Chapter 9

Schwarzschild Solution

Let us begin our study by considering one of the first problems tackled after Einstein published *The Foundation of the General Theory of Relativity* in Annalen der Physik **49**, 1916. While Karl Schwarzschild lay dying from a dreaded disease that he picked up at the front during the *war to end all wars*, he worked out the famous solution of the vacuum Einstein equations that bears his name.

Here we suppose that the only matter present is a spherically symmetric source, which may be either completely static, expanding or contracting. Then any associated spacetime curvature must reflect that inherent spherical symmetry. If one considers any spherical two-surface centered upon the symmetry point, the geometry of that two-surface will be indistinguishable from the geometry on the surface of a sphere of the same radius imbedded in flat space. But wait, what does one mean by "radius?" Without departing from the chosen two-surface, one can determine its circumference, and one can *define* r to be the circumference divided by 2π. We shall regard this r to be a label of the chosen two-surface.

Since, in the Einstein theory,

$$d\mathbb{I} = d(e^a\mathbf{e}_a) = de^a\mathbf{e}_a - e^a d\mathbf{e}_a = \{de^b - e^a\Gamma_a{}^b\}\mathbf{e}_b = 0,$$

it follows that

$$de^b = e^a\Gamma_a{}^b.$$

In the case of a two dimensional spherical space, described in terms of polar coordinates θ, ϕ, we may employ two orthonormal 1-forms

$$\begin{aligned} e^1 &= r\, d\theta, \\ e^2 &= r\, \sin\theta\, d\phi, \end{aligned}$$

where r is the radius of the sphere and $r\sin\theta$ is the radial distance outward from the axis to a line of latitude.

Evaluating the exterior derivative of each of the 1-forms, we have

$$\begin{aligned} de^1 &= 0, \\ de^2 &= r\cos\theta\, d\theta\, d\phi \\ &= r^{-1}\cot\theta\, e^1 e^2. \end{aligned}$$

Hence, one concludes that

$$-{\Gamma_1}^2 = -\Omega_{12} = \Omega_{21} = {\Gamma_2}^1 = -r^{-1}\cot\theta\, e^2 = -\cos\theta\, d\phi.$$

It is apparent that the only nonvanishing components of the Riemann curvature tensor are those related to R_{1212}, which has the value $-r^{-2}$ in the present case. In the case of a sphere the scalar curvature is independent of position on the sphere![1] The method of calculation which we have just illustrated permits us to avoid having to embed the curved space being considered in some higher dimensional flat space, an alternative procedure which is only practical when the dimensionality is low.

The complete specification of a spacetime point requires one more coordinate which we shall denote by t, the physical interpretation of which will emerge as we study the spacetime geometry. Thus, if we adapt the choice of orthonormal tetrad to the coordinates r, θ, φ, t, the basic one-forms can be expressed in the form

$$\begin{aligned} e^1 &= e^{\frac{1}{2}\lambda(r,t)}\, dr\ , \\ e^2 &= r\, d\theta\ , \\ e^3 &= r\sin\theta\, d\varphi\ , \\ e^4 &= e^{\frac{1}{2}\nu(r,t)}\, dt\ , \end{aligned}$$

where the two functions $\lambda(r,t)$ and $\nu(r,t)$ have been included to account for possible deviations from flatness.

Applying the exterior derivative operator d, we obtain

$$\begin{aligned} de^1 &= \frac{1}{2} e^{\frac{1}{2}\lambda}\lambda_t\, dt\, dr\ , \\ de^2 &= dr\, d\theta\ , \\ de^3 &= \sin\theta\, dr\, d\varphi + r\cos\theta\, d\theta\, d\varphi\ , \\ de^4 &= \frac{1}{2} e^{\frac{1}{2}\nu}\nu_r\, dr\, dt\ , \end{aligned}$$

[1] Some people employ a curvature tensor that is the negative of ours. Perhaps their motivation lies in the desire to see a positive number rather than a negative number as the result of this calculation.

