A Second Relativity Course

Ingemar Bengtsson (2017)

Contents

Preface

The first two chapters here are intentionally very sketchy. The remaining chapters will depend to some extent on the reaction to the two first, and will be written as the lectures proceed. This is to say that the aim of the first two lectures is that I should learn something about the audience.

You will find some books and papers recommended in footnotes. It is probably a good idea to have a look. Two books stand out as so useful that they need no recommendation, namely

R. M. Wald: *General Relativity*, Chicago UP, 1984.

T. W. Baumgarte and S. L. Shapiro: *Numerical Relativity*, Cambridge UP, 2010.

Many books can be recommended, including J. L. Synge: *Relativity: The General Theory* (brilliant), S. W. Hawking and G. F. R. Ellis: *The Large Scale Structure of Space-Time* (of course), M. Ludvigsen: *General Relativity: A Geometric Approach* (life made easy), and so on.

My hope is that I can 'examine' the students using hand-in exercises. They will be concerned with points that I do not discuss well enough in my lectures. I expect the solutions to be written up with some care, so that I can hand them out as supplementary reading.

I will be using a mixture of different notations. I don't apologize for this. On the contrary, that's life. All my conventions are consistent.

PS: At the end, let me admit that using a mixture of different notations was a bad idea. Due to lack of time, I never lectured on Chapters 6, 10, or 11, so I don't know if they make sense. Also, Chapters 8 and 9 are much too brief. Still, at least one person (me) enjoyed the course!

1 A confusing introduction

At great expense, a number of giant interferometers have been built recently. They have arm lengths of up to 4 kilometers. At the end of the arms, mirrors are attached to masses that are suspended from the roof so they can move freely in the horizontal plane (as long as they do not move too far in any direction). When a gravitational wave hits such an interferometer, the lengths of the arms are affected differently, and the resulting change in the interference pattern is observable.

There are a number of problems with this. For instance, in Hanford, Washington, scientists have to ask themselves how the interference pattern will be affected, should one of the arms be hit by a tumbleweed. Moreover, in that part of the US, people tend to shoot at anything they do not recognize, and special measures must be taken to deal with this difficulty. But what we would like to centre on now is a difficulty often raised by some gentleman in the back row whenever this project is presented to the public: since gravity affects everything equally, will not the gravitational wave change the length of the laser wave train in such a way that the change in the length of the arm is neutralized? Why did the detector detect anything at all, besides tumbleweed?

As theorists, we can investigate this question by hitting the detector with an exact solution of Einstein's equations, describing a plane gravitational wave oriented in such a way that the interferometer will perform in an optimal manner. That is to say, let the arms protrude along the x and y axes, and choose the polarisation of the plane wave so that

$$ds^2 = -dt^2 + dz^2 + (1 + (t - z)\Theta(t - z))^2 dx^2 + (1 - (t - z)\Theta(t - z))^2 dy^2 \ , \quad (1.1)$$

where Θ denotes Heaviside's step function. This description of a plane wave is due to Rosen, and we are looking at a particular kind of wave, called *impulsive*, where the first derivative of the metric is discontinuous. Spacetime is flat up to the null plane described by $t = z$, and then something happens. Despite appearances, it is flat on the other side of that null plane too. To see this, perform the coordinate change

$$X = (1+u)x \qquad Y = (1-u)y \qquad U = u \qquad V = v + (1+u)x^2 - (1-u)y^2 \quad (1.2)$$

where $u = t - z$ and $v = t + z$ are standard null coordinates in Minkowski space. After the wave front has passed, the metric then takes the manifestly flat form

$$ds^2 = -dUdV + dX^2 + dY^2 \ . \tag{1.3}$$

The way the two flat half spaces are glued together across the wave front—where the curvature will have a delta function spike—remains non-trivial.

Now let us think what the suspended masses at the ends of the arms are doing. Since they are freely movable in the horizontal plane they are freely falling as far as our analysis is concerned, so we can regard them as moving along geodesics. It is easy to check that the curves defined by $(x, y, z) = (a, 0, 0)$ and $(x, y, z) = (0, a, 0)$ are geodesics, and we can set the constant a to 4 km. The detector follows the geodesic $(x, y, z) = (0, 0, 0)$. Before the wave these three geodesics are at constant distance from each other, but after the wave has passed they are in relative motion as seen in the inertial coordinate system. And this relative motion (in Minkowski space!) is certainly detectable through the Doppler shift of signals exchanged between them. In the original Rosen coordinates this will appear in a way that is formally similar to the redshift familiar in Friedmann cosmology, that is the wave length at detection is related to the wavelength at emission through a scale factor that can be read off from the metric. In fact the calculations are identical, because we have the spatial Killing vector fields ∂_x and ∂_y available. If we assume that the signals are exchanged along the x–axis, and denote the Killing vector field by ξ^a, we get

$$\frac{\lambda_{\rm det}}{\lambda_{\rm em}} = \frac{\sqrt{\xi^2_{\rm det}}}{\sqrt{\xi^2_{\rm em}}} = \frac{1 + u_{\rm det}\Theta(u_{\rm det})}{1 + u_{\rm em}\Theta(u_{\rm em})} \ . \tag{1.4}$$

The conclusion is that the passage of the wave is observable.

But there is a moral to be drawn. The passage of the wave is observable only because signals are exchanged between the various test particles concerned. What is actually observed is the colour of light when it arrives at a detector. Or, as Einstein put it:[1]

All our space-time verifications invariably amount to a determination of space-time coincidences.

This is interestingly different from quantum theory, where conceptual difficulties are usually met by stressing that predictions of the theory concern the determination of measured data given a prepared state. But this is by the way. Einstein said many things in that 1916 review article of his. Here is one statement:

... this requirement of general co-variance, which takes away from space and time the last remnant of physical objectivity.

In modern parlance, this is a statement about the gauge group of relativity theory. It can be translated into: In theories where diffeomorphisms act as gauge transformations the points of the manifold are not observables—where

[1] A. Einstein, *Die Grundlage der allgemeinem Relativitätstheorie*, Ann. Phys. **49** 769 (1916).

"observable" carries the technical meaning that the word has in gauge theories. There is an argument known as Einstein's hole argument that leads to this conclusion in a very direct way. It will be worth our while to state this argument carefully, beginning with electrodynamics where the gauge group is much easier to understand. Thus, consider a "hole" in space-time. To the past and to the future of the hole, the electromagnetic field A_a as well as the current density J^a take some prescribed form. For simplicity, assume that the current vanishes inside the hole. In spite of this the situation within the hole is not determined: We can find a function Λ which is non-zero only inside the hole. Given any vector potential A_a within the hole consistent with the situation prescribed outside, we will then find that the vector potential $A_a + \partial_a \Lambda$ will be consistent with that information too.

Determinism seems to be at stake here. In electrodynamics the resolution is that only gauge invariant quantities count as observables in the theory. In electrodynamics the gauge invariant quantities are easily identified, and include the electric and magnetic fields as well as integrals of the vector potential around closed loops. Once $A_a(x)$ and $J^a(x)$ are prescribed outside the hole, the observables are determined inside it, and the apparent difficulty vanishes. There are various ways of making this work in the context of an initial value formulation. For instance, one can show that the gauge ambiguity allows us to impose the Lorenz gauge $\partial \cdot A = 0$ on the initial data. One then replaces Maxwell's equations with a new set,

$$\Box A_a - \partial_a \partial \cdot A = J_a \quad \rightarrow \quad \Box A_a = J_a \ . \tag{1.5}$$

The second set leads to a well-behaved initial value problem, and the matter is clinched by showing that the new equations imply that $\Box \partial \cdot A = 0$, so that the Lorenz gauge is preserved by the time evolution.

We are now ready to face Einstein's hole argument for gravity. In relativity theory a space-time is defined by a triplet $(\mathcal{M}, g, \phi)$, where $\mathcal{M}$ is a four dimensional manifold, g_{ab} is a Lorentzian metric, and ϕ stands for some collection of matter fields. It is assumed that Einstein's equation hold. We will adopt the modern "holographic" way of phrasing the problem, that is to say we ask whether Einstein's equations admit an initial data formulation, so that the future is determined by some kind of initial conditions on a spacelike hypersurface in spacetime.

Consider a diffeomorphism $\Phi : \mathcal{M} \rightarrow \mathcal{M}$, that is to say a map of the manifold into itself with suitable differentiability properties. Then the triplet $(\mathcal{M}, \Phi_* g, \phi \circ \Phi^{-1})$ is a spacetime too, that is to say it obeys Einstein's equations. (The funny notation will reappear in Figure 2.1, which comes later.) The entire field configuration can be moved around within $\mathcal{M}$ without disturbing the equations in any way. We choose the diffeomorphism so that it equals the identity on the initial data hypersurface, but we can choose it freely to the future. It would seem to follow, immediately, that the future cannot be predicted from the initial data!

It used to be argued that the Earth moves around the Sun, along an orbit

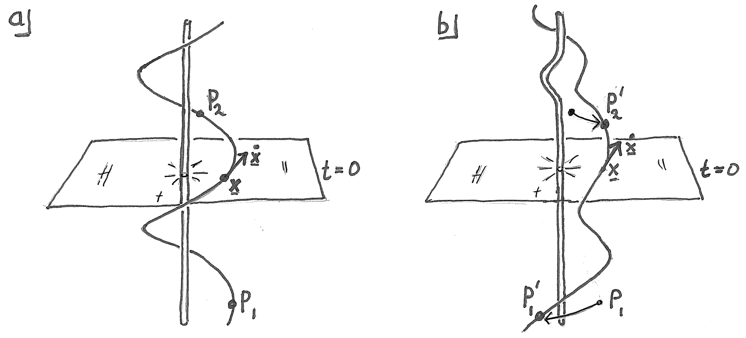

Figure 1.1. a) In Newtonian theory initial data at $t = 0$ ensure that the Earth goes around the Sun. b) In Einstein's theory initial data at $t = 0$ do no such thing, because we can perform diffeomorphisms to the past and to the future of $t = 0$ (taking the point P_1 to a point P_1' of our choosing).

that can be precisely predicted once the initial conditions of the two bodies are known. But we now see that Einstein's theory cannot agree with this, because we can choose the diffeomorphism so that the Sun is unaffected, while the future orbit of the Earth is made to coincide with any arbitrary timelike curve. As Eddington said, the Earth goes where it wants to go.[2] This is disconcerting, but on closer scrutiny not a disaster. We will still be able to define a notion of proper time on the Earth and on the Sun, and we will be able to determine the distances between events happening on these two heavenly bodies. Thus we can lay down a network of observable relations between events in spacetime, and there is hope that these relations will be determined by the initial conditions. The GPS system provides quite convincing evidence that this is true.

What the argument does is to deny any physical significance of the manifold $\mathcal{M}$, and the points of $\mathcal{M}$, as such. Once a metric tensor on $\mathcal{M}$ has been specified we can begin to define observables. This explains the Einstein quote given above. Einstein went on to say:[3]

> If we imagine the gravitational field to be removed, there does not remain a space, but absolutely *nothing*.

His arguments have often been misunderstood, because the distinction between passive and active coordinate transformations has not been drawn. Yet they are very different operations. The confusion comes about because a set of functions $x'(x)$ can be used to describe them both. The thing to remember is that a coordinate system is a map from $\mathcal{M}$ to $\mathbf{R}^4$. In a passive coordinate transformation, called simply "coordinate transformation" from now on, we are changing that map so that the coordinates used to describe the point P are changed from x to x'. In an active coordinate transformation, called a *diffeomorphism* from now on, we are mapping the point P, whose coordinates under the map to coordinate space are x, to the point P', whose coordinates

[2] He said this in *The Nature of the Physical World*, Cambridge 1928. For a version of the hole argument closer to Einstein's original, see C. Rovelli, Class. Quant. Grav. **8** 297 (1991).

[3] A. Einstein, *Relativity and the problem of space*, appendix to *Relativity*, Methuen, London 1954.

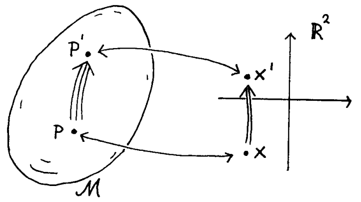

Figure 1.2. A coordinate system is a one-to-one map from (a region of) the manifold to (a region of) $\mathbf{R}^N$. The double arrow is a passive coordinate transformation, the triple arrow a diffeomorphism. Both can be described by the functions $x'(x)$.

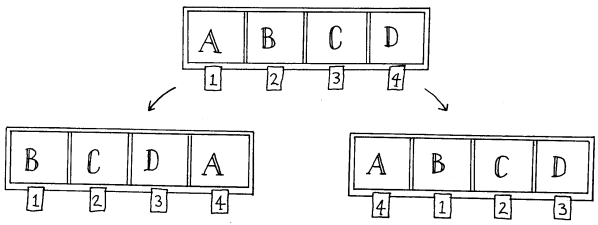

Figure 1.3. Four objects in labeled boxes. We can change ("actively") the location of the objects, or ("passively") their labels.

are x'. In the process the coordinate system is kept fixed. These are very different operations. To clarify the distinction further, think of four objects A, B, C, and D, placed in four boxes labelled 1, 2, 3, and 4. We can describe this situation by means of a function $x(A) = 1$, $x(B) = 2$, $x(C) = 3$, $x(D) = 4$. We can change this function into, say, $x'(A) = 4$, $x'(B) = 1$, $x'(C) = 2$, $x'(D) = 3$. A little reflection shows that there are two very different meanings that one can give to this change in the function, depending on whether one moves the objects or relabels the boxes. Both of these interpretations are respectable and useful, but it is important not to confuse them. Coordinate transformations can be applied to any theory. Invariance under diffeomorphisms is a deep property of some theories.

After this preamble, we can state Einstein's equations in full, in the way he first wrote them down. Given a metric tensor g_{ab}, we define the Christoffel symbols

$$\Gamma_{ab}{}^{c} = \frac{1}{2} g^{cd}(g_{da,b} + g_{db,a} - g_{ab,d}) \ , \tag{1.6}$$

the Riemann curvature tensor

$$R_{abc}{}^{d} = \partial_b \Gamma_{ac}{}^{d} - \partial_a \Gamma_{bc}{}^{d} + \Gamma_{ac}{}^{e} \Gamma_{be}{}^{d} - \Gamma_{bc}{}^{e} \Gamma_{ae}{}^{d} \ , \tag{1.7}$$

the Ricci tensor $R_{ab} = R_{acb}{}^{c}$, and the curvature scalar $R = g^{ab} R_{ab}$. Then Einstein's equations are

$$R_{ab} - \frac{1}{2}Rg_{ab} + \lambda g_{ab} = 8\pi\kappa T_{ab} \ , \tag{1.8}$$

where the right hand side is constructed out of the metric and the matter fields (if any). The arbitrary *cosmological constant* λ provides a scale for the left hand side. If both T_{ab} and λ are set to zero one finds that the equations are scale invariant, in the sense that ag_{ab} is a solution whenever g_{ab} is, for any real number a. As Eddington said, "setting the cosmical constant to zero would knock the bottom out of space". The constant $\kappa = G/c^4$ is usually set to 1 by means of a choice of units.

Considered as a coupled non-linear system of partial differential equations (PDEs) for the ten indendent functions that make up the metric tensor, this looks—to say the least—complicated. However, after a hundred years of close scrutiny, we have learned that the system has very special properties. Indeed, more is known about the long-term evolution of its solutions than what is known for the superficially much simpler Navier–Stokes equations. Moreover, Einstein's equations have given rise to amazing discoveries about things whose existence was unheard of when the equations were first written down, such as black holes. They have also given us a consistent framework for discussing the evolution of the Universe as a whole—although it has to be admitted that, at present, the correctness of the equations seems to hang on the existence of undiscovered "dark matter". Hence there are challenges too. There is probably no reason, at the moment, to think that relativity is "essentially understood".[4]

◇ **Problem 1.1** Calculate, in the Rosen coordinates u, v, x, y, the Ricci tensor of the metric (1.1). Convince yourself that there is curvature on the hyperplane where the derivative of the metric is discontinuous, but that Einstein's equations do hold.

◇ **Problem 1.2** Verify Eq. (1.4) in full detail.

[4] Experts do not think so: H. Friedrich, *Is general relativity essentially understood?*, Annalen Phys. **15** (2005) 84.

2 Manifolds, tensors, and coordinates

We have yet to define "manifolds" and "tensors". To make sure that these lectures get anywhere in the available time, I will be very sketchy about this. It is anyway the case that the definitions are best appreciated by working out examples, taking note of conceptual points as they arise. Thus I will take *manifolds* and *scalar fields* (real valued functions on the manifold) as given, and only remark that a modern way of defining a differentiable manifold starts by postulating the set (or "ring") of scalar fields one wishes to consider. In this way one prepares the ground for generalization to "non-commutative geometry", and this may turn out to become important in the future.[1]

We now focus on a point P of the manifold, and consider the set of all curves that are passing through it. At P each curve has a tangent vector, and the set of these tangent vectors form a vector space $\mathbf{T}$ called the *tangent space* at P. Linear algebra then allows us to construct the dual vector space $\mathbf{U}$ consisting of all linear maps from $\mathbf{T}$ to the real numbers, as well as—by taking tensor products of these two basic vector spaces—the vector spaces $\mathbf{T} \otimes \mathbf{T} \otimes \cdots \otimes \mathbf{T}$, $\mathbf{U} \otimes \mathbf{U} \otimes \cdots \otimes \mathbf{U}$, and mixed products such as $\mathbf{T} \otimes \mathbf{U}$. Elements of these vector spaces are called *tensors at a point*.

We must now deal with notation. Physicists usually proceed with indices (and, in the past, meant the indices to label the tensor components), while mathematicians usually prefer an index-free notation (because they do not want to commit themselves to any special coordinate system). A golden middle road is available. According to Penrose's abstract index notation the elements of these vector spaces are denoted

$$V^a \in \mathbf{T} \ , \quad U_a \in \mathbf{U} \ , \quad V^{ab} \in \mathbf{T} \otimes \mathbf{T} \ , \quad V^a_{\ b} \in \mathbf{T} \otimes \mathbf{U} \ , \tag{2.1}$$

and so on. The map from $\mathbf{V}$ to $\mathbf{R}$ provided by a vector in $\mathbf{U}$ is denoted $U_a V^a \in \mathbf{R}$. It has to be a linear map, in the sense that

$$U_a(fV_1^a + gV_2^a) = fU_aV_1^a + gU_aV_2^a \ . \tag{2.2}$$

This is the condition ensuring that U_a is a vector. At a point, f and g are

[1] I strongly recommend you to read the early parts of R. Penrose, *Structure of space-time*, in C. M DeWitt and J. A. Wheeler (eds.): Batelle Rencontres, Benjamin, New York 1967. See also chapter 4 of R. Penrose and W. Rindler: *Spinors and space-time I*, Cambridge U. P. 1984.

arbitrary numbers. More generally they are scalar fields, and then this condition may become non-trivial to check. We adopt the convention that round brackets around a set of indices denote symmetrization, and square brackets denote antisymmetrization. Thus, for a tensor that belongs to the symmetric subspace of $\mathbf{U} \otimes \mathbf{U}$ we can write

$$g_{ab} = g_{(ab)} \in (\mathbf{U} \otimes \mathbf{U})_{\text{sym}} \ . \tag{2.3}$$

Antisymmetric tensors, and complicated objects with index symmetries like

$$R_{abcd} = R_{[ab][cd]} = R_{[cd][ab]} \ , \qquad R_{[abc]d} = 0 \ , \tag{2.4}$$

are defined similarly.

It is important to realise that the abstract indices are not supposed to take one out of four possible values, and the index contraction does not denote a sum—the indices just serve as markers telling us to which vector space the object they are attached to belongs.

But we will need concrete indices also. For this purpose we introduce a set of four basis vectors in $\mathbf{T}$, and can then expand any vector in $\mathbf{T}$ according to

$$V^a = V^{\mathbf{a}} e_{\mathbf{a}}^a \ . \tag{2.5}$$

Thus $V^{\mathbf{a}}$ denotes the four components of the vector V^a, when expanded in the given basis. We also choose a dual basis in $\mathbf{U}$. It is dual in the sense that

$$e_{\mathbf{a}}^a e_a^{\mathbf{b}} = \delta_{\mathbf{a}}^{\mathbf{b}} \ , \tag{2.6}$$

and it is assumed that

$$e_{\mathbf{a}}^a e_b^{\mathbf{a}} = \delta_b^a \in \mathbf{T} \otimes \mathbf{U} \ . \tag{2.7}$$

Einstein's summation convention is in force for concrete indices.

Using these conventions we find that

$$U_a V^a = U_{\mathbf{a}} e_a^{\mathbf{a}} V^{\mathbf{b}} e_{\mathbf{b}}^a = U_{\mathbf{a}} V^{\mathbf{b}} e_a^{\mathbf{a}} e_{\mathbf{b}}^a = U_{\mathbf{a}} V^{\mathbf{b}} \delta_{\mathbf{a}}^{\mathbf{b}} = U_{\mathbf{a}} V^{\mathbf{a}} \ . \tag{2.8}$$

It may be a bit difficult to see the point, but the distinction between abstract indices—labels on an object that tell us to which vector space it belongs—and concrete indices—labels of the components of the object when it is expanded in a basis of its vector space—is important in the sense that it takes care of a conceptual point. I may not uphold the distinction between abstract and concrete indices all that strictly, but one should try to interpret any equation in terms of abstract indices whenever possible.

So far we have been sitting at a particular point in spacetime. But we want tensor fields, defined all over the manifold. The important thing is to define the *vector fields*. Everything else will follow. A standard way is to consider all possible curves passing through a point. Each has a tangent vector, obtained by differentiating along the curve, and once these tangent vectors have been divided into suitable equivalence classes we obtain the tangent space at the point. Since derivatives have entered the game, this leads to the definition of vector fields $\vec{V}$ as maps taking scalar fields to scalar fields. If k is a real number and f, g, are scalar fields, the defining rules are that

(i) $\vec{V}(k) = 0$
(ii) $\vec{V}(f+g) = \vec{V}(f) + \vec{V}(g)$
(iii) $\vec{V}(fg) = f\vec{V}(g) + g\vec{V}(f)$.

The last requirement is Leibniz' rules for derivatives, and indeed once a coordinate system is given the vector fields so defined are differential operators,

$$\vec{V}(f) = V^{\mathbf{a}}\partial_{\mathbf{a}} f \ . \tag{2.9}$$

This provides us with a convenient set of basis vectors for the tangent space $\mathbf{T}$, namely the differential operators $\partial_{\mathbf{a}}$. In abstract index notation these vector fields are denoted $\partial^a_{\mathbf{a}}$, and the components of the metric tensor in this basis are

$$g_{\mathbf{ab}} = g_{ab}\partial^a_{\mathbf{a}}\partial^b_{\mathbf{b}} \ . \tag{2.10}$$

A dual basis in the cotangent space consists of four differentials $dx^{\mathbf{a}}$, defined by

$$dx^{\mathbf{a}}(\partial_{\mathbf{b}}) = \delta^{\mathbf{a}}_{\mathbf{b}} \ . \tag{2.11}$$

If x is one of the coordinates (or any function on the manifold for that matter) we can use abstract index notation to write

$$dx_a = \nabla_a x \ , \tag{2.12}$$

where ∇_a is a derivative operator (such that $\nabla_{\mathbf{a}} f = \partial_{\mathbf{a}} f$ for all scalar functions f).

Once it is realized that vector fields are differential operators, the question whether they commute raises its head. In general they do not. Rather

$$[X, Y]^{\mathbf{a}} = X^{\mathbf{b}}\partial_{\mathbf{b}} Y^{\mathbf{a}} - Y^{\mathbf{b}}\partial_{\mathbf{b}} X^{\mathbf{a}} = \mathcal{L}_{\vec{X}} Y^{\mathbf{a}} \ , \tag{2.13}$$

where the *Lie derivative* of the vector field $\vec{Y}$ along the vector field $\vec{X}$ occurs on the right hand side. Given a vector field, we can always choose a coordinate that runs along it, so that (for instance)

$$X^a\partial_a = \partial_x \ , \tag{2.14}$$

where x is the coordinate that runs along the vector field $\vec{X}$. Given two vector fields, we can introduce two coordinates running along them if and only if they commute. In general the commutator of two vector fields is a non-vanishing vector field in itself. A key question one can ask about it is whether it belongs to the linear span of the two vector fields from which it was formed. If this is the case the two vector fields are said to be *surface forming.*[2]

The coordinate system can be changed. It used to be that tensors were defined directly in the coordinate basis, and by the way they transform under coordinate changes.[3] You can easily check that all the transformation rules come out correctly from the definitions we just gave. Thus, if we perform the coordinate transformation $x \to x'(x)$ we find that

$$V^{\mathbf{a}}\partial_{\mathbf{a}} = V^{\mathbf{b}}\frac{\partial x^{\mathbf{a}'}}{\partial x^{\mathbf{b}}}\partial_{\mathbf{a}'} \quad \Rightarrow \quad V^{\mathbf{a}'}(x') = \frac{\partial x^{\mathbf{a}'}}{\partial x^{\mathbf{b}}}V^{\mathbf{b}}(x) \ . \tag{2.15}$$

If we reinterpret this as an active coordinate transformation we see something very important, namely that contravariant tensors (with indices 'upstairs') are *pushed forward* by the map, while a similar calculation for covariant tensors shows them to be *pulled back*. Of course, diffeormorphisms are invertible, and if a metric is available this distinction can become very blurred. But it will be important in Chapter 7.

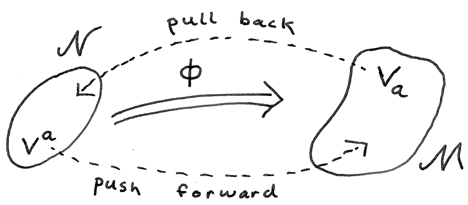

Figure 2.1. Mapping, say of $\mathcal{N}$ into $\mathcal{M}$, pullback, and pushforward. A metric g on $\mathcal{M}$ is pulled back to the metric $\Phi_* g$ on $\mathcal{N}$ by the map Φ.

In coordinate free language, let the map $\mathcal{N} \to \mathcal{M}$ be called Φ. If is clear that a function $f : \mathcal{M} \to \mathbf{R}$ will be "pulled back" by the map to give a function $f \circ \Phi : \mathcal{N} \to \mathbf{R}$. But from the abstract definition of a vector field $\vec{V}$ on $\mathcal{N}$ it is clear that this will give a vector field $\Phi^*(V^a)$ on $\mathcal{M}$, defined by

$$\Phi^*(\vec{V})(f) = \vec{V}(f \circ \Phi) \ . \tag{2.16}$$

The vector field has been pushed forward in the direction of the map Φ. Given that, a covariant vector field U_a on $\mathcal{M}$ will then give rise to a covariant vector field $\Phi_*(U_a)$ on $\mathcal{N}$, defined by how it acts on an arbitrary vector field V^a by

$$\Phi_*(U_a)V^a = U_a\Phi^*(V^a) \ . \tag{2.17}$$

The covariant vector field has been pulled back against the direction of the

[2] Here lurks a theorem due to Frobenius. See Chapter 3, or look it up in Wald's appendices.
[3] A particularly clear exposition is E. Schrödinger: *Space-Time Structure*, Cambridge U. P., 1950.

map. To remember this, observe that the usual terminology, "co"- and "contra"-variant vectors, is maximally misleading.

Metric tensors are usually given, in terms of coordinates, as quadratic forms

$$ds^2 = g_{\mathbf{ab}} dx^{\mathbf{a}} dx^{\mathbf{b}} \ . \tag{2.18}$$

Here the notation is completely different from the one we used so far. What is being meant is that the length squared of the tangent vector ds, at the point whose coordinates are $x^{\mathbf{a}}$, are given as a quadratic form in its components $dx^{\mathbf{a}}$. For some reason, this old-fashioned piece of notation has survived intact. If you do not like it, you can write—using abstract indices—

$$ds^2 = dx^2 + dy^2 \qquad \rightarrow \qquad g_{ab} = \nabla_a x \nabla_b x + \nabla_a y \nabla_b y \ . \tag{2.19}$$

In the second formulation no coordinates have been chosen, but it is assumed that we have agreed on a definition of the two scalar fields x and y.[4]

It remains to define connections, covariant derivatives, and curvature. *Covariant derivatives* ∇_a are best defined as maps from $\mathbf{V}$ to $\mathbf{U} \otimes \mathbf{V}$, subject to certain rules. First one declares that $\nabla_a f$ is the usual gradient of the function f. Then one insists that

$$\nabla_a (X^b + Y^b) = \nabla_a X^a + \nabla_a Y^b \quad \text{and} \quad \nabla_a (f X^b) = \nabla_a f X^a + f \nabla_a X^b \ . \tag{2.20}$$

Finally the action of ∇_a on arbitrary tensors is defined using Leibniz' rule. These rules do not tie down the covariant derivative uniquely however. To do so we impose the conditions that

$$\begin{aligned} &[\nabla_a, \nabla_b] f = 0 \qquad && \text{no torsion} \\ &\nabla_a g_{bc} = 0 && \text{metric compatible .} \end{aligned} \tag{2.21}$$

The resulting covariant derivative defines the *Levi-Civita connection*, and I assume that you know how to express it in terms of Christoffel symbols and so on. The *Riemann tensor* is best defined by the *Ricci identity*

$$[\nabla_a, \nabla_b] V_c = R_{abc}{}^d V_d \ . \tag{2.22}$$

This is also a convenient way of stating one's conventions for the Riemann tensor. See Problem 2.1.

We now look at coordinates in more detail. A coordinate is a real valued function defined on a manifold. In fact it is a scalar field. A *coordinate chart* on a D-dimensional manifold is a map from the manifold into $\mathbf{R}^D$ which is locally one-to-one, so we need four such functions, suitably chosen, to coordinatize spacetime. Often one needs an entire *atlas* of partially overlapping charts in order to cover the whole manifold.

[4] When I was young, George Sudarshan told me that students must get used to confusion at an early stage. So I feel free to use whatever notation I find convenient at the moment.

Choosing coordinates is an art, in which one tries to adapt them to some structure, or problem, one is interested in. The usual Cartesian coordinates (x, y) on the flat Euclidean plane provide an example. On the one hand x and y are affine parameters along two plane-filling geodetic congruences, on the other hand the coordinate vectors ∂_x and ∂_y form Killing vector fields. These coordinates are adapted both to the geodetic structure and to the symmetries of the plane.

A simple observation is that the vector field ∂_x points along the lines of constant y. Now consider the coordinate transformation

$$X = x \ , \qquad Y = x + y \ . \tag{2.23}$$

The x-coordinate has been left unchanged but the y-coordinate has changed. As a result, ∂_x changes but ∂_y remains the same:

$$\partial_X = \partial_x - \partial_y \ , \qquad \partial_Y = \partial_y \ . \tag{2.24}$$

I bring this up because it may look confusing at first sight.

Another useful coordinate system is obtained by letting one coordinate ϕ run along the Killing vector field which describes rotation around the origin, and another coordinate r run along geodesics emerging from there. This is called a geodetic polar coordinate system. We know that the metric becomes

$$ds^2 = dx^2 + dy^2 = dr^2 + r^2 d\phi^2 \ . \tag{2.25}$$

The polar coordinate system breaks down at the origin because the Killing vector field has a fixed point there, and the geodesics that are supposed to form coordinate lines meet. Playing the same game in 1+1 dimensional Minkowski space we find

$$\begin{cases} t = \tau \cosh \sigma \\ x = \tau \sinh \sigma \end{cases} \quad \Rightarrow \quad \partial_\sigma = x\partial_t + t\partial_x \tag{2.26}$$

$$ds^2 = -dt^2 + dx^2 = -d\tau^2 + \tau^2 d\sigma^2 \ . \tag{2.27}$$

This is called a Rindler coordinate system, and it has been adapted to a Lorentz boost around the origin. The coordinate τ measures the proper time from the origin to the Killing orbit it labels. We tacitly assumed that $|t| > |x|$. Another set of Rindler coordinates must be used to cover the remaining two quadrants of Minkowski space. The coordinate singularity at the origin of the polar coordinate system has grown to a coordinate singularity along the null cone emanating there. This had to be so because all points on that cone are at the same distance (namely zero) from the origin, so the coordinate τ fails there.

If such coordinates are introduced in 3 + 1 dimensions they lead to a Robertson-Walker metric known as the metric of the Milne universe, because the astrophysicist Milne thought that the real universe has to be like this. The habit of thinking that the Universe has to be in a certain way has led to many mistakes, but it has scored some successes too.

A coordinate problem also occurs with the Rosen coordinates we used to describe the plane wave (1.1). A linearly polarized plane wave is given in Rosen coordinates by

$$ds^2 = -2dudv + F^2dx^2 + G^2dy^2 \ , \qquad F = F(u) \ , \quad G = G(u) \ . \tag{2.28}$$

The Ricci tensor has only one non-vanishing component, and the metric is a solution of Einstein's equations if and only if

$$R_{uu} = -\frac{F_{,uu}}{F} - \frac{G_{,uu}}{G} = 0 \ . \tag{2.29}$$

The coordinate lines $u = \tau$, with v, x, y held constant, are null geodesics. They can meet if F or G vanish, and then the Rosen coordinates fail. The coordinate transformation (1.2) resolved this difficulty in a special case, and can easily be generalized. First we set

$$x = \frac{X}{F} \ , \quad y = \frac{Y}{G} \ , \quad u = U \ . \tag{2.30}$$

This will give rise to cross-terms in the metric that we prefer to avoid, so we follow it up with

$$v = V - \frac{1}{2}\frac{F_{,u}}{F}X^2 - \frac{1}{2}\frac{G_{,u}}{G}Y^2 \ . \tag{2.31}$$

The result is

$$ds^2 = -2dUdV + \left(\frac{F_{,uu}}{F}X^2 + \frac{G_{,uu}}{G}Y^2\right)dU^2 + dX^2 + dY^2 \ . \tag{2.32}$$

The vector field ∂_U is not a null vector field, but the Killing vector field ∂_V is.

