Chapter 7

# Introduction to general relativity

We provide a brief introduction to general relativity, from the perspective of Riemannian geometry. For an extensive treatment of general relativity, see traditional physics textbooks such as [**Wal84**, **MTW73**]. For a more mathematical treatment of general relativity that is more closely related to the materials presented here, see [**CB15**, **O'N83**, **HE73**]. For a good introduction to causality theory leading up to the Penrose incompleteness theorem (Theorem 7.29), see [**Gal14**]. For an excellent survey of mathematical relativity, see [**CGP10**].

## 7.1. Spacetime geometry

**7.1.1. Lorentzian geometry.** First, the basic setting of general relativity is *Lorentzian geometry.* Every symmetric bilinear form $A$ on $\mathbb{R}^n$ can be diagonalized in the sense that there exists a basis $e_1, \ldots, e_n$ such that if we express arbitrary vectors $u = u^i e_i$ and $v = v^i e_i$ in that basis (using Einstein notation), then $A(u, v) = \varepsilon_{ij} u^i v^j$, where $\varepsilon_{ij}$ is a diagonal matrix in which each diagonal entry is 1, $-1$, or 0. The number of 1's, $-1$'s, and 0's appearing as diagonal entries is independent of the choice of $e_1, \ldots, e_n$. (Prove these facts.) The basis $e_1, \ldots, e_n$ generalizes the concept of orthonormal basis in the case where $A$ is an inner product and $\varepsilon_{ij} = \delta_{ij}$ is the identity matrix. We say that $A$ is *nondegenerate* if none of the diagonal entries of $\varepsilon_{ij}$ are zero. The number of 1's and $-1$'s that occur for a nondegenerate symmetric bilinear form is referred to as its *signature.*

**Definition 7.1.** Let $\mathcal{M}$ be a manifold. We say that $\mathbf{g}$ is a *pseudo-Riemannian metric* on $\mathcal{M}$ if it assigns to each point $p$ a nondegenerate symmetric bilinear form on $T_p\mathcal{M}$ and does so in a way that smoothly depends on $p$. Or in other words, $\mathbf{g} \in C^\infty(T^*M \odot T^*M)$ and $\mathbf{g}$ is nondegenerate at each point. We say that such a $\mathbf{g}$ is *Lorentzian* if its signature has exactly one $-1$ in it.[1]

We will adopt the (unusual) convention of using a boldface $\mathbf{g}$ as our default variable for a Lorentzian metric and use boldface for all quantities computed using $\mathbf{g}$. The most important Lorentzian metric is the *Minkowski metric* $\eta$ on $\mathbb{R}^{n+1}$. It is convenient to index the components of $\mathbb{R}^{n+1}$ by the numbers $0, 1, \ldots, n$. We define $\eta$ to be

$$\eta = -(dx^0)^2 + (dx^1)^2 + \cdots + (dx^n)^2.$$

Note that in addition to being a Lorentzian metric, we can also think of $\eta$ as a nondegenerate symmetric bilinear form. Explicitly, for any $u \in \mathbb{R}^{n+1}$, we can write $u^\mu$ as its $\mu$th component in the standard basis. (We typically use Greek letters for indices running from 0 up to $n$, whereas Latin letters continue to be used for indices running from 1 to $n$.) Then for any $u, v \in \mathbb{R}^{n+1}$, we have $\eta(u, v) = \eta_{ij} u^\mu v^\nu$ , where

$$\eta_{\mu\nu} = \begin{cases} -1 & \text{for } \mu = \nu = 0, \\ 1 & \text{for } \mu = \nu > 0, \\ 0 & \text{for } \mu \neq \nu. \end{cases}$$

Note that every Lorentzian metric $\mathbf{g}$ is locally modeled on the Minkowski metric, in the sense that there always exists a basis at each point such that $\mathbf{g}_{\mu\nu} = \eta_{\mu\nu}$ at that point. This is analogous to how every Riemannian metric is locally modeled on the Euclidean metric and can always be written as $\delta_{ij}$ after choosing an orthonormal basis.

Since we usually think of $x^0$ as a time coordinate, sometimes we write $t$ in place of $x^0$ when the meaning is clear. In physics, we are primarily interested in the case $n = 3$ (for three space dimensions plus one time dimension), but in this book we will work in general dimension whenever possible.

*Special relativity* is essentially the principle that all physics in $\mathbb{R}^{n+1}$ should respect the bilinear form $\eta$. Or more precisely, all physics should be invariant under *Lorentz transformations*, which are defined to be elements of $\mathrm{GL}(n+1)$ that preserve $\eta$. They form the *Lorentz group* $\mathrm{O}(1, n)$. Of course, the Lorentz group contains the orthogonal group $O(n)$ acting on the spatial components. It also contains the orientation-reversing *time-reversal* symmetry that maps $t \mapsto -t$.

[1]Some physics texts instead use the convention that Lorentzian signature has exactly one 1 in it.

