14

Gravitation

In this chapter we explore the content of the gravitational field equations derived in section 13.5. In covariant notation these equations are

$$\begin{aligned} \mathcal{G}(a) - \Lambda a &= \kappa \mathcal{T}(a), \\ \mathcal{H}(a) &= \kappa \mathcal{S}(a) + \tfrac{1}{2}\kappa\big(\partial_b \cdot \mathcal{S}(b)\big)\wedge a, \end{aligned} \tag{14.1}$$

where $\kappa = 8\pi G$, Λ is the cosmological constant, $\mathcal{G}(a)$ and $\mathcal{H}(a)$ denote the Einstein and torsion tensors, and the matter sources are determined by the total energy-momentum tensor $\mathcal{T}(a)$ and the spin tensor $\mathcal{S}(a)$. Locally, the field equations define an Einstein–Cartan theory of gravitation.

We start this chapter with a discussion of the various strategies we can adopt for solving the field equations. In particular, we focus on a new technique that is unique to the gauge theory approach. Of course, the physical content of the equations does not depend on the method of solution. But the field equations have proved so resistant to analysis that it is important to have a wide range of analytical approaches at our disposal. Most of the applications of interest do not involve macroscopic spin, so the torsion is set to zero. The only exception is when we consider self-consistent cosmological models for a single spinor field in a gravitational background.

As a first application of our solution method we study spherically-symmetric, time-dependent systems. This setup is sufficiently general to use for studying non-rotating stars and black holes, and also cosmology. We study the properties of both classical and quantum matter in these backgrounds, looking in detail at scattering and absorption processes around a black hole. We then turn to static cylindrical systems. These are of limited astrophysical interest, but they do demonstrate some important features of our solution method. In particular we find that, for certain matter distributions, the gravitational fields admit closed timelike curves. These matter distributions can give rise to violations of causality, which are therefore not ruled out by the theory without further assumptions. We end this chapter with a discussion of axially-symmetric fields and the Kerr solution. We give a novel derivation of the Kerr solution, which exposes a remarkable algebraic structure hidden in other approaches. We also describe a version of the Kerr solution that illustrates many of its physical features in a straightforward manner.

14.1 Solving the field equations

The traditional approach to solving the gravitational field equations in general relativity is to start with the metric $g_{\mu\nu}$. In equation (13.143) we showed that the metric is recovered from the $\mathsf{h}(a)$ gauge field by setting

$$g_{\mu\nu} = g_\mu \cdot g_\nu = \mathsf{h}^{-1}(e_\mu)\cdot\mathsf{h}^{-1}(e_\nu), \tag{14.2}$$

where the $\{e_\mu\}$ comprise a coordinate frame. The metric $g_{\mu\nu}$ is invariant under rotation-gauge transformations, so working in terms of the metric removes this gauge freedom from the outset. The result is that the field equations become a set of non-linear, second-order differential equations for the terms in $g_{\mu\nu}$. Any metric is potentially a solution of the field equations — one where the matter energy-momentum tensor is determined by the corresponding Einstein tensor. But this is seldom useful, as what is required is a solution for a given matter distribution. This is an extremely difficult problem.

A related shortcoming of the metric approach is that it is extremely difficult to set up a consistent perturbative scheme. The problem is that the metric is gauge-dependent, so it is not apparent which quantities can be treated as small. This can only be defined consistently in terms of covariant scalars, as these are the only gauge-invariant quantities. Clearly, then, we should aim to solve the equations directly in terms of these quantities. Such a method is described here, and applied to a range of problems in this chapter.

We start by focusing on objects that transform covariantly under displacements. For ease of reference we call these *intrinsic* objects. Unlike the metric formulation, the class of intrinsic objects in the gauge treatment extends beyond scalars to include general multivectors and functions. For example, each of $\bar{\mathsf{h}}(\nabla)$, $\omega(a)$ and $\mathcal{R}(B)$ are intrinsic objects. The task is to formulate the field equations directly in terms of these objects. We assume that the spin is negligible, so that the second field equation in (14.1) states that the torsion is zero. The method we describe here can therefore be directly applied to problems in general relativity.

The torsion equation relates derivatives of $\bar{\mathsf{h}}(a)$ to the $\omega(a)$ field, where $\omega(a)$ is defined in equation (13.245). The torsion equation can be written as

$$\bar{\mathsf{h}}(\dot{\nabla})\wedge\dot{\bar{\mathsf{h}}}(c) = -\partial_d\wedge\big(\omega(d)\cdot\bar{\mathsf{h}}(c)\big), \tag{14.3}$$

which we contract with $a\wedge b$ to form

$$\begin{aligned}\langle b\wedge a\,\bar{\mathsf{h}}(\dot{\nabla})\wedge\dot{\bar{\mathsf{h}}}(c)\rangle &= -\langle b\wedge a\,\partial_d\wedge\big(\omega(d)\cdot\bar{\mathsf{h}}(c)\big)\rangle\\ &= \big(a\cdot\omega(b)-b\cdot\omega(a)\big)\cdot\bar{\mathsf{h}}(c).\end{aligned} \tag{14.4}$$

The essential operator on the left-hand side is the directional derivative $a\cdot\bar{\mathsf{h}}(\nabla)$. This turns out to be the key operator in our approach, and we write this as

$$L_a = a\cdot\bar{\mathsf{h}}(\nabla). \tag{14.5}$$

The use of L_a for this operator should not be confused with the quantum-mechanical angular momentum operators, though their properties are analysed in a similar way. In terms of L_a the torsion equation becomes

$$\big(\dot{L}_a\dot{\mathsf{h}}(b)-\dot{L}_b\dot{\mathsf{h}}(a)\big)\cdot c = \big(a\cdot\omega(b)-b\cdot\omega(a)\big)\cdot\bar{\mathsf{h}}(c), \tag{14.6}$$

where, as usual, the overdots determine the scope of a differential operator.

The information contained in the torsion equation is summarised neatly in terms of the commutator bracket of the L_a operators. We find that the commutator of L_a and L_b is

$$\begin{aligned}[L_a,L_b] &= \big(L_a\mathsf{h}(b)-L_b\mathsf{h}(a)\big)\cdot\nabla\\ &= \big(\dot{L}_a\dot{\mathsf{h}}(b)-\dot{L}_b\dot{\mathsf{h}}(a)\big)\cdot\nabla+(L_ab-L_ba)\cdot\bar{\mathsf{h}}(\nabla)\\ &= \big(a\cdot\omega(b)-b\cdot\omega(a)+L_ab-L_ba\big)\cdot\bar{\mathsf{h}}(\nabla).\end{aligned} \tag{14.7}$$

We can therefore write

$$[L_a,L_b] = L_c, \tag{14.8}$$

where

$$c = a\cdot\omega(b)-b\cdot\omega(a)+L_ab-L_ba = a\cdot\mathcal{D}\,b-b\cdot\mathcal{D}\,a. \tag{14.9}$$

This bracket structure summarises the intrinsic content of the torsion equation in a very convenient manner. If spin is present, the right-hand side of equation (14.9) is modified in a straightforward way to include spin-dependent terms.

The key to our strategy is that we delay any explicit solution for $\omega(a)$ until after further gauge fixing has been performed. Instead, we let $\omega(a)$ take on a suitably general form, consistent with the form of the $\bar{\mathsf{h}}$ function. This is often best achieved with the aid of a symbolic algebra package, though it is possible, if tedious, to perform the calculations by hand. Once a general form for $\omega(a)$ has been found, the relationship between $\bar{\mathsf{h}}(a)$ and $\omega(a)$ is then encoded intrinsically in the commutation relations of the L_a.

The next object to form is the Riemann tensor $\mathcal{R}(B)$. This is constructed in terms of abstract first-order derivatives of the $\omega(a)$ and additional quadratic terms. We see this by writing

$$\begin{aligned}\mathcal{R}(a\wedge b) &= \dot{L}_a\dot{\Omega}\big(\mathsf{h}(a)\big) - \dot{L}_b\dot{\Omega}\big(\mathsf{h}(a)\big) + \omega(a)\times\omega(b) \\ &= L_a\omega(b) - L_b\omega(a) + \omega(a)\times\omega(b) - \Omega\big(L_a\mathsf{h}(b) - L_b\mathsf{h}(a)\big),\end{aligned} \tag{14.10}$$

so that we have

$$\mathcal{R}(a\wedge b) = L_a\omega(b) - L_b\omega(a) + \omega(a)\times\omega(b) - \omega(c), \tag{14.11}$$

where c is given by equation (14.9). Equation (14.11) enables $\mathcal{R}(B)$ to be calculated entirely in terms of intrinsic quantities. Once the general form of the Riemann tensor is found, we can start to employ the rotation-gauge freedom to convert $\mathcal{R}(B)$ to a suitably simple expression. This gauge fixing is crucial in order to arrive at a set of equations that are not underconstrained. The gauge fixing is now performed directly at the level of the covariant variables. This gives the method great power, as one can motivate gauge choices on sensible physical grounds, rather than blind guesswork at the level of the metric.

With $\mathcal{R}(B)$ suitably fixed, we arrive at a set of relations between first-order abstract derivatives of the $\omega(a)$, quadratic terms in $\omega(a)$ and matter terms. The next step is to impose the Bianchi identities, which ensure overall consistency of the equations with the bracket structure. Once all this is achieved, one arrives at a fully intrinsic set of equations. Solving these equations usually involves searching for natural integrating factors. The final step is to make an explicit position gauge choice of the h function. The natural way to do this is often to ensure that the form of $\bar{\mathsf{h}}(a)$ is such that the integrating factors are expressed simply in terms of the chosen coordinates. This description is quite abstract, but in the following sections we apply this scheme to a range of physical problems. These should illustrate how the scheme is applied in practice. We start with the simplest case of spherically-symmetric, torsion-free systems.

14.2 Spherically-symmetric systems

To solve the field equations for spherically-symmetric systems, we first introduce the standard polar coordinates (t, r, θ, ϕ). In terms of the fixed $\{\gamma_\mu\}$ frame we write

$$\begin{aligned} t &= x\cdot\gamma_0, & \cos(\theta) &= \frac{x\cdot\gamma^3}{r}, \\ r &= \sqrt{(x\wedge\gamma_0)^2}, & \tan(\phi) &= \frac{x\cdot\gamma^2}{x\cdot\gamma^1}. \end{aligned} \tag{14.12}$$

The associated coordinate frame is

$$\begin{aligned}
e_t &= \gamma_0, \\
e_r &= \sin(\theta)\big(\cos(\phi)\,\gamma_1 + \sin(\phi)\,\gamma_2\big) + \cos(\theta)\,\gamma_3, \\
e_\theta &= r\cos(\theta)\big(\cos(\phi)\,\gamma_1 + \sin(\phi)\,\gamma_2\big) - r\sin(\theta)\,\gamma_3, \\
e_\phi &= r\sin(\theta)\big(-\sin(\phi)\,\gamma_1 + \cos(\phi)\,\gamma_2\big),
\end{aligned} \tag{14.13}$$

and we will also make use of the unit vectors $\hat{\theta}$ and $\hat{\phi}$ defined by

$$\hat{\theta} = \frac{1}{r} e_\theta, \qquad \hat{\phi} = \frac{r}{\sin(\theta)} e_\phi. \tag{14.14}$$

From these we define the unit bivectors

$$\sigma_r = e_r e_t, \qquad \sigma_\theta = \hat{\theta} e_t, \qquad \sigma_\phi = \hat{\phi} e_t. \tag{14.15}$$

For applications in gravity there is little reason to write these spatial bivectors in bold face, so we break the convention adopted earlier in this book and leave the unit bivectors in ordinary face. We work throughout in natural units $c = \hbar = G = 1$, so that $\kappa = 8\pi$, and in the first instance we set the cosmological constant to zero.

14.2.1 The spherical equations

Our first step towards a solution is to decide a suitable form for the $\bar{\mathsf{h}}$ function consistent with spherical symmetry. The form we use is

$$\begin{aligned}
\bar{\mathsf{h}}(e^t) &= f_1 e^t + f_2 e^r, \\
\bar{\mathsf{h}}(e^r) &= g_1 e^r + g_2 e^t, \\
\bar{\mathsf{h}}(e^\theta) &= \alpha e^\theta, \\
\bar{\mathsf{h}}(e^\phi) &= \alpha e^\phi,
\end{aligned} \tag{14.16}$$

where f_1, f_2, g_1, g_2 and α are all functions of t and r only. The only rotation-gauge freedom in this system is the freedom to perform a boost in the σ_r direction. This freedom will be employed later to simplify the equations. Our remaining position-gauge freedom lies in the freedom to reparameterise t and r, which does not affect the general form of $\bar{\mathsf{h}}(a)$. A natural parameterisation will emerge once the physical variables have been identified.

To find a general form $\omega(a)$ consistent with the $\bar{\mathsf{h}}$ function of equation (14.16), we substitute the latter into equation (13.250) and look at the general algebraic form of $\omega(a)$. Where the coefficients in $\omega(a)$ contain derivatives of terms from $\bar{\mathsf{h}}(a)$ new symbols are introduced. Undifferentiated terms from $\bar{\mathsf{h}}(a)$ appearing in $\omega(a)$ arise from frame derivatives and are left in explicitly. The result is that

	σ_r	σ_θ	σ_ϕ
$e_t \cdot \mathcal{D}$	0	$G I \sigma_\phi$	$-G I \sigma_\theta$
$e_r \cdot \mathcal{D}$	0	$F I \sigma_\phi$	$-F I \sigma_\theta$
$\hat{\theta} \cdot \mathcal{D}$	$T \sigma_\theta - S I \sigma_\phi$	$-T \sigma_r$	$S I \sigma_r$
$\hat{\phi} \cdot \mathcal{D}$	$T \sigma_\phi + S I \sigma_\theta$	$-S I \sigma_r$	$-T \sigma_r$

Table 14.1 *Covariant derivatives of the polar-frame unit timelike bivectors.*

we can write

$$\begin{aligned} \omega(e_t) &= G e_r e_t, \\ \omega(e_r) &= F e_r e_t, \\ \omega(\hat{\theta}) &= S \hat{\theta} e_t + (T - \alpha/r) e_r \hat{\theta}, \\ \omega(\hat{\phi}) &= S \hat{\phi} e_t + (T - \alpha/r) e_r \hat{\phi}, \end{aligned} \tag{14.17}$$

where G, F, S and T are functions of t and r only. The important feature of these functions is that they transform covariantly under displacements of r and t.

To define a suitable bracket structure we first introduce the operators

$$\begin{aligned} L_t &= e_t \cdot \bar{\mathsf{h}}(\nabla), \qquad & L_{\hat{\theta}} &= \hat{\theta} \cdot \bar{\mathsf{h}}(\nabla), \\ L_r &= e_r \cdot \bar{\mathsf{h}}(\nabla), \qquad & L_{\hat{\phi}} &= \hat{\phi} \cdot \bar{\mathsf{h}}(\nabla). \end{aligned} \tag{14.18}$$

Equation (14.8), together with our form for $\omega(a)$, yields the relations

$$\begin{aligned} [L_t, L_r] &= G L_t - F L_r, \qquad & [L_r, L_{\hat{\theta}}] &= -T L_{\hat{\theta}}, \\ [L_t, L_{\hat{\theta}}] &= -S L_{\hat{\theta}}, \qquad & [L_r, L_{\hat{\phi}}] &= -T L_{\hat{\phi}}, \\ [L_t, L_{\hat{\phi}}] &= -S L_{\hat{\phi}}, \qquad & [L_{\hat{\theta}}, L_{\hat{\phi}}] &= 0. \end{aligned} \tag{14.19}$$

A set of bracket relations such as these is the first step in writing the field equations in an entirely intrinsic form. The use of orthonormal vectors in expressing these relations brings out the structure most clearly.

Next we seek an intrinsic form of the Riemann tensor. This calculation is simplified by making use of the results in table 14.1. The bracket relations enable us to calculate the derivatives of α/r by writing

$$\begin{aligned} L_t(\alpha/r) &= L_t L_{\hat{\theta}} \theta = [L_t, L_{\hat{\theta}}]\theta = -S\alpha/r, \\ L_r(\alpha/r) &= L_r L_{\hat{\theta}} \theta = [L_r, L_{\hat{\theta}}]\theta = -T\alpha/r. \end{aligned} \tag{14.20}$$

Application of equation (14.11) is now straightforward, and leads to the Riemann tensor

$$\begin{aligned}
\mathcal{R}(\sigma_r) &= (L_r G - L_t F + G^2 - F^2)\sigma_r, \\
\mathcal{R}(\sigma_\theta) &= (-L_t S + GT - S^2)\sigma_\theta + (L_t T + ST - SG) I\sigma_\phi, \\
\mathcal{R}(\sigma_\phi) &= (-L_t S + GT - S^2)\sigma_\phi - (L_t T + ST - SG) I\sigma_\theta, \\
\mathcal{R}(I\sigma_\phi) &= (L_r T + T^2 - FS) I\sigma_\phi - (L_r S + ST - FT)\sigma_\theta, \\
\mathcal{R}(I\sigma_\theta) &= (L_r T + T^2 - FS) I\sigma_\theta + (L_r S + ST - FT)\sigma_\phi, \\
\mathcal{R}(I\sigma_r) &= (-S^2 + T^2 - (\alpha/r)^2) I\sigma_r.
\end{aligned} \tag{14.21}$$

We must next decide on the form of matter energy-momentum tensor that the gravitational fields couple to. We assume that the matter is modelled by an ideal fluid, as discussed in section 13.5.4, so we can write

$$\mathcal{T}(a) = (\rho + p) a \cdot v\, v - pa, \tag{14.22}$$

where ρ is the energy density, p is the pressure and v is the covariant fluid velocity ($v^2 = 1$). Radial symmetry means that v can only lie in the e_t and e_r directions, so v must take the form

$$v = \cosh(\chi)\, e_t + \sinh(\chi)\, e_r. \tag{14.23}$$

But, in restricting the $\bar{\mathsf{h}}$ function to the form of equation (14.16), we retained the gauge freedom to perform arbitrary radial boosts. This freedom can now be employed to set $v = e_t$, so that the matter energy-momentum tensor becomes

$$\mathcal{T}(a) = (\rho + p) a \cdot e_t\, e_t - pa. \tag{14.24}$$

There is no physical content in the choice $v = e_t$ as all physical relations must be independent of gauge choices. The choice simply fixes the rotation gauge in such a way that the energy-momentum tensor takes on the simplest form. This removes all rotation-gauge freedom — an essential step in the solution method, since all non-physical degrees of freedom must be removed before one can achieve a complete set of physical equations.

In section 13.6.1 we saw how to decompose the Riemann tensor into a source term and the Weyl tensor. The source term can be written

$$\mathcal{R}(a \wedge b) - \mathcal{W}(a \wedge b) = \frac{4\pi}{3}\big(3a \wedge \mathcal{T}(b) + 3\mathcal{T}(a) \wedge b - 2\mathcal{T}\, a \wedge b\big), \tag{14.25}$$

where $\mathcal{T} = \partial_a \cdot \mathcal{T}(a)$ is the trace of the matter energy-momentum tensor. With $\mathcal{T}(a)$ given by equation (14.24), $\mathcal{R}(B)$ is restricted to the form

$$\mathcal{R}(B) = \mathcal{W}(B) + \frac{4\pi}{3}\big(3(\rho + p) B \cdot e_t\, e_t - 2\rho B\big). \tag{14.26}$$

Comparing this with equation (14.21) we see that the Weyl tensor must have the general form

$$\begin{aligned} \mathcal{W}(\sigma_r) &= \alpha_1\sigma_r, & \mathcal{W}(I\sigma_r) &= \alpha_4 I\sigma_r, \\ \mathcal{W}(\sigma_\theta) &= \alpha_2\sigma_\theta + \beta_1 I\sigma_\phi, & \mathcal{W}(I\sigma_\theta) &= \alpha_3 I\sigma_\theta + \beta_2\sigma_\phi, \\ \mathcal{W}(\sigma_\phi) &= \alpha_2\sigma_\phi - \beta_1 I\sigma_\theta, & \mathcal{W}(I\sigma_\phi) &= \alpha_3 I\sigma_\phi - \beta_2\sigma_\theta. \end{aligned} \tag{14.27}$$

Here each of the α_i represents a combination of intrinsic objects.

The torsionless gravitational field equations ensure that the Weyl tensor is self-dual and symmetric. The former implies that $\alpha_1 = \alpha_4$, $\alpha_2 = \alpha_3$ and $\beta_1 = -\beta_2$, and the latter implies that $\beta_1 = \beta_2$. It follows that $\beta_1 = \beta_2 = 0$. Finally, $\mathcal{W}(B)$ must be traceless, which requires that $\alpha_1 + 2\alpha_2 = 0$. Taken together, these conditions reduce $\mathcal{W}(B)$ to the form

$$\mathcal{W}(B) = \frac{\alpha_1}{4}(B + 3\sigma_r B\sigma_r). \tag{14.28}$$

This is of Petrov type D. From the form of $\mathcal{R}(I\sigma_r)$ we can see that

$$\alpha_1 = \frac{8\pi\rho}{3} - S^2 + T^2 - \frac{\alpha^2}{r^2}. \tag{14.29}$$

If we now define β by

$$4\beta = -S^2 + T^2 - \frac{\alpha^2}{r^2}, \tag{14.30}$$

then the full Riemann tensor can be written as

$$\mathcal{R}(B) = \left(\beta + \frac{2\pi}{3}\rho\right)\big(B + 3\sigma_r B\sigma_r\big) + \frac{4\pi}{3}\big(3(\rho+p)B\cdot e_t e_t - 2\rho B\big). \tag{14.31}$$

We compare this with equation (14.21) to obtain the following set of equations:

$$\begin{aligned} L_t S &= 2\beta + GT - S^2 - 4\pi p, \\ L_t T &= S(G - T), \\ L_r S &= T(F - S), \\ L_r T &= -2\beta + FS - T^2 - 4\pi\rho, \\ L_r G - L_t F &= F^2 - G^2 + 4\beta + 4\pi(\rho + p). \end{aligned} \tag{14.32}$$

We are now close to our goal of a complete set of intrinsic equations. The remaining step is to enforce the Bianchi identities. The only identity that contains new information in our present setup is the contracted Bianchi identity defined in section 13.6.2, which guarantees covariant conservation of the energy-momentum tensor. For an ideal fluid this results in the pair of equations

$$\begin{aligned} \mathcal{D}\cdot(\rho v) + p\mathcal{D}\cdot v &= 0, \\ (\rho + p)(v\cdot\mathcal{D}v)\wedge v - (\mathcal{D}p)\wedge v &= 0. \end{aligned} \tag{14.33}$$

The quantity $(v{\cdot}\mathcal{D}v){\wedge}v$ is the covariant acceleration bivector, so the second equation relates the acceleration to the pressure gradient. For the case of spherically-symmetric fields, these equations reduce to

$$\begin{aligned} L_t\rho &= -(F+2S)(\rho+p), \\ L_r p &= -G(\rho+p). \end{aligned} \tag{14.34}$$

The latter of these identifies G as the radial acceleration. The full Bianchi identities now turn out to be satisfied as a consequence of the contracted identities and the bracket relation

$$[L_t, L_r] = GL_t - FL_r. \tag{14.35}$$

This completes our derivation of the intrinsic equations. The full set is defined by equations (14.20), (14.32), the contracted identities (14.34) and the bracket structure of equation (14.35). The equation structure is closed, as the bracket relation (14.35) is consistent with the known derivatives. The derivation of such a set of equations is the basic aim of our method. The equations deal solely with objects that transform covariantly under displacements, and many of these quantities have direct physical significance.

