Black Holes

MSc in Quantum Fields and Fundamental Forces
Imperial College London

Fay Dowker

Blackett Laboratory,
Imperial College,
London, SW7 2AZ,
U.K.

Contents

Recommended Reading:

Popular/Historical:

- K.S. Thorne, *Black Holes and Time Warps*, Picador, 1994.

- W. Israel, *Dark Stars: The Evolution of an Idea*, in *Three Hundred Years of Gravitation* (S.W. Hawking & W. Israel, eds.), Cambridge University Press, 1987.

Textbooks:

- S.W. Hawking & G. F. R. Ellis, *The Large Scale Structure of Space-Time*, Cambridge University Press, 1973.

- R.M. Wald, *General Relativity*, Chicago University Press, 1984.

- S. M. Carroll, *Spacetime and Geometry: An Introduction to General Relativity*, Addison-Wesley, 2004.

Lecture Notes:

- P. Townsend, *Black Holes*. Available online at arXiv:gr-qc/9707012.

General:

- S. W. Hawking & R. Penrose, *The Nature of Space and Time*, Princeton University Press, 1996.

  This is a series of lectures by Hawking and Penrose concluded by a final debate. Hawking's part of the lectures is available online at arXiv:hep-th/9409195.

1   The Schwarzschild Black Hole

The Schwarzschild metric (1916) is a solution to the vacuum Einstein equations $R_{\mu\nu} = 0$. It is given by

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = -\left(1 - \frac{2M}{r}\right)dt^2 + \left(1 - \frac{2M}{r}\right)^{-1}dr^2 + r^2 d\Omega_2^2, \tag{1.1}$$

where $0 < r < \infty$ is a radial coordinate and $d\Omega_2^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the round metric on the two-sphere.

The line-element (1.1) is the *unique* spherically symmetric solution to the vacuum Einstein equations. This result is known as Birkhoff's theorem and it has strong implications. For example, since (1.1) is independent of t, Birkhoff's theorem implies that *any* spherically symmetric vacuum solution must be time-independent (*static*).

To study the physics of black holes, we will start with a simple model of spherically symmetric gravitational collapse: a ball of pressure-free dust that collapses under its own gravity. Since the dust ball is spherically symmetric, Birkhoff's theorem tells us that outside of it (where the energy-momentum tensor vanishes), the metric must be Schwarzschild. Since the metric is continuous, it must be Schwarzschild *on* the surface, too. Let us follow the path of a massive particle on the surface of the dust ball. Assume it follows a radial geodesic, $d\theta = d\phi = 0$. Since the pull of gravity in the r-direction is the same for all particles at a given radius, the particle will remain at the surface throughout its trajectory.

The action of a massive relativistic particle moving in a spacetime (M, g) along a trajectory $x^\mu(\tau)$ parametrised by τ is given by

$$S = \int \mathcal{L} \, d\tau = \frac{1}{2m} \int \left(g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau} - m^2\right) d\tau. \tag{1.2}$$

An implicit assumption in this form of the action is that the parameter τ is equal to the proper time measured along the particle's worldline (more on this in section 3.2). For the radial trajectory of the surface particle we have $x^\mu(\tau) = (t(\tau), r(\tau), \theta_0, \phi_0)$. Define

$$R(t) \equiv r(\tau(t)), \tag{1.3}$$

the radius of the particle's position as a function of coordinate-time t. We want to find the equation of motion for $R(t)$. A simple way to get the first equation is from $d\tau^2 = -ds^2$, which implies

$$1 = \left[\left(1 - \frac{2M}{R}\right) - \left(1 - \frac{2M}{R}\right)^{-1}\left(\frac{dR}{dt}\right)^2\right]\left(\frac{dt}{d\tau}\right)^2. \tag{1.4}$$

The factor $dt/d\tau$ can be found from the Euler-Lagrange equations of the action:

$$\frac{\partial \mathcal{L}}{\partial x^\mu} - \frac{d}{d\tau}\left(\frac{\partial \mathcal{L}}{\partial(dx^\mu/d\tau)}\right) = 0. \tag{1.5}$$

The action of the dust particle in the Schwarzschild background takes the form

$$S = \frac{1}{2m}\int\left(-\left(1 - \frac{2M}{R}\right)\left(\frac{dt}{d\tau}\right)^2 + \left(1 - \frac{2M}{R}\right)^{-1}\left(\frac{dR}{d\tau}\right)^2 - m^2\right)d\tau. \tag{1.6}$$

Since it has no explicit dependence on the t-coordinate (it only depends on $dt/d\tau$), the Euler-Lagrange equation for the t-coordinate

$$\frac{d}{d\tau}\left(\frac{\partial \mathcal{L}}{\partial(dt/d\tau)}\right) = \frac{d}{d\tau}\left[-2\left(1 - \frac{2M}{R}\right)\frac{dt}{d\tau}\right] = 0 \tag{1.7}$$

reveals a constant of motion along the geodesic $x^\mu(\tau)$:

$$\epsilon = \left(1 - \frac{2M}{R}\right)\frac{dt}{d\tau} = const. \tag{1.8}$$

The constant ϵ is equal to the energy per unit rest mass of the particle, as measured by an inertial observer at $r = \infty$. To see this, recall that the energy of a relativistic particle is given by the t-component of the momentum (co-)vector $p_\mu = (-E, \mathbf{p}) = mg_{\mu\nu}\,dx^\mu/d\tau$. Then

$$E/m = -p_0/m = g_{00}\frac{dt}{d\tau} = \left(1 - \frac{2M}{R}\right)\frac{dt}{d\tau} = \epsilon > 0. \tag{1.9}$$

A particle starting at rest at infinity has $\epsilon = 1$ at $R = \infty$ and therefore $\epsilon = 1$ everywhere along its geodesic. Consequently, a particle starting at rest at some finite radius $R = R_{\max}$ must have $\epsilon < 1$. We will therefore assume $0 < \epsilon < 1$ from now on.

Plugging (1.8) into (1.4) we obtain

$$1 = \left[1 - \frac{2M}{R} - \left(1 - \frac{2M}{R}\right)^{-1}\dot{R}^2\right]\left(1 - \frac{2M}{R}\right)^{-2}\epsilon^2, \tag{1.10}$$

where we have defined $\dot{R} \equiv dR/dt$. This gives the following equation for $\dot{R}$:

$$\dot{R}^2 = \epsilon^{-2}\left(1 - \frac{2M}{R}\right)^2\left(\frac{2M}{R} - 1 + \epsilon^2\right). \tag{1.11}$$

We immediately see that $\dot{R}$ vanishes at

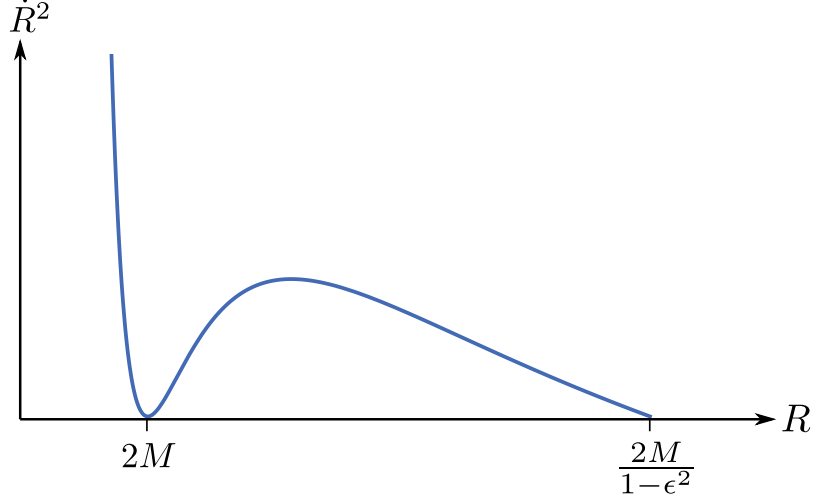$$R = R_{\max} \equiv 2M/(1 - \epsilon^2) > 2M \tag{1.12}$$

Figure 1. $\dot{R}^2$ as a function of R for the massive surface particle.

and at $R = 2M$, as illustrated in Figure 1. The zero at R_{max} simply corresponds to the initial condition that the particle starts off at rest at R_{max}. However, $\dot{R}$ approaches zero again as $R \to 2M^+$: the particle appears to *slow down* as it approaches the Schwarzschild radius.

How much time t_{2M} does it take the particle to reach $R = 2M$ from $R = R_{\text{max}}$? Integrating (1.11), we find

$$t_{2M} = \int_0^{t_{2M}} dt = -\epsilon \int_{R_{\text{max}}}^{2M} dR \left(1 - \frac{2M}{R} \right)^{-1} \left(\frac{2M}{R} - 1 + \epsilon^2 \right)^{-\frac{1}{2}} = \infty \, ! \qquad (1.13)$$

To convince yourself that this integral is infinite, look at the contribution to the integral from the region near $R = 2M$. Let $R = 2M + \rho$ with $0 < \rho \ll 2M$ and Taylor expand in ρ:

$$\begin{aligned}
t_{2M} &= \text{finite number} - \epsilon \int_\rho^0 d\rho \left(1 - \frac{2M}{2M + \rho} \right)^{-1} \left(\frac{2M}{2M + \rho} - 1 + \epsilon^2 \right)^{-\frac{1}{2}} \\
&= \text{finite number} - \epsilon \int_\rho^0 d\rho \left[\frac{2M}{\rho} + \mathcal{O}(\rho^0) \right].
\end{aligned} \qquad (1.14)$$

This is infinite since the $1/\rho$ term in the integral diverges logarithmically at 0. Hence, as far as the coordinate time t is concerned, the particle never reaches the Schwarzschid radius.

This result seems physically rather strange. The surface of a collapsing ball of dust mysteriously slows down as it approaches $R = 2M$ and never actually reaches $R = 2M$. However, notice that while t may be a meaningful measure of time for
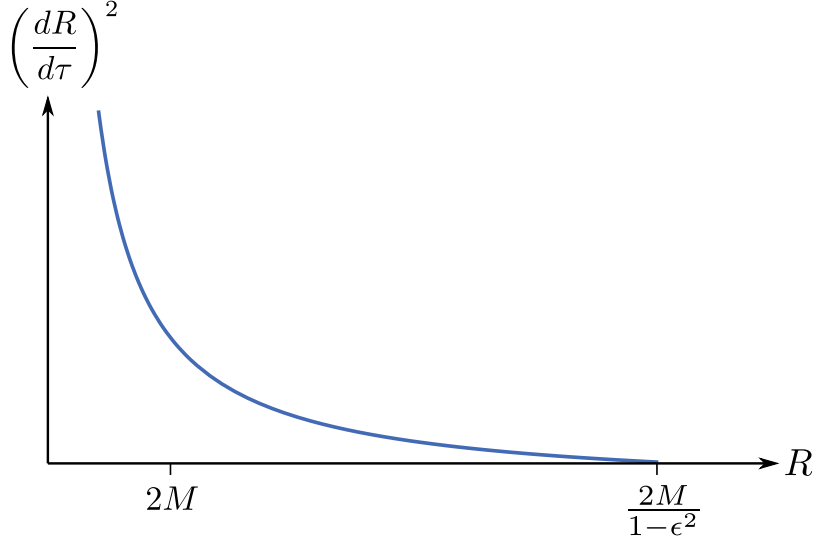
Figure 2. $(dR/d\tau)^2$ as a function of R for the massive surface particle.

inertial observers at $r = \infty$ (t is proportional to their proper time since the metric is flat at infinity), it has no physical meaning to the infalling particle itself (or to an observer falling along with it). This should not be surprising. After all, t is just a coordinate, and **coordinates in general relativity have no physical meaning.**

So let us instead work with the proper time τ of the infalling particle. How does R change as a function of τ? Using (1.8) and (1.4) we find

$$\left(\frac{dR}{d\tau}\right)^2 = (1 - \epsilon^2)\left(\frac{R_{\max}}{R} - 1\right). \tag{1.15}$$

We see that nothing curious at all happens as $R \to 2M^+$. The particle (and therefore the surface of the dust ball) passes smoothly through $R = 2M$. This is confirmed by the plot of $(dR/d\tau)^2$ against R in Figure 2.

Let us calculate the proper time τ_{2M} that elapses for the particle between $R = R_{\max}$ and $R = 2M$. Using (1.15), we find (exercise)

$$\tau_{2M} = \int_0^{\tau_{2M}} d\tau = \int_{R_{\max}}^{2M}\left(\frac{d\tau}{dR}\right) dR = \frac{2M}{(1 - \epsilon^2)^{\frac{3}{2}}}\left(\epsilon\sqrt{1 - \epsilon^2} + \arcsin \epsilon\right), \tag{1.16}$$

which is indeed finite. The strange behaviour near $R = 2M$ has disappeared — a first indication that the singularity in the metric at $R = 2M$ is due to a pathology in the time coordinate t.

The proper time taken to reach $r = 0$ is also finite:

$$\tau_0 = \int_{R_{\text{max}}}^0 \left(\frac{d\tau}{dR} \right) d\tau = \frac{\pi M}{(1 - \epsilon^2)^{\frac{3}{2}}}. \tag{1.17}$$

This means that the entire dust star collapses to the single point $r = 0$ in a finite proper time, a first sign that there is a true singularity at $r = 0$ — a *curvature singularity* — at which physical (coordinate-independent) quantities such as $R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}$ diverge.

While the calculations above have given us some useful insights, we have been a bit careless. We worked out the geodesics using the Schwarzschild metric, and traced them through $r = 2M$, even though the metric (1.1) breaks down at $r = 2M$. Furthermore, for $r < 2M$, the factor $\left(1 - \frac{2M}{r} \right)$ becomes negative, which makes t appear like a space-coordinate and r like a time-coordinate. Statements such as "the star collapses to $r = 0$ in finite time" then become somewhat suspect. To make our calculations sound, let us replace the coordinate system (t, r, θ, ϕ) by a more suitable set of coordinates.

1.1   Eddington-Finkelstein Coordinates

If we want to overcome the problems encountered in the Schwarzschild coordinate system, an obvious place to start is the time-coordinate t. We will replace t by a coordinate that is "adapted to null geodesics". For a null geodesic we have $ds^2 = 0$, which implies

$$dt^2 = \left(1 - \frac{2M}{r} \right)^{-2} dr^2. \tag{1.18}$$

First, we define the radial (Regge-Wheeler) coordinate r_* via

$$dr_*^2 \equiv \left(1 - \frac{2M}{r} \right)^{-2} dr^2, \tag{1.19}$$

so that null geodesics obey the simple equation $dt^2 = dr_*^2$. Solving (1.19) and requiring r_* to be real and increasing with r, we obtain

$$r_* = r + 2M \ln \left(\frac{r - 2M}{2M} \right). \tag{1.20}$$

The range $2M < r < \infty$ corresponds to $-\infty < r_* < \infty$, since $r_* \to -\infty$ as $r \to 2M^+$. For this reason, r_* is also known as a *tortoise coordinate*: as we approach the Schwarzschild radius, r changes more and more slowly with r_* since $dr/dr_* \to 0$.

The solutions $t \pm r_* = const.$ correspond to ingoing and outgoing null geodesics. Let us define a new pair of *lightcone coordinates*:

$$
\begin{aligned}
u &= t - r_* \\
v &= t + r_*,
\end{aligned}
\tag{1.21}
$$

such that lines of constant v correspond to ingoing null geodesics and lines of constant u to outgoing null geodesics.

1.1.1 Ingoing Eddington-Finkelstein coordinates

We now make a coordinate transformation from (t, r, θ, ϕ) to (v, r, θ, ϕ), i.e. we replace t by the coordinate v that labels ingoing radial null lines. The coordinates (v, r, θ, ϕ) are called *ingoing Eddington-Finkelstein (IEF) coordinates* and the line-element in IEF coordinates reads (exercise)

$$
ds^2 = -\left(1 - \frac{2M}{r}\right) dv^2 + 2dvdr + r^2 d\Omega_2^2.
\tag{1.22}
$$

This metric is perfectly regular at $r = 2M$ (i.e. it is non-degenerate: $\det g_{\mu\nu} \neq 0$). In fact, it is fine *everywhere* except at $r = 0$. We may therefore extend the range of r across $r = 2M$ to all $r > 0$ (you can easily check that (1.22) is a solution to the vacuum Einstein equations $R_{\mu\nu} = 0$ for all $r > 0$). We can see clearly now that there is nothing peculiar about $r = 2M$; there is nothing going on in the local physics that would tell you that you are approaching or passing through the surface $r = 2M$. The original coordinates (t, r, θ, ϕ) were simply not suitable to describe the spacetime region beyond $r = 2M$ due to their bad behaviour there.

Figure 3 shows radial ingoing ($v = const.$) and outgoing ($u = const.$) null geodesics in the IEF metric. Notice how the lightcones tip over as we cross the Schwarzschild radius: while "outgoing" photons that start off ouside the horizon eventually escape to infinity, "outgoing" photons emitted at $r = 2M$ *remain* at $r = 2M$ forever.[1] "Outgoing" photons emitted at $r < 2M$ stay within $r < 2M$ and fall towards $r = 0$. As they reach $r = 2M$, their geodesics become parallel to ingoing null geodesics.

We call the surface $r = 2M$ (a two-sphere in $3 + 1$ dimensions) the *event horizon* and the region $r < 2M$ the *black hole*. The surface dust particle is massive, so its worldline must lie between ingoing and outgoing null geodesics (inside the lightcone).

---

[1]This reveals one at first perhaps counter-intuitive aspect of the horizon, which we will see in more detail in later sections: the surface $r = 2M$ is a *null* surface (null geodesics travel along it), even though it is a surface of constant r.
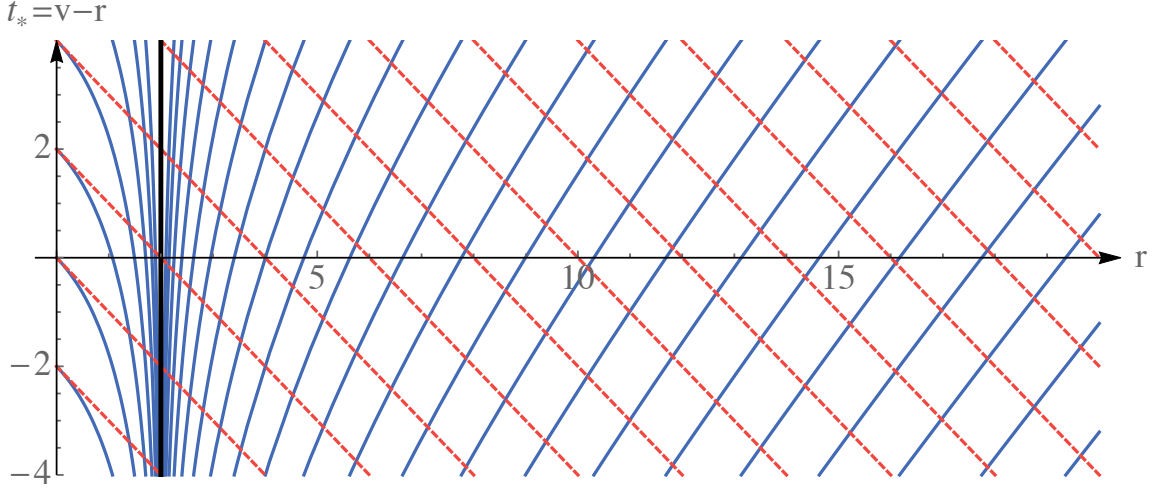
Figure 3. Ingoing (dashed red) and outgoing (solid blue) null geodesics in IEF coordinates for $M = 1$. The horizon at $r = 2M$ is highlighted in black.

As you can see in Figure 3, this means that the particle must eventually hit $r = 0$. Furthermore, no signal from an event inside the event horizon can ever escape the black hole to reach an observer at $r > 2M$.

From Figure 3 we can also get an idea how someone hovering at a fixed radius $r > 2M$ outside the black hole will perceive the infalling matter as it falls across the horizon. Imagine a signal being transmitted from the surface of the dust ball at a constant rate. Due to the bending of the lightcones in the vicinity of the horizon, the signals reaching the observer will become sparser and sparser as the surface approaches $r = 2M$. If we think of these signals as electromagnetic radiation, i.e. light, then the light is shifted toward the lower (i.e. red) end of the spectrum, and the signal received by the distant observer becomes redder and redder until it eventually disappears.

1.1.2 Outgoing Eddington-Finkelstein coordinates

In the previous section, we made a coordinate transformation by replacing t by v. Another obvious possibility is to use the outgoing EF coordinate $u = t - r_*$ instead. This leads to *outgoing Eddington-Finkelstein (OEF)* coordinates (u, r, θ, ϕ).

The metric in OEF coordinates is given by

$$ds^2 = \left(1 - \frac{2M}{r}\right) du^2 - 2dudr + r^2 d\Omega_2^2. \tag{1.23}$$
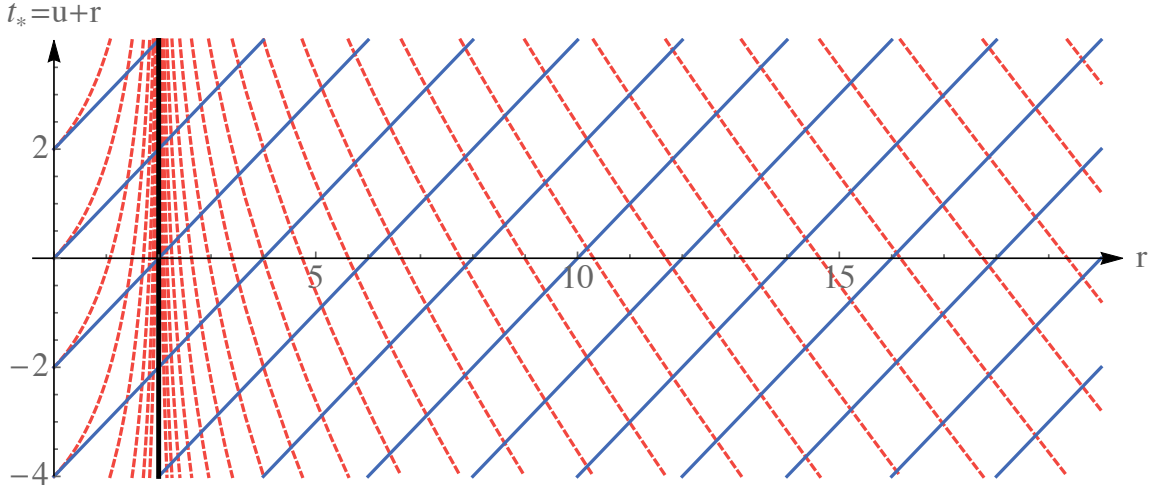
Figure 4. Ingoing (dashed red) and outgoing (solid blue) null geodesics in the OEF metric for $M = 1$. The horizon at $r = 2M$ is highlighted in black.

Note the similarity between (1.22) and (1.23): the only difference is the sign of the cross-term. However, the physical picture of the spacetime described by the line-element (1.23) turns out to be very different. To see this, let us consider again "ingoing" ($v = const.$) and "outgoing" ($u = const.$) null geodesics. They are depicted in Figure 4, where we have defined the OEF time $t_* = u + r_*$. "Ingoing" photons emitted at $r > 2M$ or $r < 2M$ never cross $r = 2M$: they approach $r = 2M$ and hover the horizon forever. Conversely, all outgoing null geodesics escape to infinity. Looking at the lightcones, we see that everything inside $r = 2M$ is ejected. The interior region $r < 2M$ is therefore called a *white hole* and $r = 2M$ the white hole horizon. It is the exact time reversal of a black hole (you can see this by turning Figure 3 upside down and comparing it with Figure 4).

What happens in the spacetime described by OEF coordinates is clearly very different from what happens in the spacetime described by IEF coordinates. This may seem contradictory, since both are obtained by coordinate transformations from the original Schwarzschild spacetime. We will learn how to make sense of this apparent contradiction by introducing yet another coordinate system.

1.2 Kruskal-Szekeres Coordinates

First, change coordinates to (u, v, θ, ϕ), replacing both Schwarzschild coordinates r and t by $u = t - r_*$ and $v = t + r_*$ (where r_* is the function of r defined in (1.20)). The metric then reads

$$ds^2 = - \left(1 - \frac{2M}{r}\right) du\, dv + r^2 d\Omega_2^2, \tag{1.24}$$

where $r = r(u, v)$ can be given in terms of u and v via (1.20) and $r_* = (v - u)/2$. This metric has a coordinate singularity at $r = 2M$, so it is only defined for $r \neq 2M$. To remove the singularity, we define a new pair of coordinates U and V in the region $r > 2M$ by

$$U = -\exp\left(-\frac{u}{4M}\right) \qquad \text{and} \qquad V = \exp\left(\frac{v}{4M}\right). \qquad (1.25)$$

Note that $U < 0$ and $V > 0$ for all values of r. The coordinates (U, V, θ, ϕ) are called *Kruskal-Szekeres (KS) coordinates*. The Schwarzschild metric in KS coordinates is (exercise):

$$ds^2 = -\frac{32M^3}{r}\exp\left(-\frac{r}{2M}\right) dU\, dV + r^2 d\Omega_2^2, \qquad (1.26)$$

where $r = r(U, V)$ is defined in terms of U and V by the implicit equation $UV = -\left(\frac{r-2M}{2M}\right)\exp\left(\frac{r}{2M}\right)$.[2]

This metric is well-defined for $r = 2M$ and indeed for all $r > 0$. Notice that for $r < 2M$, we have $UV > 0$, which is incompatible with (1.25). Let us therefore *forget* the initial definitions of U and V in terms of the Schwarzschild coordinates for now (much like we did for u and v when we went from Schwarzschild to Eddington-Finkelstein coordinates), and instead investigate the new spacetime with metric (1.26) an extended coordinate ranges $-\infty < U, V < \infty$.

Figure 5 is a picture of the spacetime described by the KS metric. Since null lines are conventionally plotted at $45°$, we define time and space coordinates $T = U + V$ and $X = T - V$, which label the vertical and horizontal axes in the figure. The U and V axes are then at $45°$ to the T and X axes. At the horizon $r = 2M$, we have $UV = 0$, which means either $U = 0$ or $V = 0$. This corresponds to the solid diagonals. The singularity $r = 0$ corresponds to the (two branches of the) hyperbola described by $UV = 1$, which is represented by a wavy line (singularities will always be represented by wavy lines). In general, surfaces of $r = const.$ correspond to hyperbolae $UV = const.$ with $UV < 1$, as shown in blue on the diagram. Spatial sections with $t = const.$ have $U/V = const.$ and $|U/V| < 1$, which corresponds to straight lines through the origin with gradient between -1 and 1. Finally, ingoing and outgoing null geodesics are respectively given by $U = const.$ and $V = const.$

We can now see the relation between the different coordinate systems. The original Schwarzschild coordinates (t, r, θ, ϕ) only cover region I. When we changed to IEF and OEF coordinates, we extended our spacetime to regions II and III, respectively. There is a new region that we only see in KS coordinates, region IV.