where λ_t stands for the t-derivative of λ, and so on. Expressing the two-forms in terms of the basis $\{e^a\}$, we obtain

$$\begin{aligned} de^1 &= \frac{1}{2}e^{-\frac{1}{2}\nu}\lambda_t\, e^4e^1 \ , \\ de^2 &= e^{-\frac{1}{2}\lambda}r^{-1}\, e^1e^2 \ , \\ de^3 &= e^{-\frac{1}{2}\lambda}r^{-1}\, e^1e^3 + r^{-1}\cot\theta\, e^2e^3 \ , \\ de^4 &= \frac{1}{2}e^{-\frac{1}{2}\lambda}\nu_r\, e^1e^4 \ . \end{aligned}$$

From these four relations we have to obtain the connection $\Gamma_a{}^b$. In four dimensions it is not quite as trivial as in the two dimensional case considered earlier. Nevertheless, the equations to be solved are linear algebraic relations, so they can be solved.

One approach which I have often used when dealing with an orthonormal tetrad is the following one:

(1) I first evaluate the two-forms

$$de^a \ (a = 1, 2, 3, 4),$$

which can be expressed in the form

$$de^a = \frac{1}{2}f_{bc}{}^a\, e^be^c.$$

(2) Then I evaluate the three-form

$$e_a\, de^a,$$

which can be expressed in the form

$$e_a\, de^a = \frac{1}{6}h_{bcd}\, e^be^ce^d.$$

(3) The Ricci rotation matrix is then given by

$$\Omega_{bc} = \{f_{bcd} - \frac{1}{2}h_{bcd}\}\, e^d.$$

Let's apply that approach here. In this application it is clear that $h_{bcd} = 0$, so it is only necessary to compute f_{bcd}. The latter quantities may be read off by inspection:

$$f_{41}{}^1 = \tfrac{1}{2}e^{-\frac{1}{2}\nu}\lambda_t \ , \ f_{12}{}^2 = e^{-\frac{1}{2}\lambda}r^{-1} \ , \ f_{13}{}^3 = e^{-\frac{1}{2}\lambda}r^{-1} \ ,$$
$$f_{23}{}^3 = r^{-1}\cot\theta \ , \ f_{14}{}^4 = \tfrac{1}{2}e^{-\frac{1}{2}\lambda}\nu_r \ .$$

Taking into account the skew symmetry, we easily obtain the Ricci rotation matrix

$$\begin{aligned}
\Omega_{12} &= e^{-\frac{1}{2}\lambda} r^{-1}\, e^2 \,, \\
\Omega_{13} &= e^{-\frac{1}{2}\lambda} r^{-1}\, e^3 \,, \\
\Omega_{14} &= -\frac{1}{2} e^{-\frac{1}{2}\nu} \lambda_t\, e^1 - \frac{1}{2} e^{-\frac{1}{2}\lambda} \nu_r\, e^4 \,, \\
\Omega_{23} &= r^{-1} \cot\theta\, e^3 \,, \\
\Omega_{24} &= 0 \,, \\
\Omega_{34} &= 0 \,.
\end{aligned}$$

Ex. 16 *Check that the* ${\Gamma_a}^b$ *constructed from this solution does indeed satisfy the relations* $de^b = e^a {\Gamma_a}^b$.

Once one has determined the Ricci rotation matrix, and hence the connection ${\Gamma_a}^c$, one can proceed to evaluate the Riemann tensor components. This entails calculating the two-form $d{\Gamma_a}^c - {\Gamma_a}^b {\Gamma_b}^c$. In connection with the evaluation of the first term, note that we may easily express Ω_{ab} in terms of coordinate differentials:

$$\begin{aligned}
\Omega_{12} &= e^{-\frac{1}{2}\lambda}\, d\theta \,, \\
\Omega_{13} &= e^{-\frac{1}{2}\lambda} \sin\theta\, d\varphi \,, \\
\Omega_{14} &= -\frac{1}{2} e^{\frac{1}{2}(\lambda-\nu)} \lambda_t\, dr - \frac{1}{2} e^{\frac{1}{2}(\nu-\lambda)} \nu_r\, dt \,, \\
\Omega_{23} &= \cos\theta\, d\varphi \,, \\
\Omega_{24} &= 0 \,, \\
\Omega_{34} &= 0 \,.
\end{aligned}$$

It is not very difficult to evaluate the differential of any of these one-forms and thus complete the calculation. Even here, however, the tediousness of curvature tensor calculations is quite apparent. You should bear in mind that the Schwarzschild problem is, by modern standards, an extremely simple problem. For this reason, I generally perform such calculations using a digital computer. Except for extremely complex calculations, even a microcomputer can be of enormous help.