14.2.2 Solving the spherical equations

To solve the intrinsic equation structure we first form the derivatives of β to obtain

$$\begin{aligned} L_t\beta + 3S\beta &= 2\pi S p, \\ L_r\beta + 3T\beta &= -2\pi T\rho. \end{aligned} \tag{14.36}$$

These results suggest that we should look for an integrating factor for the L_t+S and L_r+T operators. Such a function, X say, should have the properties that

$$L_t X = SX, \qquad L_r X = TX. \tag{14.37}$$

A function with these properties can exist only if the derivatives are consistent with the bracket relation of equation (14.35). This is checked by forming

$$\begin{aligned} [L_t, L_r]X &= L_t(TX) - L_r(SX) \\ &= X(L_t T - L_r S) \\ &= X(SG - FT) \\ &= GL_t X - FL_r X, \end{aligned} \tag{14.38}$$

which confirms that the properties of X are consistent with the bracket structure. In fact, we can see from equation (14.20) that r/α has the desired properties. Integrating factors of this type often arise as natural, intrinsically-defined coordinates, and the form of the solution is usually simplest when expressed directly in terms of these. Since the position-gauge freedom in the r direction has not yet been fixed, it is natural to set $\alpha = 1$, so that r plays the role of the integrating factor directly. We will confirm shortly that this gauge choice ensures that r is a physically meaningful quantity.

With the radial scale fixed by setting $\alpha = 1$, we can now make some further simplifications. From the form of the $\bar{\mathsf{h}}$ function in equation (14.16), together with equation (14.37), we see that

$$\begin{aligned} g_1 &= L_r r = Tr, \\ g_2 &= L_t r = Sr. \end{aligned} \tag{14.39}$$

This replaces two functions in the bivector connection in favour of terms in $\bar{\mathsf{h}}(a)$. We also define

$$M = -2r^3\beta = \frac{r}{2}\big({g_2}^2 - {g_1}^2 + 1\big), \tag{14.40}$$

which satisfies

$$\begin{aligned} L_t M &= -4\pi r^2 g_2 p, \\ L_r M &= 4\pi r^2 g_1 \rho. \end{aligned} \tag{14.41}$$

The latter suggests that M plays the role of an intrinsic mass.

So far we have defined the natural distance scale, but have not yet found a natural time coordinate. Such a coordinate is required to complete the solution, so we now look for additional criteria to motivate this choice. We are currently free to perform an arbitrary r and t-dependent displacement along the e_t direction. This gives us complete freedom in the choice of f_2 function. If we now invert equation (14.41) to find the coordinate derivatives of M we obtain

$$\begin{aligned} \frac{\partial M}{\partial t} &= \frac{-4\pi g_1 g_2 r^2(\rho + p)}{f_1 g_1 - f_2 g_2}, \\ \frac{\partial M}{\partial r} &= \frac{4\pi r^2(f_1 g_1 \rho + f_2 g_2 p)}{f_1 g_1 - f_2 g_2}. \end{aligned} \tag{14.42}$$

The second equation reduces to a simple classical relation if we choose $f_2 = 0$, as we then obtain

$$\partial_r M = 4\pi r^2 \rho. \tag{14.43}$$

This says that, at constant time t, $M(r,t)$ is determined by the amount of mass-energy in a sphere of radius r.

With f_2 set to zero we can now use the bracket structure to solve for f_1. We have

$$L_t = f_1\partial_t + g_2\partial_r, \qquad L_r = g_1\partial_r, \tag{14.44}$$

so the bracket relation of equation (14.35) implies that

$$L_r f_1 = -G f_1. \tag{14.45}$$

It follows that

$$f_1 = \epsilon(t) \exp\left(-\int^r \frac{G(s)}{g_1(s)}\, ds\right). \tag{14.46}$$

The function $\epsilon(t)$ can be absorbed by a further t-dependent rescaling along e_t, which will not reintroduce a term in f_2. In the $f_2 = 0$ gauge we can therefore reduce to a system in which

$$f_1 = \exp\left(-\int^r \frac{G(s)}{g_1(s)}\, ds\right). \tag{14.47}$$

The physical explanation for why the $f_2 = 0$ gauge is a very natural one to work in emerges when we set the pressure to zero. In this case equation (14.34) forces G to be zero, and equation (14.47) then sets $f_1 = 1$. A (free-falling) particle comoving with the fluid has covariant velocity $v = e_t$, so the trajectory of this particle is defined by

$$\dot{t} e_t + \dot{r} e_r = \mathsf{h}(e_t) = e_t + g_2\, e_r, \tag{14.48}$$

where the dots denote differentiation with respect to the proper time. Since $\dot{t} = 1$ the time coordinate t matches the proper time of all observers comoving with the fluid. In this sense, the time coordinate that has emerged behaves like a global Newtonian time on which all observers can agree (provided all clocks are correlated initially). By employing the various gauge choices outlined above, and casting the dynamics in terms of the t coordinate, we are ensuring that (when $p = 0$) the physics is formulated from the viewpoint of freely-falling observers. We then expect that the gravitational equations should take on a clear, physical form, which is indeed the case.

As a further illustration of this point, it is clear from (14.48) that g_2 represents a radial velocity for the particle. In the absence of pressure, the rate of change of mass is given by

$$\partial_t M = -4\pi r^2 g_2 \rho. \tag{14.49}$$

This equation equates the work with the rate of flow of energy density. Similarly, equation (14.40), written in the form

$$\frac{(g_2)^2}{2} - \frac{M}{r} = \frac{1}{2}\Big((g_1)^2 - 1\Big), \tag{14.50}$$

is also now familiar from Newtonian physics — it is a Bernoulli equation for zero pressure and total (non-relativistic) energy $({g_1}^2 - 1)/2$. When pressure is included, the purely Newtonian interpretation starts to break down, due mainly to the fact that pressure can act as a source of gravitation. But it remains the case that the gauge choices described here pick out what appears to be the most natural set of equations for studying spherically-symmetric systems.

The system of equations we have now derived is summarised in table 14.2. We

The $\bar{\mathsf{h}}$ field	$\bar{\mathsf{h}}(e^t) = f_1 e^t$ $\bar{\mathsf{h}}(e^r) = g_1 e^r + g_2 e^t$ $\bar{\mathsf{h}}(e^\theta) = e^\theta$ $\bar{\mathsf{h}}(e^\phi) = e^\phi$
The ω field	$\omega(e_t) = G e_r e_t$ $\omega(e_r) = F e_r e_t$ $\omega(\hat{\theta}) = g_2/r\, \hat{\theta} e_t + (g_1 - 1)/r\, e_r \hat{\theta}$ $\omega(\hat{\phi}) = g_2/r\, \hat{\phi} e_t + (g_1 - 1)/r\, e_r \hat{\phi}$
Directional derivatives	$L_t = f_1 \partial_t + g_2 \partial_r$ $L_r = g_1 \partial_r$
Equations for G and F	$L_t g_1 = G g_2$ $L_r g_2 = F g_1$ $f_1 = \exp\{\int^r -G/g_1\, ds\}$
Definition of M	$M = \frac{1}{2} r({g_2}^2 - {g_1}^2 + 1)$
Remaining derivatives	$L_t g_2 = G g_1 - M/r^2 - 4\pi r p$ $L_r g_1 = F g_2 + M/r^2 - 4\pi r \rho$
Matter derivatives	$L_t M = -4\pi r^2 g_2 p$ $L_r M = 4\pi r^2 g_1 \rho$ $L_t \rho = -(2g_2/r + F)(\rho + p)$ $L_r p = -G(\rho + p)$
Riemann tensor	$\mathcal{R}(B) = 4\pi\big((\rho + p) B \cdot e_t e_t - 2\rho/3\, B\big)$ $\quad - \frac{1}{2}(M/r^3 - 4\pi\rho/3)(B + 3\sigma_r B \sigma_r)$
Energy-momentum tensor	$\mathcal{T}(a) = (\rho + p) a \cdot e_t e_t - pa$

Table 14.2 *Gravitational equations governing a radially-symmetric perfect fluid*. An equation of state and initial data $\rho(r, t_0)$ and $g_2(r, t_0)$ determine the future evolution of the system.

refer to this system as defining the *Newtonian* gauge, since so many equations take on an almost Newtonian form. Of course, this should not distract from the fact that we have solved the full, relativistic gravitational field equations. The system of equations in table 14.2 underlies a wide range of phenomena in relativistic astrophysics and cosmology. One aspect of these equations is immediately apparent. Given an equation of state $p = p(\rho)$, and initial data in the form of the density $\rho(r, t_0)$ and the velocity $g_2(r, t_0)$, the future evolution of the system is fully determined. This is because ρ determines p and M on a time slice, and the definition of M determines g_1. The equations for $L_r p$, $L_r g_1$

and $L_r g_2$ then determine the remaining information on the time slice. Finally, the $L_t M$ and $L_t g_2$ equations can be used to update the information to the next time slice, and the process can start again. The equations can therefore be implemented numerically as a simple set of first-order update equations. This is important for a wide range of applications.

14.2.3 Static matter distributions

As a first application of the equations governing a spherically-symmetric system, we consider a static matter distribution. This solution is appropriate for a non-rotating spherical source. The density and pressure are now functions of r only. The mass is given by

$$M(r) = \int_0^r 4\pi s^2 \rho(s)\, ds \tag{14.51}$$

and it follows that

$$L_t M = 4\pi r^2 g_2 \rho = -4\pi r^2 g_2 p. \tag{14.52}$$

For any physical matter distribution ρ and p must both be positive, in which case equation (14.52) can only be satisfied if g_2 vanishes. It follows that $F = 0$ as well, so for static, extended objects we have

$$g_2 = F = 0. \tag{14.53}$$

Since g_2 is zero, g_1 is given simply in terms of $M(r)$ by

$${g_1}^2 = 1 - \frac{2M(r)}{r}. \tag{14.54}$$

For this to hold we require that $2M(r) < r$. This condition says that a horizon has not formed anywhere in the object.

The remaining equation of use is that for $L_t g_2$, which now gives

$$G g_1 = \frac{M(r)}{r^2} + 4\pi r p. \tag{14.55}$$

Equations (14.54) and (14.55) combine with that for $L_r p$ to produce the *Oppenheimer–Volkov* equation

$$\frac{\partial p}{\partial r} = -\frac{(\rho + p)(M(r) + 4\pi r^3 p)}{r(r - 2M(r))}. \tag{14.56}$$

This is the force balance equation appropriate for a relativistic matter distribution. The line element generated by our solution is

$$ds^2 = \frac{1}{(f_1)^2} dt^2 - \frac{r}{r - 2M(r)} dr^2 - r^2\, d\theta^2 - r^2 \sin^2(\theta)\, d\phi^2, \tag{14.57}$$

where f_1 is given by equation (14.47). The solution extends straightforwardly to the region outside the star. In this region M is constant, and

$$f_1 = 1/g_1 = (1 - 2M/r)^{-1/2}. \tag{14.58}$$

We therefore recover the *Schwarzschild* line element. This is the solution used for some of the most famous tests of general relativity, including those for the bending of light and the perihelion precession of Mercury. Clearly, the gauge theory framework does not alter any of these results.

14.3 Schwarzschild black holes

Perhaps the most famous solution of the Einstein equations (apart from Lorentzian spacetime) is the Schwarzschild solution for a black hole. This solution describes the gravitational fields surrounding a point source of matter, of total gravitational mass M. One form of this solution is described by the line element of equation (14.57) for the case of constant M. But this is ill defined at $r = 2M$ which, as we shall soon discover, defines an *event horizon*. This tells us that our gauge choice has not yielded a satisfactory global solution, so we must return to the field equations to discover what went wrong.

For a point source located at the origin we have $\rho = p = 0$ everywhere away from the source. The matter equations therefore reduce to

$$L_t M = L_r M = 0, \tag{14.59}$$

which tells us that the mass M is constant. The remaining equations simplify to

$$\begin{aligned} L_t g_1 &= G g_2, \\ L_r g_2 &= F g_1, \\ {g_1}^2 - {g_2}^2 &= 1 - 2M/r. \end{aligned} \tag{14.60}$$

No further equations yield new information, so we have an underdetermined system of equations. Despite all of the gauge-fixing steps taken to arrive at the set of equations summarised in table 14.2, for vacuum fields some additional gauge fixing is still required. The reason for this is that, in the vacuum region, the Riemann tensor reduces to

$$\mathcal{R}(B) = -\frac{M}{2r^3}(B + 3\sigma_r B \sigma_r). \tag{14.61}$$

This tensor is now invariant under boosts in the σ_r plane, whereas previously the presence of the fluid velocity in the Riemann tensor vector broke this symmetry. The appearance of this new symmetry in the matter-free case manifests itself as a new freedom in the choice of the $\bar{\mathsf{h}}$ function.

Given this new freedom, we can look for a choice of g_1 and g_2 which simplifies the equations. If we attempt to reproduce the Schwarzschild line element we have to set $g_2 = 0$, but then we immediately run into difficulties with g_1, which is not defined for $r < 2M$. We must therefore look for an alternative gauge choice. A suitable candidate, motivated by the pressure-free equations, is provided by the simple choice

$$g_1 = 1. \tag{14.62}$$

It follows that

$$f_1 = 1, \qquad g_2 = -\sqrt{2M/r} \tag{14.63}$$

and

$$G = 0, \qquad F = -\frac{M}{g_2 r^2} = \left(\frac{M}{2r^3}\right)^{1/2}. \tag{14.64}$$

In this gauge the $\bar{\mathsf{h}}$ function has the remarkably simple form

$$\bar{\mathsf{h}}(a) = a - \sqrt{2M/r}\, a\cdot e_r\, e_t. \tag{14.65}$$

This only differs from the identity through a single term. The line element obtained from this gauge choice is

$$ds^2 = dt^2 - \left(dr + \left(\frac{2M}{r}\right)^{1/2} dt\right)^2 - r^2(d\theta^2 + \sin^2(\theta)\, d\phi^2), \tag{14.66}$$

which is regular at the horizon ($r = 2M$) and covers all spacetime down to $r = 0$. This form of the line element was first derived by Painlevé and Gullstrand, not long after Schwarzschild's original work was published. Despite the many advantages of this form of the solution, it has not been routinely employed in solving physical problems.

The $\bar{\mathsf{h}}$ field of equation (14.65) is the form of the Schwarzschild solution we will use for studying the properties of spherically-symmetric black holes. Of course, all physical predictions must be independent of gauge, but this only reinforces the point that we should always endeavour to work in a gauge that simplifies the analysis as far as possible. The results for the extension to the action of $\bar{\mathsf{h}}$ to an arbitrary multivector A are useful in what follows. We find that

$$\bar{\mathsf{h}}(A) = A - \sqrt{2M/r}(A\cdot e_r)\wedge e_t. \tag{14.67}$$

It follows that det $(\mathsf{h}) = 1$ and the inverse of the adjoint function, as defined by equation (4.152), is given by

$$\mathsf{h}^{-1}(A) = A + \sqrt{2M/r}(A\cdot e_t)\wedge e_r. \tag{14.68}$$

It is straightforward to verify that this function recovers the line element of equation (14.66).

14.3.1 Point particle trajectories

The motion of a classical point particle in free fall is governed by the geodesic equation

$$v\cdot\mathcal{D}v = \dot{v} + \omega(v)\cdot v = 0. \tag{14.69}$$

The mass m of the particle is unimportant (provided $m \ll M$), and is set to unity throughout this section. Since $G = 0$ in our chosen gauge, we immediately see that

$$\omega(e_t) = 0. \tag{14.70}$$

It follows that $v = e_t$ is a solution of the geodesic equation. The trajectory this defines has

$$\dot{x} = \mathsf{h}(v) = \mathsf{h}(e_t) = e_t + ue_r, \tag{14.71}$$

where

$$u = \dot{r} = -\surd(2M/r). \tag{14.72}$$

Particles, or observers, following the geodesic defined by $v = e_t$ fall in radially with velocity $\dot{r}$ given by the familiar Newtonian formula. Furthermore, we see that $\dot{t} = 1$, so the time coordinate t is precisely the time measured by these infalling observers. This is, in part, why the gauge choice we have adopted turns out to simplify many calculations.

Now consider a more general trajectory, with covariant velocity

$$v = \dot{t}\, e_t + (\dot{t}\sqrt{2M/r} + \dot{r})e_r + \dot{\theta}e_\theta + \dot{\phi}e_\phi. \tag{14.73}$$

Since the $\bar{\mathsf{h}}$ function is independent of t we have, from equation (13.272),

$$\mathsf{h}^{-1}(e_t)\cdot v = (1 - 2M/r)\dot{t} - \dot{r}\sqrt{2M/r} = \text{constant}. \tag{14.74}$$

So, for particles moving forwards in time ($\dot{t} > 0$ for $r \to \infty$), we can write

$$(1 - 2M/r)\dot{t} = \alpha + \dot{r}\sqrt{2M/r}, \tag{14.75}$$

where the constant α satisfies $\alpha > 0$. The radial equation is found from the constraint that $v^2 = 1$, which gives

$$\dot{r}^2 = \alpha^2 - \big(1 - 2M/r\big)\Big(1 + r^2\big(\dot{\theta}^2 + \sin^2(\theta)\,\dot{\phi}^2\big)\Big). \tag{14.76}$$

Spherical symmetry implies that the angular velocity J is also conserved, where

$$J^2 = r^4(\dot{\theta}^2 + \sin^2(\theta)\,\dot{\phi}^2). \tag{14.77}$$

The motion of a particle around a black hole is therefore determined by the single radial equation

$$\dot{r}^2 = \alpha^2 - \big(1 - 2M/r\big)\left(1 + \frac{J^2}{r^2}\right). \tag{14.78}$$

This equation is gauge-*invariant*, as it relates local quantities. The radial coordinate r is defined locally by the magnitude of the Riemann tensor, and the dots denote the derivative with respect to (local) proper time. This transition from global to local variables is in keeping with the gauging process. The motion of a particle in spacetime is obtained by integrating equations (14.78) and (14.75). At the horizon we have $\dot{r} = -\alpha$, so there is no pole in equation (14.75), and the equations can be integrated down to the singularity.

Differentiating equation (14.78) we obtain

$$\ddot{r} = -\frac{M}{r^2} + \frac{J^2}{r^3} - \frac{3MJ^2}{r^4}. \tag{14.79}$$

The equivalent three-dimensional vector equation is

$$\ddot{\boldsymbol{x}} = -\left(\frac{M}{r^2} + \frac{3MJ^2}{r^4}\right)\hat{\boldsymbol{x}}. \tag{14.80}$$

This equation was analysed perturbatively in section 3.3.1. For stable orbits the main new effect introduced by relativity is a small perturbation of the eccentricity vector. The content of equation (14.78) can similarly be summarised in the radial effective potential (per unit mass)

$$V_{\mathit{eff}} = -\frac{M}{r} + \frac{J^2}{2r^2}\left(1 - \frac{2M}{r}\right). \tag{14.81}$$

We then have

$$\frac{\alpha^2 - 1}{2} = \frac{\dot{r}^2}{2} + V_{\mathit{eff}} \tag{14.82}$$

which identifies $m\alpha$ as the conserved relativistic energy of the particle. Bound states have $\alpha < 1$ and scattering states have $\alpha > 1$.

The effective potential differs from the Newtonian expression in the factor of $(1 - 2M/r)$ multiplying the centrifugal term. This has little effect at large distances, but dramatically alters the small-r behaviour. Inside $r = 2M$ the centrifugal term in the effective potential changes sign and becomes *attractive*. There is no longer any term in the potential applying an effective outward force, and the particle must inexorably move towards the central singularity. One can see this clearly in equation (14.73). Inside the horizon the velocity $\dot{r}$ must be negative in order for $v^2 = 1$ to remain satisfied. Once inside the horizon, no particle can escape the singularity, no matter what force is applied to attempt to counteract the gravitational pull. Eventually, the tidal forces (defined by the Riemann tensor) become so large that all objects are pulled apart into their constituent particles.

14.3.2 Photon trajectories

A full treatment of the properties of electromagnetic waves in a gravitational background involves solving the gravitationally-coupled Maxwell equations of section 13.5.8. For a range of practical problems it is sufficient to ignore the detailed properties of the electromagnetic field, and work in the geometric optics limit. In this approach, photons are treated as massless (scalar) point particles. These particles follow *null* trajectories with

$$k = \mathsf{h}^{-1}(\dot{x}), \qquad k^2 = 0. \tag{14.83}$$

The trajectories are still specified by the equation $k\cdot\mathcal{D}k = 0$. For radial *infall* we must have

$$k = \nu(e_t - e_r), \tag{14.84}$$

where $\nu = k \cdot e_t$ is the frequency measured by radially free-falling observers (at rest at infinity). The photon trajectory is independent of the frequency, as demanded by the equivalence principle. The path defined by k is given by

$$\dot{x} = \mathsf{h}(k) = \nu\big(e_t - (1 + \sqrt{(2M/r)})e_r\big). \tag{14.85}$$

It follows that

$$\frac{dr}{dt} = -(1 + \sqrt{(2M/r)}). \tag{14.86}$$

This integrates straightforwardly to give the photon path. We have therefore found the path without employing the equation of motion. This is possible because we restricted to motion in a single spacetime plane.

The equations of motion tell us how the frequency changes along the path. To find this we need

$$\omega(k) = -\nu\left(\frac{M}{2r^3}\right)^{1/2}\sigma_r, \tag{14.87}$$

from which we see that

$$\dot{\nu} = \nu^2\left(\frac{M}{2r^3}\right)^{1/2}. \tag{14.88}$$

This equation is more usefully expressed in terms of the derivative with respect to r. We use

$$\dot{r} = -\nu\big(1 + \sqrt{(2M/r)}\big) \tag{14.89}$$

to arrive at

$$\frac{1}{\nu}\frac{d\nu}{dr} = \frac{M}{r}\frac{1}{2M + \sqrt{(2Mr)}} = \frac{1}{2r}\frac{1}{\sqrt{r/r_S} + 1}, \tag{14.90}$$

where $r_S = 2M$ is the Schwarzschild radius. This equation can again be integrated straightforwardly to tell us how frequency ν changes with radius. We see that nothing untoward happens until $r = 0$ is reached.

