Earth's dragging of inertial frames, and is two orders of magnitude smaller. Both will be masked by other external perturbations on the spacecraft.

The planets also deflect light, and while that due to the Earth is negligible, that due to Jupiter is about 17 milliarcsec at its limb. Since Jupiter is too bright to allow measurements right at the limb, and since the single-observation accuracy will be of the order of 10 milliarcsec, the effect may only be observable in a few stars. Interestingly, Jupiter's dragging-of-inertial-frames deflection is only a factor of 4 smaller than the Sun's (Ref.7).

2.2 Parametrized post-Newtonian (PPN) framework

General relativity may not be the correct theory of gravity, and in fact the 1960's and '70's saw the invention of a large number of competing theories, most of them predicting slightly different corrections to Newtonian gravity (called their 'post-Newtonian' effects) in the solar system. In order to systematize these theories and make a sensible framework for comparing them with the few observations one could anticipate making in the solar system, Thorne and Will (Ref.8) devised the PPN formalism. Beginning from rather general theoretical assumptions, such as the equivalence principle (which can of course also be tested), they showed that the predictions of most theories for observable effects in the solar system involved a few functions, such as the Newtonian potential,

$$U(\underset{\sim}{x},t) = G\int\rho(\underset{\sim}{x}',t)\,|\underset{\sim}{x}-\underset{\sim}{x}'|^{-1}d^3x',$$

or a similar potential whose source was the momentum density,

$$\underset{\sim}{V}(x,t) = G\int\rho(\underset{\sim}{x}',t)\underset{\sim}{v}(\underset{\sim}{x}',t)\,|\underset{\sim}{x}-\underset{\sim}{x}'|^{-1}d^3x',$$

or terms essentially quadratic in U, like the gravitational field caused by the mass-equivalent of the gravitational potential energy,

$$G\int\rho(\underset{\sim}{x}',t)U(\underset{\sim}{x}',t)\,|\underset{\sim}{x}-\underset{\sim}{x}'|^{-1}d^3x'.$$

The different predictions of various theories arose because these functions were multiplied by different constants. The PPN formalism replaced these constants by arbitrary parameters. Then on the one hand each theory could be characterized by the values of its parameters, and on the other hand the observations could be taken as measurements of the parameters, so one could then say whether a given theory was close or not to the measured point in the 10-dimensional PPN parameter space.

It turned out that light-deflection depended upon only the potential U and therefore upon only one parameter, called γ, and the general PPN version of Eq.2 is

$$\delta = 2.03598\ (1+\gamma)\cot\frac{\alpha}{2}\ \text{milliarcsec.} \qquad (3)$$

General relativity has $\gamma=1$, Brans-Dicke theory $(1+\omega)/(2+\omega)$, where ω is an arbitrary parameter of the theory, $\omega\to\infty$ being the general-relativistic limit. The effects of dragging of inertial frames depend on a combination of other parameters. The best observations to date have $\gamma=1$ to within 1% (Ref.8), but it is likely that milliarcsec-accurate positions for 10^5 stars could substantially improve this. I would therefore strongly urge that the observable relativistic effects be incorporated into the data analysis in their PPN-parametric form. In order to do this consistently, one must also determine whether there are any other measureable PPN effects which happen to be absent from general relativity.

2.3 Direct measurement of the metric

The PPN formalism was designed for a situation in which many theories were chasing a small amount of data. With the HIPPARCOS mission we will suddenly have a vast amount of data, mostly about the light-bending effect, and it seems to me that the PPN framework may be narrower than necessary for such a situation. It should be possible to use the data to measure some aspects of the metric directly. This seems to me the least model-dependent approach to the analysis of the data, because it makes no assumptions about how the Sun generates the metric, as the PPN formalism does. It is not completely free of assumptions, of course, as I will discuss below. First, however, I want to explain the term 'metric' and discuss coordinate systems.