Using my own symbolic manipulation program, I obtained the following expressions for the non-vanishing components of the Ricci tensor R_{ab}, relative to the chosen orthonormal basis:

$$R_{14} = R_{41} = -r^{-1} e^{\frac{1}{2}(\lambda+\nu)} \lambda_t \,,$$

$$
\begin{aligned}
R_{11} &= [\frac{1}{2}(\nu_{rr} - \lambda_{tt}) + \frac{1}{4}\nu_r(\nu_r - \lambda_r)]e^{-\lambda} \\
&\quad + \frac{1}{4}\lambda_t(\nu_t - \lambda_t)e^{-\nu} - r^{-1}\lambda_r e^{-\lambda} \ , \\
R_{44} &= -[\frac{1}{2}(\nu_{rr} - \lambda_{tt}) + \frac{1}{4}\nu_r(\nu_r - \lambda_r)]e^{-\lambda} \\
&\quad - \frac{1}{4}\lambda_t(\nu_t - \lambda_t)e^{-\nu} - r^{-1}\nu_r e^{-\lambda} \ , \\
R_{22} = R_{33} &= -r^{-2}(1 - e^{-\lambda}) + \frac{1}{2}r^{-1}(\nu_r - \lambda_r)e^{-\lambda} \ .
\end{aligned}
$$

In the region outside the material source of the field, all components of the Ricci tensor must vanish.

- In particular, the equation $R_{14} = 0$ tells one immediately that the field λ must be independent of t.

- Secondly, the equation $R_{11} + R_{44} = 0$ tells one that $\lambda_r + \nu_r = 0$. Hence $\lambda + \nu$ is a function of t alone. A careful study of what happens when one carries out a transformation $t \to t' = t'(t)$ shows that one can, without loss of generality, require $\lambda + \nu = 0$. That is,

$$
e^{-\lambda} = e^{\nu} = \text{ function of } r \text{ only.}
$$

- The equation $R_{22} = 0$ may be written in the form

$$
R_{22} = -r^{-2}\frac{d}{dr}[r(1 - e^{-\lambda})] = 0.
$$

  Hence $r(1 - e^{-\lambda}) = \text{const} =: r_0$.

- Finally, one can check that the solution

$$
e^{-\lambda} = e^{\nu} = 1 - \frac{r_0}{r}
$$

  also satisfies the equation $R_{11} - R_{44} = 0$. This is the result which Schwarzschild obtained.

In conclusion, when one has any spherically symmetric source, coordinates r, θ, φ, t may be introduced such that the spacetime curvature outside that source is t-independent as well as θ and φ-independent. The final expressions for the orthonormal tetrad e^a $(a = 1, 2, 3, 4)$ are given by

$$\begin{aligned} e^1 &= \frac{dr}{\sqrt{1 - \frac{r_0}{r}}} , \\ e^2 &= r\, d\theta , \\ e^3 &= r \sin\theta\, d\varphi , \\ e^4 &= \sqrt{1 - \frac{r_0}{r}}\, dt . \end{aligned}$$

The case $r_0 = 0$ corresponds to flat spacetime.

How, in principle, might one determine the constant r_0? One might consider two identical clocks, one at $r = \infty$ and one at $r = R$, some finite value of r. Consider two successive ticks of the clock at R. This corresponds to a proper time interval

$$\Delta\tau = \sqrt{1 - \frac{r_0}{R}}\, \Delta t.$$

Thus, the two events (the two ticks of the clock) correspond to t labels differing by

$$\Delta t = \frac{\Delta\tau}{\sqrt{1 - \frac{r_0}{R}}}.$$

The clock at infinity measures Δt directly, for proper time is t time there. By comparing the two clocks, it is possible (in principle) to determine $\sqrt{1 - \frac{r_0}{R}}$ and hence the ratio r_0/R.