There is another important type of Lorentz symmetry called *boosts*. For every constant $v \in (-1, 1)$, consider the map

$$t \mapsto \frac{t - vx^1}{\sqrt{1-v^2}},$$
$$x^1 \mapsto \frac{x^1 - vt}{\sqrt{1-v^2}}$$

that leaves $x^2, \ldots, x^n$ unchanged. This is a boost in the $x^1$ direction with velocity $v$, but we can analogously define boosts in any spatial direction. Check that boosts preserve $\eta$. Less obvious is the fact that all of $\mathrm{O}(1, n)$ can be generated by $\mathrm{O}(n)$, time-reversal, and boosts. The group $\mathrm{O}(1, n)$ breaks into four connected components, one of which contains the identity, called the restricted Lorentz group $\mathrm{SO}^+(1, n)$, one of which contains time-reversal, one of which contains orientation-reversing symmetries of $\mathrm{O}(n)$, and the last of which contains the product of time-reversal with a spatial reflection. The union of the first and the fourth of these components forms the group $\mathrm{SO}(1, n)$ of orientation-preserving Lorentz transformations, while the union of the first and the third forms the group $\mathrm{O}^+(1, n)$ of orthochronous Lorentz transformations.

In particular, we can count the dimension of $\mathrm{O}(1, n)$. For example, the dimension of $\mathrm{O}(1, 3)$ is 6, with three dimensions coming from $\mathrm{O}(3)$ and the other three coming from boosts. We can also consider all maps of $\mathbb{R}^{n+1}$ that preserve the Minkowski metric on $\mathbb{R}^{n+1}$. These maps form the *Poincaré group*, which is a semidirect product of the Lorentz group and translations of $\mathbb{R}^{n+1}$. (This is analogous to how the isometry group of the Euclidean plane is a semidirect product of the orthogonal group and translations.) Note that when $n = 3$, the Poincaré group has 10 dimensions.

It is fairly intuitive that physical laws should be invariant under $\mathrm{O}(3)$ and under spatial translations, as well as under time translation and time-reversal. That is, if one performs a rotation or translation of the coordinate system, one expects to be able to use the exact same equations governing physical phenomena, as long as all quantities are rewritten in terms of the new coordinate system. In classical physics we also expect that physics should be invariant under a Galilean transformation that maps

$$t \mapsto t,$$
$$x^1 \mapsto x^1 - vt$$

and leaves $x^2$ and $x^3$ unchanged, where $v$ is a constant. That is, if you are moving at constant velocity $v$ in the $x^1$ direction and decide to use a new coordinate system in which your position is fixed, then you should be able to use the same equations in your coordinate system. This is especially important when you consider that everything is always moving relative to other things, so there should never be a preferred *reference frame* (that is, choice of coordinate system for space and time), and therefore we need some principle for moving between different reference frames. "Galilean relativity" was once believed to accomplish this.

The starting point of special relativity was the observation that Maxwell's equations for electromagnetism [**Wik**, Maxwell's_equations], in addition to being invariant under Euclidean isometries (and time translation and time-reversal), were also invariant under boosts, where we use units in which the speed of light $c$ is equal to 1. Explicitly, if one includes the factors of $c$, the boost above would be written as

$$t \mapsto \frac{t - \frac{vx^1}{c^2}}{\sqrt{1 - \left(\frac{v}{c}\right)^2}},$$

$$x^1 \mapsto \frac{x^1 - vt}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}.$$

In particular, Maxwell's equations are not invariant under Galilean transformations. Although the Lorentz invariance of Maxwell's theory was well-understood by others, one of Einstein's insights was to take the Lorentz invariance seriously as an underlying feature of reality. Note that for $v \ll c$, the Galilean transformation is a good approximation of the corresponding boost. Using boosts, one can derive basic features of special relativity such as time dilation and length contraction. One consequence of the Lorentz invariance is that although time has a "distinguished status" compared to the spatial directions, one can no longer meaningfully separate space and time, since boosts "mix" the two in some sense.

In general relativity, we take things a step further by asking for physics to be invariant under general coordinate transformations, but we still want our geometry to be locally modeled on special relativity (that is, Minkowski space). To that end, the setting of general relativity is a Lorentzian manifold $(\mathcal{M}^{n+1}, \mathbf{g})$. Given $p \in \mathcal{M}$, a tangent vector $v \in T_p\mathcal{M}$ is said to be *spacelike* if $\mathbf{g}(v, v) > 0$, *timelike* if $\mathbf{g}(v, v) < 0$, or *null* if $\mathbf{g}(v, v) = 0$. Note that the null vectors form a double-sided cone in $T_p\mathcal{M}$, called the *null cone* or *light cone*, which separates timelike vectors from spacelike ones. See Figure 7.1. In a Lorentzian manifold, we say that a curve is spacelike, timelike, or null if its tangent directions are always spacelike, timelike, or null, respectively.

**7.1.2. Causal structure and global hyperbolicity.** A physical object traces out a "worldline" in $\mathcal{M}$, that is, a curve[2] in the Lorentzian manifold.

[2] In this section, we will assume our curves to be piecewise smooth unless stated otherwise.

