

## 4. The Einstein Equations

It is now time to do some physics. The force of gravity is mediated by a gravitational field. The glory of general relativity is that this field is identified with a metric  $g_{\mu\nu}(x)$  on a 4d Lorentzian manifold that we call spacetime.

This metric is not something fixed; it is, like all other fields in Nature, a dynamical object. This means that there are rules which govern how this field evolves in time. The purpose of this section is to explore these rules and some of their consequences.

We will start by understanding the dynamics of the gravitational field in the absence of any matter. We will then turn to understand how the gravitational field responds to matter – or, more precisely, to energy and momentum – in Section 4.5.

### 4.1 The Einstein-Hilbert Action

All our fundamental theories of physics are described by action principles. Gravity is no different. Furthermore, the straight-jacket of differential geometry places enormous restrictions on the kind of actions that we can write down. These restrictions ensure that the action is something intrinsic to the metric itself, rather than depending on our choice of coordinates.

Spacetime is a manifold  $M$ , equipped with a metric of Lorentzian signature. An action is an integral over  $M$ . We know from Section 2.4.4 that we need a volume-form to integrate over a manifold. Happily, as we have seen, the metric provides a canonical volume form, which we can then multiply by any scalar function. Given that we only have the metric to play with, the simplest such (non-trivial) function is the Ricci scalar  $R$ . This motivates us to consider the wonderfully concise action

$$S = \int d^4x \sqrt{-g} R \tag{4.1}$$

This is the *Einstein-Hilbert action*. Note that the minus sign under the square-root arises because we are in a Lorentzian spacetime: the metric has a single negative eigenvalue and so its determinant,  $g = \det g_{\mu\nu}$ , is negative.

As a quick sanity check, recall that the Ricci tensor takes the schematic form (3.39)  $R \sim \partial\Gamma + \Gamma\Gamma$  while the Levi-Civita connection itself is  $\Gamma \sim \partial g$ . This means that the Einstein-Hilbert action is second order in derivatives, just like most other actions we consider in physics.

## Varying the Einstein-Hilbert Action

We would like to determine the Euler-Lagrange equations arising from the action (4.1). We do this in the usual way, by starting with some fixed metric  $g_{\mu\nu}(x)$  and seeing how the action changes when we shift

$$g_{\mu\nu}(x) \rightarrow g_{\mu\nu}(x) + \delta g_{\mu\nu}(x)$$

Writing the Ricci scalar as  $R = g^{\mu\nu} R_{\mu\nu}$ , the Einstein-Hilbert action clearly changes as

$$\delta S = \int d^4x \left( (\delta\sqrt{-g}) g^{\mu\nu} R_{\mu\nu} + \sqrt{-g} (\delta g^{\mu\nu}) R_{\mu\nu} + \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu} \right) \quad (4.2)$$

It turns out that it's slightly easier to think of the variation in terms of the inverse metric  $\delta g^{\mu\nu}$ . This is equivalent to the variation of the metric  $\delta g_{\mu\nu}$ ; the two are related by

$$g_{\rho\mu} g^{\mu\nu} = \delta_\rho^\nu \quad \Rightarrow \quad (\delta g_{\rho\mu}) g^{\mu\nu} + g_{\rho\mu} \delta g^{\mu\nu} = 0 \quad \Rightarrow \quad \delta g^{\mu\nu} = -g^{\mu\rho} g^{\nu\sigma} \delta g_{\rho\sigma}$$

The middle term in (4.2) is already proportional to  $\delta g^{\mu\nu}$ . We now deal with the first and third terms in turn. We will need the following result:

**Claim:** The variation of  $\sqrt{-g}$  is given by

$$\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g} g_{\mu\nu} \delta g^{\mu\nu}$$

**Proof:** We use the fact that any diagonalisable matrix  $A$  obeys the identity

$$\log \det A = \text{tr} \log A$$

This is obviously true for diagonal matrices. (The determinant is the product of eigenvalues while the trace is the sum of eigenvalues.) But because both the determinant and the trace are invariant under conjugation, it is also true for a diagonalisable matrix. Using this, we have,

$$\frac{1}{\det A} \delta(\det A) = \text{tr}(A^{-1} \delta A)$$

Applying this to the metric, we have

$$\delta\sqrt{-g} = \frac{1}{2} \frac{1}{\sqrt{-g}} (-g) g^{\mu\nu} \delta g_{\mu\nu} = \frac{1}{2} \sqrt{-g} g^{\mu\nu} \delta g_{\mu\nu}$$

Using  $g^{\mu\nu} \delta g_{\mu\nu} = -g_{\mu\nu} \delta g^{\mu\nu}$  then gives the result.  $\square$

So far, we have managed to write the variation of the action (4.2) as

$$\delta S = \int d^4x \sqrt{-g} \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} + \sqrt{-g} g^{\mu\nu} \delta R_{\mu\nu}$$

We now need only worry about the final term. For this, we use:

**Claim:** The variation of the Ricci tensor is a total derivative

$$\delta R_{\mu\nu} = \nabla_\rho \delta \Gamma_{\mu\nu}^\rho - \nabla_\nu \delta \Gamma_{\mu\rho}^\rho$$

where

$$\delta \Gamma_{\mu\nu}^\rho = \frac{1}{2} g^{\rho\sigma} (\nabla_\mu \delta g_{\sigma\nu} + \nabla_\nu \delta g_{\sigma\mu} - \nabla_\sigma \delta g_{\mu\nu})$$

**Proof:** We start by looking at the variation of the Christoffel symbols,  $\Gamma_{\mu\nu}^\rho$ . First note that, although the Christoffel symbol itself is not a tensor, the variation  $\delta \Gamma_{\mu\nu}^\rho$  is a tensor. This is because it is the difference of Christoffel symbols, one computed using  $g_{\mu\nu}$  and the other using  $g_{\mu\nu} + \delta g_{\mu\nu}$ . But the extra derivative term in the transformation of  $\Gamma_{\mu\nu}^\rho$  is independent of the metric and so cancels out when we take the difference, leaving us with an object which transforms nicely as a tensor.

This is a useful observation. At any point  $p \in M$  we can choose to work in normal coordinates such that  $\partial_\rho g_{\mu\nu} = 0$  and, correspondingly,  $\Gamma_{\mu\nu}^\rho = 0$ . Then, to linear order in the variation, the change in the Christoffel symbol evaluated at  $p$  is

$$\begin{aligned} \delta \Gamma_{\mu\nu}^\rho &= \frac{1}{2} g^{\rho\sigma} (\partial_\mu \delta g_{\sigma\nu} + \partial_\nu \delta g_{\sigma\mu} - \partial_\sigma \delta g_{\mu\nu}) \\ &= \frac{1}{2} g^{\rho\sigma} (\nabla_\mu \delta g_{\sigma\nu} + \nabla_\nu \delta g_{\sigma\mu} - \nabla_\sigma \delta g_{\mu\nu}) \end{aligned}$$

where we're at liberty to replace the partial derivatives with covariant derivatives because they differ only by the Christoffel symbols  $\Gamma_{\mu\nu}^\rho$  which, in normal coordinates, vanish at  $p$ . However, both the left and right-hand sides of this equation are tensors which means that although we derived this expression using normal coordinates, it must hold in any coordinate system. Moreover, the point  $p$  was arbitrary so the final expression holds generally.

Next we look at the variation of the Riemann tensor. In normal coordinates, the expression (3.39) becomes

$$R^\sigma{}_{\rho\mu\nu} = \partial_\mu \Gamma_{\nu\rho}^\sigma - \partial_\nu \Gamma_{\mu\rho}^\sigma$$

and the variation is

$$\delta R^\sigma{}_{\rho\mu\nu} = \partial_\mu \delta\Gamma^\sigma{}_{\nu\rho} - \partial_\nu \delta\Gamma^\sigma{}_{\mu\rho} = \nabla_\mu \delta\Gamma^\sigma{}_{\nu\rho} - \nabla_\nu \delta\Gamma^\sigma{}_{\mu\rho}$$

where, as before, we replace partial derivatives with covariant derivatives as we are working in normal coordinates where the Christoffel symbols vanish. Once again, our final expression relates two tensors and must, therefore, hold in any coordinate system. Contracting indices (and working to leading order), we have

$$\delta R_{\rho\nu} = \nabla_\mu \delta\Gamma^\mu{}_{\nu\rho} - \nabla_\nu \delta\Gamma^\mu{}_{\rho\mu}$$

as claimed.  $\square$

The upshot of these calculations is that

$$g^{\mu\nu} \delta R_{\mu\nu} = \nabla_\mu X^\mu \quad \text{with} \quad X^\mu = g^{\rho\nu} \delta\Gamma^\mu{}_{\rho\nu} - g^{\mu\nu} \delta\Gamma^\rho{}_{\nu\rho}$$

The variation of the action (4.2) can then be written as

$$\delta S = \int d^4x \sqrt{-g} \left[ \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) \delta g^{\mu\nu} + \nabla_\mu X^\mu \right] \quad (4.3)$$

This final term is a total derivative and, by the divergence theorem of Section 3.2.4, we ignore it. Requiring that the action is extremised, so  $\delta S = 0$ , we have the equations of motion

$$G_{\mu\nu} := R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 0 \quad (4.4)$$

where  $G_{\mu\nu}$  is the Einstein tensor defined in Section 3.4.1. These are the *Einstein field equations* in the absence of any matter. In fact they simplify somewhat: if we contract (4.4) with  $g^{\mu\nu}$ , we find that we must have  $R = 0$ . Substituting this back in, the vacuum Einstein equations are simply the requirement that the metric is Ricci flat,

$$R_{\mu\nu} = 0 \quad (4.5)$$

These deceptively simple equations hold a myriad of surprises. We will meet some of the solutions as we go along, notably gravitational waves in Section 5.2 and black holes in Section 6.

Before we proceed, a small comment. We happily discarded the boundary term in (4.3), a standard practice whenever we invoke the variational principle. It turns out that there are some situations in general relativity where we should not be quite so cavalier. In such circumstances, one can be more careful by invoking the so-called Gibbons-Hawking boundary term.

### 4.1.1 An Aside on Dimensional Analysis

As it stands, there's something a little fishy about the action (4.1): it doesn't have the right dimensions. This isn't such an issue since we have just a single term in the action and multiplying the action by a constant doesn't change the classical equations of motion. Nonetheless, it will prove useful to get it right at this stage.

If we take the coordinates  $x^\mu$  to have dimension of length, then the metric  $g_{\mu\nu}$  is necessarily dimensionless. The Ricci scalar involves two spatial derivatives so has dimension  $[R] = L^{-2}$ . Including the integration measure, the action (4.1) then has dimensions  $[S] = L^2$ . However, actions should have dimensions of energy  $\times$  time (it's the same dimensions as  $\hbar$ ), or  $[S] = ML^2T^{-1}$ . This means that the Einstein-Hilbert action should be multiplied by a constant with the appropriate dimensions. We take

$$S = \frac{c^3}{16\pi G} \int d^4x \sqrt{-g} R$$

where  $c$  is the speed of light and  $G$  is Newton's constant,

$$G \approx 6.67 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$$

This factor doesn't change the equation of motion in vacuum, but we will see in Section 4.5 that it determines the strength of the coupling between the gravitational field and matter, as we might expect.

It's no fun carrying around a morass of fundamental constants in all our equations. For this reason, we often work in "natural units" in which various constants are set equal to 1. From now on, we will set  $c = 1$ . (Any other choice of  $c$ , including  $3 \times 10^8$ , is simply dumb.) This means that units of length and time are equated.

However, different communities have different conventions when it comes to  $G$ . Relativists will typically set  $G = 1$ . Since we have already set  $c = 1$ , we have  $[G] = LM^{-1}$ . Setting  $G = 1$  then equates mass with length. This is useful when discussing gravitational phenomenon where the mass is often directly related to the length. For example, the Schwarzschild radius of black hole is  $R_s = 2GM/c^2$  which becomes simply  $R_s = 2M$  once we set  $G = c = 1$ .

However, if you're interested in phenomena other than gravity, then it's no more sensible to set  $G = 1$  than to set, say, the Fermi coupling for the weak force  $G_F = 1$ . Instead, it is more useful to choose the convention where  $\hbar = 1$ , a choice which equates energy with inverse time (also known as frequency). With this convention, Newton's

constant has dimension  $[G] = M^{-2}$ . The corresponding energy scale is known as the (reduced) Planck mass; it is given by

$$M_{\text{pl}}^2 = \frac{\hbar c}{8\pi G}$$

It is around  $10^{18}$  GeV. This is a very high energy scale, way beyond anything we have probed in experiment. This can be traced to the weakness of the gravitational force. With  $c = \hbar = 1$ , we can equally well write the Einstein-Hilbert action as

$$S = \frac{1}{2} M_{\text{pl}}^2 \int d^4x \sqrt{-g} R$$

You might be tempted to set  $c = \hbar = G = 1$ . This leaves us with no remaining dimensional quantities. It is typically a bad idea, not least because dimensional analysis is a powerful tool and one we do not want to lose. In these lectures, we will focus only on gravitational physics. Nonetheless, we will retain  $G$  in all equations.

#### 4.1.2 The Cosmological Constant

We motivated the Einstein-Hilbert action as the simplest term we could write down. While it's true that it's the simplest term that results in interesting dynamics for the gravitational field, there is in fact a simpler term which we could add to the action. This comes from multiplying the volume form by a constant. The resulting action is

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} (R - 2\Lambda)$$

Here  $\Lambda$  is referred to as the *cosmological constant*. It has dimension  $[\Lambda] = L^{-2}$ . The minus sign in the action comes from thinking of the Lagrangian as “ $T - V$ ”: the cosmological constant is like the potential energy  $V$ .

Varying the action as before now yields the Einstein equations,

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = -\Lambda g_{\mu\nu}$$

This time, if we contract with  $g^{\mu\nu}$ , we get  $R = 4\Lambda$ . Substituting this back in, the vacuum Einstein equations in the presence of a cosmological constant become

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

We will solve these shortly in Section 4.2.

## Higher Derivative Terms

The Einstein-Hilbert action (with cosmological constant) is the simplest thing we can write down but it is not the only possibility, at least if we allow for higher derivative terms. For example, there are three terms that contain four derivatives of the metric,

$$S_{4\text{-deriv}} = \int d^4x \sqrt{-g} (c_1 R^2 + c_2 R_{\mu\nu} R^{\mu\nu} + c_3 R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma})$$

with  $c_1$ ,  $c_2$  and  $c_3$  dimensionless constants. General choices of these constants will result in higher order equations of motion which do not have a well-defined initial value problem. Nonetheless, it turns out that one can find certain combinations of these terms which conspire to keep the equations of motion second order. This is known as *Lovelock's theorem*. In  $d = 4$  dimensions, this combination has a rather special topological property: a generalisation of the Gauss-Bonnet theorem states that

$$\frac{1}{8\pi^2} \int_M d^4x \sqrt{g} (R^2 - 4R_{\mu\nu} R^{\mu\nu} + R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}) = \chi(M)$$

where  $\chi(M) \in \mathbf{Z}$  is the Euler character of  $M$  that we previously defined in (2.35). In Lorentzian signature, this combination of curvature terms is also a total derivative and does not affect the classical equations of motion.

As in any field theory, higher derivative terms in the action only become relevant for fast varying fields. In General Relativity, they are unimportant for all observed physical phenomena and we will not discuss them further in this course.

### 4.1.3 Diffeomorphisms Revisited

Here's a simple question: how many degrees of freedom are there in the metric  $g_{\mu\nu}$ ? Since this is a symmetric  $4 \times 4$  matrix, our first guess is  $\frac{1}{2} \times 4 \times 4 = 10$ .

However, not all the components of the metric  $g_{\mu\nu}$  are physical. Two metrics which are related by a change of coordinates,  $x^\mu \rightarrow \tilde{x}^\mu(x)$  describe the same physical space-time. This means that there is a redundancy in any given representation of the metric, which removes precisely 4 of the 10 degrees of freedom, leaving just 6 behind.

Mathematically, this redundancy is implemented by diffeomorphisms. (We defined diffeomorphisms in Section 2.1.3.) Given a diffeomorphism,  $\phi : M \rightarrow M$ , we can use this to map all fields, including the metric, on  $M$  to a new set of fields on  $M$ . The end result is physically indistinguishable from where we started: it describes the same system, but in different coordinates. Such diffeomorphisms are analogous to the gauge symmetries that are familiar in Maxwell and Yang-Mills theory.

Let's look more closely at the implication of these diffeomorphisms for the path integral. We'll consider a diffeomorphism that takes a point with coordinates  $x^\mu$  to a nearby point with coordinates

$$x^\mu \rightarrow \tilde{x}^\mu(x) = x^\mu + \delta x^\mu$$

We could view this either as an “active change”, in which one point with coordinates  $x^\mu$  is mapped to another point with coordinates  $x^\mu + \delta x^\mu$ , or as a “passive” change, in which we use two different coordinate charts to label the same point. Ultimately, the two views lead to the same place. We'll adopt the passive perspective here, simply because we have a lot of experience of changing coordinates. Later we'll revert to the active picture.

We can think of the change of coordinates as being generated by an infinitesimal vector field  $X$ ,

$$\delta x^\mu = -X^\mu(x)$$

The metric transforms as

$$g_{\mu\nu}(x) \rightarrow \tilde{g}_{\mu\nu}(\tilde{x}) = \frac{\partial x^\rho}{\partial \tilde{x}^\mu} \frac{\partial x^\sigma}{\partial \tilde{x}^\nu} g_{\rho\sigma}(x)$$

With our change of coordinate  $\tilde{x}^\mu = x^\mu - X^\mu(x)$ , with infinitesimal  $X^\mu$ , we can invert the Jacobian matrix to get

$$\frac{\partial \tilde{x}^\mu}{\partial x^\rho} = \delta_\rho^\mu - \partial_\rho X^\mu \quad \Rightarrow \quad \frac{\partial x^\rho}{\partial \tilde{x}^\mu} = \delta_\mu^\rho + \partial_\mu X^\rho$$

where the inverse holds to leading order in the variation  $X$ . Continuing to work infinitesimally, we then have

$$\begin{aligned} \tilde{g}_{\mu\nu}(\tilde{x}) &= (\delta_\mu^\rho + \partial_\mu X^\rho) (\delta_\nu^\sigma + \partial_\nu X^\sigma) g_{\rho\sigma}(x) \\ &= g_{\mu\nu}(x) + g_{\mu\rho}(x) \partial_\nu X^\rho + g_{\nu\rho}(x) \partial_\mu X^\rho \end{aligned}$$

Meanwhile, we can Taylor expand the left-hand side

$$\tilde{g}_{\mu\nu}(\tilde{x}) = \tilde{g}_{\mu\nu}(x + \delta x) = \tilde{g}_{\mu\nu}(x) - X^\lambda \partial_\lambda \tilde{g}_{\mu\nu}(x)$$

Comparing the the different metrics at the same point  $x$ , we find that the metric undergoes the infinitesimal change

$$\delta g_{\mu\nu}(x) = \tilde{g}_{\mu\nu}(x) - g_{\mu\nu}(x) = X^\lambda \partial_\lambda g_{\mu\nu} + g_{\mu\rho} \partial_\nu X^\rho + g_{\nu\rho} \partial_\mu X^\rho \quad (4.6)$$

But this is something we've seen before: it is the Lie derivative of the metric. In other words, if we act with an infinitesimal diffeomorphism along  $X$ , the metric changes as

$$\delta g_{\mu\nu} = (\mathcal{L}_X g)_{\mu\nu}$$

This makes sense: it's like the leading term in a Taylor expansion along  $X$ .



