Fermat's principle in general relativity and light deflection by the sun

Klaus Kassner 2 January 2015

In my last post including longer derivations, I stated my surprise in discovering that Fermat's principle works in general relativity, showing its validity in curved spacetime. This result was essentially based on the fact that using the principle, I could calculate not only the part of light deflection by the sun that can be explained using the equivalence principle but also the part due to spatial curvature. What is more, checking an alternative theory of gravity (due to Brans and Dicke) which leads to a *different* spatial curvature in spherically symmetric spacetime, I got the correct non-equivalence principle contribution for that theory as well.

However, obtaining a correct result or two proves little, especially in this field, where a number of derivations have been published that get a correct final result, too, but are based on fallacious arguments. For example, Schiff claimed in 1959 to be able to derive the *full* deflection employing the equivalence principle alone [1]. He got the full deflection but used an incorrect argument that happens to give the right answer in the case of the spherically symmetric gravitational field. It fails in other cases. There was a rebuttal by Rindler [2], which is correct but discusses a much wider scope than necessary, viz. the question of how to operationally construct a *global* metric from the equivalence principle.[1]

So I was a bit worried whether the conclusion based on my amusing results was correct, in particular, because I failed to see how my version of Fermat's principle was generally covariant. Also I was unsure of how to generalize it to non-isotropic coordinates. So I decided to do a bit of research in the web. If Fermat's principle is applicable in general relativity, somebody must have published on this before. And indeed, this turned out to be true. The principle seems to be quite useful in calculating gravitational lensing.

Apparently, the first derivation of Fermat's principle in curved spacetime, then still restricted to a static spacetime, was due to Weyl [3]. In the meantime, appropriate forms of the principle have been given that are valid in arbitrary spacetimes [4]. A very nice (and short!) proof is given in [5].

The most general and manifestly covariant formulation of the principle is the following: Given an event p, a world line $\mathcal{L}$ and the set of (future-oriented) null curves starting at p and arriving on $\mathcal{L}$, the curve taken by light is stationary with respect to first-order variations of the null curve within the set.[2]

A usable form of the principle arises by specifying that the quantity that is stationary with respect to the variation considered is the arrival time of the light ray on $\mathcal{L}$, measured in some arbitrary time coordinate along the world line.

[1] What is required for the application of the equivalence principle to light bending is less – we only need a consistent *local* metric about the actual path taken by the light.

[2] A similar principle exists for massive particles, where however we can take two fixed time-like events instead of a fixed event and a fixed world line. This is due to the fact that all events that are time-like with respect to a given event fill a volume of spacetime, whereas all events that are null with respect to a given event fill only a surface of spacetime.

For our considerations, it is completely sufficient to have the mathematical formulation of the principle in stationary or static spacetimes, where it reads:

$$\delta \int_{t_{\mathrm{emit}}}^{t_{\mathrm{obs}}} \mathrm{d}t = \delta \int \frac{\mathrm{d}l}{\sqrt{g_{00}}} = 0 \,. \tag{1}$$

In the second formulation $\mathrm{d}l$ is the proper line element, in general given by

$$\mathrm{d}l^2 \equiv \gamma_{ij} \mathrm{d}x^i \mathrm{d}x^j = \left(-g_{ij} + g_{0i} g_{0j}/g_{00}\right) \mathrm{d}x^i \mathrm{d}x^j \qquad (i,j = 1,\ldots 3) \tag{2}$$

and the integration limits can be fixed points in space now.[3]

Now in my application of the principle, I used an isotropic time-orthogonal metric, i.e., a metric of the general form:

$$\mathrm{d}s^2 = g_{00}\, \mathrm{d}t^2 - f(x,y,z)^2 \left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \,, \tag{3}$$

where x, y, z are cartesian coordinates. We obviously have $\gamma_{ij} = f(x,y,z)^2 \delta_{ij}$ and the local coordinate speed of light is given by

$$\begin{aligned} 0 &= g_{00}\, \mathrm{d}t^2 - f(x,y,z)^2 \left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \\ \Rightarrow \quad c(x,y,z) &\equiv \frac{\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2}}{\mathrm{d}t} = \frac{\sqrt{g_{00}}}{f(x,y,z)} \,. \end{aligned} \tag{4}$$