For most masses encountered in practice, this procedure is completely impractical. However, in the case of dense white dwarf stars, one can actually detect a difference between the rate of ticking of clocks on the surface of such stars and the rate of ticking of clocks here on earth. What one actually does is to compare the spectral lines of light emitted by atoms on the surface of the white dwarf star with the spectral lines of light emitted by similar atoms on the surface of the earth. What is observed is a shifting of the spectral lines toward the red end of the spectrum. From the amount of the shift one can infer a value for the ratio $r_0/R \approx 2\Delta\lambda/\lambda$. Numbers such as $\frac{1}{1000}$ are not uncommon. Since white dwarf stars have radii similar to that of the earth, one can infer that for a white dwarf star r_0 is of the order of a few kilometers.

We shall want to establish a firmer connection between the parameter r_0 and the mass of the central body which gives rise to the spacetime curvature. For this we shall have to study the Kepler problem, for it is Kepler's third law that is customarily used to estimate the masses of stars that are members of double star systems.

There is another aspect of the Schwarzschild solution that bears study. For $r_0 \neq 0$ it is clear that the coordinate r must be restricted to values $r > r_0$. At this point it is unclear whether $r = r_0$ represents just the edge of a particular coordinate chart, or it represents a place where the spacetime geometry itself becomes singular. As long as the material source of the curvature occupies a larger region, there is no problem, but if one considers sources of smaller extent the answer to this question becomes very significant indeed.

The Kepler Problem

In the early days of general relativity, it was not appreciated that the motion of bodies is determined by the field equations themselves; as in classical mechanics, equations of motion were postulated independently. It was assumed that a body subjected to no external forces would always travel a *world line* such that the elapsed proper time is a MAXIMUM. (This turns out to be the same as the STRAIGHTEST path in spacetime.) We can set this up as a variational principle involving the proper time, $\Delta\tau$, where

$$d\tau^2 = -\frac{1}{c^2}\left[\frac{dr^2}{1-\frac{r_0}{r}} + r^2(d\theta^2 + \sin^2\theta\; d\varphi^2)\right] + (1-\frac{r_0}{r})dt^2 \;.$$

The way this is usually carried out is by maximizing the integral

$$I := \int d\tau \left\{-\frac{1}{c^2}\left[\frac{\dot{r}^2}{1-\frac{r_0}{r}} + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2)\right] + (1-\frac{r_0}{r})\dot{t}^2\right\} \;,$$

which is essentially equivalent to maximizing the integral

$$\int d\tau \sqrt{-\frac{1}{c^2}\left[\frac{\dot{r}^2}{1-\frac{r_0}{r}} + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2)\right] + (1-\frac{r_0}{r})\dot{t}^2} \;.$$

Here, the objective is to derive the coordinates of the moving body as a function of proper time, i.e., $r(\tau), \theta(\tau), \varphi(\tau)$ and $t(\tau)$. The effective Lagrangian

$$L := (1-\frac{r_0}{r})\dot{t}^2 - \frac{1}{c^2}\left[\frac{\dot{r}^2}{1-\frac{r_0}{r}} + r^2(\dot{\theta}^2 + \sin^2\theta\dot{\varphi}^2)\right]$$

is a function of the generalized coordinates r, θ, φ, t and the corresponding generalized velocities $\dot{r}, \dot{\theta}, \dot{\varphi}, \dot{t}$. The equations of motion are nothing but the Lagrange equations

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \ .$$

Since φ and t are *cyclic* coordinates, i.e., they don't occur in the Lagrangian, there are two obvious constants of the motion; namely,

$$\begin{aligned} \frac{\partial L}{\partial \dot{\varphi}} &= -\frac{2}{c^2} r^2 \sin^2\theta \dot{\varphi} \ , \\ \frac{\partial L}{\partial \dot{t}} &= 2(1 - \frac{r_0}{r})\dot{t} \ . \end{aligned}$$

The first of these constants of the motion has a simple interpretation. Recall that in Newtonian physics, the angular momentum of a body of mass m in a central field is conserved. In terms of spherical polar coordinates one has the Newtonian formula

$$\ell = mr^2 \sin^2\theta \dot{\varphi}$$

for the conserved angular momentum, where the dot signifies the derivative with respect to the time t in this formula. We see that the only modification in relativity theory is to replace the time-derivative by a proper time-derivative. We shall continue to refer to ℓ as the angular momentum of the moving body.