In fact, we can also massage (4.6) into a slightly different form. We lower the index on  $X^\rho$  in the last two  $\partial X^\rho$  terms by taking the metric inside the derivative. This results in two further terms in which the derivative hits the metric, and these must be cancelled off. We're left with

$$\delta g_{\mu\nu} = \partial_\mu X_\nu + \partial_\nu X_\mu + X^\rho (\partial_\rho g_{\mu\nu} - \partial_\mu g_{\rho\nu} - \partial_\nu g_{\mu\rho})$$

But the terms in the brackets are the Christoffel symbols,  $2g_{\rho\sigma}\Gamma_{\mu\nu}^\sigma$ . We learn that the infinitesimal change in the metric can be written as

$$\delta g_{\mu\nu} = \nabla_\mu X_\nu + \nabla_\nu X_\mu \tag{4.7}$$

Let's now see what this means for the path integral. Under a general change of the metric, the Einstein-Hilbert action changes as (4.3)

$$\delta S = \int d^4x \sqrt{-g} G^{\mu\nu} \delta g_{\mu\nu}$$

where we have discarded the boundary term. Insisting that  $\delta S = 0$  for *any* variation  $\delta g_{\mu\nu}$  gives the equation of motion  $G^{\mu\nu} = 0$ . In contrast, symmetries of the action are those variations  $\delta g_{\mu\nu}$  for which  $\delta S = 0$  for any choice of metric. Since diffeomorphisms are (gauge) symmetries, we know that the action is invariant under changes of the form (4.7). Using the fact that  $G_{\mu\nu}$  is symmetric, we must have

$$\delta S = 2 \int d^4x \sqrt{-g} G^{\mu\nu} \nabla_\mu X_\nu = 0 \quad \text{for all } X_\mu(x)$$

After integrating by parts, we find that this results in something familiar: the Bianchi identity

$$\nabla_\mu G^{\mu\nu} = 0$$

We already know that the Bianchi identity holds from our work in Section 3.4, but the derivation there was a little fiddly. Here we learn that, from the path integral perspective, the Bianchi identity is a result of diffeomorphism invariance.

In fact it makes sense that the two are connected. Naively, the Einstein equation  $G_{\mu\nu} = 0$  comprises ten independent equations. But, as we've seen, diffeomorphism invariance means that there aren't ten independent components of the metric, so one might worry that the Einstein equations are overdetermined. Happily, diffeomorphisms also ensure that not all the Einstein equations are independent either; they are related by the four Bianchi constraints. We see that, in fact, the Einstein equations give only six independent conditions on the six independent degrees of freedom in the metric.

## 4.2 Some Simple Solutions

We will now look for some simple solutions to the Einstein equations

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

As we will see, the solutions take a very different form depending on whether  $\Lambda$  is zero, positive or negative.

### Minkowski Space

Let's start with  $\Lambda = 0$ . Here the vacuum Einstein equations reduce to  $R_{\mu\nu} = 0$ . If we're looking for the simplest solution to this equation, it's tempting to suggest  $g_{\mu\nu} = 0$ . Needless to say, this isn't allowed! The tensor field  $g_{\mu\nu}$  is a metric and, as defined in Section 3, must be non-degenerate. Indeed, the existence of the inverse  $g^{\mu\nu}$  was assumed in the derivation of the Einstein equations from the action.

While this restriction is natural geometrically, it is rather unusual from the perspective of a physical theory. It is not a holonomic constraint on the physical degrees of freedom: instead it is an inequality  $\det g_{\mu\nu} < 0$  (together with the requirement that  $g_{\mu\nu}$  has one, rather than three, negative eigenvalues). Other fields in the Standard Model don't come with such restrictions. Instead, it is reminiscent of fluid mechanics where one has to insist that matter density obeys  $\rho(\mathbf{x}, t) > 0$ . Ultimately, it seems likely that this restriction is telling us that the gravitational field is not fundamental and should be replaced by something else in regimes where  $\det g_{\mu\nu}$  is getting small.

The restriction that  $\det g_{\mu\nu} \neq 0$  means that the simplest Ricci flat metric is Minkowski space, with

$$ds^2 = -dt^2 + d\mathbf{x}^2$$

Of course, this is far from the only metric obeying  $R_{\mu\nu} = 0$ . Another example is provided by the Schwarzschild metric,

$$ds^2 = -\left(1 - \frac{2GM}{r}\right) dt^2 + \left(1 - \frac{2GM}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (4.8)$$

which we will discuss further in Section 6. We will meet more solutions as the course progresses.

### 4.2.1 de Sitter Space

We now turn to the Einstein equations with  $\Lambda > 0$ . Once again, there are many solutions. Since it's a pain to solve the Einstein equations, let's work with an ansatz that we've already seen. Suppose that we look for solutions of the form

$$ds^2 = -f(r)^2 dt^2 + f(r)^{-2} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad (4.9)$$

We already computed the components of the Riemann tensor for such a metric in Section 3.4.3 using the technology of curvature 2-forms. From the result, given in (3.54), we can easily check that the Ricci tensor is diagonal with components

$$R_{tt} = -f^4 R_{rr} = f^3 \left( f'' + \frac{2f'}{r} + \frac{f'^2}{f} \right)$$

and

$$R_{\phi\phi} = \sin^2 \theta R_{\theta\theta} = (1 - f^2 - 2ff'r) \sin^2 \theta$$

The resulting Ricci tensor can indeed be made to be proportional to the metric, with  $R_{\mu\nu} = \Lambda g_{\mu\nu}$ . Comparing to (4.9), we see that the function  $f(r)$  must satisfy two constraints. The first comes from the  $tt$  and  $rr$  components,

$$f'' + \frac{2f'}{r} + \frac{f'^2}{f} = -\frac{\Lambda}{f} \quad (4.10)$$

The second comes from the  $\theta\theta$  and  $\phi\phi$  components,

$$1 - 2ff'r - f^2 = \Lambda r^2 \quad (4.11)$$

It's simple to see that both conditions are satisfied by the choice

$$f(r) = \sqrt{1 - \frac{r^2}{R^2}} \quad \text{with} \quad R^2 = \frac{3}{\Lambda}$$

The resulting metric takes the form

$$ds^2 = - \left( 1 - \frac{r^2}{R^2} \right) dt^2 + \left( 1 - \frac{r^2}{R^2} \right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad (4.12)$$

This is *de Sitter space*. Or, more precisely, it is the *static patch* of de Sitter space; we'll see what this latter statement means shortly.

## Geodesics in de Sitter

To interpret this metric, it's useful to understand the behaviour of geodesics. We can see immediately that the presence of the non-trivial  $g_{tt}(r)$  term means that a particle won't sit still at constant  $r \neq 0$ ; instead it is pushed to smaller values of  $g_{tt}(r)$ , or larger values of  $r$ .

We can put some more flesh on this. Because the metric (4.12) has a similar form to the Schwarzschild metric, we simply need to follow the steps that we already took in Section 1.3. First we write down the action for a particle in de Sitter space. We denote the proper time of the particle as  $\sigma$ . (In Section 1.3, we used  $\tau$  to denote proper time, but we'll need this for a different time coordinate defined below.) Working with the more general metric (4.9), the action is

$$S_{dS} = \int d\sigma \left[ -f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) \right] \quad (4.13)$$

where  $\dot{x}^\mu = dx^\mu/d\sigma$ .

Any degree of freedom which appears only with time derivatives in the Lagrangian is called *ignorable*. They lead to conserved quantities. The Lagrangian above has two ignorable degrees of freedom:  $\phi(\sigma)$  and  $t(\sigma)$ . The first leads to the conserved quantity that we call angular momentum,

$$l = \frac{1}{2} \frac{dL}{d\dot{\phi}} = r^2 \sin^2 \theta \dot{\phi}$$

where the factor of 1/2 in front of  $dL/d\dot{\phi}$  arises because the kinetic terms in (4.13) don't come with the usual factor of 1/2. Meanwhile, the conserved quantity associated to  $t(\sigma)$  is usually referred to as the energy

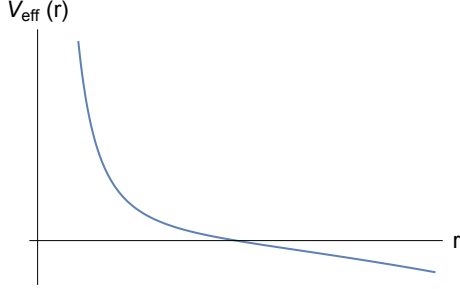
$$E = -\frac{1}{2} \frac{dL}{d\dot{t}} = f(r)^2 \dot{t} \quad (4.14)$$

The equations of motion arising from the action (4.13) should be supplemented with the constraint that tells us whether we're dealing with a massive or massless particle. For a massive particle, the constraint ensures that the trajectory is timelike,

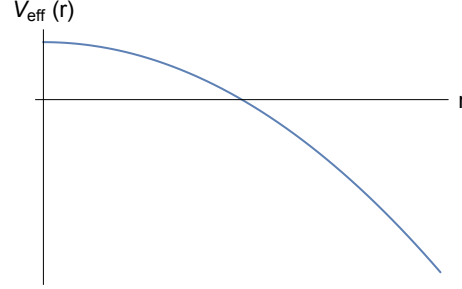
$$-f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2 (\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) = -1$$

Without loss of generality, we can restrict to geodesics that lie in the  $\theta = \pi/2$  plane, so  $\dot{\theta} = 0$  and  $\sin^2 \theta = 1$ . Replacing  $\dot{t}$  and  $\dot{\phi}$  with  $E$  and  $l$  respectively, the constraint becomes

$$\dot{r}^2 + V_{\text{eff}}(r) = E^2$$



**Figure 27:** The effective potential for a massive particle in de Sitter with angular momentum..



**Figure 28:** ...and with no angular momentum.

where the effective potential is given by

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right) f(r)^2$$

For geodesics in de Sitter, we therefore have

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right) \left(1 - \frac{r^2}{R^2}\right)$$

This is shown in the figures for  $l \neq 0$  and for  $l = 0$ . We can immediately see the key physics: the potential pushes the particle out to larger values of  $r$ .

We focus on geodesics with vanishing angular momentum,  $l = 0$ . In this case, the potential is an inverted harmonic oscillator. A particle sitting stationary at  $r = 0$  is a geodesic, but it is unstable: if it has some initial velocity then it will move away from the origin, following the trajectory

$$r(\sigma) = R\sqrt{E^2 - 1} \sinh\left(\frac{\sigma}{R}\right) \quad (4.15)$$

The metric (4.12) is singular at  $r = R$ , which might make us suspect that something fishy is going on there. But whatever this fishiness is, it's not visible in the solution (4.15) which shows that any observer reaches  $r = R$  in finite proper time  $\sigma$ .

The fishiness reveals itself if we look at the coordinate time  $t$ . This also has the interpretation of the time experienced by someone sitting at the point  $r = 0$ . Using (4.14), the trajectory (4.15) evolves as

$$\frac{dt}{d\sigma} = E \left(1 - \frac{r^2}{R^2}\right)^{-1}$$

It is simple to check that  $t(\sigma) \rightarrow \infty$  as  $r(\sigma) \rightarrow R$ . (For example, suppose that we have  $r(\sigma) = R$  at some value  $\sigma = \sigma_*$  of proper time. Then look at what happens just before this time by expanding  $\sigma = \sigma_* - \epsilon$  with  $\epsilon$  small. The equation above becomes  $dt/d\epsilon = -\alpha/\epsilon$  for some constant  $\alpha$ , telling us that  $t(\epsilon) \sim -\log(\epsilon/R)$  and we indeed find that  $t \rightarrow \infty$  as  $\epsilon \rightarrow 0$ .) This means that while a guy on the trajectory (4.15) sails right through the point  $r = R$  in finite proper time, according to his companion waiting at  $r = 0$  this will take infinite time.

This strange behaviour is, it turns out, similar to what happens at the horizon of a black hole, which is the surface  $r = 2GM$  in the metric (4.8). (We will look at more closely at this in Section 6.) However, the Schwarzschild metric also has a singularity at  $r = 0$ , whereas the de Sitter metric looks just like flat space at  $r = 0$ . (To see this, simply Taylor expand the coefficients of the metric around  $r = 0$ .) Instead, de Sitter space seems like an inverted black hole in which particles are pushed outwards to  $r = R$ . But how should we interpret this radius? We will get more intuition for this as we proceed.

### de Sitter Embeddings

We will have to wait until Section 4.4.2 to get a full understanding of the physics behind this. But we can make some progress by writing the de Sitter metric in different coordinates. In fact, it turns out that there's a rather nice way of embedding de Sitter space as a sub-manifold of  $\mathbf{R}^{1,4}$ , with metric

$$ds^2 = -(dX^0)^2 + \sum_{i=1}^4 (dX^i)^2 \quad (4.16)$$

We will now show that the de Sitter space metric (4.12) is a metric on the sub-manifold in  $\mathbf{R}^{1,4}$  defined by the timelike hyperboloid

$$-(X^0)^2 + \sum_{i=1}^4 (X^i)^2 = R^2 \quad (4.17)$$

There are a number of different ways to parameterise solutions to this constraint. Suppose that we choose to treat  $X^4$  as a special coordinate. We define the sum of the first three spatial coordinates to be

$$r^2 = (X^1)^2 + (X^2)^2 + (X^3)^2 \quad (4.18)$$

so the constraint (4.17) becomes

$$R^2 - r^2 = -(X^0)^2 + (X^4)^2$$

We can parameterise solutions to this equation as

$$X^0 = \sqrt{R^2 - r^2} \sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2} \cosh(t/R) \quad (4.19)$$

The variation is then

$$\begin{aligned} dX^0 &= \sqrt{1 - \frac{r^2}{R^2}} \cosh(t/R) dt - \frac{r}{\sqrt{R^2 - r^2}} \sinh(t/R) dr \\ dX^4 &= \sqrt{1 - \frac{r^2}{R^2}} \sinh(t/R) dt - \frac{r}{\sqrt{R^2 - r^2}} \cosh(t/R) dr \end{aligned}$$

Meanwhile the variation of  $X^i$ , with  $i = 1, 2, 3$ , is just the familiar line element for  $\mathbf{R}^3$ :  $\sum_{i=1}^3 (dX^i)^2 = dr^2 + r^2 d\Omega_2^2$  where  $d\Omega_2^2$  is the metric on the unit 2-sphere. A two line calculation then shows that the pull-back of the 5d Minkowski metric (4.16) onto the hypersurface (4.17) gives the de Sitter metric in the static patch coordinates (4.12).

The choice of coordinates (4.18) and (4.19) are not the most intuitive. First, they single out  $X^4$  as special, when the constraint (4.17) does no such thing. This hides some of the symmetry of de Sitter space. Moreover, the coordinates do not cover the whole of the hyperboloid, since they restrict only to  $X^4 \geq 0$ .

We can do better. Consider instead the solution to the constraint (4.17)

$$X^0 = R \sinh(\tau/R) \quad \text{and} \quad X^i = R \cosh(\tau/R) y^i \quad (4.20)$$

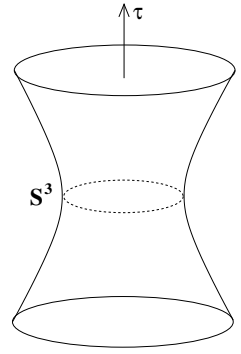
where the  $y^i$ , with  $i = 1, 2, 3, 4$ , obey  $\sum_i (y^i)^2 = 1$  and so parameterise a unit 3-sphere. These coordinates have the advantage that they retain (more of) the symmetry of de Sitter space, and cover the whole space. Substituting this into the 5d Minkowski metric (4.16) gives a rather different metric on de Sitter space,

$$ds^2 = -d\tau^2 + R^2 \cosh^2(\tau/R) d\Omega_3^2 \quad (4.21)$$

where  $d\Omega_3^2$  denotes the metric on the unit 3-sphere. These are known as *global coordinates*, since they cover the whole space. (Admittedly, any choice of coordinates on  $\mathbf{S}^3$  will suffer from the familiar problem of coordinate singularities at the poles.) Since this metric is related to (4.12) by a change of coordinates, it too must obey the Einstein equation. (We'll check this explicitly in Section 4.6 where we discuss a class of metrics of this form.)

These coordinates provide a much clearer intuition for the physics of de Sitter space: it is a time-dependent solution in which a spatial  $\mathbf{S}^3$  first shrinks to a minimal radius  $R$ , and subsequently expands. This is shown in the figure. The expansionary phase is a fairly good approximation to our current universe on large scales; you can learn more about this in the lectures on [Cosmology](#).

The cosmological interpretation of an expanding universe is much harder to glean from the static patch coordinates (4.12) in which the space appears to be unchanging in time. Indeed, de Sitter himself originally discovered the metric in the static patch coordinates. He noticed that light is redshifted in this metric, which then caused all sorts of confusion when trying to understand whether the redshift of galaxies (then known as the de Sitter effect!) should be viewed as evidence for an expanding universe. There is a lesson here: it can be difficult to stare at a metric and get a sense for what you're looking at.



**Figure 29:**

The global coordinates clearly show that there is nothing fishy happening when  $X^4 = 0$ , the surface which corresponds to  $r = R$  in (4.12). This is telling us that this is nothing but a coordinate singularity. (As, indeed, is the  $r = 2GM$  singularity in the Schwarzschild metric.) Nonetheless, there is still some physics lurking in this coordinate singularity, which we will extract over the next few sections.

#### 4.2.2 Anti-de Sitter Space

We again look for solutions to the Einstein equations,

$$R_{\mu\nu} = \Lambda g_{\mu\nu}$$

now with a negative cosmological constant  $\Lambda < 0$ . We can again use the ansatz (4.9) and again find the constraints (4.10) and (4.11). The fact that  $\Lambda$  is now negative means that our previous version of  $f(r)$  no longer works, but it's not hard to find the tweak: the resulting metric takes the form

$$ds^2 = - \left(1 + \frac{r^2}{R^2}\right) dt^2 + \left(1 + \frac{r^2}{R^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (4.22)$$

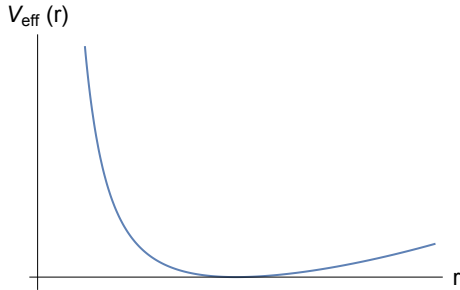
with  $R^2 = -3/\Lambda$ . This is the metric of *anti-de Sitter space*, also known simply as AdS.