The coordinates we have arrived at are known as Brinkmann coordinates. The most general plane wave spacetime is given, in terms of them and a two-by-two matrix A_{ij}, by

$$ds^2 = -2dUdV + A_{ij}x^ix^jdU^2 + dX^2 + dY^2 \ , \qquad A_{ij} = A_{ij}(U) \ . \tag{2.33}$$

The indices i, j run from 1 to 2, and $x^1 = X$, $x^2 = Y$. A calculation shows that the only non-vanishing components of the Riemann tensor are given by

$$R_{UiUj} = A_{ij} \ . \tag{2.34}$$

The Ricci tensor vanishes if the matrix A_{ij} is traceless, so the most general vacuum plane wave spacetime is given by

$$ds^2 = -2dUdV + \big(A(U)(X^2 - Y^2) + 2B(U)XY\big)\, dU^2 + dX^2 + dY^2 \ . \quad (2.35)$$

The polarization is described by two degrees of freedom, just as one would expect from the linearized approximation.

An interesting feature of the Brinkmann coordinates is that they obey

$$\Gamma^a \equiv g^{bc}\Gamma_{bc}{}^{a} = 0 \ . \quad (2.36)$$

This is the defining property of a *harmonic coordinate system*. Using

$$\Gamma_{ac}{}^{c} = \frac{\sqrt{-g}_{,a}}{\sqrt{-g}} \quad (2.37)$$

it can be rewritten as

$$\partial_c(\sqrt{-g}g^{ac}) = 0 \ . \quad (2.38)$$

It can also be written, for each of the scalar fields $x^{\mathbf{a}}$ that serve as coordinates,

$$\Box_g x^{\mathbf{a}} = g^{bc}\nabla_b\nabla_c x^{\mathbf{a}} = -\Gamma^{\mathbf{a}} = 0 \ . \quad (2.39)$$

This shows that, locally, such coordinates can always be found by solving the wave equation. It also explains why these coordinates are called “harmonic”, although “wave coordinates” would be a better name. The reason why this is interesting is that the condition $\Gamma^a = 0$ is an analogue of the Lorenz gauge in electrodynamics. If $\Gamma^a = 0$ one finds, for the Ricci tensor,

$$R_{ab} = \partial_c\Gamma_{ab}{}^{c} - \partial_a\Gamma_{bc}{}^{c} + \Gamma_{ab}{}^{e}\Gamma_{ce}{}^{c} - \Gamma_{cb}{}^{e}\Gamma_{ae}{}^{c} = -\frac{1}{2}g^{cd}g_{ab,cd} + Q_{ab}(g, \partial g) \ . \quad (2.40)$$

The point is that all the “mixed” second derivatives of the metric have gone away, and one is in the position to begin to apply general existence theorems about partial differential equations (PDEs) to this system. The analogy to the Lorenz gauge in electrodynamics should be clear, and similar consistency issues have to be dealt with in both cases.[5] Variants of this idea are widely used in numerical relativity. The first breakthrough simulation of inspiral, merger, and ringdown of two black holes used one of them.[6]

We go on to something more special. A spacetime is spherically symmetric if it admits an $SO(3)$ group of isometries. This group will act transitively on round spheres embedded in spacetime, and it can be shown that the metric can always be written on the block diagonal form

$$ds^2 = g_{AB}dx^A dx^B + r^2 d\Omega^2 \ , \quad (2.41)$$

[5] This goes back to Y. Choquet-Bruhat. For her own account of the history, see *Beginnings of the Cauchy problem*, arXiv:1410.3490.

[6] F. Pretorius, *Evolution of binary black-holes spacetimes*, Phys. Rev. Lett. **95** (2005) 121101.

where $d\Omega^2$ is the metric on the round unit sphere and g_{AB} is some 1+1 dimensional Lorentzian metric. The round spheres have radii equal to r. There are spacetimes—such as the Nariai solution of Einstein's equations including a cosmological constant—for which r is constant, but in most cases it is possible to use r as a local coordinate on the 1+1 dimensional part. We do so from now on. Then r is known as the *area coordinate*, since the area of the *surfaces of transitivity* is precisely $4\pi r^2$.

Thus the area coordinate has an invariant meaning, and so has

$$||\nabla r||^2 = g^{ab}\nabla_a r \nabla_b r = g^{rr} \ , \tag{2.42}$$

where the last equality is valid in our adapted coordinate system only. There is a coordinate singularity when the gradient vector field $\nabla_a r$ vanishes. At this point it is conventional to define the *Misner-Sharp mass function* m through

$$||\nabla r||^2 = 1 - \frac{2m}{r} \ . \tag{2.43}$$

It is a function on the 1+1 dimensional quotient space with a clear geometrical meaning. The metric on the quotient space can always be written on diagonal form, and then our metric is

$$ds^2 = -e^{2\beta}\left(1 - \frac{2m}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{2m}{r}} + r^2 d\Omega^2 \ . \tag{2.44}$$

Here $\beta = \beta(t, r)$ and $m = m(t, r)$. The function β can be changed by reparametrizations of t, and unlike m it does not have an invariant meaning.

These spacetimes admit a congruence of radially ingoing null geodesics, and it makes considerable sense to adapt the coordinate system to it. The geodesics obey

$$\frac{dt}{dr} = -\frac{e^{-\beta}}{1 - \frac{2m}{r}} \ . \tag{2.45}$$

We then introduce a coordinate that (together with the angles) labels these geodesics. This coordinate is denoted v, and is referred to as *advanced time.* Thus we perform the coordinate transformation

$$dv = dt + e^{-\beta}\frac{dr}{1 - \frac{2m}{r}} \ . \tag{2.46}$$

Then v is constant along every ingoing radial null geodesic. To solve explicitly for $v = v(t, r)$ we need to specify the arbitrary functions m and β, but we do not need to do this yet. Eq. (2.46) is enough to show that the metric takes the *Eddington-Finkelstein* form

$$ds^2 = -e^{2\beta}\left(1 - \frac{2m}{r}\right)dv^2 + 2e^{\beta}dvdr + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{2.47}$$

It is noticeable that the coordinate singularity at $r = 2m$ has evaporated. A coordinate system using *retarded time*, labelling outgoing null geodesics, is also possible.

If we impose Einstein's vacuum equations, and set the cosmological constant to zero, the only spherically symmetric solution is given by

$$m(v, r) = M = \text{constant} \;, \qquad \beta(v, r) = 0 \;. \tag{2.48}$$

This is the *Schwarzschild solution.*

Quite incidentally, if we now introduce a brand new time coordinate t such that $v = t + r$ we find that the Schwarzschild metric takes the form

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) + \frac{2M}{r}(dt + dr)^2 \;. \tag{2.49}$$

This is the form in which Eddington actually wrote it down.[7] It was later generalized to metrics of the *Kerr-Schild* form

$$g_{ab} = \eta_{ab} + F l_a l_b \;, \tag{2.50}$$

where F is some function of the coordinates, η_{ab} is the Minkowski metric, and l_a is a null vector with respect to both metrics,

$$l^a = g^{ab} l_b = \eta^{ab} l_b \;, \qquad l_a l^a = 0 \;. \tag{2.51}$$

Note that

$$g^{ab} = \eta^{ab} - F l^a l^b \;. \tag{2.52}$$

What is not so obvious, but still true, is that some physically very important metrics that do not have spherical symmetry can be written in this form.

◇ **Problem 2.1** Calculate the Riemann tensor as given by Einstein, (1.7), from the definition (2.22). Do this on three consecutive days, or for however many days it takes until you can do it in less than five minutes.

◇ **Problem 2.2** Derive the claims of equations (2.38) and (2.39). Also prove that the Brinkmann coordinates are harmonic.

[7] A. S. Eddington, *A comparison of Whitehead's and Einstein's formulæ*, Nature **113** (1924) 192.

3 Differential geometry I

Give us the tools, and we will finish the job.

Winston Churchill

This is the first of two chapters in which we will be concerned only with differential geometry. We focus on one-dimensional curves, and on congruences of curves. After a much needed break (in the form of static black holes) we will then devote ourselves to surfaces and hypersurfaces (for instance trapped surfaces, and hypersurfaces on which initial data are set) in part II.

3.1 More about diffeomorphisms

A *curve* is a map from from the real numbers to a manifold $\mathcal{M}$, given in terms of coordinates by

$$[a, b] \rightarrow \mathcal{M} \ : \qquad \tilde{\sigma} \rightarrow x^{\mathbf{a}}(\tilde{\sigma}) \ . \tag{3.1}$$

We insist that there is an everywhere non-vanishing tangent vector

$$t^{\mathbf{a}} = \frac{dx^{\mathbf{a}}}{d\tilde{\sigma}} \neq 0 \ . \tag{3.2}$$

On the other hand the curve is allowed to intersect itself, should it wish to do so. If we change the parametrization, $\tilde{\sigma} \rightarrow \sigma = \sigma(\tilde{\sigma})$, the tangent vector will change according to the chain rule. This provides us with an opportunity to make a canonical choice of parameter, but before coming to this we will discuss the important role that curves are playing as flowlines of diffeomorphisms. For this purpose we assume that we have a *congruence* of curves, one through each point of (a region in) spacetime. A point sitting on a particular curve, and labelled by the coordinates $x^{\mathbf{a}}(\tilde{\sigma}_0)$, is transformed by a diffeomorphism $\Phi_{\tilde{\sigma}}$ to another point on the same curve, labelled $x^{\mathbf{a}}(\tilde{\sigma}_0 + \tilde{\sigma})$. In continuum mechanics this is how flows of matter are described. There one also introduces a set of *Lagrangian coordinates* labeling the individual flowlines. Since the congruence is space-filling the Lagrangian coordinates, together with the parameter along the curves, form a coordinate system for spacetime.

Every (reasonable) one-parameter subgroup $\Phi_{\tilde{\sigma}}$ of the diffeomorphism group gives rise to a congruence of curves. The transformation rules for tensors under diffeomorphisms allows us to define the *Lie derivative* along the flow, at the point P, as

$$\mathcal{L}_{\tilde{t}} V^a = \lim_{\tilde{\sigma}\to 0} \frac{1}{\tilde{\sigma}} \left(V^a_{|P} - \Phi^*_{\tilde{\sigma}}(V^a)_{|P} \right) \ . \tag{3.3}$$

Here we are comparing the vector field existing at a particular point P along the flow with that pushed forward to P from a point further upstream. Professor Arnold calls it the "fisherman's derivative". The flow carries all possible differential geometric objects past the fisherman, who sits at P and differentiates them.

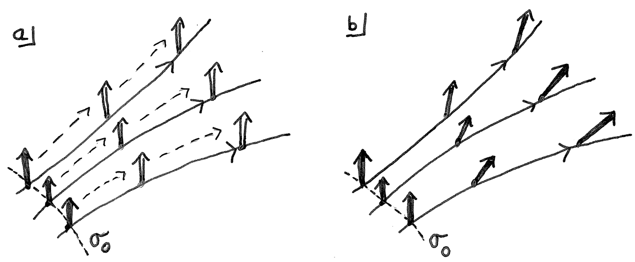

Figure 3.1. Vectors at $\tilde{\sigma} = \tilde{\sigma}_0$ are carried forwards by the flow along the congruence (in panel a). But there may already be a vector field along the congruence (as in panel b). If the two vector fields disagree, the Lie derivative along the tangent vector of the flow is non-zero.

To see what the fisherman catches it is convenient to express the diffeomorphism using a fixed coordinate system,

$$x^{\mathbf{a}} \to x^{\mathbf{a}'} = x^{\mathbf{a}'}(x) \approx x^{\mathbf{a}} + \tilde{\sigma} t^{\mathbf{a}}(x) \ . \tag{3.4}$$

Calculating to first order in $\tilde{\sigma}$ we find that the vector field is pushed forward according to

$$V^{\mathbf{a}'}(x') = \frac{\partial x^{\mathbf{a}'}}{\partial x^{\mathbf{c}}} V^{\mathbf{c}}(x) \approx V^{\mathbf{a}}(x) + \tilde{\sigma}\partial_c t^{\mathbf{a}} V^c(x) \ . \tag{3.5}$$

However, we want the transformed vector field evaluated at the point P, which by assumption has coordinates x, not x'. This is

$$V^{\mathbf{a}'}(x) \approx V^{\mathbf{a}}(x - \tilde{\sigma}t) + \tilde{\sigma}\partial_c t^{\mathbf{a}} V^c(x) \approx V^{\mathbf{a}}(x) - \tilde{\sigma} t^{\mathbf{c}}\partial_{\mathbf{c}} V^{\mathbf{a}} + \tilde{\sigma} V^{\mathbf{c}}\partial_{\mathbf{c}} t^{\mathbf{a}} \ . \tag{3.6}$$

Inserting this in the definition (3.3) we obtain

$$\mathcal{L}_{\vec{t}} V^{\mathbf{a}} = \lim_{\tilde{\sigma}\to 0} \frac{1}{\tilde{\sigma}} \Big(V^{\mathbf{a}}(x) - V^{\mathbf{a}'}(x) \Big) = t^{\mathbf{c}} \partial_{\mathbf{c}} V^{\mathbf{a}} - \partial_{\mathbf{c}} t^{\mathbf{a}} V^{\mathbf{c}} \ . \tag{3.7}$$

In coordinate-free language this is

$$\mathcal{L}_{\vec{t}} V^a = t^c \nabla_c V^a - \nabla_c t^a V^c \ . \tag{3.8}$$

It does not matter which derivative operator we use in this expression. When we replace the coordinate based partial derivative with a torsion free covariant derivative we find that the extra terms cancel because the Christoffel symbols are symmetric.

The Lie derivative of arbitrary tensors can be worked out in the same way, or more conveniently they can be found using Leibniz' rule. For instance

$$\begin{aligned} &\mathcal{L}_{\vec{t}}(U_a V^a) = t^c \nabla_c (U_a V^a) = \\ &= (t^c \nabla_c U_a + \nabla_a t^c U_c) V^a + U_a (t^c \nabla_c V^a - \nabla_c t^a V^c) = \mathcal{L}_{\vec{t}} U_a V^a + U_a \mathcal{L}_{\vec{t}} V^a \ . \end{aligned} \tag{3.9}$$

A particularly important case is the Lie derivative of the metric tensor. One finds

$$\mathcal{L}_{\vec{t}} g_{ab} = t^c \partial_c g_{ab} + \partial_a t^c g_{cb} + \partial_b t^c g_{ac} = \nabla_a t_b + \nabla_b t_a \ , \tag{3.10}$$

where the metric compatible derivative operator ∇_a was used in the second step.

An important point is that we can use the diffeomorphism to define a tensor field, starting from tensors that are defined only at $x^{\mathbf{a}}(\tilde{\sigma}_0)$. We simply insist that $\mathcal{L}_{\vec{t}} V^a = 0$, $\mathcal{L}_{\vec{t}} g_{ab} = 0$, and so on. The tensors are carried along by the flow. If, on the other hand, a tensor field such as g_{ab} is defined already, we can compare it to the tensor field created by the diffeomorphism. If they agree, it means that we have stumbled on a flow of diffeomorphisms such that $\mathcal{L}_{\vec{t}} g_{ab} = 0$. Such diffeomorphisms are called *isometries*, and their tangent vectors $\vec{t}$ are called *Killing vectors*.

Finally, let me explain the unease that some of us feel when it is argued that gravity is "emergent", in the same sense as phonons (say) are emergent in condensed matter physics. There the idea is that one relies on a fixed background metric to divide a system into small boxes. The boxes serve as a cut-off, and excitations with wavelengths that fit into a box are integrated out. Meanwhile, the Milky Way performs its stately rotation, with a period of a hundred million years or so (depending on where you are in the galaxy). But the relativist can perform a diffeomorphism causing the entire Milky Way to flow into one of the small cut-off boxes. The idea to integrate out everything that happens within the Milky Way is not so attractive.

3.2 Curves and congruences

With these matters out of the way, we take up the study of individual curves. We introduce a convenient notation for differentiation along the curve,

$$\dot{v}^a \equiv \nabla_{\vec{t}} v^a \equiv t^b \nabla_b v^a = \frac{dv^a}{d\tilde{\sigma}} + \Gamma_{bc}{}^a t^b v^c \ , \tag{3.11}$$

where $v^a = v^a(\sigma)$ is any vector field defined on the curve. Concerning the last member of the equation, it has probably dawned on you by now that the game is to hide the non-tensorial object $\Gamma_{bc}{}^a$ from sight, whenever we can. It will appear only in explicit coordinate calculations. I should also say that the dot-notation used on the left is unusual. If $\dot{\vec{v}} = 0$ the vector is said to be *parallel transported* along the curve.

Now consider the vector

$$\dot{t}^a = t^b \nabla_b t^a \ . \tag{3.12}$$

We will insist that $\vec{t} \cdot \dot{\vec{t}} = 0$. An exercise using the chain rule shows that this can always be arranged, provided we choose a canonical parameter $\sigma = \sigma(\tilde{\sigma})$ along the curve. It is assumed that we can solve for $\tilde{\sigma}(\sigma)$. Let $\dot{\vec{t}}_{\rm old}$ and $\dot{\vec{t}}_{\rm new}$ be the vector calculated using $\tilde{\sigma}$ and σ, respectively. Then the calculation goes as follows:

$$\dot{t}^{\mathbf{a}}_{\rm old} = \frac{d^2x^{\mathbf{a}}}{d\tilde{\sigma}^2} + \Gamma_{\mathbf{bc}}{}^{\mathbf{a}} \frac{dx^{\mathbf{b}}}{d\tilde{\sigma}} \frac{dx^{\mathbf{c}}}{d\tilde{\sigma}} = \frac{d\sigma}{d\tilde{\sigma}} \frac{d}{d\sigma} \left(\frac{d\sigma}{d\tilde{\sigma}} \frac{dx^{\mathbf{a}}}{d\sigma} \right) + \left(\frac{d\sigma}{d\tilde{\sigma}} \right)^2 \Gamma_{\mathbf{bc}}{}^{\mathbf{a}} \frac{dx^{\mathbf{b}}}{d\sigma} \frac{dx^{\mathbf{c}}}{d\sigma} =$$
$$= \left(\frac{d\sigma}{d\tilde{\sigma}} \right)^2 \left(\frac{d^2x^{\mathbf{a}}}{d\sigma^2} + \Gamma_{\mathbf{bc}}{}^{\mathbf{a}} \frac{dx^{\mathbf{b}}}{d\sigma} \frac{dx^{\mathbf{c}}}{d\sigma} \right) - \frac{\frac{d^2\tilde{\sigma}}{d\sigma^2}}{\left(\frac{d\tilde{\sigma}}{d\sigma} \right)^3} \frac{dx^{\mathbf{a}}}{d\tilde{\sigma}} = \left(\frac{d\sigma}{d\tilde{\sigma}} \right)^2 \dot{t}^a_{\rm new} - \frac{\frac{d^2\tilde{\sigma}}{d\sigma^2}}{\left(\frac{d\tilde{\sigma}}{d\sigma} \right)^3} t^a_{\rm old} \ . \tag{3.13}$$

If $\dot{t}^a_{\rm old}$ has a component in the tangent direction, we can ensure that $\dot{t}^a_{\rm new}$ has none by solving a differential equation. The resulting canonical parameter σ is called an *affine parameter*, and from now on we assume that this choice has been made. The only ambiguity that remains is a choice of scale and zero point of σ. Unless $\vec{t}$ is a null vector the scale is completely fixed by insisting that

$$\vec{t} \cdot \vec{t} = g_{ab} t^a t^b = \begin{cases} 1 & \text{if } \vec{t} \text{ is spacelike} \\ -1 & \text{if } \vec{t} \text{ is timelike} \ . \end{cases} \tag{3.14}$$

The curve is now parametrized using *arc length*. If the curve is timelike the preferred parameter σ is called *proper time*, and usually denoted by τ.

The curve is a *geodesic* if its tangent vector is parallel transported along itself,

$$\dot{t}^a = t^b \nabla_b t^a = 0 \ . \tag{3.15}$$

This is a system of ODEs (ordinary differential equations) that determines the curve uniquely for some interval in σ, given the initial vector $\vec{t}$. If the solution exists for all σ the geodesic is said to be *complete*. (You may know that spacetimes containing incomplete geodesics are said to be 'singular'.)

Now suppose that the curve is given, but suppose the right hand side of the equation is non-zero. Let the curve be parametrized by arclength, and write

$$\dot{t}^a = \kappa_1 n^a \ , \tag{3.16}$$

where the vector $\vec{n}$, known as the *principal normal*, is normalized so that $\vec{n} \cdot \vec{n} = \pm 1$. Actually it can be null, but we ignore this case for the time being. The function κ_1 will be referred to as the *first curvature*, or as the *proper acceleration* if the curve is timelike. If the curve changes from timelike to spacelike then κ_1 diverges, which is forbidden for the worldline of an observer. It is important to know whether $\vec{n}$ is spacelike or timelike. In $2+1$ dimensional Minkowski space, an example of a curve with a spacelike normal is obtained by intersecting two lightcones with vertices on the time axis. A circle of radius r arises this way, and incidentally it has

$$\kappa_1 = \frac{1}{r} \ . \tag{3.17}$$

A spacelike curve with constant curvature and timelike normal can be found by intersecting two lightcones with vertices on a space axis. In the $2+1$ dimensional spacetime diagram it looks like a hyperbola.

To cut down the ambiguities we momentarily assume that we are in a 3-dimensional space with a positive definite metric (so that all vectors are spacelike). Given that $\vec{n} \cdot \vec{n} = 1$ we know that $\dot{\vec{n}}$ is orthogonal to $\vec{n}$. Moreover, since $\vec{t} \cdot \vec{n} = 0$, we find that

$$\dot{n}^a t_a = -n_a \dot{t}^a = -\kappa_1 n_a n^a = -\kappa_1 \ . \tag{3.18}$$

Carrying on in this way we arrive at the *Frenet–Serret equations*

$$\begin{array}{rcccc} \dot{t}^a = & & \kappa_1 n^a & & \\ \dot{n}^a = & -\kappa_1 t^a & & +\kappa_2 b^a & \\ \dot{b}^a = & & -\kappa_2 n^a & & . \end{array} \tag{3.19}$$

The vector $\vec{b}$ is called the binormal, because one has to call it something, and the function κ_2 is the second curvature or *torsion* (since it determines how the curve is 'twisting' in space). Doing the same exercise in four dimensions would result in three curvatures. In Lorentzian cases there will be extra signs, depending on which one of the vectors is timelike—or perhaps either $\vec{t}$ or $\vec{n}$ is null, in which case one has to think again.

Once this procedure is completed we have an orthonormal triad of vectors spanning the tangent spaces that are encountered by the curve, so that the metric there can be written as

$$g^{ab} = t^a t^b + n^a n^b + b^a b^b \ , \tag{3.20}$$

with an extra sign inserted somewhere if we are in Minkowski space. In the field theory literature this expression would normally be written as

$$g^{ab} = \delta^{ij} e_i^a e_j^b \ , \tag{3.21}$$

and the triad vectors are called 'dreibeins'. However, after going through all the work needed to adapt the triad vectors to the geometric object we are studying, it seems better to give them separate names such as $\vec{t} \equiv \vec{e}_1$.

As soon as one tries to do a concrete example one realizes that quite some work is involved. Let us consider a helix spiraling around the z-axis in Euclidean space,

$$x^a(\tilde{\sigma}) = \begin{pmatrix} r\cos\tilde{\sigma} \\ r\sin\tilde{\sigma} \\ a\tilde{\sigma} + h \end{pmatrix} \ , \tag{3.22}$$

where r, h, a are constants. The first thing one finds is that $\tilde{\sigma}$ is not normalized according to the rules, and has to be replaced by the arclength parameter σ,

$$\tilde{\sigma} = \frac{\sigma}{\sqrt{r^2 + a^2}} \ . \tag{3.23}$$

This gives a tangent vector of unit length. One then has to separate out a factor κ_1 to ensure that the normal $\vec{n}$ has unit length, and finally a small calculation is needed to solve for the binormal, using

$$\kappa_2 b^a = \dot{n}^a + \kappa_1 t^a \ . \tag{3.24}$$

When all is said and done we obtain

$$\begin{gathered} t^a = \frac{1}{\sqrt{r^2+a^2}} \begin{pmatrix} -r\sin\tilde{\sigma} \\ r\cos\tilde{\sigma} \\ a \end{pmatrix} \ , \qquad n^a = \begin{pmatrix} -\cos\tilde{\sigma} \\ -\sin\tilde{\sigma} \\ 0 \end{pmatrix} \ , \\ b^a = \frac{1}{\sqrt{r^2+a^2}} \begin{pmatrix} a\sin\tilde{\sigma} \\ -a\cos\tilde{\sigma} \\ r \end{pmatrix} \ , \qquad \kappa_1 = \frac{r}{r^2+a^2} \ , \qquad \kappa_2 = \frac{a}{r^2+a^2} \ . \end{gathered} \tag{3.25}$$

We will return to this result shortly, but first we observe that the particular triad or *moving frame* that we have attached to the curve may not be the most useful one for physical applications.

Another option is the *Fermi-Walker frame*. The usual context is that of an observer following a timelike worldline in spacetime, with acceleration $\vec{a} = \kappa_1 \vec{n}$.

She wants to set up a frame $(\vec{t}, \vec{e}_1, \vec{e}_2)$ such that the spatial vectors are non-rotating, in the sense that they are changing only in the direction of the tangent vector. If we start with a Frenet frame we can make the Ansatz

$$e_1^a = \alpha n^a + \beta b^a \ , \qquad e_2^a = -\beta n^a + \alpha b^a \ , \tag{3.26}$$

and use Frenet's equations to deduce that

$$\dot{e}_1^a = \kappa_1 \alpha t^a + (\dot{\alpha} - \kappa_2 \beta) n^a + (\dot{\beta} + \kappa_2 \alpha) b^a \ . \tag{3.27}$$

Setting the unwanted terms to zero gives a pair of linear ODEs, with the solution

$$e_1^a = \cos\left(\int^\sigma \kappa_2 d\sigma\right) n^a - \sin\left(\int^\sigma \kappa_2 d\sigma\right) b^a \ . \tag{3.28}$$

Using the acceleration vector $\vec{a} = \kappa_1 \vec{n}$ we can define

$$\omega^{ab} = t^a a^b - t^b a^a \ , \tag{3.29}$$

and arrive at the attractive formulas

$$\dot{t}^a = \omega^a{}_b t^b \ , \qquad \dot{e}_1^a = \omega^a{}_b e_1^a \ , \qquad \dot{e}_2^a = \omega^a{}_b e_2^a \ . \tag{3.30}$$

We can check this, for instance

$$\dot{t}^a = t^a a_b t^b - a^a t_b t^b = a^a \ , \tag{3.31}$$

which is just right. The point is that there is no rotation in the plane spanned by $\vec{e}_1$ and $\vec{e}_2$. If the observer carries out Newton's bucket experiment she will find no effect precisely if the water does not rotate relative to a Fermi-Walker frame. (It remains to explain why distant galaxies do not rotate relative to this frame. There are solutions of Einstein's equations in which they do.[1])

Now we go back to the equation (3.22) that defines the circular helix, but we allow r and h to vary. In this way we obtain a space-filling congruence of curves labeled by the Lagrangian coordinates r and h (where h is a periodic coordinate). A relevant question about a congruence is whether it is *hypersurface orthogonal.* The idea is that, at each point in space, the tangent vector of the curve defines a hyperplane element in tangent space, spanned by vectors orthogonal to the tangent. In the case of the helix we studied, the hyperplane elements are two-dimensional planes in each tangent space, spanned by the vectors $\vec{n}$ and $\vec{b}$. Can one fit those plane elements together so that they become the tangent planes of a surface? Suppose that this surface exists, and is given by an equation $f(x) = 0$. It will have a normal vector given by the gradient $\nabla_a f$. More generally we can rescale the normal, so that the normal vector of the surface is given by $g\nabla_a f$, where g is another function. But we

[1] I. Ozsváth and E. Schücking, *The finite rotating universe*, Ann. Phys. **55** 166 (1969).

are asking for a surface with a normal in the direction of the tangent vector $\vec{t}$, so we must find functions f and g such that

$$t_a = g\nabla_a f \ . \tag{3.32}$$

The condition that this be (locally) possible is

$$t_{[a}\nabla_b t_{c]} = 0 \ . \tag{3.33}$$

We can easily apply this criterion to our congruence of helices. We find that

$$\epsilon^{abc} t_a \nabla_b t_c = 2\kappa_2 \neq 0 \ . \tag{3.34}$$

(For the calculation you may find it convenient to express the vector field $\vec{t}$ in Cartesian coordinates.) So the answer is no. The curves in the helical congruence are not everywhere orthogonal to any surface. The plane elements orthogonal to the tangent vectors are horizontal at the z-axis, but when you go out from there they twist relative each other in such a way that you cannot draw a single surface through any selection of such plane elements.

Note that you may be familiar with this kind of argument from thermodynamics. Without going into the details of the kind of differential form notation used there, one considers vectors given by

$$dU + pdV + \mu dN + \dots \ , \tag{3.35}$$

and states as a Law that they are surface forming, and can be set equal to a vector TdS, normal to surfaces given by $S =$ constant. Note also that in two dimensions the question trivializes, because then the orthogonal plane elements are one-dimensional vector fields, and vector fields can always be fitted together as tangent vectors of a family of curves. And note finally that we are dealing with a special case of *Frobenius' theorem*. It says that a collection of vector fields are surface forming if and only if their commutators belong to the subspace of tangent space that they span. The hypersurface case is the easy one, because it can be discussed in terms of the unique dual vector orthogonal to all the vectors tangent to the hypersurface.[2]

Actually we have associated three vector fields to our helical congruence, and a quick calculation confirms that

$$n_{[a}\nabla_b n_{c]} = b_{[a}\nabla_b b_{c]} = 0 \ . \tag{3.36}$$

Hence the distributions of plane elements spanned by $(\vec{t}, \vec{b})$ and by $(\vec{t}, \vec{n})$ are integrable, and do form surfaces. An effort at visualization shows that in the former case we obtain cylinders, and in the latter case we obtain surfaces known as *helicoids*. They have equations like

[2] For the details, see Wald's appendices.

$$f = z - a \arctan \frac{y}{x} = 0 \ . \tag{3.37}$$

Although this equation does not work globally, it is a useful one because it describes the surface as a graph of a function. It is easily checked that

$$t^{\mathbf{a}} \partial_{\mathbf{a}} f = n^{\mathbf{a}} \partial_{\mathbf{a}} f = 0 \ , \tag{3.38}$$

so the helicoid does have $\vec{t}$ and $\vec{n}$ as tangent vectors. We will come back to this surface, because it provides a standard example of a minimal surface. (Make a helix out of a steel wire and dip it in a soap solution. You should be able to see a helicoid when you take it out again. Although it is easier to do the locally isometric catenoid.[3])

Congruences of curves are of considerable interest, partly because they can be thought of as the flowlines of moving matter in the form of freely falling 'dust' (like the galaxies in Friedmann models of the universe), and for many other reasons. Congruences may develop caustics if followed long enough, but here we will be interested in a local description where this kind of behaviour is excluded. A typical question is this: Focus on one curve in the congruence, and at some particular parameter time surround it with a geodetic ball. This picks out a particular set of curves for consideration. Let the parameter time grow and let the ball be comoving with the flow. What shapes will then be assumed by the ball? We take the ball to be small, so it will be enough to analyse the normal vectors attached to the original curve. They define the geodesics that we are using as radii of our ball. We focus on timelike curves, because they can be thought of as representing a matter flow.[4]

For this purpose let us think of a *Jacobi field* $\vec{\eta}$, that is to say a vector field that is Lie dragged along the congruence:

$$\mathcal{L}_{\vec{t}} \eta^a = 0 \qquad \Leftrightarrow \qquad [t, \eta]^a = 0 \ . \tag{3.39}$$

Since the two vector fields commute the vector field $\vec{\eta}$ points along some Lagrangian coordinate line. Intuitively the vector $\vec{\eta}$ connects two neighbouring flowlines. Because the Lie derivative vanishes we see that

$$\dot{\eta}^a = t^b \nabla_b \eta^a = \nabla_b t^a \eta^b \equiv B^a{}_b \eta^b \ , \tag{3.40}$$

where we defined

$$B_{ab} \equiv \nabla_b t_a = t_{a;b} \ . \tag{3.41}$$

(The old-fashioned notation for covariant derivatives makes it easy to remember the ordering of the indices here.) We assume that the congruence

[3] C. V. Boys: *Soap bubbles, their colours and the forces which mold them*, Courier Corporation, 1959. The classic reference for curves and surfaces in Euclidean space is D. J. Struik: *Lectures on Classical Differential Geometry*, Dover, New York 1961.

[4] A classic description is by J. Ehlers, 1961, reprinted as *Contributions to the relativistic mechanics of continuous media*, Gen. Rel. Grav. **25** (1993) 1225.

is geodetic—it is a family of test particles in free fall, if you like. (If we do not assume this we will have to carry an extra term along.) Then

$$t^b \nabla_b t^a = 0 \quad \Rightarrow \quad t^b B_{ab} = t^b B_{ba} = 0 \ . \tag{3.42}$$

Hence, at a point, the tensor B_{ab} belongs to the vector space $\mathbf{T}^\perp \otimes \mathbf{T}^\perp$, where $\mathbf{T}^\perp$ is the orthogonal complement of the tangent vector.

The next step is to introduce a metric on the orthogonal hypersurface elements, namely

$$h_{ab} = g_{ab} + t_a t_b \ , \qquad h_{ab} t^b = 0 \ . \tag{3.43}$$

This metric tensor belongs to $(\mathbf{T}^\perp \otimes \mathbf{T}^\perp)_{\mathrm{sym}}$. We use it to decompose the tensor B_{ab} into irreducible parts:

$$B_{ab} = \frac{1}{3}\theta h_{ab} + \sigma_{ab} + \omega_{ab} \ . \tag{3.44}$$

Here θ is the *expansion*, the symmetric traceless tensor σ_{ab} (i.e. $h^{ab}\sigma_{ab} = 0$) is the *shear*, and the antisymmetric tensor ω_{ab} is the *rotation* of the congruence. The idea is that if we look at a small spherical bundle of geodesics, then we find that θ causes the sphere to grow as we move along the congruence, σ_{ab} turns it into an ellipsoid, and ω_{ab} causes it to rotate.