---

[2]We can write r explicitly as $r(U, V) = 2M\left[1 + W\left(-UV/e\right)\right]$ where e is Euler's constant and W stands for the Lambert W-function.
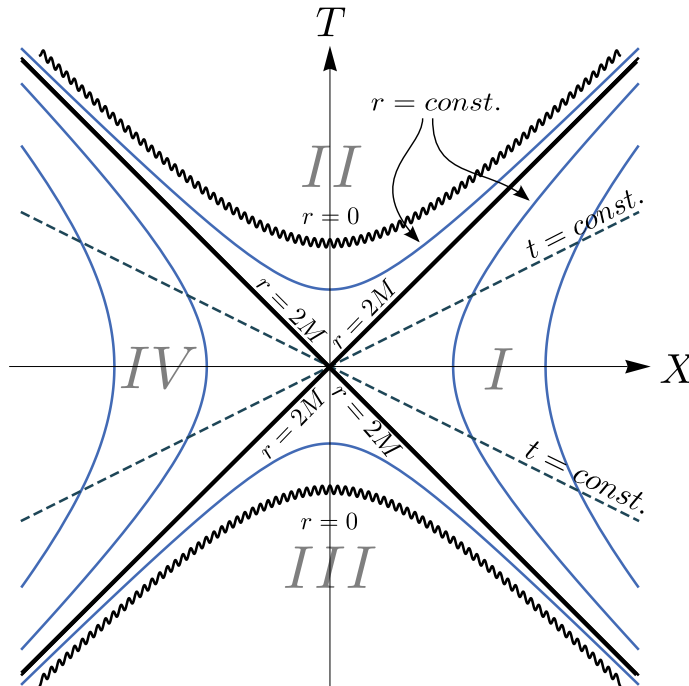
Figure 5. Kruskal spacetime. Each point represents a $2-$sphere of radius $r(T, X)$. The black lines at $45°$ correspond to $r = 2M$. The singularity $r = 0$ is represented by a wavy line. Solid blue hyperbolae correspond to $r = const.$ surfaces and straight dashed lines correspond to $t = const.$ surfaces.

This region is a mirror image of region I, which can be seen by defining

$$U = \exp\left(-\frac{u}{4M}\right) \qquad \text{and} \qquad V = -\exp\left(\frac{v}{4M}\right) \qquad (1.27)$$

on region IV and noting that the metric (1.26) will be of the same form as the Schwarzschild metric. (The proper statement is that region IV is *isometric* to region I. In general relativity, two spacetimes or regions of spacetime that are isometric are physically identical.)

The geometry of spatial sections is also worth a closer look. Consider a slice with $t = const. \iff T/X = const.$ On Figure 5, it corresponds to a straight line through the point $(T, X) = (0, 0) \iff r = 2M$, where regions I and IV attach to each other. In both regions the metric on the slice is given by

$$ds^2 = \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2 \qquad (1.28)$$

and in both regions r increases from $2M$ to ∞. For large r, we have $\left(1 - \frac{2M}{r}\right) \to 1$ and so the geometry becomes Euclidean. As we approach $r = 2M$, i.e. the centre in Figure 5, the metric (1.28) starts to deviate from the Euclidean metric. What is the
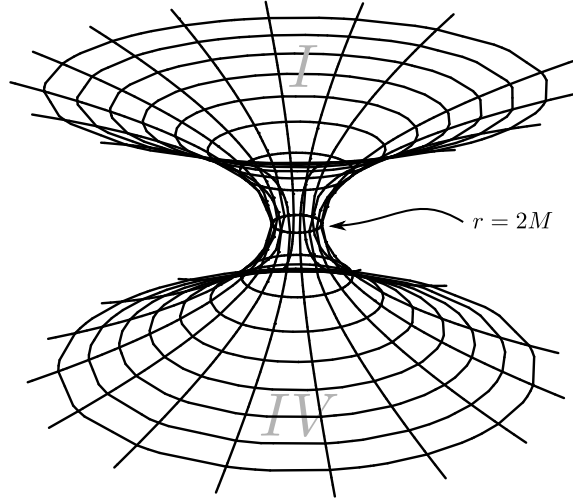
Figure 6. The Einstein-Rosen bridge with one spatial dimension suppressed. Each circle represents a two-sphere in the three-dimensional analogue.

geometry near $r = 2M$? Since we cannot draw a three-dimensional surface, let us suppress one angular coordinate by setting θ to its equatorial value $\theta = \frac{\pi}{2}$. Then the metric is $ds^2 = \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\phi^2$. It is not hard to show (exercise) that this is just the metric on a so-called quartic surface $x^2 + y^2 = (z^2/8M + 2M)^2$ embedded in three-dimensional Euclidean space $\mathbb{E}^3$ with cartesian coordinates x, y, z.[3] The geometry is shown in Figure 6, where two identical copies of the surface have been attached to each other at the circle $r = 2M$. One side corresponds to region I and the other side to region IV. This is known as the *Einstein-Rosen bridge*, which is one example of a *wormhole*. Note that no observer can ever cross the wormhole, as you can see clearly in the Kruskal diagram (and, further along in the lectures, in the Penrose diagram for Kruskal space): in order to cross the wormhole from region I to IV or vice-versa, the trajectory of the observer would have to be spacelike somewhere.

We can also get an intuitive sense for the strange sign change that appears in the original Schwarzschild metric (1.1) between $r > 2M$ and $r < 2M$, which makes r appear like a time coordinate when $r < 2M$. Indeed, if $r = 0$ were just a "position in space", as one might naively think of it, it would seem that one could simply avoid it by navigating around it. In Figure 5, we see that for anyone who has fallen across the horizon, the singularity $r = 0$ is *not* a position in space — it becomes a *moment of time*, as unavoidable as 9am tomorrow morning.

---

[3]Hint: Recall that the metric on $\mathbb{E}^3$ is $ds^2 = dx^2 + dy^2 + dz^2$. Consider the surface parametrised by $x = r\cos\phi$, $y = r\sin\phi$ and $z = \sqrt{8M(r - 2M)}$, show that it satisfies the quartic equation above and find the metric induced on it. For an illustration see http://wolfr.am/WIRjK3.

Finally, we can justify in retrospect some of the coordinate extensions (analytic continuations) that we performed rather ad-hoc in the last sections. For example, consider the extension from region I covered by Schwarzschild coordinates to regions $I + II$ covered by IEF coordinates. The original coordinate system with time coordinate $-\infty < t < \infty$ covers region I only. If we think of region I as a physical spacetime in its own right, then a particle will hit a "boundary" ($r = 2M$) in finite proper time τ_{2M}. In light of this, it seems physically quite reasonable to work instead with the spacetime $I + II$ obtained by extending the original coordinate system, where this artificial boundary disappears.

1.3   Penrose Diagrams

In this section we introduce a useful way of representing the causal structure of an infinite spacetime on a finite piece of paper. This involves performing a *conformal transformation* on the metric:

Definition. A *conformal transformation* is a map from a spacetime (M, g) to a spacetime $(M, \tilde{g})$ such that

$$\tilde{g}_{\mu\nu}(x) = \Lambda(x)^2 g_{\mu\nu}(x)$$

where $\Lambda(x)$ is a smooth function of the spacetime coordinates and $\Lambda(x) \neq 0 \; \forall \; x$.

Conformal transformations preserve the causal structure of a spacetime. To see this, consider a vector V^μ on M. Then it follows from $\Lambda(x)^2 > 0$ that

$$
\begin{aligned}
g_{\mu\nu}V^\mu V^\nu > 0 &\iff \tilde{g}_{\mu\nu}V^\mu V^\nu > 0 \\
g_{\mu\nu}V^\mu V^\nu = 0 &\iff \tilde{g}_{\mu\nu}V^\mu V^\nu = 0 \\
g_{\mu\nu}V^\mu V^\nu < 0 &\iff \tilde{g}_{\mu\nu}V^\mu V^\nu < 0.
\end{aligned}
\tag{1.29}
$$

Hence, curves that are timelike/null/spacelike with respect to g are timelike/null/spacelike with respect to $\tilde{g}$. Furthermore, by consequence of the second line, null geodesics for g correspond to null geodesics for $\tilde{g}$ (whereas timelike/spacelike geodesics for g are *not* necessarily geodesics for $\tilde{g}$).

The idea of a Penrose diagram is this. First, we use a coordinate transformation on the spacetime (M, g) to bring "infinity" to a finite coordinate distance, so that we can draw the entire spacetime on a sheet of paper. The metric will typically diverge as we approach the "points at infinity", i.e. the edges of the finite diagram. To remedy this, we perform a conformal transformation on g to obtain a new metric $\tilde{g}$ that is regular on the edges. Then $(M, \tilde{g})$ is a good representation of the original

spacetime (M, g) insofar that it has exactly the same causal structure. It is customary to add the points at infinity to M to form a new manifold $\tilde{M}$. The resulting spacetime $(\tilde{M}, \tilde{g})$ is what is called the *conformal compactification* of (M, g).

Note that the curvature tensors are in general not preserved under conformal transformations, e.g. $\tilde{R}^{\mu}_{\nu\rho\sigma} \neq R^{\mu}_{\nu\rho\sigma}$, $\tilde{R} \neq R$ etc. In that sense, the conformally compactified spacetime $(\tilde{M}, \tilde{g})$ is unphysical: it provides a good representation of the causal structure of the physical spacetime (M, g), but it should otherwise not be viewed as a picture of what is going on (for example, as noted above, the geodesics of massive test particles in $(\tilde{M}, \tilde{g})$ do not correspond to geodesics of massive test particles in (M, g)).

1.3.1 Example 1: Minkowski space in $1+1$ dimensions

The two-dimensional Minkowski metric is given by

$$ds^2 = -dt^2 + dx^2 \tag{1.30}$$

where $-\infty < t, x < \infty$. We introduce light-cone coordinates $u = t - x$ and $v = t + x$, in which the metric $g_{\mu\nu}$ takes the simple form

$$ds^2 = -dudv. \tag{1.31}$$

Now, to shrink "infinity" down to a finite coordinate distance, we define a new set of coordinates via

$$u = \tan \tilde{u}$$
$$v = \tan \tilde{v}. \tag{1.32}$$

These coordinates indeed have a finite range $-\frac{\pi}{2} < \tilde{u}, \tilde{v} < \frac{\pi}{2}$ (the range is open since points with $u, v = \tan(\pm\frac{\pi}{2}) = \pm\infty$ are *not* points in spacetime). The line-element in terms of $\tilde{u}$ and $\tilde{v}$ is

$$ds^2 = -\frac{1}{(\cos \tilde{u} \cos \tilde{v})^2} d\tilde{u}d\tilde{v}. \tag{1.33}$$

It diverges as $u, v \to \pm\infty \iff \tilde{u}, \tilde{v} \to \pm\frac{\pi}{2}$. Define a new metric $\tilde{g}_{\mu\nu}$ through a conformal transformation on $g_{\mu\nu}$:

$$d\tilde{s}^2 = (\cos \tilde{u} \cos \tilde{v})^2 ds^2 = -d\tilde{u}d\tilde{v}. \tag{1.34}$$

This metric is regular at the points at infinity where either $\tilde{u}$ or $\tilde{v}$ are equal to $\pm\frac{\pi}{2}$ and we can now add those points to the spacetime. The resulting spacetime $(\tilde{M}, \tilde{g})$ is the conformal compactification of (M, g). Both spacetimes are shown in $(\tilde{u}, \tilde{v})$ coordinates in Figure 7.
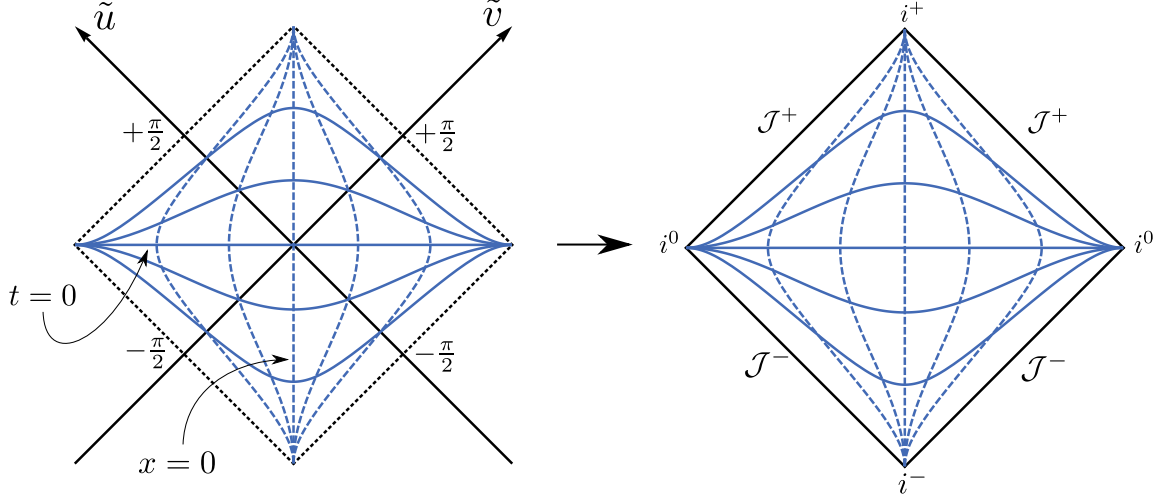
Figure 7. **Left:** Minkowski space (M, g) in $(\tilde{u}, \tilde{v})$ coordinates. The boundaries $\tilde{u}, \tilde{v} = \pm \frac{\pi}{2}$ are not part of M and g diverges there. Some timelike and spacelike geodesics of g have been included: lines with $r = const.$ are represented by dashed curves and lines with $t = const.$ are represented by solid curves. **Right:** The Penrose diagram of conformally compactified Minkowski space $(\tilde{M}, \tilde{g})$, with future/past timelike infinity $i^{\pm}$, future/past null infinity $\mathcal{J}^{\pm}$ and spacelike infinity i^0. The timelike and spacelike geodesics of Minkowski space are clearly not all geodesics for $(\tilde{M}, \tilde{g})$, since $\tilde{g}$ is flat in $(\tilde{u}, \tilde{v})$ coordinates.

The two points $(\tilde{u}, \tilde{v}) = (\frac{\pi}{2}, \frac{\pi}{2})$ and $(-\frac{\pi}{2}, -\frac{\pi}{2})$ are denoted by $i^{\pm}$. All future (past) directed timelike curves end up at i^+ (i^-), so we refer to i^+ (i^-) as *future (past) time-like infinity*. Future directed null geodesics either end up at $\tilde{v} = \frac{\pi}{2}$ with constant $|\tilde{u}| < \frac{\pi}{2}$ or at $\tilde{u} = \frac{\pi}{2}$ with constant $|\tilde{v}| < \frac{\pi}{2}$. This set of points is denoted by $\mathcal{J}^+$ ("scri-plus") and referred to as *future null infinity*. An analogous definition holds for past null infinity $\mathcal{J}^-$ ("scri-minus"). Finally, *spacelike infinity i^0* denotes the set of endpoints of spacelike geodesics, which corresponds here to $(\tilde{u}, \tilde{v}) = (\frac{\pi}{2}, -\frac{\pi}{2})$ and $(\tilde{u}, \tilde{v}) = (-\frac{\pi}{2}, \frac{\pi}{2})$.

1.3.2 Example 2: Minkowski space in $3 + 1$ dimensions

The four-dimensional Minkowski metric

$$ds^2 = -dt^2 + dx^2 + dy^2 + dz^2 \tag{1.35}$$

can be written in terms of spherical polar coordinates (t, r, θ, ϕ) by choosing an arbitrary point as the origin, say $x = y = z = 0$. We then have

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega_2^2 \tag{1.36}$$

where $d\Omega_2^2$ is the round metric on the 2-sphere. Define light-cone coordinates $u = t - r$ and $v = t + r$ and perform the same transformation (1.32) as above to bring infinity
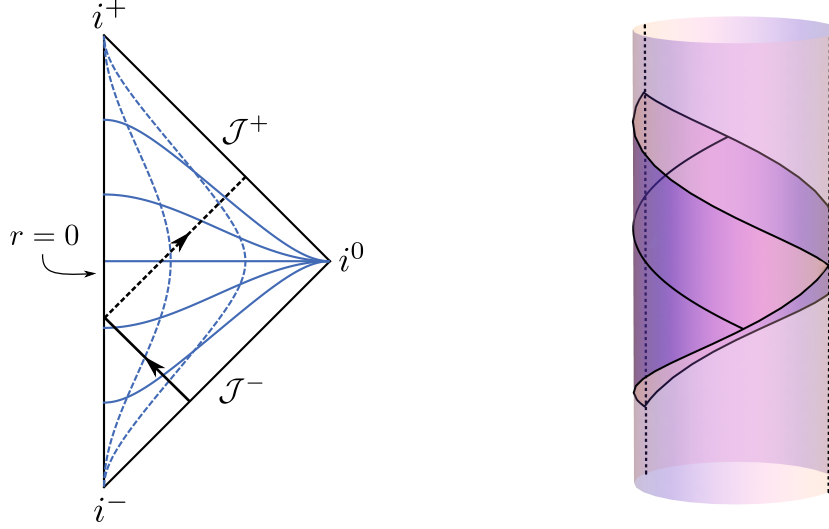
Figure 8. **Left:** The Penrose diagram of $3+1$ dimensional Minkowski space with some lines of constant r and t. Each point represents a $2-$sphere. As the null geodesic shown passes through $r = 0$, it emerges on another copy of the Penrose diagram whose points represent the antipodes on the two-spheres. **Right:** The conformal compactification drawn as a portion of the Einstein static universe with the same null geodesic.

to finite coordinate distance. The only difference to the previous analysis is that since $r \geq 0$, we have the additional constraint $u \leq v \iff \tilde{u} \leq \tilde{v}$. The metric in $(\tilde{u}, \tilde{v}, \phi, \theta)$ coordinates reads

$$ds^2 = -\frac{1}{(2\cos\tilde{u}\cos\tilde{v})^2} \left[-4d\tilde{u}d\tilde{v} + \sin^2(\tilde{v} - \tilde{u}) d\Omega_2^2 \right] \tag{1.37}$$

and its conformal compactification corresponds to the spacetime $(\tilde{M}, \tilde{g})$ with metric

$$d\tilde{s}^2 = (\cos\tilde{u}\cos\tilde{v})^2 ds^2 = -4d\tilde{u}d\tilde{v} + \sin^2(\tilde{v} - \tilde{u}) d\Omega_2^2. \tag{1.38}$$

and extended coordinate ranges $-\pi/2 \leq \tilde{u}, \tilde{v} \leq \pi/2$. On the Penrose diagram of $1+1$ dimensional Minkowski space in Figure 7, $\tilde{u} \leq \tilde{v}$ implies that we only include the part that lies to the right of the vertical line $x = 0$. The resulting diagram for the $3+1$ space is shown in Figure 8. Every point on the diagram represents a two-sphere of radius $\sin(\tilde{v} - \tilde{u})$.

There exists a somewhat more illustrative way to picture the conformal compactification. Define $\tilde{T} = \tilde{v} + \tilde{u}$ and $\chi = \tilde{v} - \tilde{u}$. The coordinate ranges are then (exercise) $-\pi < \tilde{T} < \pi$ and $0 < \chi < \pi$ and the metric reads

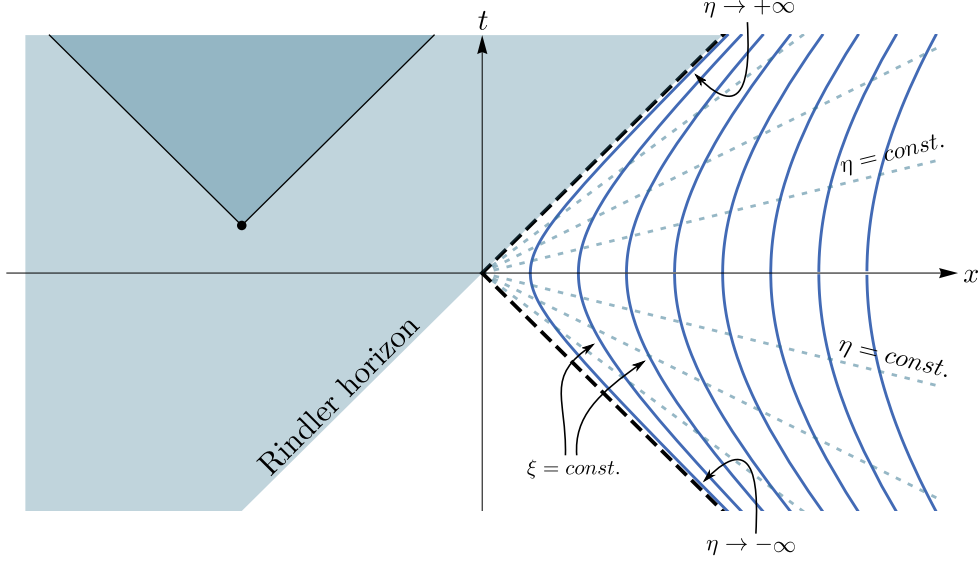$$d\tilde{s}^2 = -d\tilde{T}^2 + d\chi^2 + \sin^2\chi d\Omega_2^2. \tag{1.39}$$

Figure 9. Eternally accelerating observers in Minkowski space. Their worldlines are shown in blue and labelled by ξ. Events in the shaded region such as the black dot are hidden to them. The Rindler horizon is the boundary between the shaded and unshaded regions. Rindler space corresponds to the right wedge outlined by the dashed black null lines. The straight dotted lines are lines of constant Rindler time η.

The spatial part $d\chi^2 + \sin^2\chi \, d\Omega_2^2$ of this metric is just the round metric of a three-sphere S^3 parametrised by polar coordinates (χ, θ, ϕ). The spacetime (1.39) therefore represents a static universe with spherical spatial slices, which corresponds to a finite portion of the *Einstein static universe (ESU)* whose topology is $\mathbb{R} \times S^3$. This is shown in Figure 8, and the Penrose diagram corresponds to that part of the region wrapped around the cylinder facing out of the page.

1.3.3 Example 3: Rindler space in $1+1$ dimensions

Rindler space is a subregion of Minkowski space associated with observers eternally accelerated at a constant rate. The worldlines of these "Rindler observers" (accelerating in the positive x-direction) is given by the hyperbolae $x_\mu x^\mu = -t^2 + x^2 = \xi^2$. You can check (exercise) that the 4-acceleration $a^\mu = d^2 x^\mu / d\tau^2$ along this worldline is indeed constant: $a_\mu a^\mu = 1/\xi^2$.

Consider the one-parameter family of Rindler observers depicted in Figure 9. The region $x \leq t$ is forever hidden to them, which makes the line $x = t$ a horizon to these observers.[4] *Rindler space* corresponds to the right wedge $x > |t|$ foliated by the

---

[4]This horizon is different from the event horizon in the Schwarzschild black hole spacetime, since the Schwarzschild horizon is observer/frame-independent, while the horizon here is associated with
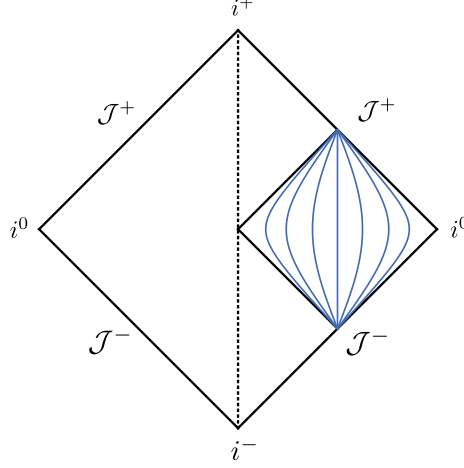
Figure 10. The Penrose diagram for $1+1$ dimensional Rindler space, seen as a portion of Minkowski space. Some accelerated worldlines (curves of constant ξ) have been drawn for $\xi = \frac{2}{9}, \frac{2}{5}, \frac{2}{3}, 1, \frac{3}{2}, \frac{5}{2}, \frac{9}{2}$. Note that all of the worldlines represent observers accelerating in the postitive $x-$direction, even though they appear to bend toward the right for $\xi > 1$. This distortion is a side effect of the coordinate transformation $(u, v) \to (\tilde{u}, \tilde{v})$.

worldlines of the accelerated observers in Figure 9. To obtain the Rindler metric, we introduce a new set of space and time coordinates ξ and η in the subregion $x > |t|$ of Minkowski space. As space coordinate we use ξ, which labels the hyperbolic worldlines $x^2 - t^2 = \xi^2$. As time coordinate we use the proper time $\eta = \tanh^{-1}(t/x)$ measured by a Rindler observer, synchronised such that $\eta = 0$ when the observer passes the x-axis. The relation to Cartesian coordinates is

$$x = \xi \cosh \eta \qquad \text{and} \qquad t = \xi \sinh \eta \tag{1.40}$$

and the coordinate ranges are $0 < \xi < \infty$ and $-\infty < \eta < \infty$. The Minkowski metric in (η, ξ) coordinates is

$$ds^2 = -\xi^2 d\eta^2 + d\xi^2. \tag{1.41}$$

The proper time measured by an accelerated observer — i.e. a stationary ($\xi = const.$) observer in Rindler coordinates — is $\tau = \xi\eta$. Since Rindler space is a subregion of Minkowski space, the Penrose diagram is just a piece of Figure 7, as shown in Figure 10.

1.3.4 Example 4: Kruskal space in $3+1$ dimensions

Recall the Kruskal metric (1.26):

$$ds^2 = -\frac{32M^3}{r} \exp\left(-\frac{r}{2M}\right) dU dV + r^2 d\Omega_2^2.$$

---

a family of special observers/frames.