2.3.1 Metric and coordinates. The metric of a space is its distance measure. It is usually expressed in a generalized form of the Pythagorean theorem, e.g. for Euclidean space in three dimensions

$$\begin{aligned} ds^2 &= dx^2 + dy^2 + dz^2 \\ &= dr^2 + r^2 (d\theta^2 + \sin^2\theta \, d\phi^2), \end{aligned} \tag{4}$$

the second form being appropriate to spherical coordinates. For special relativity there is a metric which measures time as well:

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2. \tag{5}$$

Integrated along the world line of a body, $|ds^2|^{1/2}$ gives the proper time, i.e. the time measured by a clock on the body. Eq.5 also permits the measurement of distances in spacelike directions, for which $ds^2 > 0$. For a general spacetime with coordinates $\{x^\alpha, \alpha=0,1,2,3\}$ the analogous expression is

$$ds^2 = \sum_{\alpha,\beta} g_{\alpha\beta}(x^\mu)\, dx^\alpha dx^\beta.$$

The functions $\{g_{\alpha\beta}\}$ of position x^μ are the components of the metric tensor, and because one can change one's coordinates (as in Eq.4) there is a whole family of such sets for each spacetime. If it is possible to find coordinates in which ds^2 takes the simple form in Eq.5 then we say that spacetime is flat. Otherwise it is curved.

The spacetime of the solar system is not the most general imaginable. If we can ignore Jupiter and the quadrupole moment of the Sun then we can idealize the metric as time-independent and spherically symmetric about the center of the Sun at any one moment. It should then suffice to adopt the form

$$\begin{aligned} ds^2 = g_{00}(r)\,dt^2 &+ 2g_{0\phi}(r)\,dt\,d\phi + 2g_{0\theta}(r)\,dt\,d\theta \\ &+ g_{rr}(r)\left[dr^2 + r^2(d\theta^2 + \sin^2\theta \, d\phi^2)\right]. \end{aligned} \tag{6}$$

Eq.6 involves a specific coordinate system: t, θ, and ϕ are chosen to show explicitly the time-independence and spherical symmetry of the instantaneous situation. (Compare the last term in Eq.6 with that in Eq.4.) The coordinate r is that used in the PPN formalism, and it makes the metric for space (t=const, dt=0 in Eq.6) as similar as possible to that of flat space, Eq.4.

Of course, the center of the Sun moves in an orbit around the barycenter of the solar system, and this barycenter is the natural center for a spherical coordinate system in flat space. But spacetime is curved by the Sun and so is not spherically symmetric about the barycenter: if observations were possible from there, apparent positions of stars would change with time because of the deflection caused by the Sun. This raises a question of principle, namely how we are able to define the coordinates of stars if they are not directly observable from anywhere, which I address in the next section. In practice, fortunately, this is an easy correction to make because the Sun moves slowly enough so that frame-dragging effects are negligible. One can therefore get away with using Eq.6 as the metric centered on the Sun, moving slowly with it.

2.3.2 Coordinate locations of stars and asteroids. Although we have chosen coordinates analogous to those of flat spacetime, we must not forget that spacetime is curved: no measurements can detect a flat space near the Sun, so any talk of an underlying flat space is a fiction, a coordinate-dependent concept. Yet our language does use this fiction. We speak of light deflection, as if we are measuring the angle between the true path of light and the path it would take if it could go straight. This is nonsense. The 'deflection' angle is simply the difference between the apparent position of the star seen from the Earth and that seen from the Sun, after allowance for parallax and aberration has been made. At any instant, the celestial sphere is spherically symmetric as observed from the Sun, and the 'deflection' helps us to determine these Sun-centered coordinates (Eq.6) from measurements made elsewhere. In a more strongly curved spacetime it would not even be possible to separate deflection, aberration, and parallax from one another uniquely: there would be just one grand correction. See Ref.5.)