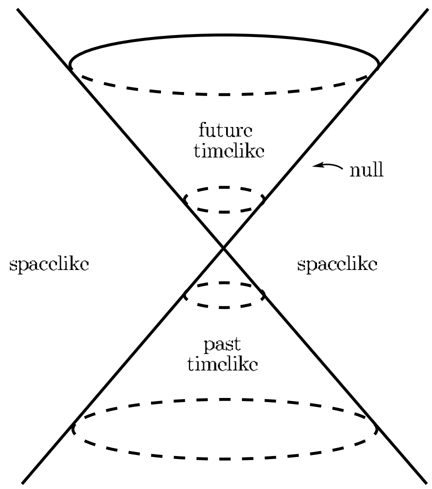


**Figure 7.1.** The null cone in $T_p\mathcal{M}$.

For massive objects, this curve is always timelike, and for massless objects (such as photons), the curve is always null (meaning that the object is traveling at the speed of light). Because of this, we use the word *causal* to describe vectors or curves that can be timelike or null. Essentially, something that happens at a certain point in $\mathcal{M}$ can affect another point in $\mathcal{M}$ only if they can be joined by a causal curve. We would like to think of one of those points as being in the future of the other. In order to do this, we must choose which side of the null cone at each $p \in \mathcal{M}$ is the future and which is the past. If such a choice can be made smoothly consistently over all of $\mathcal{M}$, that choice is called a *time-orientation* of $(\mathcal{M}, \mathbf{g})$. Note that a choice of global timelike vector field on $\mathcal{M}$ determines a time orientation. Technically, choosing a time-orientation is equivalent to reducing the structure group from $\mathrm{O}(1, n)$ down to $\mathrm{O}^+(1, n)$.

**Definition 7.2.** When we refer to a *spacetime* in this text, we will mean a connected, time-oriented Lorentzian manifold.

Given a time orientation, we can separate the nontrivial timelike and null vectors into those that are *future pointing* and those that are *past pointing*. Given a point $p \in \mathcal{M}$, we define $J^+(p)$ (respectively, $J^-(p)$) to be the *causal future* (respectively, *causal past*) of $p$, meaning the set of all points that can be reached from $p$ by future-pointing (respectively, past-pointing) causal curves. The collection of sets $J^+(p)$ and $J^-(p)$ can be referred to as the *causal structure* of a spacetime $(\mathcal{M}, \mathbf{g})$. We also define $I^+(p)$ (respectively, $I^-(p)$) to be the *chronological future* (respectively, *chronological past*) of

$p$, meaning the set of all points that can be reached from $p$ by future-pointing (respectively, past-pointing) timelike curves. Given a set $S$, we define $J^{\pm}(S) = \bigcup_{p \in S} J^{\pm}(p)$ and $I^{\pm}(S) = \bigcup_{p \in S} I^{\pm}(p)$. As a convention, we consider $p \in J^{\pm}(p)$ but $p \notin I^{\pm}(p)$.

It turns out that the causal structure of a spacetime is equivalent to the conformal structure together with the time orientation. To see this, note that the information contained in the causal structure is equivalent to knowledge of what the time-oriented null cones are at each point. Next, a simple argument (as in [**HE73**, p. 61], for example) shows that at any point $p \in \mathcal{M}$ the null cone in $T_p\mathcal{M}$ determines the Lorentzian metric on $T_p\mathcal{M}$ up to a constant (and vice versa, of course).

Given a causal curve $\gamma : [a, b] \longrightarrow \mathcal{M}$, we can define its *length* to be

$$L(\gamma) = \int_a^b \sqrt{-\mathbf{g}(\gamma'(t), \gamma'(t))}\, dt.$$

For a timelike curve, the length measures *proper time*, which is the amount of time that the object experiences, or, in other words, how much time has passed according to a clock traveling with the object. A timelike curve can be parameterized by proper time. A test particle moving under the influence of gravity and no other forces will travel along a geodesic (timelike if it is massive and null if it is massless). Note that the Levi-Civita connection and geodesics in a Lorentzian manifold are defined in the exact same way as in Riemannian manifolds. (Raising and lowering of indices is also defined the same way.) However, in contrast to Riemannian geometry, a timelike geodesic locally *maximizes* length if and only if it is free of conjugate points.

In the rest of this subsection we study the existence of causal geodesics.

**Lemma 7.3.** *Given a spacetime $(\mathcal{M}, \mathbf{g})$, if $\gamma$ is a causal curve from $p$ to $q$ in $\mathcal{M}$ that is not null everywhere, then $\gamma$ can be deformed to a timelike curve from $p$ to $q$.*

We omit the proof, but note that it ultimately rests on the fact that it holds in small balls covering the curve.

**Proposition 7.4.** *Given a spacetime $(\mathcal{M}, \mathbf{g})$, let $p, q \in \mathcal{M}$ such that $q \in J^+(p) \setminus I^+(p)$. Then there exists a future null geodesic from $p$ to $q$.*

**Idea of the proof.** Observe that Lemma 7.3 implies that if $q \in J^+(p) \setminus I^+(p)$, then they are joined by a null curve. Suppose that null curve joining $p$ to $q$ is not geodesic. In this case we can fairly explicitly deform the curve to a causal curve from $p$ to $q$ that is timelike somewhere. Think about the situation in Minkowski space in order to understand the intuition behind this claim. (If a null curve is not geodesic, then it must "spiral" as it moves upward in time.) See [**O'N83**, Proposition 46, **HE73**, Proposition 4.5.10] for details. But if there is a causal curve from $p$ to $q$ that is timelike somewhere, then Lemma 7.3 says that $q \in I^+(p)$, which is a contradiction. □

We will also be interested in various types of hypersurfaces of $(\mathcal{M}, \mathbf{g})$. We say that a hypersurface is *timelike* if its normal vector is always spacelike, *spacelike* if its normal vector is always timelike, or *null* if its normal vector is always null. Keep in mind that in the last case, that null normal vector will actually be tangent to the hypersurface. Equivalently, a hypersurface is timelike if $\mathbf{g}$ induces a Lorentzian metric on it, spacelike if $\mathbf{g}$ induces a Riemannian metric on it, or null if $\mathbf{g}$ induces a degenerate bilinear form on it.