($f(x,y,z)$ is assumed to be positive.)
Moreover, we have $\mathrm{d}l = f(x,y,z)\sqrt{\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2} \equiv f(x,y,z)\mathrm{d}\tilde{l}$, where we may consider $\mathrm{d}\tilde{l}$ a coordinate line element.[4] Then Fermat's principle (1) may be rewritten

$$0 = \delta \int \frac{\mathrm{d}l}{\sqrt{g_{00}}} = \delta \int \frac{f(x,y,z)\mathrm{d}\tilde{l}}{\sqrt{g_{00}}} = \delta \int \frac{\mathrm{d}\tilde{l}}{c(x,y,z)} \,, \tag{5}$$

which is precisely the form I used. It is remarkable that we need the coordinate line element $\mathrm{d}\tilde{l}$ here instead of the (invariant) proper length element. So my intuition about the form of Fermat's principle in these coordinates was correct, and my reasoning in that article still stands.[5]

I recall the isotropic form of the Schwarzschild metric here

$$\mathrm{d}s^2 = k_1^2\, c^2 \mathrm{d}t^2 - k_2^2 \left(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2\right) \,,$$

with

$$\begin{aligned} k_1 &= \frac{1 - \frac{GM}{2Rc^2}}{1 + \frac{GM}{2Rc^2}} \,, \\ k_2 &= \left(1 + \frac{GM}{2Rc^2}\right)^2 \,, \qquad\qquad R = \sqrt{x^2 + y^2 + z^2} \end{aligned} \tag{6}$$

[3] As we can choose the time-like world line to correspond to a fixed spatial position.
[4] $\mathrm{d}\tilde{l}$ was called $\mathrm{d}s$ in my last essay, a notation avoided here for obvious reasons.
[5] Even though I had not realized that in anisotropic coordinates, the principle would take a different form...

from which we deduce a speed of light

$$c(x,y,z) = c(R) = \frac{k_1}{k_2} c \approx c\left(1 - \frac{2GM}{Rc^2}\right) = c\left(1 + \frac{2\Phi(R)}{c^2}\right) \tag{7}$$

and obtain Fermat's principle in the form

$$\delta \int \frac{\mathrm{d}\tilde{l}}{c\left(1 + \frac{2\Phi(R)}{c^2}\right)} = 0\,. \tag{8}$$

Having the general form of the principle at hand now, we can also formulate it in traditional Schwarzschild coordinates

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

with

$$k = \sqrt{1 - \frac{2GM}{rc^2}} = \left(1 + \frac{2\Phi(r)}{c^2}\right)^{1/2}\,. \tag{9}$$

The two radial coordinates r and R are related by

$$r = R\left(1 + \frac{GM}{2Rc^2}\right)^2\,. \tag{10}$$

Using (1), we can now write down Fermat's principle in these coordinates:

$$\delta \int \frac{1}{kc}\left[\frac{1}{k^2}\mathrm{d}r^2 + r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)\right]^{1/2} = 0\,. \tag{11}$$

Introducing the weak-field form of the potential again and expanding the square root describing k, we have:

$$\delta \int \frac{1}{c\left(1 + \frac{\Phi(r)}{c^2}\right)}\left[\frac{1}{1 + 2\frac{\Phi(r)}{c^2}}\mathrm{d}r^2 + r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)\right]^{1/2} = 0\,. \tag{12}$$

It appears that the form (8) is easier to use mathematically, and this was indeed the form that I used in my previous calculation.

But the form (12) is nicer to interpret. We note first that in doing the variation, we may restrict ourselves to small variations about the true path of the light ray. Since for the part of the integral that describes the passing of the ray close to the sun, we know that the actual light ray is essentially tangential to the sun's surface, hence has only a ϑ and φ component, we may rewrite that part of the integral again using a local speed of light:

$$\int \frac{\mathrm{d}\tilde{l}_t}{c_t(r)}\,,$$

where

$$c_t(r) = c\left(1 + \frac{\Phi(r)}{c^2}\right)\,,$$

$$\mathrm{d}\tilde{l}_t = r\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\,\mathrm{d}\varphi^2\right)^{1/2}\,. \tag{13}$$

That is, c_t is the tangential speed of light[6] and $\mathrm{d}\tilde{l}_t$ is the tangential spatial line element. Note that c_t is precisely the speed of light that is predicted by application of the equivalence principle. Therefore, we can say that close to the sun, the contribution to the integral to be varied is extremely well approximated by the equivalence principle. This is physically satisfying, as the equivalence principle should contain all first-order effects of gravity and near the sun all that happens to the light ray is bending, so bending is a first-order effect and should be fully captured by the equivalence principle.

Far from the sun, the situation is different. Now the light ray is oriented radially.[7] But then, we may again rewrite the contribution to the integral in terms of a local speed of light.

$$\int \frac{1}{c}\frac{1}{1+\frac{\Phi(r)}{c^2}}\frac{\mathrm{d}r}{1+\frac{\Phi(r)}{c^2}} = \int \frac{\mathrm{d}r}{c_r(r)}\,,$$

where

$$c_r(r) = c\left(1+\frac{2\Phi(r)}{c^2}\right) \tag{14}$$

is the radial speed of light in the Schwarzschild geometry. Local bending as predicted by the equivalence principle is almost zero here, as the light ray is almost (anti)parallel to the direction of acceleration in the "equivalent elevator", so the first-order effect is a blueshift or redshift. However, there is still variation of the speed of light, which is now due to the spatial curvature. And this variation must lead to local bending according to Huygens' principle.

