Black Holes

7

- A black hole forms around an object massive enough and dense enough to fit within its event horizon, which is a one-way surface through which particles and light can only traverse inward. Thus an exterior observer cannot receive any signal sent from inside.
- In this chapter, we mostly study the spherically symmetric, nonrotating Schwarzschild black hole. The event horizon is a spherical surface of radius $r = r^* = 2G_{\mathrm{N}}M/c^2$, which is a coordinate singularity of the Schwarzschild metric. The metric elements g_{00} and g_{rr} change signs when crossing from the $r > r^*$ to the $r < r^*$ region, leading to a role reversal between space and time.
- The gravitational energy unleashed when a particle falls into a tightly bound orbit around a black hole can be enormous, more than ten times that released in a nuclear fusion reaction. This powers some of the most energetic phenomena observed in the universe.
- GR-based models show that a rotating star of sufficient final mass ($\gtrsim 3M_\odot$) after it burns out cannot support its own weight, inevitably collapsing into a rotating (Kerr) black hole.
- The physical reality of, and observational evidence for, black holes are briefly discussed.
- There is a mysterious correspondence between the laws of black hole physics and the laws of thermodynamics. In particular, the surface gravity at the event horizon behaves like the temperature of a thermodynamical system; the horizon area behaves like the entropy.
- This correspondence was greatly strengthened by the discovery of Hawking radiation. Quantum fluctuations around the event horizon bring about the thermal emission of particles and light from a black hole. This is allowed because pair-produced particles falling into the black hole can have negative energy; their partners may thereby escape with positive energy without violating energy conservation.

In Chapter 6, we began to describe the Schwarzschild geometry of the spacetime outside a spherical source. In particular, we studied the bending of a light ray by the sun and the precession of the planet Mercury's perihelion.

A College Course on Relativity and Cosmology. First Edition. Ta-Pei Cheng.

For these solar system applications, gravity is relatively weak (and therefore GR corrections are small). Here we study the spacetime structure exterior to any object whose mass is so compressed that its radius is smaller than its Schwarzschild radius $r^* = 2G_{\rm N}M/c^2$. Such objects have been given (by John Wheeler) the evocative name black holes: they are holes because radiation and matter can fall into them; they are black because nothing, not even light, can escape from them. These structures necessarily involve such strong gravity (i.e., such strongly curved spacetime) that the GR framework is indispensable for their explication.

Even in the context of Newtonian physics, one can consider a gravity so strong that light cannot escape. In the eighteenth century, John Michell (1724–1793) and (independently) Pierre-Simon Laplace (1749–1827), proposed the possibility of a "black star" whose ratio of mass to radius was so large that the required escape velocity $v_{\rm esc} = \sqrt{2G_{\rm N}M/r} = c\sqrt{r^*/r}$ exceeded the light velocity c. Of course, this speculation was based upon the (from our modern perspective) erroneous assumption that light carried a gravitational mass. GR interprets this phenomenon instead in terms of the causal structure of the spacetime outside a strong gravitational source.

Black holes manifest the full power and glory of Einstein's GR. One of its signature features is the equal treatment of space and time; hence spacetime is the natural arena for the description of physical phenomena. GR is the classical field theory of gravitation, in which curved spacetime is the gravitational field. Now, in the case of black holes, gravity is so strong and the spacetime so warped that the roles of space and time are interchanged, leading to many counterintuitive results.

7.1 Schwarzschild black holes

Coordinate singularities The Schwarzschild geometry in Schwarzschild coordinates (ct, r, θ, ϕ) has the metric

$$g_{\mu\nu} = \begin{pmatrix} -\left(1-\dfrac{r^*}{r}\right) & & & \\ & \left(1-\dfrac{r^*}{r}\right)^{-1} & & \\ & & r^2 & \\ & & & r^2\sin^2\theta \end{pmatrix}. \tag{7.1}$$

The metric and its inverse have singularities at $r=0$ and $r = r^*$, as well as $\theta = 0$ and π. We understand that $\theta = 0$ and π are coordinate singularities associated with our choice of the spherical coordinate system. They are not physical, do not show up in physical measurements at $\theta = 0$ and π, and can be removed by a coordinate transformation. However, the $r = 0$ singularity is real. This is not surprising, as the Newtonian gravitational potential ($\Phi \sim 1/r$) for a point mass already has this feature.

What about the $r = r^*$ surface? As we shall demonstrate, it is actually a coordinate singularity. We have discussed the Riemann curvature tensor $R_{\mu\nu\lambda\rho}$ (see (6.20) as well as (6.7)), a nonlinear second-derivative function of the metric that is nontrivial only in a curved spacetime. In the case of Schwarzschild geometry, the coordinate-independent product[1] $R_{\mu\nu\lambda\rho}R^{\mu\nu\lambda\rho} = 12r^{*2}/r^6$ is only singular at $r = 0$. This indicates that the singularity at $r = r^*$ must be associated with our choice of the Schwarzschild coordinate system. Namely, it is not physical and can be transformed away in suitable coordinates, for example, the Kruskal–Szekeres coordinates (see Section 7.1.2).

[1] The Ricci scalar is similarly nonsingular, as it is proportional to the trace of energy-momentum tensor $R = -\kappa T$ as discussed in Section 6.3.2.

The event horizon While physical measurements are not singular at $r = r^*$, that does not mean that this surface is not special. It is an event horizon, separating events that can be viewed from afar from those that cannot (no matter how long one waits). That is, the $r = r^*$ surface is the boundary of a region from which it is impossible to send out any signal. It is a boundary of communication, much as earth's horizon is a boundary of our vision. An event horizon is a one-way barrier: any timelike or null worldline can pass through only inward; particles and light rays cannot move outward.

7.1.1 Time measurements around a black hole

We shall begin our discussion of the causal structure of an event horizon with a simple examination of the elapsed time for a particle traveling inward across the $r = r^*$ boundary. While the proper time of the crossing particle is perfectly finite, a faraway observer sees this crossing take an infinite amount of (Schwarzschild coordinate) time. Thus no signal sent from the horizon's surface or its interior can reach such an observer.

The local proper time

We have already mentioned that there is no physical singularity at $r = r^*$. Here we will examine the time measured by an observer traveling across the Schwarzschild surface. The result shows that such a physical measurement is not singular at $r = r^*$.

Let τ be the proper time measured on the surface of a collapsing star (or, alternatively, the time aboard a spaceship traveling radially inward). Recall from Section 6.4.2 that for a particle (with mass) in the Schwarzschild spacetime, we can write a generalized energy balance equation (6.88). This equation can be simplified further for the case of a collapsing star or infalling spaceship starting from rest at $r = \infty$ (so that $\mathcal{E} = 0$), following a radial path along some fixed azimuthal angle ϕ (i.e., $d\phi/d\tau = 0$, so the angular momentum $l = 0$):

$$\frac{1}{2}\dot{r}^2 - \frac{G_N M}{r} = 0; \tag{7.2}$$

thus

$$\frac{1}{c^2}\left(\frac{dr}{d\tau}\right)^2 = \frac{2G_N M}{c^2 r} = \frac{r^*}{r}; \quad \text{hence} \quad c\,d\tau = \pm\sqrt{\frac{r}{r^*}}\,dr. \tag{7.3}$$

The plus sign corresponds to an exploding star (or an outward-bound spaceship) and the minus sign to a collapsing star (or an inward-bound probe). We pick the minus sign. A straightforward integration yields

$$\tau\,(r) = \tau_0 - \frac{2r^*}{3c}\left[\left(\frac{r}{r^*}\right)^{3/2} - \left(\frac{r_0}{r^*}\right)^{3/2}\right], \tag{7.4}$$

where τ_0 is the time when the probe is at some reference point r_0.

Thus the proper time $\tau(r)$ is perfectly smooth at the Schwarzschild surface (see Fig. 7.1). An observer on the surface of the collapsing star would not feel anything peculiar when the star passed through the Schwarzschild radius. It would then take a finite amount of proper time to reach the origin, which is a physical singularity.

