Contents

7 Black holes

7.1 Schwarzschild black hole

7.1.1 Schwarzschild coordinates

So far, we have discussed the Schwarzschild solution in the Schwarzschild coordinates, where the Einstein equation gives the solution

$$\mathrm{d}s^2 = -\left(1-\frac{r_\mathrm{s}}{r}\right)\mathrm{d}t^2 + \frac{1}{1-\frac{r_\mathrm{s}}{r}}\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2\right) . \tag{7.1}$$

Some components of this metric diverge or vanish at $r = r_\mathrm{s}$. This is not relevant if the matter generating the gravitational field extends to $r > r_\mathrm{s}$. But what if the body of matter is more compact – or there is no matter at all? We have already seen that a geodesic observer falling down reaches r_s in a finite amount of proper time measured by their own clock. However, it takes infinitely long for matter to fall down to r_s from the point of view of stationary observers, and signals sent up lose energy without limit as the sender is closer to $r = r_\mathrm{s}$. Does spacetime continue beyond r_s and what does it look like?

In local orthonormal coordinates (discussed in section 2.4.6) attached to an in-falling geodesic observer, the non-zero components of the Riemann tensor are

$$R_{\hat{0}\hat{1}\hat{0}\hat{1}} = -\frac{r_s}{r^3} \tag{7.2}$$

$$R_{\hat{0}\hat{2}\hat{0}\hat{2}} = R_{\hat{0}\hat{3}\hat{0}\hat{3}} = \frac{1}{2}\frac{r_s}{r^3} \tag{7.3}$$

$$R_{\hat{1}\hat{2}\hat{1}\hat{2}} = R_{\hat{1}\hat{3}\hat{1}\hat{3}} = -\frac{1}{2}\frac{r_s}{r^3} \tag{7.4}$$

$$R_{\hat{2}\hat{3}\hat{2}\hat{3}} = \frac{r_s}{r^3}\,. \tag{7.5}$$

So gravitational effects remain finite as the observer approaches r_s. In contrast, the curvature appears to grow without bound when approaching $r = 0$ – however, as it is not obvious that we can continue the coordinate system (7.1) below r_s, we have to go to some coordinates that are smooth at r_s and reach down to $r = 0$ to check this.

Let us preview the results: the passage through $r = r_s$ is indeed smooth, but it is not possible to turn back. The surface at $r = r_s$ is called the **event horizon** (see below for details). The curvature indeed diverges at $r = 0$, where there is a **singularity**. The part of spacetime inside the event horizon is called a **black hole**. Let us explain these terms a little.

Let us first define an event horizon. Consider a timelike curve that is maximally extended in both directions. Consider the union of all points from which null and timelike lines can intersect this curve, i.e. the region from which the observer can receive signals. The boundary of this region (if it exists) is the event horizon. As this definition makes clear, an event horizon is observer-dependent. For example, an observer travelling at constant acceleration in Minkowski space has an event horizon, even though the spacetime is flat.

Now for the singularity. Properly speaking, a singularity is not part of the spacetime. In the case of the Schwarzschild black hole we are considering, the curvature diverges at $r = 0$, which means we cannot describe spacetime at $r = 0$. (For example, a hypersurface with divergences is in general not a well-defined initial value surface for the Einstein equation, so we cannot continue time evolution beyond it.) The region where $r = 0$ is thus not part of the spacetime. The presence of absence is visible in the fact that radial geodesics cannot be maximally extended. In general, a singularity is an incompleteness of spacetime, signalled by the fact that timelike or null geodesics cannot be continue beyond some finite value of the affine parameter. If a component of the Riemann tensor or its covariant derivatives grows without bound as the geodesic approaches its terminus, we have a **curvature singularity**. If this does not happen, we have a **non-curvature singularity**.

Enough with the preliminaries, let's go below the Schwarzschild radius.

7.1.2 Tortoise coordinates

We want to switch to coordinates that are regular at $r = r_s$. We also want to understand the causal structure of spacetime, for which it convenient to use coordinates such that light travels at 45° angles. Let us implement the second property first. In the Schwarzschild coordinates, the equation $\mathrm{d}s^2 = 0$ for null lines gives, for radial motion,

$$\mathrm{d}t = \pm \frac{\mathrm{d}r}{1 - \frac{r_\mathrm{s}}{r}} \Rightarrow t = \pm \left[r + r_\mathrm{s} \ln \left(\frac{r}{r_\mathrm{s}} - 1 \right) \right] + \text{constant} \ . \tag{7.6}$$

We define a new radial coordinate r^* such that light travels at 45° angle, $t = \pm r^*$:

$$r^* \equiv r + r_\mathrm{s} \ln \left(\frac{r}{r_\mathrm{s}} - 1 \right) \ , \tag{7.7}$$

so the metric becomes

$$\mathrm{d}s^2 = \left(1 - \frac{r_\mathrm{s}}{r} \right) \left(-\mathrm{d}t^2 + \mathrm{d}r^{*2} \right) + r^2 \mathrm{d}\Omega^2 \ , \tag{7.8}$$

where we have denoted $\mathrm{d}\Omega^2 \equiv \mathrm{d}\theta^2 + \sin^2\theta \mathrm{d}\varphi^2$. The radial coordinate r^* is called the **tortoise coordinate**. Note that r is no longer a coordinate, but a function of r^* implicitly defined by (7.7).

Unlike in the Schwarzschild coordinates, in the tortoise coordinates the metric does not diverge anywhere (although some components still vanish at $r = r_\mathrm{s}$). We achieved this with because coordinate transformation (7.7) has a divergence. The radial coordinate $r = r_\mathrm{s}$ in the Schwarzschild coordinates (7.1) gets mapped to $r^* = -\infty$, so the range of the tortoise coordinate is $-\infty < r^* < \infty$. In Zeno's paradox Achilles never reaches the tortoise because the closer he approaches, the shorter the interval he moves. Likewise, the tortoise coordinates make it transparent that light never reaches the event horizon, because the closer it approaches, the shorter proper length it covers in a given time (from the point of view of a stationary observer outside).

As noted, we would like to use coordinates such that the metric does not vanish at the Schwarzschild radius, so that they apply on both sides of the event horizon. We have a few steps to go, the next being a simple rotation of the tortoise coordinates.