Sometimes this metric is written by introducing the coordinate  $r = R \sinh \rho$ , after which it takes the form

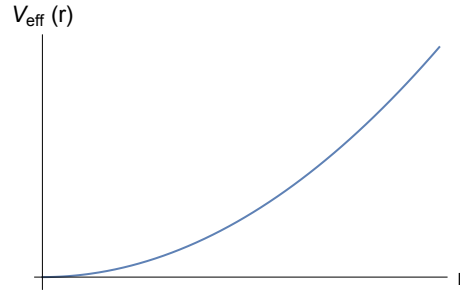
$$ds^2 = - \cosh^2 \rho dt^2 + R^2 d\rho^2 + R^2 \sinh^2 \rho (d\theta^2 + \sin^2 \theta d\phi^2) \quad (4.23)$$

Now there's no mysterious coordinate singularity in the  $r$  direction and, indeed, we will see shortly that these coordinates now cover the entire space.





**Figure 30:** The effective potential for a massive particle in anti-de Sitter with angular momentum..



**Figure 31:** ...and with no angular momentum.

### Geodesics in Anti-de Sitter

Because the anti-de Sitter metric (4.22) falls in the general class (4.9), we can import the geodesic equations that we derived for de Sitter space. The radial trajectory of a massive particle moving in the  $\theta = \pi/2$  plane is again governed by

$$\dot{r}^2 + V_{\text{eff}}(r) = E^2 \quad (4.24)$$

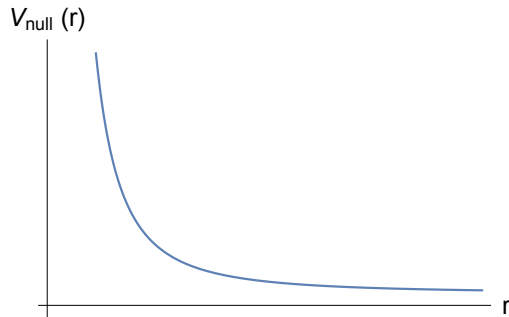
but this time with the effective potential

$$V_{\text{eff}}(r) = \left(1 + \frac{l^2}{r^2}\right) \left(1 + \frac{r^2}{R^2}\right)$$

Again,  $l = r^2 \dot{\phi}$  is the angular momentum of the particle. This potential is shown in the figures for  $l \neq 0$  and  $l = 0$ . From this, we can immediately see how geodesics behave. If there is no angular momentum, so  $l = 0$ , anti-de Sitter space acts like a harmonic potential, pushing the particle towards the origin  $r = 0$ . Geodesics oscillate backwards and forwards around  $r = 0$ .

In contrast, if the particle also has angular momentum then the potential has a minimum at  $r_\star^2 = Rl$ . This geodesic is like a motorcycle wall-of-death trick, with the angular momentum keeping the particle pinned up the potential. Other geodesics spin in the same fashion, while oscillating about  $r_\star$ . Importantly, particles with finite energy  $E$  cannot escape to  $r \rightarrow \infty$ : they are trapped by the spacetime to live within some finite distance of the origin.

The picture that emerges from this analysis is that AdS is like a harmonic trap, pushing particles to the origin. This comes with something of a puzzle however because, as we will see below (and more in Section 4.3), AdS is a homogeneous space which, roughly speaking, means that all points are the same. How is it possible that AdS acts like a harmonic trap, pushing particles to  $r = 0$ , yet is also a homogeneous space?!



**Figure 32:** The potential experienced by massless particles in AdS.

To answer this question, consider a guy sitting stationary at the origin  $r = 0$ . This is a geodesic. From his perspective, intrepid AdS explorers on other geodesics (with, say,  $l = 0$ ) will oscillate backwards and forwards about the origin  $r = 0$ , just like a particle in a harmonic trap. However, since these explorers will themselves be travelling on a geodesic, they are perfectly entitled to view themselves as sedentary, stay-at-home types, sitting perfectly still at their ‘origin’, watching the other folk fly around them. In this way, just as everyone in de Sitter can view themselves at the centre of the universe, with other observers moving away from them, everyone in anti-de Sitter can view themselves in the centre of the universe, with other observers flying around them.

We can also look at the fate of massless particles. This time the action is supplemented by the constraint

$$-f(r)^2 \dot{t}^2 + f(r)^{-2} \dot{r}^2 + r^2(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) = 0$$

This tells us that the particle follows a null geodesic. The equation (4.24) gets replaced by

$$\dot{r}^2 + V_{\text{null}}(r) = E^2$$

with the effective potential now given by

$$V_{\text{null}}(r) = \frac{l^2}{r^2} \left( 1 + \frac{r^2}{R^2} \right)$$

This potential is again shown in Figure 32. This time the potential is finite as  $r \rightarrow \infty$ , which tells us that there is no obstacle to light travelling as far as it likes: it suffers only the usual gravitational redshift. We learn that AdS spacetime confines massive particles, but not massless ones.

To solve the equations for a massless particle, it's simplest to work in the coordinates  $r = R \sinh \rho$  that we introduced in (4.23). If we restrict to vanishing angular momentum,  $l = 0$ , the equation above becomes

$$R\dot{\rho} = \pm \frac{E}{\cosh \rho} \quad \Rightarrow \quad R \sinh \rho = E(\sigma - \sigma_0)$$

where  $\sigma$  is the affine geodesic parameter. We see that  $\rho \rightarrow \infty$  only in infinite affine time,  $\sigma \rightarrow \infty$ . However, it's more interesting to see what happens in coordinate time. This follows by recalling the definition of  $E$  in (4.14),

$$E = \cosh^2 \rho \dot{t}$$

(Equivalently, you can see this by dint of the fact that we have a null geodesic, with  $\cosh \rho \dot{t} = \pm R \dot{\rho}$ .) We then find

$$R \tan(t/R) = E(\sigma - \sigma_0)$$

So as  $\sigma \rightarrow \infty$ , the coordinate time tends to  $t \rightarrow \pi R/2$ . We learn that not only do light rays escape to  $\rho = \infty$ , but they do so in a finite coordinate time  $t$ . This means that to make sense of dynamics in AdS, we must specify some boundary conditions at infinity to dictate what happens to massless particles or fields when they reach it.

Anti-de Sitter space does not appear to have any cosmological applications. However, it turns out to be the place where we best understand quantum gravity, and so has been the object of a great deal of study.

### Anti-de Sitter Embeddings

Like its  $\Lambda > 0$  cousin, anti-de Sitter spacetime also has a natural embedding in a 5d spacetime. This time, it sits within  $\mathbf{R}^{2,3}$ , with metric

$$ds^2 = -(dX^0)^2 - (dX^4)^2 + \sum_{i=1}^3 (dX^i)^2 \quad (4.25)$$

where it lives as the hyperboloid,

$$-(X^0)^2 - (X^4)^2 + \sum_{i=1}^3 (X^i)^2 = -R^2 \quad (4.26)$$

We can solve this constraint by

$$X^0 = R \cosh \rho \sin(t/R) \quad , \quad X^4 = R \cosh \rho \cos(t/R) \quad , \quad X^i = R y^i \sinh \rho \quad (4.27)$$

where the  $y^i$ , with  $i = 1, 2, 3$ , obey  $\sum_i (y^i)^2 = 1$  and so parameterise a unit 2-sphere. Substituting this into the metric (4.25) gives the anti-de Sitter metric in the coordinates (4.23).

In fact there is one small subtlety: the embedding hyperboloid has topology  $\mathbf{S}^1 \times \mathbf{R}^3$ , with  $\mathbf{S}^1$  corresponding to a compact time direction. This can be seen in the parameterisation (4.27), where the time coordinate takes values  $t \in [0, 2\pi R)$ . However, the AdS metrics (4.22) or (4.23) have no such restriction, with  $t \in (-\infty, +\infty)$ . They are the universal covering of the hyperboloid (4.26).

There is another parameterisation of the hyperboloid that is also useful. It takes the rather convoluted form

$$X^i = \frac{r}{R}x^i \text{ for } i = 0, 1, 2 \text{ , } X^4 - X^3 = r \text{ , } X^4 + X^3 = \frac{R^2}{r} + \frac{r}{R^2}\eta_{ij}x^i x^j$$

with  $r \in [0, \infty)$ . Although the change of coordinates is tricky, the metric is very straightforward, taking the form

$$ds^2 = R^2 \frac{dr^2}{r^2} + \frac{r^2}{R^2}\eta_{ij}dx^i dx^j \quad (4.28)$$

These coordinates don't cover the whole of AdS; instead they cover only one-half of the hyperboloid, restricted to  $X^4 - X^3 > 0$ . This is known as the *Poincaré patch* of AdS. Moreover, the time coordinate, which already extends over the full range  $x^0 \in (-\infty, +\infty)$ , cannot be further extended. This means that as  $x^0$  goes from  $-\infty$  to  $+\infty$  in (4.28), in global coordinates (4.22), the time coordinate  $t$  goes only from 0 to  $2\pi R$ .

Two other choices of coordinates are also commonly used to describe the Poincaré patch. If we set  $z = R^2/r$ , then we have

$$ds^2 = \frac{R^2}{z^2} (dz^2 + \eta_{ij}dx^i dx^j)$$

Alternatively, if we set  $r = Re^\rho$ , we have

$$ds^2 = R^2 d\rho^2 + e^{2\rho}\eta_{ij}dx^i dx^j$$

In each case, massive particles fall towards  $r = 0$ , or  $z = \infty$ , or  $\rho = -\infty$ .

### 4.3 Symmetries

We introduced the three spacetimes – Minkowski, de Sitter and anti-de Sitter – as simple examples of solutions to the Einstein equations. In fact, what makes them special are their symmetries.

The symmetries of Minkowski space are very familiar: they consist of translations and rotations in space and time, the latter splitting into genuine rotations and Lorentz boosts. It's hard to overstate the importance of these symmetries: on a fixed Minkowski background they are responsible for the existence of energy, momentum and angular momentum.

The purpose of this section is to find a way to characterise the symmetries of a general metric.

### 4.3.1 Isometries

Intuitively, the notion of symmetry is clear. If you hold up a round sphere, it looks the same no matter what way you rotate it. In contrast, if the sphere has dimples and bumps, then the rotational symmetry is broken. The distinction between these two should be captured by the metric. Roughly speaking, the metric on a round sphere looks the same at all points, while the metric on a dimpled sphere will depend on where you sit. We want a way to state this mathematically.

To do this, we need the concept of a *flow* that we introduced in Section 2.2.3. Recall that a flow on a manifold  $M$  is a one-parameter family of diffeomorphisms  $\sigma_t : M \rightarrow M$ . A flow can be identified with a vector field  $K \in \mathfrak{X}(M)$  which, at each point in  $M$ , points along tangent vectors to the flow

$$K^\mu = \frac{dx^\mu}{dt}$$

This flow is said to be an *isometry*, if the metric looks the same at each point along a given flow line. Mathematically, this means that an isometry satisfies

$$\mathcal{L}_K g = 0 \quad \Leftrightarrow \quad \nabla_\mu K_\nu + \nabla_\nu K_\mu = 0 \quad (4.29)$$

where the equivalence of the two expressions was shown in Section 4.1.3. This is the *Killing equation* and any  $K$  satisfying this equation is known as a *Killing vector field*.

Sometimes it is possible to stare at a metric and immediately write down a Killing vector. Suppose that the metric components  $g_{\mu\nu}(x)$  do not depend on one particular coordinate, say  $y \equiv x^1$ . Then the vector field  $X = \partial/\partial y$  is a Killing vector, since

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

However, we have met coordinates like  $y$  before: they become the ignorable coordinates in the Lagrangian for a particle moving in the metric  $g_{\mu\nu}$ , resulting in conserved quantities. We once again see the familiar link between symmetries and conserved quantities. We'll explore this more in Section 4.3.2 and again later in the lectures.

There is a group structure underlying these symmetries. Or, more precisely, a Lie algebra structure. This follows from the result (2.15)

$$\mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X = \mathcal{L}_{[X,Y]}$$

(Strictly speaking, we showed this in (2.15) only for Lie derivatives acting on vector fields, but it can be checked that it holds on arbitrary tensor fields.) This means that Killing vectors too form a Lie algebra. This is the Lie algebra of the isometry group of the manifold.

### An Example: Minkowski Space

As a particularly simple example, consider Minkowski spacetime with  $g_{\mu\nu} = \eta_{\mu\nu}$ . The Killing equation is

$$\partial_\mu K_\nu + \partial_\nu K_\mu = 0$$

There are two forms of solutions. We can take

$$K_\mu = c_\mu$$

for any constant vector  $c_\mu$ . These generate translations in Minkowski space. Alternatively, we can take

$$K_\mu = \omega_{\mu\nu} x^\nu$$

with  $\omega_{\mu\nu} = -\omega_{\nu\mu}$ . These generate rotations and Lorentz boosts in Minkowski space.

We can see the emergence of the algebra structure more clearly. We define Killing vectors

$$P_\mu = \partial_\mu \quad \text{and} \quad M_{\mu\nu} = \eta_{\mu\rho} x^\rho \partial_\nu - \eta_{\nu\rho} x^\rho \partial_\mu \quad (4.30)$$

There are 10 such Killing vectors in total; 4 from translations and six from rotations and boosts. A short calculation shows that they obey

$$\begin{aligned} [P_\mu, P_\nu] &= 0 \quad , \quad [M_{\mu\nu}, P_\sigma] = -\eta_{\mu\sigma} P_\nu + \eta_{\sigma\nu} P_\mu \\ [M_{\mu\nu}, M_{\rho\sigma}] &= \eta_{\mu\sigma} M_{\nu\rho} + \eta_{\nu\rho} M_{\mu\sigma} - \eta_{\mu\rho} M_{\nu\sigma} - \eta_{\nu\sigma} M_{\mu\rho} \end{aligned}$$

which we recognise as the Lie algebra of the Poincaré group  $\mathbf{R}^4 \times SO(1,3)$ .

## More Examples: de Sitter and anti-de Sitter

The isometries of de Sitter and anti-de Sitter are simplest to see from their embeddings. The constraint (4.17) that defines de Sitter space is invariant under the rotations of  $\mathbf{R}^{1,4}$ , and so de Sitter inherits the  $SO(1,4)$  isometry group. Similarly, the constraint (4.26) that defines anti-de Sitter is invariant under the rotations of  $\mathbf{R}^{2,3}$ . Correspondingly, AdS has the isometry group  $SO(2,3)$ . Note that both of these groups are 10 dimensional: in terms of counting, these spaces are just as symmetric as Minkowski space.

It is simple to write down the 10 Killing spinors in the parent 5d spacetime: they are

$$M_{AB} = \eta_{AC} X^C \partial_B - \eta_{BC} X^C \partial_A$$

where  $X^A$ ,  $A = 0, 1, 2, 3, 4$  are coordinates in 5d and  $\eta_{AB}$  is the appropriate Minkowski metric, with signature  $(- + + + +)$  for de Sitter and  $(- - + + +)$  for anti-de Sitter. In either case, the Lie algebra is that of the appropriate Lorentz group,

$$[M_{AB}, M_{CD}] = \eta_{AD} M_{BC} + \eta_{BC} M_{AD} - \eta_{AC} M_{BD} - \eta_{BD} M_{AC}$$

Importantly, the embedding hyperbolae (4.17) and (4.26) are both invariant under these Killing vectors, in the sense that the flows generated by  $M_{AB}$  take us from one point on the hyperbolae to another. This means that the Killing vectors are inherited by de Sitter and anti-de Sitter spaces respectively.

For example, we can consider de Sitter space in the static patch with  $r^2 = (X^1)^2 + (X^2)^2 + (X^3)^2$  and (4.19)

$$X^0 = \sqrt{R^2 - r^2} \sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2} \cosh(t/R)$$

We know that the metric in the static patch (4.12) is independent of time. This means that  $K = \partial_t$  is a Killing vector. Pushed forwards to the 5d space, this becomes

$$\frac{\partial}{\partial t} = \frac{\partial X^A}{\partial t} \frac{\partial}{\partial X^A} = \frac{1}{R} \left( X^4 \frac{\partial}{\partial X^0} + X^0 \frac{\partial}{\partial X^4} \right) \quad (4.31)$$

In fact, this Killing vector highlights a rather important subtlety with de Sitter space. As we go on, we will see that timelike Killing vectors – those obeying  $g_{\mu\nu} K^\mu K^\nu < 0$  everywhere – play a rather special role because we can use them to define energy. (We'll describe this for particles in the next section.)

In anti-de Sitter space, there is no problem in finding a timelike Killing vector. Indeed, we can see it by eye in the global coordinates (4.22), where it is simply  $K = \partial_t$ . But de Sitter is another story.

If we work in the static patch (4.12), then the Killing vector (4.31) is a timelike Killing vector. Indeed, we used this to derive the conserved energy  $E$  when discussing geodesics in de Sitter in Section 4.2.1. But we know that the static patch does not cover all of de Sitter spacetime.

Indeed, if we extend the Killing vector (4.31) over the entire space, it is not timelike everywhere. To see this, note that when  $X^4 > 0$  and  $X^0 = 0$ , the Killing vector pushes us forwards in the  $X^0$  direction, but when  $X^4 < 0$  and  $X^0 = 0$  it pushes us backwards in the  $X^0$  direction. This means that the Killing vector field points to the future in some parts of space and to the past in others! Correspondingly, if we try to define an energy using this Killing vector it will be positive in some parts of the space and negative in others. Relatedly, in parts of the space where  $X^4 = 0$  and  $X^0 \neq 0$ , the Killing vector pushes us in the  $X^4$  direction, and so is spacelike.

The upshot of this discussion is an important feature of de Sitter space: there is no global, positive conserved energy. This tallies with our metric in global coordinates (4.21) which is time dependent and so does not obviously have a timelike Killing vector. The lack of a globally defined energy is one of several puzzling aspects of de Sitter space: we'll meet more as we proceed.

### 4.3.2 A First Look at Conserved Quantities

Emmy Noether taught us that symmetries are closely related to conserved quantities. In the present context, this means that any dynamics taking place in a spacetime with an isometry will have a conserved quantity.

There are a number of different scenarios in which we can ask about conserved quantities. We could look at particles moving in a fixed background; these are the kinds of calculations that we did in Section 1. Alternatively, we could ask about fields moving in a fixed background; we will address this in Section 4.5.5. Finally, we could ask about the energy stored in the spacetime itself. We will provide a formula for this in Section 4.3.3, and also make some further comments in Section 4.5.5.