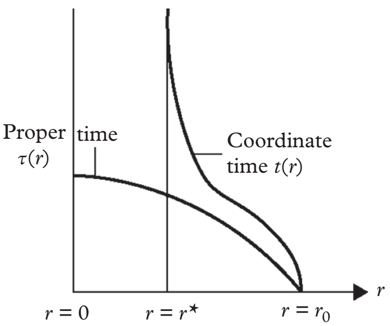

Figure 7.1 *The contrasting behavior of proper time $\tau(r)$ vs. coordinate time $t(r)$ at the Schwarzschild surface.*

Exercise 7.1 Travel time from the event horizon to the singular origin

(a) How much proper time $\Delta\tau$ (in terms of the Schwarzschild radius r^*) passes for a probe falling from the event horizon to the $r = 0$ singularity? You may assume that the probe fell in radially from rest at infinity as in the discussion above. (b) Evaluate this time interval for the case of a black hole with a mass $3M_\odot$ as well as the case of a supermassive black hole with a mass $10^9 M_\odot$.

The Schwarzschild coordinate time

While the time measured by an observer traveling across the Schwarzschild surface is perfectly finite, this is not the case for an observer far away from the source. Recall that the Schwarzschild coordinate t is the time measured by an observer far away, where the spacetime approaches the flat Minkowski limit. Here we will show that the Schwarzschild coordinate time blows up as the probe approaches the $r = r^*$ surface. To find the coordinate time as a function of the radial coordinate in the $r > r^*$ region, we start with the chain rule: $dt/dr = (dt/d\tau)/(dr/d\tau) = \dot{t}/\dot{r}$. We already have an expression for $\dot{r}$ from (7.3), while $\dot{t}$, according to (6.82), is directly related to the conserved particle energy κ, which is fixed to be c because we are considering a geodesic with zero kinetic energy at infinity, $\mathcal{E} = m(\kappa^2 - c^2)/2 = 0$. In this way, we find $dt/dr = \dot{t}/\dot{r} = -(1 - r^*/r)^{-1}/c(r^*/r)^{1/2}$, so that

$$c\,dt = -\sqrt{\frac{r}{r^*}}\frac{dr}{1 - r^*/r}, \tag{7.5}$$

which shows clearly the singularity at r^*. For $r \simeq r^*$, we can integrate $c\,dt \simeq -r^*\,dr/(r-r^*)$ to display the logarithmic singularity:

$$t - t_0 \simeq -\frac{r^*}{c} \ln \frac{r - r^*}{r_0 - r^*}. \tag{7.6}$$

It takes an infinite amount of coordinate time to reach $r = r^*$. The full function $t(r)$ in the region outside the Schwarzschild surface can be calculated[2] and is displayed in Fig. 7.1.

[2] See, e.g., Problem 8.2 (solved on p. 387) in (Cheng 2010).

Infinite gravitational redshift The above-discussed phenomenon of a distant observer seeing a collapsing star slow to a standstill can also be interpreted as an infinite gravitational time dilation. The relation (6.70) between coordinate and proper time intervals for a stationary observer is given by

$$dt = \frac{d\tau}{\sqrt{-g_{00}}} = \frac{d\tau}{\sqrt{1 - r^*/r}}. \tag{7.7}$$

Clearly, the coordinate time interval dt will blow up as r approaches r^*. In terms of wave peaks, this means that an infinite time interval passes between peaks reaching the faraway receiver. This can be equivalently described as an infinite gravitational redshift. Equation (5.45) showed that the ratio of the received frequency to the emitted frequency is

$$\frac{\omega_{\text{rec}}}{\omega_{\text{em}}} = \sqrt{\frac{(g_{00})_{\text{em}}}{(g_{00})_{\text{rec}}}} = \sqrt{\frac{1 - r^*/r_{\text{em}}}{1 - r^*/r_{\text{rec}}}} = \frac{d\tau_{\text{em}}}{d\tau_{\text{rec}}} \tag{7.8}$$

When $r_{\text{em}} \to r^*$, the received frequency ω_{rec} approaches zero, as the time between received peaks $d\tau_{\text{rec}}$ blows up to infinity. Thus no signal can be transmitted from the black hole.

7.1.2 Causal structure of the Schwarzschild surface

The phenomenon of infinite gravitational redshift implies the impossibility of any signal transmission from the $r < r^*$ region to an outside observer. The Schwarzschild surface is in fact a one-way barrier: while matter and radiation can proceed inward across the horizon, no particle, whether massive or massless, can move outward.

Role change between space and time

To gain a deeper understanding of the Schwarzschild surface as an event horizon, we need to study the causal structure of the geometry exterior to a spherical source. One of the key differences between the $r > r^*$ and $r < r^*$ regions is that the roles of space and time are interchanged, because the metric functions g_{00} and g_{rr} exchange signs at $r = r^*$:

$$r > r^* \quad \text{outside a black hole:} \quad g_{00} < 0, \quad g_{rr} > 0 \quad \text{normal metric;}$$

$$r < r^* \quad \text{inside a black hole:} \quad g_{00} > 0, \quad g_{rr} < 0 \quad \text{flipped metric.} \tag{7.9}$$

Thus, in the $r > r^*$ region, we have the familiar result that the t coordinate is timelike, while the radial r coordinate is spacelike, but inside the Schwarzschild surface, the time axis is actually spacelike and the radial axis timelike. As we discussed in the flat-spacetime diagram, Fig. 3.5, the lightcone structure normally dictates that a timelike or lightlike ($ds^2 \leq 0$) trajectory inevitably moves in the direction of ever-increasing time. You cannot stand still (much less go back) in time, but you can move freely in space. However, (7.9) tells us that inside the $r = r^*$ surface of a black hole, any particle following a timelike or lightlike worldline cannot rest at a fixed radial position (much less move outward) but must proceed toward the $r = 0$ singularity.

Schwarzschild coordinates and their limitation

As we emphasized while discussing the spacetime diagram, the lightcone structure can clarify the causal structure of spacetime. In this context, the tipping of lightcones at the event horizon illustrates the above-discussed role reversal of the Schwarzschild time and space coordinates. For $r > r^*$, the lightcones open in the future direction; for $r < r^*$, they tip toward the singularity at the origin (see Fig. 7.2). We might also like to depict the mechanism by which the event horizon acts as a one-way barrier. Unfortunately, we will see that the coordinate singularity at $r = r^*$ makes Schwarzschild coordinates unsuitable for such an investigation. The faraway observer who measures the coordinate time never sees anything cross the event horizon.

To study lightcones is to study the light geodesics that form them. Recall that in a flat spacetime the radial (i.e., with fixed θ and ϕ) lightlike worldlines, corresponding to the solutions of the condition $ds^2 = -c^2\,dt^2 + dr^2 = 0$, or $c\,dt = \pm dr$, are straight lines of unit slope in the (ct, r) spacetime diagram:

$$ct = \pm r + \text{constant.} \tag{7.10}$$

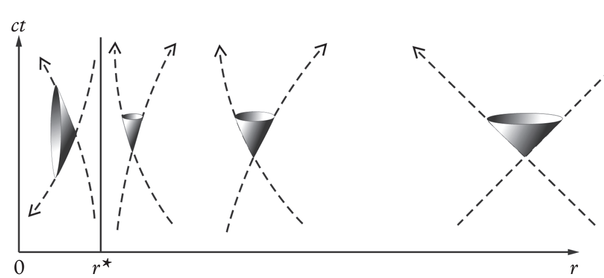

Figure 7.2 *Lightcone behavior in Schwarzschild coordinates. The dashed lines are lightlike paths. Lightcones close up when they approach the Schwarzschild surface in the Schwarzschild coordinate system. Inside the black hole, they tip over toward the singularity at r = 0.*

The plus sign is for outgoing (r increasing with t) and the minus sign for incoming light. Timelike worldlines are always contained inside the lightcone; particles with mass must proceed toward the future, cf. Fig. 3.5.