We can repeat the previous analysis for outgoing photons. For this case we have

$$k = \nu(e_t + e_r) \tag{14.91}$$

and the path is

$$\dot{x} = \mathsf{h}(v) = \nu\big(e_t + (1 - \sqrt{(2M/r)})e_r\big). \tag{14.92}$$

It follows that

$$\frac{dr}{dt} = 1 - \sqrt{(2M/r)}. \tag{14.93}$$

But now, when $r < 2M$ the path is still *inwards*. Inside $r = 2M$, not even light can escape. The surface $r = 2M$ is called the *event horizon*. It marks the boundary between two regions, one of which (the interior in this case) cannot signal to the other. We also find that

$$\frac{1}{\nu}\frac{d\nu}{dr} = \frac{M}{r}\frac{1}{2M - \sqrt{(2Mr)}} = -\frac{1}{2r}\frac{1}{\sqrt{r/r_S} - 1}, \tag{14.94}$$

which is negative outside the horizon. So, as photons climb out of a gravitational field, they are *redshifted*. This is one of the best-tested predictions of general relativity. The redshift becomes increasingly large as the horizon is approached, so photons emitted from near the horizon are strongly redshifted as they climb out to infinity. The various features of radial motion in a black hole background are shown in figure 14.1. One conclusion from this plot is that, as seen by *external observers*, any object falling through the horizon appears to hover outside the horizon and just fade out of existence as the redshift increases.

If any object collapses to within its event horizon, it must carry on collapsing to form a central *singularity*. There is no possible force capable of preventing the collapse. This is because matter is always constrained to follow timelike paths, and if the entire future light-cone points inwards towards the singularity, no matter can escape. The object remaining at the end of this process is called a *black hole*. All paths for infalling matter terminate on the singularity. There has been much research into the properties of singularities, though their nature remains enigmatic. In one sense, gravitational singularities are no more difficult to deal with than singularities in the electromagnetic field due to point sources. They can also be analysed in much the same way using integral equations. But this (classical) treatment of singularities can only contain part of the story. Quantum mechanically, black holes have an associated entropy, implying the existence of a series of microstates consistent with the macroscopic properties of the hole. It is widely believed that a more complete understanding of quantum gravity should explain this phenomenon through a detailed quantum description of the singularity.

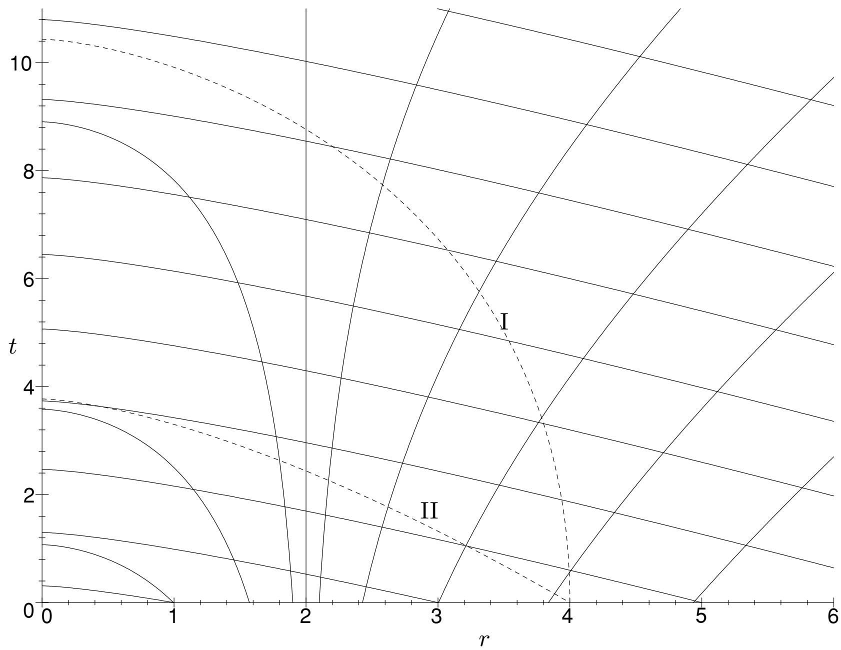

Figure 14.1 *Matter and photon trajectories in a black hole background.* The solid lines are photon trajectories, and the horizon lies at $r = 2$. Outside the horizon it is possible to send photons out to infinity, and hence communicate with the rest of the universe. As the emitter approaches the horizon, these photons are strongly redshifted and take a long time to escape. Once inside the horizon, all photon paths end on the singularity. The broken lines represent two possible trajectories for infalling matter. Trajectory I is for a particle released from rest at $r = 4$. Trajectory II is for a particle released from rest at $r = \infty$.

14.3.3 Stationary observers

It is instructive to see how physics appears from the point of view of stationary observers in a Schwarzschild background. These observers have constant r, θ, ϕ, so

$$\dot{x} = \dot{t} e_t. \tag{14.95}$$

It follows that

$$v = \dot{t}(e_t + \surd(2M/r) e_r). \tag{14.96}$$

But we require that $v^2 = 1$ for the path to be parameterised by the observer's proper time, so

$$\dot{t}^2(1 - 2M/r) = 1, \qquad \dot{t} = (1 - 2M/r)^{-1/2}. \tag{14.97}$$

This is a constant, since r is fixed for these observers. We can see immediately that it is only possible to remain at rest *outside* the horizon. This is reasonable given the preceding considerations, though the picture is not quite so clear if the black hole is rotating. For this case there is a region outside the horizon within which it is impossible to remain at rest (though it is still possible to escape).

The covariant acceleration bivector for a particle with velocity v is defined by

$$(v\cdot\mathcal{D}v)\wedge v = \dot{v}v + \omega(v)\cdot v\, v. \tag{14.98}$$

This gives the acceleration required to follow a given path. For stationary observers we have

$$(v\cdot\mathcal{D}v)\wedge v = \frac{M}{r^2(1-2M/r)^{1/2}}\sigma_r. \tag{14.99}$$

So an observer with mass m needs to apply force of $Mm/r^2 \times (1-2M/r)^{-1/2}$ to remain at rest. This is the Newtonian value multiplied by a relativistic correction term. This correction becomes increasingly large as the horizon is approached, as one would expect.

We can now look at physics from point of view of these observers, which can be viewed as both being stationary and having constant acceleration. For example, if a second observer has velocity γ_0 (so is in free fall), the relative velocity the two observers measure when their positions coincide is

$$\frac{v\wedge\gamma_0}{v\cdot\gamma_0} = \sqrt{(2M/r)}\sigma_r. \tag{14.100}$$

As we might expect, this is the Newtonian result. The only difference now lies in the interpretation of who is accelerating. The stationary observer is the one applying a force, so we now say that it is this observer that is accelerating. The observer in free fall is applying zero force, so is not accelerating. That is, we no longer view gravity as applying a force, as this would require a concept of what the particle would have done if the gravitational field were not present. Such a concept is not gauge-invariant, so is unphysical.

14.3.4 Absorption and scattering

The presence of the horizon implies that incident particles with total energy $E > mc^2$ can suffer two fates. Either they will be scattered by the gravitational fields, or they will be absorbed onto the central singularity. The crucial quantity that determines the fate of the particle is the angular velocity J. In figure 14.2 we plot the effective potential of equation (14.81) for a range of angular velocities. If J is too small there is nothing to prevent the particle hitting the singularity. As J increases, the effective potential develops a barrier. If this barrier is greater than the total (non-relativistic) energy, the particle is no longer absorbed, and instead is scattered by the black hole.

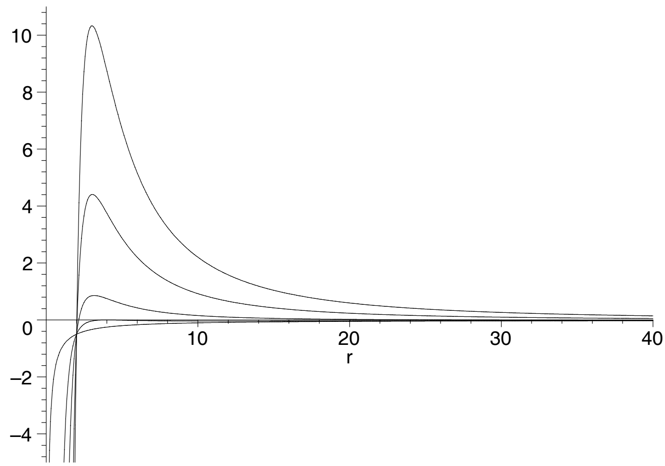

Figure 14.2 *The gravitational effective potential.* The potential for a unit mass particle is defined by equation (14.81), and units are chosen so that the horizon lies at $r = 2$. The plots are for J values of 0, 4, 8, 16 and 24. For small J nothing prevents the particle hitting the singularity. As J increases a barrier of increasing height is formed. If the particle has insufficient energy to surmount this barrier it is scattered.

For a given energy, we can determine the critical value of J that distinguishes between absorption and scattering. This is most usefully encoded in terms of an impact parameter b, as illustrated in figure 14.3. Asymptotically, the incoming particle has angular velocity

$$J = b\dot{r}(\infty). \tag{14.101}$$

But in this region the energy is determined entirely by $\dot{r}$, so the impact parameter is given by

$$b^2 = \frac{J^2}{\alpha^2 - 1}, \tag{14.102}$$

where α is the energy per unit mass of the incident particle, as defined in equation (14.75). For a fixed energy, the critical value of J therefore determines the critical value of the impact parameter. From the point of view of absorption, the black hole then appears as a disc of radius b, and the total absorption cross section is defined by

$$\sigma_{abs} = \pi b^2. \tag{14.103}$$

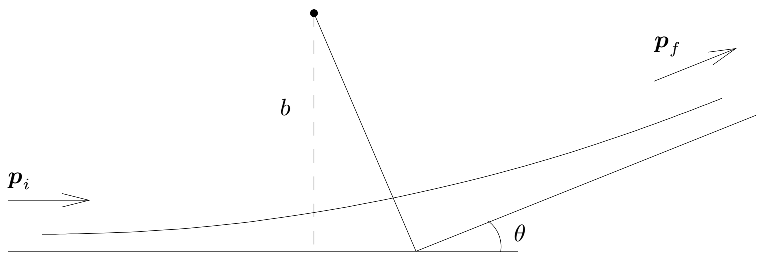

Figure 14.3 *The impact parameter.* In the asymptotic incoming region, the impact parameter b measures the distance between the incoming trajectory and a parallel radial trajectory. For a black hole there is a critical value of b inside which all geodesics terminate on the singularity. The diagram also defines the scattering angle θ.

This will be a decreasing function of energy — the faster the particle is travelling, the less likely it is to be absorbed.

The algebra needed to compute the absorption cross section is straightforward, if a little tedious. First we write $x = 1/r$, so that the effective potential becomes

$$V_{\mathit{eff}} = -Mx + \frac{b^2(\alpha^2 - 1)}{2} x^2(1 - 2Mx). \tag{14.104}$$

The turning point is at

$$x_c = \frac{1}{6M}\left(1 + \left(1 - \frac{12M^2}{(\alpha^2 - 1)b^2}\right)^{1/2}\right). \tag{14.105}$$

To find b the equation we need to solve is therefore

$$2V_{\mathit{eff}}(x_c) = \alpha^2 - 1. \tag{14.106}$$

The solution then returns the absorption cross section

$$\sigma_{abs} = \frac{\pi M^2}{2u^4}\big(8u^4 + 20u^2 - 1 + (1 + 8u^2)^{3/2}\big), \tag{14.107}$$

where we have expressed the result in terms of the velocity u:

$$u^2 = \frac{p^2}{E^2} = \frac{\alpha^2 - 1}{\alpha^2}. \tag{14.108}$$

The absorption cross section is plotted in figure 14.4. For small velocities we see that

$$\sigma_{abs} \mapsto \frac{16\pi M^2}{u^2}. \tag{14.109}$$

As the incident velocity decreases, the absorption cross section increases, as is to be expected. As the velocity increases the absorption cross section tends towards

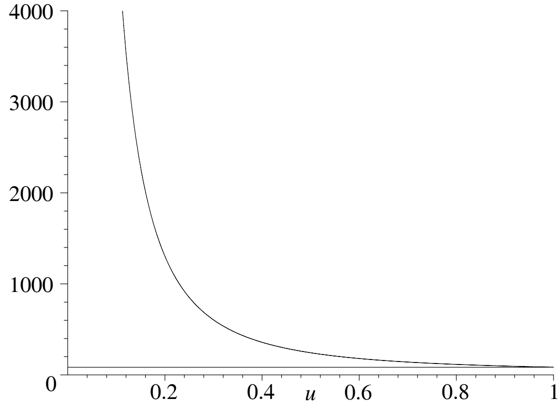

Figure 14.4 *The classical absorption cross section.* The cross section is a function of the incident velocity u (in units of c). As the velocity approaches the speed of light the cross section approaches the photon limit, as shown by the straight line. The vertical axis is in units of $(GM/c^2)^2$.

the limiting result for a massless particle. For these the effective potential is simply

$$V_{\mathit{eff}} = \frac{J^2}{2r^2}\left(1 - \frac{2M}{r}\right). \tag{14.110}$$

The turning point occurs at $r = 3M$, at which the effective potential has the value $J^2/54M^2$. Equating this with asymptotic energy $J^2/2b^2$ we see that for photons $b^2 = 27M^2$, and the photon absorption cross section is

$$\sigma_{abs} = \pi b^2 = 27\pi M^2. \tag{14.111}$$

This is the limiting value of equation (14.107) as $u \mapsto 1$. In section 14.4.3 we study how these features are modified by a more complete, quantum treatment of the absorption process.

Scattering presents a more difficult problem. The differential scattering cross section for a Newtonian $1/r$ potential is determined by the Rutherford formula

$$\frac{d\sigma}{d\Omega} = \frac{M^2}{4u^2 \sin^4(\theta/2)}, \tag{14.112}$$

where θ is the scattering angle and u is the velocity of the incident particle. This formula relates the incident cross sectional area σ to the solid angle $d\Omega$, where

$$d\sigma = 2\pi b\, db, \qquad d\Omega = 2\pi \sin(\theta)\, d\theta. \tag{14.113}$$

The Rutherford cross section formula is easily computed from the properties of hyperbolic trajectories. The relativistic corrections to the Rutherford formula are generated by the additional r^{-3} term in the potential. This term makes the problem considerably more difficult to solve, and no simple analytic formula exists for the classical scattering cross section. One problem is that it is now possible for particles to spiral around the centre before escaping. We could build up a perturbative picture of the scattering problem using the techniques described in section 3.3.1, though the resulting expressions are usually extremely complicated. A better approach to this problem is described in section 14.4.1, where the cross section is calculated using perturbative quantum theory.

14.3.5 Electromagnetism in a black hole background

Further insight into the nature and effects of a black hole is obtained by considering the electromagnetic fields surrounding charges held at rest outside the horizon. The relevant equations were obtained in section 13.5.8. We assume that the charge is placed at a distance $a > 2M$ along the z axis. The vector potential can be written in terms of a single scalar potential $V(r, \theta)$ as

$$A = V(r,\theta)\left(e_t + \frac{\sqrt{2Mr}}{r-2M}e_r\right), \tag{14.114}$$

so that

$$\mathcal{F} = -\frac{\partial V}{\partial r}e_r e_t - \frac{1}{r-2M}\frac{\partial V}{\partial \theta}\,\hat{\theta}(e_t + \sqrt{2M/r}e_r) \tag{14.115}$$

and

$$\boldsymbol{D} = -\frac{\partial V}{\partial r}e_r e_t - \frac{1}{r-2M}\frac{\partial V}{\partial \theta}\hat{\theta}e_t. \tag{14.116}$$

The Maxwell equations now reduce to the single partial differential equation

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial V}{\partial r}\right) + \frac{1}{r(r-2M)}\frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}\left(\sin(\theta)\frac{\partial V}{\partial \theta}\right) = -\rho, \tag{14.117}$$

where $\rho = q\delta(\boldsymbol{x} - \boldsymbol{a})$ is a δ-function at $z = a$. The solution (originally found by Linet) is

$$V(r,\theta) = \frac{q}{ar}\frac{(r-M)(a-M) - M^2\cos^2(\theta)}{d} + \frac{qM}{ar}, \tag{14.118}$$

where

$$d = \big(r(r-2M) + (a-M)^2 - 2(r-M)(a-M)\cos(\theta) + M^2\cos^2(\theta)\big)^{1/2}. \tag{14.119}$$

When this result is substituted back into equation (14.115) we see that the covariant field $\mathcal{F}$ is both finite and continuous at the horizon. It follows that we

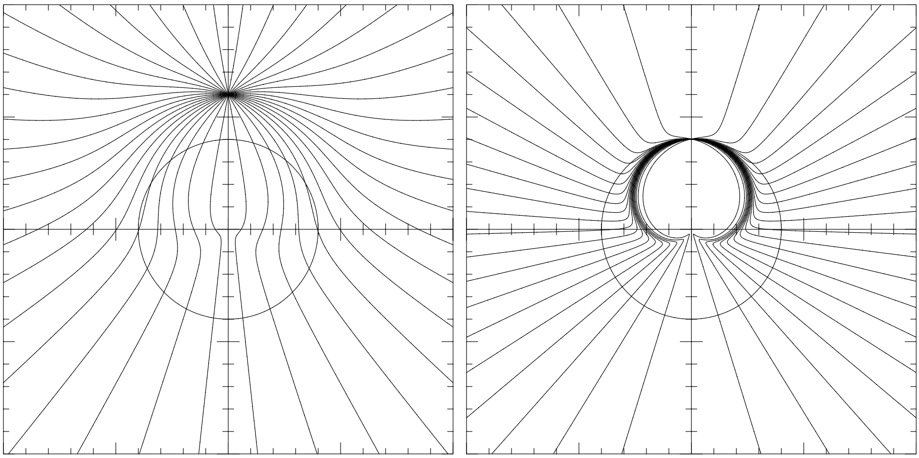

Figure 14.5 *Streamlines of the electric field in a black hole background.* The horizon lies at $r = 2$ and the charge is placed on the z axis. In the left-hand diagram the charge is held at $z = 3$, and in right-hand diagram it is at $z = 2.1$.

have found a *global* solution to the electromagnetic field equations, appropriate both inside and outside the horizon.

One way to illustrate the global properties of $\mathcal{F}$ is to plot the streamlines of $\boldsymbol{D}$. Equation (13.279) ensures that these streamlines begin and end on charges, so for our case of a single isolated charge they should therefore spread out from the charge and cover all space. Furthermore, since the distance scale r was chosen to agree with the gravitationally-defined distance, the streamlines of $\boldsymbol{D}$ convey genuine intrinsic information. The plots therefore encode gauge-invariant information about the electromagnetic field. Figure 14.5 shows streamline plots for charges held at different distances above the horizon. A polarisation charge is clearly visible at the origin, and streamlines are attracted towards this but never actually meet it. The effects of the polarisation charge can be felt outside the horizon as a repulsive force acting back on the charge. That is, less force is required to keep a charge at rest outside a black hole than is required for an uncharged particle. The fact that the origin of this effect lies inside the horizon reinforces the importance of constructing global solutions to the field equations.

14.3.6 Other gauges

Before proceeding it is useful to study the vacuum spherical equations in an arbitrary gauge. We return to the spherical equations before the $f_2 = 0$ gauge choice was made, and again impose that M is constant. The equations that remain are

$$g_1{}^2 - g_2{}^2 = 1 - 2M/r \tag{14.120}$$

and

$$\partial_r g_1 = G, \qquad \partial_r g_2 = F, \tag{14.121}$$

and all fields are functions of r only. The bracket relation of equation (14.35) gives

$$g_2 \partial_r f_2 - g_1 \partial_r f_1 = G f_1 - F f_2, \tag{14.122}$$

from which it follows that

$$\partial_r (f_1 g_1 - f_2 g_2) = \partial_r \det\,(\mathsf{h}) = 0. \tag{14.123}$$

The determinant of h is constant, and the value of this constant depends on the choice of gauge. Because the Riemann tensor falls off as r^{-3} we always choose to work in a gauge where $\bar{\mathsf{h}}$ tends to the identity as $r \mapsto \infty$. In this case we have $\det\,(\mathsf{h}) = 1$, so we can write

$$f_1 g_1 - f_2 g_2 = 1. \tag{14.124}$$

No other equations remain to fix the solution further. We therefore have two free functions in the choice of $\bar{\mathsf{h}}$ function.

A useful alternative to the Newtonian gauge chosen in this section is to write the solution in Kerr–Schild form. For this we set

$$\begin{aligned} g_1 &= 1 - M/r, \qquad & g_2 &= -M/r, \\ f_1 &= 1 + M/r, \qquad & f_2 &= M/r. \end{aligned} \tag{14.125}$$

In this case the $\bar{\mathsf{h}}$ function takes on the compact form

$$\bar{\mathsf{h}}(a) = a + \frac{M}{r} a\cdot e_-\, e_-, \qquad e_- = e_t - e_r. \tag{14.126}$$

This algebraic form has a number of convenient algebraic features. The first is that the solution is of the form of the identity plus an interaction term, as is also the case in the Newtonian gauge setup. The second is that this form of $\bar{\mathsf{h}}(a)$ is a symmetric function. Finally, e_- is a null vector that satisfies $\bar{\mathsf{h}}(e_-) = e_-$. All of these features can be employed to simplify calculations.

The line element generated by our general form of $\bar{\mathsf{h}}$ function is

$$\begin{aligned} ds^2 =& (1 - 2M/r)\, dt^2 + 2(f_1 g_2 - f_2 g_1)\, dt\, dr - (f_1{}^2 - f_2{}^2)\, dr^2 \\ & - r^2 (d\theta^2 + \sin^2(\theta)\, d\phi^2). \end{aligned} \tag{14.127}$$

This in effect contains one arbitrary function, because the constraint on the determinant fixes one of the two unknown coefficients. The remaining unspecified degree of freedom lies in the rotation gauge, which does not affect the metric. We can draw an important conclusion about the metric by considering its behaviour at the horizon. There we must have $g_1 = \pm g_2$, and we know that $f_1 g_1 - f_2 g_2 = 1$ globally. It follows that

$$f_1 g_2 - f_2 g_1 = \pm 1 \quad \text{at } r = 2M, \tag{14.128}$$

so the off-diagonal term must be either $+1$ or -1 at the horizon. The presence of the horizon must break time reversal symmetry. This is to be expected. For a black hole (corresponding to the negative solution), the horizon is the place where particles can fall in, but cannot escape. The opposite value at the horizon (corresponding to a positive value of g_2 in the Newtonian gauge) defines an object from which particles can escape, but no particle can cross the horizon. This is called a *white hole*, though it is unclear whether such a solution defines a physically relevant object.

