General relativity, 4

Orbital motion of small test masses

The starting point for analyzing free fall trajectories in the (2-space, 1-time) Schwarzschild spacetime is Equation (3) from GR 3

$$(d\tau)^2 = \left(1-\frac{r_S}{r}\right)(dT)^2 - \left(1-\frac{r_S}{r}\right)^{-1}(dr)^2 - r^2(d\phi)^2, \qquad (1)$$

in c = 1 units (in which all terms have dimensions of length^2). Worldlines for freely falling *test masses* (i.e., masses that don't contribute to the gravity described in (1)) in general relativity correspond, in some sense, to the "straightest" paths through spacetime. In special relativity, without gravity, a straight worldline corresponds to the maximum proper time path connecting two events. (Recall, for example, the discussion of the twin paradox. The time between departure and return for the twin staying on Earth [i.e., events connected by a single straight worldline] is longer than for the twin travelling out and back [which requires two different worldlines glued together to connect the same events].) By extension, the desired free-fall worldline connecting two events is the continuous sequence of events, (T, r, ϕ), such that $\Delta\tau = \int_{start}^{end} d\tau$ along the path is greater than along any other path connecting the same events. (More precisely, the path should be broken up into a sequence of infinitesimal straight line segments each labeled by a real number, λ, starting at 0 and ending at 1, and $\Delta\tau = \int_{\lambda=0}^{1} \frac{d\tau}{d\lambda} d\lambda$.)

Conserved quantities

Finding these worldlines is greatly facilitated by borrowing a result from the "calculus of variations": ***if a coordinate (not its differential) does not appear explicitly in the integrand of an integral to be maximized there will be a quantity associated with that coordinate that will remain constant at any point on the desired path***. In (1) above, r explicitly appears in $\left(1-\frac{r_S}{r}\right)$ and in r^2, but neither T nor ϕ appear explicitly. Thus, there are two associated "constants of the motion" for a test mass freely falling through the spacetime described by equation (1). Again, borrowing from the calculus of variations, the constant corresponding to the absence of T is $\left(1-\frac{r_S}{r}\right)\frac{dT}{d\tau}$ and to the absence of ϕ is $r^2 \frac{d\phi}{d\tau}$.

So, what are these constants? Let's start with $r^2 \frac{d\phi}{d\tau}$. As noted in GR3, angular momentum, $\vec{L} = \vec{r} \times \vec{p}$, is constant for masses freely falling in spherically symmetric gravity. The relativistic definition of momentum is $\vec{p} = m\, d\vec{r}/d\tau$; $d\vec{r}/d\tau$ can have two components, one in the radial direction, which contributes nothing to the cross product for $\vec{L}$, and one perpendicular to $\vec{r}$, which in spherical coordinates has a magnitude $r d\phi/d\tau$. Thus, the magnitude of $\vec{L}$ is $L = mr^2\, d\phi/d\tau$. (In c = 1 units, L has the dimensions of mass times length.

In conventional units $L = mcr^2\, d\phi/d\tau_{conv}$, where τ_{conv} is measured in units of time.) Thus, the ϕ-related constant is $r^2 \dfrac{d\phi}{d\tau} = \dfrac{L}{m}$.

This result can be used to eliminate $d\phi$ in the Schwarzschild proper time:

$$(d\tau)^2 = \left(1 - \frac{r_S}{r}\right)(dT)^2 - \left(1 - \frac{r_S}{r}\right)^{-1}(dr)^2 - \left(\frac{L}{mr}\right)^2 (d\tau)^2. \qquad (2)$$

If we multiply both sides of (2) by $\left(1 - \dfrac{r_S}{r}\right)$ and divide by $(d\tau)^2$ (and rearrange), we obtain

$$e^2 - 1 = \left(\frac{dr}{dt}\right)^2 + \left(\frac{L}{mr}\right)^2 - \frac{r_S}{r}\left[1 + \left(\frac{L}{mr}\right)^2\right], \qquad (3)$$

where $e = \left(1 - \dfrac{r_S}{r}\right)\dfrac{dT}{d\tau}$. If we multiplied both sides of (3) by $m/2$ the first term on the right would look like the test mass's "radial kinetic energy," the second would look like its " angular kinetic energy," and (because $r_S = 2GM$) the third would look like a generalization of the potential energy of interaction between M and m. Thus, the constant e describes conservation of energy for the freely falling mass.