A radial ($d\theta = d\phi = 0$) worldline for a photon in the Schwarzschild coordinates has a null line-element interval:

$$0 = ds^2 = -\left(1 - \frac{r^*}{r}\right) c^2\, dt^2 + \left(1 - \frac{r^*}{r}\right)^{-1} dr^2$$

$$= -\left(1 - \frac{r^*}{r}\right)\left(c\, dt + \frac{dr}{1 - r^*/r}\right)\left(c\, dt - \frac{dr}{1 - r^*/r}\right). \tag{7.11}$$

We then have[3]

[3] This relation differs from (7.5), because we are now considering lightlike worldlines.

$$c\, dt = \pm \frac{dr}{1 - r^*/r}, \tag{7.12}$$

which can be integrated to give

$$ct = \pm(r + r^* \ln|r - r^*| + \text{ constant}). \tag{7.13}$$

Outside the event horizon ($r > r^*$), the plus sign is for the outgoing, and the minus sign for the infalling, lightlike geodesics. In Fig. 7.2, we have drawn several representative lightcones. We note that for the region far from the source where the spacetime becomes flat, the lightcones formed by the null geodesics (7.13) approach their expected form with sides of unit slope; however, as one moves closer to the $r = r^*$ surface, the lightcones close up. Inside the event horizon ($r < r^*$), the causal structure changes as discussed above. The lightcones tip over, opening toward the $r = 0$ singularity rather than toward ever-increasing time.

The fact that the metric becomes singular at $r = r^*$ means that the Schwarzschild coordinates (t, r, θ, ϕ) are not convenient for the description of events near the Schwarzschild surface. One might get the impression from the clammed-up lightcones in Fig. 7.2 that nothing ever crosses the event horizon. In fact, all it means is that a distant observer (whose clock keeps the Schwarzschild coordinate time) never sees it happen. Better-behaved coordinates should better depict such events.

Better-behaved coordinate systems

We now search for coordinates that can describe the Schwarzschild geometry without the $r = r^*$ singularity. In such coordinates, the lightcones should tilt over smoothly.

We start with the (advanced) Eddington–Finkelstein (EF) coordinates, whose time coordinate ($c\, d\bar{t}$) is chosen to equal the distance traveled by an infalling photon ($-dr$). Recall from (7.4) that the proper time of a particle falling into a black hole is smooth for all values of r. Instead of setting up the coordinate system using a static observer far from the gravitational source (as in Schwarzschild coordinates), one could describe the Schwarzschild geometry from the viewpoint of an infalling observer. Mathematically, an even simpler choice is to use an infalling photon. While a photon cannot measure a proper time, its traveled distance could serve as a time coordinate: $c\,d\bar{t} = -dr$, which would make the infalling photon's worldline look the same as (7.10) for a flat space. This suggests that we replace the factor $[c\,dt + dr/(1 - r^*/r)]$ in (7.11) with $(c\,d\bar{t} + dr)$. Namely, we introduce a new coordinate time $c\,d\bar{t} = c\,dt + r^*\,dr/(r - r^*)$, so that, in terms of $\bar{t}$, the light geodesic condition $ds^2 = 0$ is satisfied (i.e., (7.11) is solved) by the vanishing in turn of the expressions in the last two parentheses:

$$c\,d\bar{t} = -dr \qquad \text{(incoming)}, \tag{7.14}$$

$$c\,d\bar{t} = \frac{r + r^*}{r - r^*}\,dr \qquad \text{(outgoing/incoming)}. \tag{7.15}$$

In the new spacetime diagram with $(c\bar{t}, r)$ axes (Fig. 7.3a), incoming light follows straight paths of negative unit slope (for which we rigged the time coordinate). Meanwhile, the outgoing lightlike worldlines gradually become steeper

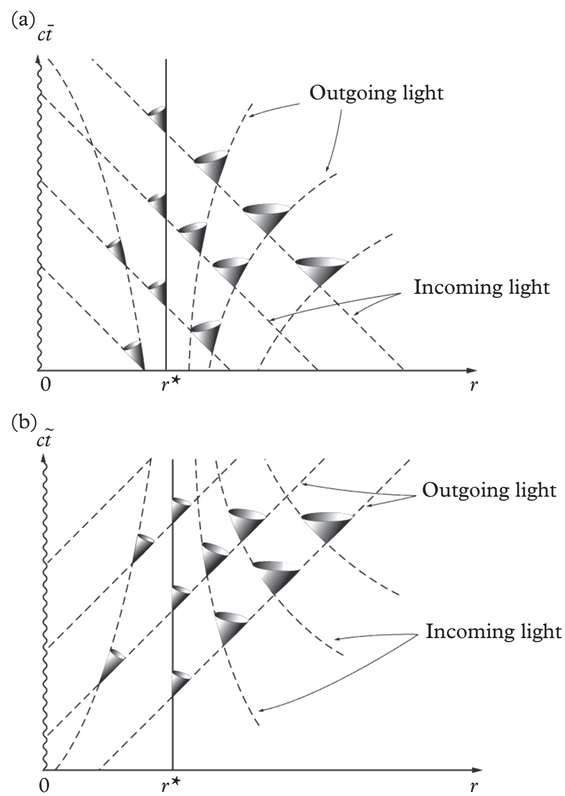

Figure 7.3 *Lightcones tilt over smoothly in Eddington-Finkelstein spacetime. (a) A black hole in advanced EF coordinates $(\bar{t}, r)$: all light geodesics inside the horizon move toward the future $r = 0$ singularity. (b) Reversing the time coordinate ($d\tilde{t} = -d\bar{t}$) yields a white hole in retarded EF coordinates $(\tilde{t}, r)$: all light geodesics inside the horizon move away from the past $r = 0$ singularity.*

as they approach the $r = r^*$ event horizon, which itself is a vertical null line, a lightlike path. Once inside the $r < r^*$ region, the coefficient in (7.15) becomes negative, just like that in (7.14). Thus both lightlike geodesics, (7.14) and (7.15), are incoming; r decreases as $\bar{t}$ increases. Namely, inside the black hole, all light cones open inward, so all timelike geodesics head for the $r = 0$ singularity. We already saw the tipping of lightcones in Schwarzschild coordinates. The improvement in EF coordinates is that the tipping is smooth. The lightcones do not clam up at the event horizon; some cross over into the black hole. Thus EF coordinates can credibly describe the event horizon.

A stationary ($dr = 0$) point on the $r = r^*$ Schwarzschild surface is the limit of solutions to (7.15) and therefore traces a null (lightlike) path in spacetime, a vertical line in Fig. 7.3(a). Thus the stationary event horizon is a null surface in spacetime, everywhere tangent to inward-pointing lightcones as shown in Fig. 7.4. This is what makes the event horizon special; this is why it is a one-way barrier. A timelike worldline passing through any point on such a null surface can only point inward toward the $r = 0$ singularity.