Let us be completely explicit about the expansion, and rewrite it in two useful ways:

$$\theta = h^{ab} B_{ab} = h^{ab} \nabla_a t_b = \begin{cases} = \nabla_a t^a \\ \\ = \frac{1}{2} h^{ab} \mathcal{L}_{\bar{t}} h_{ab} \ . \end{cases} \tag{3.45}$$

If we introduce $\sqrt{h} \equiv \sqrt{\det h_{ab}}$, measuring the volume of a 3-dimensional co-moving volume element, we can write

$$\theta = \frac{1}{2} h^{ab} \mathcal{L}_{\bar{t}} h_{ab} = \frac{\mathcal{L}_{\bar{t}} \sqrt{h}}{\sqrt{h}} \ . \tag{3.46}$$

We see that, indeed, the expansion measures how a co-moving volume expands or shrinks as it is carried along the flow.

But the expansion, shear, and rotation themselves change as we move along the congruence. To see how, observe that

$$\begin{aligned} \dot{B}_{ab} &= t^c \nabla_c \nabla_b t_a = t^c \nabla_b \nabla_c t_a + t^c [\nabla_c, \nabla_b] t_a = \\ &= \nabla_b (t^c \nabla_c t_a) - \nabla_b t^c \nabla_c t_a + t^c R_{cba}{}^d t_d = -B^c{}_b B_{ac} - R_{cbda} t^c t^d \ . \end{aligned} \tag{3.47}$$

(The geodesic equation enabled us to drop one term. Notice also the general rule: whenever two covariant derivatives act on a tensor, we can bring the Riemann tensor into the game by changing their order.) It follows that

$$\dot{\theta} = -\frac{1}{3}\theta^2 - \sigma_{ab}\sigma^{ab} + \omega_{ab}\omega^{ab} - R_{ab}t^a t^b \ . \tag{3.48}$$

This is *Raychaudhuri's equation*.

The *strong energy condition* requires that

$$R_{ab}t^a t^b \geq 0 \tag{3.49}$$

for all timelike vectors $\vec{t}$. Using Einstein's equations it can be converted into a condition on the stress-energy tensor T_{ab}, which holds for most reasonable types of matter. Anyway the conclusion is that if the strong energy condition holds, and if the cosmological constant is zero, then the expansion of a rotation free congruence can only decrease. Gravity has a focussing effect, which becomes even stronger in the presence of matter obeying the strong energy condition.

Concerning rotation: In a similar way one deduces that

$$\dot{\omega}_{ab} = -\frac{2}{3}\theta\omega_{ab} - 2\sigma^c{}_{[b}\omega_{a]c} \ . \tag{3.50}$$

If there is no rotation to start with, the congruence stays rotation free. Rotation free geodetic congruences, if they exist, have an interesting property:

$$\nabla_{[a}t_{b]} = 0 \quad \Leftrightarrow \quad t_a = \nabla_a f \ . \tag{3.51}$$

The function f is known as a velocity potential. The forwards implication is only local, but let us suppose it holds everywhere, and that the congruence fills all of spacetime. This is interesting because there is a function increasing along the timelike geodesics in the congruence. By setting its value to a constant along all the curves we obtain a spacelike hypersurface of "simultaneous" events, in other words this function serves as a *global time function*. Because the function grows monotoneously the spacelike hypersurface can be crossed in one direction only by timelike curves. In this way the existence of a global time function guarantees that there are no closed timelike curves. The converse is not true in general—relativity theory allows rotating universes that do not have closed timelike curves. But some rotating universes have them.[5]

Concerning shear:

[5] K. Gödel, *An example of a new type of cosmological solutions of Einstein's field equations*, Rev. Mod. Phys. **21** (1949) 447. Gödel's solution uses a negative cosmological constant to balance the 'centrifugal force' of the rotation. Its topology is $\mathbf{R}^4$. The Ozsváth–Schücking solution on the other hand has no CTCs.

$$\dot{\sigma}_{ab} = -\frac{2\theta}{3}\sigma_{ab} - \sigma_{ac}\sigma^{c}{}_{b} - \omega_{ac}\omega^{c}{}_{b} + \\ + \frac{1}{3}h_{ab}(\sigma_{cd}\sigma^{cd} - \omega_{cd}\omega^{cd}) + C_{cbda}t^c t^d + \frac{1}{2}\tilde{R}_{ab} \ , \tag{3.52}$$

where

$$\tilde{R}_{ab} = h_{ac}h_{bd}R^{cd} - \frac{1}{3}h_{ab}h_{cd}R^{cd} \ , \tag{3.53}$$

and C_{abcd} is the Weyl tensor. (When one splits the Riemann tensor into irreducible pieces, the Weyl tensor is the piece that is algebraically unconstrained by Einstein's equations. It is "trace free" by construction.) The message is that even in vacuum gravity (where the Ricci tensor vanishes), a congruence that starts out shear free will develop shear if there is a non-zero Weyl tensor. The shear then enters the Raychaudhuri equation and causes focussing there.

One look at Raychaudhuri's equation (3.48) reveals something disturbing. Suppose the rotation is zero. Then it stays zero. Suppose also that θ is negative to start with. When the strong energy condition holds it follows that the derivative of θ along the congruence is negative, and θ will grow more negative as the absolute value of θ grows. More precisely,

$$\dot{\theta} \leq -\frac{1}{3}\theta^2 \quad \Rightarrow \quad \frac{d}{d\tau}\left(\frac{1}{\theta}\right) \geq \frac{1}{3} \quad \Rightarrow \quad \theta(\tau) \leq \frac{1}{\theta(0) + \frac{\tau}{3}} \ . \tag{3.54}$$

For a negative initial value $\theta(0)$ it follows that $\theta \to -\infty$ in finite proper time. When this happens the geodesics in the congruence come together, and a *caustic* occurs. While this is not disturbing as such, suppose that we consider the rotation-free timelike congruence along which the galaxies in a dust filled Friedmann universe are moving. As θ diverges, the dust density ρ is also diverging, leading via Einstein's equations to a curvature singularity. Hence, without any symmetry assumptions, we can conclude (by letting the parameter grow in the negative time direction) that if the Universe is expanding at some particular moment (as defined by the global time function) then there must have been a singularity in the past.

This observation was, in fact, Raychaudhuri's original point. Some years later a much more serious singularity theorem, not relying on the details of the matter model, was proved by Penrose.[6]

◇ **Problem 3.1** Prove, by introducing extra coordinates kept constant on a curve, that

[6] For this history, see J. M. M. Senovilla and D. Garfinkle, *The 1965 Penrose singularity theorem*, Class. Quant. Grav. **32** (2015) 124008.

$$t^{\mathbf{a}} = \frac{dx^{\mathbf{a}}}{d\sigma} \quad \Leftrightarrow \quad t^a \nabla_a \sigma = 1 \ . \tag{3.55}$$

This is the coordinate independent way of introducing a parameter along the curve.

◇ **Problem 3.2** Write down the Frenet–Serret equations for a timelike curve in a four-dimensional spacetime.

◇ **Problem 3.3** Our helical congruence is not hypersurface orthogonal. If you think about it, you will realize that the set of its binormals can be regarded as the tangent vectors of another helical congruence—which is everywhere orthogonal to the helicoid. Clarify using pictures rather than formulas why this difference arises.

◇ **Problem 3.4** You obtain a cone by identifying points connected with a rotation $< 2\pi$. Misner space is obtained by identifying points (in Minkowski space) connected by a Lorentz boost. Show that Misner space contains closed spacelike curves whose normal is everywhere timelike. For definiteness, choose curves with constant first curvature.

◇ **Problem 3.5** Given $\dot{\eta}^a$ as defined in eq. (3.40), define $a^a = t^b \nabla_b \dot{\eta}^a$. Use eq. (3.39) to prove the geodetic deviation equation

$$a^a = \eta^b t^c t^d R_{bcd}{}^a \ . \tag{3.56}$$

4 Equilibrium states

An equilibrium state does not change with time. If a spacetime is to qualify as an equilibrium state it must possess an everywhere timelike Killing vector, and is then called stationary. We may also require that the spacetime in some sense describes an isolated system, which presumable translates into the requirement that it be *asymptotically flat* in a suitable sense. (It may be reasonable to ignore the cosmological constant in this context.) Finally, an equilibrium state must be stable under small perturbations if it is to be of physical relevance.

4.1 The Schwarzschild solution and its relatives

It is known that the only stationary and asymptotically flat solution of the Einstein vacuum equation (with $\lambda = 0$) on a spacetime of topology $\mathbf{R}^4$ is flat Minkowski space itself. The proof goes back to Einstein, and becomes very simple if we assume that the spacetime is not only stationary but also static. This assumption implies that one can find a coordinate system such that the metric is independent of a time coordinate t, and moreover—this is the precise meaning of static—invariant under the reflection $t \rightarrow -t$. Alternatively one demands that the Killing vector field ∂_t be hypersurface orthogonal, and everywhere orthogonal to spatial hypersurfaces defined by setting the coordinate t equal to a constant. Either way one arrives at the Ansatz

$$ds^2 = -N^2 dt^2 + \gamma_{ij} dx^i dx^j \ , \tag{4.1}$$

where $x^{\mathbf{i}}$ are coordinates on the spacelike hypersurfaces orthogonal to the Killing field, while γ_{ij} and N are independent of the coordinate t. Asymptotic flatness requires that the coordinate system can be chosen so that

$$N \rightarrow 1 \ , \qquad \gamma_{ij} \rightarrow \delta_{ab} \ , \tag{4.2}$$

at large distances. For the moment we pass lightly by the question of how fast the fall-off has to occur.

Next one writes the Einstein equations for this Ansatz. Let $\bar{\nabla}_i$ be the 3-dimensional covariant derivative defined using the metric γ_{ij}, and let $\bar{R}_{ij}$ be its Ricci tensor. Then the result is

$$R_{tt} = -N\gamma^{ij}\bar{\nabla}_i\bar{\nabla}_j N = 0 \tag{4.3}$$

$$R_{ij} = \bar{R}_{ij} + \frac{1}{N}\bar{\nabla}_i\bar{\nabla}_j N \ . \tag{4.4}$$

But the first equation here is the Laplace equation, and since we are assuming that $N \to 1$ at infinity we must have $N = 1$ everywhere. The second equation then says that the spatial Ricci tensor vanishes, which implies that space is flat. QED.

Of course the story is not quite over here. The Schwarzschild metric

$$ds^2 = -\left(1 - \frac{2m}{r}\right)dt^2 + \frac{dr^2}{1 - \frac{2m}{r}} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{4.5}$$

is an asymptotically flat solution of Einstein's equation, and it is of the form (4.1). But the topology of the manifold is not $\mathbf{R}^4$, since the metric is defined only on the coordinate range $r > 2m$. Actually $0 < r < 2m$ is a possibility too, but then the metric is not static since r will be the time coordinate in this case. The hypersurface forming Killing field has a norm given by

$$||\partial_t||^2 = -V(r) = -\left(1 - \frac{2m}{r}\right) \ . \tag{4.6}$$

The Killing vector is timelike if $r > 2m$ and spacelike if $r < 2m$. It is presumably null if $r = 2m$, but this is outside the range of the coordinate system we are using, so strictly speaking we are not allowed to draw that conclusion yet. But it can be confirmed by switching to Eddington–Finkelstein coordinates. Throughout I assume that the integration constant $m > 0$. A negative value would not help since there is a curvature singularity at $r = 0$, and again the spatial topology is not that of $\mathbf{R}^3$.

It is important to familiarize oneself with the Schwarzschild solution. The story is best told with a few pictures. I hope you can make sense of all three, although the third may be unfamilar.[1] As Figure 4.1 makes clear, the solution has some unphysical features. In particular there are two asymptotic regions, and a 'white hole' in the past. Still, like the equilibrium states in statistical physics, this is a state that can be approached (in a suitable sense) in the real world.

The non-static region II repays scrutiny. For one thing it ends with a singularity, reached by timelike geodesics in finite proper time. At constant r the spatial metric is

$$dl^2 = \left(\frac{2m}{r} - 1\right)dt^2 + r^2(d\theta^2 + \sin^2\theta) \ . \tag{4.7}$$

[1] Arguably the best explanations of how to draw Penrose diagrams were given by M. Walker, *Block diagrams and the extension of timelike two-surfaces*, J. Math. Phys. 11 (1970) 2280. For Figure 4.3, see D. Marolf, *Spacetime embedding diagrams for black holes*, Gen. Rel. Grav. 31 919, 1999.

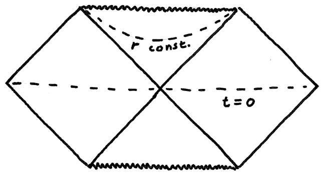

Figure 4.1. The Penrose diagram of the completed Schwarzschild manifold. Each point is a round sphere and radial light rays slope 45 degrees. Two interesting hypersurfaces are marked, and four regions labelled.

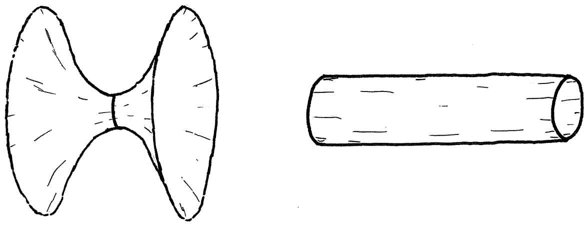

Figure 4.2. Realistic sketches of the two spatial hypersurfaces occurring in the Penrose diagram (with one dimension suppressed).

If we change r to a smaller value (that is to say, if we move to a later moment of 'time'), space expands in one direction and shrinks in two. This kind of behaviour recurs in less symmetrical models of gravitational collapse, except that there may be chaotic transitions between the expanding and the contracting directions. Another interesting observation concerns null geodesics leaving the round spheres at constant t in directions orthogonal to the spheres. That is, imagine that the round sphere is emitting two flashes of light, giving rise a moment later to two spherical wave fronts. The observation is that the area of both wave fronts are shrinking. Nothing like it happens in region I, where the ingoing wave front shrinks and the outgoing wavefront grows. A sphere such that both wave fronts shrink is said to be *trapped.* The singularity theorems proved by Penrose and others say (after the addition of some reasonable looking extra conditions) that if trapped spheres are present then there necessarily exist timelike or null geodesics that cannot be continued to infinite values of their affine parameter. As is indeed the case with the geodesics that encounter the Schwarzschild singularity.

More generally we can consider spherically symmetric and static spacetimes with metrics of the form

$$ds^2 = -V(r)dt^2 + \frac{dr^2}{V(r)} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{4.8}$$

This includes the Reissner–Nordström–(anti)–de Sitter solution, for which

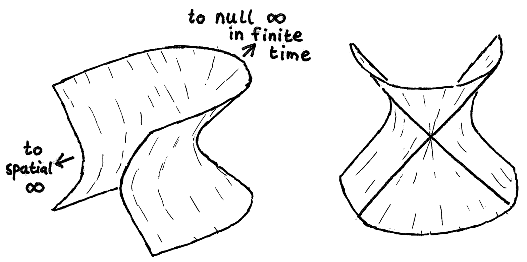

Figure 4.3. The geometry of the $1+1$ dimensional Penrose diagram, with the spheres suppressed. This surface is supposed to be embedded in a $2+1$ dimensional Minkowski space, with the singularities occurring "at infinity" in the picture (but at finite distance within the surface).

$$V(r) = 1 - \frac{2m}{r} + \frac{e^2}{r^2} - \frac{\lambda r^2}{3} = \frac{1}{r^2}\left(r^2 - 2mr + e^2 - \frac{\lambda}{3}r^4\right) . \tag{4.9}$$

This is a solution of the Einstein-Maxwell equations, with non-vanishing electric field. The parameter e is an integration constant arising when the Maxwell equations are solved, and equals the electric charge as evaluated by a surface integral at infinity. There is no charged matter anywhere. To simplify matters we set $\lambda = 0$ from now on. Then

$$V(r) = \frac{1}{r^2}(r - r_+)(r - r_-) , \qquad r_\pm = m \pm \sqrt{m^2 - e^2} . \tag{4.10}$$

If $m < |e|$ the solution is everywhere static and has a naked singularity at $r = 0$, as is in fact required by Einstein's theorem. If $m > |e|$ the solution as given splits splits into three region, and the Killing vector $\vec{\xi} = \partial_t$ is spacelike when $r_- < r < r_+$. There is an outer event horizon at $r = r_+$ and an inner horizon at $r = r_-$. Once the solution has been analytically extended as far as possible an infinite number of asymptotic regions appears. In the borderline case $m = e$ the function $V(r)$ has a double root. This case is known as the extremal Reisser-Nordström solution, and its Penrose diagram is dramatically different from that of the generic case. In particular the inner and outer horizons have merged together. The timelike singularities mean that *strong cosmic censorship* is violated in both cases (because they will affect, very adversely, the predictability of the evolution of the interior), but the *weak censor* is doing her job, in the sense that the (future) singularity is invisible for an outside observer.

Within this two-parameter family of spacetimes the spacelike singularity that occurs when $e = 0$ is clearly non-generic. However, in a wider context, it is the timelike singularity that fails to be generic. In the interior of the Reissner–Nordström solution the inner horizon is a *Cauchy horizon* beyond which predictability is lost. To a mathematical relativist, a Cauchy horizon

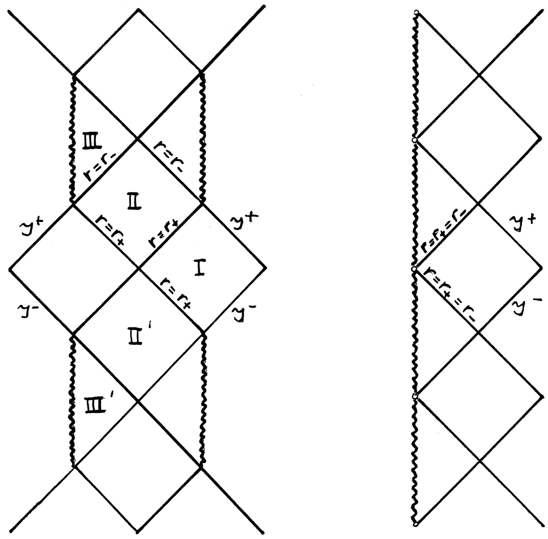

Figure 4.4. The Reissner-Nordström spacetime to the left, and the extremal Reissner-Nordström to the right. $\lambda = 0$. Ingoing Eddington–Finkelstein coordinates cover regions I, II, and III, while outgoing Eddington–Finkelstein coordinates cover regions I, II', and III'.

is much more shocking than a singularity. Fortunately this Cauchy horizon appears to be unstable under perturbations, while the event horizon is stable.[2]

From an astrophysical perspective the Reissner–Nordström solution is uninteresting. Black holes in the real world are likely to have effectively vanishing electric charge. However, from the point of view of mathematical relativity—where one of the aims is to show that the world is a 'deterministic box', albeit not a very comfortable one due to singularities—it is interesting as a toy model for the astrophysically interesting Kerr solution. The Kerr solution also has Cauchy horizons in its interior, and one aims to prove that this is an unstable feature, so that the full time evolution is predictable from initial data in the generic case.

The overall conclusion so far is that the theory admits non-trivial equilibrium states, which is surprising in view of Einstein's theorem. However—excepting flat spacetime—the static and asymptotically flat solutions of Einstein's equations that we have found are bounded inwards by a null hypersurface ruled by a Killing vector field which is null on the hypersurface itself. We need to understand this null hypersurface better.

4.2 Killing horizons and surface gravity

First some facts about null hypersurfaces in general. Locally every hypersurface is defined by setting some function of the coordinates to zero. The

[2] See M. Dafermos, *Price's law, mass inflation, and cosmic censorship*, arXiv:gr-qc/0401121.

gradient of that function defines the normal vector of the hypersurface, and at each point of the hypersurface the tangent space contains all vectors orthogonal to the normal. If the normal vector is timelike the hypersurface is spacelike, and conversely. But it can happen that the normal vector is null, in which case it is orthogonal to itself—and the normal vector is then tangential to the hypersurface. Such a hypersurface is null. At every point of a null hypersurface there is a preferred null direction within the hypersurface, along which its null normal vector points.

Interestingly the preferred null vector field is always geodetic. To see this, consider the equation $f(x) = c$, with $c \in [-\epsilon, \epsilon]$ a constant. Locally this describes a one parameter family of hypersurfaces with normal vectors $n_a = \nabla_a f$. We asssume that

$$f = 0 \quad \Rightarrow \quad n^2 = g^{ab}\nabla_a f \nabla_b f = 0 \ . \tag{4.11}$$

In words $f = 0$ describes a null hypersurface, but $f = c$ may not. We now compute

$$n^b\nabla_b n_a = n^b\nabla_b\nabla_a f = n^b\nabla_a\nabla_b f = n^b\nabla_a n_b = \frac{1}{2}\nabla_a n^2 \ . \tag{4.12}$$

Two cases arise. Either the parameter c does provide a one-parameter family of null hypersurfaces, in which case n^2 vanishes in a region. Then the right hand side is zero, and the normal vector field $\vec{n}$—which then lies within the null hypersurfaces—is geodetic. Or else we have a null hypersurface only if $c = 0$, in which case that particular hypersurface is singled out by the equation

$$n^2 = g^{ab}\nabla_a f \nabla_b f = 0 \ . \tag{4.13}$$

At the hypersurface it follows that the gradient of f and the gradient of n^2 point in the same direction, namely along the unique null direction on the hypersurface. Therefore Eq. (4.12) says that the acceleration along the vector field $\vec{n}$ is aligned with the vector field itself. As noted in Section 3.2 we can then make the acceleration vanish by means of a reparametrization of the parameter along the vector field, and again we conclude that there is a null geodesic directed along that null direction. The conclusion is that every null hypersurface is ruled by null geodesics.

The intrinsic metric on a null hypersurface is degenerate. Null hypersurfaces, and null curves, are in many ways more 'rigid' than their spacelike and timelike cousings. In a way a null hypersurface resembles a Newtonian spacetime, in which distances can be measured either in spacelike directions, or along a preferred time direction (in the null hypersurface we use an affine parameter along the null geodesics), while a proper notion of spacetime distance is missing.

But the inner boundary of the Schwarzschild solution is a very special null hypersurface since it is ruled by a null Killing vector field. It is this property that singles it out as an equilibrium state of the theory, as we will see.

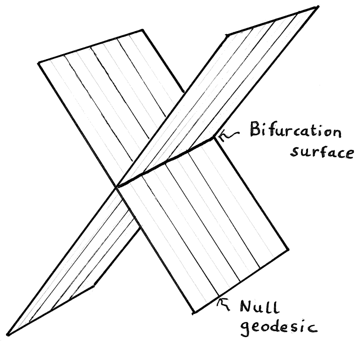

Figure 4.5. A bifurcate Killing horizon.

In all sufficiently small regions spacetime is close to flat. If spacetime has a symmetry, and if you look at it with a sufficiently strong magnifying glass, the Killing vector field will behave like some Killing vector field in Minkowski space. There are then three main cases to consider: translations, rotations, and boosts. There are also various linear combinations of these which we gloss over. The various cases are distinguished by the nature of their fixed points and by the nature of the hypersurfaces on which the norm squared of the Killing vectors vanish.

Translations have no fixed points, and their norms are always non-zero. Rotations do have fixed points—forming timelike 2-planes in Minkowski space. Boosts are more interesting. In Minkowski space, with its standard coordinate system, a typical boost is

$$\vec{\xi} = T\partial_X + X\partial_T \quad \Rightarrow \quad \vec{\xi}\cdot\vec{\xi} = T^2 - X^2 \ . \tag{4.14}$$

There is a spacelike 2-plane of fixed points at $T = X = 0$. More is true; the flow lines are lightlike on a two sheeted null hypersurface

$$T = \pm X \ . \tag{4.15}$$

More than that, the flow lines are not only lightlike, they are null geodesics. The two sheets bifurcate at the fixed 2-plane. In general a null hypersurface where a Killing vector field becomes null is called a *Killing horizon.* The Schwarzschild solution also has a Killing horizon, but its *bifurcation surface* is a 2-sphere (represented by the central point of its Penrose diagram). The Reissner–Nordström solution has several Killing horizons.

We define the *surface gravity* κ of a Killing horizon through the equation

$$\nabla_a \xi^2 = -2\kappa \xi_a \ , \tag{4.16}$$

evaluated at the horizon itself. This equation will always hold, for some function κ, because we assume that the normal vector of the hypersurface defined by $\xi^2 = 0$ is null, and we also know that the Killing vector points along that same direction. It certainly can happen that a Killing vector field becomes null on some timelike hypersurface, but then the story ends because Eq. (4.16) does not hold—there is no Killing horizon.

Due to its definition in terms of Killing vectors the surface gravity κ must be constant along each generator—obviously so because we can move in this direction using an isometry. Now let v be a *Killing parameter* along the flow lines of the Killing vector field, that is to say that

$$\xi^a \nabla_a v = 1 \quad \Leftrightarrow \quad \xi^{\mathbf{a}} = \frac{dx^{\mathbf{a}}}{dv} \ . \tag{4.17}$$

We recall that on the Killing horizon these flow lines are actually null geodesics. As such, they admit a preferred affine parameter. This means that we have two different notions of "preferred parameter" available, and it turns out that we can interpret surface gravity as a measure of to what extent the Killing parameter differs from the standard affine parameter along a geodesic. To see this, we rewrite the defining equation a little;

$$\kappa \xi_a = -\xi^b \nabla_a \xi_b = \xi^b \nabla_b \xi_a \ . \tag{4.18}$$

The conclusion, a necessary one in view of what we already know about null hypersurfaces, is that within the horizon the Killing field is aligned with a geodetic vector field. This is why there is another preferred parameter along the generators, namely the affine parameter σ.

Before we proceed, let us make sure that we know how to calculate κ in a concrete situation. It is likely that we already have a Killing parameter defined, because when a spacetime admits a Killing vector field one usually tries to place one of the coordinate lines along it. The usual Schwarzschild coordinate system employs two Killing parameters, t and ϕ, running along the Killing vector fields ∂_t and ∂_ϕ. In passing to Eddington–Finkelstein coordinates by means of Eq. (2.46), the Killing parameter t is exchanged for the Killing parameter v. They differ only by a function of r, so that on each individual Killing flowline they are shifted relative to each other by an additive constant only. Now, keeping spacetime as general as the Ansatz (4.8), and assuming that there is a Killing horizon at $r = r_H$ so that $V(r_H) = 0$, we want to calculate its surface gravity. We do this using Eddington-Finkelstein coordinates. The Killing vector is

$$\vec{\xi} = \partial_v \quad \Rightarrow \quad \xi_a = \nabla_a r \ . \tag{4.19}$$

Evidently then

$$\nabla_a \xi^2 = -\nabla_a V(r) = -V'(r) \nabla_a r = -2 \frac{V'(r)}{2} \xi_a \ . \tag{4.20}$$

This calculation shows first that the hypersurface $V(r) = 0$ is null, and second—comparing to Eq. (4.16)—it permits us to read off that

$$\kappa = \frac{V'(r_H)}{2} \ . \tag{4.21}$$

And the calculation is complete. The subscript on the argument reminds us that we must evaluate the expression at a zero of the function $V(r)$. For the event horizon in Schwarzschild we obtain

$$\kappa = \frac{M}{r_H^2} = \frac{1}{4M} \ . \tag{4.22}$$

A large black hole means a small surface gravity.

Now we go back to generalities. To find the relation $\sigma = \sigma(v)$ between the two preferred parameters along the null geodesics on the horizon, set

$$t^{\mathbf{a}} \equiv \frac{dx^{\mathbf{a}}}{d\sigma} = \frac{dv}{d\sigma}\frac{dx^{\mathbf{a}}}{dv} = \frac{dv}{d\sigma}\xi^{\mathbf{a}} = \frac{1}{\sigma'}\xi^{\mathbf{a}} \ , \tag{4.23}$$

where the prime denotes differentiation with respect to the Killing parameter. A quick calculation along the lines of (3.13) then shows that

$$\dot{x}^b\nabla_b\dot{x}^a = \frac{1}{\sigma'^{\ 2}}\left(\xi^b\nabla_b\xi^a - \frac{\sigma''}{\sigma'}\xi^a\right) \ . \tag{4.24}$$

We set this to zero. Comparing to Eq. (4.18) this leads to

$$\frac{\sigma''}{\sigma'} = \kappa \ . \tag{4.25}$$

Ignoring two arbitrary integration constants the solution is

$$\sigma = \begin{cases} e^{\kappa v} & \text{if } \kappa \neq 0 \\ v & \text{if } \kappa = 0 \ . \end{cases} \tag{4.26}$$

We see that κ enters this relation in an essential way. At the bifurcation surface the affine parameter passes through zero while the Killing parameter v goes to $-\infty$.

The affine parameter, the Killing parameter, and the surface gravity are defined only up to numerical factors. In asymptotically flat spacetimes the standard convention is to insist that the norm of the Killing field tends to one at infinity, and in any case to insist that the surface gravity is non-negative. This normalization is used for the Killing vector ∂_t in the Schwarzschild solution, but cannot be used for a boost in Minkowski space.

An explicit formula for the surface gravity is

$$\kappa^2 = -\frac{1}{2}\nabla_a\xi_b\nabla^a\xi^b \ . \tag{4.27}$$

This formula can be used also on a bifurcation surface, where the Killing vector field vanishes. To prove it is an interesting exercise. The starting point is the observation that the Killing horizon is a hypersurface, and therefore its normal vector $\vec{\xi}$ is hypersurface orthogonal at least on the horizon itself, implying that

$$\xi_{[a}\nabla_b\xi_{c]} = 0 \tag{4.28}$$

on the horizon. Next we recall that $\nabla_a\xi_b = -\nabla_b\xi_a$ for any Killing vector field. With a little dexterity, one sees that this can be used to rewrite the previous equation as

$$\xi_a\nabla_b\xi_c = -2\xi_{[b}\nabla_{c]}\xi_a \ . \tag{4.29}$$

Contracting with $\nabla^b\xi^c$, and making use of the definition of κ twice, gives the result.

With Eq. (4.27) in hand we go on to prove the important result that the surface gravity is constant on any bifurcate Killing horizon, and not only along each generator separately. To do so we first observe the fact that κ must be constant along the Killing generators, so that it is enough to show that κ is constant on any cross section of the horizon. Let us choose the bifurcation surface for this purpose. Let $\vec{s}$ be a tangent vector to the bifurcation surface. It follows that

$$\kappa s^a\nabla_a\kappa = -\frac{1}{2}s^c\nabla_c\nabla_a\xi_b\nabla^a\xi^b \ . \tag{4.30}$$

But, for any Killing vector field

$$\nabla_c\nabla_a\xi_b = -R_{abcd}\xi^d \ . \tag{4.31}$$

To prove this, observe that

$$\nabla_a\nabla_b\xi_c + \nabla_b\nabla_c\xi_a = \nabla_a\nabla_b\xi_c - \nabla_b\nabla_a\xi_c = R_{abc}{}^d\xi_d \ . \tag{4.32}$$

Write down all the cyclic permutations of this equation. Add two of them, subtract one, and use the Jacobi identity for the curvature. We can now continue the calculation in Eq. (4.30) to see that

$$\kappa s^a\nabla_a\kappa = \frac{1}{2}s^cR_{cabd}\xi^d\nabla^a\xi^b = 0 \tag{4.33}$$

on the bifurcation surface itself—because $\vec{\xi}$ vanishes there. This does it.

It is a weakness of the above proof of the constancy of κ that we had to assume that there is a bifurcation surface. An alternative proof that avoids this assumption can be given, but it relies on Einstein's equation and a condition on the stress-energy tensor known as the *dominant energy condition* (saying that if t^a is a future directed timelike vector, then $-T^a_{\ b}t^b$ is future directed timelike or null). The alternative proof is somewhat lengthy. It is worth noticing that a number of general statements, such as the theorem that the event horizon of a stationary black hole has to be a Killing horizon, depend for their proof either on some geometrical restriction or on the field equations with a suitable energy condition imposed.[3]

Since κ is constant all over the horizon it begins to feel a little like the temperature of an object in equilibrium. This analogy is strengthened when we realize that it has an absolute zero. A Killing horizon with vanishing surface gravity is said to be degenerate. Returning to our explicit example we note that as long as the zero of $V(r)$ is a simple root the surface gravity (4.21) will be non-zero. But if the root is repeated, so that $V'(r_H) = 0$, we have a degenerate Killing horizon with $\kappa = 0$. This is what happens for the extremal Reissner-Nordström black hole, see Figure 4.4. Indeed, by definition a black hole is said to be *extremal* if its event horizon is a degenerate Killing horizon. Degenerate Killing horizons, for which $\kappa = 0$ and $\sigma = v$, do not have a bifurcation surface. They consist of a single sheet. A particularly simple example of a degenerate Killing horizon is that of any null plane in Minkowski space. It is the Killing horizon of a null translation, and obviously does not have a bifurcation surface.

To summarize: The overall scale of the surface gravity is arbitrary since it can be changed by a constant renormalization of the Killing vectors, but it has an absolute zero attained by degenerate horizons. Also the surface gravity is the same all over the horizon. This is already enough to see that the surface gravity has properties reminiscent of the temperature of an object in equilibrium. A quantum field theory calculation first done by Hawking gives the interpretation

$$\kappa = 2\pi T_{\mathrm{H}} \tag{4.34}$$

to the surface gravity of a Killing event horizon, where T_{H} is the *Hawking temperature* of an evaporating black hole. An easier example of a similar nature is that of the vacuum state as observed by constantly accelerating detectors in Minkowski space. This is called the *Unruh effect*. It all stems from the fact that the discrepancy between the affine and Killing parameters gives rise to two different ways of splitting a field into positive and negative frequency modes, and hence the definition of "particles" becomes ambiguous. But we do not go into this here.[4]

It remains to explain the etymology of the name "surface gravity". It comes from an interpretation of κ valid for static black holes only. The argument behind it is delicate, so we will go slowly through it. In static spacetimes there is a natural notion of 'standing still'. It means to follow a Killing flowline. Suppose that $\xi^2 = -V(r) = -\Lambda^2$. This defines the useful redshift factor Λ. Then an observer following a Killing flowline has velocity vector $\vec{u} = \Lambda^{-1}\vec{\xi}$. This observer is subject to the acceleration

[3] For more on this, see R. M. Wald, *The thermodynamics of black holes*, Living Reviews in Relativity **4:6**, 2001.