14.4 Quantum mechanics in a black hole background

The gauge theory formulation of gravity is motivated by constructing gauge fields to ensure that the Dirac equation is covariant under local rotations and displacements. We now study the effects of the black hole gauge fields on a Dirac fermion. Assuming that no electromagnetic couplings are present, the minimally-coupled equation takes the familiar form

$$D\psi I\boldsymbol{\sigma}_3 = m\psi\gamma_0. \tag{14.129}$$

The simplicity of the $\bar{\mathsf{h}}$ field in the Newtonian gauge suggests that this will be the simplest gauge to work in. As always, we must ensure that the all physical predictions are gauge-invariant. With the gravitational fields as described in equations (14.64) and (14.65), the Dirac equation becomes

$$\nabla\psi I\boldsymbol{\sigma}_3 - \left(\frac{2M}{r}\right)^{1/2}\gamma_0\left(\frac{\partial}{\partial r}\psi + \frac{3}{4r}\psi\right)I\boldsymbol{\sigma}_3 = m\psi\gamma_0. \tag{14.130}$$

If we pre-multiply by γ_0 and employ the i symbol to represent right-sided multiplication by $I\boldsymbol{\sigma}_3$, then equation (14.130) becomes

$$i\partial_t\psi = -i\boldsymbol{\nabla}\psi + i\left(\frac{2M}{r}\right)^{1/2}\frac{1}{r^{3/4}}\frac{\partial}{\partial r}\left(r^{3/4}\psi\right) + m\bar{\psi}, \tag{14.131}$$

where $\bar{\psi} = \gamma_0\psi\gamma_0$. We see that the Newtonian gauge has enabled us to write the Dirac equation in a very straightforward Hamiltonian form. One reason for this simplicity is that the spatial sections defined by the time coordinate t are flat.

The interaction Hamiltonian in equation (14.131), with all constants included, is

$$\hat{H}_I(\psi) = i\hbar\left(\frac{2GM}{r}\right)^{1/2}\left(\frac{\partial}{\partial r} + \frac{3}{4r}\right)\psi. \tag{14.132}$$

This single term incorporates all gravitational effects exerted by a black hole on a Dirac fermion. A number of observations can be made immediately. The first is that the interaction Hamiltonian does not depend on the mass of the particle, which is how the equivalence principle is embodied in the Dirac equation. The second point is that $\hat{H}_I$ does not depend on the speed of light. The non-relativistic approximation is therefore straightforward, following the technique of section 8.3.3. To lowest order in c^{-1} we obtain the Schrödinger equation with interaction determined by $\hat{H}_I$. For stationary states this equation is

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\psi + i\hbar\left(\frac{2GM}{r}\right)^{1/2}\frac{1}{r^{3/4}}\frac{\partial}{\partial r}\left(r^{3/4}\psi\right) = E\psi, \tag{14.133}$$

where ψ now denotes the Schrödinger wave function. This equation is simplified by introducing the phase-transformed variable

$$\Psi = \psi\exp\left(-i(8r/a_G)^{1/2}\right), \tag{14.134}$$

where

$$a_G = \frac{\hbar^2}{GMm^2}. \tag{14.135}$$

The distance a_G is the gravitational analogue of the Bohr radius for the hydrogen atom. The new variable Ψ satisfies the simple equation

$$-\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2\Psi - \frac{Mm}{r}\Psi = E\Psi. \tag{14.136}$$

This is precisely the equation we would expect if we used the Newtonian gravitational potential. The solutions for Ψ are therefore Coulomb wavefunctions. The non-relativistic limit enables us to make two immediate predictions. The first is that a spectrum of bound states should exist, with similar properties to that of the hydrogen atom. The second is that, in the non-relativistic limit, the scattering cross section should be determined by the Rutherford formula. This latter prediction is confirmed in the following section.

The interaction Hamiltonian $\hat{H}_I$ hides a significant feature, which is that it is not Hermitian due to the presence of the singularity. To see this we form the difference between $\hat{H}_I$ and its adjoint. With ϕ and ψ both Dirac spinors we find that

$$\begin{aligned}\int d^3x\,\langle\phi^\dagger\hat{H}_I(\psi)\rangle_q &= \sqrt{2M}\int d\Omega\int_0^\infty dr\,\langle r^{3/4}\phi^\dagger\partial_r(r^{3/4}\psi)I\boldsymbol{\sigma}_3\rangle_q \\ &= \int d^3x\,\langle(\hat{H}_I(\phi)^\dagger\psi\rangle_q + \sqrt{2M}\int d\Omega\left[r^{3/2}\langle\phi^\dagger\psi I\boldsymbol{\sigma}_3\rangle_q\right]_0^\infty,\end{aligned} \tag{14.137}$$

where we follow the convention of section 8.1.2. We will see shortly that all wavefunctions approach the origin as $r^{-3/4}$. The boundary term at the origin therefore does not vanish, and the Hamiltonian is not Hermitian. It follows that any normalisable stationary state must have an imaginary component to its energy. This is sensible. For all states the covariant current vector is always timelike. Inside the horizon this vector must point inwards, towards the singularity, so current density is inevitably swept onto the singularity. This implies that bound states must necessarily decay, so we expect the energy to have an imaginary component.

14.4.1 Scattering

The Dirac equation (14.131) is ideally suited to a perturbative scattering calculation employing the methods of section 8.5. We seek an iterative solution to the Green's function equation

$$\left(i\hat{\nabla}_2 - \hat{B}(x_2) - m\right) S_G(x_2, x_1) = \delta^4(x_2 - x_1), \tag{14.138}$$

where

$$B(x) = i\hat{\gamma}_0 \left(\frac{2M}{r}\right)^{1/2} \left(\frac{\partial}{\partial r} + \frac{3}{4r}\right). \tag{14.139}$$

As usual, the hats denote operators which act on spinors, and in this section we retain the familiar i symbol to denote the complex structure.

The iterative solution to equation (14.138) is given by

$$\begin{aligned} S_G(x_f, x_i) &= S_F(x_f, x_i) + \int d^4x_1\, S_F(x_f, x_1) B(x_1) S_F(x_1, x_i) \\ &+ \iint d^4x_1\, d^4x_2\, S_F(x_f, x_1) B(x_1) S_F(x_1, x_2) B(x_2) S_F(x_2, x_i) + \cdots, \end{aligned} \tag{14.140}$$

where $S_F(x_2, x_1)$ is the free-field, position-space Feynman propagator. The interaction term $B(x)$ is independent of time so energy is conserved throughout the interaction. Converting to momentum space we find that the scattering multivector T_{fi}, as defined in equation (8.229), is given by

$$T_{fi} = (\hat{p}_f + m) \left(B(\boldsymbol{p}_f, \boldsymbol{p}_i) + \int \frac{d^3k}{(2\pi)^3} B(\boldsymbol{p}_f, \boldsymbol{k}) \frac{\hat{k} + m}{k^2 - m^2 + i\epsilon} B(\boldsymbol{k}, \boldsymbol{p}_i) + \cdots \right). \tag{14.141}$$

Here $B(\boldsymbol{p}_2, \boldsymbol{p}_1)$ denotes the spatial Fourier transform of the interaction term,

$$B(\boldsymbol{p}_2, \boldsymbol{p}_1) = (2M)^{1/2} i\hat{\gamma}_0 \int d^3x\, e^{-i\boldsymbol{p}_2\cdot\boldsymbol{x}} \frac{1}{r^{1/2}} \left(\frac{\partial}{\partial r} + \frac{3}{4r}\right) e^{i\boldsymbol{p}_1\cdot\boldsymbol{x}}, \tag{14.142}$$

where bold symbols refer to spatial components only. To evaluate this we first write

$$B(\boldsymbol{p}_2, \boldsymbol{p}_1) = (2M)^{1/2} i\hat{\gamma}_0 \left(\frac{3}{4} f(\boldsymbol{p}_1 - \boldsymbol{p}_2) + \left. \frac{\partial f(\lambda\boldsymbol{p}_1 - \boldsymbol{p}_2)}{\partial\lambda} \right|_{\lambda=1} \right), \tag{14.143}$$

where

$$f(\boldsymbol{p}) = \int d^3x \frac{e^{i\boldsymbol{p}\cdot\boldsymbol{x}}}{r^{3/2}} = \left(\frac{2\pi}{|\boldsymbol{p}|}\right)^{3/2}. \tag{14.144}$$

We therefore find that the momentum-space interaction is governed by the *vertex factor*

$$B(\boldsymbol{p}_2, \boldsymbol{p}_1) = 3\pi^{3/2} i(M)^{1/2} \frac{{\boldsymbol{p}_2}^2 - {\boldsymbol{p}_1}^2}{|\boldsymbol{p}_2 - \boldsymbol{p}_1|^{7/2}} \hat{\gamma}_0. \tag{14.145}$$

This factor has the unusual feature of vanishing if the ingoing and outgoing particles are on-shell, because energy is conserved throughout the process. It follows that the lowest order contribution to the scattering cross section vanishes. This is to be expected, as the vertex factor goes as $\sqrt{M}$, and we expect the amplitude to go as M to recover the Rutherford formula in the low velocity limit.

Working to the lowest non-zero order in M the scattering multivector becomes

$$T_{fi} = -9\pi^3 M(\hat{p}_f + m)\hat{\gamma}_0 I_1 \hat{\gamma}_0, \tag{14.146}$$

where

$$I_1 = \int \frac{d^3k}{(2\pi)^3} \frac{{\boldsymbol{p}_f}^2 - \boldsymbol{k}^2}{|\boldsymbol{p}_f - \boldsymbol{k}|^{7/2}} \frac{\hat{k} + m}{k^2 - m^2 + i\epsilon} \frac{\boldsymbol{k}^2 - {\boldsymbol{p}_i}^2}{|\boldsymbol{k} - \boldsymbol{p}_i|^{7/2}}. \tag{14.147}$$

Here we have explicitly included a factor of $i\epsilon$ to ensure that any poles in the complex plane are navigated in the correct manner. However, we have

$$k^2 - m^2 = E^2 - \boldsymbol{k}^2 - m^2 = \boldsymbol{p}^2 - \boldsymbol{k}^2, \tag{14.148}$$

where E is the particle energy and $\boldsymbol{p}^2 = {\boldsymbol{p}_i}^2 = {\boldsymbol{p}_f}^2$. The pole in the propagator is therefore cancelled by the vertex factors, so there is no need for the factor of $i\epsilon$ in the denominator. The integral we need to evaluate is therefore

$$I_1 = \int \frac{d^3k}{(2\pi)^3} \frac{\boldsymbol{k}^2 - \boldsymbol{p}^2}{|\boldsymbol{p}_f - \boldsymbol{k}|^{7/2}|\boldsymbol{k} - \boldsymbol{p}_i|^{7/2}} (\hat{k} + m), \tag{14.149}$$

and the result of this integral is

$$I_1 = \frac{1}{9\pi^2 \boldsymbol{q}^2}\big(2m + 3(\hat{p}_f + \hat{p}_i) - 4E\hat{\gamma}_0\big), \tag{14.150}$$

where $\boldsymbol{q} = \boldsymbol{p}_f - \boldsymbol{p}_i$. The scattering multivector is now given by

$$T_{fi} = -\frac{4\pi M}{\boldsymbol{q}^2}\big(E(2E + \boldsymbol{q}) + \boldsymbol{p}^2 + \boldsymbol{p}_f \boldsymbol{p}_i\big). \tag{14.151}$$

This should be contrasted with the equivalent expression for Coulomb scattering, given in equation (8.237). We see immediately that the coupling term goes with the particle energy, rather than its mass. This is because the interaction Hamiltonian is independent of m. The unpolarised cross section is given by

$$\begin{aligned}\frac{d\sigma}{d\Omega} &= \frac{|T_{fi}|^2}{16\pi^2} \\ &= \frac{2M^2}{\boldsymbol{q}^4}\big(m^2(E^2 - \boldsymbol{p}_f\cdot\boldsymbol{p}_i) + (2E^2 - m^2)^2 + 4E^2\boldsymbol{p}_f\cdot\boldsymbol{p}_i\big).\end{aligned} \tag{14.152}$$

If we now let $v = |\boldsymbol{p}|/E$ denote the particle velocity, and θ the scattering angle, we arrive at the simple expression

$$\frac{d\sigma}{d\Omega} = \frac{M^2}{4v^4\sin^4(\theta/2)}\big(1 + 2v^2 - 3v^2\sin^2(\theta/2) + v^4 - v^4\sin^2(\theta/2)\big). \tag{14.153}$$

As demanded by the equivalence principle, this formula depends only on the incident velocity, and not on the particle mass. This confirms that the equivalence principle is directly encoded in the Dirac equation as a consequence of minimal coupling. The final cross section formula is gauge-invariant. We can perform analogues of this calculation in a range of different gauges, and the same result is obtained in all cases. Furthermore, all terms in the result have local, gauge-invariant definitions. The mass M can be defined in terms of tidal forces, and the velocity v is that measured locally by observers in radial free fall from rest at infinity. The angle θ is the angle between asymptotic in and out states, measured locally in the asymptotic regime.

The cross section of equation (14.153) confirms that the low velocity limit recovers the Rutherford formula. The massless limit $m \mapsto 0$ is also well defined, and is obtained by setting $v = 1$. This produces the simple formula

$$\frac{d\sigma}{d\Omega} = \frac{M^2\cos^2(\theta/2)}{\sin^4(\theta/2)}. \tag{14.154}$$

The small angle limit to this gives a cross section going as $(4M)^2/\theta^4$. This recovers the classical formula for the bending of light by a massive source. While the calculation here has assumed a point mass source, the small angle limit is appropriate for any localised source of gravitational mass M. The massless limit contains a surprise in the backward direction, however. Simulations of scattering based on massless particles following null geodesics reveal a large 'glory' scattering in the backward direction. This is absent from the quantum treatment, and is a diffraction effect for massless spin-1/2 particles that is not evident at the classical level. The scheme described here can be modified to the case of a scalar field, and produces the differential cross section

$$\frac{d\sigma}{d\Omega} = \frac{M^2}{4v^4\sin^4(\theta/2)}(1 + v^2)^2. \tag{14.155}$$

Again, we see that the equivalence principle is obeyed, and the various small angle and low velocity approximations are retained. The classical cross section contains further structure, attributable to multiple orbits. In the quantum framework these effects should be present in the higher-order terms.

14.4.2 Stationary states and angular separation

The Dirac equation in the Newtonian gauge is immediately separable in space and time, and admits stationary state solutions of the form

$$\psi(x) = \psi(\boldsymbol{x}) \exp(-EtI\boldsymbol{\sigma}_3). \tag{14.156}$$

If the state is normalisable then E contains an imaginary component determined by

$$\mathrm{Im}(E) = -\frac{\sqrt{2M}}{2N} \lim_{r\to 0} r^{3/2} \int d\Omega \, \langle \psi^\dagger \psi \rangle, \tag{14.157}$$

where N is the normalisation constant

$$N = \int d^3x \, \langle \psi^\dagger \psi \rangle. \tag{14.158}$$

As expected, the sign of the imaginary component of E corresponds to a decaying wavefunction. This behaviour is independent of the sign of the real part of E, so both positive and negative energy states must decay. For scattering states we do not demand that ψ is normalisable, and can look for solutions where the energy is real, with $E > m$.

With the time dependence separated out, equation (14.131) reduces to

$$\boldsymbol{\nabla}\psi - (2M/r)^{1/2} r^{-3/4} \partial_r \big(r^{3/4}\psi\big) = iE\psi - im\bar{\psi}. \tag{14.159}$$

To solve this equation we follow the standard procedure for a central potential and separate out the angular dependence. This is achieved using the spherical monogenics, described in section 8.4.1. We assume that the wavefunction takes the standard form of

$$\psi(\boldsymbol{x}, \kappa) = \begin{cases} \psi_l^m u(r) + \sigma_r \psi_l^m v(r) I\boldsymbol{\sigma}_3 & \kappa = l+1, \\ \sigma_r \psi_l^m u(r)\boldsymbol{\sigma}_3 + \psi_l^m I v(r) & \kappa = -(l+1), \end{cases} \tag{14.160}$$

where κ is a non-zero integer and $u(r)$ and $v(r)$ are complex functions of r only. On substituting this wavefunction into the Dirac equation (14.159) we obtain the pair of coupled radial equations

$$\begin{pmatrix} 1 & -(2M/r)^{1/2} \\ -(2M/r)^{1/2} & 1 \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \boldsymbol{A} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{14.161}$$

where

$$\boldsymbol{A} = \begin{pmatrix} \kappa/r & i(E+m) - (2M/r)^{1/2}(4r)^{-1} \\ i(E-m) - (2M/r)^{1/2}(4r)^{-1} & -\kappa/r \end{pmatrix}, \tag{14.162}$$

u_1 and u_2 are the reduced functions defined by

$$u_1 = ru, \qquad u_2 = irv \tag{14.163}$$

and the primes denote differentiation with respect to r. The form of this equation should be contrasted with the hydrogen atom of section 8.4.3.

To analyse equation (14.161) we first rewrite it in the equivalent form

$$(1 - 2M/r) \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} 1 & (2M/r)^{1/2} \\ (2M/r)^{1/2} & 1 \end{pmatrix} \boldsymbol{A} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{14.164}$$

This makes it clear that the equations have regular singular points at the origin and horizon ($r = 2M$), as well as an irregular singular point at $r = \infty$. Unfortunately, the special function theory required to deal with such equations has not been developed. Hypergeometric functions are appropriate for differential equations with three regular singular points, or one regular and one irregular singular point. An attempt to generalise hypergeometric functions results in Heun's equation, but most techniques for handling this involve series solutions and numerical integration, so these are the techniques that must be applied here. The presence of the three singular points implies that any power series will have a limited radius of convergence, so typically these can only be used to define initial data for numerical integration routines.

A Frobenius series about the origin shows that both u_1 and u_2 approach the origin as $r^{1/4}$. It follows that the wavefunction goes as $r^{-3/4}$ near the origin, as was stated earlier. For normalisable states this behaviour ensures that the energy contains an imaginary decay factor. Next we construct a series about the horizon by writing

$$u_1 = \eta^s \sum_{k=0}^{\infty} a_k \eta^k, \qquad u_2 = \eta^s \sum_{k=0}^{\infty} b_k \eta^k, \tag{14.165}$$

where $\eta = r - 2M$. On substituting this series into equation (14.164), and setting $\eta = 0$, we obtain

$$\frac{s}{2M} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \kappa/(2M) & i(E+m) - (8M)^{-1} \\ i(E-m) - (8M)^{-1} & -\kappa/(2M) \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}. \tag{14.166}$$

The two values of the index s for which this has non-zero solutions are

$$s = 0 \quad \text{and} \quad s = -\tfrac{1}{2} + 4iME. \tag{14.167}$$

The $s = 0$ solution corresponds to an analytic power series with a well-defined wavefunction at the horizon. Such solutions are certainly physical. The second root gives rise to a wavefunction that is singular at the horizon, and as such is physically inadmissible. As a consequence, it is not possible to construct a complete set of outgoing modes at infinity and in any scattering process some of the wavefunction is lost. This is the quantum-mechanical description of absorption by a black hole.

Before proceeding, we should confirm that the two indicial roots at the horizon are gauge-invariant, and not an artifact of our various gauge choices. This is important because the singular index can be used to determine the Hawking temperature of the black hole. The method we use to confirm gauge invariance is quite general and can be applied to a range of situations. We start by keeping the gauge unspecified so that, after separating out the angular dependence, the Dirac equation reduces to

$$\begin{pmatrix} L_r & L_t \\ L_t & L_r \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \kappa/r - G/2 & im - F/2 \\ -im - F/2 & -\kappa/r - G/2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{14.168}$$

We can still assume that the time dependence is of the form $\exp(-iEt)$, so that equation (14.168) becomes

$$\begin{pmatrix} g_1 & g_2 \\ g_2 & g_1 \end{pmatrix} \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \boldsymbol{B} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \tag{14.169}$$

where

$$\boldsymbol{B} = \begin{pmatrix} \kappa/r - G/2 + if_2E & i(m + f_1E) - F/2 \\ -i(m - f_1E) - F/2 & -\kappa/r - G/2 + if_2E \end{pmatrix}. \tag{14.170}$$

The form of time dependence is gauge-invariant, since the time coordinate is defined by the requirement that the Riemann tensor is stationary. Given a time coordinate t, a general displacement consistent with this requirement takes the form

$$t \mapsto t' = t + \alpha(r), \tag{14.171}$$

where α is a differentiable function of r. This ensures that stationary states all go as $\exp(-iEt)$, regardless of the choice of time coordinate.

Now, since ${g_1}^2 - {g_2}^2 = 1 - 2M/r$ holds for vacuum solutions in all gauges, we obtain

$$\big(1 - 2M/r\big) \begin{pmatrix} u_1' \\ u_2' \end{pmatrix} = \begin{pmatrix} g_1 & -g_2 \\ -g_2 & g_1 \end{pmatrix} \boldsymbol{B} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}. \tag{14.172}$$

We again look for a power series solution of the form of equation (14.165), and setting $\eta = 0$ produces the indicial equation

$$\det \left[\begin{pmatrix} g_1 & -g_2 \\ -g_2 & g_1 \end{pmatrix} \boldsymbol{B} - \frac{s}{r} \mathsf{I} \right]_{r=2M} = 0, \tag{14.173}$$

where I is the identity matrix. For vacuum fields we know that

$$g_1 G - g_2 F = \tfrac{1}{2} \partial_r({g_1}^2 - {g_2}^2) = M/r^2, \tag{14.174}$$

which is gauge-invariant. It follows that the solutions to the indicial equation are

$$s = 0 \quad \text{and} \quad s = -\tfrac{1}{2} + 4iME(g_1 f_2 - g_2 f_1). \tag{14.175}$$

But, as discussed in section 14.3.6, at the horizon we have

$$(g_1 f_2 - g_2 f_1) = \pm 1, \tag{14.176}$$

with the positive sign corresponding to the black hole case. The indices of the Dirac equation are therefore gauge-invariant. Similar arguments can be applied to scalar and higher-spin fields.

14.4.3 Quantum absorption

We are now in a position to give a full, quantum-mechanical description of absorption by a black hole. At the horizon the solutions of the Dirac equation separate into two branches, one regular and one singular. The singular branch is unphysical and cannot be excited by finite incoming waves. The regular branch is finite at the horizon, with an inward-pointing current. This gives rise to absorption. To understand this process in detail we need to study the asymptotic form of the regular solutions and determine their split into incoming and outgoing modes. We can then construct an arbitrary incoming mode (typically a plane wave) and study the amount of scattered radiation. Any radiation that is not scattered is absorbed.