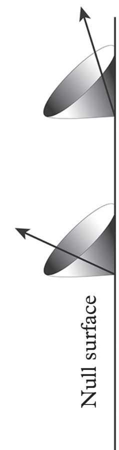

Figure 7.4 *A null surface is an event horizon. The lightcones of all points on the null surface are on one side of the surface. All timelike worldlines (samples shown as arrowed lines) are contained inside lightcones and thus can cross the null surface only in one direction. Therefore, a null surface is a one-way barrier.*

Black hole vs. white hole One may wonder whether the same procedure can be applied to the last parenthesis in (7.11) to straighten out the outgoing light geodesics. Indeed, we can define a (retarded) EF time $\tilde{t}$ with $c\,d\tilde{t} = dr$ for the outgoing light and $c\,d\tilde{t} = -[(r + r^*)/(r - r^*)]dr$ for the incoming/outgoing light. Essentially, we have just reversed the direction of time: $d\tilde{t} = -d\bar{t}$. Again we see that the lightcones, depicted in Fig. 7.3(b), tilt over smoothly across the Schwarzschild surface. But instead of tipping inward as in Fig. 7.3(a), they lean outward away from the $r = 0$ singularity. That is, while the Schwarzschild geometry depicted in the advanced EF coordinates has a future singularity at $r = 0$, the geometry depicted in retarded EF coordinates contains a past singularity at $r = 0$. Once again, the $r = r^*$ surface is a null surface, a one-way membrane allowing the transmission of particles and light only in one direction—this time outward. Thus we now have a white hole (containing the past singularity). While this time-reversed black hole is a perfectly good solution to Einstein's equation (which of course is covariant under time reversal), we have not found such a thing in our physical universe.

Exercise 7.2 Retarded EF coordinates with past $r = 0$ singularity

Above, we obtained the advanced EF coordinates $(\bar{t}, r)$ with lightlike geodesics, (7.14) and (7.15), defining lightcones tilting over smoothly inward toward a future $r = 0$ singularity. Obtain likewise the corresponding retarded EF coordinates $(\tilde{t}, r)$. Find the outgoing and incoming light geodesics that bound lightcones tilting outward away from a past $r = 0$ singularity as in Fig. 7.3(b).

Metric is singularity-free at $r = r^*$ in Kruskal–Szekeres coordinates Above, we separately straightened the light geodesics into worldlines of unit slope. Advanced EF coordinates straightened the incoming null geodesics, and retarded EF coordinates the outgoing. One can actually make a coordinate transformation that straightens both at the same time, i.e., that simplifies both parentheses in (7.11). But the metric still has the troublesome prefactor $(1 - r^*/r)$. Kruskal and Szekeres independently found a transformation involving the exponentiation of the coordinates that eliminates the singularity at $r = r^*$, leaving only the genuine one at $r = 0$. As the procedure is somewhat complicated, we shall omit its presentation.[4]

Section 7.1.3 concludes our presentation of the Schwarzschild black hole by discussing the orbits of particles around such a compact source of gravity.

[4] For an elementary introduction with some details worked out, see (Cheng 2010, Section 8.1.3).

7.1.3 Binding energy to a black hole can be extremely large

We are familiar with the fact that thermonuclear fusion, when compared with chemical reactions, is a very efficient process for releasing the rest energy of a particle. Here we show that binding a particle to a compact gravity source like a black hole can be an even more efficient mechanism. The thermonuclear reactions taking place in the sun can be summarized as fusing four protons (hydrogen nuclei, each with a rest energy of 938 MeV) into a helium nucleus (having a rest energy smaller than the sum of the four proton rest energies) with a released energy of 27 MeV, which represents $27/(4 \times 938) = 0.7\%$ of the input energy. Here we discuss the energy that can be released when a particle first falls into stable orbits around a Schwarzschild black hole before it eventually spirals through the event horizon.

Recall that the orbit can be determined from the effective 1D energy balance equation that we studied in the Chapter 6. Equation (6.88) may be written as

$$\frac{1}{2}m\dot{r}^2 + m\Phi_{\text{eff}} = \mathcal{E}, \tag{7.16}$$

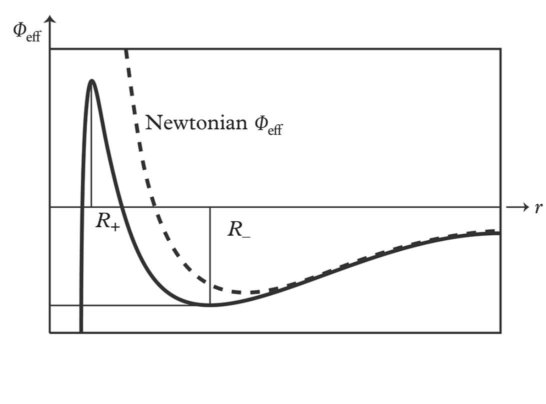

Figure 7.5 *Schwarzschild vs. Newtonian effective potential. The solid curve represents a specific choice of angular momentum $(l_0/mc)^2 \simeq 4.6r^{*2}$. For higher l, Φ_{eff} more closely tracks the Newtonian potential before falling off sharply at lower r. In the text it is shown that for l below l_0, $(l_0/mc)^2 = 3r^{*2}$, there are no maxima or minima; the potential is monotonic.*

with an effective potential

$$\Phi_{\text{eff}} = -\frac{G_{\text{N}}M}{r} + \frac{l^2}{2m^2r^2} - \frac{r^*l^2}{2m^2r^3}. \tag{7.17}$$

The first term on the right-hand side is clearly the Newtonian gravitational potential, the second term is the rotational kinetic energy (the centrifugal barrier), and the last term is the new GR contribution. It is a small correction for situations such as a planet's motion discussed in Chapter 6, but can be very important when the radial distance r is comparable to the Schwarzschild radius r^* as in the case of an orbit just outside the horizon.

The innermost stable circular orbit We can find the extrema of this potential by setting $\partial\Phi_{\text{eff}}/\partial r = 0$:

$$\frac{r^*c^2}{2r^2} - \frac{l^2}{m^2r^3} + \frac{3r^*l^2}{2m^2r^4} = 0, \tag{7.18}$$

or

$$r^2 - 2\left(\frac{l}{mc}\right)^2\frac{r}{r^*} + 3\left(\frac{l}{mc}\right)^2 = 0. \tag{7.19}$$

The solutions R_+ and R_- specify the locations where Φ_{eff} has a maximum and a minimum, respectively (see Fig 7.5):

$$R_\pm = \frac{1}{r^*}\left(\frac{l}{mc}\right)^2\left[1 \mp \sqrt{1 - 3\left(\frac{r^*mc}{l}\right)^2}\right]. \tag{7.20}$$

We note the important difference between the Schwarzschild Φ_{eff} and its Newtonian analog, whose centrifugal barrier always dominates in the small-r region ($l^2/r^3 \to \infty$ as $r \to 0$). This means that a particle in the Newtonian field cannot fall into the $r = 0$ center as long as $l \neq 0$. In the small-r region of the relativistic Schwarzschild geometry, the last (GR correction) term in (7.17) dominates ($-l^2/r^4 \to -\infty$ as $r \to 0$). When $\mathcal{E} \geq m\Phi_{\text{eff}}(R_+)$, a particle can plunge into the gravity center even if $l \neq 0$. For $\mathcal{E} = m\Phi_{\text{eff}}(R_-)$, just as in the Newtonian case, we have a stable circular orbit with radius R_-. This circular radius cannot be arbitrarily small. In order to have an orbit of any kind and not just plunge into the black hole, a particle must have enough angular momentum to create a sufficient centrifugal barrier. Equation (7.19) must have a solution, so its determinant in the square root of (7.20) must be nonnegative. This fixes a minimum angular momentum l_0:

$$3\left(\frac{r^*mc}{l_0}\right)^2 = 1, \quad \text{or} \quad \left(\frac{l_0}{mc}\right)^2 = 3\,(r^*)^2, \tag{7.21}$$

so that the innermost stable circular orbit (ISCO) has a radius

$$R_0 = \frac{1}{r^*}\left(\frac{l_0}{mc}\right)^2 = 3r^*. \tag{7.22}$$

As we reduce the angular momentum to $l = l_0$, the centrifugal barrier peak in Φ_{eff} (Fig. 7.5) at R_+ falls and the stable orbit trough at R_- rises until they meet, forming a flat point of inflection at R_0, the ISCO. For smaller angular momenta ($l < l_0$), the potential falls monotonically from $\lim_{r\to\infty} \Phi_{\text{eff}} = 0$ to $\lim_{r\to 0} \Phi_{\text{eff}} = -\infty$, so there are no orbits.[5] The plasma in an accretion disc around a black hole settles into stable orbits, but will lose its orbital angular momentum through turbulent viscosity (due to magnetohydrodynamic instability) and eventually, owing to the disappearing centrifugal barrier, spiral into the black hole.