Define a new set of null coordinates via $U = \tan\tilde{U}$ and $V = \tan\tilde{V}$, such that $-\frac{\pi}{2} < \tilde{U}, \tilde{V} < \frac{\pi}{2}$. Then the line-element becomes

$$ds^2 = (2\cos\tilde{U}\cos\tilde{V})^{-2}\left[-4\frac{32M^3}{r}\exp\left(-\frac{r}{2M}\right)d\tilde{U}d\tilde{V} + r^2\cos^2\tilde{U}\cos^2\tilde{V}d\Omega_2^2\right] \quad (1.42)$$

We perform the usual conformal transformation, which leaves us with

$$\begin{aligned}
d\tilde{s}^2 &= (2\cos\tilde{U}\cos\tilde{V})^2 ds^2 \\
&= -4\frac{32M^3}{r}\exp\left(-\frac{r}{2M}\right)d\tilde{U}d\tilde{V} + r^2\cos^2\tilde{U}\cos^2\tilde{V}d\Omega_2^2,
\end{aligned} \quad (1.43)$$

and we add the points at infinity. The curvature singularity $UV = 1$ now corresponds to

$$\tan\tilde{U}\tan\tilde{V} = 1 \iff \sin\tilde{U}\sin\tilde{V} = \cos\tilde{U}\cos\tilde{V} \iff \cos(\tilde{U} + \tilde{V}) = 0 \quad (1.44)$$

which implies $\tilde{U} + \tilde{V} = \pm\frac{\pi}{2}$, or $\tilde{T} = \pm\frac{\pi}{4}$ if we define $\tilde{T}$ and $\tilde{X}$ through $\tilde{U} = \tilde{T} - \tilde{X}$ and $\tilde{V} = \tilde{T} + \tilde{X}$ in analogy to section 1.2. The Penrose diagram including the points at infinity and the singularity is shown in Figure 11. Also shown is the curve corresponding to the surface of a collapsing star. If we only keep the region exterior to the surface, we obtain the Penrose diagram for the collapsing star on the right. The interior of the star is excluded since the stress-energy tensor is non-zero there and spacetime is not described by the Schwarzschild metric. Therefore, regions I and III are the only regions of Kruskal spacetime relevant to the description of a black hole formed through stellar collapse.

1.3.5 Example 5: de Sitter space in $3 + 1$ dimensions

The de Sitter metric is a solution of the Einstein equations in the presence of a positive cosmological constant $\Lambda > 0$:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = 0. \quad (1.45)$$

It corresponds to a universe with uniform positive energy density and negative pressure. To see this, note that (1.45) can be interpreted as the Einstein equations in the presence of an energy momentum tensor

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = T_{\mu\nu} \quad (1.46)$$

with $T_{\mu\nu} = -\Lambda g_{\mu\nu}$, which implies positive energy density $T_{00} = \Lambda$ and negative uniform pressure $T_{ii} = -\Lambda$ for $i = 1, 2, 3$.
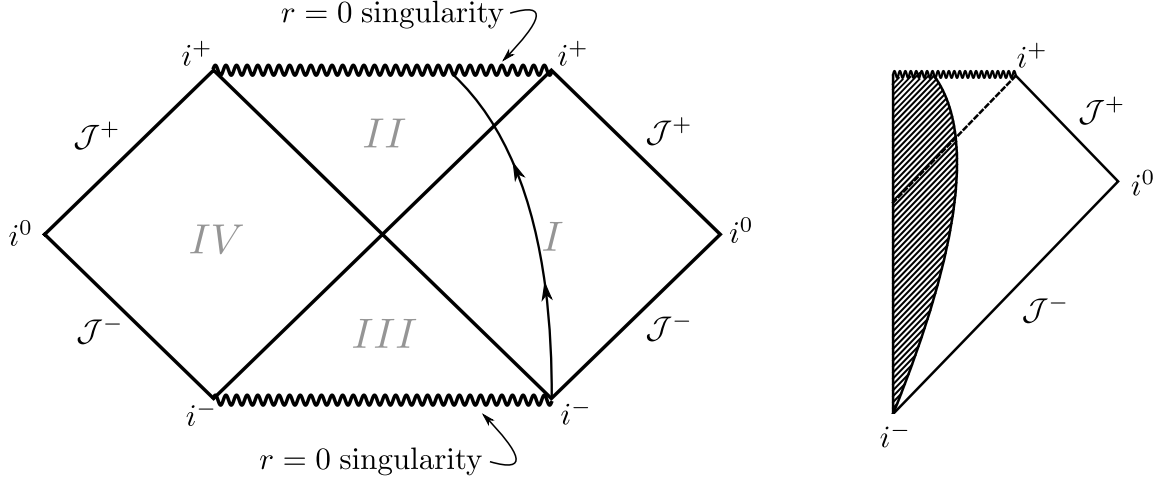
Figure 11. **Left:** The Penrose diagram for Kruskal space. The possible trajectory of the surface of a collapsing star is shown — the region to the left of the curve corresponds to the interior of the star, where spacetime is not described by the Kruskal metric. **Right:** The Penrose diagram for a collapsing star. The curved line represents the surface and the shaded region corresponds to the interior of the star. The horizon corresponds to the dashed line.

de Sitter space admits a number of different coordinate systems, some of which cover the whole space and some of which cover only part of it. The geometry of spatial sections in these coordinates can also vary (closed, open, flat). One convenient set of coordinates are closed global coordinates (t, χ, θ, ϕ) in which the line-element reads

$$ds^2 = -dt^2 + a^2 \cosh^2(t/a)d\Omega_3^2, \tag{1.47}$$

corresponding to a spacetime whose spatial sections are three-spheres with a time-dependent radius $a \cosh(t/a)$.

The best way to picture de Sitter space is as a hyperboloid embdedded in Minkowski space of one dimension higher (so $3 + 1$ dimensional de Sitter space is a hyperboloid in $4 + 1$ dimensional Minkowski space). Since our illustrations are restricted to three dimensions, let us suppress two of the angular coordinates by setting $\chi = \theta = \frac{\pi}{2}$. In that case, spatial cross sections are circles instead of three-spheres and $d\Omega_3^2 \to d\phi^2$. To see that this two-dimensional de Sitter space is a hyperboloid embedded in three-dimensional Minkowski space, consider the surface parametrised by $T = a \sinh(t/a)$, $X = a \cosh(t/a) \cos \phi$ and $Y = a \cosh(t/a) \sin \phi$, where T, X, Y are Cartesian coordinates in Minkowski space. Then it is easy to check that the embedding $2 + 1$ dimensional Minkowski metric $ds^2 = -dT^2 + dX^2 + dY^2$ reproduces (1.47) and furthermore that de Sitter space satisfies the equation of a hyperboloid: $-T^2 + X^2 + Y^2 = a^2$. This is shown in Figure 12.
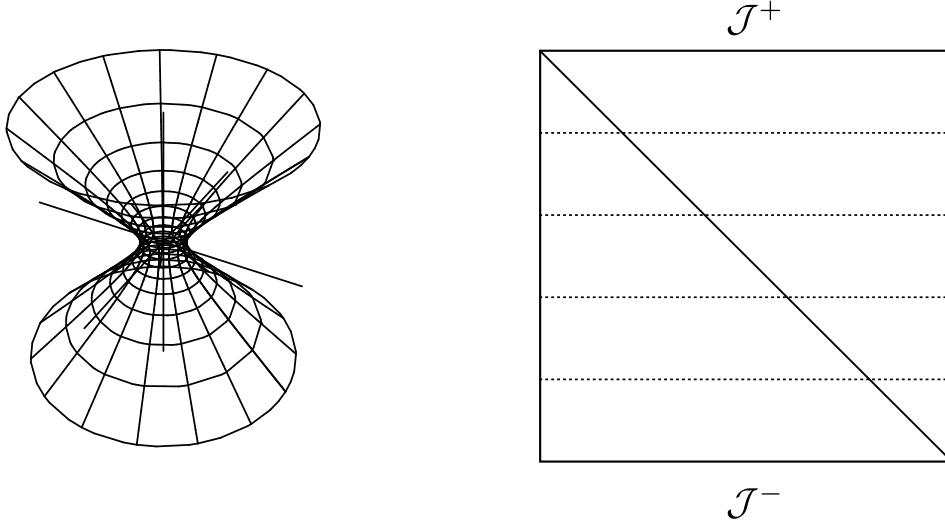
Figure 12. **Left:** The $1+1$ dimensional de Sitter hyperboloid embedded in $2+1$ dimensional Minkowski space. Some curves of constant θ and t are shown. **Right:** The Penrose diagram of de Sitter space. Dotted lines are lines of constant η. The diagonal line is the de Sitter horizon for a comoving observer at the north pole $\chi = 0$.

To construct the Penrose diagram of de Sitter space we pull out the factor of $a^2 \cosh^2(t/a)$ in the metric (1.47):

$$ds^2 = a^2 \cosh^2(t/a) \left(-d\eta^2 + d\Omega_3^2 \right), \tag{1.48}$$

and define the "conformal time coordinate" η accordingly: $d\eta^2 = dt^2/(a^2 \cosh^2(t/a))$. Integrating, we obtain $\eta = \pm 2 \tan^{-1} e^{t/a} + c$. Choose the upper sign and fix $c = -\frac{\pi}{2}$ so that η is monotonically increasing with t and has the symmetric range $\eta \in \left(-\frac{\pi}{2}, \frac{\pi}{2} \right)$. Substitution into (1.48) then yields

$$ds^2 = \frac{a^2}{\cos^2 \eta} \left(-d\eta^2 + d\Omega_3^2 \right). \tag{1.49}$$

This metric is conformal to the Einstein Static Universe, as shown in Figure 13 and the Penrose diagram is that half of the cylinder facing out of the page.

1.3.6   Example 6: Schwarzschild in $3+1$ dimensions with $M < 0$

In the Schwarzschild metric (1.1)

$$ds^2 = -\left(1 - \frac{2M}{r} \right) dt^2 + \left(1 - \frac{2M}{r} \right)^{-1} dr^2 + r^2 d\Omega_2^2$$

there is no a priori restriction on the values of M. Let's consider the case where $M < 0$ (without going into the question what "negative mass" really means). The
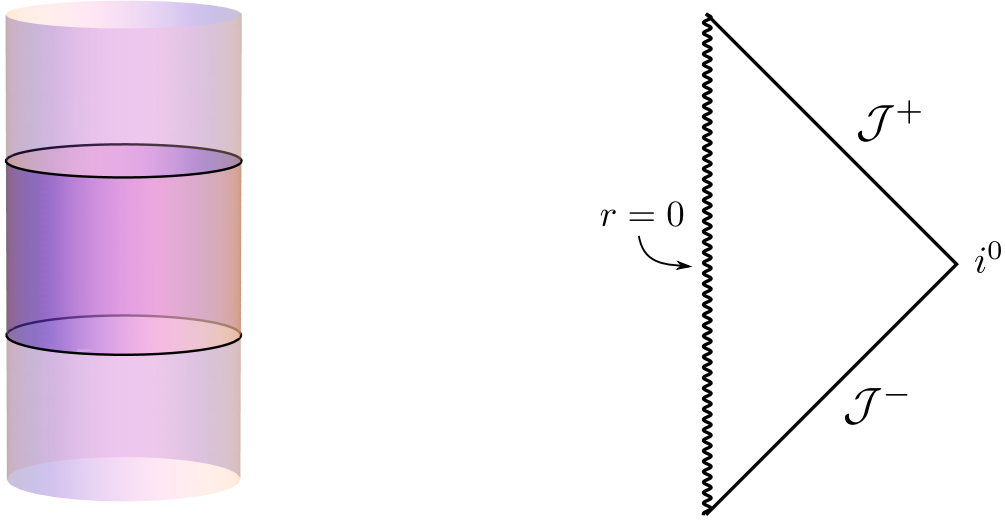
Figure 13. **Left:** The conformal compactifiction of de Sitter space viewed as a finite slab of the Einstein static universe. **Right:** The Penrose diagram for the negative mass black hole. The singularity at $r = 0$ can be seen from $\mathcal{J}^+$.

term $\left(1 - \frac{2M}{r}\right) = (1 + \frac{2|M|}{r})$ is then always positive and there are no singularities in the metric except the curvature singularity at $r = 0$.

We can construct the Penrose diagram in the same manner as in the positive mass case, by writing the metric in terms of $u = t - r_*$ and $v = t + r_*$ coordinates (with the restriction $r \geq 0 \iff v \geq u$) and proceeding with the conformal compactification as usual. The Penrose diagram is shown in Figure 13.

Notice that this singularity is different from the black hole singularity: it can be seen from $\mathcal{J}^+$. Conversely, the singularity inside the black hole is hidden behind the horizon $r = 2M$. A singularity that can be seen from $\mathcal{J}^+$ is known as a *naked singularity*. The white hole has a naked singularity too, as it can be seen in Figure 11. While naked singularities occur quite frequently in solutions to Einsteins equations, their physical status is debated. Roger Penrose has formulated the *cosmic censorship hypothesis: "Nature abhors a naked singularity"*, which encapsulates the expectation that naked singularities (except for the Big Bang) are unphysical and do not occur in the real world.

2   Charged & Rotating Black Holes

The Schwarzschild black hole is only the simplest among a number of black hole solutions to the Einstein equations. In fact, the astrophysical black holes for which

we have observational evidence all appear to be rotating, while the Schwarzschild solution has zero angular momentum. In this section we review two further, more general, black hole solutions.

2.1 The Reissner-Nordström Solution

Gravity coupled to the electromagnetic field is described by the Einstein-Maxwell action

$$S = \frac{1}{16\pi G} \int \sqrt{-g} \left(R - F_{\mu\nu} F^{\mu\nu} \right) d^4 x \tag{2.1}$$

where $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ and A_μ is the electromagnetic (four-)potential. The normalisation of the Maxwell term in (2.1) is such that the Coulomb force between two charges Q_1 and Q_2 separated by a (large enough) distance r is $G|Q_1 Q_2|/r^2$. This corresponds to *geometrised units of charge*.

The equations of motion derived from the variation of the Einstein-Maxwell action are

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 2 \left(F_{\mu\lambda} F^\lambda_{\ \nu} - \frac{1}{4} g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \right)$$
$$\nabla_\mu F^{\mu\nu} = 0. \tag{2.2}$$

They admit the spherically symmetric solution

$$ds^2 = -\left(1 - \frac{2M}{r} + \frac{Q^2}{r^2} \right) dt^2 + \left(1 - \frac{2M}{r} + \frac{Q^2}{r^2} \right)^{-1} dr^2 + r^2 d\Omega_2^2, \tag{2.3}$$

which is known as the *Reissner-Nordström solution*. The electric potential is $A_0 = \frac{Q}{r}$ and the other components of A_μ vanish. We therefore interpret Q as the charge of the black hole (by analogy with the electric potential of a point charge) and M as its mass. Without loss of generality we will assume that $Q > 0$. By a theorem analogous to Birkhoff's theorem, the Reissner-Nordström solution is the *unique* spherically symmetric solution to the Einstein-Maxwell equations.

It is convenient to introduce the function

$$\Delta = Q^2 - 2Mr + r^2 = (r - r_+)(r - r_-) \tag{2.4}$$

where $r_\pm = M \pm \sqrt{M^2 - Q^2}$. The metric then reads

$$ds^2 = -\frac{\Delta}{r^2} dt^2 + \frac{r^2}{\Delta} dr^2 + r^2 d\Omega_2^2. \tag{2.5}$$

There are three separate cases to look at: $Q > M$, $Q < M$ and $Q = M$. Let's consider them in turn.

2.1.1 Super-Extremal RN: $Q > M$

If $Q > M$ then Δ has no real roots and the metric is regular for all $r > 0$. There is a curvature singularity at $r = 0$. This is the same situation as for the negative mass black hole, and the Penrose diagram looks exactly the same (Figure 13).

Note that an electron has charge $e = 1.6 \times 10^{-19}$ C and mass $m_e = 9.1 \times 10^{-31}$ kg. In geometrised units this means $Q = e/\sqrt{4\pi\epsilon_0 G} = 1.4 \times 10^{-29}$ kg, so $Q \gg M$. Could it be that an electron is just a charged black hole? No. The electron is a quantum mechanical object, whose Compton wavelength $\lambda = h/mc = 2.4 \times 10^{-12}$ m is much larger than its Schwarzschild radius $r_s = 2Gm_e/c^2 = 1.4 \times 10^{-57}$ m.

2.1.2 Sub-Extremal RN: $Q < M$

Now Δ has two real roots $r_+ > r_-$ and there are two coordinate singularities. As always, we can remove them if we find a suitable coordinate system. Recalling our strategy with the Schwarzschild metric, let us define a tortoise coordinate r_* via

$$\frac{\Delta}{r^2} dr_*^2 = \frac{r^2}{\Delta} dr^2, \tag{2.6}$$

in terms of which the metric takes the form

$$ds^2 = -\frac{\Delta}{r^2}(dt^2 - dr_*^2) + r^2 d\Omega_2^2. \tag{2.7}$$

Radial null geodesics are then given by the simple equation $t \pm r_* = const.$ (and $\theta = \phi = const.$). A solution of (2.6) with a convenient choice of sign and integration constant is

$$r_* = r + \frac{1}{2\kappa_+} \ln\left(\frac{r - r_+}{r}\right) + \frac{1}{2\kappa_-} \ln\left(\frac{r - r_-}{r}\right), \tag{2.8}$$

where

$$\kappa_+ = \frac{r_+ - r_-}{2r_+^2} > 0 \qquad \text{and} \qquad \kappa_- = \frac{r_- - r_+}{2r_-^2} < 0. \tag{2.9}$$

Define the null coordinates $u = t - r_*$ and $v = t + r_*$ and ingoing Eddington-Finkelstein coordinates (v, r, θ, ϕ). In terms of the latter, the metric becomes

$$ds^2 = -\frac{\Delta}{r^2} dr^2 + dr dv + r^2 d\Omega_2^2, \tag{2.10}$$

which is regular for all $r > 0$, including $r = r_+$ and $r = r_-$. To understand the spacetime structure close to $r = r_\pm$ we can use two different sets of Kruskal-type coordinates at each of the two radii:

$$U^\pm = -\exp\left(-\kappa_\pm u\right) \qquad \text{and} \qquad V^\pm = \exp\left(\kappa_\pm u\right). \tag{2.11}$$

This gives rise to the Penrose diagram shown in Figure 14. Notice that a timelike trajectory can avoid $r = 0$, since the $r = 0$ singularity is timelike itself. In fact, to hit $r = 0$, one must accelerate toward it (this time it *is* like a position in space).
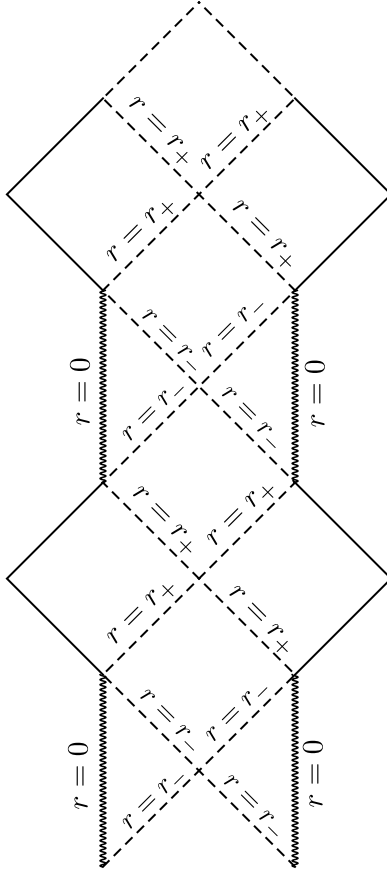
Figure 14. The Penrose diagram for the sub-extremal Reissner-Nordström solution.

2.1.3 Extremal RN: $Q = M$

The metric of the extremal Reissner-Nordström solution is

$$ds^2 = -\left(1 - \frac{M}{r}\right)^2 dt^2 + \left(1 - \frac{M}{r}\right)^{-2} dr^2 + r^2 d\Omega_2^2, \qquad (2.12)$$

which has one coordinate singularity at $r = r_+ = r_- = M$. To get rid of it, define the tortoise coordinate $dr_* = (1 - \frac{M}{r})^{-2} dr$ such that

$$ds^2 = -\left(1 - \frac{M}{r}\right)^2 (dt^2 - dr_*^2) + r^2 d\Omega_2^2, \qquad (2.13)$$

and change to ingoing Eddington-Finkelstein coordinates (v, r, θ, ϕ), where $v = t + r_*$ labels ingoing null geodesics, which should be clear from a look at (2.13). This leaves us with the improved line-element

$$ds^2 = -\left(1 - \frac{M}{r}\right) dv^2 + 2dvdr + r^2 d\Omega_2^2 \qquad (2.14)$$

which is regular at $r = M$. The inner and outer horizons have now coalesced. The result is the Penrose diagram shown in figure (to come).

2.2 Rotating Black Holes

So far we have only discussed solutions with spherical symmetry. We now introduce the *Kerr-Newman solution* (1965) to the Einstein-Maxwell equations, which describes a rotating charged black hole of mass M, charge Q and angular momentum $J \equiv Ma$ (so a is the angular momentum per unit mass). In Boyer-Lindquist coordinates (t, r, θ, ϕ), in which the black hole rotates about the polar axis, the metric part of the solution is given by

$$ds^2 = -\left(\frac{\Delta - a^2 \sin^2 \theta}{\Sigma}\right) dt^2 + \frac{\Sigma}{\Delta} dr^2 - 2\frac{a \sin^2 \theta}{\Sigma}\left(r^2 + a^2 - \Delta\right) dt d\phi$$
$$+ \Sigma d\theta^2 + \left(\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta}{\Sigma}\right) \sin^2 \theta d\phi^2 \tag{2.15}$$

where where $\Sigma = r^2 + a^2 \cos^2 \theta$ and $\Delta = r^2 - 2Mr + Q^2 + a^2$. The components of the electromagnetic potential are

$$A_t = \frac{Qr}{\Sigma}, \qquad A_\phi = -\frac{Qar \sin \theta}{\Sigma}, \qquad A_r = A_\theta = 0. \tag{2.16}$$

For $a = Q = 0$, we recover the Schwarzschild solution (1.1). For $a = 0$, we recover the Reissner-Nordström solution (2.3). Finally, the solution is symmetric under the simultaneous replacements $\phi \to -\phi$ and $a \to -a$, so we can set $a \geq 0$ without loss of generality.

When a black hole is rotating, there is no analogue of Birkhoff's theorem. This means that, during gravitational collapse with rotating matter, we cannot use the same reasoning as in the spherically symmetric case to argue that, on the surface of the collapsing matter, the metric should be of the form given above. All we can say is that, after enough time has passed and matter and spacetime have "settled down" to equilibrium, they will be described by the Kerr-Newman solution.

We will investigate the structure of the simple but illustrative special case of a rotating black hole with zero charge $Q = 0$. The metric (2.15) then reduces to the *Kerr solution* (1963):

$$ds^2 = \Sigma\left(\frac{dr^2}{\Delta} + d\theta^2\right) + (r^2 + a^2)\sin^2 \theta d\phi^2 + \frac{2Mr}{\Sigma}\left(a \sin^2 \theta d\phi - dt\right)^2 - dt^2 \tag{2.17}$$

where $\Delta = r^2 - 2Mr + a^2$ and $\Sigma = r^2 + a^2 \cos^2 \theta$. This metric is a solution to the vacuum Einstein equations. It has coordinate singularities at

$$\Delta = 0 \iff r = r_\pm = M \pm \sqrt{M^2 - a^2}. \tag{2.18}$$

Below we will show a coordinate transformation that removes them. It also has a curvature singularity at

$$\Sigma = 0 \iff r = 0 \text{ and } \cos\theta = 0. \tag{2.19}$$

The latter condition implies that the curvature singularity is only there when $\theta = \frac{\pi}{2}$, i.e. when $r = 0$ is approached along the equator. When approached from any other angle, there is no singularity at $r = 0$.

There are again three cases to consider: $M < a$, $M = a$ and $M > a$. We will concentrate on the $M > a$ solution, for which there are two coordinate singularities at r_+ (the "outer" horizon) and $r_- < r_+$ (the "inner" horizon). To remove them, we do a coordinate transformation to ingoing Kerr coordinates (v, r, θ, χ), where $v = t + r_*$ and r_* and χ are defined by

$$dr_* = \frac{r^2 + a^2}{\Delta} dr \qquad \text{and} \qquad d\chi = d\phi + \frac{a}{\Delta} dr. \tag{2.20}$$

The definition of χ implies that $\phi = const.$ does not correspond to $\chi = const.$ For example, in order to stay at $\chi = const.$ as you fall in ($dr < 0$), you need to rotate too: $d\phi = -a/\Delta dr$. In terms of ingoing Kerr coordinates the metric becomes

$$ds^2 = -\left(\frac{\Delta - a^2 \sin^2\theta}{\Sigma}\right) dv^2 + 2dvdr - 2\frac{a\sin^2\theta}{\Sigma}(r^2 + a^2 - \Delta)dvd\chi$$
$$- 2a\sin^2\theta d\chi dr + \left[\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2\theta}{\Sigma}\right]\sin^2\theta d\chi^2 + \Sigma d\theta^2. \tag{2.21}$$

There are no more factors of Δ in the numerators and the metric is regular at r_+ and r_-. The only remaining singularity is the curvature singularity at $\Sigma = 0$.

To draw the Penrose diagram is more difficult because the metric is not spherically symmetric. Since the curvature singularity at $r = 0$ only appears when $\theta = \frac{\pi}{2}$, the Penrose diagram should look very different for $\theta \neq \frac{\pi}{2}$ and $\theta = \frac{\pi}{2}$. In order to represent both cases, it is customary to draw a Penrose diagram that is an amalgam of the Penrose diagram for an observer falling in from the north pole ($\theta = 0$) and of that for an observer falling in in the equatorial plane ($\theta = \frac{\pi}{2}$) at fixed χ. Notice that $\chi = const.$ means that ϕ is not constant, so the observer falling in at $\theta = \frac{\pi}{2}$ rotates about the polar axis.