Equation (3) can be integrated to find $r(\tau)$ in terms of the input constants L and e. This can be combined with $d\phi/d\tau = L/mr^2$ to find $\phi(\tau)$, and hence the test mass' position in the (r,ϕ)-plane at any instant. It is often more useful to determine the *shape* of the test mass' trajectory, that is $r(\phi)$, rather than its instantaneous position. This can be achieved by noting that $\dfrac{dr}{d\tau} = \dfrac{dr}{d\phi}\dfrac{d\phi}{d\tau} = \dfrac{dr}{d\phi}\dfrac{L}{mr^2}$. Inserting into (3) yields

$$e^2 - 1 = \left(\frac{L}{m}\right)^2 \left[\frac{1}{r^4}\left(\frac{dr}{d\phi}\right)^2 + \frac{1}{r^2}\left(1 - \frac{r_S}{r}\right)\right] - \frac{r_S}{r}.$$

Though it's probably not obvious, the substitution $r = 1/x$ greatly simplifies the latter equation. In particular,

$$e^2 - 1 = \left(\frac{L}{m}\right)^2 \left[\left(\frac{dx}{d\phi}\right)^2 + x^2(1 - r_S x)\right] - r_S x. \qquad (4)$$

Orbits in the Newtonian limit

In Newtonian gravity, test masses travel slowly compared with c, their rest energies are much greater than their mechanical energies, and their distances from M are much greater than M's Schwarzschild radius. In the Newtonian limit, the left hand side of (4) can be replaced by $e^2 - 1 = 2\,E_{mechanical}/mc^2$, where $E_{mechanical}$ is the sum of kinetic and potential energies. Multiplying both sides of (4) by $mc^2/2$, noting that noting that $x^2 >> x^2 r_S x$, and remembering that $r_S = 2GM/c^2$ leads to

$$E_{mechanical} = \frac{L^2}{2m}\left[\left(\frac{dx}{d\phi}\right)^2 + x^2\right] - GMmx \,. \tag{5}$$

The first term on the right hand side of (5) is the total kinetic energy of the test mass (radial part plus angular part), while $-GMmx$ is just the Newtonian gravitational potential energy (remember $x = 1/r$). As $E_{mechanical}$ is constant, (5) can be differentiated with respect to ϕ to obtain

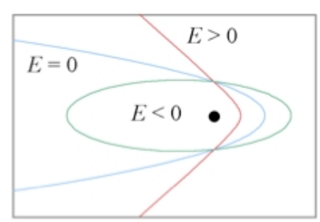
$$\frac{d^2x}{d\phi^2} = -x + \frac{GMm^2}{L} \,, \tag{6}$$

an equation that is formally identical to that of a simple harmonic oscillator (where ϕ takes the place of time), with an "angular frequency" equal to 1, and an "applied constant force" that shifts the oscillator's equilibrium position from 0 to GMm^2/L. Solutions to the simple harmonic oscillator equation are sines or cosines so, for example, $x(\phi) = x_0 + C\cos(\phi)$, where $x_0 = GMm^2/L$ and C is a constant of integration. This constant can be expressed in terms of $E = E_{mechanical}$ and L by plugging x back into the energy equation: $C = \pm\sqrt{x_0^2 + 2mE/L^2}$. Thus, we find that

$$r(\phi) = 1/\left(x_0 + C\cos(\phi)\right).$$

Let's assume that $C > 0$. When that is true the closest approach of the test mass to M (the position where r is smallest) occurs at $\phi = 0$ (where the cosine has its largest value). Because the gravitational potential energy is defined to be negative, E can be >, =, < 0.

Case (i): **Open trajectories** Suppose $E \geq 0$, i.e., where m's kinetic energy ≥ –potential energy. When $E \geq 0$, $C \geq x_0$. For both positive and negative values of ϕ the cosine is less than 1 and, in fact, can be negative. Thus, there are two values of ϕ for which $r \to \infty$, namely, where $\cos(\phi) = -x_0/C$: i.e., $\phi_C = \cos^{-1}(-x_0/C)$. For $E = 0$, $C = x_0$, and $\phi_C = \pm\pi$, whereas for $E > 0$, $|\phi_C| < \pi$. In the former case, the trajectory is an open *parabola*, and for the latter case, it's an open *hyperbola*, as shown in the figure to the right. Thus, in these cases m comes in from infinity, swings by M with a closest approach of $r(0) = 1/(x_0 + C)$, then returns to infinity.