[4] The standard reference is R. M. Wald: *Quantum Field Theory in Curves Spacetime and Black Hole Thermodynamics*, Chicago UP, 1994.

$$a_a = u^b \nabla_b u_a = \frac{1}{\Lambda}\xi^b \nabla_b \xi_a + \frac{1}{2\Lambda}\xi_a \xi^b \nabla_b(\xi_c \xi^c) = -\frac{1}{\Lambda^2}\xi^b \nabla_a \xi_b =$$
$$= -\frac{1}{2\Lambda^2}\xi^b \nabla_a (\xi^2) = \frac{1}{\Lambda}\nabla_a \Lambda = \frac{\Lambda'}{\Lambda}\nabla_a r \quad \Rightarrow \quad a = \sqrt{a_a a^a} = \Lambda' \ . \tag{4.35}$$

The Killing equation was used twice in this calculation. By definition the *energy per unit mass* of a particle following this worldline is

$$E = -u_a \xi^a = \Lambda \ . \tag{4.36}$$

At infinity the energy of a particle equals mc^2, a familiar result.

We will be interested in the work required to move a particle along some path in space. This is a differential form dW. Only shifts in the radial direction require work, so it must be that

$$dW = f_a dx^a = f_r dr \ . \tag{4.37}$$

Now work is force times distance, and the magnitude of the force acting on a stationary particle is the acceleration a computed above. Therefore, if we act with dW on a radial tangent vector of unit length, the result must be

$$dW(\Lambda \partial_r) = a \quad \Rightarrow \quad dW = a\frac{dr}{\Lambda} \ . \tag{4.38}$$

Here we assumed that the metric takes the form (4.8). This gives the work expended locally, when we move the particle.

But suppose that the particle hangs at the end of a massless inelastic string suspended from a rocket ship including a stationary observer. How much work W_2 does the observer at $r = r_2$ have to spend in order to perform an amount of work W_1 on the particle at $r = r_1$? It is part of the definition of the string that the lengths over which the two ends of the string move are equal at r_2 and r_1. But the amounts of work are not. If W_1 is subsequently transformed into pure radiation and beamed back to the observer, this energy will be redshifted, so energy conservation actually requires that $W_1 > W_2$. The exact relation is given by introducing a redshift factor:

$$\Lambda_2 W_2 = \Lambda_1 W_1 \ . \tag{4.39}$$

But the total work will be equal to force times distance. Since the distances are the same, the force applied by the observer is smaller, by the redshift factor, than the force acting on the particle:

$$F_2 = \frac{\Lambda_1}{\Lambda_2} F_1 \ . \tag{4.40}$$

Suppose the observer is at infinity ($\Lambda_2 = 1$), and the particle is kept hovering at the horizon. The force acting on the particle will then be infinite, but the redshift factor Λ_1 will vanish. The force exerted at infinity is

$$F_\infty = (\Lambda a)_{\text{hor}} = (\Lambda \Lambda')_{\text{hor}} = \frac{V'(r_H)}{2} = \kappa \ . \tag{4.41}$$

And this is why κ is referred to as "surface gravity". The name is kept also for the rotating Kerr black hole, even though the argument fails there—there the Killing vector ∂_t goes lightlike at the ergosphere, outside the horizon, and it is not possible for a stationary observer to keep the particle hovering at any point at or below the ergosphere. (It is still true that the event horizon is a Killing horizon, as we will see in Chapter 6.)

◇ **Problem 4.1** Give the equations that go into Flamm's paraboloid (shown in Figure 4.2). That is, find functions $x = x(r,\phi), y = y(r,\phi), z = z(r,\phi)$ such that

$$ds^2 = dx^2 + dy^2 + dz^2 = \frac{dr^2}{1 - \frac{2m}{r}} + r^2 d\phi^2 \ .$$

◇ **Problem 4.2** Find a timelike curve in Minkowski space that reaches 'infinity' in finite proper time. How is this relevant to Figure 4.3? Does it mean that Minkowski space is "incomplete", in the same sense as the Schwarzschild spacetime is incomplete because timelike curves can disappear at finite proper time?

◇ **Problem 4.3** In Schwarzschild, introduce new coordinates

$$\tau = \frac{2^{1/2}}{3} m^{-1/2} r^{3/2} \ , \qquad \rho^2 = \frac{3^{4/3}}{2^{2/3}} m^{2/3} \theta^2 \ , \qquad z = \frac{2^{2/3}}{3^{1/3}} m^{1/3} t \ . \tag{4.42}$$

Then take the limit $m \to \infty$. What happens?

◇ **Problem 4.4** In Minkowski space, find an example of a Killing vector field that goes lightlike on a timelike hypersurface.

◇ **Problem 4.5** Suppose you have a theory that allows you to convert particles to light and back again, and that gravity acts on the particles. Of course, constructing perfect mirrors is not a problem in this theory. Show that you either have a redshift factor, or that you can construct a *perpetuum mobile*.

5 A first look at gravitational collapse

We would now like to see a solution describing a physical system that approaches (in some sense) the Schwarzschild solution as it evolves. This can be obtained by means of a method invented by the Irish relativist Synge. *Synge's method* is as follows.[1] To solve

$$G_{ab} = 8\pi T_{ab} \ , \tag{5.1}$$

rewrite as

$$T_{ab} = \frac{1}{8\pi} G_{ab} \ , \tag{5.2}$$

choose any metric tensor g_{ab}, compute its Einstein tensor G_{ab}, and read off the stress-energy tensor T_{ab} from Eq. (5.2). The result is a solution of Eq. (5.1). To avoid any misunderstanding, Synge meant this as a joke (and he did not predict dark matter). A stress-energy tensor computed in this way is not likely to obey any of the positivity conditions that are necessary for it to qualify as physical.

Very occasionally the method works though. As our input metric we choose a spherically symmetric metric in Eddington-Finkelstein coordinates, and specialize it to

$$ds^2 = -\left(1 - \frac{2m(v)}{r}\right) dv^2 + 2dvdr + r^2 d\theta^2 + r^2 \sin^2\theta d\phi^2 \ . \tag{5.3}$$

We then arrive at

$$T_{ab} = \frac{\dot{m}}{4\pi r^2} l_a l_b \ , \tag{5.4}$$

where the dot now means differentiation with respect to v and l_a is the inwards directed null vector field

$$l_a = -\nabla_a v \quad \Leftrightarrow \quad l^a \partial_a = -\partial_r \ . \tag{5.5}$$

[1] His book, J. L. Synge: *Relativity: The General Theory*, North-Holland, Amsterdam, 1960, is not only brilliant, its preface gives a very good rationale for why the subject should be studied in the first place.

Provided that $\dot{m} \geq 0$ this is a perfectly respectable stress-energy tensor, describing a shell of incoherent electromagnetic radiation or "null dust" coming in from past null infinity. You can easily check that both the strong and the dominant energy conditions are obeyed (if you remember what those are). The spacetime itself is called the *Vaidya solution*.

This is an interesting toy model of gravitational collapse. The rate at which matter comes in is at our disposal, and we choose to set $m = 0$ when $v < 0$, then let m grow at some rate that suits us, until it reaches some finite value M at some later moment in advanced time. Thus, what we are describing is a spherically symmetric shell of matter falling into Minkowski space, eventually—since we assume that $\dot{m}$ eventually becomes zero—leaving a piece of the Schwarzschild solution behind. Actually, to ensure that the solution be asymptotically flat to the future it is enough if the total mass remains finite. One can also start from a Schwarzschild black hole, and add a Vaidya region to model black hole accretion, but we will touch only very lightly on this. Vaidya originally thought of his solution with the opposite time orientation, and referred to it as a solution for a radiating star.

With our choice the solution is divided into an initial flat region, a Vaidya region, and a final Schwarzschild region created by infall of radiation into a flat spacetime. There are pitfalls along the way: the radiation density will go to infinity at the origin. This is known as a shell focusing singularity. The geometry itself also misbehaves. The *Kretschmann scalar*—one of the scalars functions one can construct out of the curvature tensor—is

$$R_{abcd}R^{abcd} = \frac{48m^2}{r^6} \ . \tag{5.6}$$

Hence the geometry is singular at $r = 0$. The question arises whether this is due to the fact that we assumed exact spherical symmetry, or whether the singularity will be present also if the initial data are changed so that they are only approximatively spherically symmetric.

The singularity theorems—due mainly to work by Penrose, Hawking, and Geroch—state that if there exists a trapped surface in a solution of Einstein's equations, the solution will be geodesically incomplete to the future provided that the stress-energy tensor obeys a suitable positivity condition. A *trapped surface* is a closed surface (in practice, a sphere) such that both of the two orthogonal congruences of future directed null geodesics that emanate from the surface are convergent when they leave the surface. For a closed surface in Minkowski space the outgoing congruence would be divergent, i.e. outgoing wavefronts are increasing their areas, so there are no trapped surfaces in Minkowski space. On the other hand it is clear by inspection that the round spheres in a general spherically symmetric spacetime will be trapped whenever $r < 2m$: the wavefronts will again be round spheres, the area of the round spheres is $4\pi r^2$, and the coordinate r serves as a monotoneously decreasing time coordinate in such circumstances. The thing to observe is that the trapped surface condition comes in the form of an inequality, which means that it will be valid also for small perturbations away from the initial data that contain them. Hence the singularities found in spherically symmetric models are not just artefacts of the special symmetry. Geodesically incomplete means that there are geodesics (null or timelike) that disappear at a finite value of their affine parameters in an irreparable way, that is to say that it is impossible to find an extended spacetime free of this difficulty. More precisely, the theorems show that it is not possible to isometrically embed the solution as a subset of a larger geodesically complete manifold. Note that the restriction to geodesics is important, since Minkowski space contains incomplete curves—and Minkowski space certainly should count as regular. (Recall Figure 4.3.) It is expected that the incomplete geodesics will encounter regions of diverging curvature when they disappear, but this does not follow from the theorems.

Strong censorship states that in a generic spacetime no observations of a future singularity can be made, that is to say that no future directed timelike or null curves emerge from them. In effect this means that a generic spacetime is globally hyperbolic, and fully determined by initial data on a spacelike Cauchy hypersurface. Weak censorship states that no observations of a future singularity can be made close to infinity in a generic asymptotically flat spacetime, that is to say that they occur only inside the event horizon that bounds the region that can be seen from infinity. In effect a black hole forms around the singularity, so that astronomers cannot see it. Observable singularities are called naked, and would wreak havoc with the predictive power of general relativity if they occur.

These formulations are vague, in particular the meaning of the word "generic" is not specified. It is known, for instance, that the Reissner-Nordström solution contains locally naked singularities, but there are arguments to show that this part of the solution is unstable against perturbations. The formulations can be improved, but they will remain vague until the cosmic censorship hypothesis is either proved or disproved. This will probably take a long time, and meanwhile black hole physics rests on an unproved conjecture. One can try to argue that astronomers would have alerted theoreticians that naked singularities are out there—if they were.

We want to know if the singularity in the Vaidya solution is naked or not. Note at the outset that if it is, this will not count as a serious failure of cosmic censorship, but will be blamed on the matter model, which gives a poor description of real world electromagnetic fields when the density becomes very high. Nevertheless we will learn that the standard energy conditions imposed on T_{ab} are not in themselves enough to ensure cosmic censorship.

To investigate whether the Vaidya singularity is naked we study radially directed null geodesics; they obey

$$\dot{x}^2 = \dot{v}\left(2\dot{r} - \left(1 - \frac{2m}{r}\right)\dot{v}\right) = 0 \ . \tag{5.7}$$

The ingoing congruence is disposed of easily. It is described by

$$v = \text{constant} \ , \qquad r = \tau_0 - \tau \ , \tag{5.8}$$

where τ is an affine parameter along the ray. Since the Eddington-Finkelstein coordinates are perfectly adapted to the study of the Vaidya solution we will sometimes draw pictures directly in the v-r-plane. This is quite confusing at first, since the ingoing null geodesics will be parallel with the r-axis in the diagrams. It is not too hard to get used to though. See Figure 5.3 for an example.

To investigate the singularity we need the outgoing congruence, whose equation is

$$\frac{dv}{dr} = \frac{2}{1 - \frac{2m(v)}{r}} \ . \tag{5.9}$$

To solve it we must specify the mass function $m(v)$, but to begin with it is enough to observe that the forwards light cones are pointing towards decreasing r as soon as $r < 2m$. This means that if there are any signals coming out from the singularity they must come from a single point in the $v-r$-diagram, namely $(r, v) = (0, 0)$. Given that $m(0) = 0$ this is just on the boundary between the flat and the Vaidya region. We can use a variant of Synge's method to investigate what happens there.[2]

We rewrite the geodesic equation as

$$2\frac{dr}{dv} = 1 - \frac{2m}{r} \qquad \Leftrightarrow \qquad m(v) = \frac{r}{2}\left(1 - 2\frac{dr}{dv}\right) \ . \tag{5.10}$$

Let us assume that, for small v, the outgoing geodesic is given by

$$r = \beta v^a \ , \qquad \beta > 0 \ , \quad a > 0 \ . \tag{5.11}$$

Clearly a geodesic that starts at $r = 0$ and goes into the region with positive r-values must behave like this, with a positive coefficient β, to leading order in v. We find that

$$m(v) = \frac{1}{2}\beta v^a (1 - 2\beta a v^{a-1}) \ . \tag{5.12}$$

There are now three cases to investigate.

First the case when the mass function grows very slowly in the initial stages. Then

$$a > 1 \ : \qquad m(v) \sim \frac{\beta}{2} v^a \ . \tag{5.13}$$

There are geodesics coming out of the singularity, and hence the singularity is at least locally naked. The next case is

[2] Y. Kuroda, *Naked singularities in the Vaidya solution*, Prog. Theor. Phys. **72** (1984) 63.

$$a = 1 \ : \qquad m(v) = \frac{1}{2}\beta(1-2\beta)v \equiv \mu v \ . \tag{5.14}$$

Since we insist that $\dot{m} > 0$ we must set $0 < \beta < 1/2$, and we find that there is a locally naked singularity provided that

$$m(v) = \mu v \ , \qquad \mu \leq \frac{1}{16} \ . \tag{5.15}$$

Finally we have the case when the mass starts out growing quickly:

$$0 < a < 1 \ : \qquad m(v) \sim -a\beta^2 v^{2a-1} \ . \tag{5.16}$$

But this is not allowed: the assumption that outgoing geodesics exist leads to a contradiction with the condition that $m(v)$ be a positive function. The conclusion is that there are no outgoing geodesics, and hence no naked singularity, in this case.

Although convenient, Synge's method is not needed to show this. A more systematic approach is to rewrite Eq. (5.9) as an autonomous system of ordinary differential equations, namely

$$\frac{dv}{d\sigma} = 2r \ , \qquad \frac{dr}{d\sigma} = r - 2m(v) \ . \tag{5.17}$$

We are now interested in the phase portait in the r–v-plane, especially in the neighbourhood of the obvious fixed point at the origin. (Recall that $m(0) = 0$.) We linearize around the fixed point, using the definition

$$\lim_{v \to 0_+} \frac{m(v)}{v} = \mu \ . \tag{5.18}$$

We allow $\mu = 0$ and $\mu \to \infty$, so all three cases are included. Close to the fixed point we have the linear system

$$\begin{pmatrix} \dot{r} \\ \dot{v} \end{pmatrix} = \begin{pmatrix} 1 & -2\mu \\ 2 & 0 \end{pmatrix} \begin{pmatrix} r \\ v \end{pmatrix} \ . \tag{5.19}$$

We can solve this exactly, but we do not need to do this in order to sketch the picture we obtain in the $v - r$-plane. We introduce the angle-like variable

$$x = \frac{v}{r} \ , \tag{5.20}$$

and observe that

$$\dot{x} = \frac{\dot{v}r - v\dot{r}}{r^2} = 2 - x + 2\mu x^2 = 2\mu\left(\left(x - \frac{1}{4\mu}\right)^2 + \frac{16\mu - 1}{\mu}\right) \ . \tag{5.21}$$

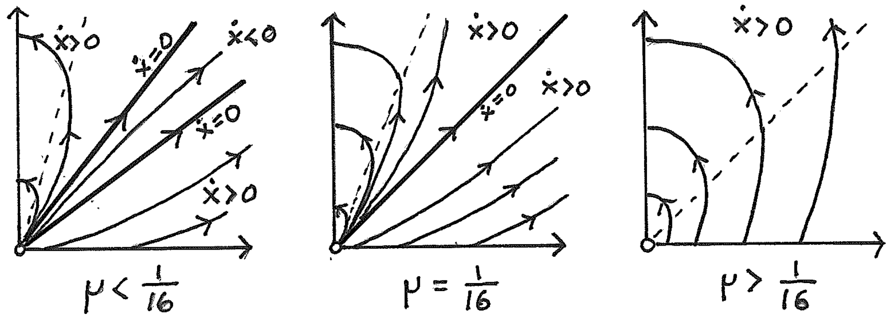

Figure 5.1. If $\mu \geq 1/16$ the null geodesics surround the fixed point, and nothing can come out. If $\mu = 1/16$ there is a spray of geodesics coming out tangential to the critical solution $x = 4$. If $\mu < 1/16$ there are two critical solutions, with null geodesics coming out in between. The dashed line is where the geodesics are directed vertically. I adjusted the scale on the vertical v-axis to make the pictures look nicer.

The value $\mu = 1/16$ is critical. If μ is larger than this, x increases monotoneously. If $\mu = 1/16$ then we find the special solution $x = 4$. If $\mu < 1/16$ there are two special solutions of this kind, and $\dot{x} < 0$ in between. To draw the picture it is useful to notice that, in all three cases,

$$\frac{dv}{dr} = \begin{cases} 2 & \text{if} \quad v = 0 \\ \infty & \text{if} \quad r = 2\mu v \\ 0 & \text{if} \quad r = 0\ . \end{cases} \tag{5.22}$$

Also

$$\frac{d^2v}{dr^2} = \frac{1}{16r}\left(\frac{dv}{dr}\right)^3 \left((4\mu x - 1)^2 + 16\mu - 1\right)\ , \tag{5.23}$$

from which we can easily read off the sign of the second derivative. We are ready to draw, and the result can be seen in Figure 5.1.

We have found a (partial) failure of cosmic censorship. Whether this is serious or not is a question we will have to think about. But so far we have addressed the issue of strong cosmic censorship only. A locally naked singularity may still be hidden behind an event horizon so that weak cosmic censorship holds. This is a more difficult question. To answer it we must specify the mass function and solve Eq. (5.9) for the outgoing geodesics also far from the fixed point. In fact, depending on the mass function all three possibilities occur: naked, locally but not globally naked, and clothed.

To exhibit the causal structure of the solution we will draw a Penrose diagram. To do so we begin with the question how to draw the singularity. It will always have a spacelike part, because radial null geodesics cannot come out from it anywhere in the region where $v > 0$. But we have just seen that in some cases a whole spray of them can come out of the point where $(r, v) = (0, 0)$.

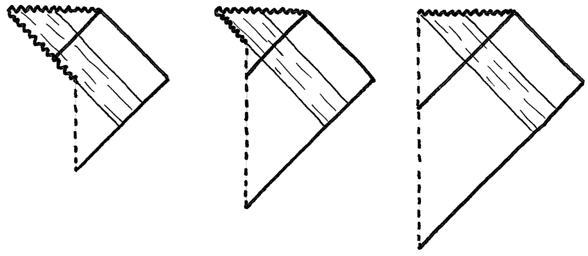

Figure 5.2. Possible Penrose diagrams for the Vaidya solution. Matter comes in in the dashed region. The singularity may be naked or clothed.

Whenever this is the case the singularity will have a null part. It cannot appear as a timelike line in the diagram, and it must be slanted as shown, because at fixed values of the angular coordinates only one incoming null geodesic can hit this point. So we have to draw the singularity either as a spacelike line or as a null line meeting a spacelike line. See Figure 5.2.

At this point you may have become nervous. I came close to saying that the singularity "sits at $r = 0$", and this cannot be quite right since this is outside the range of our coordinates. There are no such points in the spacetime manifold. But it is still true that the singularity has acquired some structure and one can in fact talk, in a meaningful way, of spacelike, timelike, or null singularities. To do this strictly one can define the "points" of the singularity as equivalence classes of those curves that are leaving the spacetime manifold. Effectively this is what we just did.

To complete our Penrose diagram we must locate the event horizon. To do so we must specify the mass function, roll up our sleeves, and solve Eq. (5.9) exactly. The easy case is

$$m = \begin{cases} 0 & \text{if} \quad v < 0 \\ \mu v & \text{if} \quad 0 < v < M/\mu \\ M & \text{if} \quad v > M \end{cases} \tag{5.24}$$

where μ is a positive constant. From our analysis of Eq. (5.9) we already know that the singularity is locally naked if $\mu \leq 1/16$. The linear mass function is distinguished since the Vaidya solution then admits the homothetic Killing field

$$\vec{\eta} = v\partial_v + r\partial_r \quad \Rightarrow \quad \mathcal{L}_{\vec{\eta}} g_{ab} = 2g_{ab} \ . \tag{5.25}$$

A *homothety* means that if you scale things up, everything remains the same. Self-similar spacetimes are rather special, and it is somewhat dangerous to draw general conclusions from them. But the extra symmetry leads to soluble equations, and this is irresistible. For the Vaidya solution we note that

$$\eta_a \eta^a = 2\mu v r \left(\left(x - \frac{1}{4\mu} \right)^2 + \frac{1}{\mu^2} \left(\mu - \frac{1}{16} \right) \right) , \tag{5.26}$$

where x is the dimensionless variable introduced above, in Eq. (5.20). The homothetic Killing vector is spacelike in the entire Vaidya region provided that $\mu > 1/16$, that is whenever the solution is free of locally naked singularities. In the Minkowski region it is timelike, and in the Schwarzschild region it does not exist.

The problem of finding the path of the null geodesics in the Vaidya region has now been reduced to that of solving the linear system (5.19). Or we can address Eq. (5.9) directly.[3] In the self-similar case, once we have rewritten it in terms of the dimensionless variable x, it becomes

$$r \frac{dx}{dr} = \frac{1}{8\mu} \frac{(4\mu x - 1)^2 + 16\mu - 1}{1 - 2\mu x} . \tag{5.27}$$

This can be solved using separation of variables, and routine calculation. Having done so we can continue the event horizon from the Schwarzschild part of spacetime (where it sits at $r = 2M$) back in time along null geodesics through the Vaidya region. Due to spherical symmetry its spatial cross-sections are all the time round spheres. When the null rays forming it emerge into Minkowski space on the other side of the shell, they form a round lightcone with its tip a point somewhere in flat space. This is the point where the event horizon first begins to form, in anticipation of the disaster that is going to befall spacetime in the future.

We have arrived at Figure 5.3. In drawing it we assumed that $\mu > 1/16$, so we are in the harmless case when no geodesics emerge from the origin of the $v - r$-diagram. Thus strong cosmic censorship holds. This is gravitational collapse, as we expect it to be.

To find the location of the event horizon we start in the Schwarzschild region, where we know it—it is the inner boundary of the region that can be seen from infinity.[4] In our diagrams it is represented by a past directed radial null geodesic. In the $v - r$-diagram it is a part of a vertical line. We match it to that null geodesic in the Vaidya region which is momentarily vertically directed at the value of v where the matching is made. When there are no globally naked singularities this null geodesic can be continued into the flat spacetime region, so a part of flat spacetime will be inside the event horizon. This flat region is the intersection of the interiors of a backwards lightcone (the inner boundary of the shell) and a forwards lightcone (the event horizon). In other words it is a *causal diamond*, and can be drawn without conformal distortion—the singularity sits at the top of the diamond and is situated at

[3] A. Papapetrou, *Formation of a singularity and causality*, in N. Dadhich et al. (eds): A Random Walk in Relativity and Cosmology, Wiley 1985.

[4] Numerical relativists consider much more complicated situations, and moreover they do not have "infinity" represented in their data. But the idea is the same. See R. A. Matzner et al., *Geometry of a black hole collision*, Science **270** (1995) 941.

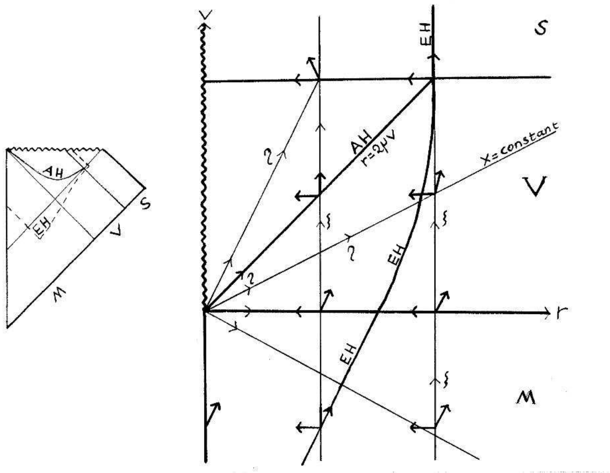

Figure 5.3. Since we use Eddington-Finkelstein coordinates for all calculations we use a $v - r$-diagram for visualization. The picture is for $\mu = 1/2$ (a clothed singularity) and includes the homothetic Killing and the Kodama vector fields, selected future light cones, the event horizon EH, and the spacelike hypersurface $r = 2m$, here referred to (with a slight abuse of terminology) as the apparent horizon AH.

finite timelike distance from any point on the axis. Now consider the time an observer can spend within this causal diamond, which is the flat part of the interior of the black hole. This time will grow the larger the mass, and the faster it comes in—because the event horizon then lies further out. Then the observer will live longer inside the black hole, but if she follows the central world line in the diagram she will never notice the incoming matter—until she is suddenly killed.

All this makes perfect sense once one realizes that the event horizon is an "upside down" concept. Its location is not determined by what has happened, it is determined by what will happen. In particular its area grows quickly in Minkowski space, its rate of growth drops when the incoming radiation crosses it, and then the area stays constant in the Schwarzschild region.

Some further structure was added to Figure 5.3. In every spherically symmetric spacetime we define the *Kodama vector field* $\vec{\xi}$ as being orthogonal to the round spheres (that have area $4\pi r^2$, whatever coordinate system we are using), and such that

$$\xi^a \nabla_a r = 0 \ , \quad \text{and} \quad \xi_a \xi^a = -g^{ab} \nabla_a r \nabla_b r \ . \tag{5.28}$$

The Kodama vector field points in the direction in which the area of the round spheres is unchanged. In the Schwarzschild spacetime it coincides with a Killing vector field. But, because the area radius r of the round spheres has a clear geometrical meaning, the Kodama vector field remains interesting also in the presence of spherically symmetric matter. We notice that it manages

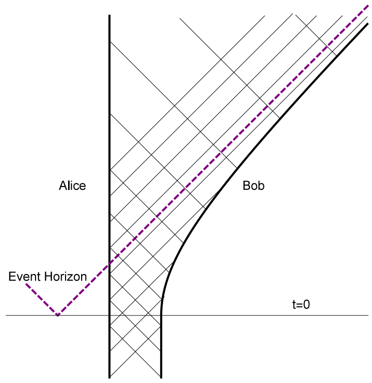

Figure 5.4. Alice and Bob are sending signals to each other, at regular proper time intervals. At $t = 0$ an event horizon forms. The prescient Bob immediately starts accelerating at a constant rate in order to avoid it. Local measurements in his spaceship suggest he is now in a constant gravitational field. The signals sent by Alice before she crosses the horizon arrive with increasing redshift, the ones sent later not at all. The signals sent by Bob arrive with modest redshift.

to stay timelike also in a part of the interior of the expanding event horizon. It becomes null on a hypersurface that we call the *apparent 3-horizon*. In the Schwarzschild part of spacetime the apparent 3-horizon coincides with the event horizon, but unlike the event horizon the apparent 3-horizon never enters Minkowski space. In the Vaidya part it is a spacelike hypersurface.

What is special about the apparent 3-horizon? Imagine that all the round spheres in the solution emit two flashes of light, one directed outwards and one directed inwards. If the Kodama vector field is spacelike, all available null directions point in the direction of shrinking round spheres. Thus you can conclude from Figure 5.3 that both of the wave fronts always shrink in area once you are inside the apparent 3-horizon. We conclude that the apparent 3-horizon is the boundary of the region where round trapped spheres exist. It is foliated by *marginally trapped* round spheres. We will return to the study of apparent horizons and trapped spheres—round or not round—in Chapter 9.

An important difference between the apparent and the event horizon is that the location of the former, on some spacelike hypersurface, can be deduced without knowing anything about the future. The event horizon is a very different kettle of fish. In the flat part of the Vaidya solution it is just an ordinary lightcone. (It is consistent with everything we know about the Universe that we live inside a converging spherically symmetric shell of collapsing null dust. If so, you may be passing the event horizon at the very moment that you read this.) It is interesting to consider the exchange of signals between an observer crossing the event horizon in free fall, and another that accelerates at a constant rate in order to avoid it. The latter may continue to receive increasingly redshifted signals from the former for all eternity, but they were all emitted before she passed the horizon. See Figure 5.4—which plays out in a part of flat spacetime.

If the mass function $m(v)$ grows slowly enough, in the self-similar case if $\mu \leq 16$, Figure 5.3 no longer applies. Null geodesics would come out of the singularity at the origin of the $v-r$-diagram. Then a Cauchy horizon appears, and cosmic censorship fails. In the self-similar case the singularity is in fact globally naked. On the face of it, this is serious: we now know for a fact that the ingredients giving rise to singularities—in particular inequalities imposed on the Ricci tensor via energy inequalities—do not suffice to cover them up. However, the outcome is not a disaster. We can try to blame it on spherical symmetry, which is a non-generic situation. We can also try to blame it on the matter model, which is not physically realistic at high densities. Indeed, because a singularity in the form of infinite dust density would occur also in a flat background, its nakedness is usually dismissed as a pathology of the matter model rather than as a threat to cosmic censorship.

Let us go back to Figure 5.2 once more. It is important that each point in the Penrose diagram is a sphere—except the dotted line which describes the spatial origin. The null part of the singularity is created as the incoming matter shells converge to that point. Such a singularity is referred to as a *shell focusing singularity.* According to our model the story ends there, but in the real world something else would happen—the shells might continue through there in a collisionless manner, or they would rebound inelastically. Hence our description there is quite suspect. The spacelike part of the singularity—if you like, the part that was successfully censored—is a different story. There the matter shells remain spherical shells, and the problem is that due to spacetime curvature the areas of the spheres are shrinking to zero.

If we reflect on our results we might reason as follows: in a fixed flat spacetime the collapse would lead to a singular mass density visible from afar. When the backreaction on the metric is taken into account the strong cosmic censor hides this singularity—but only provided the departure from flat space is strong enough. Intuitively this is a very reasonable conclusion, and it rather supports the idea that there is a mechanism that tends to hide singularities.

This ends our analysis of naked and clothed singularities in the Vaidya solution. Another exact solution in the same vein is the *Oppenheimer–Snyder* solution. To obtain it, one does not have to solve any differential equation at all. One starts from two known spherically symmetric solutions, the Schwarzschild solution, and the closed Friedmann model describing an evolving 3-sphere filled with matter in the form of dust moving on timelike geodesics. Consider initial data at $t = 0$ in the former, and at the moment of maximal expansion in the latter. Recall Figure 4.2. Glue suitable parts of these spaces together, to obtain something that looks like a badminton ball. These initial data will evolve into a spacetime that can be taken as a model of a collapsing spherically symmetric homogeneous star (where pressure is overwhelmed by the gravitational attraction, and ignored). The full solution can be obtained by matching the two solutions together across a timelike hypersurface ruled by timelike geodesics, in the manner suggested by the Penrose diagram in Figure 5.5. However, as you may recall from your electrodynamics course, matching two solutions describing two different media together requires a certain amount of delicate

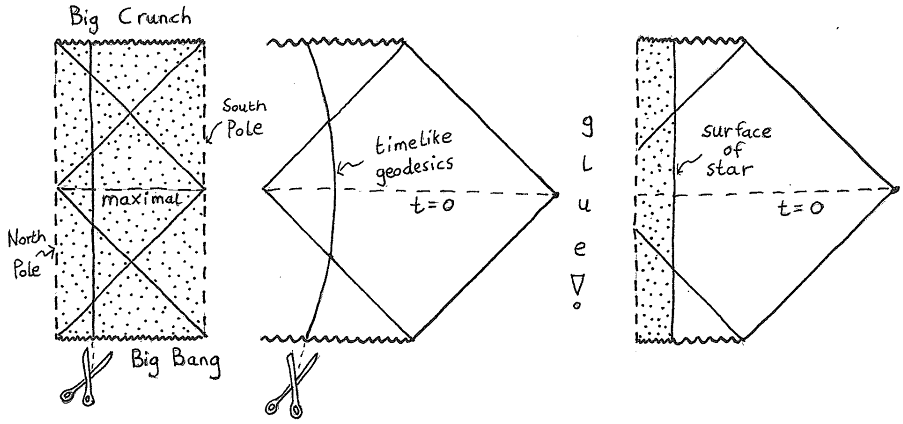

Figure 5.5. The Penrose diagram of a dust filled Friedmann model is at the left. It describes expanding and recontracting 3-spheres, and the dust particles follow timelike geodesics (directed vertically in the diagram). After cutting and gluing, a part of this spacetime can be joined to a Schwarzschild exterior to give a model of a collapsing star.

handling, and we will not have the means to discuss it until we have gone through Chapter 7.

At the end of a long path starting from here stands a theorem, due to Christodoulou, which supports cosmic censorship in spherically symmetric models with a massless scalar field in the role of matter.[5] As a matter model, the massless scalar field is above suspicion. But the restriction to spherical symmetry in itself defines a toy model of the real thing, so this does not settle the question. As soon as spherical symmetry is dropped there will be gravitational radiation filling some part of the region outside the event horizon.

◇ **Problem 5.1** Express the stress-energy tensor of the Vaidya solution in terms of an ON basis (including one timelike vector). Show that it cannot be diagonalized by means of Lorentz transformations.

◇ **Problem 5.2** Send in two self-similar Vaidya shells, one after the other with some spacetime in between. Set $\mu = 1/2$ and draw an exact picture like that of Figure 5.3, showing the event horizon as well as the location of the marginally trapped round spheres.