The second constant of the motion is related to the energy of the moving body, so we shall use the symbol E, writing

$$E = mc^2(1 - \frac{r_0}{r})\dot{t} \ .$$

This reduces to the special relativity formula $E = mc^2\dot{t}$ when $r_0 \to 0$.

Let's look at the Newtonian limit of this conservation of energy equation. For this purpose, we shall introduce the symbol

$$v^2 := \frac{(\frac{dr}{dt})^2}{1 - \frac{r_0}{r}} + r^2 \left\{ (\frac{d\theta}{dt})^2 + \sin^2\theta (\frac{d\varphi}{dt})^2 \right\} \ .$$

It is something like the square of the speed of the body. We then have

$$\left(\frac{d\tau}{dt}\right)^2 = 1 - \frac{v^2}{c^2} - \frac{r_0}{r},$$

Therefore,

$$\dot{t} = \frac{1}{\sqrt{1 - \frac{v^2}{c^2} - \frac{r_0}{r}}}.$$

Substituting this into the equation for E, we obtain

$$\begin{aligned} E &= \frac{mc^2(1 - \frac{r_0}{r})}{\sqrt{1 - \frac{v^2}{c^2} - \frac{r_0}{r}}} \\ &\approx mc^2 + \frac{1}{2}mv^2 - \frac{r_0 mc^2}{2r} . \end{aligned}$$

The first term will be recognized as the rest energy of the body, the second as the kinetic energy, and the third is clearly the gravitational potential energy; namely, $-GMm/r$, where G is the universal constant of gravitation. Thus, we have the identification of the parameter r_0:

$$r_0 = 2\frac{GM}{c^2} .$$

For the Sun, the value of r_0 is approximately 2.94 kilometers, while the radius R of the Sun is approximately 697,000 kilometers. Hence, for the Sun, $r_0/R \approx 4.2 \times 10^{-6}$.

Since the distortion of spacetime caused by the Sun is so minute, we can use perturbation theory to work out the planetary orbits, using the Newtonian orbits as the lowest approximation. To proceed, we must work out the remaining equations of motion. Consider, for example, the Lagrange equation corresponding to the generalized coordinate θ.

Ex. 17 *Show that*

$$\frac{d}{d\tau}(r^2\dot{\theta}) - r^2 \sin\theta \cos\theta \dot{\varphi}^2 = 0 ,$$

which aside from the replacement of t by the proper time τ is identical to the corresponding Newtonian equation.

This equation tells us that a particle initially moving within the plane $\theta = \frac{\pi}{2}$ will continue to move within that plane. Since the planetary motion is planar, it is obviously advantageous to select the polar axis so that $\theta = \frac{\pi}{2}$. With this choice we have the simplified equation

$$\ell = mr^2\dot{\varphi}$$

for the conserved angular momentum.

Ex. 18 *Show that the Lagrange equation corresponding to the generalized coordinate r turns out (with the specialized choice of polar axis) to be*

$$\ddot{r} - r\dot{\varphi}^2 = -\frac{GM}{r^2} - 3\frac{GM}{c^2}\dot{\varphi}^2.$$

Except for the replacement of t by the proper time τ and the appearance of the second term on the right, this equation is the same as the Newtonian equation.

A first integral of this equation is obtained in relativity theory the same way as it is obtained in Newtonian theory.

Ex. 19 *Multiply the preceding equation by $\dot{r}$ and integrate with respect to τ to obtain a relation that can be cast into the form*

$$\left(\frac{E}{mc^2}\right)^2 = \dot{r}^2 + \left(1 - \frac{2GM/c^2}{r}\right)\left(1 + \frac{\ell^2}{r^2}\right). \tag{9.1}$$

A more direct route to this result is by setting the conserved Lagrangian equal to unity. That the Lagrangian is a constant of the motion follows from two facts:

(1) Since the Lagrangian is homogeneous quadratic in the generalized velocities, the Lagrangian and the Hamiltonian $H(p, q, \tau) := \sum(p\dot{q}) - L(q, \dot{q}, \tau)$ are equal.

(2) Since it does not depend explicitly upon the proper time τ, the Hamiltonian is a constant of the motion.