[5] For a more detailed discussion of Φ_{eff}, see (Wald 1984, pp. 139–143).

The binding energy of a particle around a black hole To illustrate the energy of gravitational binding by a Schwarzschild black hole, consider a free particle that falls toward a black hole and ends up bound in the ISCO. Thus, according to (7.22) and (7.21), the particle orbits at a radial distance $r = R_0 = 3r^*$ with angular momentum $l_0 = \sqrt{3}r^* mc$. According to the energy balance equation (7.16) with $\dot{r} = 0$, we have $\mathcal{E} = m\Phi_{\text{eff}} = -mc^2/18$. This solution gives the total energy for the gravitationally bound particle:[6]

[6] Cf. (6.87) and Sidenote 34 in Chapter 6.

$$\frac{E(\infty)}{mc^2} = \frac{\kappa}{c} = \sqrt{\frac{2\mathcal{E}}{mc^2} + 1} = \sqrt{\frac{8}{9}} = 0.94. \tag{7.23}$$

That is, 6% of the rest energy is released—almost ten times larger than the 0.7% from thermonuclear fusion.[7]

[7] We should note that the gravitational binding energy of a particle around a spinning black hole is even greater; it can be as much as 42% of its rest energy!

7.2 Astrophysical black holes

So far we have concentrated on the Schwarzschild black hole: an idealized, static, spherically symmetric entity. Is it relevant for any astrophysical phenomena? In this section, we shall qualitatively answer this question on two fronts: on the theoretical side we summarize the results of studies of more realistic black hole solutions in GR; on the phenomenological side, we briefly report the present status of our search for black holes in the universe.

7.2.1 More realistic black holes

First, we present some theoretical results applicable to more realistic black holes. The Kerr solution of the Einstein equation generalizes the Schwarzschild solution to rotating sources. Model studies indicate that stellar gravitational collapse can result in rotating black holes.

Rotating black holes

Most stars rotate and thus have only axial symmetry. The simplest of such sources of gravity is characterized not just by its mass M but also by its angular momentum $\mathcal{J}$. The solution of Einstein's equation for the spacetime exterior to such a rotating source was discovered by Roy Kerr (1934–). The Kerr spacetime reduces to the Schwarzschild geometry in the limit of $\mathcal{J} = 0$, but has in general a considerably more complicated singularity structure. We shall present only a brief introduction to some of its salient features. The physical singularity is no longer a single point, but a ring perpendicular to the symmetry axis, with a radius proportional to the angular momentum of the source. Similarly, the Kerr black hole has an event horizon, a null surface, that is not spherical but ellipsoidal. While the event horizon of a Schwarzschild black hole coincides with the surface of infinite redshift, the Kerr horizon surface is enclosed inside the surface of infinite redshift, which coincides with the stationary limit surface, which we will explain below. As the source rotates, GR predicts that the spacetime will be dragged along;[8] see Fig. 7.6. If a particle (or photon) starts with a vanishing angular momentum $l = 0$, we would normally expect it to fall straight toward the center of the gravitational attraction; but with a rotating inertial frame of reference, such a zero-angular-momentum particle would still develop an angular velocity. The stationary limit surface is the boundary of a region where the frame dragging is so strong that no particle (not even light) can be stationary; everything rotates in the same direction as the source—even if it entered with great angular momentum opposing the source rotation.

An interesting feature of the rotating black hole is that one can actually extract energy from it. A physical processes (called the Penrose process) taking place in the region (called the ergosphere) between the stationary limit surface and the event horizon null surface can send particles to distant observers that carry away the rotational energy of the source. Clearly, a rotating black hole can bring about more complex physical processes than a static Schwarzschild black hole. The mathematics involved is correspondingly more complicated, so we refer interested readers to more advanced texts.[9]

[8] This GR frame-dragging prediction has been tested by the Gravity Probe B experiment. This satellite experiment managed to measure the tiny gyroscopic precession brought about by earth's rotation.

[9] See, e.g., (Hobson et al. 2006, Chapter 13) or (Cheng 2010, Section 8.4).

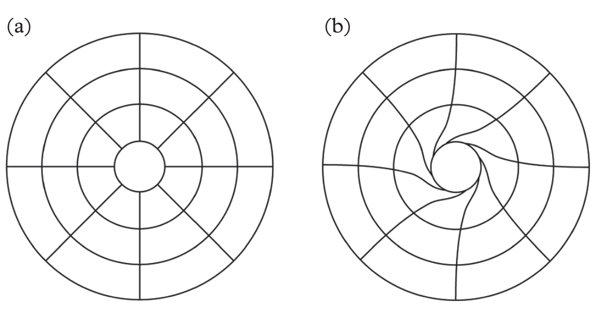

Figure 7.6 *Dragging of the inertial frame: a counterclockwise-rotating source turns (a) some initial spacetime (before source rotation) into (b) a twisted geometry (following the source rotation). Radial geodesics follow those twisted lines; they are swept along with the rotation.*

Gravitational collapse into a black hole

Our discussion of black holes has so far assumed a static source, an eternal black hole that has always existed, with its mass behind its event horizon. Naturally, one would like to know whether GR could allow the creation, through the gravitational collapse of more normal matter, of a region of spacetime with these features? Such an investigation would involve solving the Einstein equation for a nonvanishing energy/momentum source. The simplest model was the original Schwarzschild interior solution for a constant energy density; gradually, more realistic equations of state for the stellar interior were incorporated in such studies. The most influential investigations were carried out by Robert Oppenheimer (1905–1967) and his students around 1939. Their research showed analytically that a cold Fermi gas quickly collapses from a smooth initial distribution to form a black hole with the properties discussed above. The gravitational attraction causes each mass element to follow a geodesic trajectory toward the center. As the interior density increases, ever more exterior space is described by the Schwarzschild metric until all matter passes through the $r = r^*$ surface. The exterior then contains an event horizon, so a black hole is formed.

Nevertheless, the physics community remained skeptical of the reality of black holes. Their reservations were many. For example, it was questioned whether the assumption of spherical symmetry was too much of an idealization. How should one account for realistic complications such as deformation-forming lumps, shock waves leading to mass ejection, effects of electromagnetic and gravitational radiation, etc.? However, numerical calculations years later showed that any multipole distortion to the Schwarzschild metric is quickly shaken off through gravitational radiation; the source relaxes to the exact Schwarzschild black hole.

As for stellar rotation, we have already mentioned the Kerr solution found in 1963. While there does not exist an analytic solution of gravitational collapse for a rotating source analogous to the one discussed above, numerical calculations have again demonstrated that even with large distortions, a collection of matter with nonvanishing angular momentum always collapses into a Kerr black hole.

The revival of theoretical study of black holes since the 1950s was due in large part to the leadership of John Wheeler (1911–2008) in the United States and Yakov Zel'dovich (1914–1987) in the Soviet Union. The final acceptance by physicists of the GR prediction of black holes as the generic end product of gravitational collapse was brought about by the proof in the early 1960s of the singularity theorems[10] by Roger Penrose (1931–). Related to this, we note also a well-known theorem stating that all black holes can be completely characterized by their mass, angular momentum, and electric charge.[11] Their lack of any other features inspired the witty summary "Black holes have no hair."

[10] This set of theorems show in realistic situations the inevitability of the formation of an event horizon, within which always lies a singularity.

[11] The GR solution for an electrically charged source, called the Reissner-Nordström geometry, is thought to be less phenomenologically relevant.

7.2.2 Black holes in our universe

In Chapter 6, we saw that the GR predictions of bending of light and gravitational redshift have all been verified within our solar system. Can we likewise confirm the strong gravity predictions of black holes' existence?

