Chapter 11

Spherically symmetric gravitational field of isolated objects.

In this chapter we will deal with gravitational fields of spherically symmetric isolated objects, such as stars and black holes.[9] Just as in Newton's theory, this case is the easiest to visualise and to handle mathematically, and is a useful simplified model of real objects.

11.1 The curvature coordinates.

In Sec. 8.8 we showed that the metric form of a spherically symmetric spacetime, in a suitably limited class of coordinate systems, has the form

$$\mathrm{d}s^2 = \alpha(t,r)\mathrm{d}t^2 + 2\beta(t,r)\mathrm{d}t\mathrm{d}r + \gamma(t,r)\mathrm{d}r^2 + \delta(t,r)\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2\right). \tag{11.1}$$

The values of ϑ and φ fix a point on a 2-sphere, the spheres are the orbits of the symmetry group, t and r are arbitrary coordinates whose values are constant on the spheres. The general form of (11.1) and the formulae for the Killing fields (8.38) do not change under transformations that are composites of the following two classes:

- 1. The isometries generated by the Killing fields via (8.15) – (8.16). These transformations preserve not just the form (11.1), but also the values of the functions α, β, γ and δ. An example is a rotation by the angle α_0 around the axis $\{\vartheta = \pi/2, \varphi = 0\}$ (this axis passes through the center of the sphere and through the point $\varphi = 0$ on the equator). In the (ϑ, φ)-coordinates, such a rotation is given by[10]

$$\begin{aligned} \vartheta &= \arctan\left[\frac{\sqrt{\sin^2\vartheta'\cos^2\varphi' + (\cos\alpha_0\sin\vartheta'\sin\varphi' + \sin\alpha_0\cos\vartheta')^2}}{\cos\alpha_0\cos\vartheta' - \sin\alpha_0\sin\vartheta'\sin\varphi'}\right], \\ \varphi &= \arctan\left(\cos\alpha_0\tan\varphi' + \sin\alpha_0\frac{\cot\vartheta'}{\cos\varphi'}\right). \end{aligned} \tag{11.2}$$

[9] Spherically symmetric solutions of Einstein's equations are also used in cosmology; these applications will be discussed separately in chapter 15.

[10] See the hint in Exercise 1 at the end of this chapter.

- 2. The transformations that preserve the general form of (11.1), but mix the functions α, β, γ and δ. They are:

$$t = F(t', r'), \qquad r = G(t', r'), \tag{11.3}$$

where F and G are arbitrary functions subject only to $\partial(F, G)/\partial(t', r') \neq 0$. After such a transformation the functions α, β and γ change as follows

$$\widetilde{\alpha} = \alpha F,_{t'}{}^2 + 2\beta F,_{t'} G,_{t'} + \gamma G,_{t'}{}^2, \tag{11.4}$$
$$\widetilde{\beta} = \alpha F,_{t'} F,_{r'} + \beta \left(F,_{t'} G,_{r'} + F,_{r'} G,_{t'}\right) + \gamma G,_{t'} G,_{r'}, \tag{11.5}$$
$$\widetilde{\gamma} = \alpha F,_{r'}{}^2 + 2\beta F,_{r'} G,_{r'} + \gamma G,_{r'}{}^2, \tag{11.6}$$

while δ transforms like a scalar, $\widetilde{\delta}(t', r') = \delta(F, G)$.

Thus, using F and G, one can manipulate the values of α, β and γ, but not of δ. In the signature $(+ - - -)$ assumed throughout, the determinant of the metric must be negative

$$\alpha\gamma - \beta^2 < 0. \tag{11.7}$$

Because of (11.7), the equation $\widetilde{\alpha} = 0$ can always be solved for $F,_{t'}$, and the solution is

$$F,_{t'} = \frac{1}{\alpha}\left(-\beta \pm \sqrt{\beta^2 - \alpha\gamma}\right) G,_{t'}\,. \tag{11.8}$$

We can thus take an arbitrary G and choose F to obey (11.8), achieving $\widetilde{\alpha} = 0$. Consequently, we lose no generality when we begin with the coordinates in which $\alpha = 0$ already. With $\alpha = 0$, we see from (11.5) that we can achieve $\widetilde{\beta} = 0$ by solving the equation

$$\beta G,_{r'} F,_{t'} + \beta G,_{t'} F,_{r'} + \gamma G,_{t'} G,_{r'} = 0 \tag{11.9}$$

(but the new $\widetilde{\alpha}$ will be nonzero). With any G, (11.9) is a linear inhomogeneous partial differential equation for F, for which standard methods of solving exist.

We achieved $\beta = 0$ in two steps, and in both G was arbitrary. Hence, we can now choose G so as to achieve some further simplification.

This is where most textbooks make a mistake. They say: we choose G so that $\widetilde{\delta} = -r'^2$. But δ is a scalar with respect to (11.3): if $\delta =$ constant, then $\widetilde{\delta} = \delta =$ constant, and no condition can be imposed on it. Thus, further simplification of (11.1) requires care.

Suppose we wish to transform the coordinates of (11.1) so that

$$\widetilde{\beta} = 0, \qquad \widetilde{\delta} = -r'^2. \tag{11.10}$$

When is this possible? From the second of (11.10) we have

$$r' = \sqrt{-\delta(F(t', r'), G(t', r'))}. \tag{11.11}$$

Differentiating this by t' and by r' we get

$$\delta,_F F,_{t'} + \delta,_G G,_{t'} = 0, \tag{11.12}$$
$$\delta,_F F,_{r'} + \delta,_G G,_{r'} = -2\sqrt{-\delta} = -2r'. \tag{11.13}$$

While solving this set for $(G,_{t'}, G,_{r'})$, two subcases must be considered separately

- Ia. $\delta,_G \neq 0$. Then, substituting for $G,_{t'}$ from (11.12) in (11.4) we get

$$\widetilde{\alpha} = \frac{{F,_{t'}}^2}{{\delta,_G}^2}\left(\alpha{\delta,_G}^2 - 2\beta\delta,_F\,\delta,_G + \gamma{\delta,_F}^2\right). \tag{11.14}$$

  We want t' to be the time coordinate, so we must require $\widetilde{\alpha} > 0$, which means

$$\alpha{\delta,_G}^2 - 2\beta\delta,_F\,\delta,_G + \gamma{\delta,_F}^2 > 0. \tag{11.15}$$

  This is easily verified to be equivalent to

$$g^{\alpha\beta}\delta,_\alpha\,\delta,_\beta < 0, \tag{11.16}$$

  which means that the gradient of δ must be a spacelike vector.

- Ib. $\delta,_G = 0$. Then, from (11.12), either $\delta,_F = 0$, which means $\delta =$ constant (this case is discussed below), or $\delta,_F \neq 0$ and $F,_{t'} = 0$. In this second case, (11.4) implies

$$\widetilde{\alpha} = \gamma{G,_{t'}}^2. \tag{11.17}$$

  Thus, $\widetilde{\alpha} > 0$ only if $\gamma > 0$, which implies (11.16) again. (Equation (11.17) shows that in this case r is the time-coordinate before the transformation.)

Hence, the following lemma applies:

Lemma 11.1 (Case I) $\widetilde{\delta} = -r'^2$ *can be fulfilled only if* $\delta,_\mu$ *is a spacelike vector.*

But there are three other cases in which the set (11.12) – (11.13) cannot be solved at all. Proofs of the remaining lemmas are left as an exercise for the reader.

Lemma 11.2 (Case II) *If* $\delta,_\mu$ *is a timelike vector, then one can choose the coordinates of (11.1) so that*

$$\widetilde{\beta} = 0, \qquad \widetilde{\delta} = -t'^2. \tag{11.18}$$

Lemma 11.3 (Case III) *If* $\delta,_\mu$ *is a nontrivial null vector (*$g^{\mu\nu}\delta,_\mu\,\delta,_\nu = 0$ *but* $\delta,_\mu \neq 0$*), then one can choose the coordinates of (11.1) so that* $\widetilde{\delta} = -r'^2$*, but this automatically implies* $\widetilde{\alpha} = 0$*. Then it is not possible to simultaneously achieve* $\widetilde{\beta} = 0$ *because of (11.7), but it is possible to have in addition* $\widetilde{\gamma} = 0$*.*

(Case IV) When δ is constant, no condition at all can be imposed on it. The spacetime is then a Cartesian product of a sphere of radius $l = \sqrt{-\delta}$ and a 2-dimensional surface with the metric $\alpha(t,r)\mathrm{d}t^2 + 2\beta(t,r)\mathrm{d}t\mathrm{d}r + \gamma(t,r)\mathrm{d}r^2$.

The existence of regions of spacetimes with $\delta,_\mu$ being spacelike or timelike was first noted by Novikov [60], who called them *R-regions* and *T-regions*, respectively.

[60] I. D. Novikov, *Vestn. Mosk. Univ.* no. 6, 66 (1962).

Case II is ignored by most textbooks, even by those that subsequently describe the extension of the Schwarzschild manifold (see Sec. 11.7). Cases III and IV are ignored almost universally, with the notable exception of Ref. [33]. The situation of these cases versus the Einstein equations is this:

For the vacuum equations with zero cosmological constant, the cases $g^{\alpha\beta}\delta_{,\alpha}\,\delta_{,\beta} < 0$ and $g^{\alpha\beta}\delta_{,\alpha}\,\delta_{,\beta} > 0$ cover different parts of the same manifold, as we will see in Sec. 11.7.

Case III leads to a contradictory set of equations in pure vacuum [61], but the solution of the vacuum Einstein–Maxwell equations with this geometry exists ([33], eq. (15.18)). Case IV leads to a contradiction when $T_{\alpha\beta} = 0 = \Lambda$. The solution with $T_{\alpha\beta} = 0 \neq \Lambda$ was found by Nariai [62, 63]. The solutions of the Einstein–Maxwell equations with this geometry and with the source being the vacuum electrostatic field were found by Bertotti [64] and Robinson [65], they are compared with the Nariai solution in Ref. [66].

In this chapter, we will deal only with the case $g^{\alpha\beta}\delta_{,\alpha}\,\delta_{,\beta} < 0$. We choose the coordinates so that (11.10) applies and denote $\widetilde{\alpha} = \mathrm{e}^{2\nu(t,r)}$, $\widetilde{\gamma} = -\mathrm{e}^{2\mu(t,r)}$. Then

$$\mathrm{d}s^2 = \mathrm{e}^{2\nu(t,r)}\mathrm{d}t^2 - \mathrm{e}^{2\mu(t,r)}\mathrm{d}r^2 - r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2\right). \tag{11.19}$$

This form of the metric still allows a subgroup of the transformations (11.3). We have $\delta = -r^2$ before and $\widetilde{\delta} = -r'^2$ after the transformation. Since $\delta = \widetilde{\delta}$, this means $r = r'$, and, consequently, $G_{,t'} = 0$, $G_{,r'} = 1$ in (11.4) – (11.6). We also require that $\beta = 0$ before and $\widetilde{\beta} = 0$ after the transformation, so from (11.5)

$$\alpha F_{,t'}\,F_{,r'} = 0. \tag{11.20}$$

But $\alpha \neq 0$, or else the metric would be singular. Also $F_{,t'} \neq 0$ because otherwise, with $G_{,t'} = 0$, the transformation would be singular. Hence, (11.20) implies $F_{,r'} = 0$, i.e. the transformations still allowed are $t = F(t')$, where F is an arbitrary function.

The coordinates in which the metric has the form (11.19) are called **curvature coordinates** because in them the radial coordinate r is connected with the curvature $\mathcal{R}$ of the symmetry orbits in the same way as in a flat space, $\mathcal{R} = 1/r^2$.

11.2 The Schwarzschild solution

We now find for the Ricci tensor (preferably, by using an algebraic computer program):

$$R_{00} \;=\; \mathrm{e}^{2(\nu-\mu)}\left(\nu'' - \nu'\mu' + \nu'^2 + 2\nu'/r\right) - \ddot{\mu} + \dot{\nu}\dot{\mu} - \dot{\mu}^2, \tag{11.21}$$

---

[61] J. M. Foyster adn C. B. G. McIntosh; *Commun. Math. Phys.* **27**, 241 (1972).

[62] H. Nariai, *Scientific Reports of the Tôhoku University* **34**, 160 (1950); **35**, 46 (1951); reprinted in *Gen. Relativ. Gravit.* **31**, 951 and 963 (1999), with an editorial note by A. Krasiński, *Gen. Relativ. Gravit.* **31**, 945 (1999) and author's biography by K. Tomita, *Gen. Relativ. Gravit.* **31**, 948 (1999).

[63] A. Krasiński and J. Plebański, *Rep. Math. Phys.* **17**, 217 (1980).

[64] B. Bertotti, *Phys. Rev.* **116**, 1331 (1959).

[65] I. Robinson, *Bull. Acad. Polon. Sci., Ser. Mat. Fis. Astr.* **7**, 351 (1959).