The procedure is very similar to that for the sub-extremal Reissner-Nordström solution in section 2.1.2. First, perform a coordinate transformation to coordinates (u, v, θ, ϕ) where $u = t + r_*$ and $v = t - r_*$ with r_* as defined in (2.20). Then, define Kruskal-type coordinates $U^\pm$ and $V^\pm$ close to $r = r_\pm$, respectively, and draw the
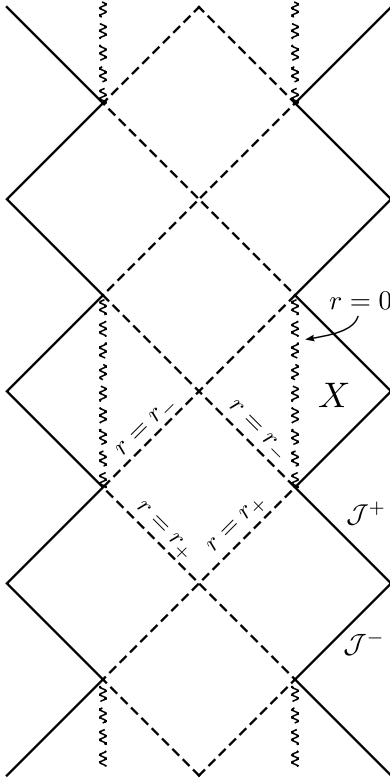
Figure 15. The Penrose diagram for the sub-extremal Kerr black hole.

Penrose diagram. This leads to the infinitely sequence of spacetime regions we saw in Figure 14. Up to this point, the analysis is identical for $\theta = 0$ and $\theta = \frac{\pi}{2}$. The only difference is that the Penrose diagram for $\theta = 0$ has a curvature singularity at $r = 0$, whereas the Penrose diagram for $\theta = \frac{\pi}{2}$ has none. In the amalgam Penrose diagram for the Kerr spacetime, we indicate this by drawing an interrupted wavy line at $r = 0$. The result is shown in Figure 15.

2.2.1 Ring singularities

How can it be that there is a singularity for $\theta = \frac{\pi}{2}$ but not otherwise? The simple explanation is that the singularity has the shape (topology) of a ring. Indeed, for fixed r, v and θ, the metric (2.17) becomes

$$ds^2 = \frac{(r^2 + a^2)^2 - \sin^2\theta(r^2 - 2Mr + a^2)a^2}{a^2\cos^2\theta}\sin^2\theta d\chi^2, \qquad (2.22)$$

which tends to $ds^2 = a^2\sin^2\theta d\chi^2$ as $r \to 0$. When $\theta = \frac{\pi}{2}$, we obtain

$$ds^2 = a^2 d\chi^2, \qquad (2.23)$$

the metric of a ring with radius a. If you travel toward $r = 0$ from any other angle than $\theta = \frac{\pi}{2}$, you will not encounter the singularity. Instead, you will fall through the interior of the "ring" and emerge in a *new* region of spacetime.

2.2.2 Closed Timelike Curves

A *closed timelike curve (CTC)*, also called a time machine, is a curve that is everywhere timelike and that eventually returns to where it started from *in spacetime*. *CTC*s have been extensively explored in the context of general relativity. In fact, they are ubiquitous and appear in a number of spacetimes that are solutions to the Einstein equations.

Kerr spacetime is one such example: region X in Figure 15 contains CTCs. To see this, consider a curve in region X at fixed t, $\theta = \frac{\pi}{2}$ and $r < 0$. Then

$$ds^2 = \left(r^2 + a^2 + \frac{2Ma}{r} \right) d\chi^2. \tag{2.24}$$

Close enough to the singularity, where r is small and negative, $r < 0$ and $|r| < 2Ma/(r^2 + a^2)$, (2.24) is negative and the curve is timelike. Since χ is a periodic coordinate with $\chi \equiv \chi + 2\pi$, the curve is also closed: it is a *CTC*.

However, it turns out that region X is unphysical. Much like in the case of the sub-extremal RN solution of section 2.1.2, the inner horizon at $r = r_-$ becomes a curvature singularity in the presence of the smallest perturbations to the Kerr metric: at the inner horizon, perturbations are infinitely blueshifted, which leads to divergences in the curvature scalars.

3 Killing Vectors & Killing Horizons

So far we have produced a number of different black hole solutions and looked at some of their features. In this section we introduce the concepts and machinery that we will need to really understand the structure of black hole spacetimes.

Notation: to denote partial derivatives we use shorthand notations such as $\partial_t = \frac{\partial}{\partial t}$ and $\partial_\mu = \frac{\partial}{\partial x^\mu}$ as well as $X_{,\mu} = \partial_\mu X$ when applied to a tensor X.

3.1 Symmetries & Killing Vectors

Let (M, g) be a Lorentzian manifold. Given a smooth vector field ξ on M, an *integral curve* of ξ is a curve $\gamma : \mathbb{R} \to M$ whose tangent vector is equal to ξ at every point

$p \in \gamma$. This can be expressed as the demand that

$$\xi_p(f) = \frac{d}{d\lambda}(f \circ \gamma(\lambda))\Big|_p \tag{3.1}$$

for all smooth functions $f : M \to \mathbb{R}$. Equivalently, given a coordinate system x^μ on M, the components of ξ in that coordinate system must satisfy

$$\xi_p^\mu = \frac{d}{d\lambda}x^\mu(\gamma(\lambda))\Big|_p, \tag{3.2}$$

where λ parametrises γ.

When ξ is smooth and everywhere non-zero, the set of integral curves form a *congruence*: every point $p \in M$ lies on a unique integral curve. Given any congruence there is an associated one-parameter family of diffeomorphisms from M onto itself, defined as follows: for each $s \in \mathbb{R}$, define $h_s : M \to M$, where $h_s(p)$ is the point parameter distance s from p along ξ, i.e. if $p = \gamma(\lambda_0)$ then $h_s(p) = \gamma(\lambda_0 + s)$. These transformations form an abelian group: the composition law is $h_s \circ h_t = h_{s+t}$, the identity is h_0 and the inverse is $(h_s)^{-1} = h_{-s}$.

This leads to the concept of the *Lie derivative* $\mathcal{L}_\xi$ along the vector field ξ. When applied to a vector V at point p it is defined as

$$(\mathcal{L}_\xi V)_p = \lim_{\delta\lambda \to 0} \frac{V_p - (h_{\delta\lambda})_* V_{h_{-\delta\lambda}(p)}}{\delta\lambda}. \tag{3.3}$$

Here $(h_s)_*$ denotes the *push-forward* associated with the group element h_s, which maps a vector defined at p to a vector defined at $h_s(p)$. It can be shown (exercise) that the Lie derivative of a vector is equal to the bracket

$$(\mathcal{L}_\xi V)_p = [\xi, V]_p$$

where $[X, Y]^\mu = X^\nu Y^\mu_{,\nu} - Y^\nu X^\mu_{,\nu}$.

The Lie derivative can be applied to any tensor on M, with appropriate definitions analogous to (3.3). In particular, given a metric tensor g on M, we can take its Lie derivative. Since the components of g transform covariantly (whereas vector fields transform contravariantly), the Lie derivative of g involves the *pull-back* h_s^*, which maps a covector at $h_s(p)$ to a covector at p:

$$(\mathcal{L}_\xi g)_p = \lim_{\delta\lambda \to 0} \frac{g_p - (h_{\delta\lambda})^* g_{h_{\delta\lambda}(p)}}{\delta\lambda}. \tag{3.4}$$

It can then be shown that

$$(\mathcal{L}_\xi g)_{\mu\nu} = \nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu. \tag{3.5}$$

If the metric does not change under the transformation h_s we say that the transformation is an *isometry* and that the metric has a *symmetry*. In this case $\mathcal{L}_\xi g = 0$, which implies

$$\nabla_\mu \xi_\nu + \nabla_\nu \xi_\mu = 0. \tag{3.6}$$

This is known as *Killing's equation* and a vector ξ that satisfies (3.6) is a *Killing vector*. Note that this equation implicitly involves the metric, which is hidden in ∇. Finding the symmetries of a spacetime amounts to finding the vectors that satisfy the Killing equation; this can be done either by inspection or by integrating (3.6).

A useful fact to know when looking for Killing vectors is the following. Given any vector field ξ we can (locally) find a coordinate system (x^1, x^2, x^3, x^4) in which ξ takes the form

$$\xi = \frac{\partial}{\partial x^1}, \tag{3.7}$$

i.e. $\xi^\mu = (1, 0, 0, 0)$. This implies that

$$(\mathcal{L}_\xi g)_{\mu\nu} = \frac{\partial g_{\mu\nu}}{\partial x^1}. \tag{3.8}$$

Hence, if we find a coordinate system in which $g_{\mu\nu}$ is independent of one of the coordinates, say y, then we know that $\frac{\partial}{\partial y}$ must be a Killing vector since

$$(\mathcal{L}_{\frac{\partial}{\partial y}} g)_{\mu\nu} = \frac{\partial g_{\mu\nu}}{\partial y} = 0. \tag{3.9}$$

The converse statement is not true: if $g_{\mu\nu}$ depends on all the coordinates, that does not mean that that g has no Killing vectors.

3.1.1 Example 1: Schwarzschild

Using the fact above, a quick look at the Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2 d\Omega_2^2$$

reveals the two Killing vectors $k = \partial_t$ and $m = \partial_\phi$. This means that Schwarzschild spacetime is *static* and *axisymmetric*:

Definition. An asymptotically flat spacetime is *stationary* if it admits a Killing vector k such that $k^2 \to -1$ asymptotically. If in addition the metric is invariant under $t \leftrightarrow -t$, the spacetime is *static*.

Definition. An asymptotically flat spacetime is *axisymmetric* if it admits a Killing vector ℓ that is spacelike asymptotically and whose integral curves are closed.

Schwarzschild spacetime is not just axisymmetric but spherically symmetric: it has two more spacelike Killing vectors, which together with ∂_ϕ generate the symmetries of a sphere — the $SO(3)$ transformations. The other two Killing vectors cannot be "read off" from the metric but they can be found by solving Killing's equation.

3.1.2 Example 2: Kerr-Newman

The Kerr-Newman solution (2.15)

$$ds^2 = -\left(\frac{\Delta - a^2 \sin^2\theta}{\Sigma}\right) dt^2 + \frac{\Sigma}{\Delta} dr^2 - 2\frac{a\sin^2\theta}{\Sigma} \left(r^2 + a^2 - \Delta\right) dt d\phi$$
$$+ \Sigma d\theta^2 + \left(\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2\theta}{\Sigma}\right) \sin^2\theta d\phi^2$$

is axisymmetric and stationary: since $g_{\mu\nu}$ is independent of t and ϕ, it has Killing vectors $k = \partial_t$ and $\ell = \partial_\phi$. However, the metric is not invariant under $t \leftrightarrow -t$ due to the $dt d\phi$ term, so it is not static.

A number of *uniqueness theorems*, proved between 1967 and 1975, have established the remarkable fact that the two-parameter Kerr family (parameters M and J) is the unique stationary asymptotically flat solution of the vacuum Einstein equations and that the three-parameter Kerr-Newman family (parameters M, J and Q) is the unique stationary asymptotically flat black hole solution of the electrovacuum Einstein-Maxwell equations.

3.2 Conservation Laws

In classical physics, the presence of symmetries is very closely tied to the existence of conservation laws. As we shall see in this section, this is also the case in general relativity. We first take a look at the geodesic motion of test particles.

Consider the action of a particle in a spacetime (M, g) moving on a curve γ with parameter λ and endpoints A and B. Pick a coordinate system x^μ and denote the coordinates of the curve by $x^\mu(\lambda)$. Then the action for γ is given by

$$I(x^\mu) = m \int d\tau = m \int_{\lambda_A}^{\lambda_B} \sqrt{-g_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda}}. \tag{3.10}$$

If we deform the curve by a small amount $\delta x^\mu(\lambda)$ and require the action to be stationary with respect to this variation, we obtain the Euler-Lagrange equations:

$$\frac{\delta I}{\delta x^\mu} = 0 \implies \nabla_{(\lambda)} \dot{x}^\mu \equiv \frac{d}{d\lambda} \dot{x}^\mu + \Gamma^\mu_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma \propto \dot{x}^\mu. \tag{3.11}$$

The dot denotes differentiation with respect to λ. The solutions to (3.11) are the *geodesics* in (M, g). If the right hand side vanishes, $\nabla_{(\lambda)}\dot{x}^\mu = 0$, the geodesic is said to be *affinely parametrised* and λ is an *affine parameter*. If we set $\lambda = \tau$ then (3.11) reduces to

$$\frac{d^2 x^\mu}{d\tau^2} + \Gamma^\mu_{\rho\sigma} \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} = 0. \tag{3.12}$$

For the purpose of finding geodesics we can cast (3.10) into a more useful form by introducing an independent function $e(\lambda)$ (called the *auxiliary field* or *einbein*):

$$I(x^\mu, e) = \frac{1}{2} \int_{\lambda_A}^{\lambda_B} d\lambda \left[e^{-1}(\lambda) g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu - m^2 e(\lambda) \right]. \tag{3.13}$$

To see that this is equivalent to (3.10), notice that from $\delta I/\delta e = 0$ we get

$$- e^{-2} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu - m^2 = 0 \implies e = \frac{1}{m} \sqrt{-g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} = \frac{1}{m} \frac{d\tau}{d\lambda} \tag{3.14}$$

and from $\delta I/\delta x^\mu$ we get

$$\frac{d}{d\lambda} \dot{x}^\mu + \Gamma^\mu_{\rho\sigma} \dot{x}^\rho \dot{x}^\sigma = (e^{-1}\dot{e})\dot{x}^\mu. \tag{3.15}$$

This equation is already the geodesic equation $\nabla_{(\lambda)}\dot{x}^\mu = (e^{-1}\dot{e})\dot{x}^\mu$ and (3.14) relates e to the choice of parameter λ. To turn this into the equation for an affinely parametrised geodesic, we need to set $\dot{e} = 0$. Then $d\tau/d\lambda = const.$, which implies that $\lambda = a\tau + b$.

Consider now an infinitesimal translation of the curve γ along a Killing vector field k (leaving e unchanged). In the coordinate chart x^μ this corresponds to

$$x^\mu \rightarrow x^\mu + \alpha k^\mu \tag{3.16}$$

where α is an infinitesimal constant. Then the action will be changed by an amount

$$\begin{aligned}
\delta I &= I(x^\mu + \alpha k^\mu, e) - I(x^\mu, e) \\
&= \frac{\alpha}{2} \int d\lambda \left[e^{-1} \left(g_{\mu\nu} \dot{k}^\mu \dot{x}^\nu + g_{\mu\nu} \dot{x}^\mu \dot{k}^\nu + g_{\mu\nu,\sigma} \dot{x}^\mu \dot{x}^\nu k^\sigma \right) \right] \\
&= \frac{\alpha}{2} \int d\lambda \left[e^{-1} \left(g_{\mu\nu} \dot{x}^\sigma \dot{x}^\nu k^\mu{}_{,\sigma} + g_{\mu\nu} \dot{x}^\mu \dot{x}^\sigma k^\nu{}_{,\sigma} + g_{\mu\nu,\sigma} \dot{x}^\mu \dot{x}^\nu k^\sigma \right) \right] \\
&= \frac{\alpha}{2} \int d\lambda \left[e^{-1} \dot{x}^\mu \dot{x}^\nu \left(g_{\sigma\nu} k^\sigma{}_{,\mu} + g_{\mu\sigma} k^\sigma{}_{,\nu} + g_{\mu\nu,\sigma} k^\sigma \right) \right] \\
&= \frac{\alpha}{2} \int d\lambda \left[e^{-1} \dot{x}^\mu \dot{x}^\nu \left(\nabla_\mu k_\nu + \nabla_\nu k_\mu \right) \right] \\
&= 0
\end{aligned} \tag{3.17}$$

where the last line follows from the fact that k satisfies Killing's equation. To get from the second to the third line we used the fact that $\dot{k}^\mu = \frac{d}{d\tau} k^\mu = \frac{dx^\nu}{d\tau} \partial_\nu k^\mu = \dot{x}^\nu k^\mu{}_{,\nu}$.

We see that k being Killing leads to a symmetry of the particle action. Associated with this symmetry is a quantity (charge) that is conserved along geodesics:

Claim. Let k be a Killing vector. Then

$$Q = k^\mu p_\mu \tag{3.18}$$

is conserved along geodesics, where p_μ is the momentum of the particle defined by

$$p_\mu = \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = e^{-1} g_{\mu\nu} \dot{x}^\nu = m g_{\mu\nu} \frac{dx^\nu}{d\tau}. \tag{3.19}$$

The last equality follows from equation (3.14) for e.

Proof. Consider a small variation $\delta x^\mu = \alpha k^\mu$ generated by the Killing vector k. Set $\lambda = \tau$ for simplicity (the proof for general λ is similar). As shown above, such variations leave the action invariant: $\delta I = 0$. Defining the Lagrangian density $\mathcal{L}$ as $I = \int \mathcal{L} d\lambda$, we then have

$$\frac{\partial \mathcal{L}}{\partial x^\mu} \delta x^\mu + \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} \delta \dot{x}^\mu = 0. \tag{3.20}$$

Using the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial x^\mu} = \frac{d}{d\tau} \frac{\partial \mathcal{L}}{\partial \dot{x}^\mu} = \dot{p}_\mu, \tag{3.21}$$

the previous equation can be written as

$$0 = \dot{p}_\mu \alpha k^\mu + p_\mu \alpha \dot{k}^\mu = \alpha \frac{d}{d\tau}(k^\mu p_\mu) = \alpha \frac{d}{d\tau} Q. \tag{3.22}$$

Therefore Q is constant along geodesics.

3.2.1 Example 1: Schwarzschild

Recall the Killing vector $k = \partial_t$ in Schwarzschild spacetime. Then $k^\mu = (1, 0, 0, 0)$ and we have

$$Q = k^\mu p_\mu = k^\mu m g_{\mu\nu} \dot{x}^\nu = m g_{00} \dot{x}^0 = -m \left(1 - \frac{2M}{r}\right) \frac{dt}{d\tau} \equiv -E, \tag{3.23}$$

where we have identified Q with minus the energy of the particle in the rest frame of the black hole. This agrees with the the definition of the constant $\epsilon = E/m$ in equation (1.9) of section 1.

The conserved quantity associated with $\ell = \partial_\phi$ is

$$Q = m g_{33} \dot{x}^3 = m r^2 \sin^2 \theta \dot{\phi} \equiv J, \tag{3.24}$$

the angular momentum of the particle along the axis of symmetry as measured by an observer at rest at infinity.

3.2.2 Example 2: Kerr

The vectors k and ℓ are also Killing vectors for the Kerr metric (2.15,2.17):

$$ds^2 = \Sigma \left(\frac{dr^2}{\Delta} + d\theta^2 \right) + (r^2 + a^2) \sin^2 \theta d\phi^2 + \frac{2Mr}{\Sigma} \left(a \sin^2 \theta d\phi - dt \right)^2 - dt^2.$$

In Schwarzschild spacetime, you (or a test particle) can travel along integral curves of $k = \partial_t$ anywhere outside the horizon — you will then appear stationary[5] to observers at infinity, since your position in space is not changing. This is possible because g_{00} is negative everywhere for $r > 2M$, so that $k^2 = g_{00}$ is negative and the integral curves of k are timelike. It turns out that this is not the case in Kerr spacetime: there is a region around the outer horizon, called the *ergosphere* or *ergoregion*, in which it is impossible for you (and any test particle) to remain stationary with respect to observers at infinity — everything rotates. This happens because

$$g_{00} = - \left(1 - \frac{2Mr}{\Sigma} \right) \tag{3.25}$$

becomes positive in the region

$$\frac{2Mr}{\Sigma} > 1 \implies \xi(r) \equiv \Sigma - 2Mr = r^2 + a^2 \cos^2 \theta - 2Mr < 0, \tag{3.26}$$

part of which lies *outside* the outer horizon $r = r_+$ when $a \neq 0$. This is easy to see by noting that the equation for $\xi(r)$ is a parabola with roots at $\tilde{r}_\pm = M \pm \sqrt{M^2 - a^2 \cos^2 \theta}$ and $\tilde{r}_+$ is bigger than $r_+ = M + \sqrt{M^2 - a^2}$ for $\theta \neq 0, \pi$. Hence g_{00} is positive in the ellipsoidal region $r_+ < r < \tilde{r}_+$, which has a maximum extent on the equator $\theta = \frac{\pi}{2}$ where $\tilde{r}_+ = M + \sqrt{M^2 + a^2}$.

In the ergoregion, orbits of ∂_t are not timelike, so you cannot travel along them and remain stationary with respect to observers at infinity. In order for a curve $x^\mu = (t, r, \theta, \phi)$ to be timelike, its tangent vector $u^\mu = dx^\mu/d\tau$ must satisfy $u^2 = -1$. But in the ergoregion, *every* term in $u^2 = g_{\mu\nu} u^\mu u^\nu$ is positive except for $g_{t\phi} u^t u^\phi$, which means that $u^\phi = d\phi/d\tau$ must be non-zero. Moreover, since $u^t > 0$ for a future-directed worldline and $g_{t\phi} < 0$, u^ϕ must be positive. Any timelike worldline is therefore dragged around in the direction of rotation of the black hole. This effect is an example of *frame dragging*.

3.2.3 Example 3: The Penrose Process

The Penrose process is a process that allows you to extract energy from a rotating black hole. Imagine sending a particle into the ergoregion. Prepare it carefully such

---

[5]The word "stationary" here has nothing to do with the definition of a stationary spacetime. When used to describe the relative motion of one observer or test particle with respect to another observer, stationary just means "not moving in space" from the point of view of the latter.

that, once in the ergoregion, it decays into two particles, one of which falls into the black hole and one of which escapes the ergoregion again. Denote the energy of the initial particle by $E = -k_\mu p^\mu$ and that of the final particles by $E_1 = -k_\mu p_1^\mu$ and $E_2 = -k_\mu p_2^\mu$, where k is the asymptotically timelike Killing vector. Conservation of four-momentum $p^\mu = p_1^\mu + p_2^\mu$ implies that $E = E_1 + E_2$. The fact that k is spacelike in the ergoregion allows you to arrange the decay such that the energy of the particle that falls into the black hole is negative with respect to you: $E_1 = k_\mu p_1^\mu < 0$. To see this, choose a coordinate system in which $k^\mu = (0, x, 0, 0)$. Prepare your decay such that $p_1^\mu = (1, y, 0, 0)$, where y is small enough for p_1 to be timelike and adjusted such that $xy > 0$. Then E_1 is negative and $E_2 > E$: the particle that reemerges from the ergoregion has more energy than the particle you sent in.

3.3 Hypersurfaces

Let $S(x)$ be a smooth function of the spacetime coordinates. Consider the family of hypersurfaces $S(x) = const.$ The normal vector to $S(x) = const.$ is given by

$$n^\mu = f(x) g^{\mu\nu} \partial_\nu S, \tag{3.27}$$

where $f(x)$ is an arbitrary (smooth) normalisation function. A hypersurface is space-like (timelike) when n is timelike (spacelike) everywhere on it. When the hypersurface is timelike (spacelike) we can always find an $f(x)$ such that $n^2 = -1$ ($n^2 = +1$). A hypersurface is null when n is null, $n^2 = 0$.

3.3.1 The induced metric

Consider a spacelike (timelike) hypersurface Σ in a spacetime (M, g). Normalise the normal n to Σ such that $n^2 = -1$ ($n^2 = +1$). Then the *induced metric h* on Σ is given by

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu. \tag{3.28}$$

The metric h is the restriction of g onto Σ: while h is degenerate on the full tangent space at any point $p \in \Sigma$, it is positive definite for spacelike Σ (Lorentzian for time-like Σ) on the subspace of the tangent space spanned by vectors tangent to curves in Σ (i.e. vectors v that satisfy $v.n = 0$).

To give an example, consider a spatial slice Σ defined by $S(x) = t = const.$ in 3+1 dimensional Minkowski space with Cartesian coordinates $(t, \mathbf{x})$. Then $n^\mu = (1, 0, 0, 0)$ and

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu = \text{diag}(-1, 1, 1, 1) + \text{diag}(1, 0, 0, 0) = \text{diag}(0, 1, 1, 1). \tag{3.29}$$

On the spatial slice Σ, h is just the positive definite three-dimensional flat Euclidean metric. Next, consider a timelike hypersurface Σ defined by $S(x) = x^1 = const.$ with

normal $n^\mu = (0, 1, 0, 0)$. The induced metric on it is $h_{\mu\nu} = \mathrm{diag}(-1, 0, 1, 1)$, which carries Lorentzian signature on the subspace of the tangent space spanned by vectors tangent to Σ.

3.3.2 Null hypersurfaces

For a null hypersurface Σ, the normal n to Σ is also tangent to Σ, since, by definition, a vector t is tangent to Σ if $t.n = 0$. This is satisfied by n itself when it is null: $n.n = 0$. When the normal is null we will henceforth use the symbol ℓ instead. The normal ℓ is tangent to curves $x^\mu(\lambda)$ in Σ: $\ell^\mu = \frac{dx^\mu}{d\lambda}$. In fact, the integral curves of ℓ are *null geodesics*:

Claim: For a null hypersurface Σ, the integral curves in Σ of the normal ℓ are null geodesics.

Proof: For $\ell^\mu = fg^{\mu\nu}\partial_\nu S$ we have

$$\begin{aligned}
\ell.\nabla\ell^\nu &= \ell^\mu\nabla_\mu(fg^{\nu\rho}\partial_\rho S) \\
&= \ell^\mu(\nabla_\mu f)g^{\nu\rho}\partial_\rho S + f\ell^\mu g^{\nu\rho}\nabla_\mu\partial_\rho S \\
&= (\ell.\nabla f)f^{-1}\ell^\nu + f\ell^\mu g^{\nu\rho}\nabla_\mu\partial_\rho S.
\end{aligned} \tag{3.30}$$

The second term reduces as follows:

$$\begin{aligned}
\ell^\mu f g^{\nu\rho}\nabla_\mu\partial_\rho S &= f\ell^\mu g^{\nu\rho}\nabla_\rho\partial_\mu S \\
&= f\ell^\mu g^{\nu\rho}\nabla_\rho(f^{-1}\ell_\mu) \\
&= fg^{\nu\rho}(\nabla_\rho f^{-1})\ell^2 + \ell^\mu g^{\nu\rho}\nabla_\rho\ell_\mu \\
&= \frac{1}{2}g^{\nu\rho}\nabla_\rho(\ell^2) \\
&\propto \ell^\nu
\end{aligned} \tag{3.31}$$

using the fact that $\ell^2 = 0$ on Σ in the third line. In the last line, note that $\ell^2 = 0$ on Σ does not necessarily imply $\nabla_\rho(\ell^2) = 0$, because ℓ^2 can be non-zero outside of Σ and hence its derivative can be non-zero. However, any non-zero contribution to $\nabla_\rho(\ell^2)$ must be proportional to the normal to Σ: $\nabla_\rho(\ell^2) \propto \ell_\rho$. Together with (3.30) this implies that $\ell.\nabla\ell^\nu \propto \ell^\nu$ and therefore the integral curves of ℓ^μ are geodesics. $\square$

If the integral curves of ℓ^μ are not affinely parametrised, then we can always find some function $h(x)$ such that $\tilde{\ell}^\mu = h(x)\ell^\mu$ has affinely parametrised integral curves, i.e. $\tilde{\ell}.\nabla\tilde{\ell}^\mu = 0$. These curves are called the *null geodesic generators* of the null hypersurface.