◇ **Problem 5.3** The Friedmann model in Figure 5.5 has the metric

$$ds^2 = a^2\left(-d\eta^2 + d\chi^2 + \sin^2\chi(d\theta^2 + \sin^2\theta d\phi^2)\right) \ , \qquad a = \frac{a_{\rm m}}{2}(1+\cos\eta) \ . \tag{5.29}$$

Consider round spheres centred at the north poles of the 3-spheres. When are they trapped?

[5] For a review see D. Christodoulou, *On the global initial value problem and the issue of singularities*, Class. Quant. Grav. **16A** (1999) 23.

6 The Kerr spacetime

The Kerr solution provides an exact description of the most general isolated stationary black hole that can exist in our Universe. The only scientific result that can be compared to it, so far, is Newton's solution of the non-relativistic gravitational two-body problem.

I will simply present the solution in a reasonable looking coordinate system, and then explain how one gradually understands its geometry.[1] But to give the story away: the solution will depend on two parameters m and a. Their eventual interpretation will be that $M = m$ is the total mass of this spacetime, and $J = am$ is its total angular momentum. If $a \leq m$ the solution describes a spinning black hole, and if $a > m$ it describes a nakedly singular spacetime. The surprise is that this two-parameter family of metrics describes the most general stationary black hole that can occur as a solution of the Einstein's vacuum equations, and that this black hole is stable under small perturbations. A full proof of non-linear stability is an open problem that—rumour has it—is close to its solution.

We begin by simply writing down the Kerr solution. In Boyer-Lindquist coordinates it is

$$ds^2 = -\frac{\Delta}{\rho^2}(dt - a\sin^2\theta d\phi)^2 + \frac{\sin^2\theta}{\rho^2}((r^2+a^2)d\phi - adt)^2 + \frac{\rho^2 dr^2}{\Delta} + \rho^2 d\theta^2 \quad (6.1)$$

where

$$\rho^2 \equiv r^2 + a^2\cos^2\theta \ , \quad (6.2)$$

$$\Delta \equiv r^2 - 2mr + a^2 = (r - r_+)(r - r_-) \ , \qquad r_\pm = m \pm \sqrt{m^2 - a^2} \ . \quad (6.3)$$

There are two free parameters m and a. The most interesting case is that of $a^2 < m^2$, because $\Delta = 0$ then has two real roots $r_\pm$. We concentrate on this case from now on.

If we introduce

[1] A very readable book is B. O'Neill: *The Geometry of Kerr Black Holes*, A K Peters, Wellesley, Massachusetts 1995.

$$F \equiv (r^2 + a^2)^2 - a^2\Delta \sin^2\theta = (\Delta + 2mr)^2 - a^2\Delta \sin^2\theta \tag{6.4}$$

the metric can also be written as

$$ds^2 = -\frac{\Delta - a^2\sin^2\theta}{\rho^2}dt^2 - \frac{4mar\sin^2\theta}{\rho^2}dtd\phi + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + \frac{F\sin^2\theta}{\rho^2}d\phi^2 \ , \tag{6.5}$$

The contravariant form of this metric is

$$\frac{\partial^2}{\partial s^2} = -\frac{F}{\Delta\rho^2}\partial_t^2 + \frac{\Delta}{\rho^2}\partial_r^2 + \frac{1}{\rho^2}\partial_\theta^2 + \frac{\Delta - a^2\sin^2\theta}{\Delta\rho^2\sin^2\theta}\partial_\phi^2 - \frac{4Mar}{\Delta\rho^2}\partial_t\partial_\phi \ . \tag{6.6}$$

From the expression for g_{rr} we see that the range of the coordinate r can be (r_+, ∞), (r_-, r_+), or $(-\infty, r_-)$. When $r \to \infty$ the solution tends to flat spacetime, and resembles the Schwarzschild solution. When $r \to -\infty$ the solution again tends to flat spacetime, but here it resembles the unphysical negative mass Schwarzschild solution. We refer to these regions as (in order) region I, II, and III:

Region I: $r_+ < r < \infty$
Region II: $r_- < r < r_+$
Region III: $-\infty < r < r_-$

The solution can be analytically extended by adding more copies of these regions—much like the Reissner-Nordström solution discussed in Chapter 4—but we will rest content if we understand the regions we already have.

There is only one cross-term in the metric, but it is still sufficiently complicated that a computer algebra system is needed to proceed.[2] The Kretschmann scalar looks a little involved, but using the dual of the Weyl tensor (see Chapter 10) we find the complex scalar quantity

$$C_{abcd}C^{abcd} + iC^{\star}_{abcd}C^{abcd} = \frac{48m^2}{(r - ia\cos\theta)^6} \ . \tag{6.7}$$

Hence there is a genuine curvature singularity at $(r, \theta) = (0, \pi/2)$. On the other hand the surfaces $\Delta = 0$ are (presumably) coordinate singularities only, and curves can pass through the timelike hypersurface $r = 0$ if they avoid the equator. Before jumping to any conclusion, note that the stability results which are the pride of the Kerr solution apply to the late time behaviour outside the event horizon, and not at all to its exotic interior.

There are two Killing vectors, ∂_t^a and ∂_ϕ^a, and as a matter of fact every Killing vector is linearly dependent on those two. For large values of r the

[2] Unless, like Kerr, you do not have one. Then smoking will do the trick. See R. P. Kerr, *Discovering the Kerr and Kerr–Schild metrics*, in D. L. Wiltshire et al. (eds.): The Kerr Spacetime: Rotating Black Holes in General Relativity, Cambridge UP, 2009.

former is timelike and the latter spacelike, so that the solution is stationary—it admits a timelike Killing vector field—and axially symmetric—the flow lines of the spacelike Killing vector field are closed. Using $t_a = g_{ab}\partial_t^b$ it is easy to check that

$$t_{[a}\nabla_b t_{c]} = 0 \quad \Leftrightarrow \quad a = 0 \ . \tag{6.8}$$

According to Frobenius' theorem this means that the timelike Killing vector field is hypersurface orthogonal if and only if $a = 0$, that is in the Schwarzschild case. Unlike the Schwarzschild spacetime the Kerr spacetime does not admit a natural split into space and time, which is one reason why it is significantly harder to understand the latter.

We observe that

$$||\partial_t||^2 = -\frac{\Delta - a^2\sin^2\theta}{\rho^2} \ , \qquad ||\partial_\phi||^2 = \frac{F\sin^2\theta}{\rho^2} \ . \tag{6.9}$$

The first of these becomes null on two hypersurfaces defined by

$$r^2 - 2mr + a^2\cos^2\theta = 0 \quad \Leftrightarrow \quad r = m \pm \sqrt{m^2 - a^2\cos^2\theta} \ . \tag{6.10}$$

The normal of this hypersurface is

$$\begin{gathered} n_a = \nabla_a r \mp \frac{a^2\cos\theta\sin\theta}{\sqrt{m^2 - a^2\cos^2\theta}}\nabla_a\theta \\ \Rightarrow \\ ||n||^2 = \frac{1}{\rho^2}\left(\Delta + \frac{a^4\cos^2\theta\sin^2\theta}{m^2 - a^2\sin^2\theta}\right) \ . \end{gathered} \tag{6.11}$$

As long as $a \neq 0$ this is a spacelike vector, hence the hypersurface is timelike. (Except if $\sin\theta = 0$, in which case the hypersurface touches the otherwise distinct hypersurface $\Delta = 0$.) It is called the *ergosphere* of the black hole. For now we just observe that inside the ergosphere (or rather in between the two ergospheres) the Killing vector ∂_t is spacelike.

We still need to know in what region of the solution there exists a timelike linear combination of the two Killing vector fields. To this end we consider the Killing bivector

$$\kappa^{ab} = \partial_t^a\partial_\phi^b - \partial_t^b\partial_\phi^a \ . \tag{6.12}$$

At each point of spacetime the Killing bivector spans a 2-plane element, and together these 2-plane elements can be fitted together into surfaces of constant r and θ. In Boyer-Lindquist coordinates the only independent non-vanishing component of κ_{ab} is

$$\kappa_{t\phi} = g_{tt}g_{\phi\phi} - g_{t\phi}^2 = -\Delta\sin^2\theta \ . \tag{6.13}$$

By the way, the condition that the Killing bivector is surface forming is

$$\kappa_{[ab}\nabla_c \partial_{t]} = \kappa_{[ab}\nabla_c \partial_{\phi]} = 0 \ , \tag{6.14}$$

and is obviously satisfied. Indeed this is the reason why this coordinate system works. Of more importance at the moment is the norm of the bivector, namely

$$\frac{1}{2}\kappa_{ab}\kappa^{ab} = -\Delta \sin^2\theta \ . \tag{6.15}$$

The 2-plane is timelike if the above quantity is negative, and spacelike if it is negative. Any vector in such a 2-plane is a Killing vector, so we conclude that there are timelike Killing vectors throughout the region where $\Delta > 0$. At $\Delta = 0$ we have—or will have, as soon as we have introduced a coordinate system that permits us to include it in our spacetime—a null hypersurface. This clearly suggests that the hypersurface $\Delta = 0$ should be the event horizon of the Kerr black hole, so this is what we will call it from now on.

It is easy to find a Killing vector field that becomes null on the horizon, and is timelike in a region outside it. It is

$$\xi_{\rm hor} = \partial_t + \frac{a}{2mr_+}\partial_\phi \ , \tag{6.16}$$

$$||\xi_{\rm hor}||^2 = -\frac{r-r_+}{4m^2\rho^2 r_+^2} f \ , \tag{6.17}$$

$$\begin{aligned} f = &\left(4m^2(2mr_+ - a^2) - a^2(r^2+2mr+a^2)\sin^2\theta + a^4\sin^4\theta\right)(r-r_-) - \\ &-4m^2a^2\sin^2\theta(r-r_+) \ . \end{aligned} \tag{6.18}$$

Hence the hypersurface $\Delta = 0$ is ruled by a null Killing field, and therefore it is a Killing horizon—as it should be for the event horizon of a stationary black hole. The horizon Killing vector field is also null whenever $f = 0$. This defines a timelike surface outside the horizon, known as the velocity-of-light surface (because an observer corotating with the event horizon would have to follow this Killing vector field). In the extremal case $a = M$ there is a complication, namely that in a region around the equator the horizon Killing vector field is actually spacelike outside the horizon itself.

It is somewhat worrying that ∂_ϕ^a, which has closed flow lines, also can change its causal character:

$$||\partial_\phi||^2 = \frac{F\sin^2\theta}{\rho^2} \ . \tag{6.19}$$

This means that there are closed timelike curves whenever

$$F = r^4 + a^2\cos^2\theta r^2 + 2ma^2\sin^2\theta r + a^4\cos^2\theta < 0 \ . \tag{6.20}$$

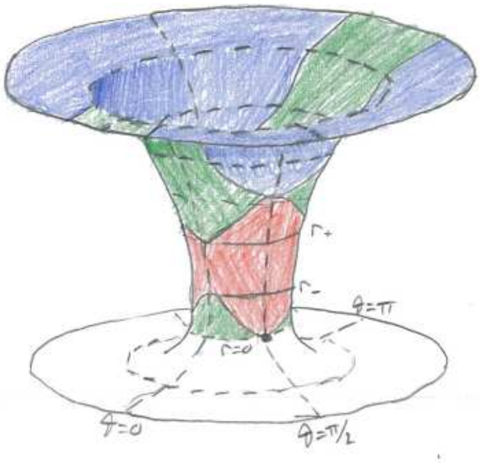

Figure 6.1. Carter's view of the three blocks of the Kerr solution, with colouring added by Helgi Freyr. The ergoregion is in red, and the horizon Killing vector field is null at the boundary of blue and green (and the value of a is such that that these surfaces happen to touch).

However, this can only happen for negative r, and then only close to the singularity, so we may perhaps ignore it. Elsewhere, in particular when $r > r_+$, the coordinate t serves as a time function. The point is that hypersurfaces of constant t have a normal vector $\nabla_a t$, with norm

$$g^{ab}\nabla_a t \nabla_b t = g^{tt} = -\frac{F}{\rho^2\Delta} \ . \tag{6.21}$$

The normal is timelike when $F > 0$. In this region these hypersurfaces are spacelike, and the function t is monotoneously increasing function along any timelike curve. This can be used to show that there are no closed timelike curves in the region where $r > r_+$, nor in region II of the solution. On the other hand there are closed timelike curves through every point in region III, basically because starting from any point there one can send a timelike curve into the region where $F < 0$, spiral back in time, and then return to the starting point along another timelike curve. This rather underscores the unphysical nature of region III.

We are now in position to draw a picture of the Kerr solution. It is impossible to improve on the one drawn by Brandon Carter, so I simply copy it.[3] It gives a rather faithful picture of the submanifold parametrized by the coordinates r and θ. To complete it you mentally add the ignorable coordinate directions parametrized by t and ϕ. The picture should not be confused with the usual picture of the Einstein-Rosen bridge at constant time t in the Schwarzschild solution. The latter also has two asymptotic regions, but is com-

[3] See the famous review by B. Carter, *Black Hole equilibrium states. Part I: Analytic and geometrical properties of the Kerr solutions*, in C. DeWitt and B. DeWitt (eds.): Black Holes—les astres occlus, Gordon and Breach, New York 1973.

pletely symmetric between the two. The lower asymptotic region in Figure 6.1 is the unphysical region III, which has no Schwarzschild counterpart.

Although strictly speaking the coordinate system we are using does not cover the event horizon it can still be used to identify its intrinsic metric. More exactly the metric on the set of its generators—which are radial outgoing geodesics—is

$$
\begin{aligned}
d\gamma^2 &= (r_+^2 + a^2\cos^2\theta)d\theta^2 + \frac{(r_+^2+a^2)^2\sin^2\theta}{r_+^2+a^2\cos^2\theta}\left(d\phi - \frac{adt}{r_+^2+a^2}\right)^2 = \\
&= (r_+^2 + a^2\cos^2\theta)d\theta^2 + \frac{(r_+^2+a^2)^2\sin^2\theta}{r_+^2+a^2\cos^2\theta}d\tilde{\phi}^2 \ ,
\end{aligned}
\tag{6.22}
$$

where we introduced a new angular coordinate $\tilde{\phi}$ in an obvious way. The area of a cross-section is easy to compute from the determinant $\det\gamma$, and it is

$$
A = \int_S (r_+^2 + a^2)d\Omega = 4\pi(r_+^2 + a^2) = 4\pi(2m^2 + 2m\sqrt{m^2 - a^2}) \ . \tag{6.23}
$$

Its Gaussian curvature is

$$
k = \frac{\bar{R}}{2} = \frac{(r_+^2 + a^2)(r_+^2 - 3a^2\cos^2\theta)}{(r_+^2 + a^2\cos^2\theta)^3} \ . \tag{6.24}
$$

As a curiousity we observe that this becomes negative at the poles if the spin is high enough, namely if $a/m > \sqrt{3}/2$.

More to the point, let us identify $M = m$, $J = am$, and also define the *entropy* of the black hole as $S = A/4\pi$. Then we obtain the fundamental thermodynamical relation of a Kerr black hole,

$$
S = 2M^2\left(1 + \sqrt{1 - \frac{J^2}{M^4}}\right) \quad \Leftrightarrow \quad M = \frac{\sqrt{S}}{2}\sqrt{1 + \frac{4J^2}{S^2}} \ . \tag{6.25}
$$

The reason why one talks about thermodynamics here becomes clear after taking the differential of M,

$$
dM = TdS + \Omega dJ \ . \tag{6.26}
$$

The explicit expressions for T and Ω are

$$
T = \frac{1}{8M}\left(1 - \frac{4J^2}{S^2}\right) \ , \qquad \Omega = \frac{J}{MS} \ . \tag{6.27}
$$

One can check that T equals the Hawking temperature $T_H = \kappa/2$, where κ is the surface gravity of the horizon.[4] The quantity thermodynamically conjugate to J,

$$\Omega = \frac{J}{MS} = \frac{a}{2mr_+} \ , \tag{6.28}$$

is known as the *angular velocity of the horizon* because Eq. (6.16) tells us that this is the rate at which the horizon generators spin with respect to the canonical time direction at infinity, where ∂_t^a becomes orthogonal to the closed flow lines generated by ∂_ϕ^a.

To increase familiarity we introduce a pair of additional coordinate systems on the Kerr spacetime. They are somewhat analogous to the Eddington–Finkelstein coordinates, and are here called Kerr coordinates. We introduce new coordinates v and $\tilde{\phi}$, such that

$$dv = dt + \frac{r^2+a^2}{\Delta}dr \qquad d\tilde{\phi} = d\phi + \frac{a}{\Delta}dr \ . \tag{6.29}$$

We do not need $v = v(t,r)$ or $\tilde{\phi} = \tilde{\phi}(\phi, r)$ explicitly. It is enough that these functions exist. Then

$$\begin{aligned} ds^2 = -\left(1 - \frac{2mr}{\rho^2}\right) dv^2 + 2dvdr + \rho^2 d\theta^2 + \frac{F\sin^2\theta}{\rho^2}d\tilde{\phi}^2 - \\ -2a\sin^2\theta dr d\tilde{\phi} - \frac{4amr}{\rho^2}\sin^2\theta dv d\tilde{\phi} \ . \end{aligned} \tag{6.30}$$

Now there are three cross terms in the metric. The inverse metric is

$$\begin{aligned} \partial_s^2 = \frac{1}{\rho^2}\Big(a^2\sin^2\theta\partial_v^2 + \Delta\partial_r^2 + \partial_\theta^2 + \frac{1}{\sin^2\theta}\partial_{\tilde{\phi}}^2 + \\ +2(r^2+a^2)\partial_v\partial_r + 2a\partial_v\partial_{\tilde{\phi}} + 2a\partial_r\partial_{\tilde{\phi}}\Big) \end{aligned} \tag{6.31}$$

These coordinates are called ingoing Kerr coordinates, and indeed they cover the event horizon at $\Delta = 0$. It will be observed that

$$a \neq 0 \quad \Rightarrow \quad g^{ab}\nabla_a v \nabla_b v = g^{vv} > 0 \tag{6.32}$$

(except at the poles). Hence v is not a null coordinate except in the Schwarzschild case. The hypersurfaces $v =$ constant do approach null surfaces at large values of r though, that is to say far away from the black hole.

It is still true that the curves defined by keeping $v, \theta, \tilde{\phi}$ constant are null geodesics, with the coordinate r serving as an affine parameter along the rays. (The geodesic equation reduces to $\ddot{r} = 0$.) They form a space-filling congruence

[4] The slight discrepancy with Eq. (4.34) is due to a new choice of Boltzmann's constant, $k = 1/\pi$.

of null geodesics that do not develop any caustics as they fall in towards the centre. This is a remarkable property of the solution—and indeed provided Roy Kerr with the handle needed to discover it in the first place. As an examination of Eqs. (6.29) shows, the values of t and $\tilde{\phi}$ diverge at the event horizon $\Delta = 0$. If an observer at infinity could watch the progress of these null geodesics she would have to wait an infinitely long time to see them reach the horizon, and meanwhile she would see them wind around an infinite number of times.

There is also an outgoing variant of the Kerr coordinates. It is

$$du = dt - \frac{r^2 + a^2}{\Delta}dr \qquad d\tilde{\phi} = d\phi - \frac{a}{\Delta}dr \ . \tag{6.33}$$

The "untwisting" of the angular coordinate is now in the opposite direction. Then

$$\begin{aligned} ds^2 = -\left(1 - \frac{2Mr}{\rho^2}\right)du^2 - 2dudr + \rho^2 d\theta^2 + \frac{F\sin^2\theta}{\rho^2}d\tilde{\phi}^2 - \\ +2a\sin^2\theta dr d\tilde{\phi} - \frac{4aMr}{\rho^2}\sin^2\theta du d\tilde{\phi} \ . \end{aligned} \tag{6.34}$$

This is the coordinate system in which Kerr first presented the solution.

For many reasons, not least those connected to astrophysics, it is important to understand the behaviour of geodesics in the Kerr spacetime. If we confine ourselves to the equatorial plane $\theta = \pi/2$ (where the accretion disk lies!) this is a straightforward albeit lengthy exercise. We have two constants of the motion

$$E = -u_a\partial_t^a = (1 - 2m/r)\dot{t} + \frac{2ma}{r}\dot{\phi} \ , \qquad L = u_a\partial_\phi^a = -\frac{2ma}{r}\dot{t} + \frac{F}{r^2}\dot{\phi} \ . \tag{6.35}$$

They represent energy and angular momentum per unit rest mass. We also have that

$$(1 - 2m/r)\dot{t}^2 + \frac{4ma}{r}\dot{t}\dot{\phi} - \frac{r^2}{\Delta}\dot{r}^2 - \frac{F}{r^2}\dot{\phi}^2 = k = \begin{cases} 1 \\ 0 \\ -1 \end{cases} , \tag{6.36}$$

depending on whether the geodesic is timelike, null, or spacelike. Solving for $\dot{t}$ and $\dot{\phi}$, and inserting, we obtain an equation for r that describes one dimensional motion in an effective potential, namely

$$\frac{\dot{r}^2}{2} + V_{\text{eff}}(r; E, L) = 0 \ , \tag{6.37}$$

where

$$V_{\text{eff}}(r) = -\frac{km}{r} + \frac{L^2}{2r^2} + \frac{k - E^2}{2}\left(1 + \frac{a^2}{r^2}\right) - \frac{m}{r^3}(L - aE)^2 \ . \tag{6.38}$$

From this much useful information can be extracted. For the Schwarzschild black hole the innermost stable circular orbit, the ISCO, lies at $r = 6m$. This is interpreted as the location of the inner edge of the accretion disk. The binding energy of a particle orbiting at this radius is about 6 % of its rest mass, meaning that this is the amount of energy that went out as radiation as the particle spiralled in to this radius. (You can read this number out of Eq. (6.47) below). For the Kerr black hole the location of the ISCO depends on the sign of L, that is on whether the particle is co- or counter-rotating with respect to the event horizon. It comes closer in the corotating case, and in fact $r_{\text{ISCO}} = m$ in the extremal case $a = m$, and then the binding energy there is about 42 %. (This does not mean that the ISCO actually lies on the event horizon however. In the extremal case the coordinate r fails badly there, and the event horizon lies infinitely far away from any exterior point.) Thus letting a particle spiral in towards a spinning black hole is a very effective way of extracting the energy hidden in $E = mc^2$. Moreover the radius of the inner edge of an accretion disk provides a signature that allows us to estimate the spin of the black hole.

Remarkably, exact solutions can be obtained also for geodesics out of the equatorial plane. The two constants of the motion contributed by the Killing vectors do not suffice for this, but it turns out that an extra constant of the motion, known as Carter's constant, is present and saves the day. This comes about because the Kerr spacetime admits a non-trivial Killing tensor, that is to say a tensor field K_{ab} such that

$$\nabla_{(a}K_{bc)} = 0 \ . \tag{6.39}$$

It is easily checked that this means that the quantity

$$K = K_{ab}u^a u^b \tag{6.40}$$

will be constant along any geodesic. As a result the Hamilton-Jacobi equation for the geodesic motion is separable, and the geodesic can be solved for.

Circular orbits in the equatorial plane are the most important ones, and we will analyze them in more detail.[5] A circular orbit will exist at values of r such that $V_{\text{eff}}(r) = V'_{\text{eff}}(r) = 0$. Stability of the orbit also requires that $V''_{\text{eff}}(r) \geq 0$. To handle these equations it is convenient to define $V = 2r^2 V_{\text{eff}}$, and begin with

$$V(r) + rV'(r) = -4mr + L^2 + (1 - E^2)(3r^2 + a^2) = 0 \ . \tag{6.41}$$

This can be solved for E. If the result is inserted in the equation $V'(r) = 0$ a fourth order equation for L results, which has two real roots. Using "computer-assisted algebraic techniques" Bardeen, Press, and Teukolsky found the solutions

[5] Following J. M. Bardeen, W. H. Press, and S. A. Teukolsky, *Rotating black holes: Locally non-rotating frames, energy extraction, and scalar synchrotron radiation*, Astrophys. J. **178** (1972) 347.

$$E = \frac{r^2 - 2mr \pm a\sqrt{mr}}{r\sqrt{r^2 - 3mr \pm 2a\sqrt{mr}}} \tag{6.42}$$

$$L = \pm\frac{\sqrt{mr}(r^2 \mp 2a\sqrt{mr} + a^2)}{r\sqrt{r^2 - 3mr \pm 2a\sqrt{mr}}} \ . \tag{6.43}$$

The upper sign is for the corotating case. The condition for stability is easily obtained by taking one more derivative of the left hand side of Eq. (6.41). It is

$$1 - E^2 \geq \frac{2m}{3r} \ . \tag{6.44}$$

If equality holds we are at the ISCO. Inserting Eq. (6.42) for E leads to

$$r^2 - 6mr \pm 8a\sqrt{mr} - 3a^2 = 0 \ . \tag{6.45}$$

This is a quartic equation for $\sqrt{r}$, and can be solved. Alternatively, we can solve it as a quadratic equation for a. In the corotating case

$$a = \frac{\sqrt{mr}}{3}\left(4 - \sqrt{\frac{3r}{m} - 2}\right) = m\frac{\sqrt{z}}{3}(4 - \sqrt{3z-2}) \ . \tag{6.46}$$

Here we introduced the dimensionless variable $z = r/m$. Inserting this solution into Eqs. (6.42-6.43) gives

$$E_{\text{ISCO}} = \sqrt{1 - \frac{2m}{3r}} = \frac{\sqrt{3z-2}}{\sqrt{3z}} \tag{6.47}$$

$$L_{\text{ISCO}} = \frac{2m}{3\sqrt{3}}\left(1 + 2\sqrt{\frac{3r}{m} - 2}\right) = \frac{2m}{3\sqrt{3}}(1 + 2\sqrt{3z-2}) \ . \tag{6.48}$$

Bardeen put Eqs. (6.46-6.48) to an interesting use.[6]

Namely, suppose that the black hole is surrounded by a corotating accretion disk, and also suppose that particles slowly drift inwards, so that there is a soft rain of particles falling from the ISCO into the black hole. The mass $M = m$ and the angular momentum $J = am$ of the black hole will then change. Provided the particles are small enough this will happen in accordance with the differential equation

$$\frac{dJ}{dM} = \frac{L_{\text{ISCO}}}{E_{\text{ISCO}}} \ . \tag{6.49}$$

[6] J. M. Bardeen, *Kerr metric black holes*, Nature **226** (1970) 64. For more details, see J. Bolin, *The Angular Momentum of Kerr Black Holes*, BSc Thesis, SU 2015.

If we introduce the dimensionless variable $a_\star = a/m = J/M^2$ this equation becomes

$$M\frac{da_\star}{dM} = M\frac{dz}{dM}\frac{da_\star}{dz} = \frac{1}{M}\frac{L_{\rm ISCO}}{E_{\rm ISCO}} - 2a_\star \ . \tag{6.50}$$

The use of the dimensionless variable z leads to some quite remarkable cancellations here. After writing it out, using Eqs. (6.46-6.48), we are left with

$$M\frac{dz}{dM} = -2z \qquad \Rightarrow \qquad z = \frac{\rm const}{M^2} \ . \tag{6.51}$$

We choose the integration constant so that $J = 0$ at $M = M_i$, and obtain the solution

$$a_\star = \sqrt{\frac{2}{3}}\frac{M_i}{M}\left[4 - \left(\frac{18M_i^2}{M^2} - 2\right)^{1/2}\right] \ , \qquad M_i \leq M \leq \sqrt{6}M_i \ . \tag{6.52}$$

In this way a Schwarzschild black hole eventually becomes an extreme Kerr black hole—and the process ends there, since the right hand side of Eq. (6.50) vanishes when $z = 1$. There is no statement about the speed of this evolution.

Finally, how fast do we expect real black holes to spin? It depends. The different ways in which black holes can be created apart, there are two main mechanisms affecting the ratio a/m of an existing black hole. Suppose matter approaches from far away in a random fashion. Then one can show that the capture cross section is greater if the angular momentum of the ingoing particles is negative, which means that the black hole tends to spin down. On the other hand the black hole may be accreting matter in an orderly way from the inner edge of its accretion disk, and Bardeen's argument shows that the changes in angular momentum and in mass will be correlated in such a way that a and m will evolve towards the extreme limit—although in reality one will not actually reach the extreme limit. Not all the matter will be accreted from the inner edge of the disk, and there are many complications including radiation emitted by the disk, magnetic fields being dragged down the hole, and more.

So it depends. Observations show that black holes with widely varying values of a/m do (seem to) exist.

◇ **Problem 6.1** Calculate $\sqrt{g}$ for the Kerr metric in Boyer–Lindquist coordinates. Also calculate, by hand, at least two non-zero Christoffel symbols. Conclusions?

◇ **Problem 6.2** Use the method of effective potentials to discuss radial timelike geodesics in the Reissner–Nordström solution.

◇ **Problem 6.3** Estimate the magnitude of the dimensionless quantity

$$a_\star = \frac{cJ}{GM^2} \tag{6.53}$$

for a) the Sun, b) the solar system, c) an electron.

7 Differential geometry II

We now continue with differential geometry, focussing on surfaces and hypersurfaces sitting inside an *ambient space*. We define the *codimension* of a submanifold to be the dimension of the full space minus its own dimension. By definition, *hypersurfaces* have codimension 1. In spacetime, surfaces have codimension 2 and curves have codimension 3. Although by now most of the literature on the subject is in English, the language used is diverse when it comes to notation, and my discussion takes this into account.

7.1 Surfaces and hypersurfaces

A general idea is that a manifold $\mathcal{N}$ of dimension n is *embedded* in a manifold $\mathcal{M}$ of dimension $m > n$ by means of a map from $\mathcal{N}$ to $\mathcal{M}$. The number $m - n$ is the codimension. Of course our $\mathcal{M}$ comes equipped with a metric tensor g_{ab}. Figure 2.1 then ticks in, and tells us that g_{ab} will be pulled back to $\mathcal{N}$ and give rise to a tensor there. It goes under the traditional name of the *first fundamental form*, and we denote it by γ_{ab}. If the submanifold $\mathcal{N}$ is provided with coordinates u^i we can make things more concrete. The map, when given in terms of coordinates, is provided by the parametric representation

$$\mathcal{N} \to \mathcal{M} \ : \qquad x^{\mathbf{a}} = x^{\mathbf{a}}(u) \ . \tag{7.1}$$

There are n distinct coordinates u^i, and as many coordinate vector fields on $\mathcal{N}$. The map pushes them forwards to n vector fields on $\mathcal{M}$,

$$e_i^{\mathbf{a}}\partial_{\mathbf{a}} = \frac{\partial x^{\mathbf{a}}}{\partial u^i}\partial_{\mathbf{a}} \ . \tag{7.2}$$

From the point of view of the ambient space this is a set of n vector fields tangent to the embedded submanifold, necessarily linearly independent because we assume that the u^i form an admissible coordinate system. We can also view $e_i^{\mathbf{a}}$ concretely as a rectangular matrix that can be used to perform projections. In particular

$$\gamma_{ij} = e_i^a e_j^b g_{ab} \ . \tag{7.3}$$

At this point the discussion splits into three cases, depending on whether γ_{ij} is positive definite, Lorenzian, or degenerate. The discussion will be quite intricate in any case. The last case includes the null hypersurfaces discussed in Chapter 4, and it will be totally ignored in this chapter.

Once it has been understood that there is a map in the background, pulling various objects back from $\mathcal{M}$ and pushing others forwards from $\mathcal{N}$, we can safely identify the embedded manifold with its image in $\mathcal{M}$. At each point on $\mathcal{N}$ the metric in the ambient space can be used to provide an orthogonal decomposition of its tangent space $\mathbf{T}$. If $\mathbf{T}^{\mathrm{T}}$ stands for the subspace spanned by vectors tangential to $\mathcal{N}$ we get

$$\mathbf{T} = \mathbf{T}^{\mathrm{T}} \oplus \mathbf{T}^{\perp} \ . \tag{7.4}$$

This is the single most important equation in this chapter, so please commit it to memory. Everything in sight will be decomposed in the same way. Unfortunately this will put quite a strain on the notation. We already started the habit of using $a, b, \ldots$ as indices in $\mathbf{T}$ and $i, j, \ldots$ as indices in $\mathbf{T}^{\mathrm{T}}$. Thus the tensor γ_{ij} belongs to the vector space $\mathbf{T}^{\mathrm{T}} \otimes \mathbf{T}^{\mathrm{T}}$. This is still true if I denote the same object by γ_{ab}, but then I have to keep the fact in my head. We could introduce a third set of indices as markers for vectors belonging to $\mathbf{T}^{\perp}$, but this is generally regarded as too much. Actually, at some point below we will plump for the index-free notation used by mathematicians.

Curves were easy because their tangent vector is unique up to normalization. For hypersurfaces there will be three linearly independent tangent vectors, but the normal vector $\vec{n}$ will be unique up to normalization. We set $n^2 = \epsilon$, where $\epsilon = +1$ if it is spacelike and $\epsilon = -1$ if it is timelike. Then

$$g_a{}^b = \gamma_a{}^b + \epsilon n_a n^b \ , \tag{7.5}$$

where the first fundamental form $\gamma_a{}^b$ appears as a projector projecting vectors in $\mathbf{T}$ to vectors in $\mathbf{T}^{\perp}$.

Surfaces of dimension two and codimension two need more work because neither the tangent direction nor the normal direction is unique. In spacetime there is the additional complication that the surface may be null, in which case there will be a normal vector that coincides with a tangent vector (as we saw in Chapter 4). Leaving this case aside, so that Eq. (7.4) applies, we will find that either $\mathbf{T}^{\mathrm{T}}$ or $\mathbf{T}^{\perp}$ is a Lorentzian vector space. To avoid too many sign choices we assume that the surface is everywhere spacelike, and then the second alternative holds. This is the most important case for physics.