[66] A. Krasiński, *Gen. Relativ. Gravit.* **31**, 945 (1999).

$$\begin{aligned} R_{01} &= 2\dot{\mu}/r, & (11.22)\\ R_{11} &= -\nu'' + \nu'\mu' - \nu'^2 + 2\mu'/r + \mathrm{e}^{2(\mu-\nu)}\left(\ddot{\mu} - \dot{\nu}\dot{\mu} + \dot{\mu}^2\right), & (11.23)\\ R_{22} &= R_{33}/\sin^2\vartheta = r\mathrm{e}^{-2\mu}(\mu' - \nu') + 1 - \mathrm{e}^{-2\mu}, & (11.24)\end{aligned}$$

where $\dot{\mu} = \partial\mu/\partial t$ and $\mu' = \partial\mu/\partial r$. We will solve the Einstein equations with $\Lambda \neq 0$, so

$$R_{\alpha\beta} - \frac{1}{2}\, g_{\alpha\beta} R + \Lambda g_{\alpha\beta} = 0. \qquad (11.25)$$

The trace of (11.25) implies $R = 4\Lambda$. Substituting this in (11.25) we obtain

$$R_{\alpha\beta} = \Lambda g_{\alpha\beta}. \qquad (11.26)$$

With the metric (11.19), eqs. (11.26) imply

$$\mathrm{e}^{-2\nu} R_{00} + \mathrm{e}^{-2\mu} R_{11} = 0. \qquad (11.27)$$

Substituting (11.21) and (11.23) in this we get $2\mathrm{e}^{-2\mu}(\mu' + \nu')/r = 0$, which solves as

$$\nu = -\mu + f(t), \qquad (11.28)$$

where $f(t)$ is an arbitrary function. Then, from $R_{01} = 0$ we get $\mu_{,t} = 0$, i.e. $\mu = \mu(r)$ (so far arbitrary). Now recall that (11.19) does not change its form under the transformations $t' = F(t)$, where $F(t)$ is arbitrary. Consequently, we transform t in (11.19) by

$$t' = \int \mathrm{e}^{f(t)} \mathrm{d}t. \qquad (11.29)$$

The result of (11.29) is as if $f(t) \equiv 0$ in (11.28). Hence, with no loss of generality $\nu = -\mu$, and the whole metric tensor becomes independent of the time coordinate.[11]

From $R_{22} = \Lambda g_{22}$ we now get, with the help of (11.24) and (11.28):

$$1 - \mathrm{e}^{-2\mu} + 2r\mathrm{e}^{-2\mu}\mu' = -\Lambda r^2, \qquad (11.30)$$

which is easily integrated with the result

$$\mathrm{e}^{-2\mu} = 1 + \frac{C}{r} + \frac{1}{3}\,\Lambda r^2 = \mathrm{e}^{2\nu}. \qquad (11.31)$$

where C is an arbitrary constant. The components $\alpha = \beta = 0$ and $\alpha = \beta = 1$ of (11.26) now turn out to be fulfilled identically.

For the strictly vacuum field ($\Lambda = 0$) we get from here

$$g_{00} = \mathrm{e}^{2\nu} = 1 + \frac{C}{r}. \qquad (11.32)$$

[11] The statement thus obtained is usually called the **Birkhoff theorem**: the spherically symmetric gravitational field in vacuum is static. However, calling it a theorem is an exaggeration. It is an intermediate result of a straightforward calculation and finds no other application. Moreover it is, in this formulation, false: it holds only for those spacetime regions where the gradient of δ in (11.1) is spacelike.

We recall that for weak gravitational fields (10.37) should hold. In Newton's theory, for a spherically symmetric gravitational field in vacuum we have

$$\phi = -GM/r, \tag{11.33}$$

where G is the gravitational constant and M is the mass of the source of the field. Eqs. (10.37) and (11.33) are consistent with (11.32) when $C = -2GM/c^2 \overset{\text{def}}{=} -2m$.

Finally then, we found the following solution:

$$\mathrm{d}s^2 = \mathrm{e}^{2\nu}\mathrm{d}t^2 - \mathrm{e}^{-2\nu}\mathrm{d}r^2 - r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2\right), \tag{11.34}$$

where

$$\mathrm{e}^{2\nu} = 1 - \frac{2m}{r} + \frac{1}{3}\,\Lambda r^2. \tag{11.35}$$

Various related solutions were derived by different authors during the first few years after general relativity was published in 1915. The case $\Lambda = 0$ was found by Schwarzschild in 1916 [67] and was historically the first exact solution of the Einstein equations ever found.[12] A large part of relativistic astrophysics and several experimental tests of general relativity are based on it. A year later, Droste [69] presented another, mathematically more elegant derivation of the same solution, accompanied by a discussion of its geometrical properties that was surprisingly insightful and much ahead of its time. It would be quite appropriate to refer to this metric as the Schwarzschild–Droste solution [70]. A generalisation to the case when the source of the gravitational field also has electric charge e was found by Reissner [71] in 1916 and Nordström [72] in 1918 (the generalised solution has $\mathrm{e}^{2\nu} = 1 - 2m/r + \frac{1}{3}\,\Lambda r^2 + e^2/r^2$, see Ref. [11] for the derivation and discussion). The subcase $m = 0$ of (11.35) was found by de Sitter [73] in 1916. The full form (11.35) was found by Kottler [74] in 1918. The Reissner – Nordström metric was generalised still further by Cahen and Defrise [75] in 1968 and by Kinnersley [76] in 1969.

We will mostly deal with the proper Schwarzschild solution for which

$$\mathrm{e}^{2\nu} = 1 - \frac{2m}{r}. \tag{11.36}$$

Note that for $r \to 2m$ this metric apparently has a singularity, since $g_{00} \to 0$ and $g_{11} \to -\infty$. However, if we substitute in (11.34) with (11.36) a value $r < 2m$, then the metric

[67] K. Schwarzschild, *Sitzungsber. Preuss. Akad. Wiss.* p. 189 (1916).

[12] An English translation of the Schwarzschild paper was published in *Gen. Relativ. Gravit.* **35**, 951 (2003), but with an erroneous commentary. A correct editorial note was published separately later:

[68] J. M. M. Senovilla, *Gen. Relativ. Gravit.* **39**, 685 (2007).

[69] J. Droste, *Koninklijke Nederlandsche Akademie van Wetenschappen Proceedings* **19**, 197 (1917); reprinted in *Gen. Relativ. Gravit.* **34**, 1545 (2002), with author's biography by C. Beenakker, *Gen. Relativ. Gravit.* **34**, 1543 (2002), and an editorial note by T. Rothman [70].

[70] T. Rothman, *Gen. Relativ. Gravit.* **34**, 1541 (2002).

[71] H. Reissner, *Ann. Physik* **50**, 106 (1916).

[72] G. Nordström, *Koninklijke Nederlandsche Akademie van Wetenschappen Proceedings* **20**, 1238 (1918).

[73] W. de Sitter, *Mon. Not. Roy. Astr. Soc.* **77**, 155 (1916–17).

[74] F. Kottler, *Ann. Physik* **56**, 410 (1918).

[75] M. Cahen and L. Defrise, *Commun. Math. Phys.* **11**, 56 (1968).

[76] W. Kinnersley, *J. Math. Phys.* **10**, 1195 (1969).

is still nonsingular, but $g_{00} < 0$ and $g_{11} > 0$, i.e. the r-coordinate becomes time, and the t-coordinate becomes a measure of the radial distance. The region $r < 2m$ is the complementary part to $r > 2m$ that was mentioned after Lemma 11.3. In that region $g^{\alpha\beta}\delta,_{\alpha}\,\delta,_{\beta} > 0$, and the "Birkhoff theorem" does not hold.[13] We will discuss the geometrical and physical interpretation of the spurious singularity at $r = 2m$ in Sec. 11.8.

11.3 Orbits of planets in the gravitational field of the Sun.

In general relativity, just as in Newton's theory, it is assumed that the planets in their motion on the orbits behave like point masses, while their own gravitational fields can be neglected. These two assumptions in fact contradict each other. The gravitational field of a point mass is singular at the mass' position, and so stronger than any exterior field. In Newton's theory this difficulty is solved by the observation that the centre of mass of any extended body follows the same trajectory as a point body with no self-gravitation. At the centre of mass, the body's own gravitational field vanishes.

In relativity, there is not even a generally accepted definition of the centre of mass, although work on this problem is being done. Thus, when we consider the orbits of point bodies in relativity, we are in fact extending the theory into the domain in which it has not been worked out yet. Nevertheless, these considerations agree with observational tests.

We assume that the gravitational field of the Sun is spherically symmetric, that the space around the Sun does not contain electromagnetic fields and that the cosmological constant is zero, so the spacetime is described by the metric (11.34) – (11.36). The orbits of the planets should be timelike geodesics in this spacetime, so should obey (5.14). Contracting (5.14) with $g_{\alpha\beta}{\rm d}x^{\beta}/{\rm d}s$ and using (5.14) in the result, we get

$$0 = \frac{D}{{\rm d}s}\left(g_{\alpha\beta}\frac{{\rm d}x^{\alpha}}{{\rm d}s}\frac{{\rm d}x^{\beta}}{{\rm d}s}\right) = \frac{\rm d}{{\rm d}s}\left(g_{\alpha\beta}\frac{{\rm d}x^{\alpha}}{{\rm d}s}\frac{{\rm d}x^{\beta}}{{\rm d}s}\right). \tag{11.37}$$

Hence, $g_{\alpha\beta}\left({\rm d}x^{\alpha}/{\rm d}s\right)\left({\rm d}x^{\beta}/{\rm d}s\right)$ is constant along every geodesic, so its sign cannot change. A geodesic is thus either timelike or null or spacelike along its whole length.

An orbit of a planet is timelike, so $g_{\alpha\beta}\left({\rm d}x^{\alpha}/{\rm d}s\right)\left({\rm d}x^{\beta}/{\rm d}s\right) > 0$. The affine parameter s is determined up to $s = as' + b$, so for timelike geodesics it can be scaled so that

$$g_{\alpha\beta}\frac{{\rm d}x^{\alpha}}{{\rm d}s}\frac{{\rm d}x^{\beta}}{{\rm d}s} = 1. \tag{11.38}$$

This affine parameter is $s = c\tau$, where c is the velocity of light, and τ is the proper time on the geodesic. We wish to use physical units, so in calculating the Christoffel symbols our time coordinate will be ct. For the metric (11.34) – (11.36) we then have:

$$\begin{Bmatrix} 0 \\ 01 \end{Bmatrix} = \frac{m}{r^2}\,\frac{1}{1 - 2m/r} = -\begin{Bmatrix} 1 \\ 11 \end{Bmatrix},$$

[13] This region is an example of a spacetime with the Kantowski–Sachs symmetry, see the footnote after (9.41). It is the unique vacuum solution of Einstein's equations with this symmetry.

$$
\begin{aligned}
\left\{ {1 \atop 00} \right\} &= \frac{mc^2}{r^2} - \frac{2m^2c^2}{r^3}, \\
\left\{ {1 \atop 22} \right\} &= 2m - r = \left\{ {1 \atop 33} \right\} \Big/ \sin^2\vartheta, \\
\left\{ {2 \atop 12} \right\} &= \frac{1}{r} = \left\{ {3 \atop 13} \right\}, \\
\left\{ {2 \atop 33} \right\} &= -\cos\vartheta \sin\vartheta, \qquad \left\{ {3 \atop 23} \right\} = \cot\vartheta.
\end{aligned} \tag{11.39}
$$

Hence, the equations of a geodesic are these:

$$c\frac{\mathrm{d}^2 t}{\mathrm{d}s^2} + \frac{2m}{r^2}\frac{1}{1-2m/r}\, c\frac{\mathrm{d}t}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} = 0, \tag{11.40}$$

$$\begin{aligned}\frac{\mathrm{d}^2 r}{\mathrm{d}s^2} &+ \left(\frac{mc^2}{r^2} - \frac{2m^2c^2}{r^3}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 - \frac{m}{r^2}\frac{1}{1-2m/r}\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 \\ &+ (2m-r)\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}s}\right)^2 + (2m-r)\sin^2\vartheta\left(\frac{\mathrm{d}\varphi}{\mathrm{d}s}\right)^2 = 0,\end{aligned} \tag{11.41}$$

$$\frac{\mathrm{d}^2\vartheta}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}\vartheta}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} - \cos\vartheta\sin\vartheta\left(\frac{\mathrm{d}\varphi}{\mathrm{d}s}\right)^2 = 0, \tag{11.42}$$

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}\varphi}{\mathrm{d}s}\frac{\mathrm{d}r}{\mathrm{d}s} + 2\cot\vartheta\frac{\mathrm{d}\vartheta}{\mathrm{d}s}\frac{\mathrm{d}\varphi}{\mathrm{d}s} = 0. \tag{11.43}$$

The last equation is easily integrated with the result

$$\frac{\mathrm{d}\varphi}{\mathrm{d}s} = \frac{J_0}{r^2\sin^2\vartheta}, \tag{11.44}$$

where J_0 is an arbitrary constant. Substituting this in (11.42) we get

$$\frac{\mathrm{d}^2\vartheta}{\mathrm{d}s^2} + \frac{2}{r}\frac{\mathrm{d}r}{\mathrm{d}s}\frac{\mathrm{d}\vartheta}{\mathrm{d}s} - \frac{{J_0}^2\cos\vartheta}{r^4\sin^3\vartheta} = 0. \tag{11.45}$$

A first integral of this equation is

$$r^4\left(\frac{\mathrm{d}\vartheta}{\mathrm{d}s}\right)^2 + {J_0}^2\cot^2\vartheta = A = \text{constant}. \tag{11.46}$$

We still have the freedom to rotate the axes of the spherical coordinates – the metric is invariant under such rotations, and we have not applied any of them so far. Rotate the $\vartheta = 0$ axis so that at s_0 the planet is on the equator, $\vartheta_0 = \pi/2$. Then we can rotate the spherical coordinates around the axis passing through $x^\alpha(s_0)$. Apply this rotation by such an angle that the initial velocity of the planet falls into the plane of the equator, $\mathrm{d}\vartheta/\mathrm{d}s|_{s=s_0} = 0$. Substituting $\vartheta_0 = \pi/2$ and $\mathrm{d}\vartheta/\mathrm{d}s|_{s=s_0} = 0$ in (11.46) we get $A = 0$. But A is a sum of two nonnegative terms, hence both of them must be zero at all times, i. e.

$$\vartheta \underset{*}{=} \pi/2 \tag{11.47}$$

on the whole orbit. Substituting (11.47) in (11.44) we get

$$\frac{\mathrm{d}\varphi}{\mathrm{d}s} = \frac{J_0}{r^2}. \tag{11.48}$$

A first integral of (11.40) is

$$\left(1 - \frac{2m}{r}\right) c\frac{\mathrm{d}t}{\mathrm{d}s} = E, \tag{11.49}$$

where E is an arbitrary constant. So far, we used (11.42), (11.40) and (11.43). Only (11.41) still remains, but we can replace it by (11.38). Using (11.47) – (11.49) it becomes

$$\frac{E^2}{1 - 2m/r} - \frac{1}{1 - 2m/r}\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 - \frac{{J_0}^2}{r^2} = 1, \tag{11.50}$$

which can be written as[14]

$$\frac{\mathrm{d}r}{\mathrm{d}s} = \left[E^2 - \left(1 + \frac{{J_0}^2}{r^2}\right)\left(1 - \frac{2m}{r}\right)\right]^{1/2}. \tag{11.51}$$

Now we have three first-order equations to solve, (11.48), (11.49) and (11.51). Most often, we are interested only in the shape of the orbit, given by (11.48) and (11.51). The third equation gives information on the position of the planet on the orbit at a given instant.