3.4 Killing Horizons

A null hypersurface Σ is a *Killing horizon* of a Killing vector ξ if ξ is normal to Σ on Σ. Let ℓ be normal to Σ and affinely parametrised such that $\ell.\nabla\ell = 0$. Then, for

some function f, $\xi = f\ell$ on Σ. It follows that ξ satisfies

$$
\begin{aligned}
\xi^\mu \nabla_\mu \xi^\nu &= f\ell^\mu \nabla_\mu (f\ell^\nu) \\
&= f\ell^\mu \ell^\nu \nabla_\mu f + f^2 \ell^\mu \nabla_\mu \ell^\nu \\
&= f\ell^\nu \ell^\mu \nabla_\mu f \\
&\equiv \kappa \xi^\nu
\end{aligned}
\tag{3.32}
$$

on Σ. In the last line we have defined κ, the *surface gravity*: $\xi^\mu \nabla_\mu \xi^\nu = \kappa \xi^\nu$. It takes its name from the fact that κ is constant over the horizon and equals the force that an observer at infinity would have to exert in order to keep a unit mass at the horizon.

3.4.1   Example 1: Schwarzschild

In Schwarzschild spacetime, the surfaces of constant r are the "cylinders" of constant $S(r) = r$ and the black hole horizon corresponds to the particular surface $S(r) = 2M$. To find the normal vector $\ell^\mu = fg^{\mu\nu}\partial_\nu S$ to the horizon, let us work in IEF coordinates. We need the inverse of the IEF metric (1.22), which is given by

$$
g^{\mu\nu} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & \left(1 - \frac{2M}{r}\right) & 0 & 0 \\ 0 & 0 & 1/r^2 & 0 \\ 0 & 0 & 0 & 1/r^2 \sin^2\theta \end{pmatrix}.
\tag{3.33}
$$

The normal vector to the surface $S(r) = 2M$ is then

$$
\ell^\mu = f(g^{vr}\partial_r S, g^{rr}\partial_r S, 0, 0) = f\left(1, \left(1 - \frac{2M}{r}\right), 0, 0\right) = f(1, 0, 0, 0) = f\partial_v.
\tag{3.34}
$$

This vector is null at $r = 2M$ since $\ell^2 = f^2 g^{rr} = f^2\left(1 - \frac{2M}{r}\right) = 0$ at $r = 2M$. Hence $r = 2M$ is a null hypersurface. $\ell = f\partial_v$ is a Killing vector (since $g_{\mu\nu}$ is independent of v) and it is timelike everywhere outside the horizon.

Remember that we already encountered a timelike Killing vector in Schwarzschild spacetime: ∂_t in Schwarzschild coordinates. In fact ∂_t and ∂_v are the same vector field. To check this, denote Schwarzschild coordinates by (t, r, θ, ϕ) and IEF coordinates by $(v, \tilde{r}, \tilde{\theta}, \tilde{\phi})$. Then

$$
\begin{aligned}
dv &= dt + \left(1 - \frac{2M}{r}\right)^{-1} dr \\
d\tilde{r} &= dr \\
d\tilde{\theta} &= d\theta \\
d\tilde{\phi} &= d\phi
\end{aligned}
\tag{3.35}
$$

and therefore

$$
\partial_t = \frac{\partial v}{\partial t}\partial_v + \frac{\partial \tilde{r}}{\partial t}\partial_{\tilde{r}} + \frac{\partial \tilde{\theta}}{\partial t}\partial_{\tilde{\theta}} + \frac{\partial \tilde{\phi}}{\partial t}\partial_{\tilde{\phi}} = \partial_v.
\tag{3.36}
$$

The surface gravity κ is evaluated most easily in IEF coordinates. The equation (3.32) defining κ is a vector equation so we only need to evaluate one of its components. For $\xi = \partial_v$, the $\mu = v$ component of $\xi^\sigma \nabla_\sigma \xi^\mu$ is

$$\xi^\sigma \nabla_\sigma \xi^v = \xi^\sigma \xi^v{}_{,\sigma} + \xi^\sigma \Gamma^v_{\sigma\rho} \xi^\rho = \Gamma^v_{vv}. \tag{3.37}$$

Now

$$\Gamma^v_{vv} = \frac{1}{2} g^{v\sigma} \left(g_{v\sigma,v} + g_{\sigma v,v} - g_{vv,\sigma} \right) = \frac{1}{2} g^{vr}(-g_{vv,r}) = \frac{M}{r^2}, \tag{3.38}$$

which, on the horizon, reduces to $\Gamma^v_{vv} = 1/4M$. Substitution into (3.37) then gives $\kappa = 1/4M$.

Note that the particular normalisation of ξ is important. Had we used a vector with a different normalisation $\tilde{\xi} = c\xi$ we would have obtained $\tilde{\xi}.\nabla \tilde{\xi}^\nu = (c\kappa)\tilde{\xi}^\nu$, leading to a different value $\tilde{\kappa} = c\kappa$ for the surface gravity. In Schwarzschild spacetime, the normalisation that leads to $\kappa = 1/4M$ is fixed by requiring that $\xi^2 = -1$ asymptotically.

3.4.2 Example 2: Kerr

The Killing horizon of a Kerr black hole is slightly more complicated in an interesting way. Recall the Killing vectors $k = \partial_t$ and $m = \partial_\phi$ of the Kerr-Newman black hole in Boyer-Lindquist coordinates (t, r, θ, ϕ) of section 3.1.2. Let's consider the zero charge case of the Kerr black hole. In Kerr coordinates (v, r, θ, χ) — defined above (2.20) — the Killing vectors correspond to $k = \partial_v = \partial_t$ and $m = \partial_\chi = \partial_\phi$. The frame dragging effect translates to the fact that k is *not* normal to the horizon. Instead, for the rotating Kerr black hole, the normal to the horizon is a combination of k and m (exercise):

$$\xi = k + \Omega_H m, \tag{3.39}$$

where $\Omega_H = a/(r_+^2 + a^2)$ is interpreted as the angular velocity of the black hole. The intuition here is that as you move forward in time along a null generator of the horizon, you also have to rotate by a certain amount. Given ξ we can evaluate the surface gravity κ_+ on the outer horizon $r = r_+$ using $\xi.\nabla \xi^\nu = \kappa \xi^\nu$, which evaluates to (exercise)

$$\kappa_+ = \frac{r_+ - r_-}{2(r_+^2 + a^2)}. \tag{3.40}$$

3.5 Black Hole Uniqueness

As already mentioned in section 3.1.2, the Kerr-Newman three-parameter family is the unique stationary black hole solution of the Einstein-Maxwell theory. An

equilibrium black hole in the presence of the electromagnetic field is therefore fully characterised by the three numbers M, J and Q.

However, the electromagnetic field is only one among many matter fields in Nature. Through a series of theorems known as "no hair" theorems it has been established that, in general, black holes have no other properties besides mass, angular momentum and charge (of whichever fields are present). To illustrate the general spirit of these proofs, we look at a simple example here.

Claim: A static black hole cannot be the source of a real (minimally coupled) scalar field.

Proof: Let ϕ be a real scalar field that satisfies the Klein-Gordon equation

$$\nabla^2\phi - m^2\phi = 0. \tag{3.41}$$

By the word "static" in the claim we mean that scalar field and spacetime (i.e. the metric field) have settled down to equilibrium, which implies that the scalar field does not vary in time anymore: "$\dot{\phi} = 0$". In covariant language the statement is that there exists a timelike Killing vector k ($= \partial_t$) of the metric such that $k^\mu\nabla_\mu\phi = 0$. Given k we can foliate the spacetime by $t = const.$ hypersurfaces whose unit normals n are proportional to k, $n \propto k$. The black hole horizon is a Killing horizon of k.

Consider the integral

$$I = \int_R \phi(\nabla^2\phi - m^2\phi)\sqrt{-g}d^4x, \tag{3.42}$$

where $V \subset M$ is the spacetime region *outside* of the horizon. From the Klein-Gordon equation (3.41) it follows immediately that $I = 0$. On the other hand, using Stokes' theorem to integrate by parts the first term, we get

$$I = \int_V (-g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi - m^2\phi^2)\sqrt{-g}d^4x + \int_{\partial V} \phi\nabla_\mu\phi dS^\mu \tag{3.43}$$

where ∂V is the boundary of the region V and dS^μ is the normal surface element on ∂V.

We now show that the surface integral in (3.43) vanishes identically. There are four contributions to ∂V: the horizon $r = 2M$, the surfaces $t = \pm\infty$ and the surface $r \to \infty$ (the "sphere at infinity"). On the horizon and on the $t = \pm\infty$ surfaces, the normal surface element is proportional to k, so the integrand vanishes: $\nabla_\mu\phi dS^\mu \propto k^\mu\nabla_\mu\phi = 0$. We are left with the boundary term at $r \to \infty$. To see that

this vanishes too, note that since the spacetime is static and asymptotically flat, the metric takes the form

$$ds^2 = f(r)dt^2 + g(r)dr^2 + r^2 d\Omega_2^2 \tag{3.44}$$

where $f(r) = -1 + f_1 r^{-1} + \mathcal{O}(r^{-2})$ and $g(r) = 1 + g_1 r^{-1} + \mathcal{O}(r^{-2})$. Using $\nabla^2 = \frac{1}{\sqrt{-g}}\partial_\mu(\sqrt{-g}g^{\mu\nu}\partial_\nu)$, a quick expansion shows that the leading terms in the Klein-Gordon equation in the limit $r \to \infty$ are

$$\partial_r^2\phi + 2r^{-1}\partial_r\phi - m^2\phi = 0, \tag{3.45}$$

which implies that the field must fall off like $\phi \sim 1/r$ or faster. For a sphere of radius r the normal surface element takes the form $dS \propto r dr \partial_r$, so the leading contribution in the integrand is $\phi\nabla_\mu\phi dS^\mu \propto r dr \phi \partial_r\phi \sim dr/r^2$. The integral of this vanishes in the limit $r \to \infty$.

Hence, the first integral in (3.43) must vanish identically. In terms of the induced metric $h_{\mu\nu} = g_{\mu\nu} - n_\mu n_\nu$ on hypersurfaces of $t = const.$ the integrand reads

$$g^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi = h^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi + (n.\nabla\phi)^2 = h^{\mu\nu}\nabla_\mu\phi\nabla_\nu\phi \geq 0 \tag{3.46}$$

having dropped the second term due to the fact that $n \propto k$ and $k.\nabla\phi = 0$. The last inequality follows from the fact that h is positive definite (see section 3.3.1). Therefore, the integrand is always negative or zero and the integral is zero only if the integrand vanishes everywhere. For $m > 0$ this implies that $\phi = 0$ everywhere. For $m = 0$ it implies that $\nabla_\mu\phi = 0 \implies \phi = const.$ everywhere, which is physically indistinguishable from $\phi = 0$ everywhere. $\qquad\square$

3.6   Komar Integrals

For a Killing vector k, the Killing equation $\nabla_\mu k_\nu + \nabla_\nu k_\mu = 0$ is the statement that the symmetric part of $\nabla_\mu k_\nu$ vanishes: $\nabla_{(\mu}k_{\nu)} = 0$. It follows that $\nabla_\mu k_\nu$ is antisymmetric:

$$\nabla_\mu k_\nu = \nabla_{[\mu}k_{\nu]} = \frac{1}{2}\left(\nabla_\mu k_\nu - \nabla_\nu k_\mu\right) \equiv K_{\mu\nu}. \tag{3.47}$$

The tensor $K_{\mu\nu}$ looks similar to the electromagnetic tensor $F_{\mu\nu} = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ and, indeed, just as Maxwell's equations lead to a conserved electric charge, Einstein's equations lead to a conserved quantity when the spacetime has a Killing vector. This is expressed most conveniently in the language of differential forms.

Lightening Review of Differential Forms

Given a d-dimensional spacetime (M, g), a *differential form* of order k (or k-form) on M is a totally antisymmetric covariant k-index tensor. By definition, zero-forms are functions and one-forms are covectors. Given a (local) coordinate chart x^μ on M, the coordinate differentials $dx^1, dx^2, \ldots, dx^d$ are examples of 1-forms and together they form a complete basis. A general one-form can therefore be written as $F = F_\mu dx^\mu$, where F_μ are the components of F. A k-form F can be written as

$$F = \frac{1}{k!} F_{\mu_1 \mu_2 \ldots \mu_k} \, dx^{\mu_1} \wedge dx^{\mu_2} \ldots \wedge dx^{\mu_k}, \tag{3.48}$$

where $\wedge$ is the antisymmetric *wedge product*: $dx^\alpha \wedge dx^\beta = -dx^\beta \wedge dx^\alpha$. The *exterior derivative* of a k-form F is defined as the $(k+1)$-form dF whose components are

$$(dF)_{\mu_1 \mu_2 \ldots \mu_{k+1}} = (k+1) \partial_{[\mu_1} \omega_{\mu_2 \mu_3 \ldots \mu_{k+1}]}. \tag{3.49}$$

A consequence of the antisymmetrisation in the definition is that $d(dF) = 0$ always. The Hodge star $\star$ is a linear map that sends k-forms to $(d-k)$-forms. If F is a k-form, then $\star F$ is defined as the $(d-k)$-form whose components are

$$(\star F)_{\mu_1 \mu_2 \ldots \mu_{d-k}} = \sqrt{-g} \, \varepsilon_{\mu_1 \mu_2 \ldots \mu_{d-k}}{}^{\sigma_1 \sigma_2 \ldots \sigma_k} F_{\sigma_1 \sigma_2 \ldots \sigma_k}, \tag{3.50}$$

where ε denotes is the Levi-Civita symbol. Stokes' theorem states that for any form F the integral of the exterior derivative dF over a region $R \subseteq M$ is equal to the integral of F over the boundary ∂R of R:

$$\int_R dF = \int_{\partial B} F. \tag{3.51}$$

The electromagnetic tensor defines the two-form $F = \frac{1}{2} F_{\mu\nu} \, dx^\mu \wedge dx^\nu$. In terms of the four-potential $A = A_\mu dx^\mu$ it is given by $F = dA$, which implies that $dF = d(dA) = 0$. Maxwell's equations are completely encapsulated in the two equations

$$\begin{aligned} d \star F &= 4\pi \star j \\ d \star j &= 0 \end{aligned} \tag{3.52}$$

where $j = j_\mu dx^\mu$. The second equation is the continuity equation. The Hodge star $\star$ exchanges magnetic and electric fields, i.e. the magnetic and electric components of F are interchanged in $\star F$. Consider the electric charge in some spacelike region B, which we define as

$$Q(B) = \int_B \star j = \frac{1}{4\pi} \int_B d \star F = \frac{1}{4\pi} \int_{\partial B} \star F. \tag{3.53}$$
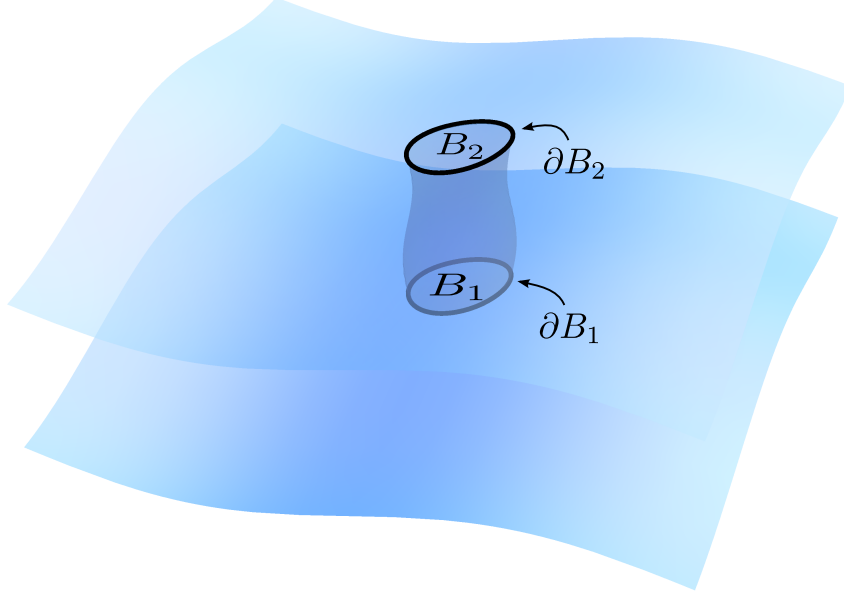
Figure 16. Two spacelike slices with subregions B_1 and B_2. V is the spacetime volume traced out between B_1 and B_2. The regions are chosen in such a way that there are no sources outside of V.

The last equality follows from Stoke's theorem. This charge is conserved, which can be shown as follows. Take any two spacelike surfaces B_1 and B_2. Consider the cylindrical region V, which is bounded by B_1 and B_2 and chosen to be large enough that all the sources (i.e. the support of j) lie inside V. In other words $j = 0$ on the cylindrical boundaries and outside of V. An illustrative example is shown in Figure 16, where V corresponds to the solid cylinder-like region between B_1 and B_2. By virtue of the continuity equation we have:

$$
\begin{aligned}
0 &= \int_V d \star j \\
&= \int_{\partial V} \star j \\
&= \int_{B_1} \star j - \int_{B_2} \star j \\
&= \frac{1}{4\pi} \int_{\partial B_1} \star F - \frac{1}{4\pi} \int_{\partial B_2} \star F \\
&= Q(B_1) - Q(B_2),
\end{aligned}
\tag{3.54}
$$

which proves that Q is conserved.

Let's turn to the analogous situation for $K_{\mu\nu}$. The associated two-form is defined as $K = \frac{1}{2}K_{\mu\nu}\,dx^\mu \wedge dx^\nu$. How does K show up in Einstein's equations? First note that $(\nabla_\mu\nabla_\nu - \nabla_\nu\nabla_\mu)k^\sigma = R^\sigma{}_{\rho\mu\nu}k^\rho$. For $\sigma = \mu$ we get $\nabla_\mu\nabla_\nu k^\mu = R_{\rho\nu}k^\rho$ or

$$\nabla_\mu K^\mu{}_\nu = R_{\rho\nu}k^\rho, \tag{3.55}$$

which looks just like $\nabla_\mu F^\mu{}_\nu = 4\pi j_\nu$. Einstein's equations $R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu}$ can be written as

$$R_{\rho\nu} = 8\pi G\left(T_{\rho\nu} - \frac{1}{2}g_{\rho\nu}T\right). \tag{3.56}$$

Let

$$\zeta_\nu \equiv \frac{1}{8\pi G}R_{\rho\nu}k^\rho \tag{3.57}$$

and $\zeta = \zeta_\nu dx^\nu$. Then Einstein's equations take the form

$$\begin{aligned} d \star K &= 8\pi G \star \zeta \\ d \star \zeta &= 0. \end{aligned} \tag{3.58}$$

In analogy to (3.53) we can then define a *Komar integral*

$$Q_k(B) \equiv \frac{c}{8\pi G}\int_{\partial B} \star K \tag{3.59}$$

for any Killing vector k. It can be shown that, for B large enough such that all matter is inside of it and spacetime is vacuum outside of it, $Q_k(B)$ is conserved.

The physical interpretation of Q_k depends on the Killing vector k. Some examples (whose proof is left as an exercise) for appropriate choices of constants c and regions B are:

- In Schwarzschild spacetime, $Q_k(B) = M$ where $k = \partial_t$.

- In Kerr spacetime, $Q_m(B) = J$ where $m = \partial_\phi$.

- In Kerr-Newman spacetime, $Q_k(B) = M$ and $Q_m(B) = J$.

4 Black Hole Thermodynamics

4.1 Overview

In 1973, Bardeen, Carter and Hawking (BCH) wrote a paper titled "The Laws of Black Hole Thermodynamics", which summarises work on black holes as a series of laws analogous to the laws of thermodynamics. The main points of this analogy are shown in table 1. BCH emphasised that black holes have zero temperature (nothing can escape black holes, so they cannot radiate) and cannot therefore have a physical

Law	Thermodynamics	Black Holes
0th	The temperature T is constant throughout a system in thermal equilibrium	The surface gravity κ is constant over the event horizon of a stationary black hole
1st	$dE = TdS + \sum_i \mu_i dN_i$	$dM = \frac{1}{8\pi}\kappa dA + \Omega_H dJ + \Psi_H dQ$
2nd	$dS \geq 0$	$dA \geq 0$
3rd	T cannot be reduced to zero by a finite number of operations	κ cannot be reduced to zero by a finite number of operations

Table 1. Analogy between the laws of thermodynamics and black holes.

entropy — in their mind, the analogy between the laws of black hole mechanics and the laws of thermodynamics was purely formal.

A young graduate student named Jacob Bekenstein disagreed. He noted that the second law of thermodynamics would be violated if black holes had no entropy, since one could throw arbitrarily entropic objects into the black hole, thereby lowering the total entropy of the exterior universe. He claimed that black holes *must* have an entropy $S_{BH} \propto A$ to save the second law of thermodynamics. Bekenstein's *generalised second law* states that

$$dS_{total} \geq 0 \tag{4.1}$$

where $S_{total} = S_{external} + S_{BH}$. In 1974, it was Hawking who announced that black holes are hot and radiate just like any hot body with a temperature

$$T_H = \frac{\hbar\kappa}{2\pi k_B}, \tag{4.2}$$

from which it follows that a black hole has an entropy given by

$$S_{BH} = \frac{A}{4G\hbar} \tag{4.3}$$

the *Bekenstein-Hawking entropy*. Since $G\hbar$ has units of length squared, we may define the *Planck length* $l_P = \sqrt{G\hbar}$ so that the Bekenstein-Hawking entropy

$$S_{BH} = \frac{1}{4}\frac{A}{l_P^2} \tag{4.4}$$

can be interpreted as a quarter of the area of the black hole, counted in units of l_P^2.

So, after all, the analogy of table 1 between black hole mechancis and thermodynamics seems to be more than a formal peculiarity: if one identifies the temperature with $\hbar\kappa/2\pi k_B$ it becomes a physical unification. In the next sections we will discover the origin of the results discussed here.

4.2 The First Law of Black Hole Mechanics

The first law encapsulates conservation of energy: if a stationary black hole with parameters M, Q and J is perturbed to a new stationary state then the changes in M, Q and J satisfy

$$dM = \frac{1}{8\pi}\kappa dA + \Omega_H dJ + \Phi_H dQ, \tag{4.5}$$

where Ω_H is the angular velocity of the black hole and Φ_H is the "electric surface potential". We will use the uniqueness theorem that the only stationary, asymptotically flat black hole solutions of the Einstein-Maxwell equations are give by the Kerr-Newman family (2.15)

$$ds^2 = -\left(\frac{\Delta - a^2 \sin^2\theta}{\Sigma}\right)dt^2 + \frac{\Sigma}{\Delta}dr^2 - 2\frac{a\sin^2\theta}{\Sigma}\left(r^2 + a^2 - \Delta\right)dtd\phi$$
$$+ \Sigma d\theta^2 + \left(\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2\theta}{\Sigma}\right)\sin^2\theta d\phi^2.$$

In order to verify (4.5), let us express κ, Ω_H, Φ_H and A in terms of M, Q and J (or $a = J/M$, equivalently):

- The surface gravity κ on the outer horizon is given by (3.40) $\kappa = (r_+ - r_-)/2(r_+^2 + a^2)$.

- The surface area A of the horizon is defined at time t_0 as the area of the intersection of the hypersurface Σ of constant $t = t_0$ with the horizon H defined by $r = r_+$. For the Kerr-Newman metric, the induced line-element on the intersection of the two hypersurfaces $t = const.$ and $r = const.$ is

$$ds^2 = h_{\mu\nu}dx^\mu dx^\nu = \Sigma d\theta^2 + \left(\frac{(r^2 + a^2)^2 - \Delta a^2 \sin^2\theta}{\Sigma}\right)\sin^2\theta d\phi^2. \tag{4.6}$$

  On $H \cap \Sigma$ we have $\Delta = 0$ and therefore $\sqrt{h} = (r_+^2 + a^2)\sin\theta$, so that

$$A = \int_{H\cap\Sigma}\sqrt{h}d\theta d\phi = \int_0^{2\pi} d\phi \int_0^\pi d\theta (r_+^2 + a^2)\sin\theta = 4\pi(r_+^2 + a^2). \tag{4.7}$$

- The angular velocity $\Omega_H = a/(r_+^2 + a^2)$ was derived in (3.39) above: it is the non-zero constant in $\chi = \partial_t + \Omega_H \partial_\phi$ that makes χ a Killing vector tangent to the generators of the horizon.

- The electric potential Φ_H is defined as the potential difference between infinity and the horizon, i.e. the work done in bringing a unit charge from $r = \infty$ to $r = r_+$:

$$\Phi_H = \left.(\chi^\mu A_\mu)\right|_{r=r_+} - \left.(\chi^\mu A_\mu)\right|_{r=\infty} = \left.(A_t + \Omega_H A_\phi)\right|_{r=r_+} = \frac{Qr_+}{r_+^2 + a^2}. \tag{4.8}$$

  The second term in the first line was dropped because A_μ vanishes as $r \to \infty$.

Putting these together, we obtain

$$A = A(M, Q, J) = 4\pi \left(2M^2 - Q^2 + 2M\sqrt{M^2 - Q^2 - a^2} \right). \tag{4.9}$$

Since M, Q and J are independent parameters, this implies that

$$dA = \frac{\partial A}{\partial M}dM + \frac{\partial A}{\partial Q}dQ + \frac{\partial A}{\partial J}dJ, \tag{4.10}$$

which can be rearranged to (exercise)

$$dM = \frac{1}{8\pi}\kappa dA + \Omega_H dJ + \Phi_H dQ. \tag{4.11}$$

The proof is deceptively simple. All the hard work goes into proving the uniqueness theorems: you need to know that the black hole settles down to another Kerr-Newman black hole and not some other spacetime. It is worth noting that there exist proofs of the first law, known as "physical process proofs", that do not assume this.