We now choose bases in both $\mathbf{T}^{\mathrm{T}}$ and $\mathbf{T}^{\perp}$. Since the surface is spacelike it has a timelike normal vector, and we can choose a basis $(\vec{n}, \vec{e})$ for $\mathbf{T}^{\perp}$ such that

$$n_a n^a = -1 \ , \qquad e_a e^a = 1 \ , \qquad n_a e^a = 0 \ . \tag{7.6}$$

Of course this basis can be changed by a Lorentz transformation. However, if

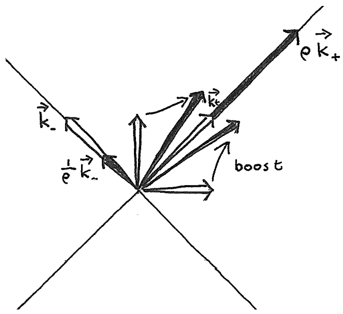

Figure 7.1. As bases in $\mathbf{T}^{\perp}$ we use either an orthogonal pair of one timelike and one spacelike vector, or two null vectors. The picture shows how one basis (white) is changed into another (black) by a Lorentz boost.

we consider spacelike surfaces embedded in spacelike hypersurfaces that are themselves embedded in spacetime, then we can choose $\vec{n}$ to be the unique normal to the hypersurface.

When $\mathbf{T}^{\perp}$ is a $(1+1)$-dimensional Minkowski space it contains two unique null directions from the origin. Hence one natural choice of basis consists of two *null normals*,

$$k_{\pm}^a = n^a \pm e^a \ , \qquad \vec{k}_+^2 = \vec{k}_-^2 = 0 \ , \quad \vec{k}_+ \cdot \vec{k}_- = -2 \ . \tag{7.7}$$

We can decide that the null normals are future directed. Still they are not fully determined by these conditions, since we can perform the change

$$k_+^a \to \rho k_+^a \ , \qquad k_-^a \to \frac{1}{\rho} k_-^a \ . \tag{7.8}$$

If we insist that the null vectors are future directed, we only allow positive functions ρ. Such a change results from a Lorentz boost in $\mathbf{T}^{\perp}$.

We can express the spacetime metric on a form analogous to Eq. (3.20), using the first fundamental form on the surface:

$$g^{ab} = \gamma^{ab} + e^a e^b - n^a n^b = \gamma^{ab} - \frac{1}{2}(k_+^a k_-^b + k_-^a k_+^b) \ . \tag{7.9}$$

Again $\gamma_a{}^b$ works as a projector onto $\mathbf{T}^{\mathrm{T}}$. But in this notation the indices do not keep track of how the vectors sit relative to the orthogonal decomposition of $\mathbf{T}$, which is a bit of a drawback.

To our unholy mixture of notations we now add the index-free one, which tends to make our calculations more understandable when the codimension exceeds one. We begin by introducing a set of vector fields X, Y, Z tangential to $\mathcal{M}$. You will simply have to remember that X, Y, Z are vectors belonging to $\mathbf{T}^{\mathrm{T}}$. Similarly, n is a normal vector. In this notation we would write, say,

$$\langle X, Y \rangle \equiv g_{ab} X^a Y^b \ , \qquad \nabla_X Y \equiv X^b \nabla_b Y^a \ . \tag{7.10}$$

Thus $\nabla_X Y$ is again a vector. We also have

$$\langle Y, n\rangle = 0 \quad \Rightarrow \quad \langle \nabla_X Y, n\rangle + \langle Y, \nabla_X n\rangle = 0 \ . \tag{7.11}$$

This is true because we assume that we are using the metric compatible connection.

The index-free notation, which may be disconcerting at first, is useful when we address the question: what kind of vector is $\nabla_X Y$? It is one of our main objectives to study the decompositions

$$\nabla_X Y = (\nabla_X Y)^{\mathrm{T}} + (\nabla_X Y)^{\perp} \tag{7.12}$$

$$\nabla_X n = (\nabla_X n)^{\mathrm{T}} + (\nabla_X n)^{\perp} \ , \tag{7.13}$$

where the derivative is the standard metric compatible covariant derivative in the ambient space, and Eq. (7.4) should be kept firmly in mind.

We first study the decomposition of $\nabla_X Y$, and begin with the tangential component $(\nabla_X Y)^{\mathrm{T}}$. That is, we first take the covariant derivative in a tangent direction of a tangent vector, and afterwards project the resulting vector back into the tangent space of the submanifold. This operation defines a covariant derivative $\bar{\nabla}_i$ on the submanifold, consistently with the rules for such things that we stated in Chapter 2. Thus

$$\bar{\nabla}_X Y = (\nabla_X Y)^{\mathrm{T}} \ . \tag{7.14}$$

But on the submanifold there already exists a unique covariant derivative compatible with the intrinsic metric γ_{ij}. Fortunately, these two covariant derivatives are identical. We prove this through the simple calculation

$$\bar{\nabla}_X \gamma_{ab} = \left(\nabla_X (g_{ab} + n_a n_b)\right)^{\mathrm{T}} = \left(n_a \nabla_X n_b + \nabla_X n_a n_b\right)^{\mathrm{T}} = 0 \ . \tag{7.15}$$

.

The point being that $(n_a)^{\mathrm{T}} = 0$, and I hope that the mixture of notations does not confuse you. For definiteness we assumed a spacelike hypersurface, but the argument goes through for all codimensions and for all kinds of hypersurfaces. The conclusion is encouraging because it is as simple as it can be.

To see what it means, consider a great circle on a sphere in Euclidean space. If X denotes its normalized tangent vector we know from the Frenet–Serret equations that $\nabla_X X$ points along a direction normal to the sphere. Therefore, when we do the projection, we find that

$$\nabla_X X = (\nabla_X X)^{\perp} \quad \Rightarrow \quad (\nabla_X X)^{\mathrm{T}} = 0 \quad \Rightarrow \quad \bar{\nabla}_X X = 0 \ . \tag{7.16}$$

So this is a geodesic in the curved intrinsic geometry of the sphere. For a circle of constant latitude (not at the equator) the normal to the curve is not normal

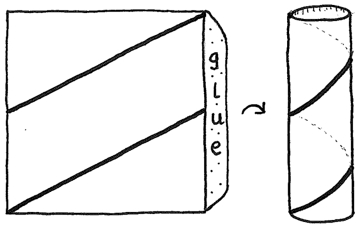

Figure 7.2. Take a rectangular piece of paper and draw geodesics on it as shown. Glue the piece of paper together so that it forms a cylinder. You will see one of the helices we discussed in Chapter 3. The normal vector of the helix is normal to the cylinder, hence the helix is a geodesic with respect to the intrinsic metric on the cylinder. Locally, the intrinsic metric of the surface is unchanged by your operations.

to the sphere, and therefore such a circle is not a geodesic on the sphere. For another example, see Figure 7.2.

Turning to the normal component we begin with the definition

$$K(X,Y) = -(\nabla_X Y)^{\perp} \ . \tag{7.17}$$

The question now is whether this expression defines a tensor. It does. What we have to prove is linearity, namely that for arbitrary scalar fields f and g there holds

$$\begin{aligned} K(fX_1 + gX_2, Y) &= fK(X_1, Y) + gK(X_2, Y) \\ K(X, fY_1 + gY_2) &= fK(X, Y_1) + gK(X, Y_2) \ . \end{aligned} \tag{7.18}$$

The first one is obvious. For the second, we note that if $\vec{n}$ is any normal vector field (obeying $\langle Y, n\rangle = 0$) the definition can be rewritten as

$$\langle K(X,Y), n\rangle = -\langle \nabla_X Y, n\rangle = \langle Y, \nabla_X n\rangle = X^a Y^b \nabla_a n_b \ . \tag{7.19}$$

Linearity, in both arguments, is now obvious. Hence we have defined a mixed tensor, known as the *Weingarten tensor*.

Looking back at Eqs. (7.12) and (7.13), we see that the Weingarten tensor accounts not only for $(\nabla_X Y)^{\perp}$ but also for $(\nabla_X n)^{\mathrm{T}}$. The second description is perhaps the more illuminating one. The Weingarten tensor determines the *extrinsic* geometry of a submanifold by encoding how its normal directions are changing in tangential directions as we move along it. Meanwhile its *intrinsic* geometry is described by the first fundamental form.

At this point it may be helpful to see a few formulas written in index notation, and indeed in concrete coordinates. We begin the translation by writing

$$K(X,Y) = X^b Y^c K_{bc}{}^a = X^i Y^j K_{ij}{}^{\mathring{a}} \ , \tag{7.20}$$

where we went to the length of introducing Swedish indices to mark vectors in $\mathbf{T}^\perp$. Using a suitable basis for T^T, consisting of a set of vectors $\vec{e}_i$, we translate Eq. (7.19) to

$$K_{ij}(n) \equiv K_{ij}{}^a n_a = -n_a e_i^b \nabla_b e_j^a = e_i^b e_j^a \nabla_b n_a \ . \tag{7.21}$$

If we have a parametric representation of the surface, so that Eq. (7.2) applies, then we can use coordinates to write

$$K_{ij}{}^a n_a = -n_a e_i^b \nabla_b e_j^a = -n_\mathbf{a} \frac{\partial^2 x^\mathbf{a}}{\partial u^i \partial u^j} - n_\mathbf{a} \Gamma_{\mathbf{bc}}{}^\mathbf{a} \frac{\partial x^\mathbf{b}}{\partial u^i} \frac{\partial x^\mathbf{c}}{\partial u^j} \ . \tag{7.22}$$

This is a computationally friendly expression, and we will soon return to it. Meanwhile, we just observe that it generalizes one of the Frenet–Serret equations for curves.

The coordinate based formula proves that the Weingarten tensor is symmetric in its lower indices. This can be proved in the index-free notation as well. Imagining that the vector fields X, Y have been extended to the ambient space in any arbitrary fashion, the calculation is

$$K(X,Y) - K(Y,X) = -(\nabla_X Y - \nabla_Y X)^\perp = -([X,Y])^\perp = 0 \ . \tag{7.23}$$

In the last step we use the fact X and Y are surface forming by construction, and then Frobenius' theorem guarantees that $[X,Y]$ belongs to $\mathbf{T}^\mathrm{T}$.[1]

Continuing the calculation begun in Eq. (7.21), now with the understanding that we are looking at a symmetric tensor, we obtain

$$K_{ij}(n) = \frac{1}{2} e_i^a e_j^b (\nabla_a n_b + \nabla_b n_a) = \frac{1}{2} e_i^a e_j^b \mathcal{L}_{\vec{n}} g_{ab} = \frac{1}{2} e_i^a e_j^b \mathcal{L}_{\vec{n}} \gamma_{ab} \ . \tag{7.24}$$

This can be cleaned up a little by observing that the Lie derivative along a normal direction applied to a vector in $\mathbf{T}^\mathrm{T}$ is again a vector in $\mathbf{T}^\mathrm{T}$, and similarly for tensors. Thus

$$K_{ij}(n) = \frac{1}{2} \mathcal{L}_{\vec{n}} \gamma_{ij} \ . \tag{7.25}$$

The Weingarten tensor tells us how the first fundamental form changes if we deform the surface in the direction of one of its normal vectors.

At this point, either you sit down to calculate the Weingarten tensor of a sphere and of a cylinder, or else you will be totally lost. As our example here we consider a round sphere embedded in a spherically symmetric spacetime

[1] Since the discussion of this theorem (in Chapter 3) was rather sketchy, you may want to spend another evening among Wald's appendices.

with metric (4.8). I will go over it very slowly. At first we assume that we are in a region where $V(r) > 0$. The sphere is defined by the equations

$$t - t_0 = 0 \ , \qquad r - r_0 = 0 \ , \tag{7.26}$$

where t_0 and r_0 are constants. A parametric representation of the surface is completed by $\theta = u^1$ and $\phi = u^2$. The first fundamental form is then

$$\gamma_{ij} = r^2(\theta_{,i}\theta_{,j} + \sin^2\theta\phi_{,i}\phi_{,j}) \ . \tag{7.27}$$

A natural basis for $\mathrm{T}^\perp$ is provided by the gradients $\nabla_a t$ and $\nabla_a r$, but we need to normalize these vectors. Thus we choose the normal (co-) vectors

$$n_a = -\sqrt{V(r)}\nabla_a t \ , \qquad e_a = \frac{1}{\sqrt{V(r)}}\nabla_a r \ . \tag{7.28}$$

The sign in front of the timelike normal ensures that the vector is future directed. Our calculational formula (7.22) now gives

$$K_{ij}(n) = -n_{\mathbf{a}}\Gamma_{\mathbf{bc}}{}^{\mathbf{a}}x^{\mathbf{b}}_{,i}x^{\mathbf{c}}_{,j} = \sqrt{V(r)}\left(\Gamma_{\theta\theta}{}^{t}\theta_{,i}\theta_{,j} + \Gamma_{\phi\phi}{}^{t}\phi_{,i}\phi_{,j}\right) = 0 \tag{7.29}$$

$$K_{ij}(e) = -\frac{1}{\sqrt{V(r)}}\left(\Gamma_{\theta\theta}{}^{r}\theta_{,i}\theta_{,j} + \Gamma_{\phi\phi}{}^{t}\phi_{,i}\phi_{,j}\right) = \frac{\sqrt{V(r)}}{r}\gamma_{ij} \ . \tag{7.30}$$

We used the fact that things were arranged so that $x^{\mathbf{a}}_{,ij} = 0$. If we had made the calculation for a sphere embedded in Minkowsi space we might have preferred to use Cartesian rather than polar coordinates. Then there would have been a contribution from the second derivatives, but on the other hand all the Christoffel symbols would have been zero. In polar coordinates we have to calculate a few of those in order to reach the same conclusion.

Are the results as expected? Yes they are. If you deform the sphere by moving it a little along the direction $\vec{n}$ it does not change, because $\vec{n}$ is in fact our familiar Killing vector field ∂_t^a. If you move it a little along the direction $\vec{e}$ you are moving it outwards, and it will grow, equally in all directions, which is precisely what the second equation says.

An alternative basis for $\mathbf{T}^\perp$ is provided by the two null vectors $\vec{k}_\pm = \vec{n} \pm \vec{e}$. Clearly

$$K_{ij}(k_\pm) = K_{ij}(n \pm e) = K_{ij}(n) \pm K_{ij}(e) = \pm K_{ij}(e) \ , \tag{7.31}$$

which is just a way of repackaging the same information. Notice that these tensors differ in their overall signs.

The above calculation suffers from the fact that we had to assume that $V(r) > 0$. At the expense of inserting some new signs we can redo it under the assumption that $V(r) < 0$, but we would like to cover also the case $V(r) = 0$. Then more radical changes are called for, because $-\nabla_a t$ and $\nabla_a r$ point in the same direction when $V(r) = 0$. So we go over to Eddington–Finkelstein coordinates, as in Eq. (2.46), meaning that

$$\nabla_a v = \nabla_a t + \frac{1}{V(r)} \nabla_a r \ . \tag{7.32}$$

Our two null normal vectors are then given by

$$k_a^- = -\sqrt{V(r)} \nabla_a v \ , \qquad k_a^+ = \frac{2}{\sqrt{V(r)}} \nabla_a r - \sqrt{V(r)} \nabla_a v \ . \tag{7.33}$$

As they stand they are useless at $V(r) = 0$. However, we noticed in Eq. (7.8) that there is some freedom left in their normalization. So we replace them with

$$k_a^- = -\nabla_a v \ , \qquad k_+^a = 2\nabla_a r - V(r) \nabla_a v \ . \tag{7.34}$$

Of course this could have been written down directly, had we started in this coordinate system. Redoing the calculation (which involves computing a few Christoffel symbols in Eddington–Finkelstein coordinates) we arrive at

$$K_{ij}(k_-) = -\frac{1}{r} \gamma_{ij} \ , \qquad K_{ij}(k_+) = \frac{V(r)}{r} \gamma_{ij} \ . \tag{7.35}$$

This time we can conclude that the first fundamental form is unchanged if it is deformed in the direction of $\vec{k}_+$ at a point where $V(r) = 0$. This is as it should be, because at such points the vector is actually pointing along the direction of a Killing vector field. Another conclusion we can draw is that if $V(r) < 0$ the sphere is shrinking when moved along any of its two future directed null normals. In this region the round spheres are trapped surfaces.

I hope you are now convinced that the Weingarten tensor means something, and that it can be calculated, so we can introduce some more terminology. Because the Weingarten tensor is symmetric it makes sense to split it into its trace and tracefree parts, using the first fundamental form for the purpose. The trace part gives the *mean curvature vector* $\vec{H}$,

$$H^a = \gamma^{ij} K_{ij}{}^a \ . \tag{7.36}$$

The mean curvature vector is normal to the surface. It is a good candidate for being the 'principal normal' of a submanifold of dimension larger than one. We also define the *null expansions* of a spacelike surface as

$$\theta_+ = \gamma^{ij} K_{ij}(k_+) \ , \qquad \theta_- = \gamma^{ij} K_{ij}(k_-) \ . \tag{7.37}$$

If both are negative the surface is trapped. If one of them is zero, the surface is said to be *marginally trapped*. If $\theta_+ = 0$, and if $\vec{k}_+$ points "outwards" (say, towards infinity) in some meaningful sense, then the surface is said to be *marginally outer trapped*, regardless of the sign of θ_-.

Now choose a particular normal vector $\vec{n}$. We can then define the *second fundamental form* associated to this normal direction by

$$K_{ij}(n) \equiv K_{ij}{}^a n_a \ . \tag{7.38}$$

If the submanifold is a hypersurface there is a unique normal direction at each point, and then we talk of *the* second fundamental form. Indeed, in this case,

$$K_{ij} \equiv K_{ij}(n) \quad \Rightarrow \quad K_{ij}{}^a = \epsilon K_{ij} n^a \ . \tag{7.39}$$

In this formula $\epsilon = +1$ if the normal is spacelike and $\epsilon = -1$ if it is timelike. Like the first fundamental form, the second fundamental form is a symmetric tensor, and not at all a differential form. They are called "forms" because they can be used to define quadratic forms in the components of a vector. As you may recall, the first fundamental form $\gamma_{\mathbf{ab}}$ is sometimes interpreted as a quadratic form ds^2 in the components $dx^{\mathbf{a}}$ of a vector. This is closely connected to the interpretation of extrinsic curvature given by Euler for the case of a surface in Euclidean space. Given a normal vector $\vec{n}$ at a point on the surface, and a tangent vector with components $dx^{\mathbf{a}}$, we single out a unique 2-plane intersecting the surface in a curve. This curve has a first curvature κ_1, and Euler showed that

$$\kappa_1 = \frac{K_{\mathbf{ab}} dx^{\mathbf{a}} dx^{\mathbf{b}}}{ds^2} \ . \tag{7.40}$$

By orienting the 2-plane suitably we find a maximum first curvature k_1 and a minimum first curvature k_2. These are called the *principal curvatures* of the surface at the point, and the corresponding curves intersect orthogonally. For a round sphere in Euclidean space Eq. (7.30) gives

$$K_{\mathbf{ab}} dx^{\mathbf{a}} dx^{\mathbf{b}} = \frac{1}{r} ds^2 \ , \tag{7.41}$$

and indeed all the sections give curves with $\kappa_1 = 1/r$ in this case.

The quick way to find the principal curvatures is to raise an index on the second fundamental form, to obtain an operator $K_i{}^j$ with eigenvalues that are, in fact, the two principal curvatures. You can now figure out why the trace $K = K_i{}^i$ of the second fundamental form is called the "mean" curvature.

There are various kinds of surfaces that deserve special attention because of the way they are embedded in the ambient space. We recognize them because their Weingarten tensors have very special properties. If the mean curvature vector (7.36) vanishes the surface is said to be *minimal*. To see why, let us move the surface a little in a normal direction, and see how its area changes. We actually know the answer from Eq. (3.45), although it is a little confusing because what is called $\vec{t}$ there (when we thought of it as a tangent vector to a congruence of curves) will now be replaced by $f\vec{n}$, where f is an arbitrary function on the surface telling us how much to deplace the surface along the

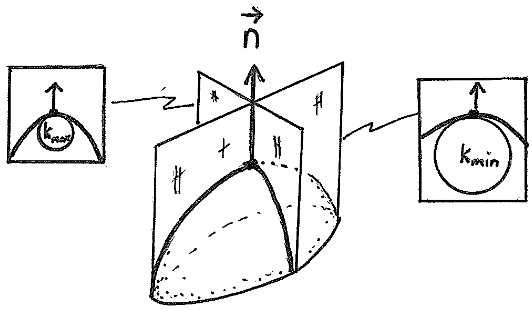

Figure 7.3. Together with a tangent vector the normal vector of a surface defines a 2-plane, and a curve in that 2-plane having a definite curvature at a given point. The second fundamental form summarizes all the relevant information.

normal direction $\vec{n}$. With these changes made, we obtain the expansion of a geodetic congruence emerging orthogonally from the surface as

$$\theta = \gamma^{ab}\nabla_a(fn_b) = f\gamma^{ab}\nabla_a n_b = fK \ . \tag{7.42}$$

From Eq. (3.46) we see that the area is unchanged to first order in the displacement if and only if $K = 0$. (If the codimension is higher than one, replace K with $\vec{H} \cdot \vec{n}$.) This means that the area of the surface is minimal, or perhaps maximal. An example of a minimal surface is the helicoid from Chapter 3. Another example is the equator on the 3-sphere. This is the actually the largest round sphere one can find embedded in the 3-sphere. Still its area grows under every deformation such that the function f is non-vanishing only in a small region of the surface, so it does deserve the name "minimal". Because a global deformation can shrink it, it is said to be an *unstable* minimal surface. Going a bit against this terminology, a spatial hypersurface with $K = 0$ embedded in a spacetime is by definition a *maximal hypersurface*.

A complementary definition is that of *totally umbilic* surfaces, for which the first and second fundamental forms are everywhere proportional. A round sphere in Euclidean space is an example, and in fact the only example one can find there. It is not easy for a surface to be umbilic at every point.

Finally we come to the case when the Weingarten tensor vanishes. Then the submanifold is called *totally geodesic*. From Gauss' formulas (7.44) we see that if X is tangent to a totally geodesic surface then

$$\nabla_X X = \bar{\nabla}_X X \ . \tag{7.43}$$

Choose a point on the surface and solve the equation $\nabla_X X = 0$ to obtain the unique geodesic that has X as its initial tangent vector. Evidently it also obeys the equation $\bar{\nabla}_X X = 0$. Therefore a geodesic that starts out tangential to a totally geodesic surface is a geodesic with respect to the intrinsic metric as well, and therefore it stays within the surface. This explains the name "totally geodesic". Totally geodesic surfaces are so special that there may well not exist any in a given curved space (except for one-dimensional ones). Higher dimensional totally geodesic surfaces arise if the space has a reflection symmetry of some sort. Thus, consider the spatial slice through the Schwarzschild solution which is depicted (in two-dimensional caricature) in Figure 4.2. Clearly there is a reflection symmetry there, and the waist of the paraboloid—which is the event horizon to be—is a fixed point set of that reflection. But then it must be totally geodesic. (Why? Because if we pick a point and a tangential vector there, there is a unique geodesic starting in that direction from that point. The point and the vector is untouched by the reflection. But if that geodesic were able to stray out of the surface it would be moved by the reflection, contradicting its uniqueness.) Similarly, the spatial slice $t = 0$ is a totally geodesic hypersurface in the Schwarzschild spacetime, because it is a fixed point set under the time reflection $t \rightarrow -t$.

The main conclusion so far is that

$$\nabla_X Y = (\nabla_X Y)^{\mathrm{T}} + (\nabla_X Y)^{\perp} = \bar{\nabla}_X Y - K(X,Y) \ . \tag{7.44}$$

This is known as *Gauss' formulas*. Our first objective has been reached. The tangential component of $\nabla_X Y$ behaves very nicely, and the perpendicular component defines an interesting geometric object. We also understand $(\nabla_X n)^{\mathrm{T}}$.

7.2 Projecting the Riemann tensor

We will use the index-free notation to tackle our second and final objective in this chapter, which is to relate the Riemann tensor in the ambient space $\mathcal{M}$ to objects that live on the submanifold. The first pieces of notation are

$$R(X,Y) \equiv X^c Y^d R_{cd}{}^a{}_b \ , \qquad R(X,Y)Z \equiv X^c Y^d R_{cd}{}^a{}_b Z^b \ . \tag{7.45}$$

If we use abstract indices, the left and right hand sides mean exactly the same thing. Note the order of the indices, which must be kept in your head if you use the index-free notation. Some formulas are messed up, notably the Ricci identity:

$$R_{cd}{}^a{}_b Z^b = [\nabla_c, \nabla_d] Z^a \ \Leftrightarrow \ R(X,Y)Z = (\nabla_X \nabla_Y - \nabla_Y \nabla_X)Z - \nabla_{[X,Y]}Z \ . \tag{7.46}$$

Once this is swallowed (preferably by checking it explicitly) the calculations are easy. The advantage is that we can (and will) take care of projections by declaring that X, Y, Z are tangential vector fields. Using Gauss' formulas we see that

$$\begin{aligned} \nabla_X(\nabla_Y Z) &= \nabla_X(\bar{\nabla}_Y Z - K(Y,Z)) = \\ &= \bar{\nabla}_X \bar{\nabla}_Y Z - K(X, \bar{\nabla}_Y Z) - \nabla_X K(Y,Z) \ . \end{aligned} \tag{7.47}$$

After some further work we find

$$
\begin{aligned}
R(X,Y)Z &= \left([\bar{\nabla}_X, \bar{\nabla}_Y] - \bar{\nabla}_{[X,Y]}\right) Z - \nabla_X K(Y,Z) + \nabla_Y K(X,Z) - \\
&\quad - K(X, \bar{\nabla}_Y Z) + K(Y, \bar{\nabla}_X Z) + K([X,Y],Z) \\
&= \bar{R}(X,Y)Z - \nabla_X K(Y,Z) + \nabla_Y K(X,Z) - \\
&\quad - K(X, \bar{\nabla}_Y Z) + K(Y, \bar{\nabla}_X Z) + K(\bar{\nabla}_X Y, Z) - K(\bar{\nabla}_Y X, Z) \ ,
\end{aligned}
\tag{7.48}
$$

where $\bar{R}$ denotes the Riemann tensor formed from the intrinsic first fundamental form. From here we can derive useful formulas for the Riemann tensor when projected into four or three tangential directions.

Equation (7.48) is a vector equation. If we take its scalar product with an arbitrary tangent vector $W \in \mathbf{T}^{\mathrm{T}}$ the terms on the second line drop out because they point in a normal direction. We also observe that

$$
-\langle \nabla_X K(Y,Z), W \rangle = \langle K(Y,Z), \nabla_X W \rangle = -\langle K(Y,Z), K(X,W) \rangle \ . \tag{7.49}
$$

With no further effort we obtain Gauss' remarkable *Theorema Egregium* :

$$
\begin{aligned}
&\langle R(X,Y)Z, W \rangle = \\
&= \langle \bar{R}(X,Y)Z, W \rangle - \langle K(Y,Z), K(X,W) \rangle + \langle K(X,Z), K(Y,W) \rangle \ .
\end{aligned}
\tag{7.50}
$$

This is the desired relation between the Riemann tensor in the ambient space and the Riemann tensor intrinsic to the submanifold.

For a hypersurface, where Eq. (7.39) allows us to trade the Weingarten tensor for a unique second fundamental form, this is

$$
R_{ijkl} = \bar{R}_{ijkl} + \epsilon (K_{il} K_{jk} - K_{ik} K_{jl}) \ . \tag{7.51}
$$

(Recall that $\epsilon = +1$ if the normal is spacelike, $\epsilon = -1$ if it is timelike.) Now

$$
\begin{aligned}
\gamma^{ac}\gamma^{bd} R_{abcd} &= (g^{ac} - \epsilon n^a n^c)(g^{bd} - \epsilon n^b n^d) R_{abcd} = R - 2\epsilon R_{ab} n^a n^b = \\
&= \epsilon (g_{ab} R - 2R_{ab}) n^a n^b = -2\epsilon G_{ab} n^a n^b \ .
\end{aligned}
\tag{7.52}
$$

Here the Einstein tensor makes one of its actually rather rare appearances in a purely geometric argument. We obtain

$$
2G_{ab} n^a n^b = -\epsilon \bar{R} + K^2 - K_{ab} K^{ab} = -\epsilon \bar{R} + 2(k_1 k_2 + k_2 k_3 + k_3 k_1) \ . \tag{7.53}
$$

In the second step we used the principal curvatures, that is to say the eigenvalues of the operator $K^a{}_b$, to express the result in a more memorable form.

Let us consider the simpler case of a surface embedded in a three dimensional flat space. Then the left hand side of Eq. (7.51) vanishes, and we see that

$$\bar{R} = \epsilon\left((\mathrm{Tr}K)^2 - \mathrm{Tr}K^2\right) = 2\epsilon k_1 k_2 \ . \tag{7.54}$$

With $\epsilon = +1$ this was the content of Gauss' original Theorema Egregium. It pleased him very much, and rightly so because the left hand side is intrinsic to the surface while the right hand side is constructed from two factors that, individually, depend on the embedding.

The intrinsic geometrical meaning of the curvature scalar R is best seen by surrounding a point by a sphere consisting of all the points at constant distance r from the point. In this way we obtain a *geodesic ball*. It is assumed that r is small enough so that the geodesics coming from the point do not intersect. We then compare the total volume $\mathrm{Vol}(\circ)$ of that ball to the corresponding volume $\mathrm{Vol}_0(\circ)$ of a ball of the same radius in flat space. One finds, if R denotes the curvature scalar evaluated at the centre of the ball and d is the dimension of the space it lives in, that

$$\frac{\mathrm{Vol}(\circ)}{\mathrm{Vol}_0(\circ)} = 1 - \frac{r^2 R}{6(d+2)} + o(r^4) \ . \tag{7.55}$$

A positive R causes the volume of the ball to grow more slowly with radius. Conversely, if $R < 0$ there is more space than there is in flat space. The proof uses Riemann's normal coordinates, see Chapter 11. In Lorentzian geometry this idea has to be modified somewhat, because a "sphere" is no longer a natural concept, but the idea nevertheless survives.[2]

The Theorema Egregium was the starting point for the idea that the geometry of a surface can be studied by means of invariant quantities—in the case of a two-dimensional surface, $\bar{R}$ is the only one—formed from the purely intrinsic first fundamental form itself. The geometry of a surface embedded in space is captured by two quantities, the *mean curvature* $k_1 + k_2$ and the *Gaussian curvature* $k = k_1 k_2$. The first depends on the embedding, the second not.

In Euclidean space $\epsilon = 1$, and we find that the surface has positive intrinsic curvature if and only if the two principal curvatures have the same sign. You can make a rough estimate of the principal curvatures by looking at the surface. Look at two osculating circles intersecting orthogonally at a point on the surface, and ask where their normal vectors point. For a sphere they point the same way, so $k > 0$. For a saddle shaped surface, $k < 0$. On a torus you will find that k changes sign. Interestingly, if you integrate the Riemann scalar over a torus you always find zero, regardless of how the torus is deformed. For a spacelike surface in Minkowski space $\epsilon = -1$, and the conclusion is

[2] One considers the volume of causal diamonds centred at the point. See J. Myrheim, *Statistical geometry*, CERN preprint CERN-TH-2538, 1978.

the opposite one: the intrinsic curvature is positive for a surface that looks saddle shaped in a spacetime diagram, while a hyperboloidal surface (say) has negative intrinsic curvature.

We return to Eq. (7.48), and look at the normal component. The term involving $\bar{R}(X,Y)$ does not contribute, but the rest of the expression is a hard nut to crack. However, let us assume that we are looking at a hypersurface. This simplifies things, because

$$\langle \nabla_X n, n\rangle = -\langle n, \nabla_X n\rangle \quad \Rightarrow \quad (\nabla_X n)^{\perp} = 0 \ . \tag{7.56}$$

There is only a tangential component. Reverting to abstract indices, we see that in the codimension one case

$$\begin{aligned} (\nabla_X K(Y,Z))^{\perp} &= \left(X^b \nabla_b (Y^c Z^d \epsilon K_{cd} n^a)\right)^{\perp} = \\ &= \left(X^b Y^c Z^d \epsilon \nabla_b K_{cd} n^a\right)^{\perp} + K(\nabla_X Y, Z) + K(Y, \nabla_X Z) \ . \end{aligned} \tag{7.57}$$

So when we project in the normal direction the terms occupying the second line in (7.48) are non-zero, but they are cancelled by terms coming from the first line. When contracting with the normal vector we are left with

$$(R_{abcd} n^c)^{\mathrm{T}} = -\bar{\nabla}_a K_{bd} + \bar{\nabla}_b K_{ad} \quad \Rightarrow \quad (R_{ac} n^c)^{\mathrm{T}} = \bar{\nabla}_b K_a{}^b - \bar{\nabla}_a K \ . \tag{7.58}$$

This is known as the *Codazzi equation* for a hypersurface.

It should not escape your attention that if we are sitting at a point in a spatial slice through spacetime, with normal vector $\vec{n}$ and tangent vectors $\vec{v}$, then the Gauss and Codazzi equations together encapsulate four of the Einstein equations,

$$\left.\begin{aligned} n^a n^b G_{ab} &= 0 \\ n^a v^a G_{ab} &= 0 \end{aligned}\right\} \quad \Rightarrow \quad \left\{\begin{aligned} &\bar{R} + K^2 - K_{ab}K^{ab} = 0 \\ &\bar{\nabla}_b K_a{}^b - \bar{\nabla}_a K = 0 \quad . \end{aligned}\right. \tag{7.59}$$

Strikingly, these components of the Einstein equations involve only quantities that can be defined using the intrinsic and extrinsic geometry of the spatial hypersurface itself. But this is another story, to be told in Chapter 8.

We have still to discuss the case of the Riemann tensor projected into two tangent vectors and two normal vectors. The equation that describes this is known as the Ricci equation. We postpone it to the next chapter, in which the codimension is one. When the codimension exceeds one the story becomes rather intricate. The normal bundle and something called the third fundamental form enter into it. The normal bundle is a vector bundle over the embedded manifold $\mathcal{M}$ such that its fibres are copies of $\mathbf{T}^{\perp}$. A cross section of the normal bundle is a particular normal vector field. The third fundamental form arises as a connection on the normal bundle.

◇ **Problem 7.1** Compute the first and second fundamental forms of the helicoid, Eq. (3.37), and verify that this is a minimal surface. Also calculate its Gaussian curvature.

◇ **Problem 7.2** A surface of revolution in Euclidean space is defined by an equation of the form $z = f(r)$. Write down their first and second fundamental forms, and show that there exists an essentially unique minimal surface of revolution. (It is called a *catenoid*.) Compute its Gaussian curvature.