If $J_0 = 0$, then, from (11.48), $\varphi = \varphi_0 =$ constant, which means that the "planet" falls directly toward the centre of the Sun. We will omit this uninteresting case, and assume $J_0 \neq 0$. Then (11.48) and (11.51) can be replaced by just one equation

$$\left(\frac{J_0}{r^2}\frac{\mathrm{d}r}{\mathrm{d}\varphi}\right)^2 = E^2 - \left(1 + \frac{{J_0}^2}{r^2}\right)\left(1 - \frac{2m}{r}\right). \tag{11.52}$$

Now we introduce the new variable $\sigma = 1/r$ and obtain in (11.52)

$$\left(J_0\frac{\mathrm{d}\sigma}{\mathrm{d}\varphi}\right)^2 = E^2 - \left(1 + {J_0}^2\sigma^2\right)(1 - 2m\sigma). \tag{11.53}$$

An exact solution of this equation would lead to elliptic integrals. We prefer to solve it approximately, up to first order in the small parameter defined below. For the perturbative calculation, it is more convenient to go one step back and differentiate (11.53) by φ:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\varphi} \times 2{J_0}^2\frac{\mathrm{d}^2\sigma}{\mathrm{d}\varphi^2} = \frac{\mathrm{d}\sigma}{\mathrm{d}\varphi}\left[-2{J_0}^2\sigma + 2m\left(1 + 3{J_0}^2\sigma^2\right)\right]. \tag{11.54}$$

One solution of this is $\mathrm{d}\sigma/\mathrm{d}\varphi = 0$, i.e. motion on a circular orbit with $r =$ constant. When $\mathrm{d}\sigma/\mathrm{d}\varphi \neq 0$, (11.54) becomes

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\varphi^2} + \sigma = \frac{m}{{J_0}^2} + 3m\sigma^2. \tag{11.55}$$

[14] We omit the minus sign in front of the right-hand side in (11.51) because the resulting solutions describe motions along these same orbits in the opposite direction. In a spherically symmetric spacetime, the reversal of the sense of orbital motion is equivalent to a coordinate transformation (reflection). This is not so when the source of the gravitational field is rotating; then the orbital motion can be **direct**, with angular velocity parallel to that of the source, or **retrograde**, with the two angular velocities anitparallel. We shall meet this problem in the Kerr spacetime, see Sec. 17.6.

This differs from the corresponding Newtonian equation by the last term that vanishes in the limit $c \to \infty$. In order that the first term on the right assumes the Newtonian form $GM\mu^2/J^2$ in this limit (where M is the mass of the Sun while μ and J are the mass and the orbital angular momentum of the planet), we must assume $J_0 = J/(\mu c)$. We denote

$$\frac{mc^2\mu^2}{J^2} \equiv \frac{GM\mu^2}{J^2} \overset{\text{def}}{=} \frac{1}{p}, \qquad 3m \equiv \frac{3GM}{c^2} \overset{\text{def}}{=} \alpha. \tag{11.56}$$

In the Newtonian limit, α becomes zero, while p is the **parameter** of the orbit: the distance from the Sun to the planet at that instant at which the Sun–planet direction is perpendicular to the Sun–perihelion direction. In this notation, eq. (11.55) becomes

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\varphi^2} + \sigma = \frac{1}{p} + \alpha\sigma^2. \tag{11.57}$$

We will solve this equation approximately, up to terms linear in α. More precisely, the small parameter will be the dimensionless number $\alpha\sigma$. For the Sun, $\alpha \approx 4.5$ km, while the radius of the Sun $\approx 7 \times 10^5$ km, so the planetary orbits have still larger radii, and $\alpha\sigma < 0.64 \times 10^{-5}$. Since σ and $1/p$ are comparable, $\alpha\sigma^2 \ll \sigma$ and $\alpha\sigma^2 \ll 1/p$, and the last term in (11.57) is indeed small compared to all other terms. However, for strongly concentrated objects such as neutron stars or black holes, the orbits may come close to $r = 0$, and the approximation used here becomes invalid. In those cases, one has to resort either to numerical computation or to calculations with elliptic functions.[15]

In the zero-th order of approximation, the solution of (11.57) is the elliptical orbit

$$\sigma_0 = \frac{1}{p} + \frac{\varepsilon}{p} \cos\left(\varphi - \varphi_0\right), \tag{11.58}$$

where ε is the **eccentricity** of the orbit, also a small quantity for solar planets (for Pluto, $\varepsilon = 0.2444$, for Mercury $\varepsilon = 0.205628$, for other planets it is still smaller). In the first approximation, we assume σ to be of the form $\sigma_{[1]} = \sigma_0 + \sigma_1$, where σ_1 is of the order of α, i.e. $\sigma_1 = 0$ when $\alpha = 0$. We substitute $\sigma_{[1]}$ in (11.57) and neglect terms of order α^2:

$$\frac{\mathrm{d}^2\sigma_1}{\mathrm{d}\varphi^2} + \sigma_1 = \frac{\alpha}{p^2}\left[1 + 2\varepsilon\cos\left(\varphi - \varphi_0\right) + \varepsilon^2\cos^2\left(\varphi - \varphi_0\right)\right]. \tag{11.59}$$

Since we assumed $\sigma_1 = 0$ when $\alpha = 0$, we must take $\sigma_{1j} = 0$ as the solution of the homogeneous part of (11.59). The problem is thus to find a special solution of the full inhomogeneous equation. This can be done by textbook methods; the solution is

$$\sigma_1 = \frac{\alpha}{p^2}\left(1 + \frac{\varepsilon^2}{3}\right) + \frac{\alpha\varepsilon}{p^2}\left(\varphi - \varphi_0\right)\sin\left(\varphi - \varphi_0\right) + \frac{\alpha\varepsilon^2}{3p^2}\sin^2\left(\varphi - \varphi_0\right). \tag{11.60}$$

However, the solution $\widetilde{\sigma}(\varphi) \overset{\text{def}}{=} \sigma_0 + \sigma_1$ implied by (11.60) disagrees with observations and with common sense: at $\varphi - \varphi_0 = n\pi$, $n = 1, 2, 3, \ldots$ the curve $\widetilde{\sigma}(\varphi)$ passes through always

[15] In real astrophysical situations, the orbits are perturbed in so many ways that numerical treatment becomes necessary anyway. Examples of perturbations: gravitation of other planets, rotation of the star, lack of spherical symmetry caused by rotation, loss of orbital angular momentum because of friction with interplanetary matter.

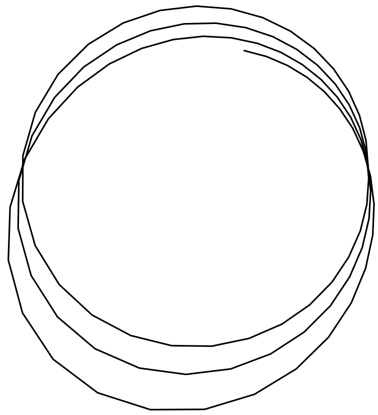

Figure 11.1: The orbit as it would look if eq. (11.60) were the correct equation. See explanation in text.

the same two points, but in the sectors where $\sin(\varphi - \varphi_0) > 0$, σ_1 increases with every revolution, with no upper limit, i.e. $r \to 0$ as $\varphi \to \infty$. In the sectors where $\sin(\varphi - \varphi_0) < 0$, σ_1 decreases with every revolution, with no lower limit, and will at a certain instant cause that $\widetilde{\sigma} = 0$, i.e. $r \to \infty$. The shape of this curve for a few first revolutions is shown in Fig. 11.1. The term responsible for this unusual behaviour is $(\alpha\varepsilon/p^2)(\varphi - \varphi_0)\sin(\varphi - \varphi_0)$, in which $(\varphi - \varphi_0)$ becomes arbitrarily large as φ increases. On the other hand, analysing the function on the right-hand side of (11.53) by the methods of classical mechanics one can prove that with $E < 1$, $J_0 > 2\sqrt{3}GM/c^2$ and the initial σ being sufficiently small, the solution $\sigma(\varphi)$ must be bounded (see Exercise 3). We conclude that the trouble-causing term must be the first term of a power series representing a bounded function, and that in the first order solution to (11.57) we must include that function in full. Identifying the unknown function by going to higher orders of approximation would be laborious since, with every added order, many new terms appear. We want the solution to reduce to (11.60) only at first order, and any bounded function with this property will do. After some guesswork, the following function is found to be the right one:

$$\frac{\varepsilon}{p} \cos\left[\left(1 - \frac{\alpha}{p}\right)(\varphi - \varphi_0)\right] \tag{11.61}$$

This is bounded for any α, but "approximating" it by the truncated Taylor formula caused the nonsensical behaviour of (11.60). We will come back to this at the end of this section.

Hence, the correct approximation to the solution of (11.57) linear in α is:

$$\sigma = \frac{1}{p} + \frac{\alpha}{p^2}\left(1 + \frac{\varepsilon^2}{3}\right) + \frac{\varepsilon}{p}\cos\left[\left(1 - \frac{\alpha}{p}\right)(\varphi - \varphi_0)\right] + \frac{\alpha\varepsilon^2}{3p^2}\sin^2(\varphi - \varphi_0). \tag{11.62}$$

Since $\varepsilon\alpha/p \ll 1$, the third term on the right dominates over the last one, which can thus be neglected. It can be seen now that σ does not return to its initial value after the planet makes one full revolution, $\varphi \to \varphi + 2\pi$. Instead, it returns (approximately) to its initial value after the planet revolves by $2\pi/(1 - \alpha/p) > 2\pi$. Consequently, the aphelion and the perihelion of the orbit move in the same direction in which the planet goes around the orbit. Two consecutive perihelia are separated by the angle $\Delta\varphi = 2\pi/(1-\alpha/p) - 2\pi = 2\pi\alpha/p + O((\alpha/p)^2)$, and the formula for $\Delta\varphi$ can be written in three equivalent ways

$$\Delta\varphi \approx \frac{2\pi\alpha}{p} = \frac{6\pi GM}{c^2 p} = \frac{6\pi GM}{c^2 b(1+\varepsilon)} = \frac{6\pi GM}{c^2 a\,(1-\varepsilon^2)}, \tag{11.63}$$

where b is the perihelion and a is the semimajor axis of the orbit. The shape of the orbit is shown in Fig. 1.1, but with a greatly exaggerated value of $\Delta\varphi$.

Note that $\Delta\varphi$ is larger for smaller b. Hence, Mercury offers the best conditions for measuring this effect. Since $\Delta\varphi$ is very small, but the perihelion advance increases with time, the most convenient unit to measure this effect is not radians per revolution, as in (11.63), but arc seconds per century. In these units, we have

$$\Delta\Phi = \frac{C}{T} \times \frac{360\cdot 60\cdot 60}{2\pi} \times \frac{6\pi GM}{c^2 a\,(1-\varepsilon^2)}, \tag{11.64}$$

where $C = 100$ years, T is the orbital period of Mercury, and the second factor converts radians to arc seconds. From the tables [77] we find $G = 6.670\times 10^{-8}$ cm^3/ (g $\times$ s^2) for the gravitational constant, $M = 1.989\times 10^{33}$g for the mass of the Sun, $c^2 = 8.987554\times 10^{20}$ cm^2/s^2 for the square of the velocity of light, $a = 57.9\times 10^{11}$cm for the semimajor axis of Mercury's orbit, $\varepsilon = 0.205628$ for the eccentricity of Mercury's orbit, and $T = 0.24085$ terrestrial years for Mercury's orbital period. Substituting all these values in (11.64) we get[16] $\Delta\Phi = 43.03$ arc seconds/century.

Equation (11.63) is the famous result of relativity (the **relativistic perihelion shift**) that predicted the orbital motion of Mercury in agreement with astronomical observations. As described in chapter 1, when this correction is added to the perturbations of Mercury's orbit by other planets, the sum is in good agreement with the observed value.

The observational tests of this effect are done by classical astronomical methods – by registering the positions of planets at different instants, reconstructing their orbits in space and determining the positions of perihelia. As mentioned in chapter 1, the perihelion shift of Mercury had been observed already in the first half of 19th century, and its Newtonian origins were identified by U. J. LeVerrier in 1859 [2, 3]. He was also the person who noted the discrepancy between observational results and the prediction of Newton's theory. In addition to perturbations of Mercury's orbit by other planets there is one more component in the observed effect: the astronomical observations are carried out from the Earth, and so their raw results are expressed in the geocentric reference system. This gives the largest contribution to the perihelion shift. According to modern observations [4, 78], the full observed perihelion shift of Mercury is $5599.74 \pm 0.41''$/century, of which approx. $5000''$/century is caused by the geocentric reference system, approx. $280''$ is caused by the gravitational perturbations from Venus, approx. $150''$ by perturbations from Jupiter and approx. $100''$ by perturbations from the remaining planets [4]. The unexplained residue is $43.11 \pm 0.45''$/century [78], which agrees with the value calculated from relativity.

[77] C. W. Allen, *Astrophysical quantities.* The Athlone Press, University of London 1973.

[16] Most pocket calculators will probably show a slightly different result, because of inconsistent roundoff errors. The value quoted here has been obtained by carefully tuning the precision of all factors.

[78] K. Lang, *Astrophysical formulae.* Springer, Berlin – Heidelberg – New York 1974.

For other planets, the relativistic perihelion shift is not larger than a few arc seconds per century [4]. Moreover, their orbits have small eccentricities, so the positions of their perihelia are known with smaller precision, thus they are less suitable for testing relativity.