These considerations confirm the qualitative picture that I have given of the light bending in a preceding answer on Research Gate.

In isotropic coordinates, we can see immediately that the bending in the Schwarzschild metric is twice that due to the equivalence principle alone, because the speed of light is $c\left(1+\frac{2\Phi(r)}{c^2}\right)$ in the first case and $c\left(1+\frac{\Phi(r)}{c^2}\right)$ in the second throughout the path of integration.

But isotropic coordinates hide the fact that space is really anisotropic in the Schwarzschild metric. This is easy to see: From the point of view of a distant observer, space is anisotropic, because there is length contraction in the radial direction and none in the orthogonal directions. From the point of view of a local observer, the situation is even more dramatic. If he is close to the event horizon of a black hole, motion along the negative r direction may mean inevitable doom, whereas motion in the tangential direction may save his life. Even worse, the positive and negative r directions are inequivalent. A less dramatic way of noticing the anisotropy, even far from the center, is to imagine our observer surrounded by close-by standard clocks or sources of light and to notice that the clocks along the radial direction tick at different rates, whereas those arranged perpendicularly to the radial direction tick at the same rate; light from the positive radial direction arrives blueshifted, light from the negative radial direction arrives redshifted and for light from tangentially

[6]Which corresponds to the speed of light orthogonal to the radial coordinate in the Schwarzschild geometry.

[7]The angle between the light ray and the radial direction is given approximately by b/r for $b \ll r$, where b is the impact parameter, i.e., the closest distance to the sun. When r is a few light minutes, this is already very small, as b is less than 3 light seconds.

displaced positions there is neither a redshift nor a blueshift. Clearly, the isotropy of the speed of light for a local observer cannot be taken as a proof of an isotropic situation as we know that this local isotropy is a tautology, following from our choice of units rather than from physics. Remember that I am not such a stubborn defender of this tautology as Robert Shuler, for example, since I believe the fact to be nontrivial that measurements of the proper length via material rulers and via light signals give the same answer. (This fact gives us a clue that our world is Lorentz invariant, and it might break down with Lorentz invariance at the Planck scale.)

If you read the Schiff paper and wonder what is wrong with his argument without wishing to delve into Rindler's elaborate answer, here is a short sketch of my view.

The equivalence principle says that we can replace the effects of a uniform gravitational field by effects of acceleration. Since uniform gravitational fields do not exist globally, we have to restrict their consideration to small spatial regions.

But how precisely do we replace a stationary observer who is in a gravity environment by one that is accelerating? Why, the accelerating observer should *instantaneously be at rest* with respect to the position of the stationary one. Then the observations of the accelerating observer should be related correctly to similar observations of an appropriate inertial observer. A direct comparison is only possible for a local inertial observer close by, so the first step should correspond to looking at this case. Later we may discuss how to relate the result to the far-away inertial observer.

Since our local inertial observer is *also* assumed to be at rest, it is clear that between our local inertial observer and the accelerated observer, there will be *neither* length contraction *nor* time dilation. Now of course, at least in order to compare clocks, we have to consider a time *interval*, and then the accelerated observer will not remain at rest but acquire a small velocity $\delta v = a\delta t$. However, both time dilation and length contraction are quadratic in δv and since the equivalence principle gets only effects to linear order correctly, we can neglect these effects.

Yet, there is an effect that *is* linear in δv. If all clocks in the accelerated system are synchronized with the clocks of the inertial observer at $t = 0$, they will, due to the Lorentz transformation $t' = \gamma(t - \frac{\delta v\, x}{c^2})$, where x is the coordinate parallel to the direction of motion, not be synchronized anymore at $t = \delta t$. The "axis of simultaneity" of the accelerating observers tilts;[8] it tilts so that a clock at positive x would have to show an earlier time to still be considered synchronous with the clock at $x = 0$. But since the time on the clock does not change – only the notion of simultaneity does – the clock actually advaces from the point of view of the accelerated observer at $x = 0$. This results in observers in the accelerated system "seeing" clocks running the faster, the farther their distance in the direction of movement. The observers may notice this via blueshift of photons sent "downward" from an observer at larger x to an observer at smaller x and a redshift of photons sent "upward" from smaller x to larger x. For the inertial observer, these wavelength shifts are explicable in terms of the Doppler effect. Because a signal will always arrive at an observer who is faster than the sender was at the time of emission of the signal, there will be a blueshift, if its motion is against the direction of acceleration and