Case (ii): Closed orbits On the other hand, when $E < 0$, we get a different story. The smallest E can be and still have a real value for C is $E_{\min} = -L^2x_0^2/2m$. For this energy $C = 0$, and r has a single value, $r = 1/x_0$, for all values of ϕ; that's a circular orbit. For $E_{\min} < E < 0$, $0 < C < x_0$, and thus r is a minimum at $\phi = 0$ and a maximum at $\phi = \pi$. Furthermore, every time ϕ is incremented by 2π, r returns to its prior value: such an orbit is a closed, periodic ellipse. See the figure. Another way of thinking about these cases is that for $E > 0$, the speed of the test mass exceeds the "escape speed," while for $E < 0$, m's speed is less than the escape speed. For the incredibly difficult to achieve special case where $E = 0$, m's speed exactly equals the escape speed. We'll come back to this "miraculous" case later when we discuss the expanding universe.

Orbits in the full Schwarzschild spacetime

1. Precession of the perihelion

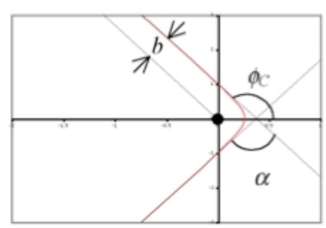
Now, making no assumption about the smallness of $x^2 r_S x$ in Equation (4), differentiation by ϕ leads to

$$\frac{d^2x}{d\phi^2} = -x + \frac{3}{2} r_S x^2 + \frac{GMm^2}{L} . \qquad (7)$$

This equation does *not* correspond to a simple harmonic oscillator; test mass trajectories for $E < 0$ do not form closed ellipses. Instead, every orbital pass finds the angular position of the minimum approach to M to be shifted somewhat. For orbits of objects in the solar system about the Sun this orbital shift is called "precession of the perihelion." (Perihelion means "closest approach to the Sun.") The various pulls of all of the planets on one another cause a much larger perihelion shift than the general relativistic effect. If one accounts for all of these, though, there's still some shift left over. The residual shift is not a lot—tens, or less, of seconds of arc per century! Only four solar system objects get close enough to the Sun for a reasonable estimate of the GR shift to be made from solar system observations. They are: Mercury (in $''$(century)$^{-1}$: 43.1±0.5 measured versus 43.0 predicted), Venus (8.4±4.8 versus 8.6), Earth (5.0±1.2 versus 3.8), and the asteroid Icarus (9.8±0.8 versus 10.3). All measured residual shifts are compatible with the predictions of GR. Note that in an elliptical orbit the speed of m varies as r varies. If we set the kinetic energy of m equal to its special relativistic form, $(\tilde{\gamma} - 1)mc^2$, Newtonian gravity will also predict a perihelion shift—but only about 1/6th as large as the GR prediction.

An even more compelling corroboration of the general relativistic orbital precession phenomenon can be found in a remarkable binary star system discovered by Joseph Taylor and Richard Hulse (then Taylor's graduate student) in 1974, using the Arecibo radio telescope. They observed a pulsar whose radiation Doppler shift alternates back and forth between "red" and "blue" in a way that can only be due to orbital motion. This pulsar orbits a second condensed stellar remnant once every 7.75 h, suggesting they are very close to one another; each of the partners has a mass of about 1.4 solar masses. Using the observed orbital characteristics of this system, general relativity (it's not exactly Schwarzschild, because both stars are moving) predicts that the periastron (closest mutual approach of the stars) of the Hulse-Taylor binary should shift by about 4.2° per year (it would take Mercury 40,000 years to achieve that much precession!), which is exactly what is seen.