Now a few comments on the procedure we used to obtain the approximate result:

1. We were able to recognise that Eq. (11.60) is incorrect because we knew from elsewhere that the orbit should be bounded. So, one must be careful about drawing conclusions from approximate results when nothing is known about the expected result.

2. While going to higher orders of approximation, an avalanche of new terms appears. This is typical of all perturbative schemes. Thus, the hope that something useful can be inferred about the exact result by considering consecutive approximations may be illusory.

3. We emphasise again that the results (11.63) – (11.64) were obtained under the assumption that the radius of the orbit, at every point, is much larger than the gravitational radius of the central body, GM/c^2. They *do not* apply to those orbits around neutron stars and black holes that come close to the central body.

11.4 Deflection of light-rays in the Schwarzschild field.

We will now calculate the path of a light ray in a spherically symmetric gravitational field. To do this, we have to solve the equations of null geodesics in the Schwarzschild metric.

The conclusion that $\vartheta = \pi/2$ on the whole orbit if coordinates are suitably chosen is valid also for null orbits. But the value of the first integral (11.38) is this time different,

$$g_{\alpha\beta}\frac{\mathrm{d}x^\alpha}{\mathrm{d}s}\frac{\mathrm{d}x^\beta}{\mathrm{d}s} = 0. \tag{11.65}$$

Moreover, now s is not time, but just an affine parameter on the geodesic.

Equations (11.48) – (11.49) still apply, but instead of (11.50) we now have

$$E^2 - \left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 - \frac{{J_0}^2}{r^2}\left(1 - \frac{2m}{r}\right) = 0. \tag{11.66}$$

With $J_0 = 0$ the geodesic is radial. With $J_0 \neq 0$, we obtain from (11.48) and (11.66):

$$\frac{\mathrm{d}r}{\mathrm{d}\varphi} = \frac{r^2}{J_0}\sqrt{E^2 - \frac{{J_0}^2}{r^2}\left(1 - \frac{2m}{r}\right)}. \tag{11.67}$$

(We take only the + sign for the square root for the same reason as before.) We will handle this equation in a similar way as in the previous section. Introducing $\sigma = 1/r$, we get

$$\left(J_0\frac{\mathrm{d}\sigma}{\mathrm{d}\varphi}\right)^2 = E^2 - {J_0}^2\sigma^2\left(1 - 2m\sigma\right). \tag{11.68}$$

From here, differentiating by φ:

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\varphi} \times \frac{\mathrm{d}^2\sigma}{\mathrm{d}\varphi^2} = \frac{\mathrm{d}\sigma}{\mathrm{d}\varphi}\left(-\sigma + 3m\sigma^2\right). \tag{11.69}$$

Again we have two cases: either a circular orbit with $\mathrm{d}\sigma/\mathrm{d}\varphi = 0$, or $\mathrm{d}\sigma/\mathrm{d}\varphi \neq 0$, and

$$\frac{\mathrm{d}^2\sigma}{\mathrm{d}\varphi^2} + \sigma = \alpha\sigma^2, \tag{11.70}$$

where $\alpha = 3m$, as before. We solve this equation by the same method as (11.57). In the zero-th (Newtonian) approximation, the solution is

$$\frac{1}{r} = \sigma_0 = \frac{1}{R}\ \cos\left(\varphi - \varphi_0\right), \tag{11.71}$$

where $R =$ constant. In the Newtonian limit, (r, ϑ, φ) are the spherical coordinates in the Euclidean space. Then (11.71) is an equation of a straight line passing at the distance R from the point $r = 0$.

In the first approximation we look for a solution in the form $\sigma_{[1]} = \sigma_0 + \sigma_1$, where $\sigma_1/\sigma_0 \propto \alpha$. The function σ_1 must then obey

$$\frac{\mathrm{d}^2\sigma_1}{\mathrm{d}\varphi^2} + \sigma_1 = \frac{\alpha}{R^2}\ \cos^2\left(\varphi - \varphi_0\right). \tag{11.72}$$

A particular integral of (11.72) is

$$\sigma_1 = \frac{\alpha}{3R^2}\ \left[1 + \sin^2\left(\varphi - \varphi_0\right)\right]. \tag{11.73}$$

The full solution of (11.72) up to terms of order α is thus

$$\frac{1}{r} = \frac{1}{R}\ \cos\left(\varphi - \varphi_0\right) + \frac{\alpha}{3R^2}\ \left[1 + \sin^2\left(\varphi - \varphi_0\right)\right]. \tag{11.74}$$

This may be solved for φ with the result

$$\varphi_\pm = \varphi_0 \pm \arccos\left[\frac{3R}{2\alpha}\left(1 - \sqrt{1 + \frac{8\alpha^2}{9R^2} - \frac{4\alpha}{3r}}\right)\right]. \tag{11.75}$$

This orbit has two asymptotes whose directions are calculated from here in the limit $r \to \infty$. The angle between them is (see Fig. 11.2)

$$\Delta\varphi = \lim_{r\to\infty}\left(\varphi_+ - \varphi_-\right) - \pi = 2\arccos\left[\frac{3R}{2\alpha}\left(1 - \sqrt{1 + \frac{8\alpha^2}{9R^2}}\right)\right] - \pi. \tag{11.76}$$

Equation (11.74) is an approximate solution to (11.70), so it will apply only with a limited precision (up to terms linear in α). Hence, we do not have to consider eq. (11.76) in its exact form – it suffices to take it with the same precision, and then it becomes

$$\Delta\varphi = \frac{4\alpha}{3R} = \frac{4GM}{c^2R}. \tag{11.77}$$

This is another famous result of relativity. But before we confront it with observations, we will first rectify a misunderstanding that still lingers in the literature.

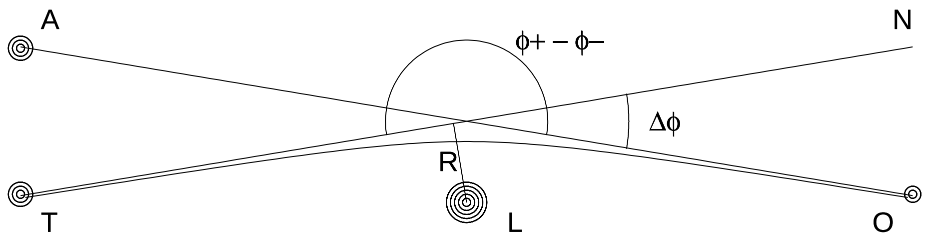

Figure 11.2: Deflection of light-rays in a spherically symmetric gravitational field. T is the true position of the source of light, A is the apparent position seen by observer O in consequence of the deflection. According to Newton's theory, the light ray emitted from T in the direction TN would move on the straight line TN. In reality, it follows the curve TO. The line TN is the asymptote of the initial part of the ray's path, AO is the asymptote of the final part. L is the deflecting body, R is the distance between the center of the body L and the Newtonian path of the light ray, $\phi_+ - \phi_-$ is the angle between the asymptotes, calculated from (11.75) as $\phi_+ - \phi_- = \lim_{r\to\infty}(\varphi_+ - \varphi_-)$, $\Delta\phi$ is the deflection angle. The origin of the polar coordinates is at the center of the deflecting body L, ϕ increases in the anticlockwise direction.

Some authors wish to calculate the angle of gravitational deflection of a light ray in a quick and easy way, and for that purpose they use a combination of special relativity and Newton's theory. They assume that a photon of energy E has the mass E/c^2 that interacts with the central body by the ordinary law of gravitation. Consequently, the orbit of a light ray is just an orbit of a particle of mass E/c^2. We will show below that the result obtained in this way is exactly one half of the correct (verified observationally!) eq. (11.77).

The Newtonian orbit in polar coordinates is given by (11.58). From there

$$\varphi_\pm = \varphi_0 \pm \arccos\left[\frac{1}{\varepsilon}\,\left(\frac{p}{r} - 1\right)\right]. \tag{11.78}$$

The deflection angle is thus

$$\Delta\varphi = \lim_{r\to\infty}(\varphi_+ - \varphi_-) - \pi = 2\arccos\left(-\frac{1}{\varepsilon}\right) - \pi = 2\arcsin\left(\frac{1}{\varepsilon}\right). \tag{11.79}$$

For a hyperbolic Newtonian orbit the eccentricity ε is

$$\varepsilon = \sqrt{1 + \frac{2E_0J^2}{G^2M^2\mu^3}}, \tag{11.80}$$

where E_0 is the total energy of the particle on the orbit, J is its angular momentum, μ is its mass. In our case the "particle" is a photon moving at the velocity c, and so

$$E_0 = \frac{1}{2}\,\mu c^2, \qquad J = \mu cR, \tag{11.81}$$

where R is, as before, the minimum distance between the orbiting particle and the center of the star. Hence in (11.80):

$$\varepsilon = \sqrt{1 + \frac{R^2c^4}{G^2M^2}} \gg 1. \tag{11.82}$$

Up to linear terms in $GM/(c^2R)$ we thus have from (11.79)

$$\Delta\varphi = \frac{2GM}{c^2R}, \tag{11.83}$$

i.e. half of the prediction of general relativity. Note that we arranged for the discrepancy to be small: had we taken the special relativistic equation $E = mc^2$ instead of (11.81), the final result would be $\Delta\varphi = \sqrt{2}GM/(c^2R)$, differing from (11.77) by still more. The lesson to be drawn is this: intuitive arguments constructed by picking out elements of different theories and doing simple arithmetics on them are not guaranteed to lead to correct results. Each physical theory has its inner logic that has to be applied consistently.

The same result, eq. (11.83), can be derived from the pure Newton theory, without mixing in special relativity, see Exercise 5.

11.5 Measuring the deflection of light-rays.

By (11.77), $\Delta\varphi \propto 1/R$, thus the deflection is stronger for those rays that approach the central star to within a smaller distance. For a light ray grazing the surface of the Sun, $R = 6.9599 \times 10^{10}$ cm. Taking this number, and the other quantities in (11.77) as in Sec. 11.3 we obtain $\Delta\varphi = 1.75''$. This is the maximum value available in observations.[17] For stars other than the Sun the necessary parameters M and R cannot be measured with sufficient precision, and the deflection angles would be much too small to be measured (see further in this section). Planets cause unmeasurably small deflection angles.

The discussion of deflection of light-rays applies to all kinds of radiation, e.g. X rays, γ rays, microwaves. But when general relativity was published in 1915, the whole of observational astronomy relied on optical observations – other ranges of electromagnetic radiation from space were impossible to detect with the technology of that time. The first attempt to measure the deflection of light in the gravitational field of the Sun was undertaken by Arthur S. Eddington in 1919 [79]. Observing the rays that graze the Sun's surface is possible only during a total eclipse. The idea was this:

1. Find two stars that will be visible at the edge of the Sun during the eclipse; best of all on opposite ends of Sun's diameter.

2. Take a photograph of these stars during the eclipse, with the darkened Sun between them.

3. Take a photograph of the same two stars several months later, when they are visible against the night sky.

---

[17] Today, gravitational deflection of light is observed also for distant galaxies, in the **gravitational lenses**, see Sec. 11.6. They are not useful for testing relativity because the mass distributions and radii of the deflecting objects are not known with sufficient precision. Also, in order to correctly calculate the deflection angle in a gravitational lens, the equation of a null geodesic should be solved in a different geometry – a model of the Universe with inhomogeneous distribution of matter. No such results are available as yet. For more on inhomogeneous models of the Universe see Chapters 15 and 16.

[79] F. W. Dyson, A. S. Eddington and C. Davidson, *Phil. Trans. Roy. Soc. London* A220, 291 (1920); cited after Ref. [4], p. 5.

4. Measure the differences in positions of the stars in the two photographs.

The geometry of this measurement is shown in Fig. 11.3. With the Sun being away, the observer O sees the two stars at their true positions T_1 and T_2, at angular separation α_1. With the Sun between the them, the stars are seen at the apparent positions A_1 and A_2, at angular separation $\alpha_2 > \alpha_1$. The angle of deflection is $\Delta\varphi = (\alpha_2 - \alpha_1)/2$.

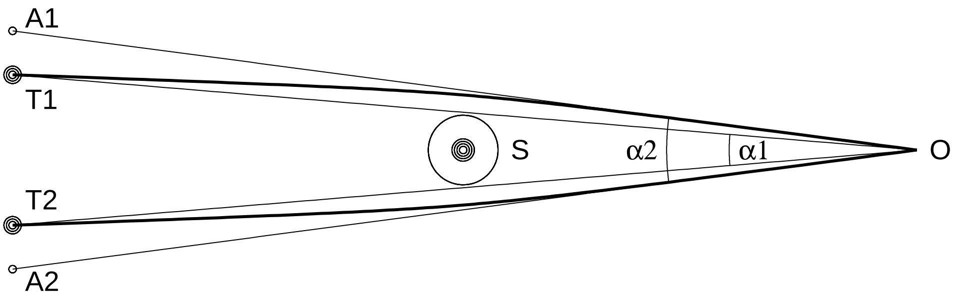

Figure 11.3: Measuring the deflection of light-rays by the Sun by Eddington's method. When the Sun is away, the observer O sees the stars at their true positions T_1 and T_2, at angular separation α_1. Then, the light from the stars reaches the observer along the straight lines T_1O and T_2O. When the Sun S is seen between the stars, their light follows the curved arcs T_1O and T_2O, and the observer sees the images of the stars at the apparent positions A_1 and A_2, at angular separation $\alpha_2 > \alpha_1$. The straight lines A_1O and A_2O are tangent to the arcs T_1O and T_2O at the point O. The prediction of relativity is $\Delta\varphi = (\alpha_2 - \alpha_1)/2 = 1.75''$.