7.1.3 Eddington–Finkelstein coordinates

Let us bring the Schwarzschild radius back to a finite coordinate value, so that we can extend the coordinates below it. The first step is to introduce the null coordinates

$$\begin{aligned} u &\equiv t - r^* \\ v &\equiv t + r^* \ , \end{aligned} \tag{7.9}$$

with range $-\infty < u < \infty$, $-\infty < v < \infty$. This is just a 45° rotation of the (t, r^*) plane. The inverse is

$$\begin{aligned} t &= \frac{1}{2}(v + u) \\ r^* &= \frac{1}{2}(v - u) \ , \end{aligned} \tag{7.10}$$

so $\mathrm{d}t = \frac{1}{2}(\mathrm{d}v + \mathrm{d}u)$ and $\mathrm{d}r^* = \frac{1}{2}(\mathrm{d}v - \mathrm{d}u)$. Inserting these into the metric (7.8), we get

$$\mathrm{d}s^2 = -\left(1 - \frac{r_\mathrm{s}}{r} \right) \mathrm{d}u\mathrm{d}v + r(u,v)^2 \mathrm{d}\Omega^2 \ . \tag{7.11}$$

These are the **Eddington–Finkelstein coordinates**. Now r is implicitly defined by the relation

$$\frac{1}{2}(v-u) = r + r_{\rm s}\ln\left(\frac{r}{r_{\rm s}} - 1\right) . \tag{7.12}$$

Let us comment on the physical meaning of the null coordinates u and v. As (7.9) shows, $u =$ constant corresponds to $r^* = t +$ constant, i.e. a light ray that rises up. Likewise, $v =$ constant corresponds to $r^* = -t +$ constant, i.e. a light ray that falls down. So whereas r is a coordinate that labels timelike observers who stay at a constant proper distance from the centre and t is a coordinate attached to such observers at asymptotic infinity, u and v are coordinates that label out- and ingoing light rays, respectively. Such coordinates are convenient for describing radial light propagation, as $u =$ constant or $v =$ constant gives $\mathrm{d}s^2 = 0$. (Note that the metric (7.11) has $g_{00} = g_{11} = 0$.) The value $u = \infty$ corresponds to a photon rising from $r_{\rm s}$, and $v = -\infty$ corresponds to a photon falling to $r_{\rm s}$.

The problem that the metric vanishes at $r = r_{\rm s}$ has been reduced to a conformal factor in front of the $\mathrm{d}u\mathrm{d}v$ part of the metric. Our next tasks are to get rid of this zero of the metric and to bring $r = r_{\rm s}$ back to a finite coordinate value. These are achieved simultaneously by exponentiating u and v.

We define new null coordinates U and V by

$$\begin{aligned} U &\equiv -e^{-\frac{u}{2r_{\rm s}}} = -e^{\frac{r^*-t}{2r_{\rm s}}} = -\sqrt{\frac{r}{r_{\rm s}} - 1}e^{\frac{r-t}{2r_{\rm s}}} \\ V &\equiv e^{\frac{v}{2r_{\rm s}}} = e^{\frac{r^*+t}{2r_{\rm s}}} = \sqrt{\frac{r}{r_{\rm s}} - 1}e^{\frac{r+t}{2r_{\rm s}}} , \end{aligned} \tag{7.13}$$

where in the second equalities we have used (7.9) and in the third equalities we have used (7.7). We see that $u = \infty$ and $v = -\infty$ (both corresponding to $r = r_{\rm s}$), are mapped to $U = 0$ and $V = 0$, respectively. So $-\infty < U < 0$, $0 < V < \infty$. We have $\mathrm{d}U = -\frac{1}{2r_{\rm s}}U\mathrm{d}u$, $\mathrm{d}V = \frac{1}{2r_{\rm s}}V\mathrm{d}v$, so

$$\begin{aligned} \mathrm{d}u\mathrm{d}v &= -4r_{\rm s}^2\frac{\mathrm{d}U\mathrm{d}V}{UV} \\ &= 4r_{\rm s}^2 e^{\frac{u-v}{2r_{\rm s}}}\mathrm{d}U\mathrm{d}V \\ &= 4r_{\rm s}^2 e^{-\frac{r}{r_{\rm s}}}\frac{1}{\frac{r}{r_{\rm s}} - 1}\mathrm{d}U\mathrm{d}V \\ &= 4\frac{r_{\rm s}^3}{r}e^{-\frac{r}{r_{\rm s}}}\frac{1}{1-\frac{r_{\rm s}}{r}}\mathrm{d}U\mathrm{d}V . \end{aligned} \tag{7.14}$$

Inserting this into the Eddington–Finkelstein metric (7.11) gives

$$\mathrm{d}s^2 = -\frac{4r_{\rm s}^3}{r}e^{-\frac{r}{r_{\rm s}}}\mathrm{d}U\mathrm{d}V + r(U,V)^2\mathrm{d}\Omega^2 , \tag{7.15}$$

where the function $r(U,V)$ is implicitly defined by the relation

$$UV = \left(1 - \frac{r}{r_{\rm s}}\right)e^{\frac{r}{r_{\rm s}}} . \tag{7.16}$$

Now the metric is non-zero and finite everywhere except at $r = 0$. To get more intuition on the spacetime, let us switch from the null coordinates to one time and one space coordinate, i.e. rotate the spacetime back 45°.

7.1.4 Kruskal–Szekeres coordinates

We define new coordinates T and R by

$$\begin{aligned} T &\equiv \frac{1}{2}(V+U) \\ R &\equiv \frac{1}{2}(V-U) \ , \end{aligned} \tag{7.17}$$

in other words

$$\begin{aligned} U &= T - R \\ V &= T + R \ . \end{aligned} \tag{7.18}$$

Implementing this change of variables in the metric (7.15) gives us the Schwarzschild solution in the **Kruskal–Szekeres coordinates** (sometimes known just as the **Kruskal coordinates**),

$$\mathrm{d}s^2 = \frac{4r_\mathrm{s}^3}{r} e^{-\frac{r}{r_\mathrm{s}}}(-\mathrm{d}T^2 + \mathrm{d}R^2) + r^2(T,R)\mathrm{d}\Omega^2 \ . \tag{7.19}$$

Naively, the range of the new coordinates seems to be $-\infty < T < \infty$, $0 < R < \infty$, i.e. the entire left side of the (R,T) plane, constrained by the condition $r(T,R) > 0$. Given that $UV = T^2 - R^2$, we can use (7.16) to express this condition as

$$R^2 - T^2 = \left(\frac{r}{r_\mathrm{s}} - 1\right) e^{\frac{r}{r_\mathrm{s}}} > -1 \ . \tag{7.20}$$

The original Schwarzschild coordinates cover the region $r > r_\mathrm{s}$, which corresponds to $-R < T < R$, $0 < R < \infty$, a wedge that covers half of the left side of the (R,T) plane. However, now there is no obstruction to extending the coordinate system everywhere on the (R,T) plane, constrained only by the condition (7.20). The Kruskal–Szekeres coordinates cover not only the whole region defined by $0 < r < \infty$ in the Schwarzschild coordinates, but also a region whose existence is not even hinted at by the original Schwarzschild coordinate patch. Let us look at the global structure of the Schwarzschild solution, starting from the region covered by the Schwarzschild coordinates.