While Proposition 7.4 can be used to establish the existence of null geodesics, it turns out that in order to construct maximizing timelike geodesics, one usually needs the hypothesis of *global hyperbolicity*, which plays a role in spacetime geometry similar to that of completeness in Riemannian geometry. Completeness is often an undesirable assumption for spacetimes, because important examples such as the Schwarzschild spacetime are incomplete.

**Definition 7.5.** Given a spacetime $(\mathcal{M}, \mathbf{g})$, we say that a smooth hypersurface $M$ is a *Cauchy hypersurface* if every inextendible causal curve must pass through $M$ exactly once. A spacetime that contains a Cauchy hypersurface is called *globally hyperbolic*.

Here, an *inextendible* causal curve $\gamma : I \longrightarrow \mathcal{M}$ is simply one that cannot be extended to a causal curve on a larger domain. Be aware that there is another common definition of Cauchy hypersurface used in the literature that is broader than this one. Interestingly, although we define a Cauchy hypersurface to be a smooth hypersurface for simplicity, more generally one can actually show that any set satisfying the Cauchy hypersurface condition is a closed $C^0$ hypersurface in $\mathcal{M}$. See [**Gal14**].

**Proposition 7.6.** *Given a globally hyperbolic spacetime* $(\mathcal{M}, \mathbf{g})$*, let* $p, q \in \mathcal{M}$ *such that* $q \in I^+(p)$*. Then there exists a length-maximizing future-pointing timelike geodesic from* $p$ *to* $q$*.*

We omit the proof, but as one might expect, it bears some similarity to the proof that there exists a length-minimizing geodesic connecting any two points in a Riemannian manifold. See [**Gal14**] or [**HE73**, Section 6.7] for details. The global hyperbolicity is used because it allows us to execute the needed compactness argument, though this is not at all obvious from the way we defined global hyperbolicity. (There is another common definition of global hyperbolicity that is known to be equivalent to ours [**Ger70**].)

The following proposition of R. Geroch [**Ger70**] gives us a simple picture of globally hyperbolic spacetimes.

**Proposition 7.7.** *If $M$ is a Cauchy hypersurface in $(\mathcal{M}, \mathbf{g})$, then there exists a homeomorphism between $\mathcal{M}$ and $\mathbb{R} \times M$ that provides a foliation of $\mathcal{M}$ by Cauchy hypersurfaces. Moreover, any Cauchy hypersurface in $\mathcal{M}$ must be homeomorphic to $M$.*

See [**Gal14**, **HE73**, Section 6.6] for details. Observe that one simple consequence of the above proposition is that globally hyperbolic spacetimes have the desirable property that there are no closed causal curves (which would violate the physical concept of "causality").

**7.1.3. Static spacetimes.** We can construct simple examples of spacetimes by simply taking a Lorentzian warped product of a Riemannian manifold with a line. That is, given a Riemannian manifold $(M, g)$ and a positive warping function $N$ on $M$, we consider the Lorentzian warped product metric

$$\mathbf{g} = -N^2 dt^2 + g \tag{7.1}$$

on $\mathbb{R} \times M$. Since $N$ and $g$ are independent of the time coordinate $t$, this may be regarded as a spacetime metric that "does not change in time," and consequently we say that $\mathbf{g}$ is *static*. Formally, we have the following definition.

**Definition 7.8.** A spacetime is *static* if it admits a global, nonvanishing timelike Killing field whose orthogonal distribution is involutive.

It is easy to see that a spacetime is static if and only if it can be locally written in the form (7.1), where the nonvanishing Killing field is just $\partial_t$, which is orthogonal to the level sets of $t$.

## 7.2. The Einstein field equations

To summarize the previous section, we view our universe as a manifold $\mathcal{M}^{n+1}$ equipped with a Lorentzian metric $\mathbf{g}$, which we regard as the "gravitational field." The data $(\mathcal{M}, \mathbf{g})$ determines how test particles move under the influence of gravity. This leads us to the natural question of what determines the spacetime metric itself. The metric ought to somehow be determined by the distribution of gravitational sources. Those sources are represented by a symmetric $(0, 2)$-tensor $T$ called the *stress-energy tensor*, rather than a single mass density function as in Newtonian gravity. If we regard a tangent vector $v \in T_p\mathcal{M}$ as an "observer" at the spacetime point $p$, then $T(v, \cdot)$ represents the energy-momentum density of the source as seen by the observer $v$. Given this tensor $T$, the spacetime metric $\mathbf{g}$ must satisfy the *Einstein (field) equations*,

$$\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g} = (n-1)\omega_{n-1}T,$$

where the Riemann, Ricci, and scalar curvatures of a Lorentzian metric $\mathbf{g}$ are defined exactly the same as for a Riemannian metric.[3] Or we could simply write $\mathbf{G} = (n-1)\omega_{n-1}T$, where $\mathbf{G}$ is the Einstein tensor of $\mathbf{g}$.