In absorption and scattering problems we are interested in states with real energy E, $E > m$. For such states the spatial current $\boldsymbol{J}$ is conserved, and for angular eigenstates we obtain the conserved Wronskian W:

$$W = g_1(u_1 u_2^\dagger + u_1^\dagger u_2) + g_2(u_1 u_1^\dagger + u_2 u_2^\dagger). \tag{14.177}$$

This measures the total outward flux over a surface of radius r, and we have written W in an arbitrary gauge. At the horizon we see that

$$W = -g_1 |u_1 - u_2|^2, \tag{14.178}$$

and so the flux is *inwards* for all regular solutions. This is to be expected, as the current must point inwards at the horizon.

For explicit calculations we return to the Newtonian gauge. The radial equation (14.164) is straightforward to integrate numerically. We start with a power series expansion around the horizon of the regular solution. This allows us to find values of u_1 and u_2 a small distance either side of the horizon. These values are then used to initiate numerical integration of the equations, both inwards and outwards. To visualise the solutions it is convenient to plot the radial density function $P(r)$:

$$P(r) = |u_1|^2 + |u_2|^2. \tag{14.179}$$

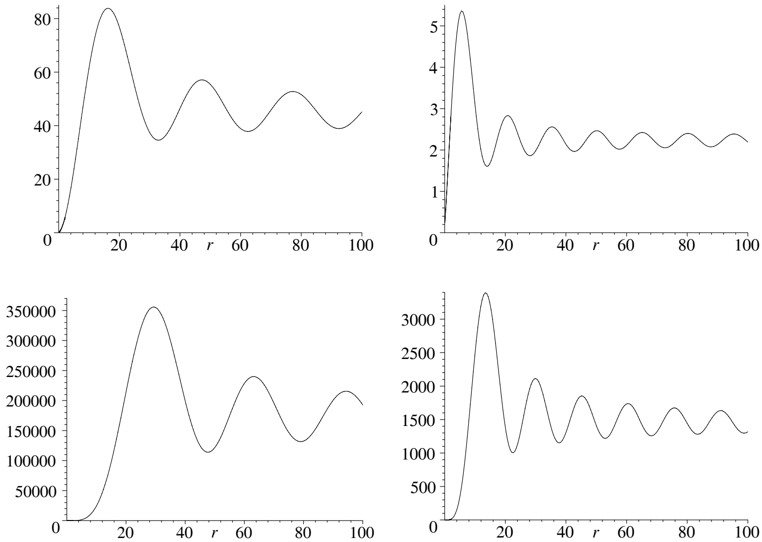

Figure 14.6 *The radial density for scattered states.* The plots show $P(r)$ as a function of radius. The horizon lies at $r = 2$, and the product mM is set to 0.01 in units of m_p^2, where m_p is the Planck mass. The modes are scaled so that the Wronskian is -1, and only the regular solution is plotted. The top two diagrams are for $\kappa = 1$, with $E = 10mc^2$ (left) and $E = 20mc^2$ (right). The bottom two diagrams are for $\kappa = 2$, with $E = 10mc^2$ (left) and $E = 20mc^2$ (right).

In physical terms $P(r)$ is r^2 times the timelike component of the Dirac current, as measured by observers in radial free fall from rest at infinity. It is only in the Newtonian gauge that this definition gives rise to the simple formula of equation (14.179).

In figure 14.6 we plot $P(r)$ for a range of energies and angula momenta. The plots are for scattering states, so the wavefunctions are unnormalised. For the sake of comparison the magnitude of each mode is fixed by setting the Wronskian to -1. The gravitational coupling is controlled by the dimensionless quantity

$$\frac{GMm}{\hbar c} = \frac{Mm}{m_p^2}, \tag{14.180}$$

where m_p is the Planck mass. In figure 14.6 we have used a dimensionless coupling of 0.01. The chosen energies of $10mc^2$ and $20mc^2$ imply that the modes are highly relativistic, and also ensure that the associated wavelengths are larger than the horizon size. To understand the asymptotic features of the plots we return to equation (14.161) and solve for the behaviour at large r. We find that the solutions behave asymptotically as

$$\begin{aligned}u_1 =&\beta \exp i\left(pr+\frac{M}{p}(m^2+2p^2)\ln(pr)\right)\mathrm{e}^{2iE(2Mr)^{1/2}}\\&+\alpha\exp -i\left(pr+\frac{M}{p}(m^2+2p^2)\ln(pr)\right)\mathrm{e}^{2iE(2Mr)^{1/2}}\end{aligned}\tag{14.181}$$

and

$$\begin{aligned}u_2 =&\frac{p\beta}{E+m}\exp i\left(pr+\frac{M}{p}(m^2+2p^2)\ln(pr)\right)\mathrm{e}^{2iE(2Mr)^{1/2}}\\&-\frac{p\alpha}{E+m}\exp -i\left(pr+\frac{M}{p}(m^2+2p^2)\ln(pr)\right)\mathrm{e}^{2iE(2Mr)^{1/2}},\end{aligned}\tag{14.182}$$

where $p^2 = E^2 - m^2$. The Wronskian is therefore equal to

$$W = -\frac{2p}{E+m}(|\alpha|^2 - |\beta|^2),\tag{14.183}$$

and the radial probability $P(r)$ is given asymptotically by

$$\begin{aligned}|u_1|^2+|u_2|^2 =&\frac{4m}{E+m}|\alpha|\,|\beta|\cos\left(2pr+\frac{2M(m^2+2p^2)}{p}\ln(pr)+\phi_0\right)\\&+\frac{2E}{E+m}\big(|\alpha|^2+|\beta|^2\big).\end{aligned}\tag{14.184}$$

The oscillations predicted by this formula are clearly visible in figure 14.6. The magnitudes of α and β determine the relative amounts of scattered and absorbed radiation present for a given mode. With the Wronskian held constant, all modes have a constant flux through the horizon onto the singularity. In the large r region $|\alpha|$ determines the amount of ingoing radiation, and $|\beta|$ the amount of outgoing radiation. As $|\alpha|$ increases, a smaller fraction of the radiation is absorbed and more is scattered. One effect that is clear in figure 14.6 is that as the angular momentum increases, for fixed energy, $|\alpha|$ also increases. That is, less radiation is absorbed for fixed energy as the angular momentum increases. This is precisely the behaviour we expect from classical considerations.

Given that each mode is normalised such that $W = -1$, then total absorption cross section is given by

$$\sigma_{abs} = \frac{\pi}{2p(E-m)}\sum_{\kappa\neq 0}\frac{|\kappa|}{|\alpha_\kappa|^2},\tag{14.185}$$

where α_κ is the value of α for each angular eigenmode. The values of α_κ are determined numerically by integrating the radial equations out to a suitable distance from the horizon and matching to the asymptotic forms of equations (14.181) and (14.182). Typically, we need to sum over a range of κ values before the sum settles down to its final result. The result of this sum, for a massive fermion, is plotted in figure 14.7. For energies close to the rest energy the absorption

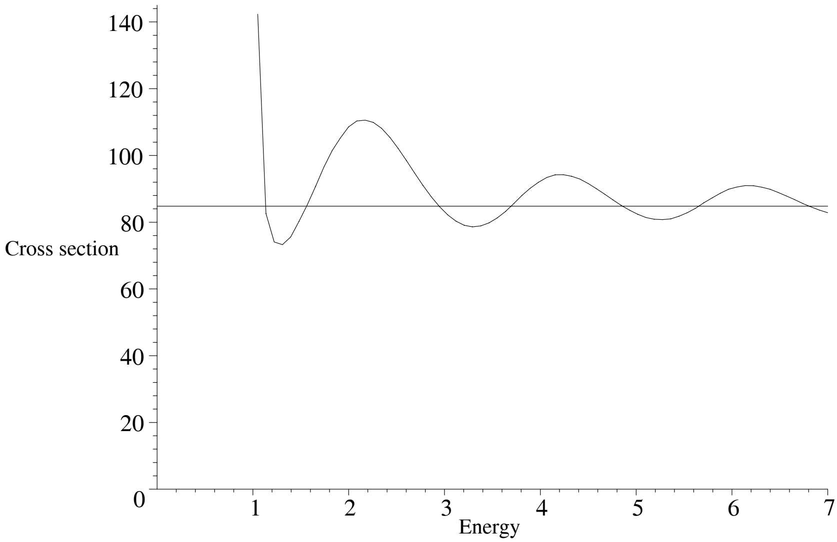

Figure 14.7 *The quantum absorption cross section.* The plot shows the total absorption cross section as a function of the incident energy. The dimensionless coupling Mm/m_p^2 is 0.1, and the energy is plotted in units of the rest energy mc^2. The horizontal line is the photon limit.

cross section follows the classical prediction. But at higher energies a series of oscillations are present as the wavelength becomes comparable with the horizon size. These oscillations take place around the photon limit of 27π, and are also present for massless particles. The precise form of these oscillations depends on the mass of the particle, so represents a quantum-mechanical violation of the equivalence principle.

14.5 Cosmology

The radial equations we have developed so far are easily adapted to the case of homogeneous, isotropic matter distributions. Such distributions provide a good model for the large scale distribution of matter in the observable universe. Before studying the field equations for such cosmological matter distributions, we must first introduce the cosmological constant. This was originally introduced by Einstein to allow the construction of static cosmological solutions, and for many years had been thought to be an unnecessary additional feature of general relativity. But experimental evidence, both from the cosmic microwave background and from distant supernovae, now favours models which do include a cosmological constant. There are also hints from quantum gravity that a cosmological constant should arise as a form of vacuum energy, though this is not well understood.

We start with the radial equations, as summarised in table 14.2. Inclusion of the cosmological constant Λ only modifies a handful of these equations. The mass function M becomes

$$M = \tfrac{1}{2}r\big({g_2}^2 - {g_1}^2 + 1 - \Lambda r^2/3\big), \tag{14.186}$$

and the derivatives of the g_1 and g_2 fields become

$$\begin{aligned} L_t g_2 &= G g_1 - M/r^2 + r\Lambda/3 - 4\pi r p, \\ L_r g_1 &= F g_2 + M/r^2 - r\Lambda/3 - 4\pi r \rho. \end{aligned} \tag{14.187}$$

The Riemann tensor is altered to

$$\begin{aligned} \mathcal{R}(B) =& 4\pi(\rho+p)B\cdot e_t\, e_t - \frac{1}{3}(8\pi\rho+\Lambda)B \\ & - \left(\frac{M}{2r^3} - \frac{2\pi}{3}\rho\right)(B + 3\sigma_r B \sigma_r) \end{aligned} \tag{14.188}$$

and we continue to assume that the matter distribution takes the form of an ideal fluid.

For cosmological models the matter distribution is assumed to be spatially homogeneous and isotropic, so that ρ and p are functions of time only. The mass function M is then given by

$$M(r,t) = \frac{4\pi}{3} r^3 \rho, \tag{14.189}$$

so the Riemann tensor also depends only on time. The equation for $L_r p$ tells us that G vanishes, and hence that

$$f_1 = 1. \tag{14.190}$$

The time coordinate t therefore measures the proper time for observers at rest with respect to the cosmological background. The derivatives of M and ρ similarly tell us that

$$F = \frac{g_2}{r} \tag{14.191}$$

and

$$\dot{\rho} = -\frac{3g_2}{r}(\rho+p). \tag{14.192}$$

For these to be consistent with the relation $L_r g_2 = F g_1$ we must have

$$F = H(t), \qquad g_2(r,t) = rH(t), \tag{14.193}$$

where $H(t)$ is a function of time only. The $L_t g_2$ equation now reduces to a simple equation for $H(t)$:

$$\dot{H} + H^2 - \frac{\Lambda}{3} = -\frac{4\pi}{3}(\rho + 3p). \tag{14.194}$$

The $\bar{\mathsf{h}}$ field	$\bar{\mathsf{h}}(a) = a + a\cdot e_r\big((g_1-1)e^r + H(t)re^t\big)$ ${g_1}^2 = 1 - kr^2 \exp\big(-2\int^t H(t')\,dt'\big)$
The ω field	$\omega(a) = H(t)a\wedge e_t - (g_1-1)/r\, a\wedge(e_r e_t)e_t$
Riemann tensor	$\mathcal{R}(B) = 4\pi(\rho+p)B\cdot e_t\, e_t - (8\pi\rho+\Lambda)/3\, B$
The density	$8\pi\rho = 3H(t)^2 - \Lambda + 3k\exp\big(-2\int^t H(t')\,dt'\big)$
Dynamical equations	$\dot{H} + H^2 - \Lambda/3 = -(4\pi/3)\,(\rho+3p)$ $\dot{\rho} = -3H(t)(\rho+p)$

Table 14.3 *Equations governing a homogeneous, isotropic perfect fluid.* The covariant vector e_t defines the rest frame of the universe. This is determined experimentally from the cosmic microwave background radiation. No other direction is contained in $\mathcal{R}(B)$, and all physical fields are functions of time only.

Finally, we are left with a pair of equations for g_1,

$$\begin{aligned} L_t g_1 &= 0, \\ L_r g_1 &= ({g_1}^2 - 1)/r. \end{aligned} \tag{14.195}$$

The second equation tells us that g_1 is of the form

$${g_1}^2 = 1 + r^2\phi(t). \tag{14.196}$$

The equation for $L_t g_1$ then tells us that $\phi(t)$ satisfies

$$\dot{\phi} = -2H(t)\phi. \tag{14.197}$$

It follows that g_1 is given by

$${g_1}^2 = 1 - kr^2 \exp\left(-2\int^t H(t')\,dt'\right), \tag{14.198}$$

where k is an arbitrary constant of integration which turns out to define the spatial geometry. The full set of equations describing a homogeneous perfect fluid are summarised in table 14.3.

14.5.1 Comparison with standard approach

The derivation of the cosmological equations presented here, as a special case of a spherical solution, differs from most presentations. To recover a more familiar set of equations we first introduce the distance function S defined by

$$H(t) = \frac{\dot{S}(t)}{S(t)}. \tag{14.199}$$

With this substitution we find g_1 is now simply

$$g_1{}^2 = 1 - kr^2/S^2. \tag{14.200}$$

Similarly, the $\dot{H}$ and density equations become

$$\begin{aligned}\frac{\ddot{S}}{S} - \frac{\Lambda}{3} &= -\frac{4\pi}{3}(\rho + 3p),\\ \frac{\dot{S}^2 + k}{S^2} - \frac{\Lambda}{3} &= \frac{8\pi}{3}\rho.\end{aligned} \tag{14.201}$$

These are the *Friedmann equations* of cosmology. Our derivation has focused attention on the *Hubble function* $H(t)$, rather than the distance scale $S(t)$. This is natural, as $H(t)$ is a directly measurable (gauge-invariant) quantity, whereas $S(t)$ is only defined up to an arbitrary scaling.

The Friedmann equations are usually derived by starting with a diagonal line element. This is obtained from the radial setup by the displacement defined by

$$f(x) = x\cdot e_t e_t + Sx\wedge e_t e_t. \tag{14.202}$$

Under this displacement, $\bar{\mathsf{h}}(a)$ transforms to

$$\bar{\mathsf{h}}'(a) = a\cdot e_t e_t + \frac{1}{S}\big((1 - kr^2)^{1/2} a\cdot e_r e^r + a\wedge\sigma_r\,\sigma_r\big), \tag{14.203}$$

and the line element this defines is

$$ds^2 = dt^2 - \frac{S^2}{1 - kr^2}dr^2 - S^2 r^2\big(d\theta^2 + \sin^2(\theta)\,d\phi^2\big). \tag{14.204}$$

In this gauge we can see clearly that S controls the distance scale, and k controls the spatial geometry. We can always choose the scale such that k is either zero or ± 1. A k of zero corresponds to a spatially flat universe, which is favoured on theoretical grounds and is consistent with observations. The non-zero values correspond to an open universe ($k < 0$, defining hyperbolic geometry) or a closed universe ($k > 0$, defining spherical geometries). These three spatial geometries are the only spatially homogeneous and isotropic models we can consider. These geometries are discussed in more detail in chapter 10. Which model is appropriate for the universe on its largest scales is determined by the present values of the density and Hubble function. Most experiments find that the universe is close to the critical density ($k = 0$), but no experiment can ever conclusively prove that k is zero. Any slight deviation in the density away from the critical value implies that k is non-zero. The fact that the universe is so close to its critical density has led theoreticians to propose a range of models which force the universe to have $k = 0$. The most popular of these is provided by *inflationary cosmology*, in which the universe passes through a stage of rapid inflation, so that all spatial sections are expanded dramatically and become essentially flat.

14.5.2 Density perturbations and cluster formation

We will not discuss the detailed solutions of the cosmological equations in this book. This is a large subject and is covered in detail in a range of modern textbooks. Here we discuss an application where the derivation from the radial equations is particularly helpful. The problem of interest is the growth of a perturbation in a cosmological background. The perturbation is assumed to be spherically-symmetric, and the coordinate system is centred on the perturbation. To simplify matters further, we ignore the cosmological constant and set the pressure to zero. We are therefore dealing with a simple model of a pressureless fluid collapsing under the influence of its own gravity.

Returning to the radial equations in table 14.2, we see that for zero pressure we have $G = 0$ and $f_1 = 1$. The matter therefore follows geodesics, and t measures the proper time for observers comoving with the matter. The mass satisfies

$$L_t M = 0, \tag{14.205}$$

which says that the mass M enclosed within radius r is conserved along the fluid streamlines. The operator L_t is clearly the comoving derivative along the fluid streamlines. The function g_1 is also conserved along a streamline, and the equations integrate straightforwardly to determine the streamlines (geodesics). The form of the geodesic depends on the value of g_1, and there are three cases to consider:

1. ${g_1}^2 < 1$. This case includes closed cosmologies, and the matter streamlines are defined by

$$\begin{aligned} r &= \frac{M}{1-{g_1}^2}(1-\cos(\eta)), \\ t - t_i &= \frac{M}{(1-{g_1}^2)^{3/2}}\big(\eta - \sin(\eta) - \eta_i + \sin(\eta_i)\big) \end{aligned} \tag{14.206}$$

where η parameterises the curve, and η_i is determined from the initial value of r at time t_i. The velocity g_2 is given by

$$g_2 = \frac{M}{r(1-{g_1}^2)^{1/2}}\sin(\eta) \tag{14.207}$$

and η_i is fixed in the range $0 < \eta_i < 2\pi$ by determining whether the initial velocity is inwards or outwards. Setting $\eta_i = \pi$ corresponds to starting from rest, and provides a simple model for black hole formation.

2. ${g_1}^2 = 1$. This case include flat cosmologies, and the equations integrate directly to give

$$t - t_i = \frac{2(r^{3/2} - r_i^{3/2})}{3(2M)^{1/2}}. \tag{14.208}$$

The velocity is chosen outwards to avoid a singularity forming instantaneously.

3. ${g_1}^2 > 1$. This case includes open cosmologies. The streamlines are parameterised by

$$\begin{aligned} r &= \frac{M}{{g_1}^2 - 1}(\cosh(\eta) - 1), \\ t - t_i &= \frac{M}{({g_1}^2 - 1)^{3/2}}(\sinh(\eta) - \eta - \sinh(\eta_i) + \eta_i) \end{aligned} \tag{14.209}$$

and the velocity is given by

$$g_2 = \frac{M}{r({g_1}^2 - 1)^{1/2}} \sinh(\eta). \tag{14.210}$$

For this case it is also necessary to start with an initial outward velocity, in order to avoid streamline crossing.

By working globally in the Newtonian gauge we keep simple control over the initial conditions. For these we wish to set up a small perturbation in a finite region, such that outside the perturbation the system evolves as a homogeneous cosmology. This will be the case provided the average density in the perturbation matches the external universe. Suppose that the perturbation initially has width r_i and the external cosmology has initial values ρ_i and H_i for the density and Hubble function respectively. We introduce the dimensionless variables

$$x = \frac{r}{r_i}, \qquad v(x) = \frac{g_2(r, t_i) - rH_i}{r_i H_i}, \qquad f(x) = \frac{\rho(r, t_i) - \rho_i}{\rho_i}. \tag{14.211}$$

The functions $f(x)$ and $v(x)$ are related by

$$x^2 f(x) = -\frac{d}{dx}(x^2 v(x)), \tag{14.212}$$

with both $f(x)$ and $v(x)$ vanishing at the boundary ($x = 1$). Equation (14.212) ensures that the model is correctly compensated, so that the perturbation has no effect on the external cosmology. (Equation (14.212) also ensures that no decaying modes are present in the perturbation left over from the linear regime.) To fix $f(x)$ and $v(x)$ we choose a parameter n, which controls the polynomial degree of the functions, and also fix the value of the velocity gradient at the origin. The function $v(x)$ is then a polynomial of degree $2n+1$, formed as follows. At the centre we set $v = 0$, and the first derivative is determined by the velocity gradient. The remaining derivatives up to order n are set to zero. Similarly, at the boundary v is chosen such that g_2 matches the exterior value of rH_i up to the first n derivatives. The result is a simple function controlling the perturbation, and for each initial value of r the fluid streamlines can be plotted easily. An example of these streamlines is shown in figure 14.8.

If the system is allowed to evolve for a suitable amount of time, it provides

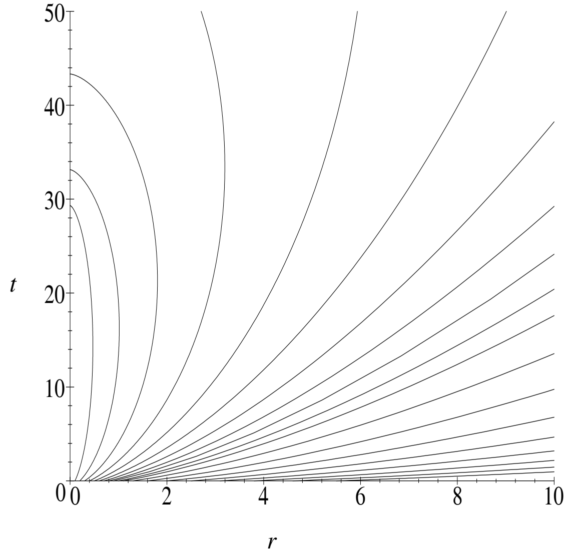

Figure 14.8 *Matter streamlines for an $n = 3$ model.* The perturbation has initial width 1, with $H_i = 1$ and $\rho_i = 3/8\pi$. The velocity gradient at the centre of the perturbation is 0.95. The central region is therefore moving inwards relative to the Hubble flow, so recollapses to a singularity after a finite time. All units are arbitrary.

a good model of a cluster of galaxies sitting inside a cosmological background. One can then study photon paths in this model, to look for lensing effects, or temperature perturbations in the cosmic microwave background. One weakness with these models is that no pressure is included, so the cluster has no means of supporting itself. This implies that a singularity forms after a finite term (determined by the central density and velocity gradient). The model then describes a black hole, sitting in an expanding universe.