Non-axial timelike geodesics

Because of the spherical symmetry of the Schwarzschild solution, one can always select the coordinate system so that initially the path of a body that is not moving purely radially can be described by $\theta = \pi/2$ and $\dot{\theta} = 0$. In this case, using the equation for $\frac{d}{d\tau}(r^2\dot{\theta})$ that you derived in one of the exercises, you can easily prove that θ will remain equal to $\pi/2$ forever; that is, the body will remain in the equatorial plane.

Equation (9.1) can be expressed in the form

$$E_{\text{class}} = \frac{1}{2}m\dot{r}^2 + V_{\text{eff}}(r),$$

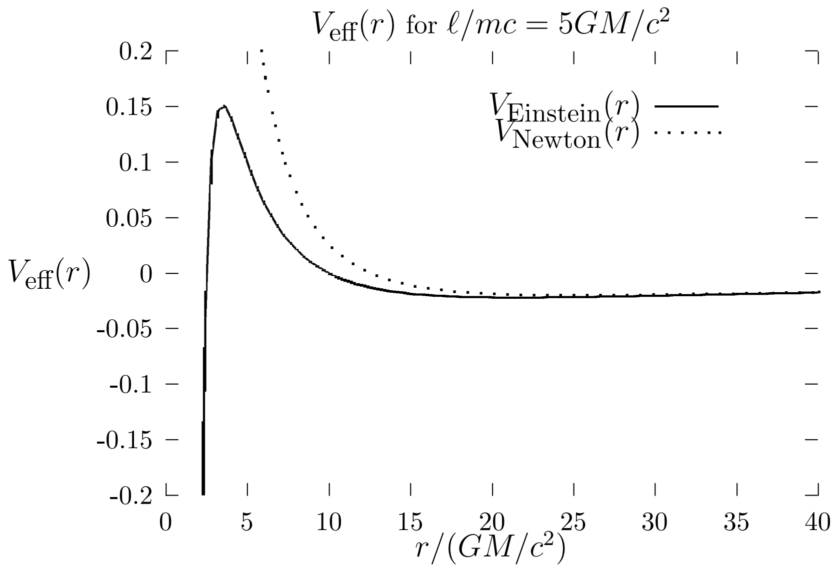

Figure 1: Comparison of the classical and relativistic effective potentials $V_{\text{eff}}(r)$ when $\ell/mc = 5GM/c^2$.

where

$$V_{\text{eff}} := -\frac{GMm}{r} + \frac{\ell^2}{2mr^2}\left(1 - \frac{2GM/c^2}{r}\right)$$

is an *effective potential*, while

$$E_{\text{class}} := \frac{1}{2}mc^2\left[(E/mc^2)^2 - 1\right], \quad \ell = mr^2\dot{\varphi}.$$

Except for the factor $1 - (2GM/c^2)/r$ in the second term, the expression for $V_{\text{eff}}(r)$ is identical to the corresponding Newtonian expression. In the accompanying figure is shown the effective potential V_{eff} both with and without the relativistic factor $1 - (2GM/c^2)/r$. The value of such a diagram lies in the fact that whenever $E_{\text{class}} \geq V_{\text{eff}}(r)$ the radial velocity $\dot{r}$ is real, and whenever $E_{\text{class}} < V_{\text{eff}}(r)$ the radial velocity is imaginary. Where the horizontal line E_{class} intersects V_{eff}, the orbital turning points occur.

A local minimum, $\approx -0.02mc^2$, of the classical effective potential energy occurs at $r = 25GM/c^2$. The total energy has to be at least this great. For $E_{\text{class}} \approx -0.02mc^2$ the orbit is a stable circle of radius $r = 25GM/c^2$. For total energies that are greater than this minimum, but which are still less than zero, the orbits are ellipses, while if the total energy is positive, the orbits are hyperbolas. The special case of zero total energy corresponds to a parabolic orbit.

In the relativistic case the minimum in the effective potential moves to $r \approx 21.5GM/c^2$, and a local maximum, $\approx +0.15mc^2$ appears at $r \approx 3.5GM/c^2$. Thus, an energy $E_{\text{class}} \approx -0.02mc^2$ corresponds to a stable circular orbit of radius $r \approx 21.5GM/c^2$. The orbits for E_{class} between the minimum allowed value and zero are again bound orbits analogous to the elliptical orbits of the Kepler problem. The orbits with positive total energy less than $+0.15mc^2$ are unbounded orbits analogous to the hyperbolic orbits of the Kepler problem, and the zero energy orbit is the analog of the parabolic orbit of the Kepler problem. However, there is a new class of orbits, those for which the total energy exceeds $+0.15mc^2$. These orbits extend to small values of r. At $E_{\text{class}} \approx +0.15mc^2$ there exists an unstable circular orbit of radius $r \approx 3.5GM/c^2$. If a body is in such a circular orbit, any slight perturbation will cause it either to accelerate out toward $r = \infty$ or to accelerate in toward small values of r.