Here, we consider a massive particle moving in a spacetime with metric  $g$ . The particle will follow some trajectory  $x^\mu(\tau)$ , with  $\tau$  the proper time. If the spacetime admits a Killing vector  $K$ , then we can construct the quantity that is conserved along the geodesic,

$$Q = K_\mu \frac{dx^\mu}{d\tau} \tag{4.32}$$



To see that  $Q$  is indeed unchanging, we compute

$$\begin{aligned}
\frac{dQ}{d\tau} &= \partial_\nu K_\mu \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} + K_\mu \frac{d^2 x^\mu}{d\tau^2} \\
&= \partial_\nu K_\mu \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} - K_\mu \Gamma_{\rho\sigma}^\mu \frac{dx^\rho}{d\tau} \frac{dx^\sigma}{d\tau} \\
&= \nabla_\nu K_\mu \frac{dx^\nu}{d\tau} \frac{dx^\mu}{d\tau} = 0
\end{aligned}$$

where in the second line we've used the geodesic equation and, in the final equality, we've used the symmetry of the Killing equation.

The derivation above looks rather different from our usual formulation of Noether's theorem. For this reason, it's useful to re-derive the Killing equation and corresponding conserved charge by playing the usual Noether games. We can do this by looking at the action for a massive particle (in the form (1.32))

$$S = \int d\tau g_{\mu\nu}(x) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}$$

Now we can play the usual Noether games. Consider the infinitesimal transformation

$$\delta x^\mu = K^\mu(x)$$

The action transforms as

$$\begin{aligned}
\delta S &= \int d\tau \partial_\rho g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} K^\rho + 2g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dK^\nu}{d\tau} \\
&= \int d\tau \partial_\rho g_{\mu\nu} \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} K^\rho + 2 \frac{dx^\mu}{d\tau} \left( \frac{dK_\mu}{d\tau} - K^\nu \frac{dg_{\mu\nu}}{d\tau} \right) \\
&= \int d\tau (\partial_\rho g_{\mu\nu} K^\rho - 2K^\rho \partial_\nu g_{\mu\rho} + 2\partial_\nu K_\mu) \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau} \\
&= \int d\tau 2\nabla_\nu K_\mu \frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}
\end{aligned}$$

The transformation is a symmetry of the action if  $\delta S = 0$ . From the symmetry of the  $\frac{dx^\mu}{d\tau} \frac{dx^\nu}{d\tau}$  terms, this is true provided that  $K_\mu$  obeys the Killing equation

$$\nabla_{(\nu} K_{\mu)} = 0$$

Noether's theorem then identifies the charge  $Q$  defined in (4.32) as the conserved quantity arising from this symmetry.

We met examples of these conserved quantities in Section 4.2 when discussing geodesics in de Sitter and anti-de Sitter spacetimes. (And, indeed, in Section 1.3 when discussing the geodesic orbits around a black hole). Both the energy  $E$  and the angular momentum  $l$  are Noether charges of this form.

Killing vectors have further roles to play in identifying conserved quantities. In Section 4.5.5, we'll describe how we can use Killing vectors to define energy and momentum of fields in a background spacetime.

### 4.3.3 Komar Integrals

If we have a Killing vector, there is a rather pretty way of associating a conserved quantity to the spacetime itself.

Given a Killing vector  $K = K^\mu \partial_\mu$ , we can construct the 1-form  $K = K_\mu dx^\mu$ . From this 1-form, we can then construct a 2-form

$$F = dK$$

Alternatively, in components, we have  $F = \frac{1}{2} F_{\mu\nu} dx^\mu \wedge dx^\nu$  with

$$F_{\mu\nu} = \nabla_\mu K_\nu - \nabla_\nu K_\mu$$

We've called this 2-form  $F$ , in analogy with the electromagnetic 2-form. Indeed, the key idea of the Komar integral is that we can think of  $F$  very much like the electromagnetic field strength. Indeed, we claim the following is true:

**Claim:** If the vacuum Einstein equations are obeyed, so  $R_{\mu\nu} = 0$ , then  $F$  obeys the vacuum Maxwell equations

$$d \star F = 0$$

Alternatively, as shown in (3.24), we can write this as

$$\nabla^\mu F_{\mu\nu} = 0$$

**Proof:** To see this, we start with the Ricci identity (3.16) which, applied to the Killing vector  $K^\sigma$ , reads

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) K^\sigma = R^\sigma_{\rho\mu\nu} K^\rho$$

Contracting the  $\mu$  and  $\sigma$  indices then gives

$$(\nabla_\mu \nabla_\nu - \nabla_\nu \nabla_\mu) K^\mu = R_{\rho\nu} K^\rho$$

But  $K^\mu$  is a Killing vector and so obeys the Killing equation  $\nabla_{(\mu}K_{\nu)} = 0$  and so  $\nabla_\mu K^\mu = 0$ . This means that the Ricci identity simplifies to

$$\nabla_\mu \nabla_\nu K^\mu = R_{\rho\nu} K^\rho \quad (4.33)$$

With this in hand, we now look at  $\nabla^\mu F_{\mu\nu}$ . We have

$$\nabla^\mu F_{\mu\nu} = \nabla^\mu \nabla_\mu K_\nu - \nabla^\mu \nabla_\nu K_\mu = -2\nabla^\mu \nabla_\nu K_\mu = -2R_{\rho\nu} K^\rho$$

where we've used the Killing equation in the second equality and the Ricci identity (4.33) in the third. This then gives the promised result:  $d \star F = 0$  provided that the Einstein equations  $R_{\rho\nu} = 0$  hold.  $\square$

Since the 2-form  $F$  obeys the vacuum Maxwell equations, we can use it to construct the *Komar charge*, or *Komar integral*. We integrate over some three-dimensional spatial submanifold  $\Sigma$ ,

$$Q_{\text{Komar}} = -\frac{1}{8\pi G} \int_\Sigma d \star F = -\frac{1}{8\pi G} \int_{\partial\Sigma} \star F = -\frac{1}{8\pi G} \int_{\partial\Sigma} \star dK$$

Here the factor of  $1/8\pi G$  is for later convenience. Because  $d \star F = 0$ , the same kind of argument that we met in Section 3.2.5 then tells us that  $Q_{\text{Komar}}$  is conserved.

Just as for the point particle discussed previously, the interpretation of the Komar integrals depends on the Killing vector at hand. For example, if  $K^\mu$  is everywhere timelike, meaning  $g_{\mu\nu}K^\mu K^\nu < 0$  at all points, then the Komar integral can be identified with the energy, or equivalently, the mass of the spacetime

$$M_{\text{Komar}} = -\frac{1}{8\pi G} \int_{\partial\Sigma} \star dK$$

In particular, we'll later see how the Komar charge can be non-vanishing, even though  $d \star F = 0$  when the vacuum Einstein equations are obeyed. Relatedly, if the Killing vector is related to rotations, the conserved charge is identified with the angular momentum of the spacetime.

At this point, it would obviously be nice to give some examples of Komar integrals. Sadly, we don't yet have any useful examples at our disposal! However, we will use this technology throughout Section 6 to identify the mass and angular momentum of black holes.

As an aside: later in Section 4.5, we will look at what happens if we couple matter to gravity. There we will learn that the Einstein equations are no longer  $R_{\mu\nu} = 0$ , but instead the right-hand side gets altered by the energy and momentum of the matter. In this case, we can again form the field strength  $F$ , but now it obeys the Maxwell equation with a source,  $d \star F = \star J$ , where the current  $J$  can be related to the energy-momentum tensor. However, it turns out that for our applications in Section 6 the case of the vacuum Einstein equations  $R_{\mu\nu} = 0$  is all we'll need.

#### 4.4 Asymptotics of Spacetime

The three solutions – Minkowski, de Sitter, and anti-de Sitter – have different spacetime curvature and differ in their symmetries. But there is a more fundamental distinction: they have different behaviour at infinity.

This is important because we will ultimately want to look at more complicated solutions. These may have reduced symmetries, or no symmetries at all. But, providing fields are suitably localised, they will asymptote to one of the three symmetric spaces described above. This gives us a way to characterise whether physics is happening “in Minkowski spacetime”, “in de Sitter”, or “in anti-de Sitter”.

It turns out that “infinity” in Lorentzian spacetimes is more interesting than you might have thought. One can go to infinity along spatial, timelike or null directions, and each of these may have a different structure. It will be useful to introduce a tool to visualise infinity of spacetime.

##### 4.4.1 Conformal Transformations

Given a spacetime  $M$  with metric  $g_{\mu\nu}$ , we may construct a new metric  $\tilde{g}_{\mu\nu}$  by a *conformal transformation*,

$$\tilde{g}_{\mu\nu}(x) = \Omega^2(x)g_{\mu\nu}(x) \tag{4.34}$$

with  $\Omega(x)$  a smooth, non-vanishing function.

Typically  $g_{\mu\nu}$  and  $\tilde{g}_{\mu\nu}$  describe very different spacetimes, with distances in the two considerably warped. However, the conformal transformation preserves angles. In particular, in a Lorentzian spacetime, this means that two metrics related by a conformal transformation have the same causal structure. A vector field  $X$  which is everywhere-null with respect to the metric  $g_{\mu\nu}$  will also be everywhere-null with respect to  $\tilde{g}_{\mu\nu}$ ,

$$g_{\mu\nu}X^\mu X^\nu = 0 \quad \Leftrightarrow \quad \tilde{g}_{\mu\nu}X^\mu X^\nu = 0$$

Similarly, vectors that are timelike/spacelike with respect to  $g_{\mu\nu}$  will continue to be timelike/spacelike separated with respect to  $\tilde{g}_{\mu\nu}$ .

A conformal transformation of the metric does not change the causal structure. However, any other change of the metric does. This fact is sometimes summarised in the slogan “the causal structure is 9/10<sup>th</sup> of the metric”. Although, taking into account diffeomorphism invariance, a better slogan would be “the causal structure is around 5/6<sup>th</sup> of the metric”.

### Conformal Transformations and Geodesics

A particle trajectory which is timelike with respect to  $g_{\mu\nu}$  will necessarily also be timelike with respect to  $\tilde{g}_{\mu\nu}$ . But because distances get screwed up under a conformal transformation, there is no reason to expect that a timelike geodesic will map to a timelike geodesic. However, it turns out that null geodesics do map to null geodesics, although the affine parameterisation gets messed up along the way.

To see this, we first compute the Christoffel symbols in the new metric. They are

$$\begin{aligned}\Gamma_{\rho\sigma}^{\mu}[\tilde{g}] &= \frac{1}{2}\tilde{g}^{\mu\nu}(\partial_{\rho}\tilde{g}_{\nu\sigma} + \partial_{\sigma}\tilde{g}_{\rho\nu} - \partial_{\nu}\tilde{g}_{\rho\sigma}) \\ &= \frac{1}{2}\Omega^{-2}g^{\mu\nu}(\partial_{\rho}(\Omega^2g_{\nu\sigma}) + \partial_{\sigma}(\Omega^2g_{\rho\nu}) - \partial_{\nu}(\Omega^2g_{\rho\sigma})) \\ &= \Gamma_{\rho\sigma}^{\mu}[g] + \Omega^{-1}(\delta_{\sigma}^{\mu}\nabla_{\rho}\Omega + \delta_{\rho}^{\mu}\nabla_{\sigma}\Omega - g_{\rho\sigma}\nabla^{\mu}\Omega)\end{aligned}$$

where, in the final line, we’ve replaced  $\partial$  with  $\nabla$  on the grounds that the derivatives are hitting a scalar function  $\Omega(x)$  so it makes no difference.

If we have an affinely parameterised geodesic in the metric  $g$

$$\frac{d^2x^{\mu}}{d\tau^2} + \Gamma_{\rho\sigma}^{\mu}[g]\frac{dx^{\rho}}{d\tau}\frac{dx^{\sigma}}{d\tau} = 0$$

then in the metric  $\tilde{g}$  we have

$$\frac{d^2x^{\mu}}{d\tau^2} + \Gamma_{\rho\sigma}^{\mu}[\tilde{g}]\frac{dx^{\rho}}{d\tau}\frac{dx^{\sigma}}{d\tau} = \Omega^{-1}(\delta_{\sigma}^{\mu}\nabla_{\rho}\Omega + \delta_{\rho}^{\mu}\nabla_{\sigma}\Omega - g_{\rho\sigma}\nabla^{\nu}\Omega)\frac{dx^{\rho}}{d\tau}\frac{dx^{\sigma}}{d\tau}$$

The right-hand side looks like a mess. And for timelike or spacelike geodesics, it is. But for null geodesics we have

$$g_{\rho\sigma}\frac{dx^{\rho}}{d\tau}\frac{dx^{\sigma}}{d\tau} = 0$$

so at least one term on the right-hand side vanishes. The others can be written as

$$\frac{d^2x^{\mu}}{d\tau^2} + \Gamma_{\rho\sigma}^{\mu}[\tilde{g}]\frac{dx^{\rho}}{d\tau}\frac{dx^{\sigma}}{d\tau} = 2\frac{dx^{\mu}}{d\tau}\frac{1}{\Omega}\frac{d\Omega}{d\tau}$$

But this is the equation for a geodesic that is not affinely parameterised, as in (1.28). So a conformal transformation does map null geodesics to null geodesics as claimed.

## The Weyl Tensor

Our favourite curvature tensors are not invariant under conformal transformations. However, it turns out that there is a combination of curvature tensors that does not change under conformal transformations. This is the *Weyl tensor*. In a manifold of dimension  $n$ , it is defined as

$$C_{\mu\nu\rho\sigma} = R_{\mu\nu\rho\sigma} - \frac{2}{n-2} (g_{\mu[\rho}R_{\sigma]\nu} - g_{\nu[\rho}R_{\sigma]\mu}) + \frac{2}{(n-1)(n-2)} R g_{\mu[\rho}g_{\sigma]\nu}$$

The Weyl tensor has all the symmetry properties of the Riemann tensor, but with the additional property that if you contract any pair of indices with a metric then it vanishes. In this sense, it can be viewed as the “trace-free” part of the Riemann tensor.

### 4.4.2 Penrose Diagrams

There are a number of interesting and deep stories associated to conformal transformations (4.34). For example, there are a class of theories that are invariant under conformal transformations of Minkowski space; these so-called conformal field theories describe physics at a second order phase transition. But here we want to use conformal transformations to understand what happens at infinity of spacetime.

The main idea is to perform a conformal transformation that pulls infinity to some more manageable, finite distance. Obviously this transformation will mangle distances, but it will retain the causal structure of the original spacetime. We can then draw this causal structure on a very finite piece of paper (e.g. A4). The resulting picture is called a *Penrose diagram*, named after its discoverers, Roger Penrose and Brandon Carter. We will illustrate this with a series of examples.

## Minkowski Space

We start with Minkowski space. It turns out that, even here, infinity is rather subtle.

It will be simplest if we first work in  $d = 1 + 1$  dimensions, where the Minkowski metric takes the form

$$ds^2 = -dt^2 + dx^2 \tag{4.35}$$

The first thing we do is introduce light-cone coordinates,

$$u = t - x \quad \text{and} \quad v = t + x$$

In these coordinates, the Minkowski metric is even simpler

$$ds^2 = -du dv$$

Both of these light-cone coordinates take values over the full range of  $\mathbf{R}$ :  $u, v \in (-\infty, \infty)$ . In an attempt to make things more finite, we will introduce another coordinate that traverses the full range of  $u$  and  $v$  over a finite interval. A convenient choice is

$$u = \tan \tilde{u} \quad \text{and} \quad v = \tan \tilde{v} \quad (4.36)$$

where we now cover the whole of Minkowski space as  $\tilde{u}, \tilde{v} \in (-\pi/2, +\pi/2)$ . Note that, strictly speaking, we shouldn't include the points  $\tilde{u}, \tilde{v} = \pm\pi/2$  since these correspond to  $u, v = \pm\infty$ .

In the new coordinates, the metric takes the form

$$ds^2 = -\frac{1}{\cos^2 \tilde{u} \cos^2 \tilde{v}} d\tilde{u} d\tilde{v}$$

Notice that the metric diverges as we approach the boundary of Minkowski space, where  $\tilde{u}$  or  $\tilde{v} \rightarrow \pm\pi/2$ . However, we can now do our conformal transformation. We define the new metric

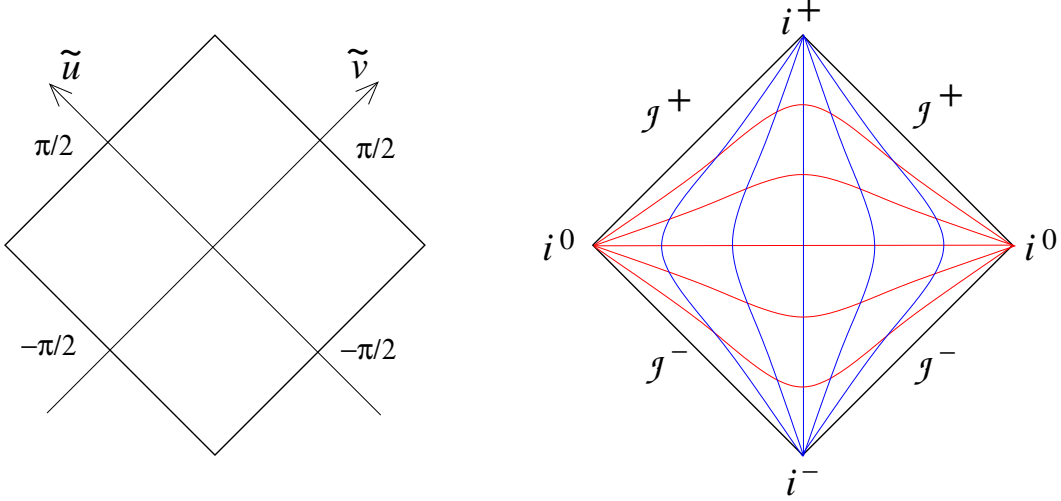
$$d\tilde{s}^2 = (\cos^2 \tilde{u} \cos^2 \tilde{v}) ds^2 = -d\tilde{u} d\tilde{v}$$

After the conformal map, nothing bad happens as we approach  $\tilde{u}, \tilde{v} \rightarrow \pm\pi/2$ . It is customary to now add in the “points at infinity”,  $\tilde{u} = \pm\pi/2$  and  $\tilde{v} = \pm\pi/2$ , an operation that goes by the name of *conformal compactification*.

The Penrose diagram is a pictorial representation of this space. As in other relativistic diagrams, we insist that light-rays go at  $45^\circ$ . We take time to be in the vertical direction, and space in the horizontal. This means that we draw the lightcone  $\tilde{u}$  and  $\tilde{v}$  coordinates at  $45^\circ$ . The resulting diagram is shown with the  $\tilde{u}$  and  $\tilde{v}$  axes on the left-hand side of Figure 33.

We can also dress our Penrose diagram in various ways. For example, we could draw geodesics with respect to the original metric (4.35). These are shown in the right-hand side of Figure 33; the verticalish blue lines are timelike geodesics of constant  $x$ ; the horizontalish red lines are spacelike geodesics of constant  $t$ . We have also listed the different kinds of “infinity” for Minkowski space. They are

- All timelike geodesics start at the point labelled  $i^-$ , with  $(\tilde{u}, \tilde{v}) = (-\pi/2, -\pi/2)$  and end at the point labelled  $i^+$  with  $(\tilde{u}, \tilde{v}) = (+\pi/2, +\pi/2)$ . In other words, this is the origin and fate of all massive particles. These points are referred to as *past* and *future timelike infinity* respectively.