With the technology of 1919, this measurement posed great technical difficulties. The expected effect was so small that mechanical deformations of photographic plates during the months when they had to be stored could disturb the result. Total eclipses of the Sun typically occur far away from well-equipped observatories. Eddington's expedition consisted of two groups, carrying out the observations in Sobral in Brazil and on Principe Island in the Guinea Bay off Africa. Reaching such a destination on time is a problem in itself. In addition, one has to carry out precise measurements under field conditions. During an eclipse the temperature of the atmosphere drops, which causes some turbulence and cooling of the telescope. The precision achievable under such conditions was limited, but the effect predicted by relativity was confirmed. The deflection angle measured at Sobral was $1.98 \pm 0.16''$, that on Principe Island was $1.61 \pm 0.40''$ [80]. Unlike the anomalous perihelion shift of Mercury that had been known for many years before relativity was created, the deflection of light-rays was *predicted* by relativity. Apart from the old and long-forgotten attempts by Cavendish and Soldner [81, 82, 83], nobody had expected gravitation to influence the propagation of light. Eddington's positive result catapulted relativity and Einstein personally to the fame and public fascination that they still enjoy today.

Results obtained later from measurements by the same method gave results between 90% and 150% of the relativity value $1.75''$, with formal errors estimated to be 6% to 25%,

[80] Ref. [4], p. 5.
[81] C. M. Will, *Am. J. Phys.* **56**, 413 (1988); cited after Ref. [83].
[82] J. Soldner, *Berliner Astronomisches Jahrbuch* 1804, str. 161; cited after Ref. [83].
[83] P. Schneider, J. Ehlers and E. E. Falco, *Gravitational lenses.* Springer, Berlin 1992.

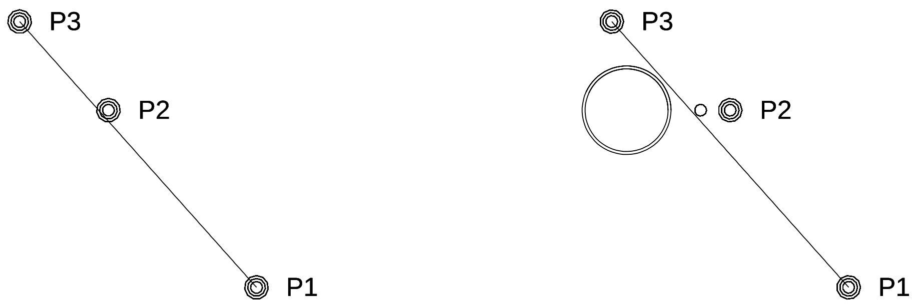

Figure 11.4: Measuring the deflection of microwaves by the Sun. When the Sun is away, the three radio sources P_1, P_2 and P_3 are seen nearly on a straight line (left). The true position of the middle source P_2 can then be measured relative to the other two. With the disc of the Sun approaching (right), the image of P_2 moves away from the Sun. Its true position (small circle) can then be calculated from the positions of P_1 and P_3 (which do not change measurably) and compared with the observed apparent position to calculate the angle of deflection. The shift of the image of P_2 is exaggerated for clarity.

and unknown systematic errors.

Today, another method is available. Three radio sources, 0119+11, 0116+08 and 0111+02, lie nearly on a straight line, inclined at a large angle to Sun's yearly trajectory through the sky. (The symbols are their coordinates on the sky, the right ascension in hours and minutes, and the declination in degrees.) Each year, for about 3 weeks in March/April, the Sun passes in front of the middle radio source, see Fig. 11.4. The other two are then so far from Sun's disc on the sky that their radiation is not measurably deflected. The position of the middle source relative to the two others is measured when the Sun is elsewhere in the sky, and compared with the observed position while Sun's edge approaches the line of sight. The Sun is a weak radio source and its own microwave radiation does not disturb the measurement; the object 0116+08 can be observed right up to the moment when it vanishes behind the Sun. The only complication is that the plasma in the solar corona also deflects the microwaves. However, the plasma deflection depends, in a known way, on the wavelength, while the gravitational deflection does not. Thus, carrying out the measurement at two wavelengths suffices to disentangle the two effects.

This measurement was first carried out in 1974 by E. B. Fomalont and R. Sramek [84] at the National Radio Astronomy Observatory in Green Bank (West Virginia). With this method, the whole observation is done in laboratory conditions and so is more precise. It has become customary to quote the observational results in the form of the ratio of the measured quantity to the value predicted by relativity. The value 1 means perfect agreement. The Fomalont – Sramek result was 1.007 ± 0.009 (standard error), so the value 1 is within the error bar.

[84] E. B. Fomalont and R. Sramek, *Astrophys. J.* **199**, 749 (1975); *Phys. Rev. Lett.* **36**, 1475 (1976); *Comments Astrophys.* **7**, 19 (1977).

11.6 Gravitational lenses.

A **gravitational lens** is a body situated relative to the observer in such a way that the deflected rays intersect at the observer's position. The theory and observations of gravitational lenses have become a science in itself [83], and we will touch on this subject here only very briefly.

A gravitational lens formed by a single star is shown in Fig. 11.5. S is the source of light, L is the deflecting star (the lens), and O is the observer; R is the radius of the star, d_S and d_O are the distances from the lens to the source of light and to the observer; $\Delta\varphi_S$ and $\Delta\varphi_O$ are the angles filled by the radius R at the position of the source and of the observer; $\Delta\varphi$ is the deflection angle of the ray that grazes the surface of the lens. Neglecting the curvature of space, we obtain, approximately for small angles

$$R = d_S \Delta\varphi_S = d_O \Delta\varphi_O. \tag{11.84}$$

On the other hand, using (11.77) and Fig. 11.5 we have

$$\Delta\varphi = \Delta\varphi_S + \Delta\varphi_O = \frac{4GM}{c^2 R}. \tag{11.85}$$

From (11.84) and (11.85) we obtain the "equation of a gravitational lens":

$$\frac{1}{d_S} + \frac{1}{d_O} = \frac{4GM}{c^2 R^2}. \tag{11.86}$$

So, unlike optical lenses, the gravitational lenses do not focus the rays: those flying farther from the optical axis are deflected by smaller angles than those closer to the axis. Hence, it is not possible to "view" anything through a gravitational lens as if it were a magnifying glass – the image is very distorted. Nevertheless, there is some intensification of light: rays that would disperse in the absence of a lens, intersect again. The gravitational lenses can thus increase the range of optical observations, and are actually used for this purpose [83].

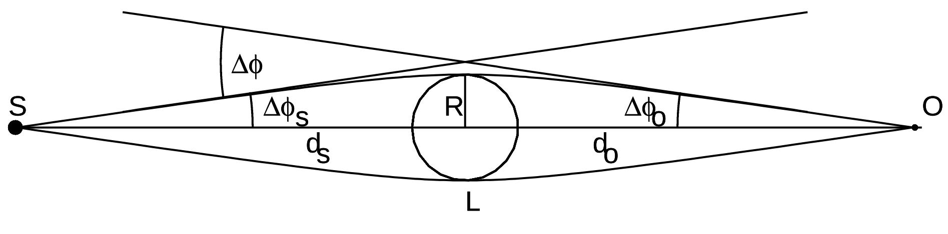

Figure 11.5: A spherical star as a gravitational lens. In the approximation used in the text, the radius of the star, R, equals the distance from the star's center to the point where the two straight lines intersect. Everything else is elementary geometry – see text.

Would it be possible for an observer on the Earth to use the Sun as a gravitational lens? Equation (11.86) implies decisively: no. The rays that graze the surface of the Sun would intersect at the distance d_O that is smaller the greater d_S is. Hence, the minimum d_O is calculated from (11.86) in the limit $d_S \to \infty$, it is $d_{\min} = (cR)^2/(4GM) = 8.2 \times 10^{10}$

km, while the radius of Earth's orbit is 1.49597892×10^8 km. (The $d_{\min}$ is more than 13 times the radius of the orbit of Pluto, i.e. far beyond the Solar System.) Likewise, there is no chance to observe the "lensing" by other stars by measuring the deflection angles. Even for the closest star, which is 4.5 light years away from the Earth, assuming it has the same radius as the Sun, the angle $\Delta\varphi_O$ would be $3.4 \times 10^{-3''}$ – too little to be measured, even in the unlikely event that another star is found behind it in the right direction.[18]

By extending the formula (11.77) to intergalactic distances (which is not really correct – see below) we can conclude that galaxies do have a chance to be gravitational lenses (and indeed many such lenses have been observed [83]). Substituting in (11.86) the mass of our Galaxy, $M = 1.4 \times 10^{11} M_\odot$ and its smallest diameter $R = 5$ kpc, with 1 kpc $= 3.0857 \times 10^{21}$ cm, we get $d_{\min} = 9.33 \times 10^2$ Mpc. According to the Hubble formula the luminosity distance is $D_L = zc/H_0$, where, according to the most recent measurement[85], $H_0 \approx 67.11$ km/(s×Mpc), this $d_{\min}$ corresponds to $z \approx 0.2$. Taking the largest diameter of our Galaxy, $R = 30$ kpc we get $z = 7.2$. The quasars are near the middle of the range $z \in (0.2, 7.2)$. For a galaxy of diameter 30 kpc at the distance 1.34×10^4 Mpc (which corresponds to $z = 3$) the angle $\Delta\varphi_O$ is $0.46''$, which is measureable. Thus, it should not be surprising that most of the observed gravitational lenses are quasars.

However, (11.77) and (11.86) do not apply to quasars. The first of them applies (approximately) only in the gravitational field of a single star, the second one is an even cruder approximation in the same geometry. The distance from the Earth to quasars is large in the cosmological scale. For calculating the deflection of light over such distances one should consider null geodesics in a model of the Universe. In astronomical practice, gravitational lenses are described by a sort of geometric optics based on the newtonian description of propagation of light [83]. This description is not a theory in the sense that the word "theory" has in physics – rather, it is a collection of heuristic rules. Nevertheless, it gives testable results that are in approximate agreement with observations.

Figure 11.6 shows the most famous example of a gravitational lens, the so-called Einstein cross,[19] also called the Huchra lens, after its discoverer John Huchra. The four peripheral bright spots are images of the quasar QSO 2237+0305 created by the nearer galaxy ZW 2237+030. The galaxy is seen as the bluish haze surrounding all the spots; the bright spot at the centre is its nucleus. The complicated image is a consequence of the lens having no symmetry; Eqs. (11.77) and (11.86) do not apply to it.

An even more interesting example of a gravitational lens is the galaxy LRG 3-757 shown in Fig. 11.7.[20] The lens lies nearly on the straight line connecting the Earth with the

[18] Such "microlensing" has been successfully observed by measuring the changes of intensity of light from more distant stars when eclipsed by the lenses.

[85] Planck collaboration, Planck 2013 results. XVI. Cosmological parameters. *Astron. Astrophys.* **571**, A16 (2014).

[19] This image was copied from

[86] https://upload.wikimedia.org/wikipedia/commons/c/c8/Einstein_cross.jpg

According to the Wikimedia Commons license

[87] http://creativecommons.org/licenses/by-sa/3.0/,
https://en.wikipedia.org/wiki/Wikipedia:
Text_of_Creative_Commons_Attribution-ShareAlike_3.0_Unported_License

it does not need the author's permission to be disseminated.

[20] This image was copied from

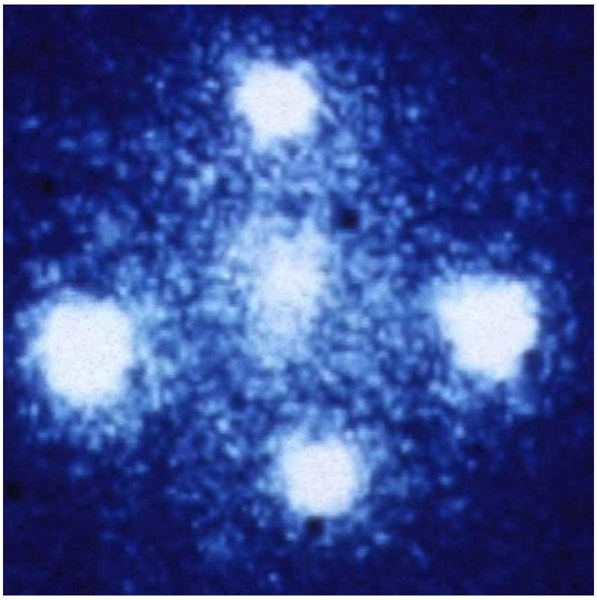

Figure 11.6: The most famous example of a gravitational lens – the "Einstein cross". See explanation in the text.

farther galaxy, which is not seen directly. In consequence of the nearly axially symmetric arrangement, the light of the far galaxy forms a nearly full ring around the nearer one.

11.7 The spurious singularity of the Schwarzschild solution at $r = 2m$.

The Schwarzschild solution given by eqs. (11.34) and (11.36) has a singularity at $r = 2m = 2GM/c^2$: as $r \to 2m$, $g_{00} \to 0$ while $g_{11} \to \infty$. However, for the determinant of the metric, $g = -r^4 \sin^4 \vartheta$, the value $r = 2m$ is not in any way special. Also, components of the Riemann tensor projected on the vectors of the orthonormal basis

$$\begin{aligned} e_{\hat{0}}{}^{\alpha} &= \mathrm{e}^{-\nu}\delta^{\alpha}{}_0, \qquad e_{\hat{1}}{}^{\alpha} = \mathrm{e}^{-\mu}\delta^{\alpha}{}_1, \\ e_{\hat{2}}{}^{\alpha} &= \frac{1}{r}\,\delta^{\alpha}{}_2, \qquad e_{\hat{3}}{}^{\alpha} = \frac{1}{r\sin\vartheta}\,\delta^{\alpha}{}_2. \end{aligned} \tag{11.87}$$

are regular at $r = 2m$: they are all of the form $\alpha m/r^3$, where $\alpha = 0, \pm 1, \pm 2$. (The indices with hats label vectors and are not affected by coordinate transformations.) These facts suggest that the singularity of (11.34) & (11.36) at $r = 2m$ is spurious – created by the coordinates being used, and not by the geometry of the manifold. To see that this is

[88] https://en.wikipedia.org/wiki/Gravitational_lens#/media/ File:A _Horseshoe _Einstein _Ring _from _Hubble.JPG.
It was made accessible in the public domain by its owner.