Black holes are faraway, small, black discs in the sky; it would seem rather hopeless to ever observe them. But, by accounting for the gravitational effects of an object on its surroundings (such as gravitational lensing, the orbits of nearby stars, and the accretion of hot gas), one can estimate the mass of the object. If a highly compact source has a mass greater than about $3M_{\odot}$, it is a strong candidate to be a black hole—since no known mechanism can stop such a massive system from gravitationally collapsing into a black hole.

The Chandrasekhar limit What is the basis of this $3M_{\odot}$ limit? In an ordinary star, the gravitational attraction is balanced by the outward pressure from thermonuclear burning in its interior. When the fuel is spent, what else can prevent gravitational collapse? One possibility is the quantum mechanical repulsive force due to Pauli exclusion (called the degeneracy pressure or Pauli blocking) among particles of half-integer spin (fermions). This is the source of stability for white dwarfs (the fermions being electrons) and neutron stars (the fermions being neutrons). In 1930, Subrahmanyan Chandrasekhar (1910–1995) used the new quantum mechanics to show that for stellar masses $M > 1.4M_{\odot}$, the electrons' degeneracy pressure is not strong enough to stop the gravitational contraction. Thus he was the first one to make the radical suggestion that massive enough stars would collapse into black holes (decades before such nomenclature was invented). In 1932, James Chadwick (1891–1974) discovered the neutron, and, soon after, in 1934, Fritz Zwicky (1898–1974) suggested that the remnant of a supernova explosion, associated with the final stage of gravitational collapse, was a neutron star. Because of the strong nuclear force, there exists no simple analytic calculation for the corresponding limit for neutron stars, but numerical estimates of the neutron degeneracy pressure all point to a value not much more than $3M_{\odot}$.

Black holes in X-ray binaries The majority of stars are members of binary systems orbiting each other. If a black hole is in a binary system with another visible star, then, by observing the Kepler motion of the visible companion, one can obtain some limit on the mass of the invisible companion. If it exceeds $3M_{\odot}$, it is a black hole candidate. Even better, if the visible star produces significant gas (as in the case of solar flares), the infall of such gas (called accretion) into the black hole will produce intense X-rays. A notable example is Cygnus X-1, which is now generally accepted as a black hole binary system, in which the visible companion is a supergiant star[12] of mass $M_{\text{vis}} \simeq 30M_{\odot}$, and the invisible compact object, presumably a black hole, has a mass $M \geq 8M_{\odot}$. Altogether, close to ten such binary candidate black holes have been identified in our galaxy.

[12] The supergiant star, having a radius of about $20R_{\odot}$, cannot be the source of the observed X-rays.

Supermassive black holes It has also been discovered (again by detecting the gravitational influence on nearby visible matter) that at the centers of most galaxies are supermassive black holes, with masses ranging from 10^6 to $10^{12} M_\odot$. Even though the initial findings were a great surprise, once this discovery had been made, it was not too difficult to understand why we should expect such supermassive centers. The gravitational interaction between stars is such that they swing-and-fling past each other: one flies off outward while the other falls inward. Thus we can expect stars and dust to be driven inward toward the galactic core, producing a supermassive gravitational aggregate. Some of these galactic nuclei emit huge amounts of X-rays and visible light a thousand times brighter than the stellar light of a galaxy. This is interpreted as light emitted from gas heated as it is funnels into the central black hole.[13] Such galactic centers are called AGN (active galactic nuclei). The well-known astrophysical objects, quasars (quasi-stellar objects), are interpreted as AGN in the early stage of galactic evolution. Observations suggest that an AGN is composed of a massive center surrounded by a molecular accretion disk. It is thought to be powered by a rotating supermassive black hole at the core of the disk. Such huge emissions require extremely efficient mechanisms for releasing the energy associated with the matter surrounding the black hole. Recall our discussion in Section 7.1.3 of the huge gravitational binding energy of particles orbiting close to a black hole horizon. Thus, besides the electromagnetic extraction of rotational energy as alluded to above, an important vehicle is gravitational binding of accreting matter. The gravitational energy is converted into radiation when free particles fall into lower-energy centrally bound states in the formation of the accretion disk around the black hole. From a whole host of such observations and deductions, we conclude that galactic centers contain objects of tens of millions of solar masses. They must be black holes, because no other known object could be so massive and so small.

[13] When stars pass close enough to a massive black hole, tidal forces rip them apart, producing streams of debris that then swirl around and ultimately get swallowed by the black hole.

7.3 Black hole thermodynamics and Hawking radiation

The "no-hair" theorem suggests that black holes, being characterized only by mass, spin, and charge, are extraordinarily simple entities. One might conclude that a black hole has vanishing entropy ($S = 0$). This immediately runs into a contradiction: the process of matter falling into a black hole would then be an entropy-decreasing process! In the early 1970s, Jacob Bekenstein (1947–) pointed out that a black hole must have nonvanishing entropy proportional to its horizon area: $S \propto A^*$. Here we shall restrict ourselves to the simplest case of a Schwarzschild black hole, whose area is simply proportional to M^2:

$$A^* = 4\pi r^{*2} = 16\pi c^{-4} G_N^2 M^2. \tag{7.24}$$

As matter falls into a black hole, the horizon area (and hence the entropy) is clearly ever-increasing, as $(M + dM)^2 > M^2 + dM^2$. The area- and entropy-increasing theorem implies that while two black holes can join to make a bigger black hole, one black hole can never split into two, because $M_1^2 + M_2^2 < (M_1 + M_2)^2$.

7.3.1 Laws of black hole mechanics and thermodynamics

There is in fact a deep analogy between the laws of black hole mechanics and the laws of thermodynamics. After noting the surface gravity of a black hole,[14]

$$\sigma^* = \frac{G_N M}{r^{*2}} = \frac{c^4}{4G_N M}, \tag{7.25}$$

we list the four laws of black hole mechanics:

0th law: The surface gravity σ^* has the same value everywhere on the event horizon.

1st law: The change in mass of a black hole is proportional to the surface gravity times the change in area:

$$dM = \frac{\sigma^*}{8\pi G_N} dA^*. \tag{7.26}$$

2nd law: The surface area of the event horizon of a black hole can only increase, never decrease.

3rd law: It is impossible to lower the surface gravity to zero through any physical process.

Exercise 7.3 Change of BH mass is proportional to BH surface gravity and change of area

Use the definition of BH surface gravity (7.25) and area (7.24) to derive the mass/area relation shown in (7.26).

These laws of black hole mechanics are closely analogous to the four laws of thermodynamics:

0th law: The temperature T of a system in thermal equilibrium has the same value everywhere in the system.

1st law: The change in energy of a system is proportional to its temperature times the change in entropy: $dE = T\,dS$.

2nd law: The entropy of a system can only increase, never decrease.

3rd law: It is impossible to lower the temperature of a system to zero through any physical process.

[14] The surface gravity of a Schwarzschild black hole is the limit of the weight (per unit mass) of a stationary object near the event horizon, as measured by a distant observer (perhaps holding the object on a long massless string). You can see from (7.25) that it is analogous to the Newtonian acceleration at the event horizon.

Thus we have the following correspondence between black hole physics and the laws of thermodynamics:

Black holes		Thermodynamics
Mass M	$\Longleftrightarrow$	Energy E
Surface gravity at horizon σ^*	$\Longleftrightarrow$	Temperature T
Area of event horizon A^*	$\Longleftrightarrow$	Entropy S

This correspondence, except for that between mass and energy, is apparently mysterious. A black hole is a piece of spacetime geometry, not a container of gas and liquid—why should there be such a correspondence to a thermodynamical system? So far, all one can say is that these two sets of laws appear similar. This suggests the proportionality of entropy to area ($S \propto A^*$) and temperature to surface gravity ($T \propto \sigma^*$), but we do not know their proportionality constants. Such questions were partially answered in 1973 by the discovery of Hawking radiation, from the application of quantum mechanics to black holes.