7.1.5 Global structure of the Schwarzschild solution

Combining the expression (7.16) for the null coordinates U and V in terms of the Schwarzschild coordinates t and r and the expression (7.17) for T and R in terms of U and V, the new coordinates T and R in the original region $r > r_\mathrm{s}$ read

$$\begin{aligned} T &= \sqrt{\frac{r}{r_\mathrm{s}} - 1}\, e^{\frac{r}{2r_\mathrm{s}}} \sinh\left(\frac{t}{2r_\mathrm{s}}\right) \\ R &= \sqrt{\frac{r}{r_\mathrm{s}} - 1}\, e^{\frac{r}{2r_\mathrm{s}}} \cosh\left(\frac{t}{2r_\mathrm{s}}\right) \ . \end{aligned} \tag{7.21}$$

This region covered by the original Schwarzschild coordinates is called region I.

Lines of constant t are mapped to straight lines with $T = R\tanh\left(\frac{t}{2r_\mathrm{s}}\right)$ that pass through the origin $(R,T) = (0,0)$. Lines of constant r are mapped to hyperbola:

$$\begin{cases} T = \text{constant} \cdot \sinh\left(\dfrac{t}{2r_\mathrm{s}}\right) \\ R = \text{constant} \cdot \cosh\left(\dfrac{t}{2r_\mathrm{s}}\right) \ . \end{cases} \tag{7.22}$$

These lines are illustrated in figure 1.

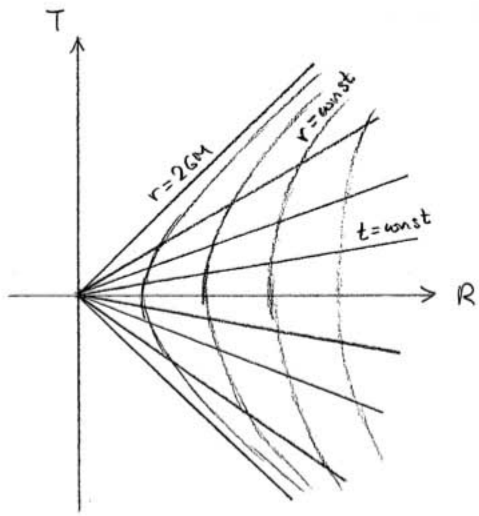

Figure 1: Region I with lines of constant t and r.

The region $r = r_\mathrm{s}$ in the Schwarzschild coordinates is mapped onto the point $(R,T) = (0,0)$. However, the lines $T = \pm R$ also correspond to $r = r_\mathrm{s}$. We approach these lines in the original coordinates when taking the simultaneous limits $t \to \pm\infty$, $r \to r_\mathrm{s}$. As the metric (7.19) neither diverges nor vanishes in any of these limits, there is no reason for the spacetime to end there. The only divergence is at $R^2 - T^2 = -1$, which corresponds to $r = 0$. So the spacetime ends at $r = 0$. But on the (R,T) plane there are two curves that solve the equation $R^2 - T^2 = -1$, not one. So the singularity at $r = 0$ is mapped both to an upper and a lower edge of the (R,T) plane.

All in all, we have expressions for T and R in terms of t and r like those in (7.21), just with different signs in front of T and/or R. This gives four regions, labeled I, II, III and IV, which together constitute the maximal extension of the Schwarzschild spacetime. The regions are illustrated in the sketch in figure 2 and the **Kruskal diagram** in figure 1.

In each of the four regions, we can reintroduce Schwarzschild coordinates:

$$\mathrm{d}s^2 = -\left(1 - \frac{r_\mathrm{s}}{r}\right)\mathrm{d}t^2 + \frac{1}{1 - \dfrac{r_\mathrm{s}}{r}}\mathrm{d}r^2 + r^2\mathrm{d}\Omega^2 \qquad \text{(regions I and IV)} \ , \tag{7.23}$$

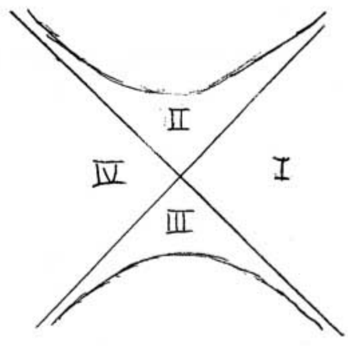

Figure 2: The four regions that make up the full Schwarzschild spacetime.

and

$$\begin{aligned} ds^2 &= -\frac{1}{\frac{r_s}{r}-1}dr^2 + \Big(\frac{r_s}{r}-1\Big)dt^2 + r^2 d\Omega^2 \\ &= -\frac{1}{\frac{r_s}{\tilde{t}}-1}d\tilde{t}^2 + \Big(\frac{r_s}{\tilde{t}}-1\Big)d\tilde{r}^2 + \tilde{t}^2 d\Omega^2 \qquad \text{(regions II and III)} \ , \quad (7.24) \end{aligned}$$

where we have on the second line relabeled $\tilde{t} \equiv r$ and $\tilde{r} \equiv t$ to make it more transparent that the coordinate r is timelike and t is spacelike (meaning the curve where only r varies is timelike and the curve where only t varies is spacelike). The range of the new time coordinate is $0 < \tilde{t} < r_s$. So we can use the original Schwarschild coordinates, properly interpreted, everywhere, but we need two copies to cover the entire spacetime. Let us now look at the different regions.

Region I is the original spherically symmetric static spacetime, with $r > r_s$. The original Schwarzschild coordinates (t, r) cover only this region.

Region II is the inside of the black hole. Here r is a time coordinate, so the metric depends on time. Observers in this region will move towards smaller values of r for the same reason that observers outside move towards large values of t. (As noted before, we don't actually know why this is the case.) Once you enter, you can never leave. An observer inside the Schwarzschild radius will hit the singularity in a finite time. The geometry of the spacelike slice $r = \tilde{t} =$ constant is that of a line segment times a two-sphere, which shrinks to zero and becomes singular as time r decreases to zero.