Figure 11.7: Another conspicuous example of a gravitational lens – LRG 3-757. See explanation in the text.

possible, let us look at the transformation:

$$g_{\alpha'\beta'} = \frac{\partial x^\alpha}{\partial y^{\alpha'}} \frac{\partial x^\beta}{\partial y^{\beta'}} g_{\alpha\beta}. \tag{11.88}$$

If $g_{\alpha\beta}$ is regular in the old coordinates, but some of the functions $\partial x^\alpha / \partial y^{\alpha'}$ have a singularity at $x^\alpha = x_0^\alpha$, then some components of $g_{\alpha'\beta'}$ will be singular at $y^\alpha = y^\alpha(x_0)$. Such a singularity can be removed by the inverse transformation $y^\alpha \to x^\alpha$.

There is no general criterion that would allow us to distinguish a "true" singularity of the Riemannian geometry from one introduced through the coordinates.[21] However, if a coordinate transformation that removes the singularity is found, then this proves that the singularity was spurious. For the Schwarzschild solution several such transformations exist. The most general of them is the one found (independently and almost simultaneously) by M. Kruskal [89] and by G. Szekeres [90] in 1960.[22]

[21] If scalars connected with the geometry of the manifold become infinite at certain points, then this is an indication that the singularity is genuine. For example, it will be seen in the following that the singularity of the Schwarzschild solution at $r = 0$ is a real one. (It occurs beyond the range of the curvature coordinates, and we will be able to deal with it only after introducing suitable other coordinates.)

[89] M. Kruskal, *Phys. Rev.* **119**, 1743 (1960).

[90] G. Szekeres, *Publicationes Mathematicae Debrecen* **7**, 285 (1960); reprinted in *Gen. Relativ. Gravit.* **34**, 2001 (2002), with an editorial note by P. Szekeres, *Gen. Relativ. Gravit.* **34**, 1995 (2002) and author's (auto)biography by G. Szekeres, *Gen. Relativ. Gravit.* **34**, 1999 (2002).

[22] The Kruskal–Szekeres transformation was a crowning of a whole series of less general results. The probably oldest proof that the singularity at $r = 2m$ is spurious was given by Lemaître

[91] G. Lemaître, *Ann. Soc. Sci. Bruxelles* **A53**, 51 (1933); reprinted in: *Gen. Relativ. Gravit.* **29**, 641

Let us write the Schwarzschild solution in the form[23]

$$ds^2 = (1 - 2m/r)\left[dt^2 - \frac{dr^2}{(1 - 2m/r)^2}\right] - r^2\left(d\vartheta^2 + \sin^2\vartheta d\varphi^2\right), \tag{11.89}$$

and transform the coordinates as follows:

$$\widetilde{p} = t + \xi, \qquad \widetilde{q} = t - \xi, \tag{11.90}$$

where

$$\xi \overset{\text{def}}{=} \int \frac{dr}{1 - 2m/r} = r + 2m \ln\left(\frac{r}{2m} - 1\right). \tag{11.91}$$

Then the metric becomes

$$ds^2 = \left(1 - \frac{2m}{r}\right) d\widetilde{p} d\widetilde{q} - r^2\left(d\vartheta^2 + \sin^2\vartheta d\varphi^2\right), \tag{11.92}$$

where now $r = r(\xi)$ is the function inverse to the $\xi(r)$ defined in (11.91). Since $\xi,_r > 0$ for $r > 2m$, the transformation (11.91) is invertible in the whole region in which the Schwarzschild solution was originally defined. The function (11.91) can be formally extended to the values $r < 2m$, by writing $\xi = r + 2m \ln |r/(2m) - 1|$, Fig. 11.8 shows the graph of the extension. The inverse functions exist in both domains, $r > 2m$ and $r < 2m$, but there exists no single inverse function in the domain $0 < r < \infty$.

Now let us carry out the next coordinate transformation

$$(p, q) = \left(e^{\widetilde{p}/a}, e^{-\widetilde{q}/a}\right) \Longleftrightarrow (\widetilde{p}, \widetilde{q}) = a(\ln p, -\ln q), \tag{11.93}$$

where a is a constant whose value will be chosen later. In these coordinates we have

$$ds^2 = -\left(1 - \frac{2m}{r}\right)\frac{a^2}{pq} dp dq - r^2\left(d\vartheta^2 + \sin^2\vartheta d\varphi^2\right), \tag{11.94}$$

where this time $r = r(p, q)$; the transformation $(\widetilde{p}, \widetilde{q}) \longrightarrow (p, q)$ is invertible at all $p > 0$ and $q > 0$. Now let us substitute for pq using (11.93) and (11.90). The result is:

$$ds^2 = -\left(1 - \frac{2m}{r}\right) a^2 e^{-2r/a}\left(\frac{r}{2m} - 1\right)^{-4m/a} dp dq - r^2\left(d\vartheta^2 + \sin^2\vartheta d\varphi^2\right). \tag{11.95}$$

(1997), with an editorial note and author's biography by A. Krasiński, *Gen. Relativ. Gravit.* **29**, 637 and 639 (1997),
who noticed that in coordinates connected with freely falling observers the surface $r = 2m$ is no obstacle to their motion. In fact, this interpretation of the Lemaître coordinates was provided much later by Novikov:

[92] I. D. Novikov, *Soobshcheniya GAISh* **132**, 3 (1964); reprinted in: *Gen. Relativ. Gravit.* **33**, 2259 (2001), with an editorial note by A. Krasiński, *Gen. Relativ. Gravit.* **33**, 2255 (2001), and author's (auto)biography by I. D. Novikov, *Gen. Relativ. Gravit.* **33**, 2258 (2001).

Other transformations removing the spurious singularity at $r = 2m$ were found by Raychaudhuri in 1953

[93] A. K. Raychaudhuri, *Phys. Rev.* **89**, 417 (1953).

and by Finkelstein in 1958

[94] D. Finkelstein, *Phys. Rev.* **110**, 965 (1958);

the Kruskal–Szekeres transformation is a generalisation of the latter.

[23] Even though the set $r = 2m$ is the boundary of the region covered by the curvature coordinates, Eq. (11.89) still makes sense when $0 < r < 2m$. In that region, t becomes a spacelike coordinate and r becomes time. The resulting metric is the vacuum Kantowski–Sachs (K–S) spacetime [15] (see Secs. 8.8 and 9.8). This is the case that is commonly overlooked in textbooks, mentioned in Sec. 11.1, Eq. (11.18).

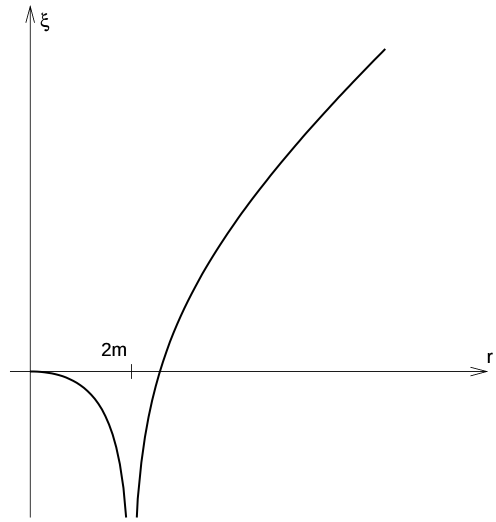

Figure 11.8: The graph of the function $\xi(r) = r + 2m \ln \left| \frac{r}{2m} - 1 \right|$.

If we choose $a = 4m$, then the singular factor cancels out and the result is

$$\mathrm{d}s^2 = -\frac{32m^3}{r}\mathrm{e}^{-r/(2m)}\mathrm{d}p\mathrm{d}q - r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right). \tag{11.96}$$

The hypersurfaces of constant p and of constant q are null, which is not always convenient. In order to go back to time-space coordinates, we introduce v and u by

$$\begin{aligned} v = \frac{1}{2}(p-q) &= \mathrm{e}^{r/(4m)}\sqrt{\frac{r}{2m}-1}\sinh\left(\frac{t}{4m}\right) \\ u = \frac{1}{2}(p+q) &= \mathrm{e}^{r/(4m)}\sqrt{\frac{r}{2m}-1}\cosh\left(\frac{t}{4m}\right). \end{aligned} \tag{11.97}$$

The metric (11.96) now becomes

$$\mathrm{d}s^2 = \frac{32m^3}{r}\mathrm{e}^{-r/(2m)}\left(\mathrm{d}v^2 - \mathrm{d}u^2\right) - r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right). \tag{11.98}$$

By following all the intermediate steps, it is seen that the transformation $(t, r) \to (v, u)$ is single-valued and nonsingular, thus invertible, for all $r \geq 2m$. At $r = 2m$, it is also well-defined, but $\partial v/\partial r$ and $\partial u/\partial r$ go to infinity as $r \to 2m$. This is the reason why (11.89) has a singularity at $r = 2m$.

In (11.98), r can be arbitrarily near 0, and the scalar components of the Riemann tensor go to infinity as $r \to 0$. This shows that the set $r = 0$ is a singularity of the geometry.

For $r < 2m$, the transformation (11.97) cannot be carried out, but it can be formally extended to all $r > 0$ by writing it in the form

$$u^2 - v^2 = \mathrm{e}^{r/(2m)}\left(\frac{r}{2m} - 1\right), \tag{11.99}$$

$$v/u = \tanh\left(\frac{t}{4m}\right). \tag{11.100}$$

In the (v, u) surface, the lines $r =$ constant are hyperbolae, the set $r = 2m$ is the pair of straight lines $u = \pm v$, and the lines $t =$ constant are straight lines passing through the point $u = v = 0$ with different slopes; $\lim_{t\to\pm\infty} v/u = \pm 1$. The singularity $r = 0$ is the pair of hyperbolae $u^2 - v^2 = -1$. The (u, v) surface thus looks as in Fig. 11.9. The radial null geodesics, which are $u \pm v =$ constant in (11.89), are in this figure straight lines inclined at the angle 45° to the axes. The straight lines $v/u = \pm 1$ divide the figure into four sectors. The curvature coordinates cover only sectors I or III, but not both at once. The region $r < 2m$ corresponds to the sectors II or IV, but again not to both of them at once.

Figure 11.9 is called the **Kruskal – Szekeres diagram**, after the two discoverers of the (v, u) coordinates [89, 90]. It provides not only useful coordinates for the Schwarzschild metric; it is at the same time the **maximal extension** of the Schwarzschild spacetime. The incompleteness of the representation in the curvature coordinates was recognised when it turned out that there exist timelike and null geodesics that escape the range of the Schwarzschild map without hitting any singularity. The Kruskal – Szekeres representation does not have this defect – any geodesic can either be continued to infinite values of the affine parameter, or hits the singularity at $r = 0$. Contrary to what the original Schwarzschild representation would have us expect, the radial null geodesics reaching the observer in sector I come from a different part of the manifold than the one to which the observer can send her light signals, and the singularity set consists of two separate parts. An observer in sector I can never get an answer to a signal that she would send through the hypersurface $r = 2m$, because any such answer will hit the future singularity before reaching $r = 2m$. However, observers in sector IV do have such a possibility: after moving into sector II, they can get answers from sectors I and III to the signals they had sent in sector IV, but their replies will have no chance to leave sector II.

11.8 Interpretation of the spurious singularity at $r = 2m$; black holes.

The set $r = 2m$ in the Schwarzschild spacetime does have some special properties, even though it is nonsingular. Equation (11.51), rewritten in the form

$$\left(\frac{\mathrm{d}r}{\mathrm{d}s}\right)^2 = E^2 - \left(1 + \frac{{J_0}^2}{r^2}\right)\left(1 - \frac{2m}{r}\right), \tag{11.101}$$

shows that $\mathrm{d}r/\mathrm{d}s \neq 0$ for all $r \leq 2m$. Hence, if a body crosses the hypersurface $r = 2m$ with decreasing r ($\mathrm{d}r/\mathrm{d}s < 0$), then it will not be able to turn back and will continue to move until it hits the singularity at $r = 0$, also when $J_0 \neq 0$. If, however, $\mathrm{d}r/\mathrm{d}s > 0$ at an initial point in the region $r < 2m$, then the body will move outwards until it enters the region $r > 2m$, and only then it can possibly turn back.

Equation (11.66) shows that light-rays have the same property.