7.3.2 Hawking radiation: quantum fluctuation around the horizon

Here we shall offer some brief comments on the interplay between black holes and quantum physics. Any detailed discussion of these advanced topics is beyond the scope of this introductory exposition. Our purpose here is merely to alert the readers to the existence of a vast body of knowledge on such topics, which are at the forefront of current research.

The Planck scale GR is a classical macroscopic theory. For a microscopic description, we would need to combine GR with quantum mechanics into a theory of quantum gravity. The natural scale for such a quantum description of gravity is the Planck scale.

Soon after his 1900 discovery of the eponymous Planck's constant $\hbar$ in fitting the blackbody spectrum, Max Planck (1857–1947) noted that a self-contained system of natural units of mass–length–time can be defined by various combinations of Newton's constant $G_{\rm N}$ (gravity), Planck's constant $\hbar$ (quantum theory), and the speed of light c (relativity). When we recall from Newtonian theory that $[G_{\rm N}\cdot (\text{mass})^2 \cdot (\text{length})^{-1}]$ has units of [energy], and from relativistic quantum theory that the natural scale of [energy· length] is $\hbar c$, we can obtain the natural mass scale for quantum gravity, the Planck mass,

$$M_{\rm Pl} = \left(\frac{\hbar c}{G_{\rm N}}\right)^{1/2}. \tag{7.27}$$

From this, we can immediately deduce the other Planck scales:

$$\begin{aligned}
\text{Planck energy} \quad E_{\text{Pl}} &= M_{\text{Pl}}c^2 = \left(\frac{\hbar c^5}{G_{\text{N}}}\right)^{1/2} = 1.22 \times 10^{22}\,\text{MeV}, \\
\text{Planck length} \quad l_{\text{Pl}} &= \frac{\hbar c}{E_{\text{Pl}}} = \left(\frac{\hbar G_{\text{N}}}{c^3}\right)^{1/2} = 1.62 \times 10^{-35}\,\text{m}, \\
\text{Planck time} \quad t_{\text{Pl}} &= \frac{l_{\text{Pl}}}{c} = \left(\frac{\hbar G_{\text{N}}}{c^5}\right)^{1/2} = 5.39 \times 10^{-44}\,\text{s}, \\
\text{Planck temperature} \quad T_{\text{Pl}} &= \frac{E_{\text{Pl}}}{k_{\text{B}}} = \left(\frac{\hbar c^5}{G_{\text{N}} k_{\text{B}}^2}\right)^{1/2} = 1.42 \times 10^{32}\,\text{K},
\end{aligned} \tag{7.28}$$

where we have also used Boltzmann's constant k_{B} to define a natural temperature scale. Such extreme scales are utterly beyond the reach of any laboratory setup. (Recall that the rest energy of a nucleon is about 1 GeV and that the highest energies probed by the current generation of accelerators are on the order of 10^4 GeV.) The only natural phenomena that can reach such an extreme scale are the physical singularities in GR: the endpoints of gravitational collapse hidden inside black hole horizons and the origin of the cosmological big bang. It is expected that quantum gravity will modify such singularity features of GR.

Quantum field theory vs. quantum gravity Quantum field theory (QFT) is the union of SR with quantum mechanics. Namely, SR describes classical fields (such as Maxwell's electromagnetic fields); a quantum description of fields (such as quantum electrodynamics) is a QFT. The quanta of a field are generally viewed as particles. For example, the quanta of an electromagnetic field are photons, the quanta of an electron field are electrons, etc. The central claim of QFT is that the essence of reality is a set of fields, subject to the rules of quantum mechanics and SR; all observed phenomena are consequences of the quantum dynamics of these fields. QFT is the natural language to describe interactions that include the possibility of particle creation and annihilation allowed by the relativistic energy and mass relation $E = mc^2$. QFTs of nongravitational interactions need not operate at the extreme Planck scale. Quantum gravity is the quantum theory of the gravitational field. As in other QFTs, the gravitational field has its quantum particle, the graviton. Gravitons interact with the energy–momentum 4-tensor as photons interact with the charge-current 4-vector. GR, with its geometric interpretation and warped spaces, must emerge[15] as the macroscopic (low-energy) limit of quantum gravity in the same way that Maxwell's electromagnetic theory emerges from quantum electrodynamics. But, as mentioned above, the natural scale of quantum gravity is the Planck scale; thus, in this context, all observable phenomena may be macroscopic. Gravity is too weak at our preferred scales to reveal its quantum nature.

[15] There are suggestions that space and time themselves are emergent concepts from quantum gravity.

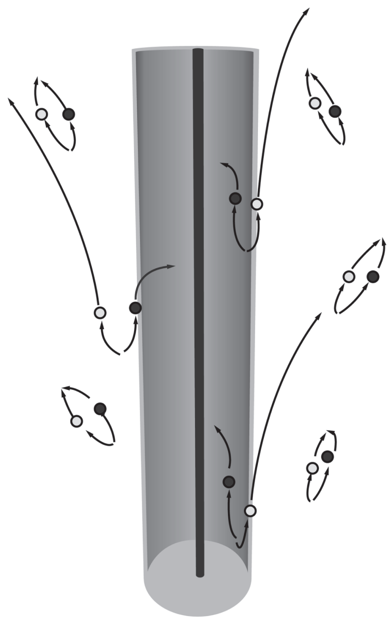

Figure 7.7 *Spacetime diagram of Hawking radiation. The cylinder is the (2+1)-dimensional worldsheet of the Schwarzschild horizon, with the time axis in the vertical direction. Quantum fluctuation causes particles and antiparticles to pop in and out of the vacuum. If one of the particles is inside the horizon and has negative energy, the other one of the pair can reach infinity with a positive energy. The black hole thereby radiates.*

Hawking radiation

The surprising theoretical discovery by Stephen Hawking (1942–) that a black hole can radiate (contrary to the general expectation that nothing can escape from it) was made in the context of a quantum description of particle fields in a background Schwarzschild geometry. That is, the relevant theoretical framework involves only a partial unification of gravity with quantum theory: while the fields of photons, electrons, etc., are treated as quantized systems, gravity is still described by the classical (nonquantum) theory of GR. Thus, the relevant theoretical framework is quantum field theory in a curved spacetime.

The quantum uncertainty principle of energy and time, $\Delta E \Delta t \gtrsim \hbar$, implies that processes can temporarily violate energy conservation, provided they do so for a sufficiently short time interval Δt. Such quantum fluctuations turn empty space into a medium with particle and antiparticle pairs appearing and disappearing. In normal circumstances, such energy-nonconserving processes cannot survive on a macroscopic timescale. Hence the temporarily created and destroyed particles are called virtual particles. However, Hawking showed that if such random quantum fluctuations take place near the event horizon of a black hole, the virtual particles can become real because in such a situation energy conservation can be maintained permanently.

Qualitative explanation Consider the simplest quantum fluctuation: the creation of a particle–antiparticle pair from the vacuum. If the pair are to persist and not promptly annihilate back into the vacuum, energy conservation requires that

$$0 = E(\infty) + \tilde{E}(\infty). \tag{7.29}$$

If both particles could reach $r = \infty$, then $E(\infty)$ and $\tilde{E}(\infty)$ would be their energies measured by observers at infinity. Such energies must be positive, so the equality (7.29) cannot be satisfied. However, if this quantum fluctuation takes place sufficiently close to the event horizon of a black hole, one particle can be outside (and eventually travel to $r = \infty$) and the other particle can be inside[16] the event horizon (and fall into the singularity); see Fig. 7.7. Recall our discussion of the causal structure of the event horizon. The roles of space and time are interchanged at the $r = r^*$ Schwarzschild surface. In the same way, the energy component[17] inside the horizon takes on the properties of momentum; in particular, it is possible for the energy of a particle to be negative. If $\tilde{E}(\infty)$ is negative, then the conservation relation (7.29) can be satisfied on a macroscopic timescale. To a distant observer, the black hole emits a particle with positive energy, while losing a corresponding amount by swallowing its partner with negative energy. This is known as the Hawking effect or Hawking radiation.