Region III is the inside of a **white hole**, which is a time-reversed black hole. You can't get in, but you will necessarily come out, into region I, II, or IV.

Region IV is another asymptotically flat spacetime, just like the original one. You cannot enter it from region I, but travellers from region I and region IV can meet each other in the black hole before encountering the singularity. Region IV is connected to region I only by spacelike curves. You end up there if you extend the $t =$ constant coordinate line beyond the $r = r_s$ line. In that case you do not find $r < r_s$, but another region with $r > r_s$. The geometry of the $t =$ constant, $\theta = \pi/2$ equatorial plane can be extended into region IV. The resulting geometry is called an

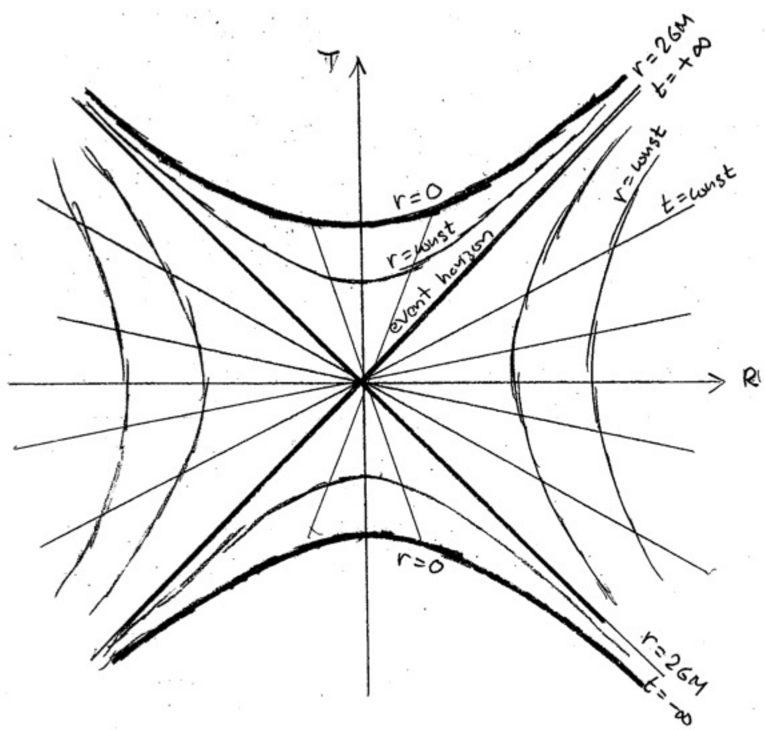

Figure 3: The Kruskal diagram. Each point is a two-sphere covered by (θ, φ).

Einstein–Rosen bridge. It cannot be travelled, as it is crossed only by spacelike curves.

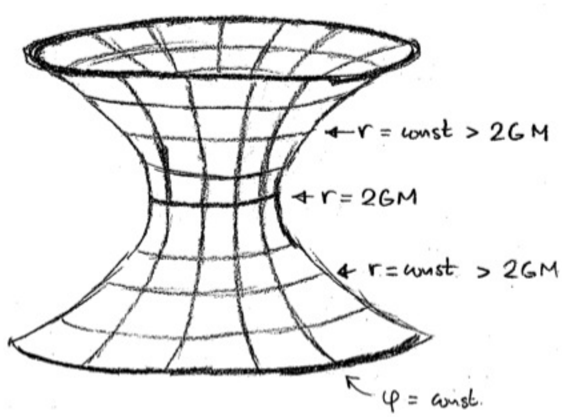

Figure 4: Einstein–Rosen bridge.

7.1.6 Penrose diagram

Penrose diagram (also called **Penrose–Carter diagram**) is a representation of the causal structure of spacetime where light travels in 45° angles and as many coordinates as possible have finite range. In the Kruskal–Szekeres coordinates, light travels at 45° angles, so all that remains to be done is to change to coordinates that have a finite range. We already played with taking the Schwarzschild radius to coordinate infinity and back. Now we want to pull physical infinity down to a finite coordinate value.

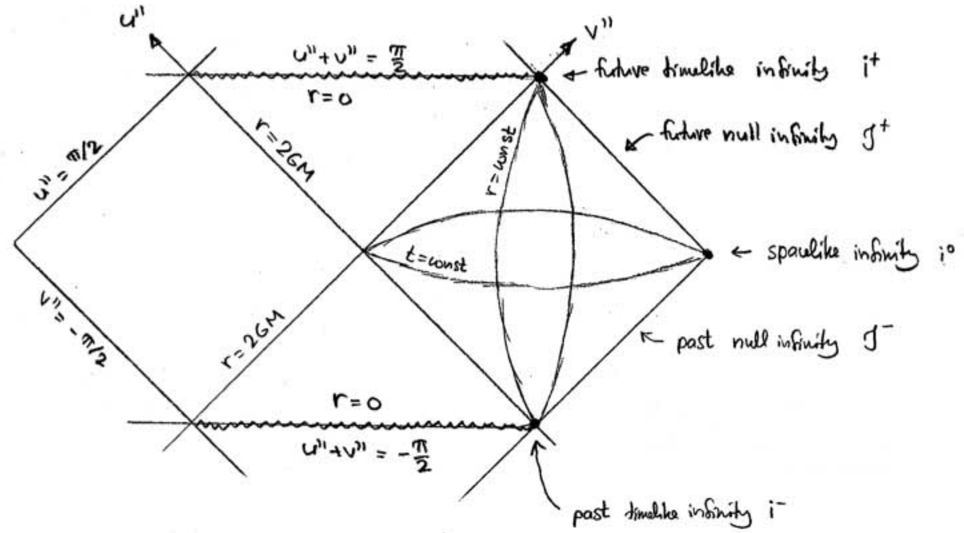

Figure 5: The Penrose diagram of the Schwarzschild spacetime. (The coordinates u'' and v'' correspond to $\tilde{U}$ and $\tilde{V}$ in the text, respectively.)