For ordinary astronomical objects the quantity $2m = 2GM/c^2$, called the **gravitational radius**, is much smaller than the physical radius. For example, for the Sun $2m \approx 2.95$ km, for the Earth $2m = 0.89$ cm (the physical radii are 696 000 km and

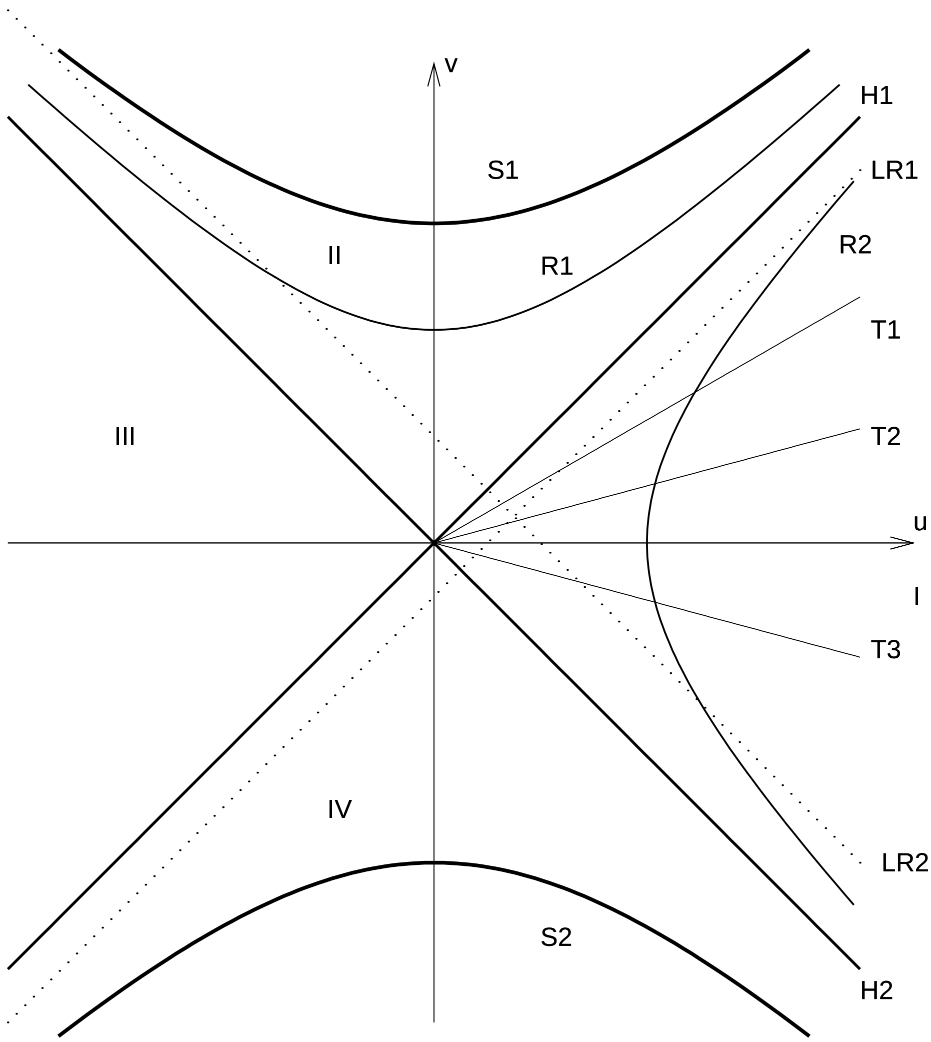

Figure 11.9: The Kruskal – Szekeres diagram of the maximally extended Schwarzschild spacetime. S_1 and S_2 are the singularities at $r = 0$, the regions above and below them are not parts of the spacetime. The straight lines H_1 and H_2 are the event horizons; they divide the plane into four sectors denoted I, ..., IV. The curvature coordinates cover only sector I or sector III, but not both of them at once. Curvature coordinates extended into the horizons cover sectors II and IV, although one does not see then that the two sectors do not coincide. The dotted lines LR_1 and LR_2 are paths of radial light-rays; LR_1 gets out from the past singularity S_2 and from under the horizon and escapes to infinity, LR_2 falls from infinity and is trapped inside the horizon, to eventually hit the future singularity S_1. The hyperbolae R_1 and R_2 are lines of constant r, one inside the horizon (with $r < 2m$) , and the other outside ($r > 2m$). T_1, T_2 and T_3 are lines of constant t. On the horizons $r = 2m$ and $|t| = \infty$, with $t = +\infty$ on H_1 and $t = -\infty$ on H_2. This had misled some early researchers into believing that the surface $r = 2m$ can never be reached by any material object. But the value of t is not the physical time, t is just a badly behaving parameter used to measure time.

6378 km, respectively). The surface $r = 2m$ for such an object is hidden deep inside it, where the spacetime metric is different from Schwarzschild's, so it has no physical meaning. This is why its existence was considered merely a mathematical curiosity well into the 1960s. If the Sun would collapse to a size $r_0 \leq 2m$, then its mean mass density would go up to $\rho_0 \geq 1.84 \times 10^{16}$ g/cm^3, which is 100 times the mean density inside a uranium nucleus and 10 times the mass density in a proton. However, $2m$ is proportional to M, while the physical radius r_p is approximately proportional to $M^{1/3}$ (because $M = (4/3)\pi\overline{\rho}{r_p}^3$, where $\overline{\rho}$ is the mean mass density in the object). Hence, for every finite density, at sufficiently large M, $2m$ will become greater than r_p, and the whole object will be inside the surface $r = 2m$. For $\overline{\rho} = 1$ g/cm^3 (the density of water) this mass is 2.7×10^{38} kg $\approx 1.36 \times 10^8$ solar masses. The corresponding $r_p = 2.68$ astronomical units, which is less than Jupiter's orbital radius.

Objects with physical radius smaller than $2m$ are called **black holes** because everything that falls into them will never get back out (see Ref. [95] for a theory of black holes in different geometries and an overview of their observations). However, the inverse behaviour is theoretically also possible: matter can be emitted from $r < 2m$ as long as there is any supply inside. These (hypothetical) objects are called **white holes**. They can be imagined as the initial cosmological singularity still going off at isolated locations; see below.

Astronomers have many candidates for black holes among observed objects. It is believed today that every galaxy contains a large-mass black hole at its center. Signatures of such candidate black holes are a high energy output by X rays (caused by particles spiralling into the black hole and radiating out energy in consequence of collisions) or a large mass concentrated in a small volume (detected through high orbital velocities of stars). We will come back to this subject in Section 15.9 and in Chapter 17.

So far, the only observed white hole is the whole Universe. In the currently accepted cosmological models, the evolution of the Universe begins with an explosion called the **Big Bang** (BB, see Chapters 12 and 14). The most popular among them are the Robertson–Walker models, in which the BB is a single event in the spacetime. The process of expansion away from the BB is the time-reverse of the collapse to a singularity, i.e. precisely what a white hole should do. In models more general than R–W the BB is not a single event, but a process extended in time; see Chapter 15. In such models, "lagging cores" of expansion may exist that would be visible to distant observers as localised white holes. They were once proposed as the explanation of the energy source in quasars [96, 97], but this explanation was later abandoned in favour of black holes with orbiting disks of matter.

In Newton's theory, the escape velocity from the surface of a spherical object of mass M and radius r is ${v_e}^2 = 2GM/r$. Note that $v_e \to c$ at $r \to r_g \overset{\text{def}}{=} 2GM/c^2$. This observation was first made by P. S. Laplace in 1795 [98]. Hence, the notion of a black hole in fact appeared already then, although the name was coined only in the 1960s.[24]

---

[95] V. P. Frolov and I. D. Novikov, *Black Hole Physics: Basic Concepts & New Developments.* Kluwer, Amsterdam 1998, 770 pp.

[96] I. D. Novikov, *Astron. Zh.* **41**, 1075 (1964); English translation: *Sov. Astr. A. J.* **8**, 857 (1965).

[97] Y. Neeman and G. Tauber, *Astrophys. J.* **150**, 755 (1967).

[98] P. S. Laplace, *Exposition du système du monde*, 1795; cited after [83].

[24] The analogy between a black hole and Laplace's "dark star" must not be taken literally. For a body that fell into a black hole it is impossible to get back out. In Newton's theory, no object can escape from

The Kruskal – Szekeres diagram allows us to follow the object falling into a black hole. The lines $u \pm v =$ constant are intersections of the diagram plane with light cones (two of them are shown with dotted lines in Fig. 11.9). Any timelike line (not necessarily geodesic) in that diagram must thus have its tangent inclined to the v axis at an angle smaller than 45° at every point. Imagine a body proceeding toward the black hole and emitting light pulses at regular intervals of its proper time s. Suppose they are picked up by an observer resting at $r = r_0 \gg 2m$; the hyperbola R_2 in Fig. 11.9 may represent her world line. When the emitter approaches the surface $r = 2m$, the observer receives the pulses at increasing intervals of coordinate time t (see (11.49) which tend to infinity as $r_{\rm sender} \to 2m$. The pulse sent at $r = 2m$ will stay in the set $r = 2m$ for ever and will never reach the observer. For this reason, this set is called the **event horizon**. All signals sent at $r < 2m$ will hit the singularity at $r = 0$. Thus, from the point of view of a distant observer, the process of falling into a black hole lasts infinitely long. The same is true for an object (a star or a galaxy) that collapses to form a black hole. Hence, one should not imagine a black hole as a "finished" object existing somewhere out there. A distant observer only has a chance to see a black hole being formed, as an object whose light is becoming constantly darker and redder. Bodies falling into a black hole will disappear from sight *before* they hit the event horizon.[25]

For an observer who decided to fall into a black hole, the proper time needed to reach the event horizon is finite. This is seen from (11.101), which, for radial motion toward the center (i.e. with $J_0 = 0$), may be written as

$$\frac{{\rm d}r}{{\rm d}s} = -\left(E^2 - 1 + \frac{2m}{r}\right)^{1/2}, \tag{11.102}$$

and then the time of flight from a surface $r = r_0$ to the horizon $r = 2m$ is

$$s_h = -\int_{r_0}^{2m} \left(E^2 - 1 + \frac{2m}{r}\right)^{-1/2} {\rm d}r. \tag{11.103}$$

This is an elementary integral that has a finite value for every $r_0 < \infty$.

11.9 The Schwarzschild solution in other coordinate systems.

The curvature coordinates were the oldest introduced for the Schwarzschild solution, and in a way the most natural. However, for investigating some advanced topics, in particular

---

the "dark star" *to infinity*, but the surface $r = 2GM/c^2$ is freely traversable in both directions.

[25] Because of high symmetry of the Schwarzschild metric, the event horizon in it coincides with two other entities that are in general distinct. One of them is the *apparent horizon*, AH. The future AH is a hypersurface in spacetime within which all null geodesics can proceed only toward the singularity, never away. In nonstatic spacetimes the AH and the event horizon in general do not coincide (see chapter 15). In gravitational fields of rotating bodies, the *stationary limit hypersurface* (SLH) does not in general coincide with the event horizon (see chapter 17). From the SLH, light signals arrive to infinity being infinitely redshifted, and yet the SLH is traversable both ways for material particles and light-rays.

the relation of the Schwarzschild solution to other solutions of Einstein's equations, other coordinates are sometimes more useful. Two coordinate systems are used most often.

In **isotropic coordinates**, the metric of the subspace $t =$ constant is proportional to the Euclidean metric. Take the Schwarzschild metric (11.34) & (11.36), and transform the r-coordinate as follows:

$$r = r' \left(1 + \frac{m}{2r'}\right)^2 . \tag{11.104}$$

The resulting metric is:

$$\mathrm{d}s^2 = \frac{\left(1 - \frac{m}{2r'}\right)^2}{\left(1 + \frac{m}{2r'}\right)^2} \, \mathrm{d}t^2 - \left(1 + \frac{m}{2r'}\right)^4 \left[\mathrm{d}r'^2 + r'^2 \left(\mathrm{d}\vartheta^2 + \sin^2 \mathrm{d}\varphi^2\right)\right] . \tag{11.105}$$

The spurious singularity now appears at $r' = m/2$. In these coordinates, a large number of nonvacuum generalisations of the Schwarzschild solution was found [99].

Also frequently used are comoving coordinates of a congruence of radially freely falling observers. They are called **geodesic**, **Lemaître-** or **Novikov coordinates**. The Schwarzschild metric in them is:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{{R,_r}^2 \mathrm{d}r^2}{1 + 2E(r)} - R^2(t, r) \left(\mathrm{d}\vartheta^2 + \sin^2 \mathrm{d}\varphi^2\right) , \tag{11.106}$$

where the function $R(t, r)$ obeys the equation:

$${R,_t}^2 = \frac{2m}{R} + 2E(r), \tag{11.107}$$

$E(r)$ being an arbitrary function. A subcase of this coordinate system, corresponding to $E = 0$, was first introduced by Lemaître [91][26]. The nonzero values of E were investigated, and the physical interpretation of all cases was provided by Novikov [92]. Equation (11.107) is the equation of radial free fall in the spherically symmetric gravitational field generated by the mass m, with $E(r)$ being the kinetic energy of the observer at infinity (and the total conserved energy of the motion). An observer at fixed r (note that r is in fact arbitrary in this metric, transformations of the form $r = f(r')$ do not change it) proceeds with time to other values of R, i.e. is either receding from or approaching the centre of symmetry, in accordance with the equation of free fall. When $E > 0$, the observers can recede to infinity and have nonzero kinetic energy there. With $E = 0$, they can recede to infinity, but their kinetic energy decreases to zero. With $E < 0$, the observers can fly away from the centre of symmetry only out to a finite distance, and then fall back. For each sign of E, (11.107) may be explicitly solved for $t(R)$, but for $E \neq 0$ the solutions cannot be inverted to define an explicit elementary function $R(t, r)$.

With $E \geq 0$, the metric (11.106) has no singularity anywhere apart from $R = 0$. As mentioned before, this is how Lemaître first noticed that the singularity of the Schwarzschild solution at $r = 2m$ is only an artefact of the coordinates used.

---

[99] A. Krasiński, *Inhomogeneous cosmological models.* Cambridge University Press 1997.

[26] Lemaitre's form was somewhat unreadable because he parametrised his metric in such a way that the limit of zero cosmological constant could not be directly taken.

The Schwarzschild metric in the form (11.106) – (11.107) emerges as the vacuum limit of the **Lemaître–Tolman** cosmological model that will be discussed in chapter 15. Equation (11.107) frequently appears in cosmology – it governs the evolution of the Lemaître–Tolman and Friedmann models (the latter is the spatially homogeneous limit of the former). It is in fact simpler to obtain by assuming geodesic coordinates (in which $g_{00} = 1$, $g_{01} = 0$, other components as in (8.50)) and solving the Einstein equations.

11.10 The equation of hydrostatic equilibrium

We will now investigate the Einstein equations *inside* a spherically symmetric body of perfect fluid, with the additional assumption that matter is at rest in the coordinates of Section 11.1. Then the velocity field has only the t-component:

$$u^\alpha = \mathrm{e}^{-\nu}\delta^\alpha{}_0, \qquad u_\alpha = \mathrm{e}^{\nu}\delta^0{}_\alpha, \tag{11.108}$$

where $\nu(t, r)$ is an unknown function, while the pressure and density of the perfect fluid are constant along the lines of flow, i.e. depend only on r. Using (11.108) and (11.21) – (11.24), we obtain the following set of equations:[27]

$$G_{00} = \mathrm{e}^{2(\nu-\mu)}\left(\frac{2}{r}\mu' - \frac{1}{r^2}\right) + \frac{\mathrm{e}^{2\nu}}{r^2} = \frac{8\pi G}{c^4}\mathrm{e}^{2\nu}\epsilon, \tag{11.109}$$

$$G_{01} = \frac{2}{r}\,\dot{\mu} = 0, \tag{11.110}$$

$$G_{11} = \frac{2}{r}\,\nu' + \frac{1}{r^2} - \frac{\mathrm{e}^{2\mu}}{r^2} = \frac{8\pi G}{c^4}p\mathrm{e}^{2\mu}, \tag{11.111}$$

$$G_{22} = G_{33}/\sin^2\vartheta = r^2\mathrm{e}^{-2\mu}\left(\nu'' - \mu'\nu' + \nu'^2 + \frac{\nu' - \mu'}{r}\right) - r^2\mathrm{e}^{-2\nu}\left(\ddot{\mu} - \dot{\mu}\dot{\nu} + \dot{\mu}^2\right) = \frac{8\pi G}{c^4}pr^2. \tag{11.112}$$

Equation (11.110) implies that $\dot{\mu} = 0$, like in the Schwarzschild solution. Since $\dot{p} = \dot{\epsilon} = 0$, differentiating (11.111) by t we obtain $\dot{\nu}' = 0$, i.e. $\nu = f(t) + g(r)$. Then, the coordinate transformation $t' = \int \mathrm{e}^{f(t)}\mathrm{d}t$ gives $f = 0$ in the new coordinates, and $\nu = \nu(r)$.