14.5.3 The Dirac equation in a cosmological background

A good illustration of the full gravitational equations, with torsion included, is provided by the case of a Dirac field coupled self-consistently to gravity. The equations governing this system are

$$\begin{aligned} \mathcal{H}(a) &= 4\pi(\psi I\gamma_3\tilde{\psi})\cdot a, \\ \mathcal{G}(a) - \Lambda a &= 8\pi\langle a\cdot D\,\psi I\gamma_3\tilde{\psi}\rangle_1, \\ D\psi I\gamma_3 &= m\psi. \end{aligned} \tag{14.213}$$

This system of equations is highly non-linear and extremely difficult to analyse in all but the simplest of situations. Here we are interested in cosmological solutions, for which all fields are functions of time only. We also restrict our discussion to the spatially flat case ($k = 0$), so that we can write

$$\bar{\mathsf{h}}(a) = a + rH(t)a\cdot e_r\, e_t. \tag{14.214}$$

The ω function is given by

$$\omega(a) = H(t)a\wedge e_t + \tfrac{1}{2}\kappa a\cdot\mathcal{S}, \tag{14.215}$$

where $\kappa = 8\pi$ and $\mathcal{S}$ denotes the spin trivector:

$$\mathcal{S} = \tfrac{1}{2}\psi I\gamma_3\tilde{\psi}. \tag{14.216}$$

After a little work, the Einstein tensor evaluates to

$$\mathcal{G}(a) = 2\dot{H}a\wedge e_t\, e_t + 3H^2 a - \tfrac{1}{2}\kappa a\cdot(\mathcal{D}\cdot\mathcal{S}) + \tfrac{1}{2}\kappa^2 a\cdot\mathcal{S}\,\mathcal{S} - \tfrac{3}{4}\kappa^2\mathcal{S}^2 a, \tag{14.217}$$

and the matter energy-momentum tensor is

$$\mathcal{T}(a) = \langle a\cdot e_t\dot{\psi}I\gamma_3\tilde{\psi} + Ha\wedge e_t\,\mathcal{S} + \tfrac{1}{2}\kappa a\cdot\mathcal{S}\,\mathcal{S}\rangle_1. \tag{14.218}$$

Finally, the Dirac equation is now

$$(e_t\partial_t + \tfrac{3}{2}He_t + \tfrac{3}{4}\kappa\mathcal{S})\psi I\gamma_3 = m\psi, \tag{14.219}$$

which has the unusual feature of being nonlinear, due to the presence of the spin term.

We will construct the simplest solution to this system by setting the spinor ψ equal to a magnitude and phase only:

$$\psi = \rho(t)^{1/2}\mathrm{e}^{-I\boldsymbol{\sigma}_3\chi(t)}. \tag{14.220}$$

The Dirac equation therefore reduces to the pair of equations

$$\begin{aligned}\dot{\rho} &= -3\rho H,\\ \dot{\chi} &= 3\pi\rho + m.\end{aligned} \tag{14.221}$$

The Einstein equation yields the final pair of equations

$$\begin{aligned}3H^2 - 12\pi^2\rho^2 - 8\pi m\rho - \Lambda &= 0,\\ 2\dot{H} + 3H^2 + 12\pi^2\rho^2 - \Lambda &= 0.\end{aligned} \tag{14.222}$$

The second of these follows from the first and the equation for $\dot{\rho}$. These equations are solved by

$$\rho = \frac{\beta^2}{6\pi\sinh(\beta t)\big(m\sinh(\beta t) + \beta\cosh(\beta t)\big)}, \tag{14.223}$$

where

$$\beta = \frac{\sqrt{3\Lambda}}{2}. \tag{14.224}$$

The initial singularity is chosen to correspond to $t = 0$. The Hubble function is similarly given by

$$H(t) = \frac{\beta^2 + 2\beta \sinh^2(\beta t) + 2m\beta \sinh(\beta t)\cosh(\beta t)}{3\sinh(\beta t)\big(m\sinh(\beta t) + \beta\cosh(\beta t)\big)}. \tag{14.225}$$

The limit $\Lambda \mapsto 0$ is easily taken and gives rather simpler behaviour in the absence of a cosmological constant:

$$\rho(t) = \frac{1}{6\pi t(1+mt)}, \qquad H(t) = \frac{1+2mt}{3t(1+mt)}. \tag{14.226}$$

Antiparticle solutions can also be found, though these can have unusual properties. At large times the Hubble function tends to a constant value of $(\Lambda/3)^{1/2}$. This behaviour is typical of Λ cosmologies and leads to the surprising prediction that the universe will keep accelerating. The presence of a non-zero spin vector implies that these models break isotropy, but this fact is hidden from the line element, which remains isotropic. The spin direction is only seen by particles with non-zero spin, which interact directly with the torsion tensor.

14.6 Cylindrical systems

We now turn our attention to a different class of exact solutions — those exhibiting cylindrical symmetry. Such solutions can provide models for stringlike configurations, and some of the solutions are also appropriate for gravity in (2+1) dimensions. We first introduce cylindrical polar coordinates (t, ρ, ϕ, z), where

$$\rho = \big((x^1)^2 + (x^2)^2\big)^{1/2}, \qquad \tan(\phi) = \frac{x^2}{x^1} \tag{14.227}$$

and $x^\mu = \gamma^\mu \cdot x$. We use the symbol ρ for the cylindrical distance to avoid confusion with the radial coordinate r used throughout this chapter. When we come to describe the matter, the energy density is denoted ε in this section. The coordinate frame defined by cylindrical polar coordinate is

$$\begin{aligned} e_t &= \gamma_0, & e_\phi &= \rho(-\sin(\phi)\,\gamma_1 + \cos(\phi)\,\gamma_2),\\ e_\rho &= \cos(\phi)\,\gamma_1 + \sin(\phi)\,\gamma_2, & e_z &= \gamma_3, \end{aligned} \tag{14.228}$$

and we continue to write $\hat{\phi}$ for the unit vector e_ϕ/ρ. As a bivector basis we use the set $\{\sigma_\rho, \sigma_\phi, \sigma_3\}$, where

$$\sigma_\rho = e_\rho e_t, \qquad \sigma_\phi = \hat{\phi} e_t, \quad \sigma_3 = e_z e_t. \tag{14.229}$$

We are interested in stationary fields that exhibit cylindrical symmetry. For these we can write a general $\bar{\mathsf{h}}$ function as

$$\begin{aligned} \bar{\mathsf{h}}(e^t) &= f_1 e^t + \rho f_2 e^\phi, & \bar{\mathsf{h}}(e^\rho) &= g_1 e^\rho,\\ \bar{\mathsf{h}}(e^\phi) &= \rho h_1 e^\phi + h_2 e^t, & \bar{\mathsf{h}}(e^z) &= e^z, \end{aligned} \tag{14.230}$$

where all of the arbitrary functions depend on ρ only. A suitable ω field consistent with this $\bar{\mathsf{h}}$ field is given by

$$\begin{aligned}
\omega_t &= \omega(e_t) = -T\sigma_\rho + (K + h_2)I\sigma_3, \\
\omega_\rho &= \omega(e_\rho) = \bar{K}\sigma_\phi, \\
\omega_{\hat{\phi}} &= \omega(\hat{\phi}) = K\sigma_\rho + (h_1 - G)I\sigma_3, \\
\omega_z &= \omega(e_z) = 0.
\end{aligned} \tag{14.231}$$

Again, the new scalar functions appearing here (T, K, $\bar{K}$, G) are functions of ρ alone. Since all expressions involving L_z must vanish, there are only three non-vanishing commutation relations to construct. These are

$$\begin{aligned}
[L_\rho, L_t] &= TL_t + (K + \bar{K})L_{\hat{\phi}}, \\
[L_\rho, L_{\hat{\phi}}] &= -(K - \bar{K})L_t - GL_{\hat{\phi}}, \\
[L_t, L_{\hat{\phi}}] &= 0.
\end{aligned} \tag{14.232}$$

Since neither L_t nor $L_{\hat{\phi}}$ contains derivatives with respect to ρ, the bracket relations immediately yield

$$\begin{aligned}
L_\rho f_1 &= Tf_1 + (K + \bar{K})f_2, \\
L_\rho f_2 &= -Gf_2 - (K - \bar{K})f_1, \\
L_\rho h_1 &= -Gh_1 - (K - \bar{K})h_2, \\
L_\rho h_2 &= Th_2 + (K + \bar{K})h_1.
\end{aligned} \tag{14.233}$$

The cylindrical derivative L_ρ is given by $L_\rho = g_1(\rho)\partial_\rho$. We can always make the position gauge choice $g_1 = 1$, though this is not always the simplest gauge to work with.

The Riemann tensor takes the general form

$$\begin{aligned}
\mathcal{R}(\sigma_\rho) &= \alpha_1\sigma_\rho + \beta I\sigma_3, \\
\mathcal{R}(I\sigma_3) &= \alpha_2 I\sigma_3 - \beta\sigma_\rho, \\
\mathcal{R}(\sigma_\phi) &= \alpha_3\sigma_\phi,
\end{aligned} \tag{14.234}$$

where the scalar functions are defined by

$$\begin{aligned}
\alpha_1 &= -L_\rho T + T^2 - K(K + 2\bar{K}), \\
\alpha_2 &= L_\rho G + G^2 - K(K - 2\bar{K}), \\
\alpha_3 &= K^2 - GT, \\
\beta &= L_\rho K + G(K + \bar{K}) - T(K - \bar{K}).
\end{aligned} \tag{14.235}$$

The same functions appear in the Einstein tensor,

$$\begin{aligned}\mathcal{G}(e_t) &= -\alpha_2 e_t - \beta\hat{\phi},\\ \mathcal{G}(e_\rho) &= -\alpha_3 e_\rho,\\ \mathcal{G}(\hat{\phi}) &= -\alpha_1\hat{\phi} + \beta e_t,\\ \mathcal{G}(e_z) &= -(\alpha_1+\alpha_2+\alpha_3)e_z.\end{aligned} \tag{14.236}$$

It is a feature of gravity in $(2+1)$ dimensions that all of the information in the Riemann tensor is also contained in the Einstein tensor. That is, there is no Weyl tensor in three dimensions. It also turns out that no additional new information is obtained from the Bianchi identities, which are satisfied automatically from the equations we have already constructed.

The $\bar{\mathsf{h}}$ function of equation (14.230) contains a single rotational gauge freedom, which is the freedom to boost in the σ_ϕ plane. If we make the physical assumption that the matter energy-momentum tensor has a future-pointing timelike eigenvector, the gauge freedom can be used to set this eigenvector to the e_t direction. Once this is done all the rotational gauge freedom in the problem has been removed, and we are left with a complete set of field equations. These are

$$\begin{aligned}-L_\rho G - G^2 + K(K-2\bar{K}) &= 8\pi\varepsilon,\\ K^2 - GT &= 8\pi P_\rho,\\ -L_\rho T + T^2 - K(K+2\bar{K}) &= 8\pi P_\phi,\\ L_\rho K + G(K+\bar{K}) - T(K-\bar{K}) &= 0,\end{aligned} \tag{14.237}$$

where ε is the matter density, and P_ρ and P_ϕ are the radial and azimuthal pressures respectively. The coefficient of $\mathcal{G}(e_z)$ is determined algebraically by the other three coefficients, and the same must therefore be true of the matter energy-momentum tensor. It follows that the z-component of the Einstein equations contains no new information. Of course, if we were working in a genuine $(2+1)$ system, the e_z equation would not be present.

14.6.1 Vacuum solutions

In the vacuum region all of the scalars $\{\alpha_1, \alpha_2, \alpha_3, \beta\}$ are zero, so we are still free to perform an ρ-dependent boost in the σ_ϕ direction. This freedom can be employed to set $\bar{K}$ to zero. It is also useful in this region to work in a gauge where $g_1 = 1$. In this case the vacuum region is described by the simple pair of equations

$$\begin{aligned}\partial_\rho G + G^2 - GT &= 0,\\ \partial_\rho T - T^2 + GT &= 0,\end{aligned} \tag{14.238}$$

with K determined by $K^2 = GT$. On subtracting these equations and integrating we see that

$$G - T = 1/(\rho + \rho_0), \tag{14.239}$$

where ρ_0 is an arbitrary constant of integration. Similarly, adding the equations and integrating yields

$$G + T = c/(\rho + \rho_0), \tag{14.240}$$

where c is a second constant of integration.

The restriction that $GT = K^2 > 0$ means that $c^2 > 1$, and we can set

$$c = \pm \cosh(2\alpha). \tag{14.241}$$

There are two distinct vacuum configurations, depending on which sign is chosen for c. In either case, the constant α can be gauged to zero with a further constant boost in the σ_ϕ direction (which does not reintroduce a $\bar{K}$ term). The two vacuum sectors are therefore characterised by the solutions

$$\begin{aligned} &\text{type I:} && G = \frac{1}{\rho + \rho_0}, \quad T = K = \bar{K} = 0, \\ &\text{type II:} && T = -\frac{1}{\rho + \rho_0}, \quad G = K = \bar{K} = 0. \end{aligned} \tag{14.242}$$

All other vacuum solutions can be reached from this pair by ρ-dependent boosts in the σ_ϕ direction. No globally-defined gauge transformation exists between these solution classes. For both solutions the Riemann tensor vanishes, since there is no Weyl tensor for three-dimensional systems. It is therefore possible locally to gauge transform all of these fields to zero, but this is not possible globally. In this sense the solutions represent two distinct topological structures.

14.6.2 Physical properties of matter solutions

The key physical properties associated with matter solutions are the acceleration, vorticity, shear and angular momentum of the string. Given that we have chosen a gauge where the timelike eigenvector of the energy-momentum tensor is e_t, the acceleration vector w is defined by

$$w = e_t \cdot \mathcal{D}\, e_t = -T e_\rho. \tag{14.243}$$

This measures the extent to which particles comoving with the matter (with velocity e_t) depart from geodesic motion. The vorticity bivector ϖ is defined by

$$\varpi = \mathcal{D} \wedge e_t + w \wedge e_t = -(K - \bar{K}) I \sigma_3. \tag{14.244}$$

The definition ensures that ϖ satisfies $e_t \cdot \varpi = 0$. To define the shear tensor we require the linear function H that projects vectors into the 3-space orthogonal to e_t,

$$\mathsf{H}(a) = a - a\cdot e_t\, e_t. \tag{14.245}$$

In terms of this function the shear tensor $\sigma(a)$ is defined by

$$\begin{aligned}\sigma(a) &= \tfrac{1}{2}\big(\mathsf{H}(a)\cdot\mathcal{D}e_t + \mathsf{H}(\partial_b)(b\cdot\mathcal{D}e_t)\cdot a\big) - \tfrac{1}{3}\mathsf{H}(a)\mathcal{D}\cdot e_t \\ &= -\tfrac{1}{2}(K+\bar{K})(a\cdot e_\rho\,\hat{\phi} + a\cdot\hat{\phi}\,e_\rho).\end{aligned} \tag{14.246}$$

This is a symmetric, traceless linear function. We see that acceleration is controlled by T, the vorticity by $(K-\bar{K})$ and the shear by $(K+\bar{K})$. In the matter region all of these scalar quantities are physically measurable functions. The same is true of the fourth function, G, which can be determined from the radial pressure.

The remaining physical property of relevance is the angular momentum contained in the fields. The vector g_ϕ is a Killing vector for cylindrical solutions, so the vector $\mathcal{T}(g_\phi)$ is covariantly conserved. It follows that

$$\nabla\cdot\big(\mathsf{h}(\mathcal{T}(g_\phi))\det\,(\mathsf{h})^{-1}\big) = 0. \tag{14.247}$$

The total conserved angular momentum per unit length in the e_t frame is therefore given by the expression

$$J_S = \int_0^{\rho_s} d^2x\, g^t\cdot\mathcal{T}(g_\phi)\det\,(\mathsf{h})^{-1}, \tag{14.248}$$

where ρ_s is the string radius. In the $g_1 = 1$ gauge this expression evaluates to give

$$J_S = -2\pi\int_0^{\rho_s} d\rho\,(\varepsilon + P_\phi)f_1 f_2(f_1h_1 - f_2h_2)^{-2}, \tag{14.249}$$

which shows that a non-zero f_2 is required for angular momentum to be present.

14.6.3 Cosmic strings

Cosmic strings are an example of topological defects that can occur as a remnant of symmetry breaking processes in the early universe. They have zero radial and azimuthal pressures. It follows that there is a negative pressure along the length of the string — they are under *tension*. The energy-momentum tensor is

$$\mathcal{T}(a) = \tfrac{1}{2}\varepsilon(a - I\sigma_3\,a\,I\sigma_3). \tag{14.250}$$

From the Einstein equations we see that $\alpha_1 = \alpha_3 = \beta = 0$, and the Riemann tensor therefore has the compact form

$$\mathcal{R}(B) = 8\pi\varepsilon\langle BI\sigma_3\rangle I\sigma_3. \tag{14.251}$$

Tidal forces are only exerted in the $I\sigma_3$ plane and are controlled by the density.

The Einstein equations tell us that $T = K = \bar{K} = 0$, so all that remains is the single equation

$$L_\rho G + G^2 = -8\pi\varepsilon. \tag{14.252}$$

The full solution is then recovered by integrating the bracket equations (14.233). These imply that both f_1 and h_2 are constant. A global rotation can therefore be performed to transform to a gauge where $f_1 = 1$ and $h_2 = 0$. The remaining equations are

$$L_\rho h_1 = -Gh_1, \qquad L_\rho f_2 = -Gf_2. \tag{14.253}$$

It follows that $f_2 = \lambda h_1$, where λ is an arbitrary constant. But ρh_1 must tend to 1 as $\rho \mapsto 0$ so that $\bar{\mathsf{h}}(a)$ is well defined on the axis. It follows that h_1, and hence f_2, must diverge as ρ^{-1}. For f_2 this would imply that $\bar{\mathsf{h}}(e^t)$ is singular on the axis, which is not permitted. It follows that the constant λ must be zero, so the string has no angular momentum. This agrees with the fact that the shear and vorticity are both zero. Pressure is necessary for strings to have any angular momentum.

We have now restricted $\bar{\mathsf{h}}(a)$ to the simple form

$$\bar{\mathsf{h}}(a) = a + (g_1 - 1)a\cdot e_\rho\, e^\rho + (\rho h_1 - 1)a\cdot e_\phi\, e^\phi, \tag{14.254}$$

and the remaining equations are

$$L_\rho h_1 = -Gh_1, \qquad L_\rho G = -8\pi\varepsilon - G^2, \tag{14.255}$$

with $L_\rho = g_1\partial_\rho$. To complete the solution we must make a gauge choice for g_1. An obvious choice is to set $g_1 = 1$, so that ρ measures the proper radial distance from the string. A slightly simpler alternative is to choose a gauge such that $\bar{\mathsf{h}}(e^\phi) = e^\phi$. This requires that

$$h_1 = 1/\rho \tag{14.256}$$

and it follows that

$$G = g_1/\rho. \tag{14.257}$$

The equations now integrate to give

$${g_1}^2 = 1 - \int_0^\rho 16\pi s\varepsilon(s)\, ds, \tag{14.258}$$

where the constant of integration is chosen so that $\bar{\mathsf{h}}(a)$ is well defined on the axis. On defining

$$M(\rho) = \int_0^\rho 2\pi s\varepsilon(s)\, ds, \tag{14.259}$$

the solution can be summarised neatly by

$$\bar{\mathsf{h}}(a) = a + \big((1 - 8M(\rho))^{1/2} - 1\big)a\cdot e_\rho\, e^\rho. \tag{14.260}$$

The choice of density function is arbitrary, provided $8M(\rho) < 1$. In the vacuum region outside the string we have $T = K = \bar{K} = 0$, so the vacuum region is described by a solution in the gauge class of type I. This can be described in terms of a flat spacetime, with a wedge of spacetime removed and the edges identified. This topological picture of a string defect can be used to provide a qualitative understanding of many of the string's properties.

14.6.4 Rigidly rotating strings

The simplest models that include pressure are those for a two-dimensional ideal fluid, with $P_\rho = P_\phi = P$. The two natural physical models to consider are those where the fluid is vorticity-free ($\bar{K} = K$) or shear-free ($\bar{K} = -K$). The latter case corresponds to a rigidly rotating string, and is the situation we analyse here. The equations governing this setup are (in the $g_1 = 1$ radial gauge)

$$\begin{aligned} \partial_\rho K - 2KT &= 0, \\ \partial_\rho G + G^2 &= -8\pi\varepsilon + 3K^2, \\ \partial_\rho T - T^2 &= -8\pi P + K^2, \\ K^2 - GT &= 8\pi P. \end{aligned} \tag{14.261}$$

These can be solved once the density distribution has been specified. A choice of density that produces a straightforward solution is

$$8\pi\varepsilon = 3K^2 + \lambda^2, \tag{14.262}$$

where λ is an arbitrary positive constant. This ansatz ensures that the density is always positive. The equations for G and T can be solved immediately to give

$$G = \frac{\lambda\cos(\lambda\rho)}{\sin(\lambda\rho)}, \qquad T = \frac{\lambda\sin(\lambda\rho)}{\cos(\lambda\rho) + A}, \tag{14.263}$$

where A is a constant satisfying $A < -1$.

We next solve for K to obtain

$$K = \frac{B}{(A + \cos(\lambda\rho))^2}, \tag{14.264}$$

where B is a further constant. The density and pressure can now be recovered from equations (14.261). The boundary of the string occurs where the pressure vanishes, and this must be reached before $\rho = \pi/\lambda$. Finally, we return to equations (14.233) to find a suitable form for the $\bar{\mathsf{h}}$ function. First we see that f_1/h_2 is a constant, so that a gauge transformation can be performed to set $h_2 = 0$.