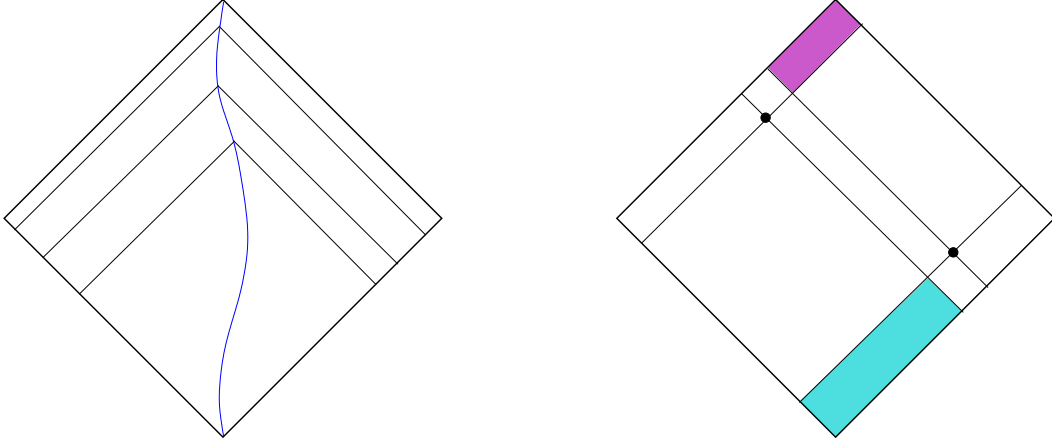
**Figure 33:** The Penrose diagram for  $d = 1 + 1$  Minkowski space.

- All spacelike geodesics begin or end at one of the two points labelled  $i^0$ , either  $(\tilde{u}, \tilde{v}) = (-\pi/2, +\pi/2)$  or  $(\tilde{u}, \tilde{v}) = (+\pi/2, -\pi/2)$ . These points are *spacelike infinity*.
- All null curves start on the boundary labelled  $\mathcal{I}^-$ , with  $\tilde{u} = -\pi/2$  and arbitrary  $\tilde{v}$ , or  $\tilde{v} = -\pi/2$  and arbitrary  $\tilde{u}$ . This boundary is pronounced “scri-minus” and known, more formally, as *past null infinity*. Such null curves end on the boundary labelled  $\mathcal{I}^+$ , with  $\tilde{u} = +\pi/2$  and arbitrary  $\tilde{v}$ , or  $\tilde{v} = +\pi/2$  and arbitrary  $\tilde{u}$ . This is pronounced “scri-plus” and known as *future null infinity*.

We see from the picture that there are more ways to “go to infinity” in a null direction than in a timelike or spacelike direction. This is one of the characteristic features of Minkowski space.

The Penrose diagram allows us to immediately visualise the causal structure of Minkowski space. For example, as timelike curves approach  $i^+$ , their past lightcone encompasses more and more of the spacetime, as shown in the left-hand side of Figure 34. This means that an observer in Minkowski space can see everything (in principle) as long as they wait long enough. Relatedly, given any two points in Minkowski space, they are causally connected in both the past and future, meaning that their past and future lightcones necessarily intersect, as shown in the Mondrian painting on the right-hand side of Figure 34. This means that there was always an event in the past that could influence both points, and always an event in the future that can be influenced





**Figure 34:** On the left: you will eventually see everything. On the right: every two points share some of their causal past and some of their causal future.

by both. These comments may seem trivial, but we will soon see that they don't hold in other spacetimes, including the one we call home.

Let's now repeat the analysis for Minkowski space in  $d = 3 + 1$  dimensions, with the metric

$$ds^2 = -dt^2 + dr^2 + r^2 d\Omega_2^2$$

where  $d\Omega_2^2 = d\theta^2 + \sin^2 \theta d\phi^2$  is the round metric on  $\mathbf{S}^2$  (and is not to be confused with the conformal factor  $\Omega(x)$  that we introduced earlier). Again we introduce lightcone coordinates

$$u = t - r \quad \text{and} \quad v = t + r$$

and write this metric as

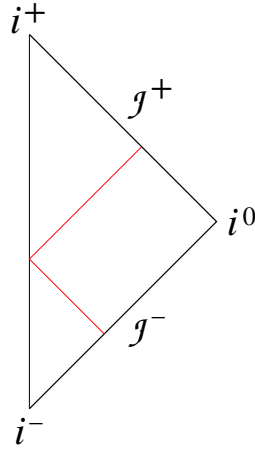
$$ds^2 = -du dv + \frac{1}{4}(u - v)^2 d\Omega_2^2$$

In the finite-range coordinates (4.36), the metric becomes

$$ds^2 = \frac{1}{4 \cos^2 \tilde{u} \cos^2 \tilde{v}} (-4d\tilde{u} d\tilde{v} + \sin^2(\tilde{u} - \tilde{v})d\Omega_2^2)$$

Finally, we do the conformal transformation to the new metric

$$d\tilde{s}^2 = -4d\tilde{u} d\tilde{v} + \sin^2(\tilde{u} - \tilde{v})d\Omega_2^2 \quad (4.37)$$



**Figure 35:** The Penrose diagram for 4d Minkowski spacetime.

There is an additional requirement that didn't arise in 2d: we must insist that  $v \geq u$  so that  $r \geq 0$ , as befits a radial coordinate. This means that, after a conformal compactification,  $\tilde{u}$  and  $\tilde{v}$  take values in

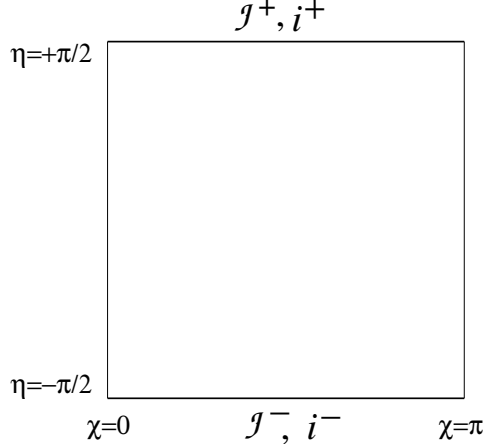
$$-\frac{\pi}{2} \leq \tilde{u} \leq \tilde{v} \leq \frac{\pi}{2}$$

To draw a diagram corresponding to the spacetime (4.37), we're going to have to ditch some dimensions. We chose not to depict the  $\mathbf{S}^2$ , and only show the  $\tilde{u}$  and  $\tilde{v}$  directions. The resulting Penrose diagram for  $d = 3 + 1$  dimensional Minkowski space is shown in Figure 35.

Every point in the diagram corresponds to an  $\mathbf{S}^2$  of radius  $\sin(\tilde{u} - \tilde{v})$ , except for the left-hand line which sits at  $\tilde{u} = \tilde{v}$  where this  $\mathbf{S}^2$  shrinks to a point. This is not a boundary of Minkowski space; it is simply the origin  $r = 0$ . To illustrate this, we've drawn a null geodesic in red in the figure; it starts at  $\mathcal{I}^-$  and when it hits the  $r = 0$  vertical line, it simply bounces off and ends up at  $\mathcal{I}^+$ .

The need to draw a 4d space on a 2d piece of paper is something of a limitation of Penrose diagrams. It means that they're only really useful for spacetimes that have an obvious  $\mathbf{S}^2$  sitting inside them that we can drop. Or, to state it more precisely, spacetimes that have an  $SO(3)$  isometry. But these spacetimes are the simplest and tend to be the most important.

We have seen that Minkowski space has a null boundary, together with a couple of points at spatial and temporal infinity. This naturally lends itself to asking questions



**Figure 36:** The Penrose diagram for de Sitter.

about scattering of massless fields: we set up some initial data on  $\mathcal{I}^-$ , let it evolve, and read off its fate at  $\mathcal{I}^+$ . In quantum field theory, this is closely related to the object we call the S-matrix.

### de Sitter Space

The global coordinates for de Sitter space are (4.21),

$$ds^2 = -d\tau^2 + R^2 \cosh^2(\tau/R) d\Omega_3^2$$

To construct the Penrose diagram we work with *conformal time*, defined by

$$\frac{d\eta}{d\tau} = \frac{1}{R \cosh(\tau/R)}$$

The solution is

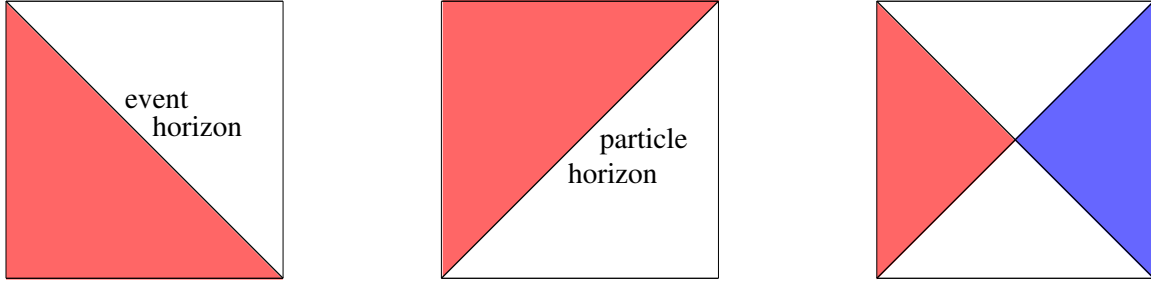
$$\cos \eta = \frac{1}{\cosh(\tau/R)} \tag{4.38}$$

with  $\eta \in (-\pi/2, +\pi/2)$  as  $\tau \in (-\infty, +\infty)$ . In conformal time, de Sitter space has the metric

$$ds^2 = \frac{R^2}{\cos^2 \eta} (-d\eta^2 + d\Omega_3^2)$$

We write the metric on the  $\mathbf{S}^3$  as

$$d\Omega_3^2 = d\chi^2 + \sin^2 \chi d\Omega_2^2 \tag{4.39}$$



**Figure 37:** On the left: an observer at the north pole does not see everything. She has an event horizon. In the middle: Nor can she influence everything: she has a particle horizon. On the right: the causal diamond for an observer at the north pole (in red) and at the south pole (in blue).

with  $\chi \in [0, \pi]$ . The de Sitter metric is conformally equivalent to

$$d\tilde{s}^2 = -d\eta^2 + d\chi^2 + \sin^2 \chi d\Omega_2^2$$

After a conformal compactification,  $\eta \in [-\pi/2, +\pi/2]$  and  $\chi \in [0, \pi]$ . The Penrose diagram is shown in Figure 36.

The two vertical lines are not boundaries of the spacetime; they are simply the north and south poles of the  $\mathbf{S}^3$ . The boundaries are the horizontal lines at the top and bottom: they are labelled both as  $i^\pm$  and  $\mathcal{I}^\pm$ , reflecting the fact that they are where both timelike and null geodesics originate and terminate.

We learn that de Sitter spacetime has a *spacelike*  $\mathbf{S}^3$  boundary. (The normal to this boundary is timelike.)

The causal structure of de Sitter spacetime is very different from Minkowski. It is not true that if an observer waits long enough then she will be able to see everything that's happening. For example, an observer who sits at the north pole (the left-hand side of the figure) will ultimately be able to see exactly half the spacetime, as shown in the left-hand side of Figure 37. The boundary of this space (the diagonal line in the figure) is her *event horizon*. It is similar to the event horizon of a black hole in the sense that signals from beyond the horizon cannot reach her. However, as is clear from the picture, it is an observer-dependent horizon: someone else will have an entirely different event horizon. In this context, these are sometimes referred to as *cosmological horizons*.

Furthermore, the observer at the north pole will only be able to communicate with another half of the spacetime, as shown in the middle of Figure 37. The boundary of the region of influence is known as the *particle horizon*. You should think of it as the furthest distance light can travel since the beginning of time. The intersection of these two regimes is called the (northern) *causal diamond* and is shown as the red triangle in the right-hand figure. An observer sitting at the southern pole also has a causal diamond, shown in blue in the right-hand side of Figure 37. It is causally disconnected from the northern diamond.

This state of affairs was nicely summarised by Schrödinger who, in 1956, wrote

“It does seem rather odd that two or more observers, even such as sat on the same school bench in the remote past, should in future, when they have followed different paths in life, experience different worlds, so that eventually certain parts of the experienced world of one of them should remain by principle inaccessible to the other and vice versa.”

In asymptotically de Sitter spacetimes, it would appear that the natural questions involve setting some initial conditions on spacelike  $\mathcal{I}^-$ , letting it evolve, and reading off the data on  $\mathcal{I}^+$ . One of the lessons of the development of quantum mechanics is that we shouldn't talk about things that cannot, even in principle, be measured. Yet in de Sitter space we see that no single observer has an overview of the whole space. This causes a number of headaches and, as yet, unresolved conceptual issues when we try to discuss quantum gravity in de Sitter space.

Finally, we can use the Penrose diagram to answer a lingering puzzle about the static patch of de Sitter, in which the metric takes the form (4.12)

$$ds^2 = - \left(1 - \frac{r^2}{R^2}\right) dt^2 + \left(1 - \frac{r^2}{R^2}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (4.40)$$

The question is: how should we interpret the divergence at  $r = R$ ?

To answer this, we will look at where the surface  $r = R$  sits in the Penrose diagram. First, we look at the embedding of the static patch in  $\mathbf{R}^{1,4}$ , given in (4.19)

$$X^0 = \sqrt{R^2 - r^2} \sinh(t/R) \quad \text{and} \quad X^4 = \sqrt{R^2 - r^2} \cosh(t/R)$$

Naively the surface  $r = R$  corresponds to  $X^0 = X^4 = 0$ . But that's a little too quick. To see this, we consider what happens as we approach  $r \rightarrow R$  by writing  $r = R(1 - \epsilon^2/2)$ , with  $\epsilon \ll 1$ . We then have

$$X^0 \approx R\epsilon \sinh(t/R) \quad \text{and} \quad X^4 \approx R\epsilon \cosh(t/R)$$

We can now send  $\epsilon \rightarrow 0$ , keeping  $X^0$  and  $X^4$  finite provided that we also send  $t \rightarrow \pm\infty$ . To do this, we must ensure that we keep the combination  $\epsilon e^{\pm t/R}$  finite. This means that we can identify the surface  $r = R$  with the lines  $X^0 = \pm X^4$ .

Now we translate this into global coordinates. These were given in (4.20),

$$X^0 = R \sinh(\tau/R) \quad \text{and} \quad X^4 = R \cosh(\tau/R) \cos \chi$$

where  $\chi$  is the polar angle on  $\mathbf{S}^3$  that we introduced in (4.39). After one further map to conformal time (4.38), we find that the lines  $X^0 = \pm X^4$  become

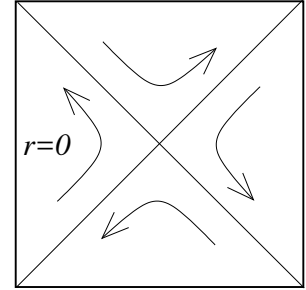
$$\sin \eta = \pm \cos \chi \quad \Rightarrow \quad \chi = \pm(\eta - \pi/2)$$

But these are precisely the diagonal lines in the Penrose diagram that appear as horizons for people living on the poles.

It's also simple to check that the point  $r = 0$  in the static patch corresponds to the north pole  $\chi = 0$  in global coordinates and, furthermore,  $t = \tau$  along this line.

The upshot is that the static patch of de Sitter (4.40) provides coordinates that cover only the northern causal diamond of de Sitter, with the coordinate singularity at  $r = R$  coinciding with the past and future observer-dependent horizons.

One advantage of the static patch coordinates is that they clearly exhibit a timelike Killing vector,  $K = \partial_t$ . This moves us from a surface of constant  $t$  to another surface of constant  $t$ . But we argued in Section 4.3.1 that there was no global timelike Killing vector field in de Sitter since, in  $\mathbf{R}^{1,4}$ , the Killing vector is given by (4.31). The Penrose diagram makes this simpler to visualise. If we extend the Killing vector beyond the static patch, it acts as shown in the figure. It is timelike and future pointing only in the northern causal diamond. It is also timelike in the southern causal diamond, but points towards the past. Meanwhile it is a spacelike Killing vector in both the upper and lower quadrants.



**Figure 38:** Killing vectors in de Sitter

## Anti-de Sitter Space

The global coordinates for anti-de Sitter space are (4.23),

$$ds^2 = -\cosh^2 \rho dt^2 + R^2 d\rho^2 + R^2 \sinh^2 \rho d\Omega_2^2$$

with  $\rho \in [0, +\infty)$ . To construct the Penrose diagram, this time we introduce a “conformal radial coordinate”  $\psi$ , defined by

$$\frac{d\psi}{d\rho} = \frac{1}{\cosh \rho}$$

This is very similar to the conformal map of de Sitter space, but with time replaced by space. The solution is

$$\cos \psi = \frac{1}{\cosh \rho}$$

One difference from the de Sitter analysis is that since  $\rho \in [0, \infty)$ , the conformal coordinate lives in  $\psi \in [0, \pi/2)$ . The metric on anti-de Sitter becomes

$$ds^2 = \frac{R^2}{\cos^2 \psi} (-d\tilde{t}^2 + d\psi^2 + \sin^2 \psi d\Omega_2^2) = \frac{R^2}{\cos^2 \psi} (-d\tilde{t}^2 + d\Omega_3^2) \quad (4.41)$$

where we’ve introduced the dimensionless time coordinates  $\tilde{t} = t/R$ . We learn that the anti-de Sitter metric is conformally equivalent to

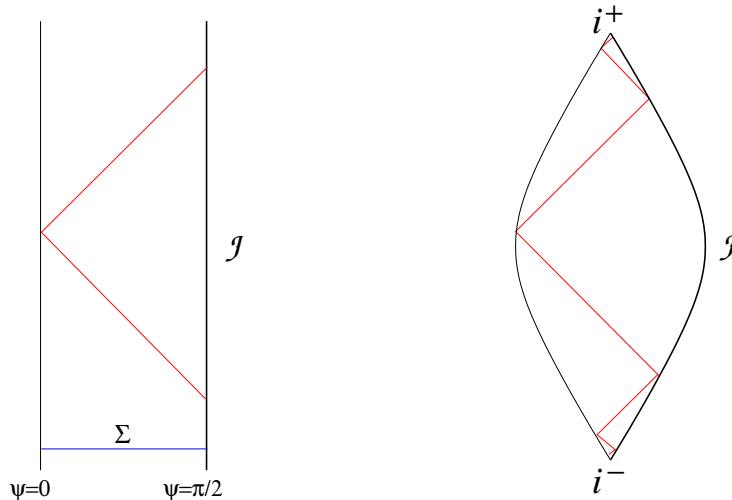
$$d\tilde{s}^2 = -d\tilde{t}^2 + d\psi^2 + \sin^2 \psi d\Omega_2^2$$

where, after a conformal compactification,  $\tilde{t} \in (-\infty, +\infty)$  and  $\psi \in [0, \pi/2]$ . The resulting Penrose diagram is shown in the left-hand of Figure 39. It is an infinite strip. The left-hand edge at  $\psi = 0$  is not a boundary: it is the spatial origin where the  $\mathbf{S}^2$  shrinks to zero size. In contrast, the right-hand edge at  $\psi = \pi/2$  is the boundary of spacetime.