Equation (11.109) can now be formally integrated. Assuming $\epsilon(0) < \infty$ we obtain

$$\mathrm{e}^{-2\mu} = 1 - \frac{8\pi G}{c^4 r}\int_0^r \epsilon(r')r'^2\mathrm{d}r'. \tag{11.113}$$

We denote

$$M(r) \stackrel{\text{def}}{=} \frac{4\pi}{c^2}\int_0^r \epsilon(r')r'^2\mathrm{d}r'. \tag{11.114}$$

This quantity has the dimension of mass and plays in (11.19) the same role as the mass parameter played in the Schwarzschild metric: if $\epsilon(r) \equiv 0$ for $r \geq R$ (i.e. if the surface

[27] By choosing the metric in the form (11.19) we left aside the cases when $\delta_{,\alpha}$ in (8.50) is a timelike or null vector, or is zero. Those cases do not contain static sources for the Schwarzschild metric, but they cannot be neglected when considering nonstatic sources.

of the matter distribution is at $r = R$), then $M(R)$ is equal to the mass parameter of the Schwarzschild metric. Consequently, we interpret $M(r)$ as the mass within the sphere of radius r. Note that $M(r)$ is smaller than the sum of rest masses of all the particles inside this sphere. The density of rest mass is $\rho(r) = \epsilon(r)/c^2$, so the sum of rest masses is

$$M_{\text{rest}} = \int_{Vol_3(r)} \rho(r')\sqrt{g_3}\, \mathrm{d}_3 x = \frac{4\pi}{c^2}\int_0^r \epsilon(r') r'^2 \mathrm{e}^{\mu(r')}\mathrm{d}r'. \tag{11.115}$$

Since $\mathrm{e}^{\mu} > 1$ (from (11.113)), we have $M_{\text{rest}} > M(r)$. This is the large-scale (astrophysical) analogue of the 'mass defect' known from nuclear/elementary particle physics: the mass of a bound object is smaller than the sum of masses of its components. The mass defect multiplied by c^2 is equal to the energy that would have to be supplied in order to break the object up into separate particles.

Using the notation (11.114), we obtain in (11.113)

$$\mathrm{e}^{-2\mu} = 1 - \frac{2GM(r)}{c^2 r}. \tag{11.116}$$

The equations of motion $T^{\alpha\beta}{}_{;\beta} = 0$ must hold in consequence of the field equations. For $\alpha = 0, 2, 3$ they are fulfilled identically, while $T^{1\beta}{}_{;\beta} = 0$ implies

$$\nu' = -\frac{p'}{\epsilon + p}. \tag{11.117}$$

This can be integrated after an equation of state $\epsilon = \epsilon(p)$ is assumed.

Using (11.116) and (11.117), Eq. (11.111) can now be written as

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{G\left(\rho + p/c^2\right)\left[M(r) + 4\pi r^3 p/c^2\right]}{r^2\left[1 - 2GM(r)/\left(c^2 r\right)\right]}. \tag{11.118}$$

Equation (11.112) is now fulfilled by virtue of those already solved.

Equation (11.118) is called the **equation of hydrostatic equilibrium**. In the Newtonian limit $c \to \infty$ it becomes

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{G\rho M(r)}{r^2}. \tag{11.119}$$

Comparing (11.119) with (11.118) we see that in relativity pressure increases the gravitational attraction because it appears as a positive contribution to the mass and mass density. Given the same mass, density and pressure, the *gradient* of pressure, opposing the gravitational attraction, is greater in (11.118) than in (11.119). The greater pressure gradient implies faster growth of pressure towards the centre of the object, and this increased pressure then requires a still greater gradient. One can thus imagine a situation in which the equilibrium maintained over some time by a certain process (such as energy production in a star) is perturbed, and then the growth of pressure will lead to a loss of stability: no pressure gradient will be able to keep the object static again. A collapse to a size smaller than the gravitational radius $r_g = 2GM/c^2$ will then occur, and a black hole will be created. The Newtonian equilibrium equation (11.119) does not allow such a situation: the gradient of pressure needed to maintain equilibrium is determined by the mass density at the distance r from the centre of the object and does not depend on the value of pressure, so the solution exists for every density and mass.

11.11 The 'interior Schwarzschild solution'

Equation (11.117) can be integrated after an equation of state, or a distribution of mass density inside the object, has been defined. Various equations of state and mass distributions intended to imitate the real conditions inside various stellar objects are discussed in relativistic astrophysics. Typically, such general equations of state lead to complicated forms of (11.117) that can be integrated only numerically. We will deal here with an example that is unrealistic, but it illustrates in a simple way the problems encountered while matching matter solutions with vacuum solutions. We assume that $\epsilon =$ constant, a constant mass density inside the object. From (11.112) – (11.114) we then obtain:

$$\mathrm{e}^{-2\mu} = 1 - \frac{8\pi G}{3c^4}\epsilon r^2 \stackrel{\text{def}}{=} 1 - Dr^2, \tag{11.120}$$

$$M(r) = \frac{4\pi}{3c^2}\epsilon r^3 = \frac{c^2}{2G}Dr^3, \tag{11.121}$$

where we used the abbreviation $D \stackrel{\text{def}}{=} 8\pi G\epsilon/(3c^4) = \text{constant} > 0$. Subtracting (11.111) from (11.112) and using (11.120) we obtain

$$\left(1 - Dr^2\right)\left(\nu'' + \nu'^2\right) - Dr\nu' - \frac{1}{r}\left(1 - Dr^2\right)\nu' = 0, \tag{11.122}$$

which is integrated with the result $\mathrm{e}^\nu = C - B\sqrt{1 - Dr^2}$, where B and C are constants. The pressure can now be calculated from (11.111):

$$p = \frac{1}{3}\,\epsilon\,\frac{3B\sqrt{1 - Dr^2} - C}{C - B\sqrt{1 - Dr^2}}, \tag{11.123}$$

and it obeys (11.118). Hence, we have obtained a complete solution of (11.109) – (11.112):

$$\mathrm{d}s^2 = \left(C - B\sqrt{1 - Dr^2}\right)^2 \mathrm{d}t^2 - \frac{\mathrm{d}r^2}{1 - Dr^2} - r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right). \tag{11.124}$$

This is called the **interior Schwarzschild solution** [100].

Now let us verify the matching conditions between (11.124) and the vacuum Schwarzschild solution, (11.34) and (11.36), at the hypersurface Σ given by $r = R =$ constant. The coordinate systems on both sides of Σ are already adapted, so we can use the formalism of Section 10.13. Here, $x^4 = r$. The continuity of g_{ab}, $a \neq 4 \neq b$, requires

$$\left(C - B\sqrt{1 - DR^2}\right)^2 = 1 - \frac{2GM}{c^2R}, \tag{11.125}$$

and then the continuity of the derivatives of g_{ab} in directions tangent to Σ is guaranteed. Following the recipe of Section 10.13, we require the continuity of $\left[\mathrm{d}(g_{ab})/\mathrm{d}r\right]/\mathcal{N}$ at $r = R$, where $\mathcal{N} = 1/\sqrt{1 - 2GM/(c^2R)}$ in the exterior metric and $\mathcal{N} = 1/\sqrt{1 - DR^2}$ in the interior metric. This imposes two more equations:

$$\sqrt{1 - DR^2} = \sqrt{1 - \frac{2GM}{c^2R}}, \tag{11.126}$$

[100] K. Schwarzschild, *Sitzber. Preuss. Akad. Wiss.* p. 424 (1916).

$$2DBR\left(C - B\sqrt{1 - DR^2}\right) = \frac{2GM}{c^2R^2}\sqrt{1 - \frac{2GM}{c^2R}}. \tag{11.127}$$

(The first one results from $\mathrm{d}(g_{22})/\mathrm{d}r$ being equal on both sides of Σ, which implies $\mathcal{N}$ being the same on both sides.) The solution of Eqs. (11.125) – (11.127) is

$$D = \frac{2GM}{c^2R^3}, \qquad B = \frac{1}{2}, \qquad C = \frac{3}{2}\sqrt{1 - \frac{2GM}{c^2R}}. \tag{11.128}$$

This guarantees that $p(R) = 0$; see (11.123).

11.12 Exercises.

1. Prove that (11.2) is indeed a rotation around the axis $\{\vartheta = \pi/2, \varphi = 0\}$. One way to do this is to find the corresponding Killing field and transform it to Cartesian coordinates in which it should reduce to the generator of rotations around the x-axis, $\underset{[y,z]}{k^\mu} = y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}$. Verify that the components of (11.1) do not change after the transformation (11.2).

2. Prove Lemmas 11.2 and 11.3.

3. Prove that with $E < 1$, $J_0 > 2\sqrt{3}GM/c^2$ and the initial σ being sufficiently small, the solution $\sigma(\varphi)$ of (11.53) must be bounded (i.e. with the initial r being sufficiently large and the other conditions fulfilled, $r(\varphi)$ cannot go to zero or to infinity).

Hint: Equation (11.53) expresses the energy conservation on a Newtonian orbit in the potential $[-\left(1 + {J_0}^2/r^2\right)(1 - 2m/r)]$. Consider the possible shapes of this potential depending on the range of values of J_0. Draw the graphs of the potential for all the cases and consider the orbital motion in various ranges of the total energy E.

4. For a light ray on a circular orbit $\mathrm{d}\sigma/\mathrm{d}\varphi \equiv 0$ in (11.68), so the radius of this orbit must obey

$$E^2r^3 - {J_0}^2(r - 2m) = 0. \tag{11.129}$$

This equation has at most two solutions (because the third one is always negative, while r must be positive). However, most of these orbits are unstable: any smallest perturbation will cause the light ray to go around an orbit on which r oscillates between a maximal and a minimal value. Find the condition that J_0, m and E must obey for circular photon orbits to exist. Show that only one of them is stable, with radius $r = 3m$.

Hint: Use the same method as in the previous exercise. For a stable circular orbit the function $E^2 - {J_0}^2\sigma^2(1 - 2m\sigma)$ has a zero and a local minimum at the same radius.

5. Prove that the same (incorrect) result (11.83) for the deflection of light-rays can be obtained by pure Newtonian methods.

Hint: Consider a hyperbolic orbit in polar coordinates, (11.58), where the geometric parameters p and ε are connected with the physical parameters by (11.80) and

$$p = \frac{J^2}{GM\mu^2}. \tag{11.130}$$

Calculate the deflection angle as in (11.79), then replace the geometric parameters by the physical ones. Note that $J = mRv_R$, where R is the smallest distance between the particle and the central star, and v_R is the velocity at the point of smallest distance. Calculate v_R by equating the energy at that point to the energy at infinity, $\frac{1}{2}\mu {v_\infty}^2$. Note that the value of the deflection angle does not depend on the mass μ. Assume that the "particle" is a photon and that $v_\infty = c$. Finally, calculate the result up to linear terms in GM/c^2.

Remark: This is how the angle of deflection of a light ray was calculated by Cavendish already in 18th century, in unpublished notes [81], and by Johann von Soldner in 1804 [82, 83]. Einstein himself found at first this incorrect result before he formulated his field equations (10.22).

6. Calculate the scalar components R_{ijkl} of the Riemann tensor for the Schwarzschild solution (11.34) – (11.35) in the basis (11.87) and see that they are all regular at $r = 2m$.

7. Calculate the integral in eq. (11.103) and verify that it is finite for every $r_0 < \infty$.

8. Assume geodesic coordinates in (8.50), so that $\mathrm{d}s^2 = \mathrm{d}t^2 - S^2(t,r)\mathrm{d}r^2 - R^2(t,r) \left(\mathrm{d}\vartheta^2 + \sin^2 \mathrm{d}\varphi^2\right)$ (make sure first that the transformations (8.51) do really allow such a choice!). Then solve the vacuum Einstein equations $G_{\mu\nu} = 0$ for this metric and verify that (11.106) – (11.107) is the general solution that results when $R,_r \neq 0$.

9. Prove by direct coordinate transformation that (11.106) – (11.107) is indeed a representation of the Schwarzschild solution (11.34) & (11.35).

Hint: Take the Schwarzschild solution (11.34) & (11.36) and transform the coordinates by $t = f(\tau, u)$, $r = R(\tau, u)$, where f and R are functions as yet unknown. Then demand that in the new coordinates $g_{\tau\tau} = 1$ and $g_{\tau u} = 0$. Solve the resulting equations algebraically for the derivatives of f. Since the solution has the form $f,_\tau =$ (given) and $f,_u =$ (given), the integrability condition $f,_{\tau u} = f,_{u\tau}$ must be imposed. The integrability condition turns out to be a differential equation defining R. You will be surprised how simply this all works. (The τ and u are the (t,r) coordinates of eqs. (11.106) – (11.107).)

10. Verify that eq. (11.112) is indeed fulfilled in virtue of (11.109) – (11.111).

Hint: Differentiate (11.111) by r, then manipulate eqs. (11.109) – (11.112) and (11.117) until you reproduce (11.112).