We start from the null coordinates U and V defined in (6.15)–(6.16). In the full Schwarzschild spacetime, they both range from $-\infty$ to ∞, with the constraint $UV = T^2 - R^2 < 1$. We bring this to a finite range by defining the new coordinates

$$\begin{aligned} \tilde{U} &\equiv \arctan U \\ \tilde{V} &\equiv \arctan V \ . \end{aligned} \tag{7.25}$$

This coordinate transformation maps the infinite (U, V) plane onto the finite region $\tilde{U}, \tilde{V} = (-\pi/2, \pi/2) \times (-\pi/2, \pi/2)$. The boundary $UV = 1$ maps to $\tilde{U} + \tilde{V} = \pm\pi/2$. The resulting diagram is shown in figure 5.

So the maximal extension of the Schwarzschild spacetime contains "another universe" in addition to ours (although it could be more appropriate to speak of a "causally disconnected part of the universe") and a white hole. This global structure is not obvious from the solution in the original Schwarzschild coordinates.

When solving the metric from the Einstein equation, we have to choose a coordinate system. However, there is no guarantee that this coordinate system 1) covers all of spacetime or 2) makes the spacetime symmetries transparent. In general, there is no coordinate system that does either. In the case of the Schwarzschild solution, we can choose either 1) or 2). The Kruskal–Szekeres coordinates cover the whole spacetime, but they are not independent of time; the Schwarzschild coordinates (in regions I and IV) are time-independent, but cover only part of the spacetime.

We have good reasons to think black holes are abundant in the universe. Do the causally disconnected parts of the universe and the white holes exist? Probably not, because these features of the maximal extension are not stable to small perturbations. If we add a little bit of rotation to the black hole, the global structure changes qualitatively. Also, the Schwarzschild solution is eternal both to the past and to the future, unlike real black holes. The spacetime of a real star before and after collapse into a black hole is illustrated in figure 6. Before the collapse, the Schwarzschild solution does not apply below the material surface of the star at $r_{\mathrm{B}} > r_{\mathrm{s}}$. During the collapse, an event horizon and a singularity form.

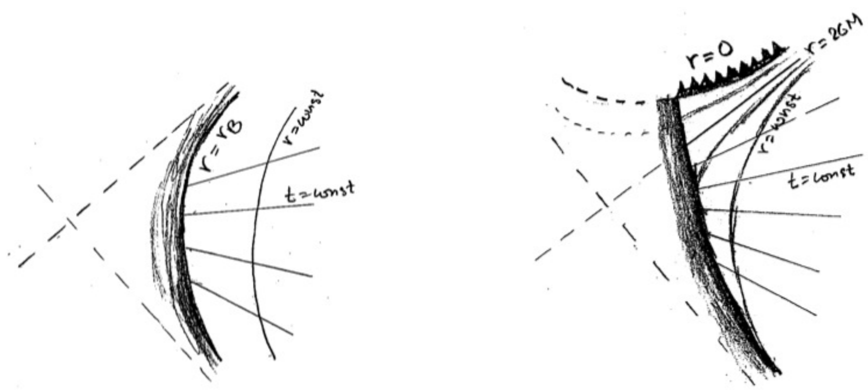

Figure 6: Spacetime around a star before and after collapse into a black hole.

Black holes are also not unchanging into the future. We will discuss Hawking radiation, via which black holes are thought to evaporate. But even classically, real black holes are not isolated, and can collide. Spacetime changes radically as black holes collide and merge, and afterwards the new black hole quickly settles to the Schwarzschild solution (or, for the general case of spinning black holes, to the rotating Kerr solution described below). The first black hole collision was observed via gravitational waves by the LIGO collaboration in 2015. A Nobel prized was awarded for this discovery in 2017. From its third observational run, which ended in March 2020, LIGO publicised 90 gravitational wave detections, shown in figure 9, the overwhelming majority of which are from colliding black hole pairs. During the ongoing fourth observation run, scheduled to go on until January 2026, the LIGO, Virgo and KAGRA detectors have observed about 10-15 events per month.

Let us first look at the solutions for rotating and charged black holes, and then comment on black holes in the real universe.

7.2 Charged and rotating black holes

7.2.1 No-hair conjecture

Black holes are not restricted to the spherically symmetric Schwarzschild case. If enough mass is brought inside a small enough radius, the result is often a singularity hidden by an event horizon. The so-called **no-hair theorem** states that in four spacetime dimensions the end result of such a collapse is characterised by just three numbers: mass M, electric charge Q and angular momentum J. If there were other long-range forces with conserved charges than electromagnetism, those charges would be added to the list. The theorem holds asymptotically in time: after a black hole forms, it settles down to such a simple solution. This has been directly observed in black hole mergers: after two black holes collide, the produced black hole quickly settles to the simple configuration, radiating away excess structure in a process called **ringdown**.

There is no full proof of the theorem, meaning a proof where you physically understand all the underlying assumptions and know what are the necessary conditions (sufficient conditions are known). So it would be more proper to call talk about the **no-hair conjecture**, but physicists are a loose bunch. Counterexamples are known in more than 4 dimensions, and also when the **null energy condition** is violated.[1]

Assuming the theorem holds, non-Schwarzschild black holes can be divided into three cases: those with $Q \neq 0, J = 0$, those with $Q = 0, J \neq 0$, and those with $Q \neq 0, J \neq 0$.

7.2.2 Reissner–Nordström metric

The simplest generalisation of the Schwarzschild black hole is the non-rotating charged black hole solution ($Q \neq 0, J = 0$), called the **Reissner–Nordström metric** after Hans Reissner (who found it in 1916) and Gunnar Nordström (who found it in 1918):

$$\mathrm{d}s^2 = -f\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2 \ , \tag{7.26}$$

where

$$f \equiv 1 - \frac{r_{\mathrm{s}}}{r} + \frac{G_{\mathrm{N}}Q^2}{r^2} \ , \tag{7.27}$$

where Q is the electric charge of the black hole, and $r_{\mathrm{s}} = 2G_{\mathrm{N}}M$ as usual. The spacetime is not empty, but contains an electromagnetic field. The gravitational field generated by the electric charge falls off faster than the gravitational field generated by mass (and faster than the electric field generated by charge). The ratio of the absolute value of the charge and mass terms in (7.27) is $2\pi\frac{r_{\mathrm{s}}}{r}Q^2\frac{M_{\mathrm{Pl}}^2}{M^2}$, where $M_{\mathrm{Pl}} \equiv 1/\sqrt{8\pi G_{\mathrm{N}}}$ is the Planck mass. For macroscopic bodies, the charge is negligible compared to the mass, and the electromagnetic contribution is tiny.