The boundary is labelled  $\mathcal{I}$ . In terms of our previous notation, it should be viewed as a combination of  $\mathcal{I}^-$ ,  $\mathcal{I}^+$  and  $i^0$ , since null paths begin and end here, as do spacelike paths. The boundary is now timelike (it has spacelike normal vector), and has topology

$$\mathcal{I} = \mathbf{R} \times \mathbf{S}^2$$

with  $\mathbf{R}$  the time factor.



**Figure 39:** Penrose diagrams for AdS. On the left, we still have an infinite time coordinate; on the right this too has been conformally compactified.

The Penrose diagram allows us to immediately see that light rays hit the boundary in finite conformal time, confirming the calculation that we did in Section 4.2.2. If we want to specify physics in AdS, we need to say something about what happens at the boundary. For example, in the figure we have shown a light ray simply emerging from the boundary at one time and absorbed at some later time. Another choice would be to impose reflecting boundary conditions, so that the light ray bounces back and forth for ever. In this way, anti-de Sitter space is very much like a box, with massive particles trapped in the interior and massless particles able to bounce off the boundary.

In field theoretic language, we could start with initial data on some  $d = 3$  dimensional spacelike hypersurface  $\Sigma$  and try to evolve it in time. This is what we usually do in physics. But in AdS, this information is not sufficient. This is because we can find points to the future of  $\Sigma$  which are in causal contact with the boundary. This means that what happens there depends on the choices we make on the boundary. It's not particularly difficult to specify what happens on the boundary: for example, we could impose a version of reflecting boundary conditions, so that everything bounces back. But this doesn't change the fact that we have to specify something and, for this reason, the dynamical evolution is not determined by the initial data alone. In fancy language, we say that AdS is not *globally hyperbolic*: there exists no Cauchy surface on which we can specify initial data .

AdS is the setting for our best-understood theories of quantum gravity. It turns out



that gravitational dynamics in asymptotically AdS spacetimes is entirely equivalent to a quantum field theory living on the boundary  $\mathcal{I}$ . This idea goes by many names, including the *AdS/CFT correspondence*, *gauge-gravity duality*, or simply *holography*.

Unlike our other Penrose diagrams, our diagram for AdS still stretches to infinity. We can do better. We play our usual trick of introducing a coordinate which runs over finite values, now for time

$$\tilde{t} = \tan \tau \quad \Rightarrow \quad d\tilde{t} = \frac{d\tau}{\cos^2 \tau}$$

The metric (4.41) then becomes

$$ds^2 = \frac{R^2}{\cos^2 \psi \cos^4 \tau} (-d\tau^2 + \cos^4 \tau d\Omega_3^2)$$

where  $\tau \in (-\pi/2, +\pi/2)$ . Now we see that AdS is conformally equivalent to

$$d\tilde{s}^2 = -d\tau^2 + \cos^4 \tau (d\psi^2 + \sin^2 \psi d\Omega_2^2)$$

Ignoring the spatial  $\mathbf{S}^2$ , we can draw the resulting Penrose diagram as shown in the right-hand side of Figure 39. Now the spatial  $\mathbf{S}^3$  grows and shrinks with time, giving the strange almond-shape to the Penrose diagram. Again, we see that there is a timelike boundary  $\mathcal{I}$ , although now we can also show the future and past timelike infinity,  $i^\pm$ . The diagram again makes it clear that a lightray bounces back and forth an infinite number of times in AdS.

## 4.5 Coupling Matter

Until now, we've only discussed the dynamics of vacuum spacetime, with matter consigned to test particles moving on geodesics. But matter is not merely an actor on the spacetime stage: instead it backreacts, and affects the dynamics of spacetime itself.

### 4.5.1 Field Theories in Curved Spacetime

The first question we should ask is: how does matter couple to the spacetime metric? This is simplest to describe when matter takes the form of fields which themselves are governed by a Lagrangian. (We will look at what happens when matter is made of particles, albeit ones that form fluids, in Section 4.5.4.)

## Scalar Fields

As a simple example, consider a scalar field  $\phi(x)$ . In flat space, the action takes the form

$$S_{\text{scalar}} = \int d^4x \left( -\frac{1}{2} \eta^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) \right) \quad (4.42)$$

with  $\eta^{\mu\nu}$  the Minkowski metric. The minus sign in front of the derivative terms follows from the choice of signature  $(-+++)$ . This differs from, say, the lectures on [quantum field theory](#), but ensures that the action takes the form “kinetic energy minus potential energy”, with the spatial gradient terms counting towards the potential energy.

It is straightforward to generalise this to describe a field moving in curved spacetime: we simply need to replace the Minkowski metric with the curved metric, and ensure that we’re integrating over a multiple of the volume form. In practice, this means that we have

$$S_{\text{scalar}} = \int d^4x \sqrt{-g} \left( -\frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - V(\phi) \right) \quad (4.43)$$

Note that we’ve upgraded the derivatives from  $\partial_\mu$  to  $\nabla_\mu$ , although in this case it’s redundant because, on a scalar field,  $\nabla_\mu \phi = \partial_\mu \phi$ . Nonetheless, it will prove useful shortly.

Note, however, that curved spacetime also introduces new possibilities for us to add to the action. For example, we could equally well consider the theory

$$S_{\text{scalar}} = \int d^4x \sqrt{-g} \left( -\frac{1}{2} g^{\mu\nu} \nabla_\mu \phi \nabla_\nu \phi - V(\phi) - \frac{1}{2} \xi R \phi^2 \right) \quad (4.44)$$

for some constant  $\xi$ . This reduces to the flat space action (4.42) when we take  $g_{\mu\nu} = \eta_{\mu\nu}$  since the Ricci scalar is then  $R = 0$ , but it gives different dynamics for each choice of  $\xi$ . To derive the equation of motion for  $\phi$ , we vary the action (4.44) with respect to  $\phi$ , keeping  $g_{\mu\nu}$  fixed for now

$$\begin{aligned} \delta S_{\text{scalar}} &= \int d^4x \sqrt{-g} \left( -g^{\mu\nu} \nabla_\mu \delta\phi \nabla_\nu \phi - \frac{\partial V}{\partial \phi} \delta\phi - \xi R \phi \delta\phi \right) \\ &= \int d^4x \sqrt{-g} \left[ \left( g^{\mu\nu} \nabla_\mu \nabla_\nu \phi - \frac{\partial V}{\partial \phi} - \xi R \phi \right) \delta\phi - \nabla_\mu (\delta\phi \nabla^\mu \phi) \right] \end{aligned}$$

Notice that although the covariant derivatives  $\nabla_\mu$  could be replaced by  $\partial_\mu$  on the first line, they’re crucially important on the second where we needed the fact that  $\nabla_\mu g_{\rho\sigma} = 0$

to do the integration by parts. The final term is a boundary term (using the divergence theorem proven in Section 3.2.4) and can be discarded. This leaves us with the equation of motion for a scalar field in curved spacetime,

$$g^{\mu\nu}\nabla_\mu\nabla_\nu\phi - \frac{\partial V}{\partial\phi} - \xi R\phi = 0$$

Again, the covariant derivatives are needed here: we could write  $\nabla_\mu\nabla_\nu\phi = \nabla_\mu\partial_\nu\phi$  except it looks stupid. But  $\nabla_\mu\nabla_\nu\phi \neq \partial_\mu\partial_\nu\phi$ .

## Maxwell Theory

We already met the action for Maxwell theory in Section 3.2.5 as an example of integrating forms over manifolds. It is given by

$$S_{\text{Maxwell}} = -\frac{1}{2} \int F \wedge \star F = -\frac{1}{4} \int d^4x \sqrt{-g} g^{\mu\rho} g^{\nu\sigma} F_{\mu\nu} F_{\rho\sigma} \quad (4.45)$$

with  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \nabla_\mu A_\nu - \nabla_\nu A_\mu$ . (The equivalence of these two expressions follows because of anti-symmetry, with the Levi-Civita connections in the final term cancelling.) This time, the equations of motion are

$$\nabla_\mu F^{\mu\nu} = 0$$

Indeed, this is the only covariant tensor that we can write down that generalises the flat space result  $\partial_\mu F^{\mu\nu} = 0$ .

### 4.5.2 The Einstein Equations with Matter

To understand how fields backreact on spacetime, we just need to consider the combined action

$$S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} (R - 2\Lambda) + S_M$$

where  $S_M$  is the action for matter fields which, as we have seen above, depends on both the matter fields and the metric. We know what happens when we vary the Einstein-Hilbert action with respect to the metric. Now we care about  $S_M$ . We define the *energy-momentum tensor* to be

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}} \quad (4.46)$$

Notice that  $T_{\mu\nu}$  is symmetric, a property that it inherits from the metric  $g_{\mu\nu}$ . If we vary the full action with respect to the metric, we have

$$\delta S = \frac{1}{16\pi G} \int d^4x \sqrt{-g} (G_{\mu\nu} + \Lambda g_{\mu\nu}) \delta g^{\mu\nu} - \frac{1}{2} \int d^4x \sqrt{-g} T_{\mu\nu} \delta g^{\mu\nu}$$

From this we can read off the equations of motion,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu} \quad (4.47)$$

These are the full Einstein equations, describing gravity coupled to matter.

There are a number of different ways of writing this. First, the cosmological constant is sometimes absorbed as just another component of the energy-momentum tensor,

$$(T_{\mu\nu})_\Lambda = -\frac{\Lambda}{8\pi G} g_{\mu\nu} \quad (4.48)$$

One reason for this is that the matter fields can often mimic a cosmological constant and it makes sense to bundle all such terms together. (For example, a scalar field sitting at an extremal point of a potential is indistinguishable from a cosmological constant.) In this case, we just have

$$G_{\mu\nu} = 8\pi G T_{\mu\nu}$$

where  $T_{\mu\nu}$  now includes the cosmological term.

Taking the trace (i.e. contracting with  $g^{\mu\nu}$ ) then gives

$$-R = 8\pi G T$$

with  $T = g^{\mu\nu} T_{\mu\nu}$ . We can use this to directly relate the Ricci tensor to the energy momentum

$$R_{\mu\nu} = 8\pi G \left( T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu} \right) \quad (4.49)$$

This form will also have its uses in what follows.

### 4.5.3 The Energy-Momentum Tensor

The action  $S_M$  is constructed to be diffeomorphism invariant. This means that we can replay the argument of Section 4.1.3 that led us to the Bianchi identity: if we vary the metric by a diffeomorphism  $\delta g_{\mu\nu} = (\mathcal{L}_X g)_{\mu\nu} = 2\nabla_{(\mu} X_{\nu)}$ , then we have

$$\delta S_M = -2 \int d^4x \sqrt{-g} T_{\mu\nu} \nabla^\mu X^\nu = 0 \quad \text{for all } X^\mu$$

This tells us that the energy momentum tensor is necessarily covariantly conserved,

$$\nabla_\mu T^{\mu\nu} = 0 \quad (4.50)$$

Of course, this was necessary to make the Einstein equation (4.47) consistent, since we know that  $\nabla_\mu G^{\mu\nu} = 0$ . Indeed, viewed from the action principle, both the Bianchi identity and  $\nabla_\mu T^{\mu\nu} = 0$  have the same origin.

Although we've introduced the energy-momentum tensor as something arising from curved spacetime, it is also an important object in theories in flat space that have nothing to do with gravity. In that setting, the energy-momentum tensor arises as the Noether currents associated to translational invariance in space and time.

A hint of this is already apparent in (4.50) which, restricted to flat space, gives the expected conservation law enjoyed by Noether currents,  $\partial_\mu T^{\mu\nu} = 0$ . However, there is a rather slick argument that makes the link to Noether's theorem tighter.

In flat space, the energy-momentum tensor comes from invariance under translations  $x^\mu \rightarrow x^\mu + X^\mu$ , with constant  $X^\mu$ . There's a standard trick to compute the Noether current associated to any symmetry which involves promoting the symmetry parameters to be functions of the spacetime coordinates, so  $\delta x^\mu = X^\mu(x)$ . The action restricted to flat space is not invariant under such a shift. But it's simple to construct an action that is invariant: we simply couple the fields to a background metric and allow that to also vary. This is precisely the kind of action we've been considering in this section. The change of the action in flat space where we don't let the metric vary must be equal and opposite to the change of the action where we let the metric vary but don't change  $x^\mu$  (because the combination of the two vanishes). We must have

$$\delta S_{\text{flat}} = - \int d^4x \left. \frac{\delta S_M}{\delta g^{\mu\nu}} \right|_{g_{\mu\nu}=\eta_{\mu\nu}} \delta g^{\mu\nu}$$

But the variation of the metric *without* changing the point  $x^\mu$  is  $\delta g_{\mu\nu} = \partial_\mu X_\nu + \partial_\nu X_\mu$ . (The Christoffel symbols in the more familiar expression with  $\nabla_\mu$  come from the  $\partial g_{\mu\nu}$  term in (4.7), and this is precisely the term we neglect.) We have

$$\delta S_{\text{flat}} = -2 \int d^4x \left. \frac{\delta S_M}{\delta g^{\mu\nu}} \right|_{g_{\mu\nu}=\eta_{\mu\nu}} \partial^\mu X^\nu = -2 \int d^4x \partial^\mu \left( \left. \frac{\delta S_M}{\delta g^{\mu\nu}} \right|_{g_{\mu\nu}=\eta_{\mu\nu}} \right) X^\nu$$

But we know that  $\delta S_{\text{flat}} = 0$  whenever  $X^\mu = \text{constant}$ , since this is precisely what it means for the theory to be translationally invariant. We learn that the conserved Noether current in flat space is

$$T_{\mu\nu} \Big|_{\text{flat}} = -2 \left. \frac{\delta S_M}{\delta g^{\mu\nu}} \right|_{g_{\mu\nu}=\eta_{\mu\nu}}$$

which is the flat space version of (4.46).

## Examples of the Energy-Momentum Tensor

It is straightforward to compute the energy-momentum tensor for a scalar field. We take the action (4.43) and vary with respect to the metric. We will need the result  $\delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{\mu\nu}\delta g^{\mu\nu}$  from Section 4.1. We then find

$$\delta S_{\text{scalar}} = \int d^4x \sqrt{-g} \left( \frac{1}{4}g_{\mu\nu}\nabla^\rho\phi\nabla_\rho\phi + \frac{1}{2}g_{\mu\nu}V(\phi) - \frac{1}{2}\nabla_\mu\phi\nabla_\nu\phi \right) \delta g^{\mu\nu}$$

where the first two terms come from varying  $\sqrt{-g}$  and the third comes from varying the metric in the gradient term. This gives us the energy momentum tensor

$$T_{\mu\nu} = \nabla_\mu\phi\nabla_\nu\phi - g_{\mu\nu} \left( \frac{1}{2}\nabla^\rho\phi\nabla_\rho\phi + V(\phi) \right) \quad (4.51)$$

If we now restrict to flat space, with  $g_{\mu\nu} = \eta_{\mu\nu}$ , we find, for example,

$$T_{00} = \frac{1}{2}\dot{\phi}^2 + \frac{1}{2}(\nabla\phi)^2 + V(\phi)$$

where  $\nabla$  is the usual 3d spatial derivative. We recognise this as the energy density of a scalar field.

We can play the same game with the Maxwell action (4.45). Varying with respect to the metric, we have

$$\delta S_{\text{Maxwell}} = -\frac{1}{4} \int d^4x \sqrt{-g} \left( -\frac{1}{2}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma} + 2g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma} \right) \delta g^{\mu\nu}$$

So the energy momentum tensor is given by

$$T_{\mu\nu} = g^{\rho\sigma}F_{\mu\rho}F_{\nu\sigma} - \frac{1}{4}g_{\mu\nu}F^{\rho\sigma}F_{\rho\sigma} \quad (4.52)$$

In flat space, with  $g_{\mu\nu} = \eta_{\mu\nu}$ ,

$$T_{00} = \frac{1}{2}\mathbf{E}^2 + \frac{1}{2}\mathbf{B}^2$$

with  $F_{0i} = -E_i$ , the electric field, and  $F_{ij} = \epsilon_{ijk}B_k$  the magnetic field. Again, we recognise this as the energy density in the electric and magnetic fields. You can read more about the properties of the Maxwell energy-momentum tensor in the lecture on [Electromagnetism](#).

### 4.5.4 Perfect Fluids

Take any kind of object in the universe. Throw a bunch of them together, heat them up, and gently splash. The resulting physics will be described by the equations of fluid dynamics.

A perfect fluid is described by its *energy density*  $\rho(\mathbf{x}, t)$ , pressure  $P(\mathbf{x}, t)$  and a velocity 4-vector  $u^\mu(\mathbf{x}, t)$  such that  $u^\mu u_\mu = -1$ . The pressure and energy density are not unrelated: there is an identity between them that is usually called the *equation of state*,

$$P = P(\rho)$$

Common examples include *dust*, which consists of massive particles floating around, moving very slowly so that the pressure is  $P = 0$ , and *radiation*, which is a fluid made of many photons for which  $P = \rho/3$ .

The energy-momentum tensor for a perfect fluid is given by

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + P g^{\mu\nu} \quad (4.53)$$

If we are in Minkowski space, so  $g_{\mu\nu} = \eta_{\mu\nu}$  and the fluid is at rest, so  $u^\mu = (1, 0, 0, 0)$ , then the energy momentum tensor is

$$T^{\mu\nu} = \text{diag}(\rho, P, P, P)$$

We see that  $T^{00} = \rho$ , as expected for the energy density. More generally, for a moving fluid we have  $T_{\mu\nu}u^\mu u^\nu = \rho$ , which means that  $\rho$  is the energy density measured by an observer co-moving with the fluid.

The energy-momentum tensor must obey

$$\nabla_\mu T^{\mu\nu} = 0$$

A short calculation shows that this is equivalent to two relations between the fluid variables. The first is

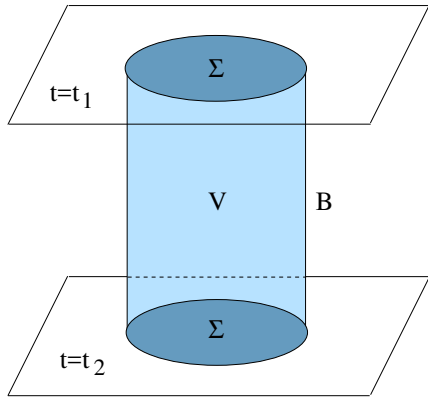
$$u^\mu \nabla_\mu \rho + (\rho + P) \nabla_\mu u^\mu = 0 \quad (4.54)$$

This is the relativistic generalisation of mass conservation for a fluid. Here “mass” has been replaced by energy density  $\rho$ . The first term,  $u^\mu \nabla_\mu \rho$  calculates how fast the energy density is changing as we move along  $u^\mu$ . The second term tells us the answer: it depends on  $\nabla_\mu u^\mu$ , the rate of flow of fluid out of a region.