◇ **Problem 7.3** Can you embed a metrical 2-sphere in 3-dimensional Minkowski space? There is an elegant solution provided one cuts the sphere open along a meridian and embeds its infinite covering space. A particular $SO(2)$ subgroup of the isometry group of the sphere will be decompactified, and realized as a Lorentz boost in the embedding space. Find this embedding explicitly.

◇ **Problem 7.4** Consider spatial hypersurfaces of constant r in the interior of the Schwarzschild spacetime. Is there a maximal hypersurface (with $K = 0$) among them?

8 Initial data

An initial value formulation of Einstein's equations can be set up once we have understood how to describe spacelike hypersurfaces in spacetime. Once we have it, we can change perspective and regard spacetimes as something that grows out of suitable data set on spacelike 3-manifolds.

8.1 $3+1$ decompositions

We are concerned with a foliation of spacetime by hypersurfaces. We assume that there exists a function t on spacetime such that the individual hypersurfaces—the leaves of the foliation—are given by setting t equal to some constant. Each leaf is equipped with a metric γ_{ij} and a symmetric tensor field K_{ij}, and the question is what conditions these tensors have to obey if they are to arise as the first and second fundamental forms induced on the hypersurfaces when the spacetime in which they are embedded obeys Einstein's equations. For simplicity we assume that we are in vacuum, so that the Einstein tensor of the spacetime metric vanishes. From Chapter 7 we already know one part of the answer: the Gauss-Codazzi equations (7.59) must hold. But this answer is only partial, because we will need to know how γ_{ab} and K_{ab} change as we move from one hypersurface to another.

It will be useful to begin by studying the normal vector to a given hypersurface. Let it be given by

$$n_a = -N\nabla_a t \ , \qquad n_a n^a = -1 \ , \qquad N > 0 \ . \tag{8.1}$$

Note that we are assuming very little about the function t, only that it should be possible to find a function N, known as the *lapse function*, such that the the normal vector n_a is a timelike future pointing unit vector. We do assume that the hypersurface forms a leaf of a foliation of spacetime, which means that we get, at least locally, a congruence of curves pointing along the vector field $\vec{n}$. We define the "acceleration vector"

$$a_a = n^b \nabla_b n_a \ \in \ \mathbf{T}^{\mathrm{T}} \ . \tag{8.2}$$

(Equation (7.4) must be kept firmly in mind throughout this chapter.) It follows that

$$\nabla_a n_b = K_{ab} - n_a a_b \ . \tag{8.3}$$

(To see this, contract the left hand side first with n^a and then with n^b.) Comparing to Eq. (3.44) we see that the second fundamental form takes care of the expansion and shear of the congruence. There is no rotation because the congruence is hypersurface forming by assumption, but there is an extra term because the curves do not have to be geodesics. Making use of Eq. (8.1) it is not difficult to show that

$$a_a = \frac{1}{N}\left(\nabla_a N + n_a n^b \nabla_b N\right) = \frac{1}{N}\bar{\nabla}_a N \ , \tag{8.4}$$

and as an easy consequence

$$\bar{\nabla}_a a_b = \frac{1}{N}\bar{\nabla}_a \bar{\nabla}_b N - a_a a_b \ . \tag{8.5}$$

These equations will become useful very soon.

We now want to know how γ_{ab} and K_{ab} change as we move away from a hypersurface. From Eqs. (7.25) and (7.39) we already know that

$$\mathcal{L}_{\vec{n}} \gamma_{ab} = 2K_{ab} \ . \tag{8.6}$$

It remains to compute

$$\mathcal{L}_{\vec{n}} K_{ab} = n^c \nabla_c K_{ab} + \nabla_a n^c K_{cb} + \nabla_b n^c K_{ac} \ . \tag{8.7}$$

We can bring the Riemann tensor projected into two normal directions into this formula by focussing on the first term on the right hand side,

$$n^c \nabla_c (\nabla_a n_b + n_a a_b) = n^c R_{cab}{}^d n_d + n^c \nabla_a \nabla_c n_b + n^c \nabla_c (n_a a_b) \ . \tag{8.8}$$

Furthermore

$$n^c \nabla_a \nabla_c n_b = \nabla_a a_b - \nabla_a n^c \nabla_c n_b \ . \tag{8.9}$$

Some welcome cancellations occur when we make use of Eq. (8.3), so that our evolution equation becomes

$$\mathcal{L}_{\vec{n}} K_{ab} = n^c R_{cab}{}^d n_d + \nabla_a a_b + a_a a_b + K_b{}^c K_{ac} + n^c n_a \nabla_c a_b + n_b a^c K_{ac} \ . \tag{8.10}$$

We test this expression by contracting it, first with n^a and then with n^b. The right hand side is zero in both cases, which means that we can project the whole formula into $\mathbf{T}^{\mathrm{T}}$ without losing any information. Then we obtain

$$\mathcal{L}_{\vec{n}} K_{ab} = n^c R_{cab}{}^d n_d + \bar{\nabla}_a a_b + a_a a_b + K_b{}^c K_{ac} \ . \tag{8.11}$$

Finally we make use of Eq. (8.5) to arrive at the attractive form

$$\mathcal{L}_{\vec{n}} K_{ab} = n^c R_{cab}{}^d n_d + \frac{1}{N}\bar{\nabla}_a \bar{\nabla}_b N + K_b{}^c K_{ac} \ . \tag{8.12}$$

This is known as the *Ricci equation*. Deriving a similar formula for the Riemann tensor projected into two normal vectors is a considerably more involved affair if the codimension of the embedded surface exceeds one, but since we are concerned with hypersurfaces in this chapter we rest content with the result as it stands.

8.2 The action principle

We gain an interesting perspective on Einstein's equations if we start from a variational principle, and more precisely from the Hilbert action

$$S[g] = \int R = \int \mathrm{d}^4x\sqrt{-g}g^{ab}(\partial_c\Gamma_{ab}{}^c - \partial_a\Gamma_{cb}{}^c + \Gamma_{ab}{}^e\Gamma_{ce}{}^c - \Gamma_{ca}{}^d\Gamma_{bd}{}^c)\ . \quad (8.13)$$

We want to show that Einstein's equations are the Euler–Lagrange equations following from this action principle. This is actually very easy if done in the right way, especially if we ignore surface terms so that we can perform partial integrations without comment.[1]

We are looking for the extremals over a very infinite-dimensional set of objects, and it is occasionally useful to know what the notation we are going to use actually means. We will look at one-parameter families of metrics $g_{ab}(s)$, and for each such family we define

$$\delta g_{ab} \equiv \frac{d}{ds}g_{ab|s=0}\ . \quad (8.14)$$

The variation of the action will be required to vanish for all such families. More mundanely, recall that

$$\delta g = gg^{ab}\delta g_{ab} = -gg_{ab}\delta g^{ab} \quad \Rightarrow \quad \delta\sqrt{-g} = -\frac{1}{2}\sqrt{-g}g_{ab}\delta g^{ab}\ . \quad (8.15)$$

It is also useful to observe that while the affine connection $\Gamma_{ab}{}^c$ is not a tensor, the 'difference' between two such objects is, that is to say that $\delta\Gamma_{ab}{}^c$ is a tensor.

With this understanding, we obtain

$$\begin{aligned} \delta S &= \int \mathrm{d}^4x\sqrt{-g}\left(\delta g^{ab}(R_{ab} - \frac{1}{2}Rg_{ab}) + g^{ab}(\nabla_c\delta\Gamma_{ab}{}^c - \nabla_a\delta\Gamma_{cb}{}^c)\right) \\ &= \int \mathrm{d}^4x\left(\sqrt{-g}\delta g^{ab}(R_{ab} - \frac{1}{2}Rg_{ab}) + \nabla_a(\sqrt{-g}g^{cb}\delta\Gamma_{cb}{}^a - \sqrt{-g}\delta\Gamma_{cb}{}^c)\right)\ . \end{aligned} \quad (8.16)$$

The last group of terms is the covariant divergence of a vector density, hence an ordinary divergence, hence a surface term. We drop it, and conclude that

$$\delta S = 0 \quad \Rightarrow \quad G_{ab} = R_{ab} - \frac{1}{2}g_{ab}R = 0\ . \quad (8.17)$$

This is Einstein's equations.

For the initial data formulation of Einstein's equations we aim to express the integrand of the action in terms of the first and second fundamental forms of foliating hypersurface, with unit timelike normals n^a. To do this, we observe that

$$R = 2(G_{ab} - R_{ab})n^an^b = \bar{R} + K^2 - K_{ab}K^{ab} - 2R_{ab}n^an^b\ . \quad (8.18)$$

[1] It is fortunate that the right way is known. It appears that Hilbert himself found the calculation somewhat baffling. See L. Corry, J. Renn, and J. Stachel, *Belated decision in the Hilbert–Einstein priority dispute*, Science **278** (1997) 1270.

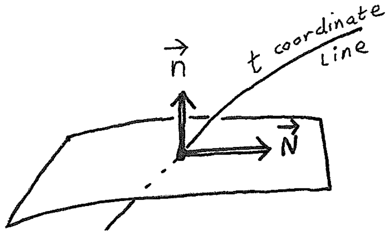

Figure 8.1. There is typically no reason to choose the t-coordinate lines to be orthogonal to the hypersurface, necessitating the introduction of lapse and shift.

Here we found a use for the Theorema Egregium. To deal with the last term we note that

$$[\nabla_a, \nabla_b]n^b = R_{ab}{}^b{}_c n^c = -R_a{}^c n_c \;\Rightarrow\; R_{ab}n^a n^b = -n^a(\nabla_a\nabla_b - \nabla_b\nabla_a)n^b \ . \quad (8.19)$$

Hence

$$R_{ab}n^a n^b = -\nabla_a(n^a\nabla_b n^b) + \nabla_a n^a \nabla_b n^b + \nabla_b(n^a\nabla_a n^b) - \nabla_b n^a \nabla_a n^b \ . \quad (8.20)$$

We are beginning to collect some total derivative terms. The second term is simply equal to K^2. Concerning the last term it helps that we are in the codimension one case. Thus $n_a\nabla_b n^a$ vanishes. But this means that the covariant derivative acting on n^b is projected tangentially, and from Eq. (7.56) we know that the normal component of $\nabla_X n^b$ vanishes. Thus we can write

$$R_{ab}n^a n^b = \nabla_a(-n^a\nabla_b n^b + n^b\nabla_b n^a) + K^2 - K_{ab}K^{ab} \ . \quad (8.21)$$

Putting everything together we obtain

$$R = \bar{R} + K_{ab}K^{ab} - K^2 + 2\nabla_a(n^a\nabla_b n^b + n^b\nabla_b n^a) \ . \quad (8.22)$$

This is precisely what we need for our purposes.

Now we introduce a time coordinate. We have assumed that the hypersurfaces are given by holding a time function t constant. Now we choose t to be a coordinate. The spacetime metric is decomposed as

$$g_{ab} = \gamma_{ab} - n_a n_b \ , \quad (8.23)$$

where γ_{ab} is the first fundamental form of a hypersurface. The vector pointing along the t-coordinate lines is

$$\partial_t^a = Nn^a + N^a \ , \qquad N^a n_a = 0 \ . \quad (8.24)$$

See Figure 8.1. The vector field N^a is known as the *shift vector*. The lapse function N was introduced in Eq. (8.1). Choosing $N = 1$ and $N^a = 0$ leads to *Gaussian normal coordinates*, but typically this is a problematic choice because the congruence defined by the normal vectors may develop caustics as we move away from a given hypersurface. So the freedom offered by the lapse and shift is quite important.

We introduce a coordinate basis ∂_i^a also on the hypersurface, and then we see what the metric looks like in the resulting coordinate system:

$$g_{tt} = g_{ab}\partial_t^a \partial_t^b = -N^2 + N_c N^c \tag{8.25}$$

$$g_{ti} = g_{ab}\partial_t^a \partial_i^b = g_{ab} N^a \partial_i^b = N^j g_{ab} \partial_j^a \partial_i^b = N_i \ . \tag{8.26}$$

In effect then

$$g_{\mathbf{ab}} = \left(\begin{array}{c|c} -N^2 + N^k N_k & N_j \\ \hline N_i & \gamma_{ij} \end{array} \right) \ , \quad g^{\mathbf{ab}} = \left(\begin{array}{c|c} -\frac{1}{N^2} & \frac{N^j}{N^2} \\ \hline \frac{N^i}{N^2} & \gamma^{ij} - \frac{N^i N^j}{N^2} \end{array} \right) \ , \tag{8.27}$$

where indices on the shift vector are raised and lowered with the first fundamental form γ_{ij}. This is known as the ADM decomposition of the spacetime metric.[2]

We found the inverse metric basically by inspection. Its determinant can be found in the same way if you recall Cramer's rule for how to invert a matrix, which here implies that

$$-\frac{1}{N^2} = \frac{\det \gamma}{\det g} \quad \Rightarrow \quad g = -N^2 \gamma \quad \Rightarrow \quad \sqrt{-g} = N\sqrt{\gamma} \ . \tag{8.28}$$

All the ingredients in the Einstein–Hilbert action have now been accounted for.

We have arrived at the 3+1 decomposition

$$S = \int d^4x \ \sqrt{-g} R = \int d^4x N \sqrt{\gamma}(\bar{R} + K_{ab}K^{ab} - K^2) + \text{surface term} \ , \tag{8.29}$$

where $\bar{R}$ is the curvature scalar of the first fundamental form γ_{ab}. We ignore the surface term in this all too brief account. From Eq. (7.25) we know that the second fundamental form can be expressed as a Lie derivative of the first. Defining

$$\dot{\gamma}_{ab} \equiv \partial_t \gamma_{ab} = \mathcal{L}_{\vec{\partial}_t} \gamma_{ab} \tag{8.30}$$

we obtain

$$2NK_{ab} = \dot{\gamma}_{ab} - \mathcal{L}_{\vec{N}} \gamma_{ab} = \dot{\gamma}_{ab} - \bar{\nabla}_a N_b - \bar{\nabla}_b N_a \ . \tag{8.31}$$

We can now proceed with the Legendre transform and derive the Hamiltonian formulation of the field equations. They will turn out to be equivalent to the constraint and evolution equations that we derived in Section 8.1.

However, here we just make three remarks. The first is that once the Legendre transformation is completed you see that the time-time and time-space components of the spacetime metric—the lapse and shift—enter the action as Lagrange multipliers only. They are not dynamical fields. In a numerical simulation one can choose lapses and shifts at will, so that the hypersurface

[2] To see what the acronym stands for, consult R. Arnowitt, S. Deser, and C. W. Misner, *The dynamics of general relativity*, in L. Witten (ed.): Gravitation: An Introduction to Current Research, Wiley 1962.

evolves at different rates in different parts of the spacetime to be. The second remark is that we will obtain the Gauss and the Codazzi equations, Eqs. (7.59), when we vary with respect to the Lagrange multipliers. These equations therefore arise as constraints on the initial data set on the hypersurface. They have to be solved before the evolution can start. The third remark is that electrodynamics provides a simple analogy of what is going on here. The time component of the vector potential enters the action as a Lagrange multiplier, and Gauss' law $\nabla_a E^a = \rho$ is a constraint on the initial data. From this point of view general relativity is at the pinnacle of all gauge theories, but has much in common with what happens on the plains.

8.3 Geometrostatics

With the constraint equations in hand, we have to ask how to solve them. What are they equations for? There is no unique answer to this question. Here we will consider a simple special case, called *geometrostatics*, where a particular strategy turns out to be quite superior.[3] The same strategy is one of the leading options in the general case too, but we pass lightly over that. Thus, suppose that we look for initial data such that $K_{ij} = 0$. In the completed spacetime this results in a totally geodesic hypersurface, and the resulting spacetime will have a reflection symmetry (in time) leaving this hypersurface invariant. The moment of maximum expansion in a closed and recollapsing Friedman cosmology can serve as an example of this. The solution is certainly not static, but it is "momentarily" static.

If we look for vacuum solutions the constraint equations on a totally geodesic hypersurface collapse to the single equation

$$\bar{R} = 0 \ . \tag{8.32}$$

This is still a highly non-linear equation for the metric γ_{ij}. To solve it—and, in the first place, to find something "inside" γ_{ij} to solve for—we will perform a *conformal rescaling*. This means that we consider a different metric $\hat{\gamma}_{ij}$ related to the true metric by

$$\gamma_{ij} = \omega^4 \hat{\gamma}_{ij} \ . \tag{8.33}$$

The exponent 4 on the *conformal factor* ω is chosen for later convenience. The important thing is that ω is nowhere zero. Not surprisingly, if you work out the Riemann tensors of two conformally related metrics, you will find that they are closely related. In particular, for the curvature scalars one finds

$$\bar{R} = \frac{1}{\omega^4}\left(\hat{R} - \frac{8}{\omega}\hat{\gamma}^{ij}\hat{\nabla}_i\hat{\nabla}_j\omega\right) \ , \tag{8.34}$$

where hatted quantities are calculated using $\hat{\gamma}_{ij}$, unhatted using γ_{ij}.[4] We can

[3] The case for studying geometrostatics was made by C. W. Misner, *The method of images in geometrostatics*, Ann. Phys. (N. Y.) **24** (1963) 102.

[4] The exact formula depends on dimension. For the details, see Wald's appendices—and beware that he uses the exponent 2 on the conformal factor, as usual in many contexts.

always choose the conformal factor so that $\det \hat{\gamma} = 1$. Then, when we impose $\bar{R} = 0$, we obtain an elliptic equation for the conformal factor ω while the hatted first fundamental form remains as free data to be specified. In other words, now we have something definite to solve for.

Here we are not asking for the general solution of the constraint equation. Any solution simple enough to play with is good enough. Therefore we choose $\hat{\gamma}_{ij}$ to be the flat metric, in which case $\hat{R} = 0$ and the equation reduces to

$$\triangle\omega = 0 \ , \tag{8.35}$$

where $\triangle$ denotes the flat space Laplacian. This we know how to solve. If we impose the boundary condition that $\omega \to 1$ at infinity (in order to obtain an asymptotically flat solution), we know that we have to accept some kind of singular behaviour in the interior. The simplest possibility is

$$\omega = 1 + \frac{e}{r} \quad \Rightarrow \quad ds^2 = \left(1 + \frac{e}{r}\right)^4 \left(dx^2 + dy^2 + dz^2\right)\right) \ . \tag{8.36}$$

What is this? We can see that, as $r \to \infty$,

$$\gamma_{ij} = \delta_{ij} + \frac{4e}{r}\delta_{ij} + O(1/r^2) \ . \tag{8.37}$$

We recognize this behaviour from the Schwarzschild solution, provided we set $m = 2e$. We can also see that there are round spheres present, with areas

$$\mathrm{Vol}(S^2) = 4\pi r^2 \left(1 + \frac{e}{r}\right)^4 \ . \tag{8.38}$$

But this does not shrink to zero as $r \to 0$. On the contrary, the area takes its minimum value $4\pi \cdot 16e^2 = 4\pi \cdot (2m)^2$ at $r = e$. It is clear that we are looking at a slice through the Schwarzschild solution. If we evolve these initial data we will find that the bifurcation surface in the event horizon sits at $r = e = 2m$.

To clinch the argument we observe that our spatial slice has a reflection symmetry under

$$x^i \to x^{i'} = -\frac{e^2 x^i}{r^2} \quad \Rightarrow \quad r \to r' = \frac{e^2}{r} \ . \tag{8.39}$$

A quick calculation confirms that

$$ds^2 = \left(1 + \frac{e}{r}\right)^4 \left(dr^2 + r^2 d\Omega^2\right) = \left(1 + \frac{e}{r'}\right)^4 \left(dr'^{\,2} + r'^{\,2} d\Omega^2\right) \ , \tag{8.40}$$

and that the special sphere at $r = e$ is left invariant by the isometry. So the two 'ends' of the space, at $r \to 0$ and at $r \to \infty$, are truly identical.

Of course, observing that our solution is spherically symmetric and provides legitimate initial data for the Einstein equations, we could have referred to Jebsen's theorem in order to say that the solution we found must be a slice through the Schwarzschild solution. But in this initial data formulation we can easily break out of the spherically symmetric straightjacket. Thus we can set

$$\omega = 1 + \frac{e_1}{|\mathbf{r} - \mathbf{a}_1|} + \frac{e_2}{|\mathbf{r} - \mathbf{a}_2|} \ , \tag{8.41}$$

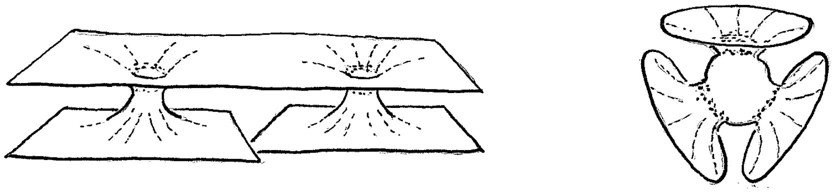

Figure 8.2. Initial data for two black holes. Which of the pictures that applies depends on the distance between the throats. On the left, a is large, and the individual black holes do not affect each other very much.

presumably corresponding to two black holes (and definitely to three asymptotic regions). If we want we can include an arbitrary number of black holes at arbitrary "positions" in the flat "background" space, but two black holes are enough in order to make some interesting observations about how they are distorted by each other. It is clear that the mass of the two, when read off at $r \to \infty$, equals $2e_1 + 2e_2$. To read off the mass in one of the other asymptotic regions we translate the solution so that the black hole in question 'sits' at $\mathbf{a}_1 = \mathbf{0}$, and the other at $\mathbf{a}_2 = (0, 0, a)$. Then we use the coordinates $x'_i = -e_1^2 x_i / r^2$ to describe the asymptotic region corresponding to $r \to 0$. We find (dropping the primes) that

$$ds^2 = \left(1 + \frac{r}{e_1} + \frac{e_2}{\sqrt{\frac{e_1^4 x^2}{r^4} + \frac{e_1^2 y^2}{r^4} + (-\frac{e_1^2 z}{r^4} - a)^2}} \right)^4 \frac{e_1^4}{r^4}(dx^2 + dy^2 + dz^2) =$$

$$= \left(\frac{e_1}{r} + 1 + \frac{1}{r} \frac{e_1 e_2}{\sqrt{\frac{e_1^4 x^2}{r^4} + \frac{e_1^2 y^2}{r^4} + (-\frac{e_1^2 z}{r^4} - a)^2}} \right)^4 (dx^2 + dy^2 + dx^2) \ . \tag{8.42}$$

In the limit when $r \to \infty$ this becomes

$$ds^2 \approx \left(1 + \frac{4}{r} \left(e_1 + \frac{e_1 e_2}{a} \right) \right) (dx^2 + dy^2 + dz^2) \ . \tag{8.43}$$

Thus the asymptotic behaviour is that of a Schwarzschild black hole with mass

$$M_1 = 2e_1 \left(1 + \frac{e_2}{a} \right) \ . \tag{8.44}$$

Let us set $e_1 = e_2$ for simplicity. Then the masses, as read off in the three asymptotic regions, are

$$M_0 = 4e \ , \qquad M_1 = M_2 = 2e \left(1 + \frac{e}{a} \right) \ . \tag{8.45}$$

We observe that $M_0 \leq M_1 + M_2$. Although we found the solution by superposing two solutions of a linear Laplace equation, the individual black hole geometries are not just superposed on each other.[5]

[5] For more on this, see D. R. Brill and R. W. Lindquist, *Interaction energy in geometrostatics*, Phys. Rev. **131** (1963) 471.

◇ **Problem 8.1** Perform the Legendre transformation of (8.29) to obtain a Hamiltonian formulation of Einstein's equations.

9 Trapped surfaces

I have already dropped a number of hints suggesting that trapped surfaces are important in gravitational collapse. The concept arises only in Lorentzian geometry, where there are different kinds of surfaces depending on whether the mean curvature vector is spacelike, timelike, or null. We insist that the surface is a spacelike surface of codimension 2, and also that it is a closed surface (probably a sphere—other topologies are less important), because then the distinction becomes physically important. Thus a closed surface is said to be *untrapped* if $\vec{H}$ is everywhere spacelike, *trapped* if $\vec{H}$ is everywhere timelike, and *marginally trapped* if $\vec{H}$ is everywhere either null or zero. In formulas, we expand the Weingarten tensor using a pair of null vectors as a basis for $\mathbf{T}^{\perp}$,

$$K_{ij}{}^{a} = -\frac{1}{2}K_{ij}(k_+)k_-^a - \frac{1}{2}K_{ij}(k_-)k_+^a \ . \tag{9.1}$$

When contracting with the first fundamental form to obtain the mean curvature vector we also define the two *null expansions* θ_+ and θ_- through

$$H^a = \gamma^{ij}K_{ij}{}^{a} = -\frac{1}{2}\theta_+ k_-^a - \frac{1}{2}\theta_- k_+^a \ . \tag{9.2}$$

The surface is trapped if the null expansions have the same sign. For definiteness we will always talk about the *future-trapped* case, when θ_+ and θ_- are both negative, so that wavefronts leaving the surface shrink in both of the two possible directions.

Important variations on the theme exist, especially if there is a natural definition of 'outwards' defined. This would be the case if the surface sits inside some spacelike hypersurface extending all the way to infinity. Let $\vec{k}_+$ be outwards directed. Then we can define *outer trapped* surfaces by saying that $\theta_+ < 0$, irrespective of the sign of θ_-. Similarly, a *marginally outer trapped* surface (a MOTS) is defined by the single condition $\theta_+ = 0$.

The importance of trapped surfaces stems from the fact that they are defined by an inequality, and also from the fact that they can be detected without knowing anything about the future evolution of spacetime. They can be detected by studying initial data set on some spacelike hypersurface, and if these initial data are perturbed a bit, the trapped surface may change its location and its null expansions a little, but there will still be a trapped surface in the perturbed initial data. The overriding reason why they are important is that their presence in the initial data signals (provided some further assumptions hold) that the time evolved spacetime will be singular in the sense of being geodesically incomplete. I have no intention to go into the singularity theorems here, but let me quote one of them for concreteness:

Theorem: No space-time M can satisfy all of the following three requirements together:
(1) M contains no closed timelike curves,
(2) every inextendible causal geodesic in M contains a pair of conjugate points,
(3) there exists a future- (or past-) trapped set $S \in M$.

I bring this up only to show that trapped surfaces are important.[1] A trapped surface is one example of a *trapped set*. Another example is that of a point such that its future lightcone eventually starts to reconverge along every direction. Concerning the second clause in the theorem, a pair of *conjugate points* is (roughly speaking) a pair of points where two neighbouring geodesics intersect and (more precisely) a pair of points such that if you send out a congruence of geodesics from one of them, the expansion diverges to $-\infty$ at the other. Hawking and Penrose supplemented the statement above with a proof that the Raychaudhuri equation will force conjugate points to appear on every complete causal geodesic if the strong energy condition and a certain (modest) genericity condition hold. Then, if clause (1) and (2) hold, it follows that incomplete causal geodesics must exist. In this sense the spacetime is singular.

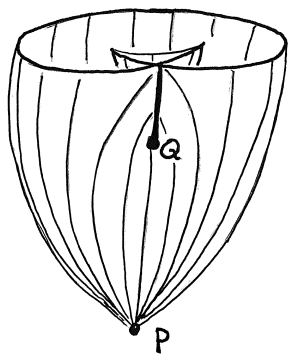

Figure 9.1. P and Q are conjugate points. The congruence emerging from P develops a line of cross-over and a caustic starting at Q.

Before we can claim that singularities is a generic feature of spacetimes it must be shown that trapped sets arise in all circumstances where matter is strongly concentrated, or when the gravitational field becomes in some sense strong. We want to show that it can happen even if the initial data we start

[1] This particular theorem is quoted from S. W. Hawking and R. Penrose, *The singularities of gravitational collapse and cosmology*, Proc. Roy. Soc. Lond. **A314** (1970) 529.

out from are free of trapped sets. If the initial data are supposed to be reasonably generic ones then this is a question about the long-time behaviour of Einstein's equations, and one would expect it to be simply too difficult to handle. Actually it is not, but we will have to leave it to the experts.[2]

On the other hand it is known how to look for trapped surfaces on a given spacelike hypersurface, notably one that has arisen from a numerical calculation. Moreover it is known that the boundary of the region in a given spacelike hypersurface where trapped surfaces occur is itself a marginally outer trapped surface, and efficient algorithms exist for how this MOTS is to be localized in numerical similations. This bounding MOTS is known as the *apparent horizon*. To locate it, no information about the future of the spacelike hypersurface is needed. It is known that if cosmic censorship holds then the apparent horizon lies inside the event horizon.[3] This means that future-trapped surfaces can never be observed by sensible observers, who avoid jumping into black holes. They are, however, observed by numerical relativists, who use them to establish that black holes are emerging from their calculations. Indeed this is of practical importance, because once an apparent horizon has been detected at some stage of the evolution, the numerical relativist can safely relax the precision of his calculation inside it, since any error introduced thereby will be subject to cosmic censorship and will leave the predictions concerning outgoing gravitational waves unaffected.

When the data on the slice containing the apparent horizon are evolved, a *marginally trapped tube* (MTT) foliated by marginally trapped surfaces will arise in spacetime. It can be spacelike, timelike, or null. In the Vaidya solution we found a spacelike MTT, which we called the apparent 3-horizon in Chapter 5. Spacelike MTTs are also known as *dynamical horizons*. If you did Problem 5.3 you found a timelike MTT in the Oppenheimer–Snyder solution. If you did Problem 5.2 you found a null MTT lying in between the two dust shells. The area of the marginally trapped surfaces foliating the latter are the same, whatever cross-section we take of it, so this is an example of a *non-expanding horizon* (NEH). With some further conditions added, to make it look more like a piece of a Killing horizon, a NEH becomes an *isolated horizon* (IH), and these extra conditions are fulfilled in the Vaidya example. As you can see, this subject is full of acronyms.[4]

It is tempting to regard the MTT, rather than the teleologically determined event horizon, as the true boundary of the black hole. The problem with this is that the resulting boundary will depend strongly on how spacetime is sliced into spatial hypersurfaces. Another choice of lapse and shift will result in different spatial slices, different apparent horizons, different marginally trapped tubes, and different suggestions for what the spacetime boundary of the black hole should be. Moreover, although the notions of trapped surfaces and MTTs are *quasi-local*—unlike the event horizon their location is determined solely by the spacetime geometry in the region that contains them—they are not

[2] D. Christodoulou: *The Formation of Black Holes in General Relativity*, EMS 2009.
[3] This is Proposition 9.2.8 in Hawking and Ellis.
[4] For more, see S. Hayward (ed.): *Black Holes: New Horizons*, World Scientific 2013.

local. If you have have an MTT inside a black hole, and if a piece of additional matter falls in from one side, the MTT will jump outwards also on the other side of the black hole because the changed geometry on one side will cause some locally marginally trapped surface extending to the other side to change from open to closed.[5]

To get some substance into this discussion we will study the non-uniqueness of marginally trapped tubes by perturbing the marginally trapped surfaces sitting in them. Then we have to understand not only the first variation of their area (which is given by the expansion), but also their second variation. Deriving the appropriate formula is a difficult matter, and the formula for $\delta\theta_+$ is a long one. In the special case of round marginally outer trapped surfaces in spherically symmetric spacetimes the result is at least easy to state, so we confine ourselves to this case. Thus we assume that the surface S obeys

$$\theta_+ = 0 \ , \qquad R_S = \frac{2}{r^2} \ , \tag{9.3}$$

where r is the usual area radius coordinate. We are going to deform the surface in a normal direction $\vec{n}$. It is convenient to introduce a basis for $\mathbf{T}^\perp$ such that $\vec{n}$ is one the basis vectors, but we do not wish to restrict its causal character (that is, we prefer to leave the sign of n^2 open for the moment). So we set

$$n^a = -\frac{1}{2}k_-^a + n^2 k_+^a \ , \qquad \vec{n}\cdot\vec{k}_+ = 1 \tag{9.4}$$

$$u^a = \frac{1}{2}k_-^a + n^2 k_+^a \ , \qquad \vec{u}\cdot\vec{n} = 0 \ , \quad u^2 = -n^2 \ . \tag{9.5}$$

Every normal vector except $\vec{k}_+$ can be expressed in this way. We illustrate this (and a little more) in Figure 9.2.

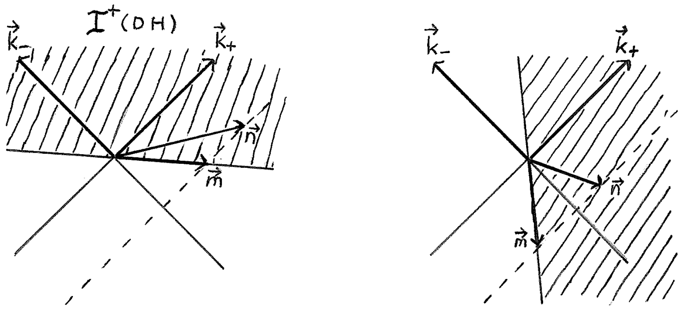

Figure 9.2. The vectors $\vec{k}_+, \vec{k}_-, \vec{n}, \vec{m}$ that appear in the perturbation argument. To the left $\vec{m}$ is spacelike, to the right $\vec{m}$ is timelike. Round trapped spheres occur in the shaded region. Exactly where, along the dashed line, the vector $\vec{n}$ is pointing is left open by Eq. (9.4).

[5] For an exact solution in a simple toy model see E. Jakobsson, *How trapped surfaces jump in* $2+1$ *dimensions*, Class. Quant. Grav. **30** (2013) 065022.

Restricting the result of Andersson et al.[6] to this case we learn—and I am asking you to simply accept it—that the second variation is given by

$$\delta_{f\vec{n}}\theta_+ = -\triangle_S f + \left(\frac{R_S}{2} - G_{ab}k_+^a u^b\right) f \ , \tag{9.6}$$

where the Laplacian on the sphere occurs in the first term on the right hand side. The function f is at our disposal, so that the magnitude and the sign of the deformation can vary freely over the sphere. It is convenient to rewrite the formula as

$$\delta_{f\vec{n}}\theta_+ = \left(\frac{1+L^2}{r^2} - \frac{1}{4}G_{ab}k_+^a k_-^b - \frac{n^2}{2}G_{ab}k_+^a k_+^b\right) f \ , \tag{9.7}$$

where

$$L^2 = -\left(\partial_\theta^2 + \cot\theta\partial_\theta + \frac{1}{\sin^2\theta}\partial_\phi^2\right) \ . \tag{9.8}$$

This one of the rather few purely geometric formulas where the Einstein tensor plays a role.