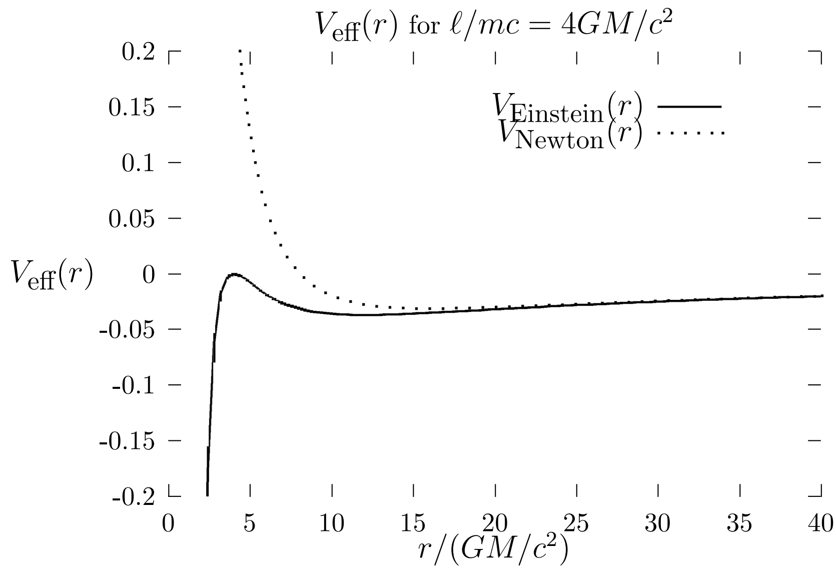

Figure 2: Comparison of the classical and relativistic effective potentials $V_{\text{eff}}(r)$ when $\ell/mc = 4GM/c^2$.

One case of special interest occurs when $\ell/mc = \pm 4.0GM/c^2$. Here, as is seen in the second figure, the local maximum in the relativistic case moves to $r = 4GM/c^2$ and the value of V_{eff} there is zero. All the orbits with

$E_{\text{class}} > 0$, analogous to the hyperbolic orbits of the Kepler problem, extend to small values of r. A body in the unstable zero-energy circular orbit at $r = 4GM/c^2$, if perturbed, will either accelerate out toward $r = \infty$ or in toward the event horizon. In the former case, the orbit at large values of r will closely resemble a parabolic orbit.

Precession of the Perihelion

In the case of the Schwarzschild spacetime that is associated with the Sun, $GM/c^2 \approx 1.47$ km, while the radius of the Sun is $R \approx 600,000$ km. Thus, all the planets are far out in the region where Newtonian physics applies. Nevertheless, there are minute effects of relativity that show up even here.

Our immediate objective is to derive an equation for the orbit of a planet encircling a central mass M. In Newtonian physics the orbit of such a planet is an ellipse with the central mass at one of the foci of the ellipse. The equation of the ellipse is usually expressed in the form

$$\frac{d^2u}{d\varphi^2} + u = GM\left(\frac{m}{\ell}\right)^2 ,$$

where $u = 1/r$. We may derive the relativistic analog of this equation from Eq. (9.1), by dividing the latter equation by $r^4\dot{\varphi}^2 = (\ell/m)^2$, thus obtaining

$$\left(\frac{du}{d\varphi}\right)^2 = \left(\frac{mc}{\ell}\right)^2 [(\frac{E}{mc^2})^2 - 1 + \frac{2GM}{c^2}u] - u^2(1 - \frac{2GM}{c^2}u) .$$

Differentiating with respect to φ, we obtain the desired generalization of the classical formula,

$$\frac{d^2u}{d\varphi^2} + u = GM\left(\frac{m}{\ell}\right)^2 + \frac{3GM}{c^2}u^2 . \qquad (9.2)$$

It is the second term on the right side of the orbital equation which gives rise to the precession of the perihelion (point of closest approach to the Sun) of planets such as Mercury.