The remaining functions are easily found by integration:

$$
\begin{aligned}
f_1 &= \frac{1+A}{\cos(\lambda\rho)+A}, \\
h_1 &= \frac{\lambda}{\sin(\lambda\rho)}, \\
f_2 &= \frac{-B({f_1}^2-1)}{\lambda(A+1)\sin(\lambda\rho)}.
\end{aligned}
\tag{14.265}
$$

For f_1 the arbitrary time-scale factor has been used to set $f_1 = 1$ on the axis. It is simple to verify that this solution is well defined on the axis of the string. For completeness, the corresponding line element is

$$
\begin{aligned}
ds^2 = {}& \frac{\big(\cos(\lambda\rho)+A\big)^2}{(1+A)^2}\,dt^2 + \frac{2B}{\lambda^2(A+1)^3}\big(1-\cos(\lambda\rho)\big)\big(2A+1+\cos(\lambda\rho)\big)\,dt\,d\phi \\
& - \frac{\sin^2(\lambda\rho)}{\lambda^2}\left(1 - \frac{B^2\big(1-\cos(\lambda\rho)\big)^2\big(2A+1+\cos(\lambda\rho)\big)^2}{\lambda^2\sin^2(\lambda\rho)(1+A)^4\big(A+\cos(\lambda\rho)\big)^2}\right)d\phi^2 - d\rho^2 - dz^2.
\end{aligned}
\tag{14.266}
$$

The exterior vacuum fields can be found simply by returning to the vacuum equations, and solving these in the case where $K + \bar{K} = 0$. The general form of vacuum fields outside a rigidly rotating string is then given by

$$
\begin{aligned}
G &= \frac{-\alpha^2}{(\rho+\rho_0)\big((\rho+\rho_0)^2-\alpha^2\big)}, \\
T &= -\frac{\rho+\rho_0}{(\rho+\rho_0)^2-\alpha^2}, \\
K &= \frac{\alpha}{(\rho+\rho_0)^2-\alpha^2},
\end{aligned}
\tag{14.267}
$$

where ρ_0 and α are constants to be determined by the fields at the boundary. This solution falls into the second class of vacuum solutions, as defined by equation (14.242). The $\bar{\mathsf{h}}$ function is determined by

$$
\begin{aligned}
f_1 &= -(1+A)(\alpha/B)^{1/2}\big((\rho+\rho_0)^2-\alpha^2\big)^{-1/2}, \\
h_1 &= (\alpha/B)^{1/2}\lambda^2\,\frac{\big((\rho+\rho_0)^2-\alpha^2\big)^{1/2}}{(\rho+\rho_0)}, \\
f_2 &= \frac{\alpha}{f_1(\rho+\rho_0)}\big(f_1^2-1\big).
\end{aligned}
\tag{14.268}
$$

These fields have an unusual property. At large distances, f_1 falls off as ρ^{-1}, whereas f_2 tends to a constant value. Beyond the point where the magnitude of f_2 overtakes that of f_1, a closed circular path orbiting the string becomes *timelike*. This solution admits closed timelike curves, even out at infinity. Such solutions are often thought of as unphysical, due to the bizarre acausal effects they would allow. But there is nothing outrageous in the matter distribution used to generate the solution, and it is difficult to pin down a precise statement of what constitutes a 'physically acceptable' matter distribution.

14.7 Axially-symmetric systems

As a further application of the gauge theory treatment of gravity, we now turn to the equations governing a stationary axisymmetric system. Such fields are produced by rotating stars, galaxies and black holes, and as such are of considerable importance in astrophysics. The prototype axisymmetric configuration is described by the *Kerr solution*, which uniquely describes the fields produced by an uncharged rotating black hole. The more complicated problem of finding the fields outside a rotating massive object such as a star or planet has yet to be fully solved. Here we discuss two forms of the Kerr solution. The first continues the solution strategy adopted in the cylindrical setup, and can be generalised to include matter fields. The second form generalises the Newtonian gauge for the Schwarzschild solution, and has a number of significant features.

14.7.1 Intrinsic form of the axisymmetric equations

We employ a standard spherical-polar coordinate system to describe axisymmetric fields, and the notation is precisely as defined at the start of section 14.2. A suitable form of the $\bar{\mathsf{h}}$ function consistent with axial symmetry is

$$\begin{aligned}
\bar{\mathsf{h}}(e^t) &= f_1 e^t + f_4 e^\phi, \\
\bar{\mathsf{h}}(e^r) &= g_1 e^r + g_3 e^\theta, \\
\bar{\mathsf{h}}(e^\theta) &= i_1 e^\theta + i_3 e^r, \\
\bar{\mathsf{h}}(e^\phi) &= h_1 e^\phi + h_2 e^t,
\end{aligned} \tag{14.269}$$

where all of the variables $\{f_1, \dots, i_3\}$ are scalar functions of r and θ. The labelling convention for the $\{f_i, \dots, i_i\}$ is chosen to allow for a more general parameterisation appropriate for time-dependent systems. We have ignored the possibility of any coupling between the e^t and e^r, so strictly speaking are looking for the fields outside an extended source with no horizon present. On solving the vacuum field equations we will construct a form of the Kerr solution, which will turn out to be ill defined at the horizon. As with the Schwarzschild solution, the singular nature of the fields is a consequence of a bad gauge choice, rather than an intrinsic property of the fields. In section 14.7.3 we give a form of the Kerr solution which avoids this problem.

A suitably general form of ω function consistent with the $\bar{\mathsf{h}}$ field of equation (14.269) is given by

$$
\begin{aligned}
\omega(e_t) &= -(T+IJ)e_r e_t - (S+IK)\hat{\theta} e_t + h_2 I\sigma_3, \\
\omega(e_r) &= (S'+IK')e_r\hat{\theta} - i_3 e_r\hat{\theta}, \\
\omega(\hat{\theta}) &= (G'+IJ')e_r\hat{\theta} - (i_1/r)e_r\hat{\theta}, \\
\omega(\hat{\phi}) &= (H+IK)\hat{\theta}\hat{\phi} + (G+IJ)e_r\hat{\phi} + h_1/(r\sin(\theta))\, I\sigma_3.
\end{aligned}
\tag{14.270}
$$

The variables written in capitals are also functions of r and θ, except for the pseudoscalar I. The reason for the labelling scheme will become clearer when the final set of equations is derived. There are 40 independent scalar variables in gravity, so it is difficult to construct a labelling scheme that does not conflict with existing conventions somewhere. A significant feature of our scheme is that a complex structure naturally emerges, generated by the pseudoscalar I. It is a well-known feature of the Kerr solution that it is underpinned by a complex analytic structure. The origin of this lies in the natural complex structure of spacetime bivectors. Throughout this section we use *complex* to refer to a combination of scalar and pseudoscalar quantities.

The bracket structure defined by our choice of the ω function is

$$
\begin{aligned}
&[L_t, L_r] = -TL_t - (K+K')L_{\hat{\phi}}, && [L_r, L_{\hat{\theta}}] = -S'L_r - G'L_{\hat{\theta}}, \\
&[L_t, L_{\hat{\theta}}] = -SL_t + (J-J')L_{\hat{\phi}}, && [L_r, L_{\hat{\phi}}] = -(K-K')L_t - GL_{\hat{\phi}}, \\
&[L_t, L_{\hat{\phi}}] = 0, && [L_{\hat{\theta}}, L_{\hat{\phi}}] = (J+J')L_t - HL_{\hat{\phi}}.
\end{aligned}
\tag{14.271}
$$

The Riemann tensor generated by these fields is complicated and, rather than giving its full algebraic expression, it is simpler to consider the general form. This can be written as

$$
\begin{aligned}
&\mathcal{R}(\sigma_r) = \alpha_1\sigma_r + \beta_1\sigma_\theta, && \mathcal{R}(I\sigma_r) = \alpha_4 I\sigma_r + \beta_4 I\sigma_\theta, \\
&\mathcal{R}(\sigma_\theta) = \alpha_2\sigma_\theta + \beta_2\sigma_r, && \mathcal{R}(I\sigma_\theta) = \alpha_5 I\sigma_\theta + \beta_5 I\sigma_r, \\
&\mathcal{R}(\sigma_\phi) = \alpha_3\sigma_\phi, && \mathcal{R}(I\sigma_\phi) = \alpha_6 I\sigma_\phi,
\end{aligned}
\tag{14.272}
$$

where each of the α_i and β_i is a complex combination. If we now specialise to the case of vacuum solutions, so that the Riemann tensor is determined solely by the Weyl tensor, the duality relation $\mathcal{W}(IB) = I\mathcal{W}(B)$ immediately sets

$$
\alpha_1 = \alpha_4, \quad \alpha_2 = \alpha_5, \quad \alpha_3 = \alpha_6, \quad \beta_1 = \beta_4 \quad \beta_2 = \beta_5. \tag{14.273}
$$

In addition, for a vacuum solution $\mathcal{R}(B)$ must be symmetric and traceless. The most general form of tensor consistent with this requirement is

$$
\begin{aligned}
\mathcal{R}(\sigma_r) &= \alpha_1\sigma_r + \beta\sigma_\theta, \\
\mathcal{R}(\sigma_\theta) &= \alpha_2\sigma_\theta + \beta\sigma_r, \\
\mathcal{R}(\sigma_\phi) &= -(\alpha_1+\alpha_2)\sigma_\phi,
\end{aligned}
\tag{14.274}
$$

with α_i and β complex combinations.

Next we consider the rotational gauge freedom in our choice of axisymmetric fields. We are free to perform a rotation in the $I\sigma_\phi$ plane, and a boost in the σ_ϕ direction. These can be summarised in the single rotor R:

$$R = \exp(wI\sigma_\phi/2), \tag{14.275}$$

where the scalar + pseudoscalar quantity w is an arbitrary function of (r, θ). This gauge freedom can be employed to diagonalise the Riemann tensor by setting $\beta = 0$. This removes all of the gauge freedom present, and enables us to write

$$\mathcal{R}(\sigma_r) = \alpha_1\sigma_r, \quad \mathcal{R}(\sigma_\theta) = \alpha_2\sigma_\theta, \quad \mathcal{R}(\sigma_\phi) = -(\alpha_1+\alpha_2)\sigma_\phi. \tag{14.276}$$

The form of the Riemann tensor for the Schwarzschild solution is algebraically special, in that two of its eigenvalues are degenerate. This is referred to as having Petrov type D. There is no reason to expect the same to be true for axisymmetric fields, and the field outside a general rotating star is almost certainly not of type D. But it turns out that, if a horizon is present, the solution must be of type D. As we are interested here in deriving the Kerr solution, we therefore impose the additional condition that the Riemann tensor is degenerate, with the general algebraic form

$$\mathcal{R}(B) = \frac{\alpha}{2}(B + 3\sigma_r B\sigma_r), \tag{14.277}$$

with α a scalar + pseudoscalar quantity. This final restriction on the form of $\mathcal{R}(B)$ is *not* a gauge choice — it is a restriction on the form of solution we can construct.

Comparing the general form of equation (14.277) with the explicit Riemann tensor constructed from the ω field, we establish that

$$\alpha = (G+IJ)(T+IJ) + (S+IK)(H+IK). \tag{14.278}$$

The remaining identities reduce to a series of equations, an example of which is

$$\begin{aligned} L_r(G+IJ) =& (S'+IK'-S-IK)(H+IK) - I(K-K')(S+IK) \\ &- (G+T+IJ)(G+IJ). \end{aligned} \tag{14.279}$$

In all there are ten equations of this type. They all relate intrinsic derivatives of the variables in the ω field to quadratic combinations of the same variables. By forming suitable combinations of these equations we find that

$$L_r\alpha = -3\alpha(G+IJ), \qquad L_{\hat{\theta}}\alpha = -3\alpha(S+IK), \tag{14.280}$$

so the intrinsic derivatives of α are quite simple.

Next we must consider the Bianchi identities. These contain higher order consistency relations between the $\bar{\mathsf{h}}$ and ω fields. For the Schwarzschild and cylindrical cases these contained no new information, but this is not the case for the axisymmetric setup. If we consider the equation

$$\mathcal{D}\mathcal{R}(\sigma_r) - \partial_a \mathcal{R}(a \cdot \mathcal{D}\sigma_r) = 0 \tag{14.281}$$

we obtain the pair of equations

$$\begin{aligned} L_r \alpha &= -\frac{3}{2}\alpha(G + IJ + G' + IJ'), \\ L_{\hat{\theta}} \alpha &= -\frac{3}{2}\alpha(S + IK + S' + IK'), \end{aligned} \tag{14.282}$$

Comparing these with equation (14.280), we see that

$$G' + IJ' = G + IJ, \qquad S' + IK' = S + IK. \tag{14.283}$$

This simplification for type D fields explains our choice of notation of primed and unprimed variables.

With four of the variables now solved for, the remaining equations simplify to

$$\begin{aligned} L_r(G + IJ) &= -(G + IJ)^2 - T(G + IJ), \\ L_r(T + IJ) &= (S + IK)^2 - \big(2(G + IJ) - T\big)(T + IJ) \\ &\quad - 2S(H + IK), \\ L_r(S + IK) &= -IJ(S + IK) - 2IK(G + IJ), \\ L_r(H + IK) &= -(G + IJ)(S + IK) - G(H + IK) \end{aligned} \tag{14.284}$$

and

$$\begin{aligned} L_{\hat{\theta}}(S + IK) &= (S + IK)^2 + H(S + IK), \\ L_{\hat{\theta}}(H + IK) &= -(G + IJ)^2 + \big(2(S + IK) - H\big)(H + IK) \\ &\quad + 2G(T + IJ), \\ L_{\hat{\theta}}(G + IJ) &= IK(G + IJ) + 2IJ(S + IK), \\ L_{\hat{\theta}}(T + IJ) &= (G + IJ)(S + IK) + S(T + IJ). \end{aligned} \tag{14.285}$$

These equations are all consistent with the bracket structure, which now takes the form

$$[L_r, L_{\hat{\theta}}] = -SL_r - GL_{\hat{\theta}}. \tag{14.286}$$

Our set of equations is now complete. We have explicit forms for the intrinsic derivatives of all of our variables; these are all consistent with the bracket structure, and the full Bianchi identities are all satisfied. We have achieved the first main goal of the intrinsic method.

14.7.2 The Kerr solution

The vacuum equations summarised in equations (14.284) and (14.285) display a number of remarkable features. They are naturally complex, with the spacetime pseudoscalar as the unit imaginary, and there is a clear symmetry between the r and $\hat{\theta}$ equations. We now demonstrate that, subject to certain boundary conditions, these equations admit a unique, two-parameter family of solutions. This is the Kerr solution. The proof is constructive, but it is slightly involved and we will skip some of the details.

The first step in solving a set of intrinsic equations is the identification of suitable integrating factors. To find the first of these consider the function

$$Z = Z_0 \alpha^{-1/3}, \tag{14.287}$$

where Z_0 is an arbitrary complex constant. The function Z satisfies

$$L_r Z = (G + IJ)Z, \qquad L_{\hat{\theta}} Z = -(S + IK)Z. \tag{14.288}$$

On separating Z into modulus X and argument χ,

$$Z = X \mathrm{e}^{I\chi} \tag{14.289}$$

we find that

$$L_r X = GX, \qquad L_{\hat{\theta}} X = -SX. \tag{14.290}$$

It follows that X acts as an integrating factor for G and S. But if we recall the bracket of equation (14.286), we see that

$$[XL_r, XL_{\hat{\theta}}] = 0. \tag{14.291}$$

We have therefore constructed a pair of commuting derivations. This is sufficient to ensure that we can fix our displacement gauge freedom by setting $g_3 = i_3 = 0$. With this done, we can then write

$$XL_r = g(r)\partial_r, \quad XL_{\hat{\theta}} = i(\theta)\partial_\theta, \tag{14.292}$$

where $g(r)$ and $i(\theta)$ are arbitrary functions that we can choose with further gauge fixing.

More generally, if a pair of variables A and B satisfy the equation

$$L_{\hat{\theta}} A - L_r B = GB + SA \tag{14.293}$$

then an integrating factor C exists defined (up to an arbitrary magnitude) by

$$L_r C = AC, \qquad L_{\hat{\theta}} C = BC. \tag{14.294}$$

One such pair is T and $-H$. For these we define the integrating factor F, satisfying

$$L_r F = TF, \qquad L_{\hat{\theta}} F = -HF. \tag{14.295}$$

With the integrating factors X, Z and F at our disposal, we can considerably simplify our equations for $G+IJ$ and $S+IK$ to obtain

$$\begin{aligned} L_r\big(FZ(G+IJ)\big) &= 0, & L_{\hat\theta}\big(XZ(G+IJ)\big) &= -2XZ(SG+JK), \\ L_{\hat\theta}\big(FZ(S+IK)\big) &= 0, & L_r\big(XZ(S+IK)\big) &= 2XZ(SG+JK). \end{aligned} \tag{14.296}$$

These equations focus attention on the quantity $SG+JK$. On forming the derivatives of this quantity we see that

$$L_r\big(XF(SG+JK)\big) = L_{\hat\theta}\big(XF(SG+JK)\big) = 0, \tag{14.297}$$

and it follows that $XF(SG+JK)$ is a constant. For the Schwarzschild solution this constant is zero. We therefore expect that this term should also vanish for a rotating source since, at large distances, the fields should tend to the Schwarzschild case. It turns out that one can construct solutions with $XF(SG+JK) \neq 0$, but these are appropriate for an infinite disc of matter and not a localised source. As we are looking for the fields outside a localised rotating source, we can set

$$SG+JK = 0. \tag{14.298}$$

It follows that

$$XFZ^2(G+IJ)(S+IK) = C_1, \tag{14.299}$$

where C_1 is an arbitrary complex constant.

Remarkably, we are now close to a complete solution to the problem. Equation (14.296) tells us that we can set

$$FZ(G+IJ) = W(\theta), \qquad FZ(S+IK) = U(r), \tag{14.300}$$

where U and W are complex functions of r and θ respectively. If we now form

$$\frac{W(\theta)}{U(r)} = \frac{G+IJ}{S+IK} = I\frac{SJ-GK}{S^2+K^2}, \tag{14.301}$$

we see that the result is a pure imaginary quantity. It follows that W and U are $\pi/2$ out of phase and, since U and W are separately functions of r and θ, their phases must be constant. Next we construct the derivatives of Z to obtain

$$XL_rZ = g(r)\partial_r Z = XZ(G+IJ) = C_1/U(r) \tag{14.302}$$

and

$$XL_{\hat\theta}Z = i(\theta)\partial_\theta Z = XZ(S+IK) = C_1/W(\theta). \tag{14.303}$$

It follows that Z must be the sum of a function of r and a function of θ. Furthermore, these functions must also have constant phases, $\pi/2$ apart. Since the overall phase of Z is arbitrary (Z was defined up to an arbitrary complex scale factor), we can write

$$Z = R(r) + I\Psi(\theta), \tag{14.304}$$

where $R(r)$ and $\Psi(\theta)$ are real functions. These satisfy the equation

$$XL_r\big(XL_r \ln(Z)\big) + XL_{\hat{\theta}}\big(XL_{\hat{\theta}} \ln(Z)\big) = \alpha = \frac{Z_0^3}{Z^3}, \tag{14.305}$$

which is to be solved for R and Ψ.

There is considerable gauge freedom in equation (14.305), since we are free to choose the functions $g(r)$ and $i(\theta)$. The most convenient choice of gauge is to set

$$Z = r - Ia\cos(\theta). \tag{14.306}$$

The remaining functions are then found by integration. The end result, after a series of further gauge choices, is the Kerr solution in the form

$$\begin{aligned}
\bar{\mathsf{h}}(e^t) &= g^t = \frac{r^2+a^2}{\rho\Delta^{1/2}}e^t + \frac{ar\sin^2(\theta)}{\rho}e^\phi, \\
\bar{\mathsf{h}}(e^r) &= g^r = \frac{\Delta^{1/2}}{\rho}e^r, \\
\bar{\mathsf{h}}(e^\theta) &= g^\theta = \frac{r}{\rho}e^\theta, \\
\bar{\mathsf{h}}(e^\phi) &= g^\phi = \frac{r}{\rho}e^\phi + \frac{a}{\rho\Delta^{1/2}}e^t,
\end{aligned} \tag{14.307}$$

where

$$\rho^2 = X^2 = r^2 + a^2\cos^2(\theta) \tag{14.308}$$

and

$$\Delta = r^2 - 2Mr + a^2. \tag{14.309}$$

The mass is given by M, and the angular momentum by aMc. The quantity a is the angular momentum per unit mass, and has dimensions of distance. The limit $a \mapsto 0$ recovers the Schwarzschild solution in the form appropriate for the exterior of a non-rotating star. The reciprocal vectors are

$$\begin{aligned}
g_t &= \frac{\Delta^{1/2}}{\rho}e_t - \frac{a}{r\rho}e_\phi, \\
g_r &= \frac{\rho}{\Delta^{1/2}}e_r, \\
g_\theta &= \frac{\rho}{r}e_\theta, \\
g_\phi &= \frac{r^2+a^2}{r\rho}e_\phi - \frac{a\Delta^{1/2}\sin^2(\theta)}{\rho}e_t.
\end{aligned} \tag{14.310}$$

The variables controlling the ω field are given by

$$
\begin{aligned}
G + IJ &= \frac{\Delta^{1/2}}{\rho\big(r - Ia\cos(\theta)\big)}, \\
S + IK &= \frac{-Ia\sin(\theta)}{\rho\big(r - Ia\cos(\theta)\big)}, \\
T - G &= -\frac{r-M}{\rho\Delta^{1/2}}, \\
H - S &= \frac{\cos(\theta)}{\rho\sin(\theta)}.
\end{aligned}
\tag{14.311}
$$

The equation for T shows that a horizon exists where $\Delta = 0$. The fact that the solution is singular there is a reflection of our choice of time coordinate. This measures the time for observers at a constant distance from the source. Such observers cannot exist inside the horizon, and the solution breaks down there. As with the Schwarzschild system, the resolution of this problem is to express the fields in terms of a different time coordinate.

The Riemann tensor for the Kerr solution can now be written in the compact form

$$
\mathcal{R}(B) = -\frac{M}{2\big(r - Ia\cos(\theta)\big)^3}(B + 3\sigma_r B\sigma_r). \tag{14.312}
$$

This is obtained from the Schwarzschild solution by simply replacing r by the scalar + pseudoscalar combination $r - Ia\cos(\theta)$. Precisely such a replacement can be used to generate the Kerr solution using a 'complex coordinate transformation' in the Newman–Penrose formalism. This transformation does produce the Kerr solution, but there is no *a priori* reason to expect that such a transformation applied to a vacuum solution will generate a new vacuum solution. Our extremely compact form of the Riemann tensor for the Kerr solution is a significant advantage of the gauge theory approach to gravitation advocated in this book. The comparison with the standard tensor formulation of general relativity is dramatic — most textbooks devote nearly a page to listing all of the components of the Riemann tensor, if they list them at all.

14.7.3 A Newtonian gauge for the Kerr solution

The form of the Kerr solution developed in the preceding section gives rise to a metric that expresses the geometry in terms of Boyer–Lindquist coordinates. Such a form is only appropriate for the region outside an extended object. If a horizon has formed we must find an alternative gauge choice which covers the horizon smoothly. From our discussion of the Schwarzschild solution, we would like to find an analogue of the Newtonian gauge appropriate for rotating black holes. Such a gauge does exist, though it is not straightforwardly obtained from the Boyer–Lindquist setup.