4.3 Working up to Hawking's Area Theorem

Recall that the integral curves of a smooth non-zero vector field form a congruence: a collection of curves through a region of spacetime such that every point in the region lies on exactly one of the curves. Every congruence naturally defines an associated coordinate system (λ, y^a) with $a = 1, 2, 3$, where λ is the parameter such that the vector $\partial_\lambda = t$ is tangent to the curve of the congruence at the point (λ, y^a). In other words, the indices y^a label the curves and λ is a parameter along the curves. The vectors $\eta_a = \partial_{y^a}$ are called *connecting vectors* (they connect neighbouring integral curves).

We shall assume that the curves are geodesics and that λ is an affine parameter so that $t.\nabla t^\mu = 0$. The congruence is then called an (affinely parametrised) *geodesic congruence*. For a null congruence we have $t^2 = 0$ and for a timelike congruences we can set $t^2 = -1$. The vectors t and η_a commute because they form a coordinate basis:

$$[t, \eta_a] = 0 \implies t^\mu \nabla_\mu \eta_a^\nu - \eta_a^\mu \nabla_\mu t^\nu = 0 \tag{4.12}$$

for $a = 1, 2, 3$. Since $t^\mu \nabla_\mu = \frac{d}{d\lambda}$ is the directional derivative along the congruence, this just says that the rate of change of the connecting vectors along a curve of the congruence is given by

$$\frac{d}{d\lambda}\eta_a^\nu = t^\mu \nabla_\mu \eta_a^\nu = \eta_a^\mu \nabla_\mu t^\nu \equiv B_\mu^\nu \eta_a^\mu. \tag{4.13}$$

The tensor B_μ^ν measures the *geodesic deviation*, i.e. the extent to which the connecting vector fails to be parallelly transported along curves of the congruence. When

η_a is parallelly transported, B^ν_μ is zero.

Note that there is an ambiguity in the definition of the connecting vectors: if η_a is a connecting vector from a curve γ to a curve $\tilde{\gamma}$, then $\eta_a + \zeta t$ with $\zeta \in \mathbb{R}$ is also a connecting vector from γ to $\tilde{\gamma}$. We may use this freedom to choose a convenient set of connecting vectors. For example, when the congruence is timelike, it is always possible to choose all three connecting vectors η_a to be orthogonal to t everywhere: $t.\eta_a = 0$ for $a = 1, 2, 3$. This condition will be preserved along the congruence since

$$\frac{d}{d\lambda}(t.\eta_a) = t^\mu \nabla_\mu (t_\nu \eta^\nu_a) = (t^\mu \nabla_\mu t^\nu)\eta^\nu_a + t^\mu t_\nu \nabla_\mu \eta^\nu_a$$
$$= t_\nu \eta^\mu_a \nabla_\mu t^\nu = \frac{1}{2}\eta^\mu_a \nabla_\mu (t^2) = 0. \tag{4.14}$$

When the geodesic congruence is null, we denote the null tangent vector by ℓ. In that case, the orthogonality condition $\eta_a.\ell = 0$ cannot be satisfied for three connecting vectors η_a that are all orthogonal to ℓ because the three-dimensional subspace of the tangent space orthogonal to ℓ includes ℓ itself: $\ell.\ell = 0$. We may still choose two spacelike connecting vectors η_1 and η_2 that are orthogonal to ℓ, i.e. $\ell.\eta_i = 0$ for $i = 1, 2$. We then need one more connecting vector, which we can choose such that it has the following properties: (i) n is null: $n^2 = 0$, (ii) n is parallelly transported along the congruence: $(\ell.\nabla)n^\nu = 0$, (iii) and n satisfies $\ell.n = -1$. These requirements are consistent since the conditions (i) and (iii) are conserved along the congruence when n is parallelly transported: $\ell.\nabla(n.n) = 0$ and $\ell.\nabla(n.\ell)$. Choosing three such η_1, η_2, n is just a particular (and convenient) choice of basis for the space of connecting vectors and it has no particular physical consequence. Once a choice has been made for n, it fixes the two-dimensional subspace of the tangent space orthogonal to both ℓ and n. We require the spacelike connecting vectors η_i to be spacelike to both ℓ and n: $\eta_i.n = \eta_i.\ell = 0$ for $i = 1, 2$. The *projector* onto the two-dimensional spacelike subspace of the tangent space spanned by η_1 and η_2 is given by

$$P^\mu_\nu = \delta^\mu_\nu + n^\mu \ell_\nu + \ell^\mu n_\nu. \tag{4.15}$$

It is a quick exercise to check that P has the right properties: $P^2 = P$, $P^\mu_\nu n^\nu = P^\mu_\nu \ell^\nu = 0$ and $P^\mu_\nu \eta^\nu_i = \eta^\nu_i$.

We are interested in the geodesic deviation of the spacelike connecting vectors η_i, which is given by

$$\frac{d}{d\lambda}\eta^\mu_i = B^\mu_\nu \eta^\nu_i. \tag{4.16}$$

The tensor B^μ_ν, however, is a map on the whole tangent space. It can be turned into a map on the subspace spanned by the connecting vectors η_i, by noting that

$$\frac{d}{d\lambda}\eta^\mu_i = \ell.\nabla \eta^\mu_i = \ell.\nabla(P^\mu_\nu \eta^\nu_i) = P^\mu_\nu \ell.\nabla \eta^\nu_i = P^\mu_\nu B^\nu_\rho \eta^\rho_i = P^\mu_\nu B^\nu_\rho P^\rho_\sigma \eta^\sigma_i \equiv \hat{B}^\mu_\sigma \eta^\sigma_i, \quad (4.17)$$

where $\hat{B}^\mu_\nu \equiv P^\mu_\rho B^\rho_\sigma P^\sigma_\nu$ is the projection of B^μ_ν into the two-dimensional subspace. In the third manipulation we used the fact that $\ell.\nabla P^\mu_\nu = 0$, which follows from the definition (4.15) and the fact that n and ℓ are parallelly transported.

Any linear transformation on a vector space can be decomposed into three components representing stretch, rotation and shear. For $\hat{B}^\mu_\nu$ this can be done by setting

$$\hat{B}^\mu_\nu = \frac{1}{2}\theta P^\mu_\nu + \hat{\sigma}^\mu_\nu + \hat{\omega}^\mu_\nu \tag{4.18}$$

where

$$\theta = \mathrm{Tr}\hat{B} \qquad\qquad \text{represents stretch}$$

$$\hat{\sigma}_{\mu\nu} = \hat{B}_{(\mu\nu)} - \frac{1}{2}P_{\mu\nu}\mathrm{Tr}\hat{B} \quad \text{is the symmetric, trace-free part that represents shear}$$

$$\hat{\omega}_{\mu\nu} = \hat{B}_{[\mu\nu]} \qquad\qquad \text{is the antisymmetric part that represents rotations}$$

and $\mathrm{Tr}\hat{B} \equiv \hat{B}^\mu_\mu$ is the trace of $\hat{B}^\mu_\nu$. The three components of $\hat{B}^\mu_\nu$ contain information about the geometry of the congruence. For example, it can be shown that $\hat{\omega} = 0$ implies that the tangent vector field ℓ is normal to a family of null hypersurfaces (the converse statement holds too). We will prove this below, but in order to do so, we need two results first.

Theorem *(Frobenius)*: A vector field χ is hypersurface orthogonal if and only if

$$\chi_{[\mu}\nabla_\nu\chi_{\omega]} = 0. \tag{4.19}$$

We omit the proof — for more detail and references containing proofs see section B.3 in Wald.

Lemma: $\ell_{[\rho}\hat{B}_{\mu\nu]} = \ell_{[\rho}B_{\mu\nu]}$.
Proof: Schematically,

$$\hat{B}_{\mu\nu} = P^\rho_\mu B_{\rho\sigma}P^\sigma_\nu = B_{\mu\nu} + \ell_\mu(\dots) + (\dots)\ell_\nu \tag{4.20}$$

using the definition of P^μ_ν in (4.15). The result then follows immediately due to the total antisymmetrisation on the indices. $\qquad\square$

Using these results the proposition above can be proven.

Claim: The tangent field ℓ is normal to a family of null hypersurfaces if and only if $\hat{\omega} = 0$.
Proof: First, assume that $\hat{\omega} = 0$. Then $\hat{B}_{[\mu\nu]} = 0$, which together with the previous lemma implies

$$0 = \ell_{[\mu}\hat{B}_{\nu\rho]} = \ell_{[\mu}B_{\nu\rho]} = \ell_{[\mu}\nabla_\nu\ell_{\rho]} \tag{4.21}$$

and by Frobenius' theorem ℓ is hypersurface orthogonal. Conversely, assume that ℓ is hypersurface orthogonal. Then it follows from Frobenius' theorem that $\ell_{[\mu}\hat{B}_{\nu\rho]} = \ell_{[\mu}B_{\nu\rho]} = 0$ and therefore

$$\ell_\mu\hat{\omega}_{\rho\sigma} + \ell_\rho\hat{\omega}_{\sigma\mu} + \ell_\sigma\hat{\omega}_{\mu\rho} = 0. \tag{4.22}$$

Contracting with n^μ and using $n^\mu\ell_\mu = 1$ and $n^\mu\hat{\omega}_{\mu\rho} = n^\mu\hat{\omega}_{\sigma\mu} = 0$ (since $\hat{\omega}$ contains the projection operator), we conclude that $\hat{\omega}_{\rho\sigma} = 0$. $\quad\square$

4.4   The Raychaudhouri Equation

The Raychaudhuri equation tells us how the area of the two-dimensional surface elements spanned by η_1 and η_2 changes along the congruence. Formally, the magnitude a of the area element spanned by η_1 and η_2 is defined by

$$a = \epsilon_{\mu\nu\rho\sigma}n^\mu\ell^\nu\eta_1^\rho\eta_2^\sigma. \tag{4.23}$$

Then it can be shown that (see e.g. Townsend section 6.1.1)

$$\frac{da}{d\lambda} = \theta a \tag{4.24}$$

where θ is the "stretch" term in $\hat{B}^\mu_\nu$ defined in (4.18). From (4.24) we see that θ indeed measures the expansion of the geodesics in the congruence: when $\theta > 0$ the geodesics *diverge*, and when $\theta < 0$ the geodesics *converge*. Raychoudhuri's equation is a statement about the rate of change of θ:

$$
\begin{aligned}
\frac{d\theta}{d\lambda} &= \ell.\nabla\theta = \ell.\nabla\hat{B}^\mu_\mu = \ell.\nabla(P^\mu_\rho B^\rho_\sigma P^\sigma_\mu) = \ell.\nabla(B^\rho_\sigma P^\mu_\rho P^\sigma_\mu) \\
&= \ell.(B^\rho_\sigma P^\sigma_\rho) && \text{(using } P^2 = P\text{)} \\
&= P^\sigma_\rho \ell^\mu \nabla_\mu B^\rho_\sigma && \text{(using } \ell.\nabla P = 0 \text{ from (4.15))} \\
&= P^\sigma_\rho \ell^\mu \nabla_\mu(\nabla_\sigma \ell^\rho) \\
&= P^\sigma_\rho \ell^\mu (\nabla_\mu\nabla_\sigma\ell^\rho - \nabla_\sigma\nabla_\mu\ell^\rho + \nabla_\sigma\nabla_\mu\ell^\rho) \\
&= P^\sigma_\rho [\ell^\mu R^\rho_{\ \nu\mu\sigma}\ell^\nu + \nabla_\sigma(\ell^\mu\nabla_\mu\ell^\rho) - (\nabla_\sigma\ell^\mu)\nabla_\mu\ell^\rho] \\
&= (\delta^\sigma_\rho + \ell^\sigma n_\rho + n^\sigma\ell_\rho)\ell^\mu R^{[\rho}_{\ \nu\mu\sigma]}\ell^\nu - P^\sigma_\rho(\nabla_\sigma\ell^\mu)\nabla_\mu\ell^\rho && \text{(using } \ell.\nabla\ell^\rho = 0\text{)} \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - P^\sigma_\rho B^\mu_\sigma B^\rho_\mu \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - P^\sigma_\rho B^\rho_\alpha \delta^\alpha_\mu B^\mu_\sigma \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - P^\sigma_\rho B^\rho_\alpha(P^\alpha_\mu - \ell^\alpha n_\mu - n^\alpha\ell_\mu)B^\mu_\sigma \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - P^\sigma_\rho B^\rho_\alpha P^\alpha_\mu B^\mu_\sigma && \text{(using } \ell_\mu B^\mu_\sigma = B^\rho_\alpha\ell^\alpha = 0\text{)} \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - P^\sigma_\gamma P^\gamma_\rho B^\rho_\alpha P^\alpha_\beta P^\beta_\mu B^\mu_\sigma && \text{(using } P = P^2 \text{ twice)} \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - \hat{B}^\mu_\nu\hat{B}^\nu_\mu \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - \left(\tfrac{1}{2}\theta P^\mu_\nu + \hat{\sigma}^\mu_\nu + \hat{\omega}^\mu_\nu\right)\left(\tfrac{1}{2}\theta P^\nu_\mu + \hat{\sigma}^\nu_\mu + \hat{\omega}^\nu_\mu\right) \\
&= -R_{\mu\nu}\ell^\mu\ell^\nu - \tfrac{1}{2}\theta^2 + \hat{\omega}^2 - \hat{\sigma}^2.
\end{aligned}
\tag{4.25}
$$

We also made use of the symmetries of the Riemann tensor. To summarise, the *Raychaudhuri equation* for null geodesic congruences is:

$$\frac{d\theta}{d\lambda} = -R_{\mu\nu}\ell^\mu\ell^\nu - \frac{1}{2}\theta^2 - \hat{\sigma}^2 + \hat{\omega}^2. \tag{4.26}$$

So far, this is a purely geometrical statement about null geodesic congruences in a Lorentzian manifold — it contains no physics. The physics comes from Einstein's equations

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G T_{\mu\nu} \implies R_{\mu\nu}\ell^\mu\ell^\nu = 8\pi G T_{\mu\nu}\ell^\mu\ell^\nu, \tag{4.27}$$

from which we obtain

$$\frac{d\theta}{d\lambda} = -8\pi G T_{\mu\nu}\ell^\mu\ell^\nu - \frac{1}{2}\theta^2 + \hat{\omega}^2 - \hat{\sigma}^2. \tag{4.28}$$

The first term in this equation is constrained when matter satisfies the *Weak Energy Condition (WEC)*:

$$T_{\mu\nu}t^\mu t^\nu \geq 0 \qquad \text{for all timelike } t, \tag{4.29}$$

which says that the energy density in the frame of any timelike observer is non-negative. It then follows by continuity that $T_{\mu\nu}\ell^\mu\ell^\nu \geq 0$ for the null vector ℓ. This has an important consequence for the expansion parameter θ:

Claim: The expansion θ of the congruence of null geodesic generators of a null surface obeys the differential inequality

$$\frac{d\theta}{d\lambda} \leq -\frac{\theta^2}{2} \tag{4.30}$$

when matter satisfies the WEC.

Proof: In the Raychaudhuri equation (4.26)

$$\frac{d\theta}{d\lambda} = -R_{\mu\nu}\ell^\mu\ell^\nu - \frac{1}{2}\theta^2 - \hat{\sigma}^2 + \hat{\omega}^2,$$

the third term satisfies $\hat{\sigma}^2 = h^{\rho\sigma}h^{\mu\nu}\hat{\sigma}_{\sigma\mu}\hat{\sigma}_{\nu\rho} \geq 0$ because the induced metric on the subspace of the tangent space spanned by η_1 and η_2 is positive-definite. The last term vanishes, $\hat{\omega} = 0$, because ℓ is hypersurface orthogonal. When the WEC is satisfied, the first term in (4.26) is smaller than or equal to zero and therefore

$$\frac{d\theta}{d\lambda} \leq -\frac{\theta^2}{2}, \tag{4.31}$$

which proves the claim. $\qquad\square$

Claim: If $\theta = \theta_0 < 0$ at a point p along a null geodesic generator γ, then $\theta \to -\infty$ at a point q along γ with finite affine parameter-distance $\leq 2/|\theta_0|$ from p.

Proof: By rearranging (4.30) we obtain

$$\frac{d\theta^{-1}}{d\lambda} = -\frac{1}{\theta^2}\frac{d\theta}{d\lambda} \geq \frac{1}{2}. \tag{4.32}$$

Choose $\lambda = 0$ at p on γ. Integrating (4.32) and setting $\theta(\lambda = 0) = \theta_0$, we find that $\theta^{-1} \geq \lambda/2 + \theta_0^{-1}$. Hence $\theta^{-1} = 0$ at some point q along γ at affine parameter-distance $\lambda \leq -2/|\theta_0|$ from p. $\qquad\square$

Hence, if the geodesics are converging at any point on the congruence, they must continue to converge and the expansion goes to $-\infty$. This signals a breakdown of the congruence and neighbouring geodesics then meet at some point q, called a *caustic* or *focus*.

In the special case of a Killing horizon, the null geodesic generators are integral curves of the Killing vector, which generates a symmetry of the spacetime. In that case $\theta = 0$ everywhere (this can be proven by studying $\hat{B}_{\mu\nu}$ directly when the surface is a Killing horizon).

4.5   Causal Structure

Given a spacetime (M, g), a *timelike curve* in M is a curve $\gamma \in M$ with an everywhere timelike tangent vector. A *causal curve* is a curve with a nowhere spacelike tangent vector. Clearly, a timelike curve is also causal. If there exists a future directed $\left\{\begin{matrix} \text{causal} \\ \text{timelike} \end{matrix}\right\}$ curve γ from p to q we say that p is to the $\left\{\begin{matrix} \text{causal} \\ \text{chronological} \end{matrix}\right\}$ past of q and we write $\left\{\begin{matrix} p \prec q \\ p \ll q \end{matrix}\right\}$. Given a subset $U \subset M$ we define the

$$\begin{matrix} \textit{chronological past of U:} & I^-(U) = \{x \in M | \exists\, u \in U \text{ s.t. } x \ll u\} \\ \textit{causal past of U:} & J^-(U) = \{x \in M | \exists\, u \in U \text{ s.t. } x \prec u\}. \end{matrix} \tag{4.33}$$

Analogous definitions can be given for the chronological future $I^+(U)$ and the causal future $J^+(U)$. For any set S, we denote the (topological) *closure* of S (i.e. S together with its limit points) by $\overline{S}$, the interior of S (i.e. the largest open set contained in S) by $\mathring{S}$ and the boundary of S by $\partial S = \overline{S} \setminus \mathring{S}$.

The chronological past $I^-(U)$ (and future) of a set U is always open. To see this, notice that timelike curves are defined by the property that the norm of their tangent vector at every point is bounded away from 0. This means that, for any point p in $I^-(U)$ that lies on some timelike curve γ from p to some $u \in U$, we can

always find a small neighbourhood N around p such that every point in N can be connected to u by a timelike curve $\tilde{\gamma}$ obtained by a small deformation of γ. This is the defining property of an open set: every point in the set has a neighbourhood contained in the set. On the other hand, the causal past $J^-(U)$ may or may not be closed. For example, consider a point p in Minkowski space. Then $J^-(p)$ is closed (it contains all its limit points). However, if we remove a point r from Minkowski space that lies on the past light-cone of p, $J^-(p)$ will not be closed since the points on the null geodesic beyond r, which extends that from p to r, are not in $J^-(p)$ whereas they are in $\partial J^-(p)$.

The following relations between the chronological and causal past are always true:

$$
\begin{aligned}
I^-(U) &= \mathring{J}^-(U) \\
\overline{I^-(U)} &= \overline{J^-(U)} \\
\partial\overline{I^-(U)} &= \partial\overline{J^-(U)} \\
J^-(U) &\subset \overline{I^-(U)}.
\end{aligned}
\tag{4.34}
$$

By consequence of the equivalence principle, every spacetime looks locally like Minkowski space: for any point $p \in M$, there exists a so-called *convex normal neighbourhood* in which the local causal structure is that of Minkowski space.

Some facts about the causal and chronological relations:

$$
\begin{aligned}
x \ll y \quad &\text{and} \quad y \ll z \quad \implies \quad x \ll z \\
x \prec y \quad &\text{and} \quad y \prec z \quad \implies \quad x \prec z \\
x \ll y \quad &\text{and} \quad y \prec z \quad \implies \quad x \ll z
\end{aligned}
\tag{4.35}
$$

Furthermore, when x, y, z do *not* lie on the same geodesic, then $x \prec y$ and $y \prec z$ together imply $x \ll z$.

The boundary of the causal past $\partial J^-(U) \setminus U$ is always a null hypersurface. Indeed, if it were timelike somewhere, one could find points $p \prec q$ with $q \in J^-(U)$ and $p \notin J^-(U)$, which is a contradiction. If it were spacelike somewhere, one could find a point $p \in J^-(U)$ such that all timelike future directed curves γ through p would leave $J^-(U)$ — a contradiction, too. The surface $\partial J^-(U)$ is generated by null geodesics. A null geodesic generator cannot leave $\partial J^-(U)$, although it can join $\partial J^-(U)$. This is a consequence of the following theorem:

Theorem *(Penrose):* A null geodesic generator of $\partial J^-(U)$ cannot have a future endpoint on $\partial J^-(U)$.

Proof: The typical arguments in such proofs go as follows. Let γ be a null geodesic, p a point in the causal past $p \in J^-(U)$ and $r \prec p$, which implies that $r \in J^-(U)$. Consider a CNN of p and a sequence of points $p_i \to p$ with p as accumulation point such that $p_i \in I^-(U)$. Then there exist timelike curces μ_i from p_i to U. Denote the point where μ_i leaves the CNN of p by q_i. Since the boundary of the CNN is compact, the points q_i have an accumulation point q. The curves μ_i corresponding to the convergent subsequences of the points q_i also converge to a limit curve μ from p to q. Since the curves μ_i are timelike, μ cannot be spacelike by continuity (but it could be null). Now $q \in \partial J^-(U)$ because $q \in I^-(U)$, and therefore $p \prec q$ and $q \ll U$, which implies $p \ll U$. This contradicts the assumption that $q \in \partial J^-(U)$. $\qquad \square$

4.6   Hawking's Area Theorem

The area theorem is a fully dynamical result about black hole spacetimes, which means that it cannot make use of such concepts as Killing horizons, which rely on particular spacetime symmetries. Hence we need another way to define the event horizon of a black hole. A natural definition can be made in terms of the spacetime's causal structure: for an asymptotically flat spacetime (M, g), the *future horizon H^+* is defined as the boundary of the causal past of future null infinity, $\partial J^-(\mathcal{J}^+)$. To give an example, the horizon in the the Penrose diagram of the collapsing star in Figure 11 corresponds to the dashed black line.

Note that, strictly speaking, $\mathcal{J}^+$ is a set of points that is *not* part of the original spacetime manifold (M, g). However, $\mathcal{J}^+$ is contained in the conformal compactification $(\tilde{M}, \tilde{g})$, and since the causal structure is invariant under conformal transformations, the set of points H^+ obtained by taking the boundary of the causal past of $\mathcal{J}^+$ will be contained in the original spacetime manifold (M, g).

Since the horizon H^+ is the boundary of the causal past of a set, it is a null hypersurface and its null geodesic generators cannot leave H^+ by the theorems in the previous section. We are now in a position to prove Hawking's area theorem.

Theorem *(Hawking):* The area of H^+ cannot decrease if the weak energy condition holds *and* cosmic censorship holds (no singularities on or outside the horizon).

Proof: By the area of H^+ we mean the area of the intersection of H^+ with a spacelike hypersurface. Now H^+ is a null geodesic congruence and if we can show that the expansion θ satisfies $\theta \geq 0$ everywhere on H^+ then the result follows immediately. Suppose that $\theta = \theta_0 < 0$ at some $p \in H^+$. If the WEC holds, there is a conjugate point q at finite affine parameter-distance along the geodesic γ through p and all the points $r \in \gamma$ beyond q are timelike related to p by theorem 9.3.10 in Wald. So γ must

have left H^+ at q. But that contradicts the Penrose theorem that null generators of $\partial J^-(\mathcal{J}^+)$ can have no future endpoints on $\partial J^-(\mathcal{J}^+)$. Hence $\theta \geq 0$ everywhere on H^+ and the theorem follows. $\square$

This theorem is fully dynamical: it holds whatever far-from-equilibrium processes are taking place. However, it is purely classical and does not hold when quantum effects are important.

What does the area theorem imply for the black hole spacetimes we have come across already? For a Schwarzschild black hole, the area is proportional to the mass: $A = 4\pi(2M)^2 = 16\pi M^2$. Hence, if a physical process starts with a Schwarzschild black hole and ends with another, the mass cannot decrease. That this relies on the WEC is clear: the situation would be different if you were able to throw negative energy into the black hole. For a Kerr black hole, $A = 4\pi(r_+^2 + a^2)$ where $r_+ = M + \sqrt{M^2 + a^2}$. Hence, the mass of a Kerr black hole could in principle decrease in the course of a physical process, but it would have to be compensated by an increase in the angular momentum.

5 Hawking Radiation

In this section, we will study the physics of black holes taking into account quantum mechanical effects. In order to do so, we will need to become acquainted with the methods of *quantum field theory (QFT) in curved spacetime*. First, a quick review of (free, scalar) QFT in flat space. Remember that four-vectors are denoted by standard letters and (spatial) three-vectors are denoted by bold-face letters, e.g. p stands for the four-momentum and $\mathbf{p}$ stands for the three-momentum.