2. Bending of "starlight"



The Schwarzschild spacetime predicts that light also gravitates. Multiply both sides of Equation (4) by $(m/L)^2$ and take the test body to be a photon with $m \to 0$:

$(m/L)^2(e^2 - 1) \to (E/L)^2 = (hf/hb/\lambda)^2 = 1/b^2$ and $(m/L)^2 r_S x \to 0$.

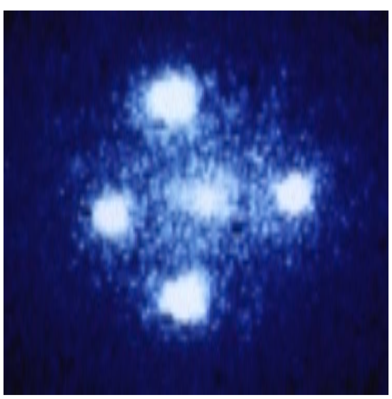
In these limits, the photon energy, hf , and momentum, h/λ , are used and the photon's orbital angular momentum is hb/λ , where b is defined as in the figure above. The resulting photon equation of motion is

$$\left[\left(\frac{dx}{d\phi}\right)^2 + x^2\left(1-r_S x\right)\right] = \frac{1}{b^2} \tag{8}$$

It is possible to integrate the photon orbit equation directly (though you need Maple or Mathematica or something to help).

We are interested in the angle α in the figure above. That's the angle of deflection of a photon falling in from infinity, swinging around the Sun, and heading back out to infinity. Plugging in values of M and b (= the solar radius, for grazing incidence) for the Sun we find $r_S / b \approx 4.2\text{x}10^{-6}$. In this weak gravity limit, the integration of the orbit equation yields $\alpha = 2r_S / b \approx 8.4\text{x}10^{-6}$ radians, or 1.75″ of arc. Despite its small value, this result is now well confirmed. The first observational confirmation of this prediction of general relativity in 1919 by Eddington during a total eclipse of Sun is one of the most celebrated results of 20th century science. Announcement of Eddington's observation appeared on the front page of the NY Times and certainly helped establish Einstein's fame as a legendary intellect. Unfortunately, these observations had large measurement uncertainty. On the other hand, recent (1996) observations of the bending of emissions from radio sources (3C273B and 3C279) using very-long-baseline-interferometry confirm GR's prediction with an uncertainty no worse than 1 part in 10^3. Incidentally, if light were treated as a mass falling in Newtonian gravity the prediction of deflection would be half the observed values, well outside of the actual observational uncertainty. Thus, Newtonian gravity supplemented by special relativity makes qualitatively correct predictions for both perihelion precession and for light bending, but the quantitative values are demonstrably incorrect.

3. Gravitational lensing



Of course, any large mass will cause light passing it to deflect. A large M can act as a not-so-great, **gravitational converging lens**. Images collected from the Hubble telescope, for example, show both multiple copies of the light from one galaxy passing on its way to earth past another galaxy and also many greatly distorted galactic images presumably due to some intervening large mass. The leftmost image above shows the "Einstein cross." The outer four blobs are images of the same distant quasar; the center blob is a relatively nearby galaxy (http://hubblesite.org/newscenter/archive/releases/1990/20/image/a/).

The center image above shows marked elongation of several very distant galaxies (small circular arcs) as their light passes through the "Abell galactic cluster." (http://hubblesite.org/newscenter/archive/releases/2003/01/image/a.) Since the light deflection is proportional to the intervening mass, it is determined that *the visible matter in the Abell cluster is a factor of five times too small to cause the observed elongation pattern*. This invisible stuff is called **dark matter** (not because it's dark, but rather because *it isn't bright* like the glowing galaxies). An even more dramatic "image" of dark matter constructed using gravitational lensing is shown on the right above. In the image, two galactic clusters are colliding. The hydrogen gas clouds each cluster carries with it are crashing into each other and emitting short wavelength electromagnetic radiation that is color-coded as pink. This radiation is produced when "colliding" atoms interact electromagnetically. The clusters also carry dark matter, which only interacts gravitationally. Thus, in the collision the ordinary matter is decelerating while the dark matter keeps on going (and hence separates from the cluster it used to travel along with). The blue color-coded blobs of dark matter are inferred by gravitational distortion of light from galaxies behind the collision (http://hubblesite.org/newscenter/archive/releases/2008/32/image/a/).