The first step in expressing the Kerr solution in a Newtonian gauge is the introduction of spheroidal coordinates $(\bar{r}, \bar{\theta}, \phi)$, as described in section 6.2.2. The spheroidal coordinates are related to their spherical counterparts (r, θ, ϕ) as follows:

$$\begin{aligned} (\bar{r}^2 + a^2)^{1/2} \sin(\bar{\theta}) &= r \sin(\theta), \\ \bar{r} \cos(\bar{\theta}) &= r \cos(\theta). \end{aligned} \tag{14.313}$$

The scalar parameter a is the same as that controlling the angular momentum. In the limit $a \mapsto 0$, the barred coordinates reduce to their unbarred spherical-polar equivalents. Surfaces of constant $\bar{r}$ are ellipses in flat space, though a statement such as this relates to the properties of the coordinate system, and not necessarily to physically measurable features. It is convenient to introduce the hyperbolic coordinate u, defined by

$$a \sinh(u) = \bar{r}. \tag{14.314}$$

The coordinate frame vectors are given by

$$\begin{aligned} e_{\bar{r}} &= \tanh(u) \sin(\bar{\theta}) \big(\cos(\phi)\, \gamma_1 + \sin(\phi)\, \gamma_2\big) + \cos(\bar{\theta})\, \gamma_3, \\ e_{\bar{\theta}} &= a \cosh(u) \cos(\bar{\theta}) \big(\cos(\phi)\, \gamma_1 + \sin(\phi)\, \gamma_2\big) - a \sinh(u) \sin(\bar{\theta})\, \gamma_3 \end{aligned} \tag{14.315}$$

with e_ϕ unchanged from its spherical definition. We also define the unit vectors

$$\hat{e}_{\bar{r}} = \frac{a \cosh(u)}{\bar{\rho}} e_{\bar{r}}, \qquad \hat{e}_{\bar{\theta}} = \frac{1}{\bar{\rho}} e_{\bar{\theta}}, \tag{14.316}$$

where $\bar{\rho}$ is defined by

$$\bar{\rho}^2 = a^2 \sinh^2(u) + a^2 \cos^2(\bar{\theta}) = \bar{r}^2 + a^2 \cos^2(\bar{\theta}). \tag{14.317}$$

The unit frame vectors satisfy

$$e_t \hat{e}_{\bar{r}} \hat{e}_{\bar{\theta}} \hat{\phi} = I. \tag{14.318}$$

The Newtonian gauge form of the Schwarzschild solution, defined in equation (14.65), contains the unit vectors e_t and e_r. The generalisation of this function to the Kerr solution is given by

$$\bar{\mathsf{h}}(n) = n - \left(\frac{2M\bar{r}}{\bar{r}^2 + a^2}\right)^{1/2} n \cdot \hat{e}_{\bar{r}}\, v, \tag{14.319}$$

where the vector argument is denoted by n to avoid confusion with the scalar parameter a. The timelike velocity vector v is defined by

$$v = \cosh(\beta)\, e_t + \sinh(\beta)\, \hat{\phi} \tag{14.320}$$

where

$$\tanh(\beta) = \frac{\sin(\bar{\theta})}{\cosh(u)} = \frac{ar\sin(\theta)}{\bar{r}^2 + a^2}. \tag{14.321}$$

It follows that

$$\cosh(\beta) = \frac{a\cosh(u)}{\bar{\rho}}, \qquad \sinh(\beta) = \frac{a\sin(\bar{\theta})}{\bar{\rho}}. \tag{14.322}$$

Comparison with equation (14.65) shows how the various terms are generalised in moving from the Schwarzschild to the Kerr solution.

The $\omega(a)$ function generated by equation (14.319) has

$$\begin{aligned}
\omega(e_t) &= 0, \\
\omega(\hat{e}_{\bar{r}}) &= -\frac{M}{\alpha(\bar{r} - Ia\cos(\bar{\theta}))^2}\,\hat{e}_{\bar{r}}\wedge v, \\
\omega(\hat{e}_{\bar{\theta}}) &= \frac{\alpha}{\bar{r} - Ia\cos(\bar{\theta})}\,\hat{e}_{\bar{\theta}}\wedge v, \\
\omega(\hat{\phi}) &= \frac{\alpha}{\cosh(\beta)(\bar{r} - Ia\cos(\bar{\theta}))}\,\sigma_\phi,
\end{aligned} \tag{14.323}$$

where

$$\alpha = -\frac{(2M\bar{r})^{1/2}}{\bar{\rho}}. \tag{14.324}$$

The terms in the ω function also neatly generalise their counterparts in the Schwarzschild solution. In particular, the fact that $\omega(e_t)$ vanishes implies that e_t satisfies the geodesic equation. The trajectories defined by this velocity define a family of observers whose proper time is given by t.

The remaining covariant object to construct is the Riemann tensor. If we define the unit bivector

$$\hat{N} = \hat{e}_{\bar{r}}\wedge v, \tag{14.325}$$

then the Riemann tensor takes on the simple form

$$\mathcal{R}(B) = -\frac{M}{2\big(\bar{r} - Ia\cos(\bar{\theta})\big)^3}(B + 3\hat{N}B\hat{N}). \tag{14.326}$$

This is obtained from the form of equation (14.312) by a displacement (taking the unbarred to the barred coordinates) and a boost from e_t to v. Both are gauge transformations, so the intrinsic information in equations (14.312) and (14.326) is precisely the same. The same transformations are involved in taking the $\bar{\mathsf{h}}(a)$ function from the form of equation (14.307) to that of equation (14.319). In addition, further (singular) transformations are also required to convert t to the time measured by a set of infalling observers with covariant velocity e_t.

14.7.4 Geodesics and the horizon

The $\bar{\mathsf{h}}$ and ω fields for the Newtonian form of the Kerr solution are well defined over all spacetime, down to the ring $\bar{r} = \cos(\bar{\theta}) = 0$. There are no problems with motion through the horizon, and infalling observers reach the central singularity in a finite coordinate time. This is because the coordinate t now measures the proper time for a family of free-falling observers with covariant velocity e_t. The trajectories defined by this velocity have

$$x' = \mathsf{h}(e_t) = e_t - \alpha \hat{e}_{\bar{r}} = e_t - \dot{\bar{r}} e_{\bar{r}}. \tag{14.327}$$

This defines a family of observers all infalling along directions with constant $\bar{\theta}$ and ϕ, and with infall velocity

$$\dot{\bar{r}} = \left(\frac{2M\bar{r}}{\bar{r}^2 + a^2} \right)^{1/2}. \tag{14.328}$$

This family neatly generalises the observers in radial free fall from rest at infinity employed in the Schwarzschild solution. As in the spherical case, many physical phenomena are simplest to interpret when expressed in terms of observers with covariant velocity e_t. A curious feature of these observers is that they appear to 'slow down' as the singularity is approached, though they do reach $\bar{r} = 0$ in a finite proper time.

The next task is to locate the horizon in our new form of the Kerr solution. A horizon marks the boundary between regions where one cannot signal to the other. This occurs where it is no longer possible to send null photons outwards. If k denotes the covariant photon velocity, with $k^2 = 0$, a horizon will occur when it is no longer possible to satisfy

$$\hat{e}_{\bar{r}} \cdot \mathsf{h}(k) < 0. \tag{14.329}$$

The left-hand side of this inequality can also be written as

$$\bar{\mathsf{h}}(\hat{e}_{\bar{r}}) \cdot k = \left(\hat{e}_{\bar{r}} + \left(\frac{2M\bar{r}}{\bar{r}^2 + a^2} \right)^{1/2} v \right) \cdot k. \tag{14.330}$$

It is not possible for two future-pointing null vectors to have an inner product less than 0, so the horizon occurs at

$$\frac{2M\bar{r}}{\bar{r}^2 + a^2} = 1. \tag{14.331}$$

This defines a quadratic equation, with two solutions when $a < M$, one when $a = M$ and no solutions for $a > M$. In the case where $a < M$, the outer horizon defines an event horizon. Photons can cross this on an inward trajectory, but no photons can escape. The inner horizon is slightly different. On the inside of the inner horizon it is possible for photons to travel outwards, but they cannot cross the horizon. Instead, they pile up just inside the boundary, forming an unstable *Cauchy horizon*.

Instead of considering observers attempting to exit to infinity, suppose instead that we look for observers at rest with respect to the background $(\bar{r}, \bar{\theta}, \phi)$ coordinates. Such observers can be constructed from observations of distant stars, for example. These observers have covariant velocity

$$\mathsf{h}^{-1}(\dot{x}) = \dot{t}\mathsf{h}^{-1}(e_t) = \dot{t}\left(e_t + \frac{2M\bar{r}}{\bar{r}^2 + a^2}\cosh(\beta)\,\hat{\phi}\right), \tag{14.332}$$

and the condition that this is a unit timelike vector forces

$$\dot{t}^2\left(1 - \frac{2M\bar{r}}{\bar{r}^2 + a^2\cos^2(\bar{\theta})}\right) = 1. \tag{14.333}$$

The surface within which it is not possible to remain at rest is called the *ergosphere*. For non-rotating black holes the horizon and ergosphere coincide. But for rotating black holes the ergosphere is defined by

$$\bar{r}^2 + a^2\cos^2(\bar{\theta}) - 2M\bar{r} = 0. \tag{14.334}$$

This surface lies outside the horizon, and touches the horizon at the poles. In the intervening region it is impossible to remain at rest, but it is still possible to escape. One can think of this in terms of the angular momentum of the hole dragging observers around with it.

To gain some further insight into the properties of the Kerr solution, consider circular orbits in the equatorial plane ($\bar{\theta} = \pi/2$). For these we have

$$\big(\mathsf{h}^{-1}(\dot{x})\big)^2 = \dot{t}^2 - (\bar{r}^2 + a^2)\dot{\phi}^2 - \frac{2M}{\bar{r}}(\dot{t} - a\dot{\phi})^2 = 1. \tag{14.335}$$

The $\bar{r}$ derivative of this expression must vanish for a circular orbit, which tells us that

$$\bar{r}^3 = M\left(\frac{\dot{t}}{\dot{\phi}} - a\right)^2. \tag{14.336}$$

If we let Ω denote the angular momentum measured by our set of preferred infalling observers (which are at rest at infinity), we have

$$\Omega = \frac{\dot{\phi}}{\dot{t}}. \tag{14.337}$$

It follows that, for circular orbits,

$$\Omega = \frac{M^{1/2}}{aM^{1/2} \pm \bar{r}^{3/2}}. \tag{14.338}$$

For a given distance, there are two possible values of the angular velocity for circular orbits. The larger value of Ω is for a particle corotating with the black hole, and the smaller for a counterrotating orbit. Again, this effect can be understood in terms of the black hole dragging matter around with it. The larger angular velocity for corotating orbits means it is possible to form stable orbits much closer to the event horizon than for the Schwarzschild case.

14.7.5 The Dirac equation in a Kerr background

As a final illustration of the utility of the Newtonian gauge form of the Kerr solution, we return to the Dirac equation. We first form

$$\partial_b \omega(b) = \frac{M}{\alpha \bar{\rho}^2} v - \left(\frac{2M\bar{r}}{\bar{r}^2 + a^2}\right)^{1/2} e_t \frac{1}{\bar{r} - Ia\cos(\bar{\theta})}. \tag{14.339}$$

The Dirac equation in the Newtonian gauge can therefore be written

$$\nabla\psi - (2M\bar{r})^{1/2}\left(\frac{v}{\bar{\rho}}\frac{\partial}{\partial \bar{r}}\psi + \frac{v}{4\bar{r}\bar{\rho}}\psi + e_t \frac{1}{2(\bar{r}^2 + a^2)^{1/2}\big(\bar{r} - Ia\cos(\bar{\theta})\big)}\psi\right) = -m\psi I\gamma_3. \tag{14.340}$$

If we again multiply through by e_t, we arrive at an interaction Hamiltonian of the form

$$\hat{H}_K\psi = \frac{i(2M)^{1/2}}{\bar{\rho}^2}\bigg((\bar{r}^3 + a^2\bar{r})^{1/4}\frac{\partial}{\partial \bar{r}}\big((\bar{r}^3 + a^2\bar{r})^{1/4}\psi\big) - a\cos(\bar{\theta})\,\bar{r}^{1/4}\,\sigma_\phi \frac{\partial}{\partial \bar{r}}\big(\bar{r}^{1/4}\psi\big) + \frac{a\bar{r}^{1/2}\cos(\bar{\theta})}{2(\bar{r}^2 + a^2)^{1/2}} I\psi\bigg), \tag{14.341}$$

where we continue to use i for the quantum imaginary. This Hamiltonian is (almost) Hermitian when integrated over flat three-dimensional space, because the measure in oblate spheroidal coordinates is

$$d^3x = \bar{\rho}^2 \sin(\bar{\theta})\, d\bar{r}\, d\bar{\theta}\, d\phi. \tag{14.342}$$

Our form of the Kerr solution therefore does generalise the many attractive features of the Newtonian gauge for the Schwarzschild solution. As in the Schwarzschild case, the Hamiltonian is not self-adjoint when acting on normalised wavefunctions. For the Kerr case a boundary term arises at $\bar{r} = 0$, which now defines a disc of radius a.

The Dirac equation (14.340) is separable in spheroidal coordinates, though the details of this separation are quite complicated. One problem is that the angular separation constant depends on the energy. This makes scattering calculations far more difficult than in the spherical case, as the separation constant must be recalculated for each energy. A considerable amount of work remains to be done in extending the detailed understanding of quantum theory in a Schwarzschild background to the Kerr case.

14.8 Notes

Many of the applications discussed in this chapter are covered in greater detail in the papers by Doran, Lasenby, Gull and coworkers. The solution method described in this chapter was first proposed in the paper 'Gravity, gauge theories and geometric algebra' by Lasenby, Doran & Gull (1998). This method should be compared with the spin coefficient formalism of Newman & Penrose (1962). The advantages of the Newtonian gauge for spherically-symmetric systems have been promoted by a handful of authors, most notably in the papers by Gautreau (1984), Gautreau & Cohen (1995), and by Martel & Poisson (2001).

The problem of the electromagnetic fields created by a point charge at rest outside a Schwarzschild black hole was first tackled by Copson (1928), who obtained a solution that was valid locally in the vicinity of the charge, but contained an additional pole at the origin. Linet (1976) modified Copson's solution by removing the singularity at the origin to obtain the potential described in section 14.3.5. Similar plots to those presented in section 14.3.5 were first obtained by Hanni & Ruffini (1973), though these authors did not extend their plots through the horizon. A popular means of interpreting these plots in terms of effects entirely around the horizon is advanced in *The Membrane Paradigm* by Thorne, Price & Macdonald (1986). We believe that a better understanding is gained by considering the global properties of fields, both inside and outside the horizon.

Scattering and absorption processes by black holes have been widely discussed by many authors. Summaries of this work can be found in the books by Futterman, Handler & Matzner (1988) and Chandrasekhar (1983), or the article by Andersson and Jensen (2000). The first attempt at a quantum calculation of the scattering cross section was by Collins, Delbourgo & Williams (1973), though their derivation did not employ a consistent perturbation scheme. The calculation described in this chapter was first published in the paper 'Perturbation theory calculation of the black hole elastic scattering cross section' by Doran and Lasenby (2002). Classical and quantum absorption processes are discussed in detail by Sanchez (1977, 1978) and Unruh (1976).

Cylindrical systems are discussed by Deser, Jackiw & 't Hooft (1984) and Jensen & Soleng (1992). The properties of cosmic strings are described in *Cosmic Strings and Other Topological Defects* by Vilenkin & Shellard (1994). The solutions described in this chapter were developed in the paper 'Physics of rotating cylindrical strings' by Doran, Lasenby & Gull (1996). The form of the fields outside a rotating black hole was first discovered by Kerr (1963), and has been widely discussed since. A fairly complete summary of this work in contained in Chandrasekhar's *The Mathematical Theory of Black Holes* (1983). The complex coordinate transformation trick for deriving the Kerr solution was discovered by Newman & Janis (1965), and later explained by Schiffer et al.(1973). The uniqueness theorem for black holes was developed by Carter (1971) and Robinson (1975). The analogue of the Newtonian gauge for the Kerr solution was discovered by Doran (2000).

The applications of the gauge theory approach to gravity discussed in this chapter have concentrated on the simplest Einstein–Cartan theory. Modern developments in quantum gravity have suggested a number of modifications to this theory. Two of the most common ideas include the introduction of local scale invariance, and the inclusion of higher order terms in the Lagrangian. The geometric algebra gauge theory approach is equally applicable in these settings. Some preliminary work on this subject is described by Lewis, Doran & Lasenby (2000). This field is developing rapidly, driven in part by developments in inflationary theory and observations of the cosmic microwave background. These observations could well revolutionise our understanding of gravitation in future years.

14.9 Exercises

14.1 Spherical symmetry of the h function can be imposed by demanding that

$$R\bar{\mathsf{h}}_{x'}(\tilde{R}aR)\tilde{R} = \bar{\mathsf{h}}(a),$$

where R is a constant spatial rotor ($Re_t\tilde{R} = e_t$), and $x' = \tilde{R}xR$. Prove that this symmetry implies that the $\{e^r, e^t\}$ and $\{e^\theta, e^\phi\}$ pairs decouple from each other. Show further that we must have

$$\bar{\mathsf{h}}(\hat{\theta}) = \alpha\hat{\theta} + \beta\hat{\phi},$$
$$\bar{\mathsf{h}}(\hat{\phi}) = \alpha\hat{\phi} - \beta\hat{\theta},$$

and explain why we can always set $\beta = 0$ with a suitable gauge choice.

14.2 The energy-momentum tensor for an ideal fluid is

$$\mathcal{T}(a) = (\rho + p)a\cdot vv - pa.$$

Show that covariant conservation of the energy-momentum tensor results in the pair of equations

$$\mathcal{D}\cdot(\rho v) + p\mathcal{D}\cdot v = 0,$$
$$(\rho + p)(v\cdot\mathcal{D}v)\wedge v - (\mathcal{D}p)\wedge v = 0.$$

Give a physical interpretation of these equations.

14.3 The Schwarzschild line element is defined by

$$ds^2 = \Big(1 - \frac{2M}{r}\Big)dt^2 - \frac{r}{r - 2M}dr^2 - r^2\, d\theta^2 - r^2\sin^2(\theta)\, d\phi^2.$$

Find the equation for the free-fall time as measured by radially-infalling observers, starting from rest at infinity. Express the line element in terms of this new time coordinate to obtain the Painlevé-Gullstrand form

$$ds^2 = dt^2 - \left(dr + \left(\frac{2M}{r}\right)^{1/2} dt\right)^2 - r^2(d\theta^2 + \sin^2(\theta)\, d\phi^2).$$

14.4 Prove that the total absorption cross section for a spherically-symmetric black hole of mass M is given by

$$\sigma_{abs} = \frac{\pi M^2}{2u^4}\big(8u^4 + 20u^2 - 1 + (1+8u^2)^{3/2}\big)$$

where u is the incident velocity.

14.5 The covariant electromagnetic field generated by a charge at rest on the z axis outside a Schwarzschild black hole is defined by

$$\mathcal{F} = -\frac{\partial V}{\partial r} e_r e_t - \frac{1}{r-2M}\frac{\partial V}{\partial \theta}\,\hat{\theta}\big(e_t + \sqrt{2M/r}e_r\big),$$

where

$$V(r,\theta) = \frac{q}{ar}\frac{(r-M)(a-M) - M^2\cos^2(\theta)}{D} + \frac{qM}{ar}$$

and

$$D = \big(r(r-2M) + (a-M)^2 - 2(r-M)(a-M)\cos(\theta) + M^2\cos^2(\theta)\big)^{1/2}.$$

Prove that $\mathcal{F}$ is finite and continuous at the horizon.

14.6 In calculating the scattering cross section from a black hole we need to compute the integral

$$I_1 = \int \frac{d^3k}{(2\pi)^3}\frac{\boldsymbol{k}^2 - \boldsymbol{p}^2}{|\boldsymbol{p}_f - \boldsymbol{k}|^{7/2}|\boldsymbol{k} - \boldsymbol{p}_i|^{7/2}}(\hat{k} + m).$$

Evaluate this integral by first displacing the origin in $\boldsymbol{k}$-space by the amount $(\boldsymbol{p}_f + \boldsymbol{p}_i)/2$, and then introducing spheroidal coordinates

$$\begin{aligned}
k_1 &= \alpha\sinh(u)\,\sin(v)\,\cos(\phi),\\
k_2 &= \alpha\sinh(u)\,\sin(v)\,\sin(\phi),\\
k_2 &= \alpha\cosh(u)\,\cos(v),
\end{aligned}$$

where $0 \le u < \infty$, $0 \le v \le \pi$, $0 \le \phi < 2\pi$ and $\alpha = |\boldsymbol{q}|/2$.

14.7 The Kerr–Schild form of the Schwarzschild solution is defined by

$$\bar{\mathsf{h}}(a) = a + \frac{M}{r} a\cdot e_-\, e_-, \qquad e_- = e_t - e_r.$$

Construct the Dirac equation in this gauge, and find the interaction vertex factor in momentum space. Calculate the differential scattering cross section for a fermion in this gauge, and verify that it is the same as found in the Newtonian gauge.

14.8 Prove that det (h) is constant for spherically-symmetric vacuum gravitational fields.

14.9 For a particle in a circular orbit around a Schwarzschild black hole, prove that the non-relativistic binding energy (as defined by the effective potential) is given by $(G = c = 1)$

$$E_b = -\frac{M}{2r}\frac{r-4M}{r-3M}.$$

14.10 Derive the full set of time-dependent radial equations with the cosmological constant Λ included.

14.11 A spherically-symmetric distribution of dust is released from rest, with the initial density distribution chosen so that streamlines do not cross. Prove that a singularity forms at the origin after a time

$$t_f = \left(\frac{3\pi}{32\rho_0}\right)^{1/2},$$

where ρ_0 is the central density.

14.12 Solve the Dirac equation in a cosmological background with $k \neq 0$. Is the Dirac field homogeneous? Can you construct self-consistent solutions to this system of equations?

14.13 Construct a matched set of interior and exterior gravitational fields around a rigidly-rotating cylindrical string. Do closed timelike curves exist in this geometry?

14.14 Verify that the Kerr solution defined by equation (14.307) satisfies the vacuum field equations.

14.15 The Riemann tensor for the Kerr solution can be written as

$$\mathcal{R}(B) = -\frac{M}{2\big(r - Ia\cos(\theta)\big)^3}(B + 3\sigma_r B\sigma_r).$$

Prove that this satisfies $\partial_b \mathcal{R}(b \wedge c) = 0$ and interpret both parts of this result.

14.16 The Newtonian gauge form of the Kerr solution involves the spheroidal coordinates $\bar{r}$ and $\bar{\theta}$. Prove that $\bar{r} = \cos(\bar{\theta}) = 0$ defines a ring.