Lightening Review of Free Scalar QFT in Flat Space

The equation of motion for a real scalar field ϕ is the Klein-Gordon equation

$$\Box\phi - m^2\phi = 0 \tag{5.1}$$

where $\Box = \eta^{\mu\nu}\partial_\mu\partial_\nu$. A special set of solutions to (5.1) are the "positive-frequency" plane waves

$$\psi_p = N_p e^{ip_\mu x^\mu} \tag{5.2}$$

where $p^2 + m^2 = 0$. The positive-frequency condition is the requirement that the p^0-component of p be positive: $p^0 = +\sqrt{\mathbf{p}^2 + m^2} > 0$.[6] The set of all positive-frequency

---

[6]Note that due to the signature of the metric $p^0 > 0 \iff p_0 < 0$.

plane wave solutions together with their complex conjugates, which we will denote by $\{\psi_p, \psi_p^*\}$, is a complete set of functions (a *basis*). Hence any field configuration can be expanded as

$$\phi(x) = \int d^3p \left(a_p \psi_p(x) + a_p^* \psi_p^*(x)\right),$$ (5.3)

where a_p and a_p^* are the expansion coefficents (the coefficients of ψ_p and ψ_p^* must be complex conjugates of each other in order for $\phi(x)$ to be real). Upon second quantisation, the coefficients in (5.3) become operators a_p and $a_p^\dagger$ (we denote the Hermitian conjugate by a dagger and we will omit the hats on operators in most cases for notational convenience, but from now on the field and its coefficients should be thought of as operators). In order that the standard canonical commutation relations for the field $[\phi(t, \mathbf{x}), \dot{\phi}(t, \mathbf{y})] = i\delta(\mathbf{x} - \mathbf{y})$ and $[\phi(t, \mathbf{x}), \phi(t, \mathbf{y})] = 0$ translate to the simple relations $[a_p, a_q^\dagger] = \delta(\mathbf{p} - \mathbf{q})$ and $[a_p, a_q] = 0$ on the operator coefficients, we normalise the mode functions using the *Klein-Gordon inner product*

$$(f, g) = i \int d^3x\, f^* \overleftrightarrow{\partial_t} g$$ (5.4)

where $f^* \overleftrightarrow{\partial_t} g = f^*(\partial_t g) - (\partial_t f^*)g$. The required normalisation conditions are then $(\psi_p, \psi_q) = \delta(\mathbf{p} - \mathbf{q})$ and $(\psi_p, \psi_q^*) = 0$, fixing $N_p = 1/\sqrt{2p^0}(2\pi)^{\frac{3}{2}}$ up to an arbitrary complex phase. The vacuum state is defined by $a_p|0\rangle = 0 \ \forall \ p$.

At first sight it seems that this definition of the vacuum state depends on a particular choice of frame: we chose an inertial frame with coordinates x^μ in which we imposed the positive-frequency condition $p^0 > 0$. This condition will look different in a new coordinate system $\tilde{x}^\mu$. However, it turns out that for all *inertial* reference frames, i.e. those related to x^μ by a Lorentz transformation $\tilde{x}^\mu = \Lambda^\mu_\nu x^\nu$, the positive-frequency condition leads to the same vacuum. To see this, consider positive-frequency mode functions in a new frame $\tilde{x}^\mu$:

$$\tilde{\psi}_p = N_p e^{ip_\mu \tilde{x}^\mu}.$$ (5.5)

The field then has an expansion

$$\phi(\tilde{x}) = \int d^3p \left(\tilde{a}_p \tilde{\psi}_p + \tilde{a}_p^\dagger \tilde{\psi}_p^*\right)$$ (5.6)

in terms of these modes and the new vacuum state will be defined by $\tilde{a}_p|\tilde{0}\rangle = 0 \ \forall \ p$. To prove that the two vacua are the same, $|\tilde{0}\rangle = |0\rangle$, we just need to show that

$$a_p|0\rangle = 0 \ \forall \ p \implies \tilde{a}_p|0\rangle = 0 \ \forall \ p.$$ (5.7)

(The converse implication then follows by symmetry.) It follows from $\tilde{p}_\mu = p_\nu \Lambda^\nu_{\ \mu}$ that $\tilde{\psi}_p \propto \psi_{\tilde{p}}$:

$$\tilde{\psi}_p = \frac{1}{\sqrt{2p^0}(2\pi)^{\frac{3}{2}}}e^{ip_\mu \tilde{x}^\mu} = \left(\frac{\tilde{p}^0}{p^0}\right)^{\frac{1}{2}}\frac{1}{\sqrt{2\tilde{p}^0}(2\pi)^{\frac{3}{2}}}e^{i\tilde{p}_\mu x^\mu} = \left(\frac{\tilde{p}^0}{p^0}\right)^{\frac{1}{2}}\psi_{\tilde{p}}. \qquad (5.8)$$

Furthermore, the mode $\psi_{\tilde{p}}$ is a positive-frequency mode because $p^0 > 0 \implies \tilde{p}^0 > 0$.[7] Hence, the coefficient in front of the positive-frequency mode $\tilde{\psi}_p$ with momentum p in frame $\tilde{x}^\mu$ will be the same as the coefficient in front of the positive-frequency mode $\psi_{\tilde{p}}$ with momentum $\tilde{p}$ in frame x^μ: $\tilde{a}_p = a_{\tilde{p}}$. By consequence, we have

$$a_p|0\rangle = 0\ \forall\ p \implies a_{\tilde{p}}|0\rangle = 0\ \forall\ \tilde{p} \implies \tilde{a}_p|0\rangle = 0\ \forall\ p, \qquad (5.9)$$

which proves that $|0\rangle = |\tilde{0}\rangle$. Hence, the vacuum state in Minkowski spacetime is independent of which inertial frame you use to define it. If the vacuum appears empty of particles to observers in one inertial frame, it will appear empty of particles to observers in all inertial frames. This is a direct consequence of the fact that positive-frequency modes in one inertial frame are linear combinations (in this case particularly simple ones) of positive-frequency modes in another inertial frame.

5.1 Quantum Field Theory in Curved Spacetime

Fix some background spacetime (M, g). We will assume that the spacetime satisfies the strongest of all global causality conditions: *global hyperbolicity*. A spacetime is globally hyperbolic if it admits a *Cauchy surface*, which is a three-dimensional submanifold of M such that every past and future inextendible causal curve intersects the hypersurface exactly once. When a manifold admits a Cauchy surface, it can be foliated entirely by a family Cauchy surfaces. In the context of field theory, the presence of a Cauchy surface is relevant because once a Cauchy surface has been picked and initial data has been specified on it, the corresponding solution to the equations of motion is completely determined on the whole of the spacetime: much like in the case of flat spacetime, everything is predictable from a set of initial data. Some examples of spacetimes that are globally hyperbolic:

- Minkowski space (any spacelike hypersurface extending out to infinity in all directions is a Cauchy surface).

---

[7]This is true because by Lorentz transformations we really mean (proper) orthochronous Lorentz transformations — orthochronous meaning that they preserve the direction of time. To see more explicitly why this implies that the sign of the time-component of the four-momentum is preserved, recall that for a (proper orthochronous) Lorentz transformation, $\Lambda^0_{\ 0} = \gamma = 1/\sqrt{1 - |\mathbf{v}|^2/c^2}$ and $\Lambda^0_{\ i} = -\gamma v_i/c$, where $\mathbf{v}$ is the relative speed between the two reference frames. Then $\tilde{p}^0 = \Lambda^0_{\ \mu}p^\mu = \gamma(\sqrt{m^2 + |\mathbf{p}|^2} - \mathbf{v}\cdot\mathbf{p}/c) \geq \gamma|\mathbf{p}|(1 - |\mathbf{v}|\cos\theta/c) > 0$ as required. We used the fact that γ is positive, $|\mathbf{v}| < c$ and $\cos\theta \leq 1$.

- FLRW spacetimes.

- The collapsing star spacetime.

Some examples of spacetimes that are *not* globally hyperbolic:

- Minkowski space with a point removed. For any infinite spacelike hypersurface, the removed point will either be to the future or past of it. But then there will exist past/future inextendible causal curves that end at the removed point and therefore fail to intersect the surface. This reflects the fact that specifying initial conditions on the surface is not enough to fix a consistent solution to the equation of motion on the whole spacetime: anything can happen at the removed point (it could be a source or a sink to the field) and so additional boundary conditions need to be specified there.

- Spacetimes with a naked singularity (such as the $M < 0$ Schwarzschild black hole). In that case, there will be timelike curves that start and/or end on the singularity and avoid crossing any candidate Cauchy surface we might think of. Again this reflects the fact that we don't know what happens at the singularity — we would have to specify additional boundary conditions on it in order to have a well-posed problem.

Assuming that (M, g) admits a Cauchy surface we study a free, real scalar field on it which satisfies the Klein-Gordon equation

$$\Box \phi - m^2 \phi = 0 \tag{5.10}$$

where $\Box = \nabla^\mu \nabla_\mu$ is now in general more complicated due to the coordinate-dependence of the metric. The Klein-Gordon inner product generalises to

$$(f, g) = i \int_\Sigma dS^\mu f^* \overleftrightarrow{\partial_\mu} g \tag{5.11}$$

where Σ is a Cauchy surface and $dS^\mu = n^\mu dS$, n^μ is the unit normal to Σ and dS is the induced volume element on Σ. It can be shown that the Klein-Gordon inner product between two solutions of the Klein-Gordon equation will be the same no matter which Cauchy surface you choose to evaluate it on. Hence, specifying a condition like "ψ must have unit norm $(\psi, \psi) = +1$" with respect to one Cauchy surface ensures that the condition will hold for all Cauchy surfaces, and hence that it will hold consistently throughout the whole spacetime.

Deep and difficult theorems in the theory of differential equations say that, when there is a Cauchy surface, we can always find some (highly non-unique) orthonormal

basis of solutions $\{\psi_i, \psi_i^*\}$ such that

$$(\psi_i, \psi_j) = +\delta_{ij}$$
$$(\psi_i^*, \psi_j^*) = -\delta_{ij} \tag{5.12}$$
$$(\psi_i, \psi_j^*) = 0.$$

Since there exists a plethora of such orthonormal bases, the "vacuum" state obtained by doing standard second quantisation with one such basis in general has no physical meaning.

However, if there exists a timelike Killing vector $k = \partial_t$, i.e. we are dealing with a stationary spacetime, then k commutes with the Klein-Gordon operator $\Box - m^2$ and one can find simultaneous eigenmodes of the two operators. The second quantisation procedure can then be used to define a vacuum with a meaningful physical interpretation, by designating a special subset of these modes as "positive-frequency modes". This special set is found as follows. First, note that k is antihermitian: $(f, kg) = -(kf, g)$. Hence it has imaginary eigenvalues $ku_i = -i\omega_i u_i$ with $\omega_i \in \mathbb{R}$, and eigenfunctions with distinct values are orthogonal.[8] We can therefore find a basis $\{\psi_i\}$ satisfying (5.12) such that $k\psi_i = -i\omega_i\psi_i$ with $\omega_i > 0$ and use this condition as the generalised "positive-frequency" condition with respect to k (this of course reduces to the original condition in Minkowski spacetime, since there $k\psi_p = \partial_t\psi_p = ip_0\psi_p = -ip^0\psi_p$ with $p^0 > 0$). If we second quantise with these modes, then the state defined by $a_i|0\rangle = 0 \,\forall\, i$, where a_i are the operator coefficients of the positive-frequency modes in the expansion of the field operator $\phi(x)$, is the ground state of a physically meaningful Hamiltonian

$$\hat{H} = \int_{t=const.} dS^\mu \hat{T}_{\mu\nu} k^\nu, \tag{5.13}$$

which can be interpreted as the operator corresponding to the total energy on the surface $t = const.$ So the state $|0\rangle$ is the minimum-energy state on the surface $t = const.$ in the sense that it is the minimum-eigenvalue eigenstate of $\hat{H}$. In flat spacetime this just reduces to

$$\hat{H} = \int d^3x \hat{T}_{00}. \tag{5.14}$$

The Hamiltonian $\hat{H}$ dictates the time-evolution of all operators through the Heisenberg equation of motion:

$$\frac{d\hat{\mathcal{O}}}{dt} = i\left[\hat{H}, \hat{\mathcal{O}}\right], \tag{5.15}$$

i.e. it is the operator that generates time-translations.

---

[8]As is customary, we discard the pathological "zero mode" $\omega = 0$.

We now apply this formalism to a simple curved spacetime. Consider a "sandwich spacetime" (M, g), which is made up of three regions, region B (for bottom), region C (for centre) and region T (for top). The Klein-Gordon equation holds throughout the whole spacetime. Region B is stationary and it admits a timelike Killing vector k^B. Region C is not stationary and all sorts of dynamical processes might take place in it (as long as the whole of M remains globally hyperbolic). In region T the metric is again stationary, with a new timelike Killing vector k^T. If we want to quantise the field in region B, we choose a set of modes $\{f_i, f_i^*\}$ that satisfy $k^B f_i = -i\omega_i f_i$ with $\omega_i > 0$. On the other hand, in region T, we choose another set of modes $\{g_i, g_i^*\}$ that satisfy $k^T g_i = -i\tilde{\omega}_i g_i$ with $\tilde{\omega}_i > 0$. Note that even though the positive-frequency conditions are imposed using the Killing vectors in the two separate regions B and T, the modes f_i and g_i all extend throughout the whole of spacetime. In the two cases, the respective expansions of the field are then

$$\phi(x) = \sum_i \left(a_i f_i + a_i^\dagger f_i^* \right) = \sum_i \left(b_i g_i + b_i^\dagger g_i^* \right) \tag{5.16}$$

where the modes have been normalised with respect to the Klein-Gordon inner product in each case such that $[a_i, a_j^\dagger] = \delta_{ij}$ and $[b_i, b_j^\dagger] = \delta_{ij}$. Now, since $\{f_i\}$ is a basis we can also expand any function g_i in terms of it:

$$g_i = \sum_i A_{ij} f_j + B_{ij} f_j^*. \tag{5.17}$$

The coefficients A_{ij} and B_{ij} are called *Bogoliubov coefficients* and the transformation between different bases is called a *Bogoliubov transformation*. Using the normalisation conditions (5.12) it can then be shown that the Bogoliubov coefficients satisfy the following relations:

$$\sum_k A_{ik} A_{jk}^* - B_{ik} B_{jk}^* = \delta_{ij}$$
$$\sum_k A_{ik} B_{jk} - B_{ik} A_{jk} = 0. \tag{5.18}$$

In matrix notation this can be written as $\mathbf{A}\mathbf{A}^\dagger - \mathbf{B}\mathbf{B}^\dagger = \mathbf{1}$ and $\mathbf{A}\mathbf{B}^T - \mathbf{B}\mathbf{A}^T = 0$. We can also relate the different operator coefficients to each other, e.g. $b_i = (g_i, \phi) = \sum_j A_{ij}^* a_j - B_{ij}^* a_j^\dagger$.

The procedure above defines a vacuum associated with the modes $\{f_i, f_i^*\}$ — call it the "*in*-vacuum" — as the state that satisfies $a_i |0\rangle_{in} = 0 \ \forall \ i$. In a stationary reference frame in region B, i.e. a frame that follows an integral curve of k^T, this vacuum will appear empty. What about region T? More precisely, what is the expected number $\langle N_i \rangle$ of particles of species (or momentum) i in the state $|0\rangle_{in}$, if evaluated in a stationary reference frame in region T? It is given by the expectation
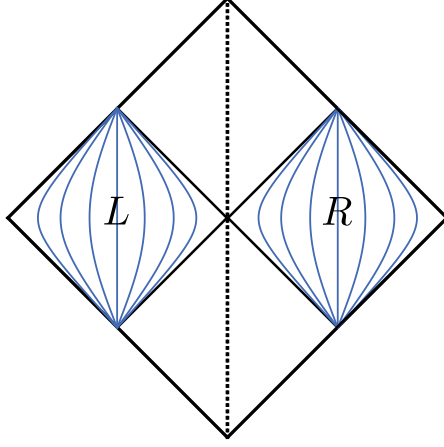
Figure 17. Minkowski space in $1+1$ dimensions and the left and right Rindler wedges. Some accelerated worlines (lines of constant ξ_L and ξ_R) have been drawn.

value of the number operator $N_i = b_i^\dagger b_i$ (no summation over i), which is (exercise):

$$\langle N_i \rangle = {}_{in}\langle 0|b_i^\dagger b_i|0\rangle_{in} = \sum_j B_{ij} B_{ij}^* = (\mathbf{BB}^\dagger)_{ii} \quad \text{(no summation over } i). \tag{5.19}$$

If this is non-zero, there is particle production. Another way to say this is that the *in*-vacuum is different from the "*out*-vacuum" defined by $b_i|0\rangle_{out} = 0 \; \forall \; i$. Hence, a changing spacetime geometry generically causes particle production.

5.2   QFT in Rindler Space in $1+1$ dimensions

We will first look at QFT in Rindler space in $1+1$ dimensions, since it features many of the aspects that will be relevant in the discussion of Hawking radiation. By QFT in Rindler space we really mean the field quantisation as it would be carried out by eternally accelerated observers in Minkowski space. We shall see that such observers, when travelling through the Minkowski vacuum, will feel themselves immersed in a thermal bath of particles.

In light-cone coordinates $\bar{u} = t - x$ and $\bar{v} = t + x$, the line-element of $1 + 1$-dimensional Minkowski space is $ds^2 = -d\bar{u}d\bar{v}$. In section 1.3.3 we constructed the Penrose diagram for Rindler space and we considered only the "right Rindler wedge": $\bar{u} < 0, \bar{v} > 0$. This corresponds to the region R in Figure 17. There also exists a left Rindler wedge $\bar{u} > 0, \bar{v} < 0$, denoted by L in Figure 17. In the right Rindler wedge, define the coordinates ξ and η via

$$\begin{aligned} t &= a^{-1} e^{a\xi} \sinh a\eta \\ x &= a^{-1} e^{a\xi} \cosh a\eta \end{aligned} \tag{5.20}$$

such that $-\infty < \xi, \eta < \infty$ and

$$ds^2 = e^{2a\xi}(-d\eta^2 + d\xi^2). \tag{5.21}$$

The proper time measured by a Rindler observer (i.e. along a $\xi = const.$ line) is $\tau = e^{2a\xi}\eta$ if we synchronise it to $\tau = 0$ as the observer passes the x-axis. Hence, the coordinate η is a physical time-coordinate for the Rindler observer travelling at constant ξ: it ticks at a rate proportional to the observer's watch. Now define the null coordinates $u = \eta - \xi$ and $v = \eta + \xi$, in terms of which the Rindler metric reads

$$ds^2 = -e^{2a\xi}dudv. \tag{5.22}$$

The relation between the Minkowski and Rindler null coordinates is

$$\begin{aligned} \bar{u} &= -a^{-1}e^{-au} \\ \bar{v} &= a^{-1}e^{av}, \end{aligned} \tag{5.23}$$

which implies that

$$\begin{aligned} \bar{u} \to 0 &\iff u \to +\infty \\ \bar{v} \to 0 &\iff v \to -\infty, \end{aligned} \tag{5.24}$$

as shown in Figure 17. Notice the analogy to the relationship between the null Eddington-Finkelstein coordinates (denoted above as u, v) and the Kruskal coordinates U, V (1.25): $U = -e^{-u/4M}$ and $V = e^{v/4M}$.

Analogous definitions can be made for coordinates that cover the left Rindler wedge. The coordinates and line elements for the left and right Rindler wedge are shown in table 2, where we've used subscripts to distinguish between the two wedges. *In this section, Rindler coordinates without subscript always refer to the right wedge.*

In fact, the time coordinates $\eta_L = \eta_R = \tanh^{-1}(t/x)$ are exactly the same function of Minkowski coordinates — they just take different ranges in L and R. We can therefore drop the subscripts on η completely, as long as we remember that $\eta < 0$ corresponds to L and $\eta > 0$ to R. Lines of $\eta = const.$ then correspond to lines through the origin that extend over both L and R. These lines are Cauchy surfaces for Minkowski spacetime. The vector ∂_η is a timelike Killing vector (exercise) in both L and R, but it is future-pointing in R and past-pointing in L. We now have three globally hyperbolic manifolds with global future-pointing timelike Killing vectors: Minkowski space with ∂_t, the left Rindler wedge with $-\partial_\eta$, and the right Rindler wedge with ∂_η. In each of these the canonical quantisation procedure outlined in the preceding section can be applied.

Left Rindler Wedge L	Right Rindler Wedge R
$t = -a^{-1}e^{a\xi_L}\sinh a\eta_L$	$t = a^{-1}e^{a\xi_R}\sinh a\eta_R$
$x = -a^{-1}e^{a\xi_L}\cosh a\eta_L$	$x = a^{-1}e^{a\xi_R}\cosh a\eta_R$
$u_L = \eta_L - \xi_L$	$u_R = \eta_R - \xi_R$
$v_L = \eta_L + \xi_L$	$v_R = \eta_R + \xi_R$
$\bar{u} = a^{-1}e^{-au_R}$	$\bar{u} = -a^{-1}e^{-au_L}$
$\bar{v} = -a^{-1}e^{av_R}$	$\bar{v} = a^{-1}e^{av_L}$
$ds^2 = e^{2a\xi_L}(-d\eta_L^2 + d\xi_L^2)$	$ds^2 = e^{2a\xi_R}(-d\eta_R^2 + d\xi_R^2)$
$\quad = -e^{2a\xi_L}du_L dv_L$	$\quad = -e^{2a\xi_R}du_R dv_R$

Table 2. Coordinates in the left and right Rindler wedges and their relations to Minkowski coordinates.

In Minkowski coordinates, the Klein-Gordon equation for the massless scalar field is $(-\partial_t^2 + \partial_x^2)\phi = 0$ and there exists a set of positive frequency "Minkowski" modes

$$f_k = \frac{1}{\sqrt{4\pi\omega_k}}\exp\left(-i\omega_k t + ikx\right) \tag{5.25}$$

with $\omega_k = |k| > 0$. It is useful to make the distinction between "left-moving" and "right-moving" modes, which take a simple form in terms of Rindler light-cone coordinates:

$$f_k = \begin{cases} \dfrac{1}{\sqrt{4\pi\omega_k}}\exp\left(-i\omega_k\bar{u}\right) & k > 0 \quad \text{``right-moving''} \\ \dfrac{1}{\sqrt{4\pi\omega_k}}\exp\left(-i\omega_k\bar{v}\right) & k < 0 \quad \text{``left-moving''}. \end{cases} \tag{5.26}$$

The field expansion is

$$\phi(x) = \sum_k a_k f_k(x) + a_k^\dagger f_k(x)^* \tag{5.27}$$

and the Minkowski vacuum is defined by $a_k|0_M\rangle = 0 \,\forall\, k$.

In R, the Klein-Gordon equation becomes $\Box\phi = e^{-2a\xi}(-\partial_\eta^2 + \partial_\xi^2)\phi = 0$, which admits plane wave solutions

$$g_k^R = \frac{1}{\sqrt{4\pi\omega_k}}\exp\left(-i\omega_k\eta + ik\xi\right) \tag{5.28}$$

with $\omega_k = |k| > 0$. These solutions are positive frequency with respect to ∂_η. Note that the functions g_k^R as defined in (5.28) are only specified for $R \subset M$. An eternally accelerated observer in R might use these modes to carry out the quantisation of the scalar field in R, which would lead to a field expansion

$$\phi(x) = \sum_k b_k g_k^R(x) + b_k^\dagger g_k^R(x)^* \tag{5.29}$$

and a vacuum $b_k |0_R\rangle = 0 \ \forall \ k$. This vacuum has no (obvious) relation to the Minkowski vacuum and indeed pertains to an entirely different spacetime manifold (the Rindler wedge). However, we are interested in a different question here: how would an eternally accelerated observer travelling through Minkowski spacetime perceive the *Minkowski* vacuum, i.e. the Poincaré-invariant ground state of Minkowski spacetime? In order to answer this, we need to decide what the appropriate positive frequency condition is for such an observer. The Killing vector ∂_t is not natural to Rindler observers, since the coordinate t is not proportional to the physical proper time measured by a Rindler observer's watch. Instead, we've seen above that the proper time of a Rindler observer is proportional to η. Hence, the vector ∂_η is the natural Killing vector with respect to which a Rindler would measure frequencies and define a notion of positive frequency.

We need to extend our definition of the Rindler modes because as defined in (5.28) they are not specified on a Cauchy surface for M. For example, they are only defined on half of the Cauchy surface $\eta = 0$ (for $x > 0$). This is easily remedied by defining the *global* modes g_k^R as follows:

$$g_k^R = \begin{cases} \dfrac{1}{\sqrt{4\pi\omega_k}} \exp\left(-i\omega_k\eta + ik\xi\right) & \text{in } R \\ 0 & \text{in } L. \end{cases} \tag{5.30}$$

These modes are defined on an entire Cauchy surface (e.g. $\eta = 0$). The set $\{g_k^R, g_k^{R*}\}$ does *not* form complete basis for the functions on M yet. For this reason, we also need to consider the positive-frequency modes in the left wedge L, which are associated with left-accelerating Rindler observers. These modes are given by

$$g_k^L = \begin{cases} \dfrac{1}{\sqrt{4\pi\omega_k}} \exp\left(+i\omega_k\eta + ik\xi\right) & \text{in } L \\ 0 & \text{in } R. \end{cases} \tag{5.31}$$

and are positive frequency with respect to $-\partial_\eta$ in L (remembering that the *future-pointing* timelike Killing vector in L is $-\partial_\eta$). Taken together, the positive-frequency modes in L and R form a complete set, and so a field configuration $\phi(x)$ in Minkowski space may be expanded as

$$\phi(x) = \sum_k b_k g_k^R(x) + c_k g_k^L(x) + h.c. \tag{5.32}$$

where $h.c.$ denotes Hermitian conjugate. We can now ask: How many R-particles are expected to be seen in the Minkowski vaccum — how many particles does the externally accelerated observer in the right wedge see? In other words, what is $\langle 0_M |^R N_k| 0_M \rangle$, where $^R N_k = b_k^\dagger b_k$? To simplify things, we specialise our question to right-movers, for which $k > 0 \implies \omega = k$. We need the Bogoliubov coefficients

$$g_\omega^R = \int_0^\infty d\omega' (A_{\omega\omega'} f_{\omega'} + B_{\omega\omega'} f_{\omega'}^*) \tag{5.33}$$

where $f_{\omega'} = (4\pi\omega')^{-\frac{1}{2}} \exp(-i\omega'\bar{u})$. Notice the resemblance of this integral with the Fourier transform of g_ω^R:

$$g_\omega^R(\bar{u}) = \frac{1}{2\pi} \int_{-\infty}^\infty d\omega' e^{-i\omega'\bar{u}} \tilde{g}_\omega(\omega') \tag{5.34}$$

where

$$\tilde{g}_\omega(\omega') = \int_{-\infty}^\infty d\bar{u} \, e^{i\omega'\bar{u}} g_\omega^R(\bar{u}) \tag{5.35}$$

(for notational convenience we omit the superscript R on the Fourier transform). Our strategy is now to massage (5.34) into a form that allows us to write the Bogoliubov coefficients in (5.33) in terms of the Fourier transform $\tilde{g}_\omega$. To that end, we split up the integral in (5.34) as follows:

$$g_\omega^R(\bar{u}) = \frac{1}{2\pi} \int_0^\infty d\omega' e^{-i\bar{u}\omega'} \tilde{g}_\omega(\omega') + \frac{1}{2\pi} \int_0^\infty d\omega' e^{i\bar{u}\omega'} \tilde{g}_\omega(-\omega'). \tag{5.36}$$

We flipped the sign of the integration variable $\omega' \to -\omega'$ in the second term. Comparing with (5.33), we find

$$A_{\omega\omega'} = \sqrt{\frac{\omega'}{\pi}} \tilde{g}_\omega(\omega') \qquad \text{and} \qquad B_{\omega\omega'} = \sqrt{\frac{\omega'}{\pi}} \tilde{g}_\omega(-\omega'). \tag{5.37}$$

In order to find the Bogoliubov coefficients (5.37), we don't actually need to evaluate the Fourier transforms. It is enough if we can relate $\tilde{g}_\omega(\omega')$ to $\tilde{g}_\omega(-\omega')$, since the additional condition $\mathbf{AA}^\dagger - \mathbf{BB}^\dagger = \mathbf{1}$ then fixes $A_{\omega\omega'}$ and $B_{\omega\omega'}$ completely. To do so we prove the following claim:

Claim: $\tilde{g}_\omega(-\omega') = -e^{-\pi\omega/a} \tilde{g}_\omega(\omega')$ if $\omega' > 0$.