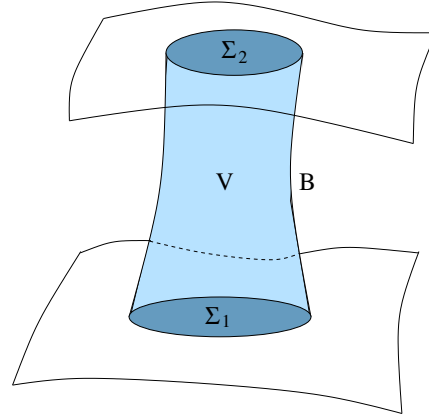
The second constraint is

$$(\rho + P)u^\mu \nabla_\mu u^\nu = -(g^{\mu\nu} + u^\mu u^\nu) \nabla_\mu P \quad (4.55)$$

This is the relativistic generalisation of the Euler equation, the fluid version of Newton’s second law “ $F=ma$ ”. The left-hand side of the equation should be viewed as “mass  $\times$  acceleration”, the right-hand side is the force which, in a fluid, is due to pressure differences. You can learn more about these equations in the relativistic context and their solutions in chapter 3 of the lectures on [Cosmology](#).



**Figure 40:** Charge conservation in flat spacetime...



**Figure 41:** ...and in curved spacetime.

#### 4.5.5 The Slippery Business of Energy Conservation

In flat space, the existence of an energy-momentum tensor ensures that we can define the conserved quantities, energy and momentum. In curved spacetime, things are significantly more subtle.

To see this, it's useful to compare the energy-momentum tensor  $T^{\mu\nu}$  with a current  $J^\mu$  which arises from a global symmetry (such as, for example, the phase rotation of a complex scalar field). In flat space, both obey current conservation

$$\partial_\mu J^\mu = 0 \quad \text{and} \quad \partial_\mu T^{\mu\nu} = 0 \quad (4.56)$$

From these, we can construct conserved charge by integrating over a spatial volume  $\Sigma$

$$Q(\Sigma) = \int_\Sigma d^3x J^0 \quad \text{and} \quad P^\mu(\Sigma) = \int_\Sigma d^3x T^{0\mu}$$

To see that these are indeed conserved, we simply need to integrate over a spacetime volume  $V$ , bounded by  $\Sigma$  at past and future times  $t_1$  and  $t_2$ . We then have

$$0 = \int_V d^4x \partial_\mu J^\mu = \Delta Q(\Sigma) + \int_B d^3x n_i J^i$$

where  $\Delta Q(\Sigma) = [Q(\Sigma)](t_2) - [Q(\Sigma)](t_1)$  and  $n^i$  is the outward-pointing normal to  $B$ , the timelike boundary of  $V$ , as shown in the left-hand figure. Provided that no current flows out of the region, meaning  $J_i = 0$  when evaluated on  $B$ , we have  $\Delta Q(\Sigma) = 0$ . (Often, we take  $\Sigma = \mathbf{R}^3$  so that  $B = \mathbf{S}_\infty^2 \times I$  with  $I$  an interval, and we only have to require that there are no currents at infinity.) This is the statement of charge conservation.



In Minkowski space, this same argument works just as well for  $P^\mu(\Sigma)$ , meaning that we are able to assign conserved energy and momentum to fields in some region, providing that no currents leak out through the boundary.

Let's now contrast this to the situation in curved spacetime. The conservation laws (4.56) are replaced by their covariant versions

$$\nabla_\mu J^\mu = 0 \quad \text{and} \quad \nabla_\mu T^{\mu\nu} = 0$$

We can replay the argument above, now invoking the divergence theorem from Section 3.2.4

$$0 = \int_V d^4x \sqrt{-g} \nabla_\mu J^\mu = \int_{\partial V} d^3x \sqrt{|\gamma|} n_\mu J^\mu$$

with  $\gamma_{ij}$  the pull-back of the metric to  $\partial V$  and  $n^\mu$  the normal vector. We consider a spacetime volume  $V$  with boundary

$$\partial V = \Sigma_1 \cup \Sigma_2 \cup B$$

Here  $\Sigma_1$  and  $\Sigma_2$  are past and future spacelike boundaries, while  $B$  is the timelike boundary as shown in the right-hand figure. If we again insist that no current flows out of the region by requiring that  $J^\mu n_\mu = 0$  when evaluated on  $B$ , then the expression above becomes

$$Q(\Sigma_2) = Q(\Sigma_1)$$

where the charge  $Q(\Sigma)$  evaluated on a spacelike hypersurface  $\Sigma$  is defined by

$$Q(\Sigma) = \int_\Sigma d^3x \sqrt{\gamma} n_\mu J^\mu$$

This means that, for a vector field, covariant conservation is the same thing as actual conservation. The story above is a repeat of the one we told using differential forms in Section 3.2.5.

Now let's try to tell a similar story for the energy-momentum tensor. In analogy to the derivation above, it's clear that we should try to manipulate the integral

$$0 = \int_V d^4x \sqrt{-g} \nabla_\mu T^{\mu\nu}$$

The problem is that we don't have a divergence theorem for integrals of this kind because of the hanging  $\nu$  index on the energy momentum tensor. The key to deriving the divergence theorem for  $J^\mu$  was the expression

$$\nabla_\mu J^\mu = \partial_\mu J^\mu + \Gamma_{\mu\rho}^\mu J^\rho = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} J^\mu)$$

This allows us to turn a covariant derivative into a normal derivative which gives a boundary term in the integral. However, the same expression for the energy-momentum tensor reads

$$\nabla_\mu T^{\mu\nu} = \partial_\mu T^{\mu\nu} + \Gamma_{\mu\rho}^\mu T^{\rho\nu} + \Gamma_{\mu\rho}^\nu T^{\mu\rho} = \frac{1}{\sqrt{-g}} \partial_\mu (\sqrt{-g} T^{\mu\nu}) + \Gamma_{\mu\rho}^\nu T^{\mu\rho}$$

That extra term involving the Christoffel symbol stops us converting the integral of  $\nabla_\mu T^{\mu\nu}$  into a boundary term. Indeed, we can rewrite  $\nabla_\mu T^{\mu\nu} = 0$  as

$$\partial_\mu (\sqrt{-g} T^{\mu\nu}) = -\sqrt{-g} \Gamma_{\mu\rho}^\nu T^{\mu\rho} \quad (4.57)$$

then the right-hand side looks like a driving force which destroys conservation of energy and momentum. We learn that, for a higher tensor like  $T^{\mu\nu}$ , covariant conservation is not the same thing as actual conservation!

### Conserved Energy from a Killing Vector

We can make progress by introducing one further ingredient. If our spacetime has a Killing vector field  $K$ , we can construct a current from the energy-momentum tensor by writing

$$J_T^\nu = -K_\mu T^{\mu\nu}$$

Taking the covariant divergence of the current gives

$$-\nabla_\nu J_T^\nu = (\nabla_\nu K_\mu) T^{\mu\nu} + K_\mu \nabla_\nu T^{\mu\nu} = 0$$

where the first term vanishes by virtue of the Killing equation, with  $T^{\mu\nu}$  imprinting its symmetric indices on  $\nabla_\nu K_\mu$ , and the second term vanishes by (4.50).

Now we're in business. We can construct conserved charges from the current  $J_T^\mu$  as explained above,

$$Q_T(\Sigma) = \int_\Sigma d^3x \sqrt{\gamma} n_\mu J_T^\mu$$

The interpretation of these charges depends on the properties of the Killing vector. If  $K^\mu$  is everywhere timelike, meaning  $g_{\mu\nu} K^\mu K^\nu < 0$  at all points, then the charge can be identified with the energy of the matter

$$E = Q_T(\Sigma)$$

If the Killing vector is everywhere spacelike, meaning  $g_{\mu\nu} K^\mu K^\nu > 0$  at all points, then the charge can be identified with the momentum of the matter.

## Conserved Energy Without a Killing Vector?

There are situations where spacetime does not have a Killing vector yet, intuitively, we would still like to associate something analogous to energy. This is where things start to get subtle.

A simple situation where this arises is two orbiting stars. It turns out that the resulting spacetime does not admit a timelike Killing vector. As we will describe in some detail in Section 5.3, as the stars orbit they emit gravitational waves, losing energy, which causes them to slowly spiral towards each other. This fits nicely with the lack of a timelike Killing vector, since we wouldn't expect to define a conserved energy for the stars.

However, it certainly feels like we should be able to define a conserved energy for the total system which, in this case, means stars together with the gravitational waves. In particular, we would like to quantify the amount of energy lost by the stars and carried away by the gravitational waves. But this requires us to define something new, namely the energy density in the gravitational field. And this is where the trouble starts!

There's an obvious way to proceed, one that starts by returning to our original definition of the energy-momentum tensor (4.46)

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}} \frac{\delta S_M}{\delta g^{\mu\nu}}$$

A naive guess would be to include the action for both matter and gravity in this definition, giving an energy momentum tensor which includes both matter and gravity

$$T_{\mu\nu}^{\text{total}} = -\frac{2}{\sqrt{-g}} \left( \frac{1}{16\pi G_N} \frac{\delta S_{EH}}{\delta g^{\mu\nu}} + \frac{\delta S_M}{\delta g^{\mu\nu}} \right)$$

But this gives

$$T_{\mu\nu}^{\text{total}} = -\frac{1}{8\pi G_N} G_{\mu\nu} + T_{\mu\nu} = 0$$

which vanishes by the Einstein field equations. The idea that the total energy of the universe vanishes, with negative gravitational energy cancelling the positive energy from matter sounds like it might be something important. It turns out to be very good for selling pseudo-scientific books designed to make the world think you're having deep thoughts. It's not, however, particularly good for anything to do with physics. For example, it's not as if electrons and positrons are suddenly materialising everywhere in space, their mass energy cancelled by the gravitational energy. That's not the way the universe works. Instead the right way to think about this equation is to simply appreciate that energy in general relativity is subtle.

Clearly, we should try to do better to understand the energy carried in the gravitational field. Unfortunately, it turns out that doing better is challenging. There are compelling arguments that show there is no tensor that can be thought of as the local energy density of the gravitational field. Roughly speaking, these arguments start from the observation that the energy in the Newtonian gravitational field is proportional to  $(\nabla\Phi)^2$ . We should therefore expect that the relativistic version of energy density is proportional to the first derivative of the metric. Yet the equivalence principle tells us that we can always find coordinates – those experienced by a freely falling observer – which ensure that the first derivative of the metric vanish at any given point. But a tensor that vanishes in one coordinate system also vanishes in another.

We'll confront these issues again in Section 5.3 when we try to answer the question of how much energy emitted in gravitational waves by a binary star system. We'll see there that, for this case, there are simplifications that mean we can converge on a sensible answer.

#### 4.5.6 Spinors

In flat space, fermions transform in the spinor representation of the  $SO(1, 3)$  Lorentz group. Recall from the lectures on [Quantum Field Theory](#) that we first introduce gamma matrices obeying the Clifford algebra

$$\{\gamma_a, \gamma_b\} = 2\eta_{ab} \tag{4.58}$$

Notice that we've put indices  $a, b = 0, 1, 2, 3$  on the gamma matrices, rather than the more familiar  $\mu, \nu$ . This is deliberate. In  $d = 3 + 1$  dimensions, each of the  $\gamma_a$  is a  $4 \times 4$  matrix.

We can use these gamma matrices to construct generators of the Lorentz group,

$$S_{ab} = \frac{1}{4}[\gamma_a, \gamma_b]$$

These define the spinor representation of the Lorentz group. We write a Lorentz transformation  $\Lambda$  as

$$\Lambda = \exp\left(\frac{1}{2}\lambda^{ab}M_{ab}\right)$$

with  $M_{ab}$  the usual Lorentz generators (defined, for example, in (4.30)) and  $\lambda^{ab}$  a choice of 6 numbers that specify the particular Lorentz transformation. Then the corresponding transformation in the spinor representation is given by

$$S[\Lambda] = \exp\left(\frac{1}{2}\lambda^{ab}S_{ab}\right)$$

A Dirac spinor field  $\psi_\alpha(x)$  is then a 4-component complex vector that, under a Lorentz transformation, changes as

$$\psi(x) \rightarrow S[\Lambda] \psi(\Lambda^{-1}x) \quad (4.59)$$

In Minkowski space, the action for the spinor is

$$S_{\text{Dirac}} = \int d^4x \, i (\bar{\psi} \gamma^a \partial_a \psi + m \bar{\psi} \psi)$$

with  $\bar{\psi} = \psi^\dagger \gamma^0$ . The magic of gamma matrices ensures that this action is invariant under Lorentz transformations, despite having just a single derivative. Our task in this section is to generalise this action to curved spacetime.

We can already see some obstacles. The gamma matrices (4.58) are defined in Minkowski space and it's not clear that they would retain their magic if generalised to curved space. Furthermore, what should we do with the derivative? We might suspect that it gets replaced by a covariant derivative, but what connection do we choose?

To answer these questions, we will need to invoke the vierbeins and connection one-form that we met in Section 3.4.2. Recall that the vierbeins  $\hat{e}_a = e_a^\mu \partial_\mu$  are a collection of 4 vector fields, that allow us “diagonalise” the metric. We define  $e^a_\mu$  to be the inverse of  $e_a^\mu$ , meaning it satisfies  $e^a_\mu e_b^\mu = \delta_b^a$  and  $e^a_\mu e_a^\nu = \delta_\mu^\nu$ . The metric can then be written

$$g_{\mu\nu} = e^a_\mu e^b_\nu \eta_{ab}$$

These formula are really telling us that we should raise/lower the  $\mu, \nu$  indices using the metric  $g_{\mu\nu}$  but raise/lower the  $a, b$  indices using the Minkowski metric  $\eta_{ab}$ .

The formalism of vierbeins allowed us to introduce the idea of a *local Lorentz transformation*  $\Lambda(x)$ , defined in (3.43), which acts on the vierbeins as

$$e_a^\nu \rightarrow \tilde{e}_a^\nu = e_b^\nu (\Lambda^{-1})^b_a \quad \text{with} \quad \Lambda_a^c \Lambda_b^d \eta_{cd} = \eta_{ab}$$

This local Lorentz transformation can now be promoted to act on a spinor field as (4.59), again with  $S[\Lambda]$  depending on the coordinate  $x$ .

We want our action to be invariant under these local Lorentz transformations. In particular, we might expect to run into difficulties with the derivative which, after a Lorentz transformation, now hits  $\Lambda(x)$  as well as  $\psi(x)$ . But this is exactly the kind of problem that we've met before when writing down actions for gauge theories, and we know very well how to solve it: we simply need to include a connection in the action

that transforms accordingly. To this end, we construct the covariant derivative acting on a spinor field

$$\nabla_\mu \psi = \partial_\mu \psi + \frac{1}{2} \omega_\mu^{ab} S_{ab} \psi$$

with  $\omega_\mu^{ab}$  the appropriate connection. But what is it?

The right choice is the connection one-form, also known as the spin connection, that we met in Section 3.4.2. From (3.44) and (3.47), we have

$$(\omega^a_b)_\mu = \Gamma_{cb}^a e^c_\mu = e^a_\rho \nabla_\mu e_b^\rho$$

This does the trick because of its inhomogeneous transformation (3.45)

$$(\omega^a_b)_\mu \rightarrow \Lambda^a_c (\omega^c_d)_\mu (\Lambda^{-1})^d_b + \Lambda^a_c (\partial_\mu \Lambda^{-1})^c_b$$

This cancels the contribution from the derivative in the same way as the covariant derivative in a non-Abelian gauge theory. The generalisation of the Dirac action to curved space is then simply

$$S_{\text{Dirac}} = \int d^4x \sqrt{-g} i (\bar{\psi} \gamma^a e_a^\mu \nabla_\mu \psi + m \bar{\psi} \psi)$$

There are a number of reasons to be interested in coupling fermions to gravity. First, and most obviously, both are constituents of our universe and its important to understand how they fit together. Second, they are important for more formal aspects of mathematical physics: they are the key component in Witten's simple proof of the positive mass theorem, and there are reasons to suspect that the quantisation of gravity ultimately requires supersymmetry at a high energy scale.

However, there is one thing that you probably shouldn't do with them, which is put them on the right-hand side of the Einstein equation and solve them. This is because fermions are quantum fields and do not have a macroscopic, classical analog.

Of course, all fields are, at heart, quantum. But for bosonic fields, it makes sense to think of them classically where they can be viewed as quantum fields with high occupation number. This is familiar in electromagnetism, where the classical electric and magnetic field can be thought of containing many photons. This means that it makes sense to find spacetimes which solve the Einstein equations  $G_{\mu\nu} = 8\pi G T_{\mu\nu}$  where the curvature is supported by a profile for scalar fields on the right-hand-side.

In contrast, there is no classical limit of fermionic fields. This is because the Pauli exclusion principle prohibits a large occupation number. If you therefore attempt to find a spacetime supported by a fermionic  $T_{\mu\nu}$ , you are really looking for a gravitational solution sourced by precisely one quantum excitation. Given the feebleness of gravity on the microscopic scale, this is unlikely to be interesting.

This is not to say that fermions don't affect gravity. Important examples for gravitating fermionic systems include white dwarfs and neutron stars. But in each of these cases there is a separation of scales where one can first neglect gravity and find an effective equation of state for the fermions, and subsequently understand how this backreacts on spacetime. If you want to understand the spacetime directly from the Dirac equation than you have a complicated many-body problem on your hands.

#### 4.5.7 Energy Conditions

If we know the kind of matter that fills spacetime, then we can just go ahead and solve the Einstein equations. However, we will often want to make more general statements about the allowed properties of spacetime without reference to any specific matter content. In this case, it is useful to place certain restrictions on the kinds of energy-momentum tensor that we consider physical.

These restrictions, known as *energy conditions*, capture the rough idea that energy should be positive. A number of classic results in general relativity, such as the singularity theorems, rely on these energy conditions as assumptions.

There are a bewildering number of these energy conditions. Moreover, it is not difficult to find examples of matter which violate most of them! We now describe a number of the most important energy conditions, together with their limitations.

- **Weak Energy Condition:** This states that, for any timelike vector field  $X$ ,

$$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X_\mu X^\mu < 0$$

The idea is that this quantity is the energy seen by an observer moving along the timelike integral curves of  $X$ , and this should be non-negative. A timelike curve can get arbitrarily close to a null curve so, by continuity, the weak energy condition can be extended to timelike and null curves

$$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X_\mu X^\mu \leq 0 \tag{4.60}$$

To get a feel for this requirement, let's first impose it on the energy-momentum tensor for a perfect fluid (4.53). We will consider timelike vectors  $X$  normalised

to  $X \cdot X = -1$ . We then have

$$T_{\mu\nu}X^\mu X^\nu = (\rho + P)(u \cdot X)^2 - P \geq 0$$

We work in the rest frame of the fluid, so  $u^\mu = (1, 0, 0, 0)$  and consider constant timelike vector fields,  $X^\mu = (\cosh \varphi, \sinh \varphi, 0, 0)$ . These describe the worldlines of observers boosted with rapidity  $\varphi$  with respect to the fluid. The weak energy condition then gives us

$$(\rho + P) \cosh^2 \varphi - P \geq 0 \quad \Rightarrow \quad \begin{cases} \rho \geq 0 & \text{when } \varphi = 0 \\ P \geq -\rho & \text{as } \varphi \rightarrow \infty \end{cases}$$

The first condition  $\rho \geq 0$  is what we expect from the weak energy condition: it ensures that the energy density is positive. The second condition  $P \geq -\rho$  says that negative pressure is acceptable, just as long as it's not too negative.