If Q is large enough, $Q > \sqrt{G_{\mathrm{N}}}M$, the metric is regular all the way to the singularity at $r = 0$. In this case there is no event horizon to veil the singularity, so rather than a black hole, we have a **naked singularity**.

7.2.3 Kerr metric

If the black hole has no electric charge but rotates ($Q = 0, J \neq 0$), it is described by the **Kerr metric**, found by Roy Kerr in 1963. In **Boyer–Lindquist coordinates**, it reads

$$\mathrm{d}s^2 = -\frac{\Delta}{\rho^2}\left(dt - a\sin^2\theta\mathrm{d}\varphi\right)^2 + \frac{\sin^2\theta}{\rho^2}\left[(r^2+a^2)\mathrm{d}\varphi - a\,dt\right]^2 + \frac{\rho^2}{\Delta}\mathrm{d}r^2 + \rho^2\mathrm{d}\theta^2 \ , \tag{7.28}$$

[1] The null energy condition states that $T_{\alpha\beta}k^\alpha k^\beta \geq 0$ for every future-pointing null vector field k^α. For an ideal fluid this is equivalent to $\rho + P \geq 0$.

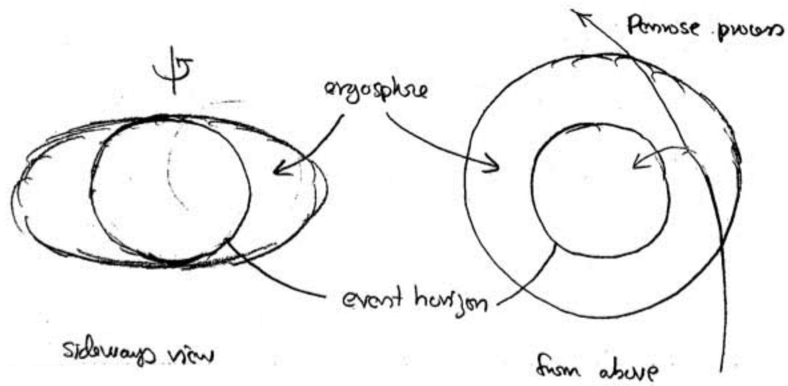

Figure 7: The Penrose process.

where

$$\begin{aligned} \Delta &\equiv r^2 - r_s r + a^2 \\ \rho^2 &\equiv r^2 + a^2\cos^2\theta \ , \end{aligned} \tag{7.29}$$

and $a \equiv J/M$ is angular momentum per mass. When the black hole rotates rapidly enough, $a > \frac{1}{2}r_s$, the singularity is naked. The Kerr spacetime is not spherically symmetric nor static, but it is axially symmetric and stationary. The latter property corresponds to the fact that even though the spacetime is rotating ($g_{t\varphi} \neq 0$), the rotation is constant ($\partial_t g_{\alpha\beta} = 0$). (We will give a coordinate-independent characterisation of stationary and static spacetimes in chapter 9.)

The maximal extension of the Schwarzschild spacetime provided a surprise in the form of the white hole and the other, causally disconnected, asymptotically flat copy of the universe. The maximal extension of the Kerr spacetime goes further. It involves an infinite sequence of asymptotically flat spacetimes, now connected by timelike curves. You can go into a black hole, emerge from a white hole, and then in that universe enter into another black hole and go to a third universe, and so on. The Kerr solution, like the Schwarzschild solution, is eternal to the past and future, so this global structure probably has as little to do with reality as the maximal extension of the Schwarzschild solution.

However, even the region of the Kerr spacetime described by the metric (7.28) has novel features. In addition to the event horizon, below which it is not possible to go up in the radial direction, it has an **ergosphere** inside which it is not possible to go in the angular direction opposite to the rotation direction of the black hole. There an observer necessarily co-rotates with the black hole. The ergosphere makes it possible to extract energy from the black hole: it is possible to go inside the ergosphere, dump part of your mass (say, a spaceship's cargo) into the black hole and come out at such a large velocity that your energy is larger than when you went in. This is illustrated in figure 7. The angular momentum and mass of the black hole correspondingly decrease. This is called the **Penrose process**, after **Roger Penrose** who suggested it in 1969. The mass of the black hole cannot be reduced arbitrarily this way, because the angular momentum will be depleted before the mass runs out.

7.2.4 Kerr–Newman metric

A charged rotating black hole ($Q \neq 0, J \neq 0$) is described by the **Kerr–Newman metric**, found by Ezra Newman in 1965. The metric has the same form as the Kerr metric (7.28), but the function Δ is different:

$$\Delta \equiv r^2 - r_{\rm s} r + a^2 + G_{\rm N} Q^2 \ . \qquad (7.30)$$

According to the no-hair conjecture, the Kerr–Newman metric is the most general isolated black hole solution asymptotically in time in 4 spacetime dimensions.

7.2.5 Time travel

The Kerr and Kerr–Newman solutions feature the exotic possibility of travelling back in time. In the case where there is a naked singularity, it is possible to go down close to the singularity, rotate in the direction opposite to the black hole, and come back up at an earlier value of t than when you left. It is possible to intersect your own geodesic (i.e. meet your past version), creating a closed timelike curve.

We have imposed the condition that motion along timelike and null lines is only possible in one direction, towards the local future. (Recall that we don't know why: GR, like electrodynamics, does not prevent sending signals to the past. This is just ruled out by hand, because we do not see it happen in the real universe.) But this local condition says nothing about the global shape of timelike or null curves. In the case of the Kerr–Newman solution with a naked singularity, they can bend backwards. However, the solution with a naked singularity is unstable to perturbations. If you take into account the effect of the traveller on the solution (instead of treating them as a test particle), the solution completely changes, and there is no time travel.

The Kerr–Newman metric is not the only known solution where time travel is possible, however. In general, any solution with rapid enough rotation enables time travel, and such solutions can be stable and can be sourced by ordinary matter.