An important generalization of the Einstein equations is the *Einstein field equations with cosmological constant* $\Lambda$,

$$\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g} + \Lambda\mathbf{g} = (n-1)\omega_{n-1}T,$$

where $\Lambda$ is a constant. In this book, we will focus on the case $\Lambda = 0$.

**7.2.1. Motivation for the Einstein equations.** We will first give some heuristic physical reasoning to motivate the Einstein equations. (Our discussion here loosely follows that of [**Ger13**].) We have decided that in general relativity, we want test particles to follow geodesics. We know that curvature shows up when we look at variations of geodesics. That is, if $\gamma_s(t)$ is a family of geodesics through $\gamma_0 = \gamma$ and $X = \left.\frac{\partial\gamma}{\partial s}\right|_{s=0}$ is its first-order variation, then the *Jacobi equation* [**Wik**, Jacobi_field] states that

$$\langle X'', Y\rangle = -\mathbf{Riem}(X, \gamma', Y, \gamma'), \tag{7.2}$$

where $Y$ is any other vector field along $\gamma$, and the primes denote (covariant) differentiation in the $t$-variable.

Let us compare this to what happens in Newtonian gravity (in the $n = 3$ case). Recall (from Section 3.1.3 or otherwise) that in Newtonian gravity, there is a gravitational potential function $V : \mathbb{R}^3 \longrightarrow \mathbb{R}$ satisfying $\Delta V = 4\pi\rho$, where $\rho$ represents the mass density of all sources. The acceleration of any test particle is $-\nabla V$. Therefore, if we imagine a family of Newtonian trajectories $\gamma_s(t)$ through $\gamma_0 = \gamma$ with first-order variation $X = \left.\frac{\partial\gamma}{\partial s}\right|_{s=0}$, then it is straightforward to derive that

$$X'' \cdot Y = -\nabla_Y \nabla_X V,$$

where $Y$ is any vector field along $\gamma$ and the left side is just dot product. In Newtonian gravity, we know that the trace of the right side of the above equation (over the $X$ and $Y$ slots) is $-4\pi\rho$. If we take the trace of equation (7.2) (noting that the full trace over $X$ and $Y$ will give the same result as

[3]When $n = 3$, the constant $(n-1)\omega_{n-1}$ is just $8\pi$. For $n > 3$, there does not seem to be a universally accepted convention for what this constant should be, but ours is chosen to be consistent with our earlier definition of ADM mass (Definition 3.9). Unfortunately, Exercise 7.9 suggests a different convention based on physics, but we have decided not to use this convention because it would require an extra factor of $(n-2)$ in many of our formulas.

tracing over the spatial directions orthogonal to $\gamma'$), we see that the quantity $\mathbf{Ric}(\gamma', \gamma')$ should correspond to $4\pi\rho$.

Of course, $\mathbf{Ric}(\gamma', \gamma')$ depends on the observer $\gamma'$ while $4\pi\rho$ does not. This dependence on observer is where the stress-energy tensor $T$ comes in. As described above, the energy density of the sources, as seen by the observer $\gamma'$, is $T(\gamma', \gamma')$, suggesting that $\rho$ should be identified with $T(\gamma', \gamma')$. However, one could also reasonably identify $\rho$ with $-\operatorname{tr} T$ in the Newtonian limit. Compromising between these two candidates for a timelike observer suggests that we take

$$\mathbf{Ric} = 4\pi(rT + (1-r)(\operatorname{tr} T)\mathbf{g})$$

for some constant $r$. From a physical perspective, we would like $T$ to be a conserved quantity, that is, $\operatorname{div} T = 0$.

**Exercise 7.9.** Show that, in order for $\operatorname{div} T = 0$ to be consistent with the contracted second Bianchi identity (Exercise 1.10), we should choose $r = 2$, and hence

$$\mathbf{Ric} - \frac{1}{2}\mathbf{R}\mathbf{g} = 8\pi T.$$

**7.2.2. Lagrangian formulation and matter models.** The simplest case of the Einstein equations is the *vacuum* case when $T$ is identically zero. This means that there are no sources at all for the gravitational field. Minkowski space is our basic model of empty space (with $\Lambda = 0$), but one of the most important features of general relativity is that there is an incredibly rich theory even in the vacuum case.

The vacuum Einstein field equations can also be derived from an action principle. The *Einstein-Hilbert action* is the functional

$$S[\mathbf{g}] = \int_{\mathcal{M}} \mathbf{R}\, d\mu_{\mathbf{g}}.$$

(Note that for a Lorentzian metric, the expression for $d\mu_{\mathbf{g}}$ with respect to a local frame involves $\sqrt{-\det \mathbf{g}}$ instead of $\sqrt{\det \mathbf{g}}$, since the $\det \mathbf{g}$ will be negative.) It is easy to see from a formal computation that the Euler-Lagrange equations for the Einstein-Hilbert action are precisely the vacuum Einstein field equations. That is, the variation of $S$ at $\mathbf{g}$ in the direction of every compactly supported variation of $\mathbf{g}$ vanishes if and only if $\mathbf{G} = 0$. (Check that this follows immediately from Exercise 1.18 and Proposition 1.3.) The Einstein field equations with cosmological constant $\Lambda$ arise from using the action $S[\mathbf{g}] = \int_{\mathcal{M}} (\mathbf{R} - 2\Lambda)\, d\mu_{\mathbf{g}}$.