[8]This is best visualized in a Minkowski diagram.

a redshift if it is parallel. The accelerated observers may interpret this as a time dilation effect, an explanation that is possible, since their axis of simultaneity tilts. Not all Doppler shifts may be explained via time dilation.[9]

Now the speed of light in the frame of the local inertial observer is isotropic. For him it is actually precisely c, but if we wish to compare it with that of a very distant inertial observer, we would say that the two observers will disagree about the local speed of light by a factor at most, which is the unit of speed measurement. For the accelerated observers, their speed of light will vary, for the short time interval considered, only due to time dilation, which actually is a consequence of the continual desynchronization of clocks.[10] The variation will be by a factor $1 + \left|\frac{\delta v}{\delta t}\right| \frac{\Delta x}{c^2}$ between two observers that are a distance Δx apart in the direction of acceleration. For the observer that is exactly at the same position as our local inertial observer, the speed of light will be c, which translates to $c(r)$ when expressed as a coordinate speed of the far away inertial observer.[11] Noticing that the local acceleration is $a = \left|\frac{\delta v}{\delta t}\right| = \Phi'(r)$, we may obtain $c(r)$ via integration from the actual r to ∞ and the requirement that the integration constant for $r \to \infty$ is c. This yields $c(r) = c\left(1 + \frac{\Phi(r)}{c^2}\right)$.

So Schiff's explanation of time dilation is wrong, because he introduces superfluous time and length dilation factors at a point where they do not arise. (We can always assume the local inertial observer to be at rest with respect to the one stationary in the acceleration field, and then the only effect that produces time dilation is clock desynchronization. There seems to be no analogous effect for lengths.)

In fact, the situation in the elevator is well described by the Rindler metric, which has time dilation but does not have length contraction and reproduces all the predictions that follow from the equivalence principle alone. In the true Schwarzschild metric, there *is* length contraction (and Schiff skillfully twisted his arguments to obtain it[12]), but it is a consequence of spacetime curvature.

[9]Here is a simple counterexample. Consider a rotating disk with its center at rest in some inertial system. Let Cheryl be the observer at this center, Abby a disk observer sitting at a fixed radius, and Bella an inertial observer, instantaneously at rest with respect to Abby. Since Abby considers herself to be at lower gravitational potential on the disk than Cheryl, she will "see" Cheryl's clock running fast, and Cheryl will agree that Abby's clock is slower than hers. Cheryl will consider Bella's clock running at the same rate as Abby's, but Bella, being inertial just as Cheryl, will view Cheryl's clock as running slow. Time dilation is mutual for inertial observers. Now let Cheryl send a light signal, containing the hydrogen spectrum, to Abby and Bella, timed so that it arrives when the two of them are at the same location. Abby clearly sees this spectrum blueshifted. Cheryl's clock runs faster than hers. Alternatively, she can construct an argument about the loss in potential energy of the photons, increasing their frequency. Since reception of a light signal is a completely local measurement (measurement of time dilation is not), Bella also has to see the spectrum blueshifted, despite the fact that Cheryl's time is slow with respect to hers. But this can be explained easily. Cheryl and Bella are on antiparallel trajectories and when the light signal arrives, they are at their closest separation (according to Bella). So at the time the signal was sent, Cheryl was moving *towards* Bella, which means there is a Doppler shift of the signal to higher frequencies that overcompensates the time dilation effect. But this Doppler shift cannot then be explained as a consequence of time dilation by Bella...

[10]This also explains why the time dilation is not mutual as usually the case in special relativity. All internal observers along the x axis agree in which direction time becomes faster and in which direction it slows down.

[11]x is a coordinate along the r direction. We have $\Delta x = \mp\Delta r$ for infalling and outgoing light rays, respectively.

[12]He certainly *believed* to get it from the equivalence principle, but whether he would have obtained a correct result without knowing the answer ahead of time, is an open question to me.

References

[1] L. I. Schiff, On Experimental Tests of the General Theory of Relativity, Am. J. Phys,**28**, 340 – 343 (1960).

[2] W. Rindler, Counterexample to the Lenz-Schiff argument, Am. J. Phys. **36**, 540 – 544 (1968).

[3] H. Weyl, Zur Gravitationstheorie, Ann. Phys. **54** (or **359**), 117 (1917).

[4] I. Kovner, Fermat principle in arbitrary gravitational fields, Astrophys. J. **351**, 114 – 120 (1990).

[5] R. Nityananda and J. Samuel, Fermat's principle in general relativity, Phys. Rev. D **45**, 3862 (1992).