In the vacuum case we observe that

$$\delta_{f\vec{n}}\theta_+ = \frac{1+L^2}{r^2} f \ . \tag{9.9}$$

To create a spherically symmetric MTT the variation must vanish for constant perturbations f. At first sight this seems impossible; the resolution is that such a perturbation is possible in the only direction that is not included here, namely along the outward null normal direction $\vec{k}_+$. We conclude that the only spherically symmetric MTT must be a null NEH in the vacuum case. It is also immediate that

$$\int_{\mathbf{S}^2} f\delta_{f\vec{n}}\theta_+ = \int_{\mathbf{S}^2} (f^2 + \mathbf{L}f\cdot\mathbf{L}f) \geq 0 \ . \tag{9.10}$$

If the variation has a definite sign, we can conclude that if the perturbation is outwards ($f > 0$) the surface can only be untrapped, and it can only be trapped if the perturbation is inwards ($f < 0$). If there is a cosmological constant we find, for a constant perturbation, that

$$\delta_{f\vec{n}}\theta_+ = 0 \quad \Leftrightarrow \quad \frac{1}{r^2} = \lambda \ . \tag{9.11}$$

This has a solution only if $\lambda > 0$, and then only for very special spheres—namely those that sit on the *cosmological horizon*.

In the non-vacuum case we assume the dominant energy condition, which

[6] You will find the details in L. Andersson, M. Mars, and W. Simon, *Stability of marginally trapped surfaces and existence of marginally outer trapped tubes*, Adv. Theor. Math. Phys. **12** (2008) 853.

implies that $G_{ab}k_+^a k_+^b > 0$. If we again do a spherically symmetric perturbation and make sure that the expansion remains zero to first order, we find an equation that must be obeyed by the radial tangent vector of the resulting MTT, namely

$$\delta_{f\vec{m}}\theta_+ = 0 \quad \Leftrightarrow \quad m^2 = \frac{2}{G(k_+, k_+)}\left(\frac{1}{r^2} - \frac{1}{2}G_{ab}k_+^a k_-^b\right) . \tag{9.12}$$

Whether the tangent vector $\vec{m}$ is spacelike or timelike depends on the magnitude of $G_{ab}k_+^a k_-^b \geq 0$. The difference we observed between the round MTTs in the Vaidya and Oppenheimer-Snyder solutions hinge on that.

We concentrate on the spacelike case, which is likely to be the most interesting one. (There are arguments why the MTT should become spacelike at least where it joins the event horizon.) First we consider a deformation with constant f in the direction of the so far unspecified vector $\vec{n}$. The vector points into the region where the round spheres are trapped if $\delta\theta_+ < 0$, which can now be translated into $n^2 - m^2 > 0$. Allowing for arbitrary functions f we see that if we wish to perturb into the future $\mathcal{I}^+(DH)$ of a spacelike dynamical horizon we must have either $f > 0$ and $n^2 > m^2$, or $f < 0$ and $n^2 < m^2$. For easy control of these signs, we observe that

$$n^2 - m^2 = \frac{2}{r^2 G(k_+, k_+)}\left(\frac{L^2 f - r^2\delta_{f\vec{n}}\theta_+}{f}\right) . \tag{9.13}$$

If f is constant we see that the sign of $\delta\theta_+$ is determined by the sign of $n^2 - m^2$: if the marginally trapped round spheres are perturbed to other round spheres, they will become trapped in one direction and untrapped in the other. Another interesting conclusion presents itself if we observe that

$$G(k_+, k_+)\int_{\mathbf{S}^2} f(n^2 - m^2) = 2\int_{\mathbf{S}^2}(L^2 f - r^2\delta_{f\vec{n}}\theta_+) = -2r^2\int_{\mathbf{S}^2}\delta_{f\vec{n}}\theta_+ . \tag{9.14}$$

I took $G(k_+, k_+)$ out of the integral since it is a function of v and r only. It follows that the deformed surface can be trapped—with a negative definite sign of the variation—only if $f(n^2 - m^2)$ is somewhere positive, and it can be untrapped only if $f(n^2 - m^2)$ is somewhere negative. Hence a trapped surface must lie at least partly in the region where the round spheres are trapped.

But we are interested in whether one can deform the surface so that it becomes a trapped surface 'sticking out' of the MTT into the region where the round spheres are untrapped. We then choose a sign for $n^2 - m^2$. A glance at Figure 9.2 shows that the first thing to try is $n^2 - m^2 < 0$, because it is only in this case that we can reach the past of the round trapped surface. The deformation will take us out of the MTT if $f > 0$. We assume that $\delta_{f\vec{n}}\theta_+ < 0$, but the form of this positive function is at our disposal. Equation (9.13) now implies that there exist positive functions (pos) and (pos') such that

$$L^2 f + (\text{pos'}) = -(\text{pos})f \quad \Leftrightarrow \quad \triangle f = (\text{pos'}) + (\text{pos})f . \tag{9.15}$$

This implies that the function is f is convex in the region where it is positive. But this means that it cannot have a maximum there, so this case is ruled out.

We will have better luck if we choose $n^2 - m^2 > 0$. Then the vector $\vec{n}$ points into the region where the round spheres are trapped, which means that we must have $f < 0$ in order to make the deformed surface stick out. This time Eq. (9.13) implies that there exist positive functions (pos) and (pos') such that

$$L^2 f - r^2 \delta_{f\vec{n}}\theta_+ = L^2 f + (\text{pos'}) = (\text{pos})f \ . \tag{9.16}$$

Here we assumed that $\delta\theta_+ < 0$, but its form is arbitrary. Given such positive functions the deformed surface does stick out at all points where $f < 0$. To simplify Eq. (9.16), define

$$g = f - a_0 \quad \Leftrightarrow \quad f = g + a_0 \ , \tag{9.17}$$

with a_0 a positive constant. The form of the function (pos'), which determines the trapping, is at our disposal. We choose

$$(\text{pos'}) = -r^2 \delta_{f\vec{n}}\theta_+ = a_0(\text{pos}) \ , \tag{9.18}$$

and are left to find solutions to

$$L^2 g = (\text{pos})g \ . \tag{9.19}$$

We can for instance set g equal to an eigenfunction of the Laplacian. Thus we conclude that

$$f(\theta) = a_0 + a_l P_l(\cos\theta) \tag{9.20}$$

leads to $(\text{pos}) = l(l+1)$, and

$$\delta_{f\vec{n}}\theta_+ = -\frac{a_0 l(l+1)}{r^2} \ . \tag{9.21}$$

The deformed, and trapped, sphere sticks out of the MTT if $f < 0$ for some θ. This can always be arranged by choosing a_0 and a_l suitably. We can also concoct functions g such that f is positive in an arbitrarily small region on the undeformed sphere, and negative elsewhere. Then the deformed trapped surface will lie almost entirely outside the MTT. The overall conclusion is that trapped surfaces can stick out of the dynamical horizon, but that there are some strong restriction on how they do it.

In none of the above did we pay any attention to the inner null expansion θ_-. Of course, if it is negative to start with, it will stay negative under a small perturbation, but it is hard to control. In fact most of the available theorems on the behaviour of marginally trapped tubes are concerned with MOTTs—surfaces foliated by marginally outer trapped surfaces. Similarly, the *horizon finders* used by numerical relativists to detect trapped surfaces in their data are in fact designed to find MOTS—marginally outer trapped surfaces. Genuinely trapped surfaces do have some advantages though, because they are not dependent on any notion of 'inwards' and 'outwards' inherited from a spatial slice. You look at the surface, and in particular at the causal character of its mean curvature vector, and this decides the matter.

To illustrate this, let us give a simple proof of the fact that trapped surfaces cannot exist in a spacetime that admits a timelike Killing vector field, or in a region of a spacetime where such a Killing field exists (such as the exterior of the Schwarzschild solution).[7] First of all, if $\vec{\xi}$ is an arbitrary vector field with both tangential and a normal components we find when we project its covariant derivative in the tangential directions that

$$e_i^a e_j^b \nabla_a \xi_b = e_i^a e_j^b \nabla_a (\xi_b^{\mathrm{T}} + \xi_b^{\perp}) = \bar{\nabla}_i \bar{\xi}_j + K_{ij}{}^a \xi_a \ . \tag{9.22}$$

A bar over an object means that it has been projected into the tangential directions. In the last term only the normal component of $\vec{\xi}$ contributes. We contract this equation with the first fundamental form, and observe that the left hand side then vanishes because $\vec{\xi}$ is a Killing field. Thus

$$0 = \bar{\nabla}_i \xi^j + H^a \xi_a \ . \tag{9.23}$$

But by assumption both $\vec{\xi}$ and $\vec{H}$ are timelike. If we integrate over the closed surface, we obtain

$$0 = \oint_S \bar{\nabla}_i \xi^j \mathrm{d}S = -\oint_S H^a \xi_a \neq 0 \ . \tag{9.24}$$

The point is that the scalar product of two timelike vectors cannot vanish, and therefore the final integrand cannot change its sign anywhere on the surface. We have obtained a contradiction, confirming that the mean curvature vector has to be spacelike somewhere.

It is interesting to compare to the situation for minimal surfaces in Riemannian geometry. It is well known that Euclidean space does not contain a single closed minimal surface, while the 3-sphere does. The argument, in the former case, is very simple. A surface S is minimal if it gives an extremum of the action functional

$$A = \int_S \mathrm{d}S = \int_S \sqrt{\gamma}\mathrm{d}^2 u \quad \Rightarrow \quad \delta A = \int_S \frac{1}{2}\sqrt{\gamma}\gamma^{ij}\delta\gamma_{ij} \ . \tag{9.25}$$

If the ambient space is flat we have

$$\gamma_{ij} = \mathbf{x}_{,i} \cdot \mathbf{x}_{,j} \ . \tag{9.26}$$

It easily follows that setting δA to zero for all possible variations implies that the components of the vector $\mathbf{x}$ have to obey the intrinsic Laplace equation. There are no non-zero solutions to this equation on a topological sphere. On

[7] This argument is due to J. M. M. Senovilla.

the 3-sphere the Laplace equation gets replaced by the Helmholtz equation, and solutions can be found. A recent result says that generic compact Riemannian manifolds contain infinitely many closed minimal hypersurfaces. Moreover their union is dense.[8]

◇ **Problem 9.1** Show that there is a closed trapped surface passing through every point in de Sitter space (whose definition you can find in many places). Guess which clause in the Hawking–Penrose theorem is not fulfilled, in this obviously geodesically complete spacetime.

[8] K. Irie, F. C. Marques, and A. Neves, *Denseness of minimal hypersurfaces for generic metrics*, arXiv:1710.10752.

10 Isolated systems

Progress in science often depends on the idea that a part of the world can be described as an isolated system, while the rest of the world can be ignored. Whether a theory describing the world has to admit such a separation is perhaps not so clear. In gravity theory one tries to perform the separation by assuming that space-time has a simple structure “far away” from the region of interest. Not too simple though—we insist that gravitational radiation can be detected also from very large distances. Roughly speaking, a gravitational wave detector detects displacements whose amplitude decrease like $1/r$, where r is—in some rough sense—the distance from the source. Hence our definition of an isolated system must admit deviations from flat space with this kind of fall-off behaviour.

In the mathematical idealization that we will use, “far away” will be interpreted as “at infinity”, and we begin by recalling how we look at “infinity” in the complex plane. For this purpose it is helpful to introduce the complex coordinate

$$z = x + iy = e^{i\phi} \tan \frac{\theta}{2} \ . \tag{10.1}$$

We multiply the flat metric g with a factor that goes to zero as $|z| \to \infty$, to obtain the conformally related metric

$$d\hat{s}^2 = \Omega^2 ds^2 = \frac{4}{(1+|z|^2)^2} dz d\bar{z} = d\theta^2 + \sin^2 d\phi^2 \ . \tag{10.2}$$

The conformally related metric $\hat{g}$ is the metric on the unit sphere, and we see that infinity in the flat plane appears as a single point ∞, situated at the south pole of the sphere (at $\theta = \pi$). Of course the procedure is arbitrary to some extent, since we can change the conformal factor Ω by multiplying it with any function that is well behaved on the sphere. However, this would not change the fact that the conformal boundary of the flat plane is a single point. Indeed this is a dimension independent statement, and remains true of any Euclidean space. For a space of constant negative curvature the procedure gives a different result. The nature of the conformal boundary of a space (if any) does capture some of its geometry.[1]

Looking at "infinity" of Minkowski space leads to a more complex picture. It is convenient to begin by writing the Minkowski metric in terms of null coordinates $u = t - r$, $v = t + r$, as

$$ds^2 = -dt^2 + dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) = \tag{10.3}$$

$$= -du^2 - 2dudr + r^2(d\theta^2 + \sin^2\theta d\phi^2) = \tag{10.4}$$

$$= -dudv + \left(\frac{v-u}{2}\right)^2 (d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{10.5}$$

We now perform a conformal rescaling of the metric, beginning with the third of these forms. The aim is to make the rescaled metric behave itself in the limits $u, v \to \infty$. To this end, define Ω_E by

$$\Omega_E^2 = \frac{4}{(1+u^2)(1+v^2)} \ . \tag{10.6}$$

Ω_E is a well defined function on all of Minkowski space, and it tends to zero as $u, v \to \pm\infty$. Consider the conformally related metric

$$d\hat{s}_E^2 = \Omega_E^2 ds^2 = -\frac{4dudv}{(1+u^2)(1+v^2)} + \frac{(v-u)^2}{(1+u^2)(1+v^2)}(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{10.7}$$

To see what goes on at "infinity" we perform the coordinate change

$$\tan p = u \ , \quad \tan q = v \ , \qquad -\pi/2 < p, q < \pi/2 \ . \tag{10.8}$$

The metric then takes the form

$$d\hat{s}^2 = -4dpdq + \sin^2(q-p)(d\theta^2 + \sin^2\theta d\phi^2) = \tag{10.9}$$

$$= -dt'^{\,2} + dr'^2 + \sin^2 r'(d\theta^2 + \sin^2\theta d\phi^2) \ . \tag{10.10}$$

If the range of the coordinate $t' = p + q$ is extended to the range $(-\infty, \infty)$ this is precisely the static Einstein universe. In particular, the spatial sections are unchanging three dimensional spheres.

The conformally rescaled Minkowski space sits inside the Einstein universe, surrounded by the boundaries $p = -\pi/2$ and $q = \pi/2$. These are in fact lightcones, with a past vertex i^- at $(p, q) = (-\pi/2, -\pi/2)$, one vertex i^0 at $(p, q) = (-\pi/2, \pi/2)$, and a future vertex i^+ at $(p, q) = (\pi/2, \pi/2)$. The vertex

[1] See Valentina di Carlo, *Conformal compactification and anti-de Sitter space*, MSc Thesis, KTH/SU 2007, for more on this.

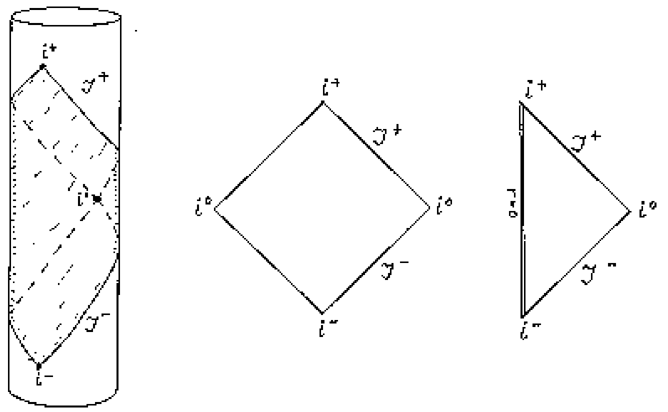

Figure 10.1. Here we see 1+1 dimensional Minkowski space, embedded in the Einstein universe, as a conformal diagram, and as a Carter-Penrose diagram. The latter is valid in any dimension.

i^0 sits at spacelike separation from any point inside the Minkowski region, and represents the conformal boundary of any spacelike slice therein. So in fact there are two disconnected null cones bounding Minkowski space, called $\mathcal{J}^+$ and $\mathcal{J}^-$. Here $\mathcal{J}$ is a script I, and pronounced *scri*.[2]

Again there is some arbitrariness in the picture. As long as we are only interested in $\mathcal{J}^+$ we can standardize the picture by insisting that the lightcone at infinity does not expand or contract. For this purpose we define a new conformal factor

$$\Omega = \omega\Omega_E = \frac{\Omega_E}{\sin(q-p)} = \frac{1}{r} \ . \tag{10.11}$$

Using Ω as a coordinate, and reverting to the original null coordinate u, we find that

$$d\hat{s}^2 = \omega^2 d\hat{s}_E^2 = \Omega^2 ds^2 = -\Omega^2 du^2 + 2dud\Omega + d\theta^2 + \sin^2\theta d\phi^2 \ . \tag{10.12}$$

$\mathcal{J}^+$ is the non-expanding null hypersurface sitting at $\Omega = 0$. This could have been obtained directly from the form (10.4). Doing so means that one misses i^0 and i^+, but these points will be rather singular whenever we try to conformally compactify a curved spacetime, and would need separate attention anyway. On the positive side, the coordinate r has a natural interpretation as an affine parameter on the radially directed null geodesics obtained by holding u, θ, ϕ constant.

The same procedure works also for the Schwarzschild, and indeed for the Kerr spacetime. Write the metric for the latter in the original outgoing Kerr

[2] Hans Rudberg studied the lightcone at infinity in his PhD thesis, Uppsala 1957. Then the idea was taken up by Roger Penrose. See his Batelle Rencontres lectures. In Polish, "skraj" means a boundary or edge ("skraj lasu" is the edge of a wood).

coordinates, (6.34). Again r is an affine parameter along a family of outgoing null geodesics, so it serves as a good (if rough) measure of the distance from the black hole. As our conformal factor we again choose

$$\Omega = \frac{1}{r} \ . \tag{10.13}$$

Trading r for Ω as a coordinate, and performing a Taylor expansion in Ω, we find that

$$\begin{aligned} d\hat{s}^2 = \Omega^2 ds^2 = 2dud\Omega - 2a\sin^2\theta d\Omega d\phi + d\theta^2 + \sin^2\theta d\phi^2 + \\ +\Omega^2(-du^2 + a^2\cos^2\theta d\theta^2 + a^2\sin^2\theta d\phi^2) + \\ +\Omega^3(2mdu^2 - 4amdud\phi) + \dots \ . \end{aligned} \tag{10.14}$$

Setting $\Omega = 0$ defines a null hypersurface whose intrinsic geometry is identical to that of the Minkowski $\mathcal{I}^+$.

Now comes the brave step. We will say that a system is isolated only if its space-time metric g admits a conformal rescaling to another Lorentzian metric $\hat{g}$,

$$\hat{g}_{ab} = \Omega^2 g_{ab} \ , \tag{10.15}$$

in such a way that the hypersurface defined by $\Omega = 0$ becomes the boundary of a region inside a larger unphysical space-time with metric $\hat{g}$. Of course we do not want arbitrary hypersurfaces in the unphysical space-time to define isolated systems, so we have to build some more physics into the definition. To do this, we observe that both g and $\hat{g}$ come with Riemann tensors of their own, the latter one being hatted. We will insist that $R_{ab} = 0$ in a neighbourhood of the boundary. (This can be relaxed a little, if electromagnetic fields are present, but we do not go into this here.) By the way we slip in the assumption that the boundary has topology $\mathbf{S}^2 \times \mathbf{R}$.

To see that the boundary is already severely constrained by the assumptions, we must study how various geometrical objects behave under conformal rescalings. We first define

$$\hat{g}_{ab} = \Omega^2 g_{ab} \ , \qquad \hat{g}^{ab} = \frac{1}{\Omega^2} g^{ab} \ , \tag{10.16}$$

$$n_a = \hat{n}_a = \nabla_a \Omega = \hat{\nabla}_a \Omega = \Omega_{,a} \ , \qquad \hat{n}^a = \hat{g}^{ab} n_b = \frac{1}{\Omega^2} n^a \ . \tag{10.17}$$

With two metrics in the game, we have to employ the convention about how indices are raised and lowered with some discernment. For hatted quantities the hatted metric performs this role.

It is then a matter of straightforward calculation to show that

$$\hat{\Gamma}_{ab}{}^{c} = \Gamma_{bc}{}^{a} + \frac{1}{\Omega}(\delta^c_a\Omega_{,b} + \delta^c_b\Omega_{,a} - g_{ab}g^{cd}\Omega_{,d}) \ , \tag{10.18}$$

$$R_{ab}{}^{cd} = \Omega^2\hat{R}_{ab}{}^{cd} + 4\left(\Omega\hat{\nabla}_{[a}\hat{n}^{[c} - \frac{1}{2}\hat{n}^2\delta^{[c}_{[a}\right)\delta^{d]}_{b]} \ . \tag{10.19}$$

We introduce the decomposition of the Riemann tensor into the Weyl tensor and the rest, and also define the *Schouten tensor* P_{ab}, as follows:

$$P_{ab} \equiv \frac{1}{2}R_{ab} - \frac{1}{12}g_{ab}R \ , \tag{10.20}$$

$$R_{ab}{}^{cd} = C_{ab}{}^{cd} + 4P_{[a}{}^{[c}\delta^{d]}_{b]} \ . \tag{10.21}$$

Then we find

$$C_{ab}{}^{cd} = \Omega^2\hat{C}_{ab}{}^{cd} \ , \qquad C_{abc}{}^{d} = \hat{C}_{abc}{}^{d} \ , \tag{10.22}$$

$$P_{ab} = \hat{P}_{ab} + \frac{1}{\Omega}\hat{\nabla}_a\hat{\nabla}_b\Omega - \frac{1}{2\Omega^2}\hat{g}_{ab}\hat{g}^{cd}\hat{\nabla}_c\Omega\hat{\nabla}_d\Omega \ . \tag{10.23}$$

The point about the Schouten tensor is that it behaves somewhat more elegantly under conformal rescalings, compared to the Ricci tensor itself.

Einstein's vacuum equations say that

$$P_{ab} = \frac{\lambda}{6}g_{ab} \ , \tag{10.24}$$

where, for once, we include a cosmological constant λ (to stress that a cosmological constant really makes a difference). By assumption these equations hold on one side of $\mathcal{I}$. For the conformally related geometry they imply, via Eq. (10.23), that

$$\hat{P}_{ab} + \frac{1}{\Omega}\hat{\nabla}_a\hat{\nabla}_b\Omega - \frac{1}{2\Omega^2}\hat{g}_{ab}\hat{g}^{cd}\hat{\nabla}_c\Omega\hat{\nabla}_d\Omega = \frac{\lambda}{6}g_{ab} = \frac{\lambda}{6}\frac{1}{\Omega^2}\hat{g}_{ab} \tag{10.25}$$

We multiply this equation with Ω^2 and then go to the limit $\Omega = 0$. The result is

$$\hat{n}_a\hat{n}^a \equiv \hat{g}^{cd}\hat{\nabla}_c\Omega\hat{\nabla}_d\Omega \mathrel{\hat{=}} -\frac{\lambda}{3} \ . \tag{10.26}$$

By convention a hatted equality is an equality that holds on $\mathcal{I}$. Since $\hat{n}^a$ is the normal vector of the hypersurface defined by $\Omega = 0$, the conclusion is that this hypersurface must be null if $\lambda = 0$, spacelike if $\lambda > 0$, and timelike if $\lambda < 0$. This is already something.

From now on we set $\lambda = 0$. We go back to Eq. (10.25), multiply with a single factor of Ω, and again set $\Omega = 0$. Then

$$\hat{\nabla}_a \hat{n}_b \hat{=} \hat{g}_{ab} \frac{\hat{g}^{cd}\hat{n}_c\hat{n}_d}{2\Omega^2} \ . \tag{10.27}$$

The traceless part of the right hand side vanishes. Recalling that $\hat{n}^a$ are the tangent vectors to the generators of the null surface $\mathcal{I}$, we conclude that these generators are rotation free (obviously, since they form a hypersurface), and also shear free (a more dramatic statement). So we know quite a bit about $\mathcal{I}$ already.

By means of a suitable function $\omega > 0$, non-vanishing also on $\mathcal{I}$, we can set the expansion to zero. That is to say, we rescale

$$\hat{g}_{ab} \to \omega^2 \hat{g}_{ab} \ . \tag{10.28}$$

We can then arrange that

$$\hat{\nabla}_a \hat{n}^a \hat{=} 0 \ . \tag{10.29}$$

The point is that we can always arrange that.

The conclusion so far is that, for any spacetime obeying Einstein's vacuum equation outside a compact region, with $\lambda = 0$, the mere assumption that $\mathcal{I}$ exists implies that it must be a shear free null hypersurface, whose expansion can be chosen to vanish.

We now come to the question of the asymptotic symmetry group, that is to say the group that leaves our requirements at $\mathcal{I}$ invariant. Because we admit deviations from Minkowski space this group is larger than the Poincaré group, that is than the symmetry group of Minkowski space itself. In fact it is much larger. It is called the *BMS group*.[3]

To begin with, the group that preserves the conformal structure on $\mathcal{I}$ is an infinite dimensional group known as the Newman-Unit group, consisting of the transformations

$$z \to z' = \frac{az+b}{cz+d} \ , \qquad u \to u' = F(u, z, \bar{z}) \ . \tag{10.30}$$

Here F is an arbitrary function of three variables. We also admit Möbius transformations acting on the generators; they form a group which is, by a seeming accident, isomorphic to the Lorentz group. These are the most general one-to-one conformal transformations of a sphere onto itself. The most general conformal transformation of a sphere onto itself is given by

$$z \to z' = z'(z) \quad \Rightarrow \quad dzd\bar{z} \to dz'd\bar{z}' = \left|\frac{dz'}{dz}\right|^2 dzd\bar{z} \ . \tag{10.31}$$

The result is a conformal transformation of the sphere, for all choices of the

[3] For Bondi, Metzner, and Sachs.

function $z'(z)$. But we restrict ourselves to one-to-one Möbius transformations. One finds that

$$z \to z' = \frac{az+b}{cz+d} \qquad \Rightarrow \qquad dz' = dz \to \frac{dz}{(cz+d)^2} \tag{10.32}$$

(where we made use of $ad-bc=1$), and hence (after a small calculation) that the intrinsic metric on $\mathcal{I}$ transforms according to

$$\frac{4dzd\bar{z}}{(1+|z|)^2} \to \frac{4dz'd\bar{z}'}{(1+|z'|)^2} = K^2\frac{4dzd\bar{z}}{(1+|z|)^2} \ , \tag{10.33}$$

where the conformal factor $K = K(z,\bar{z})$ is

$$K = \frac{1+|z|^2}{|az+b|^2+|cz+d|^2} = \frac{|-cz'+a|^2+|dz'-b|^2}{1+|z'|^2} \ . \tag{10.34}$$

This group is considered too large to be interesting, however.

To restrict it one can insist on a special scaling of the parameter u along the generators,

$$\hat{g}^{ab}\hat{\nabla}_a\Omega\hat{\nabla}_b u = \hat{n}^a\hat{\nabla}_a u = 1 \ . \tag{10.35}$$

Now we observe that, on $\mathcal{I}$ where $\Omega = 0$, a conformal rescaling of the metric results in a rescaling of the normal vector $\hat{n}^a$. Indeed

$$\Omega \to K\Omega \qquad \Rightarrow \qquad \begin{cases} \hat{g}_{ab} \to K^2\hat{g}_{ab} \\ \\ \hat{n}^a = \hat{g}^{ab}\hat{\nabla}_a\Omega \to \frac{1}{K}\hat{n}^a \quad . \end{cases} \tag{10.36}$$

This means that the tensor $\hat{n}^a\hat{n}^b\hat{g}_{cd}$ is invariant under conformal transformations. In effect any conformal transformation of the metric is counteracted by a rescaling of the parameter u that is parametrizing the null generators.

To preserve the special Bondi scaling of u we must therefore restrict our transformations to be

$$z \to z' = \frac{az+b}{cz+d} \ , \qquad\qquad u \to u' = K\left[u + a(z,\bar{z})\right] \ . \tag{10.37}$$

These transformations realize the BMS group. It is an infinite dimensional group due to the free function $a = a(z,\bar{z})$, which gives rise to what one calls supertranslations along the generators.

◇ **Problem 10.1** Use Eddington–Finkelstein coordinates to add $\mathcal{I}^+$ to the Schwarzschild spacetime, as in (10.14). Extend the conformal structure analytically to negative values of $1/r$. Show that what you get on the other side is a conformal copy of the negative mass Schwarzschild spacetime. (This exercise should give you a healthy respect for the point i^0.)

11 Special topic I: Riemann's normal coordinates

My conventions for the Christoffel symbols and the Riemann tensor are

$$\Gamma_{abc} = g_{ad}\Gamma^{d}{}_{bc} = \frac{1}{2}(g_{ab,c} + g_{ac,b} - g_{bc,a}) \tag{11.1}$$

$$[\nabla_a, \nabla_b]V_c = R_{abc}{}^{d}V_d \quad \Rightarrow \quad R_{abc}{}^{d} = \partial_b\Gamma^{d}{}_{ac} - \partial_a\Gamma^{d}{}_{bc} + \Gamma^{e}{}_{ac}\Gamma^{d}{}_{be} - \Gamma^{e}{}_{bc}\Gamma^{d}{}_{ae} \ . \tag{11.2}$$

Indices are placed on Γ in a way that is supposed to minimize confusion. These particular combinations of the metric and its derivatives are of course there for a reason. Riemann's original problem was to see to what extent the coordinate transformation

$$g_{ab}(x) \to g_{a'b'}(x') = \frac{\partial x^a}{\partial x^{a'}}\frac{\partial x^b}{\partial x^{b'}}g_{ab}(x) \tag{11.3}$$

can be used to bring the metric tensor into a standard form. Clearly we can make it diagonal at any given point. What is the obstruction that prevents us to do this in a region? The answer turns out to be that the first derivatives of the metric can always be set to zero at the given point (equivalently, the Christoffel symbols can be made to vanish there), but the second derivatives cannot be set to zero unless the metric tensor is a very special one. The obstruction is precisely the Riemann curvature tensor.

Riemann's procedure is of great interest since it is based on the introduction of a preferred coordinate system which is then used to Taylor expand various quantities of physical interest. Converting the resulting expressions to tensor form will be easy because all the Christoffel symbols vanish in this coordinate system. These coordinates are called *Riemann normal coordinates*, and are based on the idea of making the description of the geodesics from the chosen point as simple as it can be. This does involve a considerable amount of work, but we will at least make a start here. We begin with the transformation rule for the connection,

$$\Gamma^{a'}{}_{b'c'} = \frac{\partial x^{a'}}{\partial x^a}\frac{\partial x^b}{\partial x^{b'}}\frac{\partial x^c}{\partial x^{c'}}\Gamma^{a}{}_{bc} + \frac{\partial x^{a'}}{\partial x^e}\frac{\partial^2 x^e}{\partial x^{b'}\partial x^{c'}} \ . \tag{11.4}$$

Then we perform the coordinate change

$$x^{a'} = C^{a'}_a(x^a - x^a_0) + \frac{1}{2}C^{a'}_a\Gamma^a{}_{bc}(x_0)(x^b - x^b_0)(x^c - x^c_0) + \dots \ , \tag{11.5}$$

where x^a_0 are the original coordinates at the point P, $C^{a'}_a$ is a constant matrix that can be used to diagonalize the metric at P, and the higher order terms in the transformation are left undetermined. It follows that

$$\frac{\partial^2 x^{a'}}{\partial x^b \partial x^c}_{|P} = C^{a'}_a \Gamma^a{}_{bc}(x_0) \ . \tag{11.6}$$

On the other hand we see from the transformation rule (11.4) that

$$\frac{\partial x^{b'}}{\partial x^b}\frac{\partial x^{c'}}{\partial x^c}\Gamma^{a'}{}_{b'c'} = \frac{\partial x^{a'}}{\partial x^a}\Gamma^a{}_{bc} + \frac{\partial x^{a'}}{\partial x^e}\frac{\partial x^{c'}}{\partial x^c}\frac{\partial}{\partial x^b}\frac{\partial x^e}{\partial x^{c'}} = \frac{\partial x^{a'}}{\partial x^a}\Gamma^a{}_{bc} - \frac{\partial^2 x^{a'}}{\partial x^b \partial x^c} \ . \tag{11.7}$$

The condition (11.6) now ensures that $\Gamma^{a'}{}_{b'c'} = 0$ at the chosen point. Since the higher order terms in (11.5) were left open there is plenty of freedom left, and indeed the argument can be extended to show that the first derivatives of the metric can be made to vanish along any given curve.

Riemann normal coordinates are introduced by means of the exponential map from the tangent space $\mathbf{T}_o$ at a point o into the manifold. We define $\exp_p(V)$ as the point p reached by an affinely parametrized geodesic starting at the origin, at parameter value 0, with tangent vector V, and reaching p at parameter value 1. The point p is then assigned the components of V as coordinates. Usually it is understood that the standard Minkowski coordinates are used in $\mathbf{T}_o$. The exponential map is into the manifold if the latter is geodesically complete, but may fail to be one-to-one. A normal neighbourhood is a neighbourhood of the origin in which the exponential map is one-to-one. In a normal neighbourhood every point is reached by a unique geodesic from the origin. Such a neighbourhood always exists, and it may be convenient to cut it down so that it is convex.

In a general coordinate system the equation for a geodesic is

$$\ddot{x}^a + \Gamma^a{}_{bc}\dot{x}^b\dot{x}^c = 0 \ . \tag{11.8}$$

Let us now consider a point P and a normal neighbourhood around it. In this region the tangent vectors at P serve as coordinates, and the equation for a geodesic takes the simple form

$$v^a(\tau) = v^a\tau \ . \tag{11.9}$$

The coordinates v^a are precisely the Riemann normal coordinates. In these coordinates the Christoffel symbols vanish at the origin.