Finally, observe that the vacuum Einstein equations $\mathbf{G} = 0$ imply that $\mathbf{R} = 0$, and thus the *vacuum Einstein equations* can be rewritten as

$$\mathbf{Ric} = 0.$$

More generally, if one wants to build a model of general relativity in which other fields interact with gravity, then one would add extra terms to the Einstein-Hilbert action for those fields, and then the resulting Euler-Lagrange equations will determine what the stress-energy tensor $T$ is, as well as give equations governing those other fields: the Einstein equations coupled with other field equations such as Maxwell, Klein-Gordon, Yang-Mills, Vlasov, etc. These other fields are generally referred to as "matter fields."

As a quick example of how this works, if we wish to include the electromagnetic field, we introduce an electric potential 1-form $A$ and define the electromagnetic field to be $F = dA$. If we define an action

$$S[\mathbf{g}, A] = \int_M (\mathbf{R} - F_{\mu\nu}F^{\mu\nu})\, d\mu_{\mathbf{g}},$$

then its Euler-Lagrange equations from the variation of $\mathbf{g}$ will be

$$\mathbf{G}_{\mu\nu} = 2(F_{\mu\alpha}F_\nu^{\ \alpha} - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\mathbf{g}_{\mu\nu}),$$

so that we can identify the right side as $8\pi T_{\mu\nu}$. The Euler-Lagrange equations from the variation of $A$ will be

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

These equations are together known as the *Einstein-Maxwell* equations, and solutions are called *electrovacuum* since there are no sources.

**7.2.3. The Schwarzschild spacetime.** Selecting a specific matter field model is necessary if one wishes to *solve* the Einstein field equations. Historically, the first and most important nontrivial solution of the vacuum Einstein equations was the Schwarzschild metric. The Schwarzschild metric comes about in a natural way by looking for a solution with a lot of symmetry. Specifically, we can look for a solution that is both static and spherically symmetric. Or in other words, we consider the ansatz

$$\mathbf{g} = -N(r)^2dt^2 + V(r)^{-1}dr^2 + r^2d\Omega^2, \tag{7.3}$$

where $d\Omega^2$ is the standard metric on a unit $S^{n-1}$ sphere, and we try to find $N$ and $V$ such that $\mathbf{Ric} = 0$. Observe that a constant $t$ slice is totally geodesic in the ambient spacetime metric $\mathbf{g}$. By the traced Gauss equation (Corollary 2.7), it follows that the Riemannian metric $g = V(r)^{-1}dr^2 + r^2d\Omega^2$ is scalar-flat. By Exercise 3.2, it follows that $V(r) = 1 - \frac{2m}{r^{n-2}}$ for some parameter $m$. Next we solve for $N$. We use the following general fact about static metrics.

**Exercise 7.10.** Use Proposition 1.13 to show that the static metric $\mathbf{g} = -N^2dt^2 + g$ solves the vacuum Einstein equations if and only if the pair $(g, N)$ is vacuum static initial data in the sense of Definition 6.7 and $N$ is strictly positive. Note that this explains why we used the words *vacuum static* in Definition 6.7.

By the exercise above, we just need to find $N(r)$ such that $\mathrm{Hess}_g N = N\mathrm{Ric}_g$, where $g = V(r)^{-1}dr^2 + r^2 d\Omega^2$. In particular, $\Delta_g N = 0$. Writing this out as a second-order ODE for $N(r)$, it is fairly straightforward to see that $N = \sqrt{V}$ is a solution. (It is also the unique one with the property that $\mathbf{g}$ approaches Minkowski space as $r \to \infty$.) Thus we have determined what $N$ and $V$ must be.

**Exercise 7.11.** Use Proposition 1.13 to check that with this choice of $N$ and $V$, we have $\mathrm{Hess}_g N = N\mathrm{Ric}_g$.

Thus we define the *Schwarzschild spacetime metric of mass* $m$ to be

$$\mathbf{g}_m = -\left(1 - \frac{2m}{r^{n-2}}\right) dt^2 + \left(1 - \frac{2m}{r^{n-2}}\right)^{-1} dr^2 + r^2 d\Omega^2. \tag{7.4}$$

This metric solves the vacuum Einstein equations in the region where $r$ is greater than $r_0 := (2m)^{\frac{1}{n-2}}$, which is called the *Schwarzschild radius.* Our argument above proves that the Schwarzschild metric is the only spherically symmetric, *static* spacetime metric that solves the vacuum Einstein equations, but more generally Birkhoff's theorem states that it is the only spherically symmetric solution of the vacuum Einstein equations [**Bir23**]. This result holds even locally.

In the following we will assume that $m$ is positive (in order to avoid an undesirable spacelike singularity at $r = 0$). Just as was the case for the Riemannian Schwarzschild metric, the singularity at $r = r_0$ is only a coordinate singularity rather than a true geometric singularity. To see why this is the case, we rewrite the Schwarzschild metric in terms of null coordinates. If we define a new coordinate $r^*$ by

$$(dr^*)^2 = \left(1 - \frac{2m}{r^{n-2}}\right)^{-2} dr^2$$

and then define

$$u = t - r^*,$$
$$v = t + r^*,$$

then we have

$$\mathbf{g} = -\left(1 - \frac{2m}{r^{n-2}}\right) du\, dv + r^2 d\Omega^2.$$

In the literature, $r^*$ is called the Regge-Wheeler radial coordinate or tortoise coordinate, and $u$ and $v$ are called the outgoing (or retarded) and ingoing (or advanced) Eddington-Finkelstein coordinates, respectively. Note that $u$

and $v$ are null coordinates in the sense that $\partial_u$ and $\partial_v$ are null (when put together with coordinates on $S^{n-1}$ to create a coordinate system).