But what about the barycentric coordinates which have been used until now? Observations from the barycenter itself are never undistorted by deflections, so what can the celestial coordinates mean? In fact, what one is really using is a coordinate system that reflects the spherical symmetry of the distant spacetime, say on the fringes of the solar system. There space is (ideally) spherically symmetric and the coordinates one uses ought to become spherically symmetric far away. They can do anything one likes in the middle of the solar system. So the celestial coordinates are not barycentric at all, they are simply asymptotically spherical coordinates at large distances from the Sun. If we used Sun-centered coordinates then the apparent positions of the stars would be a different mix of true positions, parallax, aberration, and deflection than if we used barycentric coordinates. Observations from the barycenter give the least time-dependent apparent positions of all possible observation positions, so this point is preferred. But in fact we observe from a satellite far from the barycenter and we want to construct the asyptotically spherical coordinates: the barycenter is irrelevant to this.

For stars with parallaxes, relativity creates no ambiguity in the distances we ultimately assign them -- the coordinate distance r of a star is not measureably different from its proper distance $\int |g_{rr}|^{1/2} dr$ from the barycenter. But for asteroids this is not true, since the two will differ by perhaps 10 km, which subtends an angle of the order of 10 milliarcsec at certain phases of observation.

For these reasons, the data analysts should make a clear choice of their coordinate system, and publish it. That in Eq.6 has already been used for other solar-system experiments, such as radar-ranging and satellite-tracking.

2.3.3 Model-dependence. Apart from the assumptions of symmetry and time-independence already noted, the use of Eq.6 involves the more fundamental assumption that the motion of light is determined exclusively by this thing called the metric. There are strong constraints in this direction from terrestrial experiments (Refs4,8) and I believe it would be very difficult to build a so-called non-metric theory with observable consequences for Hipparcos that still fit the terrestrial experiments.

2.3.4 Parametrization of the metric. If we are to widen our model beyond the PPN metric to the general metric in Eq.6, then we have to decide what we can measure. One possible choice is simply to expand the unknown metric terms in power series in r:

$$g_{rr}(r) = Ar^2 + Br + C + Dr^{-1} + Er^{-2} + \ldots \quad (7)$$

and so on. (Terms that grow with r^2 might indicate a non-zero cosmological constant.) The coefficients A, B, ... would then be fit to the data, although only certain combinations can be determined from light-deflection data. If general relativity is right then few of them will be nonzero to the accuracy of the observations. Existing experiments of course already constrain the coefficients. One side benefit of this approach to the data analysis has been suggested by M. Cruise: it might provide a test for some kinds of systematic errors in the measurements, for example if thermal effects caused a systematic error correlated with the position of the Sun. In this view an anomalous coefficient in Eq.7 would be cause for concern about the data, not about general relativity!

3. SELECTION OF TARGET STARS

Obviously, knowledge of repeated accurate positions of stars has many non-astrometric uses. There are several relativistic effects which one might consider looking for in the data, transient effects on a star's apparent position or magnitude or even multiplicity caused by gravitational lensing or deflection due to compact objects such as black holes or neutron stars in the neighborhood of the star or in the intervening interstellar space. I shall discuss only one, transient deflections caused by chance passage of a black hole near the line of sight to a programme star, but better calculations are needed of this and of other possible effects.

3.1 Random black-hole deflections

In Fig.3 I have drawn the geometry for deflection of a light ray by a black hole directly on the line of sight to a star. The hole is a distance d_1 from us, the star d_1+d_2. The light is deflected by an angle δ, given by Eq.1 in terms of the impact parameter r, and the star's apparent position is thereby deflected by an angle β. Using $\alpha=r/d_2$, $\beta=r/d_1$, and the geometrical relation $\alpha+\beta=\delta$, one can easily show that $\beta/\delta=d_2/(d_1+d_2)$. So if the hole doesn't pass too near the star, the observed change in position β will be of the same order as the deflection δ. I shall therefore make estimates based on the assumption that $\beta=\delta$, i.e. that we will observe a change in the star's position if the hole comes near enough to the line of sight to deflect light by an angle greater than the single-observation accuracy of 10 milliarcsec. This means that each hole of mass M has a 'deflection cross-section' of