Proof: The integral defining $\tilde{g}_\omega$ uses g_ω^R in terms of $\bar{u}$. Since $u = -a^{-1} \ln(-a\bar{u})$, the right-moving R-modes take the form

$$g_\omega^R = \begin{cases} \dfrac{1}{\sqrt{4\pi\omega_k}} e^{-i\frac{\omega}{a} \ln(-a\bar{u})} & \text{for } \bar{u} < 0 \\ 0 & \text{for } \bar{u} > 0. \end{cases} \tag{5.38}$$
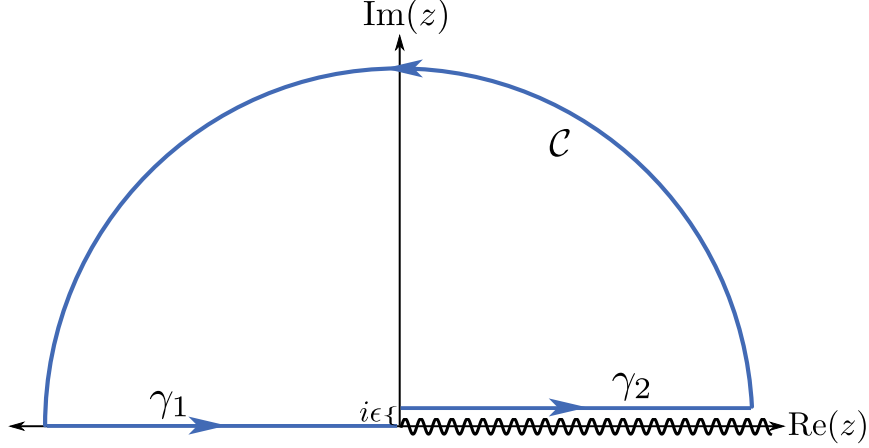
Figure 18. The contour integral used in the evaluation of $g_\omega^R(-\omega')$ with a branch cut of $\ln(-z)$ on the positive real axis.

The function $\ln(-z)$ is multi-valued on the complex plane so we will give it a branch cut on the positive real axis. Then

$$
\begin{aligned}
\tilde{g}_\omega(-\omega') &= \int_{-\infty}^{\infty} d\bar{u}\, e^{-i\omega'\bar{u}} g_\omega(\bar{u}) \\
&= \frac{1}{\sqrt{4\pi\omega}} \int_{-\infty}^{0} d\bar{u}\, e^{-i\omega'\bar{u}} e^{-i\frac{\omega}{a}\ln(-a\bar{u})}.
\end{aligned}
\tag{5.39}
$$

Now we deform the path of integration, using the contour shown in Figure 18. The contour is closed in the upper complex half-plane because the integral over the arc segment $\mathcal{C}$ then vanishes for $\omega' > 0$ in the limit as the radius of the arc goes to infinity. In order to avoid the branch cut we choose γ_2 to lie just above the real axis by an infinitesimal amount $i\epsilon$ (which we send to zero at the end). Since the integrand has no poles inside the closed contour (i.e. it is holomorphic there), Cauchy's integral theorem implies

$$
0 = \oint d\bar{u}\, e^{-i\omega'\bar{u}} e^{-i\frac{\omega}{a}\ln(-a\bar{u})} = \left\{ \int_{\gamma_1} + \int_{\gamma_2} + \int_{\mathcal{C}} \right\} d\bar{u}\, e^{-i\omega'\bar{u}} e^{-i\frac{\omega}{a}\ln(-a\bar{u})}.
\tag{5.40}
$$

Strictly speaking, there is another contribution on the right hand side for the integral from $i\epsilon$ to 0 (see Figure 18), but this will vanish when we send ϵ to zero at the end so we can forget about it here. Since the integral over the arc $\mathcal{C}$ vanishes, it follows

from (5.40) that $\int_{\gamma_1} = -\int_{\gamma_2}$, which implies

$$
\begin{aligned}
\tilde{g}_\omega(-\omega') &= -\frac{1}{\sqrt{4\pi\omega}} \int_{i\epsilon}^{\infty+i\epsilon} d\bar{u}\, e^{-i\omega'\bar{u}} e^{-i\frac{\omega}{a}\ln(-a\bar{u})} \\
&= -\frac{1}{\sqrt{4\pi\omega}} \int_{-\infty-i\epsilon}^{-i\epsilon} d\bar{u}\, e^{i\omega'\bar{u}} e^{-i\frac{\omega}{a}\ln(a\bar{u})} \qquad (\bar{u} \leftrightarrow -\bar{u}) \\
&= -\frac{1}{\sqrt{4\pi\omega}} \int_{-\infty}^{0} d\bar{u}\, e^{i\omega'\bar{u}} e^{-i\frac{\omega}{a}[\ln(-a\bar{u})-i\pi]} \qquad (\epsilon \to 0) \\
&= -e^{-\pi\omega/a}\tilde{g}_\omega(\omega')
\end{aligned}
\tag{5.41}
$$

as claimed. □

Hence we have $A_{\omega\omega'} = -e^{-\pi\omega/a}B_{\omega\omega'}$, which together with $|A_{\omega\omega}|^2 - |B_{\omega\omega}|^2 = 1$ implies

$$
\langle 0_M|{}^R N_\omega|0_M\rangle = (\mathbf{BB}^\dagger)_{\omega\omega} = |B_{\omega\omega}|^2 = \frac{1}{e^{2\pi\omega/a} - 1}.
\tag{5.42}
$$

This corresponds to a *Planck spectrum* with temperature $T = \frac{\hbar a}{2\pi k_B}$.

In fact, we can show more: the state $|0_M\rangle$ *is* the thermal state with respect to the Rindler Hamiltonian $\hat{H}_R$, i.e. the Hamiltonian associated with a hypersurface $\eta = const.$ in the right Rindler wedge (in other words, $\hat{H}_R$ generates time-translations along ∂_η). Mathematically this statement is

$$
\mathrm{Tr}_L\left(|0_M\rangle\langle 0_M|\right) = \frac{e^{-\beta\hat{H}_R}}{\left\langle e^{-\beta\hat{H}_R}\right\rangle}
\tag{5.43}
$$

where $\beta = (k_b T)^{-1} = 2\pi k_b/\hbar a$. To show this, it is convenient to introduce so-called *Unruh modes*, which are positive frequency with respect to ∂_t, i.e. they are linear combinations of f_k-modes. The Unruh mode of frequency k is defined as

$$
z_k = \frac{1}{\sqrt{2\sinh\omega\pi/a}}\left(e^{\pi\omega/2a}\, g_k^R + e^{-\pi\omega/2a}\, g_{-k}^L\right).
\tag{5.44}
$$

In order to form a complete set, they must be complemented by the modes

$$
z_k' = \frac{1}{\sqrt{2\sinh\omega\pi/a}}\left(e^{-\pi\omega/2a}\, g_{-k}^R + e^{\pi\omega/2a}\, g_k^L\right).
\tag{5.45}
$$

The modes z_k and z_k' together with their complex conjugates form a complete basis for functions on Minkowski space. Hence, the field can be equally well expanded in terms of them:

$$
\phi(x) = \sum_k d_k z_k(x) + d_k' z_k'(x) + h.c.
\tag{5.46}
$$

Comparing this with (5.32) we can read off

$$b_k = \frac{1}{\sqrt{2\sinh\omega\pi/a}}\left(e^{\pi\omega/2a}d_k + e^{-\pi\omega/2a}d'^{\dagger}_{-k}\right)$$

$$c_k = \frac{1}{\sqrt{2\sinh\omega\pi/a}}\left(e^{\pi\omega/2a}d'_k + e^{-\pi\omega/2a}d^{\dagger}_{-k}\right)$$

(5.47)

Since the Unruh modes are positive frequency with respect to ∂_t, the Minkowski vacuum satisfies $d_k|0_M\rangle = d'_k|0_M\rangle = 0 \ \forall \ k$. Now define the *Rindler vacuum* $|0_{Rd}\rangle$ by $c_k|0_{Rd}\rangle = b_k|0_{Rd}\rangle = 0 \ \forall \ k$ and further the (unnormalised) left- and right-wedge n-particle states

$$\left|{}^R n, k\right\rangle = (b_k^{\dagger})^n |0_{Rd}\rangle \qquad \text{and} \qquad \left|{}^L n, k\right\rangle = (c_k^{\dagger})^n |0_{Rd}\rangle. \tag{5.48}$$

Then it can be shown that

$$|0_M\rangle = \prod_k (\cosh\phi_k)^{-1} \sum_{n=0}^{\infty} e^{-n\pi\omega/a} \left|{}^R n, k\right\rangle \left|{}^L n, k\right\rangle \tag{5.49}$$

where ϕ_k is defined by $\tanh\phi_k = e^{-\pi\omega/a}$. This state is perfectly entangled between the left and right Rindler wedges: if a measurement is done in R in which n right-wedge particles are found, then we know with certainty that the post-measurement state has n left-wedge particles. To find the state ρ_R in R, we need to trace over the states in L via the "partial trace" Tr_L. Defining $E_n = n\omega$, we obtain

$$\rho_R = \text{Tr}_L |0_M\rangle\langle 0_M| = \sum_{n=0}^{\infty} \prod_k \frac{e^{-\beta E_n}}{\sum_{m=0}^{\infty} e^{-\beta E_m}} \left|{}^R n, k\right\rangle \left\langle {}^R n, k\right| = \frac{e^{-\beta\hat{H}_R}}{\langle e^{-\beta\hat{H}_R}\rangle} \tag{5.50}$$

as claimed above. This reveals something perhaps surprising about the Minkowski vacuum: it is not a local quantity at all, i.e. it is not a state of the form

$$(\text{vacuum at } x_1) \otimes (\text{vacuum at } x_2) \otimes (\text{vacuum at } x_3) \otimes \dots$$

Instead, it is a highly entangled state whose fluctuations are correlated throughout spacetime.

The actual physical meaning of the results above may perhaps seem a bit obscure. We considered the quantisation of a quantum field "appropriate to a Rindler observer" in Minkowski space, which led us to conclude that a Rindler observer is immersed in a thermal bath of radiation. How convincing were our arguments for choosing some appropriate quantisation scheme? How would an eternally accelerated observer actually *measure* the radiation? To put the results derived above on more solid grounds, Bill Unruh constructed a simple model of a detector (nowadays known as an Unruh-deWitt detector) and showed that an accelerated detector in Minkowski space behaves as if it were in a thermal bath of particles. This effect is known as the *Unruh effect*. Hawking radiation, which we discuss next, is a phenomenon closely related to this effect.

5.3 QFT in the spacetime of a spherical collapsing star: Hawking Radiation

Consider the spacetime that corresponds to a spherically symmetric collapsing star, heavy enough to form a black hole. The Penrose diagram is shown in Figure 19. This is a curved spacetime that is globally hyperbolic (for instance $\mathcal{J}^-$ is a Cauchy surface). Even though the Schwarzschild black hole solution is a static spacetime, the collapsing star spacetime is *not*: the formation of the black hole involves complicated dynamics. However, the spacetime is approximately stationary in the far asymptotic past (i.e. near $\mathcal{J}^-$) and in the far asymptotic future (i.e. near $\mathcal{J}^+$). Therefore we can perform a second quantisation associated with stationary observers near $\mathcal{J}^-$ (which leads to the so-called *in*-vacuum), and a second quantisation associated with stationary observers near $\mathcal{J}^+$. We can then ask the question we asked in the example of the "sandwich spacetime" above: will observers in the far future see particles in the *in*-vacuum?

The field expansion defining the *in*-vacuum can be constructed by specifying a complete set of positive-frequency modes on $\mathcal{J}^-$. For the quantisation in the far future, it is not sufficient to define a complete set on $\mathcal{J}^+$, because $\mathcal{J}^+$ is not a Cauchy surface for the spacetime (for example, causal curves that end on the singularity never register on $\mathcal{J}^+$). However, the surface $H^+ \cup \mathcal{J}^+$ is a Cauchy surface, so we can quantise the field in the far future by specifying a complete set on it. Hence there will be three sets of modes:

$$
\begin{aligned}
f_i &: \text{positive frequency on } \mathcal{J}^- \\
g_i &: \text{positive frequency on } \mathcal{J}^+ \text{ and zero on } H^+ \\
h_i &: \text{positive frequency on } H^+ \text{ and zero on } \mathcal{J}^+.
\end{aligned}
\tag{5.51}
$$

(Strictly speaking, there is no timelike Killing vector on H^+ — recall that ∂_t is null on the horizon — so the term "positive frequency" is somewhat misleading there. However, the outcome of the calculation does not actually depend on the choice of modes h_i, so we can just choose an arbitrary set of modes on H^+ and declare those as the "positive frequency" modes, i.e. attach them to annihilation operators in the field expansion, as long as together with g_i they form a complete set. For more detail, see Wald §14.3.) The set $\{f_i, f_i^*\}$ is complete. The sets $\{g_i, g_i^*\}$ and $\{h_i, h_i^*\}$ are not complete by themselves but together they form a complete set. We can expand the field in two ways:

$$
\phi(x) = \sum_i a_i f_i(x) + h.c. = \sum_i b_i g_i(x) + c_i h_i(x) + h.c.
\tag{5.52}
$$

As before, to find out what observers at late times on $\mathcal{J}^+$ see in the state $|0_{in}\rangle$ defined by $a_i |0_{in}\rangle = 0 \; \forall \; i$, we will evaluate $\langle N_i \rangle = (\mathbf{B}\mathbf{B}^\dagger)_{ii}$, for which we need the
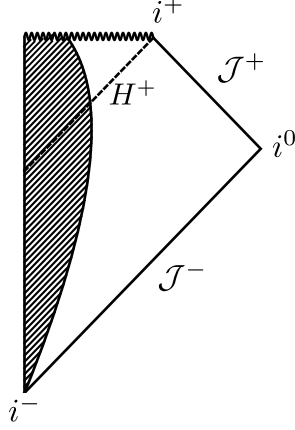
Figure 19. Penrose diagram of the spherically symmetric collapsing star.

Bogoliubov coefficients in the expansion

$$g_i = \sum_j A_{ij} f_j + B_{ij} f_j^*. \tag{5.53}$$

If we could find the exact solutions to the Klein-Gordon equation, this would be a fairly straight-forward task: once the positive-frequency solutions g_i on $\mathcal{J}^+$ are specified, we can do the analog of a Fourier transform to find their expansion in terms of f_i and f_i^* on $\mathcal{J}^-$ (such an expansion always exists since $\{f_i, f_i^*\}$ is a complete set of functions) and read off the Bogoliubov coefficients. However, we will see in a moment that the Klein-Gordon equation is complicated in Schwarzschild spacetime, which prevents us from finding analytic solutions. Instead, we will ask: if a given solution to the Klein-Gordon equation is asymptotically positive-frequency at $\mathcal{J}^+$, then what is its form at $\mathcal{J}^-$? In other words, we will impose a boundary condition to the solution at $\mathcal{J}^+$ and investigate what its corresponding form must be on $\mathcal{J}^-$. This amounts to "tracing back in time" the solution from $\mathcal{J}^+$ to $\mathcal{J}^-$.

The metric of the Schwarschild black hole spacetime with coordinates (t, r_*, θ, ϕ) reads

$$ds^2 = \left(1 - \frac{2M}{r}\right)(-dt^2 + dr_*^2) + r^2 d\Omega_2^2. \tag{5.54}$$

We will also use the light-cone coordinates $u = t - r_*$ and $v = t + r_*$ below. Using the handy formula

$$\Box = \nabla^\mu \nabla_\mu = \frac{1}{\sqrt{-g}} \partial_\mu \left(\sqrt{-g} g^{\mu\nu} \partial_\nu\right) \tag{5.55}$$

we can find the Klein-Gordon equation for the field $\phi(t, r_*, \theta, \phi)$. Expanding the solution in spherical harmonics $\phi(t, r_*, \theta, \phi) = \chi_l(r_*, t) Y_{lm}(\theta, \phi)$ we find

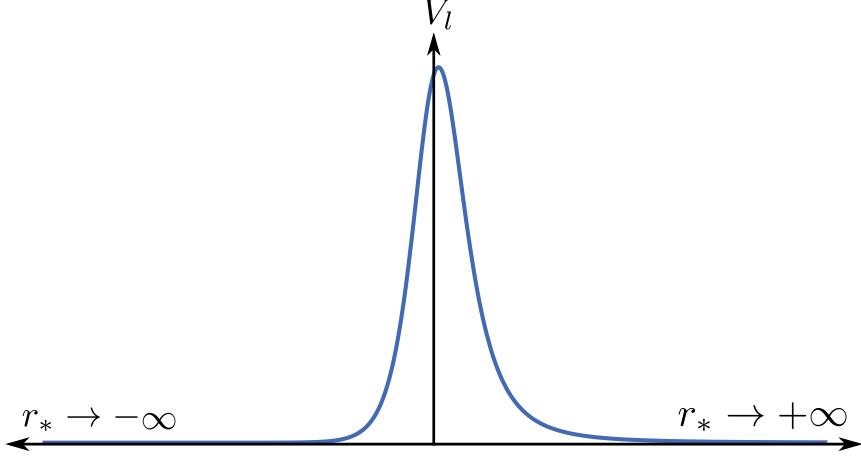$$\left[\partial_t^2 - \partial_{r_*}^2 + V_l(r_*)\right] \chi_l = 0 \tag{5.56}$$

Figure 20. The potential V_l as a function of r_* (shown here for $l = 1$).

where
$$V_l(r_*) = \left(1 - \frac{2M}{r}\right)\left[\frac{l(l+1)}{r^2} + \frac{2M}{r^3}\right]. \tag{5.57}$$

Set $\chi_l(r_*, t) = e^{-i\omega t}R_{l\omega}(r_*)$ so that

$$(\partial_{r_*}^2 + \omega^2)R_{\omega l} = V_l R_{\omega l}. \tag{5.58}$$

This equation can be cast into a form known as the *confluent Heun equation*, whose solutions have been studied to some extent but are complicated. We can gain some insight into the solutions by looking more closely at the potential $V_l(r_*)$. Both near the horizon H^+ ($r \to 2M \iff r_* \to -\infty$) and near $\mathcal{J}^\pm$ ($r \to \infty \iff r_* \to \infty$), the potential tends to zero — it takes the form of a *potential barrier*, as shown in Figure 20. Hence, if we consider how any particular solution to (5.58) evolves in time, it will be partly transmitted and partly reflected as it comes in from $r_* = \infty$.

Near $\mathcal{J}^\pm$, the solutions to (5.56) are just plane waves. The equivalent of right-moving and left-moving for a radial coordinate is outgoing and ingoing (corresponding, respectively, to r_* increasing or decreasing with time) so we define "early modes"

$$\begin{aligned}
f_{lm\omega+} &= \frac{1}{\sqrt{2\pi\omega}}e^{-i\omega u}\frac{Y_{lm}}{r} &\text{(outgoing)} \\
f_{lm\omega-} &= \frac{1}{\sqrt{2\pi\omega}}e^{-i\omega v}\frac{Y_{lm}}{r} &\text{(ingoing)}
\end{aligned} \tag{5.59}$$

at $\mathcal{J}^-$, and "late modes"

$$\begin{aligned}
g_{lm\omega+} &= \frac{1}{\sqrt{2\pi\omega}}e^{-i\omega u}\frac{Y_{lm}}{r} &\text{(outgoing)} \\
g_{lm\omega-} &= \frac{1}{\sqrt{2\pi\omega}}e^{-i\omega v}\frac{Y_{lm}}{r} &\text{(ingoing)}
\end{aligned} \tag{5.60}$$

at $\mathcal{J}^+$. We will be interested mainly in ingoing early modes and outgoing late modes, so we will use the shorthand notation

$$f_\omega \sim f_{lm\omega-}$$
$$g_\omega \sim g_{lm\omega+}.$$

(5.61)

We need to express g_ω in terms of $f_{\omega'}$ and $f_{\omega'}^*$ on $\mathcal{J}^-$. First, note that plane waves such as g_ω are in fact completely delocalised, i.e. they have support everywhere on $\mathcal{J}^+$. However, using a superposition of such positive frequency plane waves, we can always construct a localised "wave packet" on $\mathcal{J}^+$ (for example a Gaussian packet) and we can choose it to be peaked around some momentum ω_0 and around some coordinate value u_0. Hence, when we speak of the "outgoing mode at late times" we really mean a linear combination of outgoing modes on $\mathcal{J}^+$. Keeping this in mind, we phrase our argument below in terms g_ω for simplicity.

In order to find the expansion of the late mode in terms of early modes, we can trace the solution back in time. As the wave travels inwards from $\mathcal{J}^+$ (toward decreasing values of r_*), it will encounter the potential barrier. One part of the wave, call it $g_\omega^{(r)}$, will be reflected and end up on $\mathcal{J}^-$ with the same frequency ω. This will correspond to a term of the form $A_{\omega\omega'} \propto \delta(\omega - \omega')$ in (5.53). The remaining part $g_\omega^{(t)}$ of g_ω will be transmitted through the barrier and will enter the collapsing matter. In that region, the precise geometry of spacetime is unknown. However, since we are interested in a packet peaked at late times (i.e. large u_0) and at some finite frequency ω_0, we know that the packet will be peaked at a very high frequency as it enters the collapsing matter due to the gravitational blueshift. This leads to an important simplification: the packet will obey the *geometric optics approximation*, which means that g_ω takes the form $A(x)e^{iS(x)}$ where $A(x)$ is slowly varying compared to S. Substituting this into the Klein-Gordon equation gives $\nabla_\mu S \nabla^\mu S = 0$, which means that surfaces of constant phase are null. Given a wave, we can therefore trace its surfaces of constant phase back in time by following null geodesics.

Consider tracing back the wave along a particular null geodesic γ which starts off at some $u = u_0$ at $\mathcal{J}^+$ and hits $\mathcal{J}^-$ at $v = v_0$. as in Figure 21. We denote by γ_H a null generator of the horizon H^+, which has been extended into the past until it hits $\mathcal{J}^-$ at some value of v. We may set this value to $v = 0$ without loss of generality, since the spacetime is invariant under shifts $v \to v + c$. In that case we have $v_0 < 0$ for the geodesic γ. Let n be a connecting vector between the two curves as shown in Figure 21. Fix its normalisation by requiring $n.\ell = -1$, where ℓ is a generator of the Killing horizon H^+. Near the horizon, the Kruskal coordinate $U = -e^{-\kappa u}$ is an affine distance along n and we can use it to measure the distance between γ and γ_H (1.25). In order to find the form of the wave at $\mathcal{J}^-$, we need to understand how the affine distance along the connecting vector n will change by the
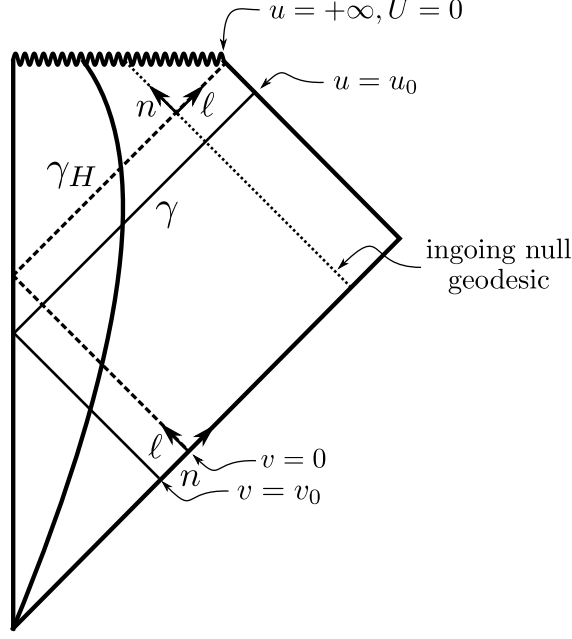
Figure 21. The BLA.

time γ reaches $\mathcal{J}^-$. At $\mathcal{J}^-$, the coordinate v is an affine parameter along the null geodesic integral curves of n. If $U_0 = 0$ (corresponding to $u_0 = \infty$) then the affine distance is zero at $\mathcal{J}^-$. Hence we can expand the affine distance between γ and γ_H at $\mathcal{J}^-$ in powers of U_0: $v = cU_0 + \mathcal{O}(U_0^2)$ for some constant $c > 0$ (since $U, v < 0$). Using $u = -\kappa^{-1}\ln(-U) = -\kappa^{-1}\ln(-cv)$, we can conclude that if a mode takes the form $g_\omega \sim e^{-i\omega u}$ on $\mathcal{J}^+$, the transmitted part $g_\omega^{(t)}$ on $\mathcal{J}^-$ will take the form

$$g_\omega^{(t)} \sim \begin{cases} e^{i\frac{\omega}{\kappa}\ln(-v)} & \text{for} \quad v < 0 \\ 0 & \text{for} \quad v > 0 \end{cases} \tag{5.62}$$

up to a constant phase. This is exactly analogous to the Rindler modes in the previous section, with $\kappa \leftrightarrow a$. We have $A_{\omega\omega'} = e^{-\pi\omega/\kappa}B_{\omega\omega'}$ and therefore

$$\langle N_\omega \rangle \propto \frac{1}{e^{\hbar\omega/k_B T} - 1} \tag{5.63}$$

where the *Hawking temperature* is given by $T = \frac{\hbar\kappa}{2\pi k_B}$. Since this temperature is inversely proportional to the mass, the black hole heats up as it evaporates.