Result obtained from QFT in curved spacetime Any radiation field, whether quantum or classical, can be decomposed into plane waves. The (Fourier) coefficients of expansion obey the simple harmonic equation. Thus radiation

[16] During a fluctuation, one cannot locate any particle to such precision. If one insists on a classical picture with exact particle locations, one can say that after the particles' creation, one of them, during the fluctuation time Δt, travels across the event horizon—or, in quantum mechanical language, one particle tunnels across the horizon.

[17] Recall that just as time and space intervals are components of the same vector (ct, x^i), so are energy and momentum $(E/c, p^i)$.

can be viewed as a collection of harmonic oscillators.[18] A QFT treats these field oscillators according to quantum mechanics, so their energy spectra have discrete, evenly spaced levels, states with a particular number of energy quanta (particles). These quantum states are related by the so-called raising and lowering ladder operators. Thus QFT provides a natural language to describe the creation and annihilation of particles. In a curved spacetime, the natural decomposition is still into plane waves, but they are plane waves with respect to the coordinates that encode the curvature. Near the horizon of a Schwarzschild black hole, for example, the Kruskal–Szekeres coordinates are convenient. The vacuum (i.e., the lowest-energy state) of such a field system, when viewed by a distant observer in flat spacetime, is a state containing a distribution of particles. In this way, Hawking obtained the particle distribution result

$$|\Psi|^2 = \langle n \rangle = (e^{2\pi cE/\hbar\sigma^*} - 1)^{-1}. \tag{7.30}$$

This has the form of a thermal number distribution, $\langle n \rangle = (e^{E/k_B T} - 1)^{-1}$, with k_B being Boltzmann's constant. In this way, one can identify the black hole's surface gravity σ^* of (7.25) as temperature:

$$k_B T = \frac{\hbar\sigma^*}{2\pi c} = \frac{\hbar c^3}{8\pi G_N M} = \frac{\hbar c}{4\pi r^*}. \tag{7.31}$$

Namely, due to quantum effects in the surrounding space, a black hole of surface gravity σ^* should radiate particles as a perfect blackbody of temperature T, with (7.31) giving us the desired proportionality constant relating σ^* to T. Once derived, this expression for thermal energy appears reasonable on dimensional grounds. It is a relativistic quantum effect, hence the presence of $\hbar c$, which has the dimensions of [energy · length]; the only lengthscale available is the Schwarzschild radius r^*. Equivalently, this Hawking thermal energy can be expressed in terms of the natural quantum gravity unit of Planck energy:

$$\frac{k_B T}{E_{Pl}} = \frac{T}{T_{Pl}} = \frac{l_{Pl}}{4\pi r^*} = \frac{M_{Pl}}{8\pi M}. \tag{7.32}$$

In short, Hawking radiation shows that black holes radiate like blackbodies; smaller and hotter black holes should evaporate completely.

Black hole entropy From the expression for temperature and noting that $E = Mc^2$, we can immediately deduce the black hole's entropy through the definition $dS = T^{-1}\,dE$ (cf. Exercise 7.4). In this way, we find that the entropy is indeed proportional to the horizon area A^* of (7.24):

$$\frac{S}{k_B} = \frac{A^*}{4l_{Pl}^2}, \tag{7.33}$$

where l_{Pl}^2 is the Planck length squared. This is a shocking result! Entropy is an extensive variable, so one would expect it to be proportional to the volume, not

[18] For a simple discussion, see (Cheng, 2013, Sections 3.1 and 6.4).

the area, of the system. This result inspired a proposal that in quantum gravity the description of a volume of space is somehow encoded on its boundary; for instance, all the information about a black hole is encoded on its event horizon. Much like Einstein's proposal of the equivalence principle in GR, this idea has been elevated to a fundamental principle of quantum gravity: the holographic principle.[19]

[19] The holographic principle was first proposed by Gerard 't Hooft (1946–) and later formulated in the context of string theory by Leonard Susskind (1940–) and others.

Exercise 7.4 From Hawking temperature to the proportionality of black hole entropy to area

By a simple integration of $dS = T^{-1}\,dE$, derive the proportionality of black hole entropy and area shown in (7.33).

Deciphering the meaning of black hole entropy

One of the great achievements of Ludwig Boltzmann (1844–1906) was to show that the second law of thermodynamics was amenable to precise mathematical treatment. The macroscopic notion of entropy S could be related to a counting of the corresponding microscopic states, the complexions W:

$$S = k_{\mathrm{B}} \ln W. \tag{7.34}$$

What would be the microscopic statistical description of a black hole that corresponds to the just-obtained entropy? The traditional approach would suggest that to do this counting, we need a microscopic theory of quantum gravity. In fact, black hole entropy would provide a check on the viability of any proposed quantum gravity theory. Currently, the most developed theory is superstring theory. Indeed, string theorists have been greatly encouraged by some success in recovering the entropy (7.33) by counting the number of ways a black hole can be formed in superstring theory. Nevertheless, it is still not a total success; the black holes being studied can be properly described only as black-hole-like entities. We are still far from having a realistic quantum description of a black hole.

Entropic gravity? Some have suggested[20] that the entropy result is even more fundamental, that gravity is an entropic force. Just like the restoring force of a rubber band, which can be viewed as resulting from entropy maximization, the gravitational force can be derived (by reverse engineering) from extremizing entropy based on some conjectured principle governing the quantized spacetime. But this may not be such an interesting approach if one is ultimately interested in the detailed microscopic description of spacetime that quantum gravity promises to provide.[21]

[20] The reader interested in finding out more about such an approach may wish to start with (Jacobson 1995), (Padmanabhan 2010), and (Verlinde 2010).

[21] As in the case of a rubber band, we are ultimately more interested in the atomic/molecular explanation of its contracting force.

Review questions

1. What does it mean that the Schwarzschild surface at $r = r^*$ is only a coordinate singularity?
2. What is the event horizon associated with a black hole?
3. At the $r = r^*$ surface of a Schwarzschild black hole, the proper time is finite, while the coordinate time is infinite. To what time measurements do we refer in these two descriptions? In terms of light frequency, what is an alternative (but equivalent) description of this phenomenon of infinite coordinate time dilation?
4. An event horizon is a null 3-surface. What is a null surface? Why does it allow particles and light to traverse only in one direction?
5. What is the basic property of the time coordinate in the advanced EF coordinate system that allows the lightcone to tip over smoothly inward across the $r = r^*$ surface? Answer the same question for the retarded EF system (with outward-tipping lightcones). Such properties of the coordinates allow their respective spacetime diagrams to display the black hole and white hole solutions. What is a white hole?
6. The effective potential for a particle in Schwarzschild spacetime has the form

$$V_{\text{eff}} = -\frac{A}{r} + \frac{Bl^2}{r^2} - \frac{Cl^2}{r^3} \tag{7.35}$$

with positive coefficients A, B, and C. Use this expression to explain why, unlike in the Newtonian central force problem, a particle can spiral into the center even with nonzero angular momentum $l \neq 0$.

7. Black holes are linked with many of the most energetic phenomena observed in the cosmos. What is the energy source associated with a black hole that can power such phenomena?
8. List three or more astrophysical phenomena that are thought to be associated with black holes.
9. Explain why one expects that stars with a final mass $> 3M_\odot$ after they have burnt out must undergo gravitational collapse into black holes?
10. There is a correspondence between black hole physics and the laws of thermodynamics. While it is not surprising that black hole mass behaves like energy, what properties of the black hole behave like temperature and entropy?
11. Hawking radiation is understood in the context of quantum mechanics applied to black hole physics. But we say it only represents a partial union of quantum theory and GR. Why so?