Recalling that all planetary orbits about the Sun are nearly circular, we may write $u = R^{-1} + \delta u$, where R is the approximate radius of the orbit. Neglecting terms of order $(\delta u)^2$, we obtain

$$\frac{d^2(\delta u)}{d\varphi^2} + [1 - \frac{6GM}{c^2R}]\delta u = 0 ,$$

showing that the angle (in radians) between one perihelion and the next is equal to $2\pi/\sqrt{1-\frac{6GM}{c^2R}}$, or approximately $2\pi(1+\frac{3GM}{c^2R})$. Using the known values of GM and R together with the orbital period of Mercury, one can find the angular precession per century; namely, 43 seconds of arc.

The observed precession of the perihelion of Mercury is about 5600 seconds of arc, of which about 5557 seconds of arc can be attributed to such things as the perturbation of Mercury's orbit by the other planets and the rotation of the Sun. The discrepancy of 43 seconds of arc had been known long before Einstein finally explained it. In fact, many attempts had been made to discover a tenth planet somewhere between Mercury and the Sun, as a possible explanation of the discrepancy.

Bending of Light Rays

Another of the traditional tests of general relativity concerned the bending of the paths of light rays by the Sun's gravitational field. Here one is concerned with a null geodesic rather than a timelike geodesic. That is, the interval between two events connected by a light ray vanishes.

It is unnecessary to go through the entire orbital calculation again, for we may use the equation which we derived for the orbit of a planet, providing we take an appropriate limit of that equation. If $m \to 0$ in Eq. (9.2), we obtain the equation of a light ray; namely,

$$\frac{d^2u}{d\varphi^2}+u=\frac{3GM}{c^2}u^2\ . \tag{9.3}$$

In the absence of the relativistic term on the right side of Eq. (9.3), light travels in straight lines, i.e.,

$$u = R^{-1}\cos\varphi$$

or

$$r\cos\varphi = R,$$

where R is a constant signifying the distance of closest approach to the center of the Sun.

One can use standard methods of perturbation theory to get an approximate solution of the equation with the relativistic term present. Into the right hand side of Eq. (9.3) one substitutes the approximate solution

$u = R^{-1}\cos\varphi$. Then one solves the resulting inhomogeneous second order differential equation, with the result

$$u = R^{-1}\cos\varphi + \frac{GM}{c^2R^2}(\cos^2\varphi + 2\sin^2\varphi)\ .$$

If there were no relativistic correction, the angle φ would run from $-\pi/2$ to $+\pi/2$. With the correction φ runs from $-(\pi/2+\epsilon)$ to $+(\pi/2+\epsilon)$, where ϵ is a very small angle. The full angle of deflection of a light ray is given by 2ϵ.

Setting $u = 0$, and treating ϵ as small, we obtain the quadratic equation

$$\epsilon^2 - \frac{c^2R}{GM}\epsilon - 2 = 0$$

for ϵ. Thus,

$$\epsilon = -\frac{c^2R}{2GM}\left\{1 \pm \sqrt{1+8(GM/c^2R)^2}\right\}\ .$$

As one can see by expanding the square root as a series in GM/c^2R, the small positive root is approximately equal to $2(GM/c^2R)$. Therefore,

$$\text{Angle of deflection in radians } = \frac{4GM}{c^2R}\ ,$$

where M is the mass of the Sun and (for grazing rays) R is the radius of the Sun. The numerical values are $GM/c^2 = 1.47$ kilometers and $R = 697{,}000$ kilometers. This corresponds to a deflection of 8.4×10^{-6} radians, or 1.75 seconds of arc, or about one thousandth of the angle subtended by the Sun's image as viewed from the earth.

Eddington led an expedition to photograph the star field during a solar eclipse in 1919. The observations were very difficult, and the results not very conclusive. However, it was felt that the observations favored general relativity over the competing theories.

Not long ago a striking confirmation of this prediction of Einstein's theory was found in photographs of distant galaxies, where a whole galaxy was responsible for bending light rays emitted by stars in a still more distant galaxy, resulting in multiple images of that more distant galaxy.