**Exercise 7.12.** For $n = 3$, solve for $r^*$ explicitly. Let $U = -e^{-u/4m}$ and $V = e^{v/4m}$. Show that we can choose the integration constant in the definition of $r^*$ so that $UV = \left(1 - \frac{r}{2m}\right) e^{r/2m}$, and show that the Schwarzschild metric becomes

$$\mathbf{g}_m = -\frac{32m^3}{r} e^{-r/2m} dU\, dV + r^2 d\Omega^2,$$

where we now regard $r$ as the function of $U$ and $V$ implicitly defined above. Observe that the original region $r > r_0$ corresponds to $U < 0$ and $V > 0$ in the new coordinates. The coordinates $U$ and $V$ are called *Kruskal-Szekeres coordinates.*

Note that in Kruskal-Szekeres coordinates, the metric is not singular when $U = 0$ or $V = 0$. However, one can show that there really is a singularity at $r = 0$ (because the curvature blows up there), which corresponds to where $UV = 1$. This allows us to naturally extend the Schwarzschild metric to the product $\{(U, V) \in \mathbb{R}^2 \,|\, UV < 1\} \times S^2$. See Figure 7.2. We will refer to this vacuum spacetime as the *Schwarzschild spacetime of mass $m$*, though it is sometimes called the Kruskal extension of Schwarzschild or the Kruskal-Szekeres spacetime in the literature. This spacetime can be thought of as a "maximal extension" of the metric (7.4) that was defined in the region $r > r_0$ (region I in Figure 7.2), in the sense that it is a simply connected vacuum extension such that every geodesic can either be extended to a complete geodesic, or else it hits the singularity at $UV = 1$.

The construction generalizes to higher dimensions, with the main difference being that one no longer has a simple formula for $r^*$. However, if one defines $U = -\exp\left(\frac{-(n-2)u}{2(2m)^{1/(n-2)}}\right)$ and $V = \exp\left(\frac{(n-2)v}{2(2m)^{1/(n-2)}}\right)$, then the extension works out in essentially the same way. See [**Chr15**, Remark 1.2.6] for details.

Observe that the region where $U > 0$ and $V < 0$ (region III in Figure 7.2) is just an isometric copy of the "original" region where $U < 0$ and $V > 0$, and thus the Schwarzschild spacetime should be thought of as having two "asymptotically flat" ends since these two regions resemble Minkowski space for large $r$. In particular, any negative constant $U/V$ slice (including the sphere at $U = V = 0$) of the Schwarzschild spacetime (which extends a constant $t$ slice of the original $r > r_0$ region) is precisely the two-ended Riemannian Schwarzschild space described in Chapter 3. The metric in the regions where $U > 0$ and $V > 0$ (region II), and where $U < 0$ and $V < 0$ (region IV), which are isometric to each other, can be identified with the metric (7.4) in the region $0 < r < r_0$, where the $t$ coordinate becomes spacelike and the $r$ coordinate becomes timelike.

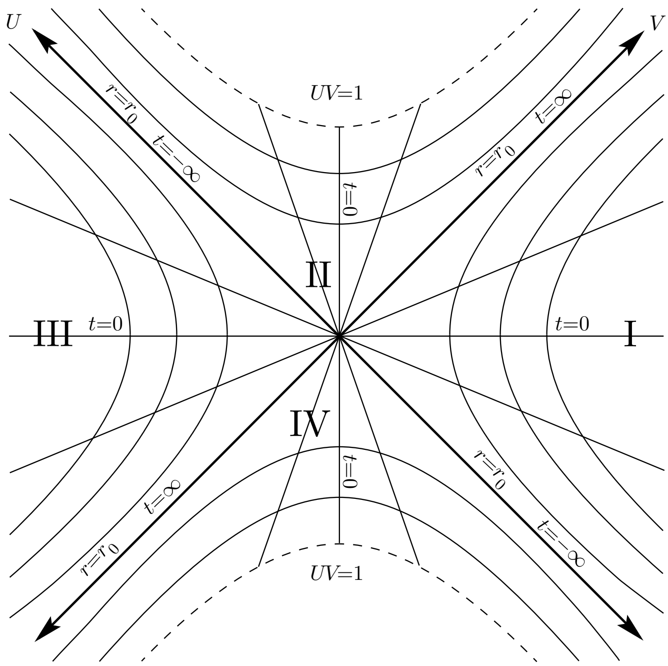


**Figure 7.2.** The Schwarzschild spacetime (aka the Kruskal-Szekeres spacetime), with some of the level sets of $r$ and $t$ drawn in. Each point in the diagram represents a sphere.