Exercise. In the game Quantum Break, you may come across the following metric that describes a general stationary, axisymmetric spacetime with rotation, in cylindrical coordinates (ϕ has a period of 2π as usual):

$$\mathrm{d}s^2 = -A(r,z)\mathrm{d}t^2 + B(r,z)^2\mathrm{d}r^2 + C(r,z)\mathrm{d}\phi^2 + 2D(r,z)\mathrm{d}\phi\mathrm{d}t + B(r,z)^2\mathrm{d}z^2 \ . \qquad (7.31)$$

Show that an observer moving on a circular curve (with constant r and z) can go backward in time, if certain conditions on the metric functions A, C and D are satisfied. What are those conditions? Show that the maximum amount of time that an observer can travel into the past when going around one full circle is $\Delta t = \frac{2\pi|C|}{D+\sqrt{D^2-A|C|}}$. (Hints: Start by demanding that the metric is Lorentzian, $\det(g_{\alpha\beta}) < 0$. Demand then that the curve is timelike, to obtain a condition for the angular velocity $\mathrm{d}\phi/\mathrm{d}t$. Show that for certain metric functions the time gain $\mathrm{d}t$ always has the same sign as $\mathrm{d}\phi$, so going in the negative direction in ϕ means travelling back in time. Find the most negative value of Δt.)

Even though time travel is possible in GR, it is not clear whether time travel falls within the domain of validity of GR. For comparison, in Newtonian mechanics it is possible for observers to travel at arbitrarily large velocities.[2] There is no indication within Newtonian gravity that this is not possible, only special relativity shows that Newtonian solutions with velocities close to or above the speed of light do not describe reality. It may be that the same is true with regard to time travel and GR.

Stephen Hawking proposed the **chronology protection conjecture**, according to which physics beyond GR prevents time travel. This idea has been studied in **semiclassical quantum gravity**. In this theory, matter is described by quantum field theory, but spacetime is classical. The energy-momentum tensor in the Einstein equation is replaced by its expectation value. The idea of the chronology protection conjecture is that the quantum fields are affected by the possible opening of a route backwards in time, and react in a way that prevents it from opening. However, it has turned out that the details depend on the quantum state in such a way that it is impossible to say within the context of semiclassical quantum gravity whether this happens or not. The answer to the question whether time travel is possible awaits deeper understanding of quantum gravity.

7.2.6 Cosmic censorship conjecture

The **cosmic censorship conjecture** says that all singularities from collapsing matter are hidden behind event horizons. (As we will see in chapter 9, in realistic cosmological models all timelike and null geodesics trace back to the cosmological singularity at the beginning of time, which is not hidden.) We have seen that solutions with naked singularities exist, so a first attempt at formulating the theorem in a form that is not obviously false could be that naked singularities never form from collapse of matter. However, this version of the theorem is non-obviously false. It has been shown that collapsing matter can lead to a naked singularity. So a refined version of the theorem is that the cases where a naked singularity forms are a set of measure zero in the space of initial conditions. It is not known whether this version of the theorem is true or false.

Singularity theorems proven by Roger Penrose and Stephen Hawking in the 1960s show that singularities are a general feature of spacetimes that satisfy some rather general conditions, including the null energy condition, lack of symmetry and absence of closed timelike curves (i.e. going back in time). Spacetimes with singularities are the rule, not the exception. GR fails at a singularity: if a spacelike surface contains a singularity, it is not a proper initial value surface for the Einstein equation, and time evolution cannot continue. So from the point of view of GR, singularities are catastrophic. However, GR is not the last word about spacetime, just an approximation of some more general theory. We can say that GR is efficient in predicting limits to its own domain if validity. From this point of view, singularities are not a problem but an opportunity: we expect to see physics beyond GR close to where GR predicts a singularity.

[2] In fact, because the Newtonian gravitational potential is unbounded from below, it is possible to extract in infinite amount of energy and accelerate to infinite velocity in a finite time. GR cures this problem, because point particles are replaced by black holes.

7.3 Astrophysical black holes

There is a lot of observational evidence that black holes are abundant in the real universe, although exotic alternatives have been proposed. The 2021 Nobel prize in physics was awarded 1/2 to Roger Penrose for "for the discovery that black hole formation is a robust prediction of the general theory of relativity". (For the other 1/2, see below.) Observed black holes can be divided into two groups: those with masses in the range of a few to about a hundred $M_\odot$, and those with masses in the range 10^6 to 10^9 $M_\odot$.

Black holes with mass in the $M_\odot$ range are expected to form in stellar collapse: after a star exhausts its nuclear fuel, it collapses as photon pressure decreases. If the initial mass of the star is of the order 10 $M_\odot$ or more, it will end up as a black hole. The precise mass is somewhat uncertain, because part of the mass is violently ejected in the final stages of the star's active life, and how much is left to collapse depends on the details. These black holes merge to produce heavier black holes.

Black holes with masses around $10^6 M_\odot$ and higher are called **supermassive black holes** and reside in the centres of galaxies. Our Milky Way hosts a $4 \times 10^6 M_\odot$ black hole. Accreting supermassive black holes from billions of years ago are visible on the sky as **quasars**, highly energetic sources. In today's universe, they have exhausted most of the fuel in the vicinity, and are more quiet. The origin of supermassive black holes is not clear. They may start out as collapsed stars or gas clouds, and grow via mergers and accretion. However, it is not fully understood how supermassive black holes manage to grow so large so fast. One possibility is that they have been seeded by **primordial black holes**, speculative objects formed in the early universe, before the birth of star. This is an area of active investigation. It has also been suggested that **dark matter** consists partially or completely of primordial black holes. Observationally the only mass ranges that are not ruled out if black holes are to explain all of the dark matter is $10^{-16} \ldots 10^{-11} M_\odot$ (this is the mass scale of asteroids), and light relics of the order of the Planck mass, discussed below. At the moment, there is no evidence for black holes of primordial origin.

We have constraints on the spacetime around black holes both from the observation of electromagnetic radiation and from gravitational waves. For example, detailed orbits of stars around the black hole at the centre of the Milky Way have been measured for many years, see `https://galacticcenter.astro.ucla.edu/animation.html`. The observational astrophysicists Reinhard Genzel and Andrea Ghez were awarded 1/2 of the 2020 physics Nobel prize for these observations. (For the other 1/2, see above.) The closest to the event horizon we have gotten with electromagnetic waves are the pictures taken by the Event Horizon Telescope of the black hole in the centre of galaxy M87 (published in 2019) and the Milky Way (published in 2022), shown in figure 8. The M87 central black hole mass is $7 \times 10^9 M_\odot$. The event horizon is about a fifth of the size of the bright ring, smaller than the dark area in the middle of the picture. It is not possible to accurately determine the rotation parameter a from this observation. More indirect observations of other black holes give better constraints.