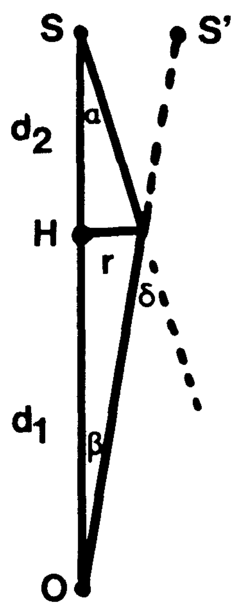

Figure 3. Simplified geometry for a black-hole deflection of a stellar position. The star is S, the hole is H, the observer is O, and the star's apparent position is S'. (See Ref.9 .)

radius (see Eq.1)

$$r_c \simeq 10^{13}\ M/M_\odot \ \text{cm}.$$

The deflection will not be noticed unless it changes during the lifetime of the satellite. Its timescale for change is simply the time of passage of the deflection disc across the line of sight, r_c/v, where v is the hole's speed:

$$T = 12\ \left(\frac{M}{M_\odot}\right)\ v_2^{-1}\ \text{days},$$

where v_2 is the hole's speed in units of 100 km/s. Therefore if the hole is a halo object with $v_2 \sim 3$, then we may expect to be able to detect holes with masses of the order of 10 $M_\odot$.

How many detections might we reasonably expect? If holes of mass 10 $M_{10} M_\odot$ make up a galactic halo of uniform density equal to $10^{-24}\rho_{24}$ g cm^{-3} (which for $\rho_{24}=1$ corresponds to 3 x 10^{11} $M_\odot$ within 15 kpc of the Galactic center), then there are roughly 1.5 x 10^6 $\rho_{24} M_{10}^{-1}$ holes per cubic kpc. This is surely an upper limit. Out to a distance R_3 kpc, their deflection disks cover a fraction f of the sky, roughly

$$f \sim 5 \times 10^{-9}\ R_3 M_{10} \rho_{24}.$$

If the halo extents to $R_3=50$ and if there are 100 extragalactic programme stars (including, say, Magellanic Cloud stars and bright unresolved objects in other galaxies) then the chance that any one observation would show a deflection is of the order of 2 x 10^{-5}. Of course, it is observable if it occurs in the first, say, 50 observations, so the chance of such a detection rises to 10^{-3}. This is small but not negligible, given the great importance that the discovery of a black hole would have. Once found in this fashion, it could be looked for by other telescopes, since it may be accreting matter and emitting a characteristic spectrum of radiation.

The deflection by a black hole would leave a characteristic signature in the residuals of a star's fit to its best position, shown in Fig.4 for a hole passing below the line of sight to a star: the star executes a small loop as the hole passes. The message to the data analysts: don't throw away your residuals!

Figure 4. As a black hole passes beneath the line of sight to the star, the star appears to travel on a loop, keeping away from the hole.

4. REFERENCES

1. McCrea W H 1979, Einstein: relations with the Royal Astronomical Society, Q. J. R. astr. Soc. 20, 251.

2. Shapiro I I 1967, New method for the detection of light deflection by solar gravity, Science 157, 806-808.

3. Ward W 1970, General-relativistic light deflection for the complete celestial sphere, Astrophys. J. 162, 345-348.

4. Misner C W, Thorne K S, & Wheeler J A 1973, Gravitation, San Francisco, W H Freeman & Co.

5. Murray C A 1981, Relativistic astrometry, Mon. Not. Roy. astr. Soc. 195, 639-648.

6. Dicke R H & Goldenberg H M 1967, Solar oblateness and general relativity, Phys. Rev. Lett. 18, 313-316.

7. Richter G W & Matzner R A 1981, Gravitational deflection of light at 1½ PPN order, Astrophys. & Sp. Sci. 79, 119-127.

8. Will C M 1979, The confrontation between gravitation theory and experiment, in Hawking S W & Israel W, General Relativity, an Einstein Centenary Survey, Cambridge, Cambridge University Press.

9. Refsdal S 1964, The gravitational lens effect, Mon. Not. Roy. astr. Soc. 128, 295-306.