There are, however, situations in which negative energy density makes physical sense. Indeed, we've met one already: if we view the cosmological constant as part of the energy momentum tensor, as in (4.48), then any  $\Lambda < 0$  violates the weak energy condition. Viewed this way, anti-de Sitter spacetime violates the weak energy condition.

We can also look at how this condition fares for scalar fields. From the energy-momentum tensor (4.51), we have

$$(X^\mu \partial_\mu \phi)^2 + \frac{1}{2} \partial_\mu \phi \partial^\mu \phi + V(\phi) \geq 0 \quad (4.61)$$

The first term is positive, but the second term can have either sign. In fact, it turns out that the first and second term combined are always positive. To see this, define the vector  $Y$  orthogonal to  $X$

$$Y_\mu = \partial_\mu \phi + X_\mu (X^\nu \partial_\nu \phi)$$

This satisfies  $X_\mu Y^\mu = 0$ : it is the projection of  $\partial_\mu \phi$  onto directions orthogonal to  $X$ . Because  $X$  is timelike,  $Y$  must be spacelike (or null) and so obeys  $Y_\mu Y^\mu \geq 0$ . The weak energy condition (4.61) can be rewritten as

$$\frac{1}{2} (X^\mu \partial_\mu \phi)^2 + \frac{1}{2} Y_\mu Y^\mu + V(\phi) \geq 0$$

Now the first two terms are positive. We see that the weak energy condition is satisfied provided that  $V(\phi) \geq 0$ . However, it is violated in any classical theory with  $V(\phi) \leq 0$  and there's no reason to forbid such negative potentials for a scalar field.



- **Strong Energy Condition:** There is a different, less immediately intuitive, energy condition. This is the requirement that, for any timelike vector field  $X$ ,

$$R_{\mu\nu}X^\mu X^\nu \geq 0$$

This is the *strong energy condition*. It is poorly named. The strong energy condition is neither stronger nor weaker than the weak energy condition: it is simply different. It turns out that the strong energy condition ensures that timelike geodesics converge, which can be viewed as the statement that gravity is attractive. (This connection is made using something called the Raychaudhuri equation.)

Using the form of the Einstein equations (4.49), the strong energy condition requires

$$\left(T_{\mu\nu} - \frac{1}{2}Tg_{\mu\nu}\right)X^\mu X^\nu \geq 0$$

for all timelike vector fields  $X$ . As before, continuity ensures that we can extend this to timelike and null vector fields,  $X \cdot X \leq 0$ .

If we take  $X \cdot X = -1$  then, applied to a perfect fluid (4.53), the strong energy condition requires

$$(\rho + P)(u \cdot X)^2 - P + \frac{1}{2}(3P - \rho) \geq 0$$

As before, we consider the fluid in its rest frame with  $u^\mu = (1, 0, 0, 0)$  and look at this condition for boosted observers with  $X^\mu = (\cosh \varphi, \sinh \varphi, 0, 0)$ . We have

$$(\rho + P) \cosh^2 \varphi + \frac{1}{2}(P - \rho) \geq 0 \quad \Rightarrow \quad \begin{cases} P \geq -\rho/3 & \varphi = 0 \\ P \geq -\rho & \varphi \rightarrow \infty \end{cases}$$

Once again, it is not difficult to find situations where the strong energy condition is violated. Most strikingly, a cosmological constant  $\Lambda > 0$  is not compatible with the strong energy condition. In fact, we may have suspected this because neighbouring, timelike geodesics in de Sitter space are pulled apart by the expansion of space. In fact, the strong energy condition forbids any FRW universe with  $\ddot{a} > 0$ , but there are at least two periods when our own universe underwent accelerated expansion: during inflation, and now.

Finally, it's not hard to show that any classical scalar field with a positive potential energy will violate the strong energy condition.

- **Null Energy Condition:** The null energy condition

$$T_{\mu\nu}X^\mu X^\nu \geq 0 \quad \text{for all } X \text{ with } X \cdot X = 0$$

This is implied by both weak and strong energy conditions, but the converse is not true: the null energy condition is strictly weaker than both the weak and strong conditions. This, of course, means that it is less powerful if we wield it to prove various statements. However, the null energy condition has the advantage that it is satisfied by any sensible classical field theory and any perfect fluid that obeys  $\rho + P \geq 0$ .

- **Dominant Energy Condition:** There is also an energy condition which is stronger than the weak condition. For any future-directed timelike vector  $X$ , we can define the current

$$J^\mu = -T^{\mu\nu}X_\nu \tag{4.62}$$

This is energy density current as seen by an observer following the lines of  $X$ . The dominant energy condition requires that, in addition to the weak energy condition (4.60), the current is either timelike or null, so

$$J_\mu J^\mu \leq 0$$

This is the reasonable statement that energy doesn't flow faster than light.

One can check that the extra condition (4.62) is satisfied for a scalar field. For a perfect fluid we have

$$J^\mu = -(\rho + P)(u \cdot X)u^\mu - PX^\mu$$

It's simple to check that the requirement  $J_\mu J^\mu \leq 0$  is simply  $\rho^2 \geq P^2$ .

The validity of the various energy conditions becomes murkier still in the quantum world. We can consider quantum matter coupled to a classical yet dynamical spacetime through the equation

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G \langle T_{\mu\nu} \rangle$$

where  $\langle T_{\mu\nu} \rangle$  is the expectation value of the energy-momentum tensor. Each of the energy conditions listed above is violated by fairly standard quantum field theories.

There is, however, a somewhat weaker statement that holds true in general. This is the *averaged null energy condition*. It can be proven that, along an infinite, *achronal* null geodesic, any reasonable quantum field theory obeys

$$\int_{-\infty}^{+\infty} d\lambda \langle T_{\mu\nu} \rangle X^\mu X^\nu \geq 0$$

Here  $\lambda$  is an affine parameter along the null geodesic and the vector  $X^\mu$  points along the geodesic and is normalised to  $X^\mu \partial_\mu \lambda = 1$ . Here the word “achronal” means that no two points on the geodesic can be connected by a timelike curve. (As a counterexample, consider an infinite null ray on  $M = \mathbf{R} \times \mathbf{S}^1$  which continually orbits the spatial circle. This geodesic is not achronal and the averaged null energy condition is not, in general, obeyed along this geodesic.)

## 4.6 A Taste of Cosmology

There are surprisingly few phenomena in Nature where we need to solve the Einstein equations sourced by matter,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}$$

However there is one situation where the role of  $T_{\mu\nu}$  on the right-hand side is crucial: this is cosmology, the study of the universe as a whole.

### 4.6.1 The FRW Metric

The key assumption of cosmology is that the universe is spatially homogeneous and isotropic. This restricts our choices of spatial geometry to one of three: these are

- Euclidean Space  $\mathbf{R}^3$ : This space has vanishing curvature and the familiar metric

$$ds^2 = dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2)$$

- Sphere  $\mathbf{S}^3$ : This space has uniform positive curvature and metric

$$ds^2 = \frac{1}{1-r^2} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2)$$

With this choice of coordinates, we have implicitly set the radius of the sphere to 1.

- Hyperboloid  $\mathbf{H}^3$ : This space has uniform negative curvature and metric

$$ds^2 = \frac{1}{1+r^2} dr^2 + r^2(d\theta^2 + \sin^2 \theta d\phi^2)$$

The existence of three symmetric spaces is entirely analogous to the three different solutions we discussed in Section 4.2. de Sitter and anti-de Sitter both have constant *spacetime* curvature, supplied by the cosmological constant. The metrics above have constant *spatial* curvature. Note, however, that the metric on  $\mathbf{S}^3$  coincides with the spatial part of the de Sitter metric in coordinates (4.12), while the metric on  $\mathbf{H}^3$  coincides with the spatial part of the anti-de Sitter metric in coordinates (4.22)

We write these spatial metrics in unified form,

$$ds^2 = \gamma_{ij} dx^i dx^j = \frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad \text{with } k = \begin{cases} +1 & \mathbf{S}^3 \\ 0 & \mathbf{R}^3 \\ -1 & \mathbf{H}^3 \end{cases}$$

In cosmology, we wish to describe a spacetime in which space expands as the universe evolves. We do this with metrics of the form

$$ds^2 = -dt^2 + a^2(t)\gamma_{ij}dx^i dx^j \tag{4.63}$$

This is the Friedmann-Robertson-Walker, or FRW metric. (It is also known as the FLRW metric, with Lemaître’s name a worthy addition to the list.) The dimensionless scale factor  $a(t)$  should be viewed as the “size” of the spatial dimensions (a name which makes more sense for the compact  $\mathbf{S}^3$  than the non-compact  $\mathbf{R}^3$ , but is mathematically sensible for both.) Note that de Sitter space in global coordinates (4.21) is an example of an FRW metric with  $k = +1$ .

### Curvature Tensors

We wish to solve the Einstein equations for metrics that take the FRW form. Our first task is to compute the Ricci tensor. We start with the Christoffel symbols: it is straightforward to find  $\Gamma_{00}^\mu = \Gamma_{i0}^0 = 0$  and

$$\Gamma_{ij}^0 = a\dot{a}\gamma_{ij} \quad , \quad \Gamma_{0j}^i = \frac{\dot{a}}{a}\delta_j^i \quad , \quad \Gamma_{jk}^i = \frac{1}{2}\gamma^{il}(\partial_j\gamma_{kl} + \partial_k\gamma_{jl} - \partial_l\gamma_{jk})$$

To compute the Ricci tensor, we use the expression

$$R_{\mu\nu} = \partial_\rho\Gamma_{\nu\mu}^\rho - \partial_\nu\Gamma_{\rho\mu}^\rho + \Gamma_{\nu\mu}^\lambda\Gamma_{\rho\lambda}^\rho - \Gamma_{\rho\mu}^\lambda\Gamma_{\nu\lambda}^\rho \tag{4.64}$$

which we get from contracting indices on the similar expression (3.39) for the Riemann tensor

It's not hard to see that  $R_{0i} = 0$ . The quick argument is that there's no covariant 3-vector that could possibly sit on the right-hand side. The other components need a little more work.

**Claim:**

$$R_{00} = -3\frac{\ddot{a}}{a}$$

**Proof:** Using the non-vanishing Christoffel symbols listed above, we have

$$R_{00} = -\partial_0\Gamma_{i0}^i - \Gamma_{i0}^j\Gamma_{j0}^i = -3\frac{d}{dt}\left(\frac{\dot{a}}{a}\right) - 3\left(\frac{\dot{a}}{a}\right)^2$$

which gives the claimed result □

**Claim:**

$$R_{ij} = \left(\frac{\ddot{a}}{a} + 2\left(\frac{\dot{a}}{a}\right)^2 + 2\frac{k}{a^2}\right)g_{ij}$$

**Proof:** This is straightforward to show for  $k = 0$  FRW metrics where the spatial metric is flat. It's a little more annoying for the  $k = \pm 1$  metrics. A trick that simplifies life is to compute the components of  $R_{ij}$  at the spatial origin  $\mathbf{x} = 0$  where the spatial metric is  $\gamma_{ij} = \delta_{ij}$ , and then use covariance to argue that the right result must have  $R_{ij} \sim \gamma_{ij}$ . In doing this, we just have to remember not to set  $\mathbf{x} = 0$  too soon, since we will first need to differentiate the Christoffel symbols and then evaluate them at  $\mathbf{x} = 0$ .

We start by writing the spatial metric in Cartesian coordinates, on the grounds that it's easier to differentiate in this form

$$\gamma_{ij} = \delta_{ij} + \frac{kx_ix_j}{1 - k\mathbf{x} \cdot \mathbf{x}}$$

The Christoffel symbols depend on  $\partial\gamma_{ij}$  and the Ricci tensor on  $\partial^2\gamma_{ij}$ . This means that if we want to evaluate the Ricci tensor at the origin  $\mathbf{x} = 0$ , we only need to work with the metric to quadratic order in  $x$ . This simplifies things tremendously since

$$\gamma_{ij} = \delta_{ij} + kx_ix_j + \mathcal{O}(x^4)$$

Similarly, we have

$$\gamma^{ij} = \delta^{ij} - kx^ix^j + \mathcal{O}(x^4)$$

where  $i, j$  indices are raised and lowered using  $\delta^{ij}$ . Plugging these forms into the expression for the Christoffel symbols gives

$$\Gamma_{jk}^i = kx^i \delta_{jk} + \mathcal{O}(x^3)$$

With this in hand, we can compute the Ricci tensor

$$\begin{aligned} R_{ij} &= \partial_\rho \Gamma_{ij}^\rho - \partial_j \Gamma_{\rho i}^\rho + \Gamma_{ij}^\lambda \Gamma_{\rho\lambda}^\rho - \Gamma_{\rho i}^\lambda \Gamma_{j\lambda}^\rho \\ &= (\partial_0 \Gamma_{ij}^0 + \partial_k \Gamma_{ij}^k) - \partial_j \Gamma_{ki}^k + (\Gamma_{ij}^0 \Gamma_{k0}^k + \Gamma_{ij}^k \Gamma_{lk}^l) - (\Gamma_{ki}^0 \Gamma_{j0}^k + \Gamma_{0i}^k \Gamma_{jk}^0 + \Gamma_{li}^k \Gamma_{jk}^l) \end{aligned}$$

We can drop the  $\Gamma_{ij}^k \Gamma_{lk}^l$  term since it vanishes at  $\mathbf{x} = 0$ . Furthermore, we can now safely replace any undifferentiated  $\gamma_{ij}$  in the Christoffel symbols with  $\delta_{ij}$ . What's left gives

$$\begin{aligned} R_{ij} &= (\partial_0(a\dot{a}) + 3k - k + 3\dot{a}^2 - \dot{a}^2 - \dot{a}^2) \delta_{ij} + \mathcal{O}(x^2) \\ &= (a\ddot{a} + 2\dot{a}^2 + 2k) \delta_{ij} + \mathcal{O}(x^2) \end{aligned}$$

We now invoke the covariance argument to write

$$R_{ij} = (a\ddot{a} + 2\dot{a}^2 + 2k) \gamma_{ij} = \frac{1}{a^2} (a\ddot{a} + 2\dot{a}^2 + 2k) g_{ij}$$

as promised □

With these results, we can now compute the Ricci scalar: it is

$$R = 6 \left( \frac{\ddot{a}}{a} + \left( \frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} \right)$$

Finally, the Einstein tensor has components

$$G_{00} = 3 \left( \left( \frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} \right) \quad \text{and} \quad G_{ij} = - \left( 2 \frac{\ddot{a}}{a} + \left( \frac{\dot{a}}{a} \right)^2 + \frac{k}{a^2} \right) g_{ij}$$

Our next task is to understand the matter content in the universe.

#### 4.6.2 The Friedmann Equations

We take the universe to be filled with perfect fluids of the kind that we introduced in Section 4.5.4. The energy momentum tensor is

$$T^{\mu\nu} = (\rho + P)u^\mu u^\nu + P g^{\mu\nu}$$

But we assume that the fluid is at rest in the preferred frame of the universe, meaning that  $u^\mu = (1, 0, 0, 0)$  in the FRW coordinates (4.63). As we saw in Section 4.5.4, the constraint  $\nabla_\mu T^{\mu\nu} = 0$  gives the condition (4.54)

$$u^\mu \nabla_\mu \rho + (\rho + P) \nabla_\mu u^\mu = 0 \quad \Rightarrow \quad \dot{\rho} + \frac{3\dot{a}}{a}(\rho + P) = 0$$

where we've used  $\nabla_\mu u^\mu = \partial_\mu u^\mu + \Gamma_{\mu\rho}^\mu u^\rho = \Gamma_{i0}^i u^0$ , and the expression (4.64) for the Christoffel symbols. This is known as the *continuity equation*: it expresses the conservation of energy in an expanding universe. You can check that the second constraint (4.55) is trivial when applied to homogeneous and isotropic fluids. (It plays a role when we consider the propagation of sound waves in the universe.)

To make progress, we also need the equation of state. The fluids of interest have rather simple equations of state, taking the form

$$P = w\rho$$

with constant  $w$ . Of particular interest are the cases  $w = 0$ , corresponding to pressureless dust, and  $w = 1/3$  corresponding to radiation.

For a given equation of state, the continuity equation becomes

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}$$

So we learn that the energy density  $\rho$  dilutes as the universe expands, with

$$\rho = \frac{\rho_0}{a^{3(1+w)}} \tag{4.65}$$

with  $\rho_0$  an integration constant. For pressureless dust, we have  $\rho \sim 1/a^3$  which is the expected scaling of energy density with volume. For radiation we have  $\rho \sim 1/a^4$ , which is due to the scaling with volume together with an extra factor from redshift.

Now we can look at the Einstein equations. The temporal component is

$$G_{00} + \Lambda g_{00} = 8\pi G T_{00} \quad \Rightarrow \quad \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{k}{a^2} \tag{4.66}$$

This is the *Friedmann equation*. In conjunction with (4.65), it tells us how the universe expands.

We also have the spatial components of the Einstein equation,

$$\begin{aligned}
 G_{ij} + \Lambda g_{ij} = 8\pi G T_{ij} &\Rightarrow 2\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} - \Lambda = -8\pi G P \\
 &\Rightarrow \frac{\ddot{a}}{a} - \frac{\Lambda}{3} = -\frac{4\pi G}{3}(\rho + 3P)
 \end{aligned}
 \tag{4.67}$$

This is the acceleration equation, also known as the *Raychaudhuri equation*. It is not independent of the Friedmann equation; if you differentiate (4.66) with respect to time, you can derive Raychaudhuri.

There is plenty of physics hiding in these equations. Some particularly simple solutions can be found by setting  $k = \Lambda = 0$  and looking at a universe dominated by a single fluid with energy density scaling as (4.65). The Friedmann equation becomes

$$\left(\frac{\dot{a}}{a}\right)^2 \sim \frac{1}{a^{3(1+w)}} \Rightarrow a(t) = \left(\frac{t}{t_0}\right)^{2/(3+3w)}$$

Picking  $w = 1/3$  we have  $a(t) \sim t^{1/2}$  which describes the expansion of our universe when it was dominated by radiation (roughly the first 50,000 years). Picking  $w = 0$  we have  $a(t) \sim t^{2/3}$  which describes the expansion of our universe when it was dominated by matter (roughly the following 10 billion years). You can find many more solutions of the Friedmann equations and a discussion of the relevant physics in the lectures on [Cosmology](#).