Direct observations of the gravitational waves from the collision, merger, and ringdown of two black holes provide a detailed picture of black holes. Simulations

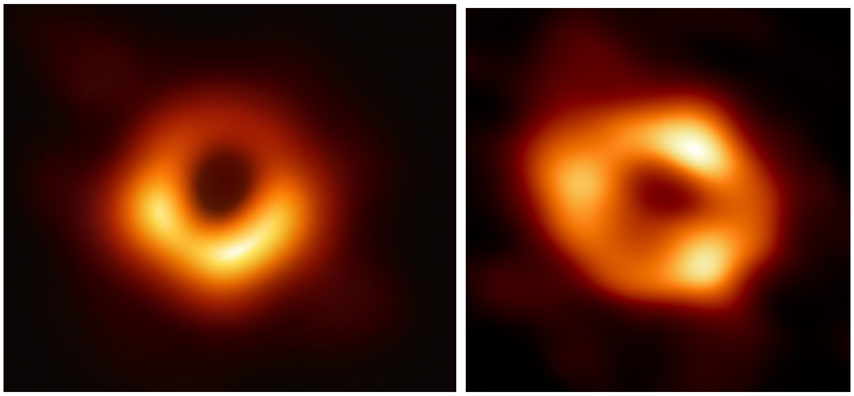

Figure 8: Left: the black hole at the centre of galaxy M87. Right: the black hole at the centre of the Milky Way. (Sources: https://eventhorizontelescope.org/press-release-april-10-2019-astronomers-capture-first-image-black-hole, https://eventhorizontelescope.org/blog/astronomers-reveal-first-image-black-hole-heart-ou

(which closely match observations) of the deformation of the event horizons and the gravitational wave signals, which we will calculate in the next chapter, can be viewed at https://www.youtube.com/user/SXSCollaboration. So far, all black holes observed are consistent with the Kerr solution. As with other objects of astrophysical size, the charge of observed black holes is negligible compared to the mass. Figure 9 shows the sources of the 90 gravity waves that the LIGO and Virgo instruments detected in their first three observing runs, involving black holes from a few to 200 solar masses.

7.4 Hawking radiation

Semiclassical quantum gravity predicts that a Schwarzschild black hole radiates like a blackbody with temperature

$$T = \frac{1}{8\pi G_{\rm N} M} = \frac{\hbar c^3}{8\pi G_{\rm N} k_{\rm B} M} = \frac{M_{\rm Pl}^2}{M} \,, \tag{7.32}$$

where in the second equality we have restored dimensional constants to illustrate how the **Hawking temperature** brings together quantum physics, relativity and thermodynamics. In Planck units, where $8\pi G_{\rm N} = 1$, we have simply $T = 1/M$. This radiation implies that black holes are not unchanging and eternal even if they are isolated, and their lifetime may be finite.

The black hole mass decreases as energy is radiated away by this **Hawking radiation**, so the temperature rises: black holes become hotter as they evaporate, not colder like normal matter. For stellar mass black holes, Hawking radiation is insignificant. However, it can be important if small primordial black holes are formed in the early universe. As the Hawking temperature rises without limit as $M \to 0$, the semiclassical calculation breaks down at some point. It is not known what is the endpoint of black hole evaporation. It is possible black holes disappear completely, or that they leave relics with mass of the order of the Planck scale. Such relics are one of the many candidates for dark matter.

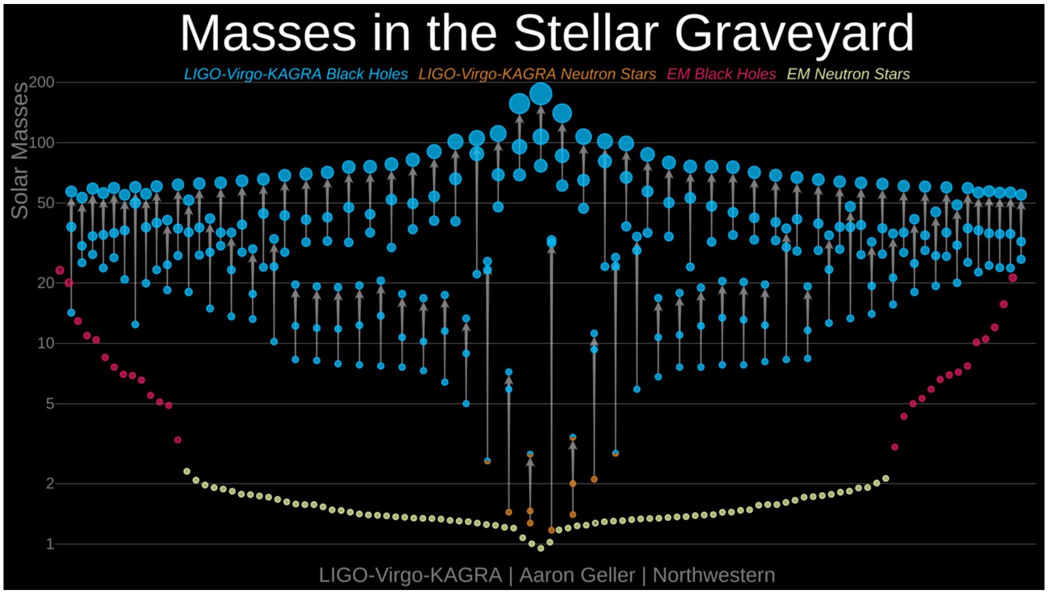

Figure 9: The 90 gravitational wave detections reported by the LIGO-Virgo-Kagra collaboration in their first three observational runs. Most of the sources are binary black holes, some are binary neutron star or black hole-neutron star pairs, and the identity of some is unclear. https://www.ligo.caltech.edu/image/ligo20211107a)

There have been suggestions to detect the analogue of Hawking radiation in fluids, where it is possible to create **dumb holes**, regions where the fluid flows so rapidly that sound waves can propagate only in one direction, so that there is an effective horizon. However, this effect, too, is quite weak and has not been observed. So Hawking radiation remains a solid prediction of semiclassical quantum gravity, but one that has not been verified by experiment. (In contrast, the approach to gravity where both matter and metric are quantised, but only perturbatively, has led to observational predictions of quantum gravity that have been confirmed in models of **cosmic inflation**.)

Exercise. a) Find the lifetime of a black hole with initial mass M.

b) What is the temperature and lifetime of a stellar scale ($M = 10M_\odot$) black hole? What is the mass and lifetime of a room temperature ($T = 295$ K) black hole?

c) If primordial black holes with different masses were produced in the very early universe (14×10^9 years ago), what is the minimum initial mass that would have survived to the present?