Another particularly important family of vacuum solutions in the $n = 3$ case especially is the *Kerr spacetime*, which is axisymmetric and comes in a family with two parameters—the mass $m > 0$ and the angular momentum $a$ with $0 \leq a \leq m$. *Axisymmetric* here means that there is an $S^1 \cong \mathrm{SO}(2)$ group of isometries. In contrast, the spherical symmetry of the Schwarzschild spacetime with $n = 3$ means that there is a full SO(3) group of isometries. Although the Kerr metric can be written down explicitly, we will not need this formula. See [**Wik**, Kerr_metric] for details. When $a = 0$, the formula reduces to the one for the Schwarzschild metric in (7.4). The Kerr spacetime has the property that it is *stationary*, meaning that there exists a global Killing field that is timelike near infinity. (Note that this is a much looser condition than being static, but it only really makes sense as a global condition since it asks us to look near infinity.) Thus, the Kerr spacetime has a two-parameter family of isometries. There is a higher-dimensional version of the Kerr spacetime known as the Myers-Perry spacetime [**MP86**]. The Kerr family of vacuum solutions also generalizes to the Kerr-Newman family [**Wik**, Kerr-Newman_metric] of electrovacuum solutions, which is also axisymmetric and stationary but carries an additional charge parameter.

As a special case of Kerr, the Schwarzschild spacetime is also stationary: the Killing vector field $\partial_t$ for the metric (7.4) naturally extends across the set where $UV = 0$ to a global Killing field on the entire Schwarzschild spacetime. Explicitly, it is $\frac{(n-2)}{2(2m)^{1/(n-2)}}(-U\partial_U + V\partial_V)$. However, note that the Killing field becomes spacelike in the region $UV > 0$ and is null where $UV = 0$.

## 7.3. The Einstein constraint equations

**7.3.1. The Einstein equations as an initial data problem.** It is natural to want to formulate and solve an initial value problem for the vacuum Einstein equations. That is, given knowledge of the metric at some fixed time, we would like to know how the solution must develop as we move forward in time. Since "fixed time" does not carry direct meaning in general relativity, this is usually taken to mean a fixed spacelike slice of the spacetime $\mathcal{M}$. This basic problem was essentially solved in the pioneering work of Yvonne Choquet-Bruhat [**FB52**]. If one writes out the Einstein equations in local coordinates, one can see that they do not quite form a hyperbolic system of equations. The underlying reason why they cannot be hyperbolic is that the Einstein equations are invariant under diffeomorphisms (or as physicists might say, they are "gauge invariant"). Choquet-Bruhat discovered that if one chooses to write the equations in "wave coordinates," that is, coordinates that solve the wave equation for the metric $\mathbf{g}$, then the equations become hyperbolic, and therefore they could be solved (for a short time) using existence theory for hyperbolic systems. In this formulation, one does not require all of $\mathbf{g}$ and its time derivative at $M$, but rather only the induced metric $g$ and its time derivative. However, in order for the "wave coordinate" condition to be preserved as one solves for $\mathbf{g}$ forward in time, it is necessary for $g$ and its time derivative to satisfy certain equations. Explicitly, we have the following theorem.

**Theorem 7.13** (Choquet-Bruhat [**FB52**])**.** *Let $(M^n, g)$ be a Riemannian manifold, and let $k$ be a symmetric $(0,2)$-tensor on $M$ such that $(g,k) \in W^{m+1,2} \times W^{m,2}$ for some $m > n/2$. Suppose that the following equations hold:*

$$R_g + (\operatorname{tr}_g k)^2 - |k|^2 = 0,$$
$$\operatorname{div}_g k - d(\operatorname{tr}_g k) = 0.$$

*Then there exists a spacetime $(\mathcal{M}, \mathbf{g})$ solving the vacuum Einstein equations such that $(M, g)$ isometrically embeds into $(\mathcal{M}^{n+1}, \mathbf{g})$ as a Cauchy hypersurface with second fundamental form $k$.*

When we say that the second fundamental form is $k$, we mean that if $\vec{n}$ is the future-pointing timelike unit normal to $M$, then $k = \langle \mathbf{A}, -\vec{n} \rangle$, where $\mathbf{A}$ is the second fundamental form of $M$ in $(\mathcal{M}, \mathbf{g})$, defined precisely as in Chapter 2. This is essentially a short-time existence theorem in the sense that $\mathcal{M}$ need only be a small neighborhood of $M$. It helps explain the reason why we use the vocabulary *Cauchy hypersurface*. The following theorem establishes existence of a unique solution that is maximal in some sense and illustrates the importance of the globally hyperbolic condition.

**Theorem 7.14** (Choquet-Bruhat–Geroch [**CBG69**]). *Let $(M^n, g)$ be a smooth Riemannian manifold, and let $k$ be a smooth symmetric $(0,2)$-tensor on $M$. Suppose that the following equations hold:*

$$R_g + (\operatorname{tr}_g k)^2 - |k|_g^2 = 0,$$
$$\operatorname{div}_g k - d(\operatorname{tr}_g k) = 0.$$

*Then there exists a spacetime $(\mathcal{M}^{n+1}, \mathbf{g})$ solving the vacuum Einstein equations such that $(M, g)$ isometrically embeds into $(\mathcal{M}, \mathbf{g})$ as a Cauchy hypersurface with second fundamental form $k$. Moreover, this solution is the unique (up to isometry) maximal globally hyperbolic solution. By this we mean that $(\mathcal{M}, \mathbf{g})$ does not sit inside any larger globally hyperbolic spacetime.*

*The spacetime $(\mathcal{M}, \mathbf{g})$ is called the* vacuum development *of the initial data $(M, g, k)$.*

**Exercise 7.15.** Let $(M^n, g)$ be a Riemannian manifold isometrically embedded in a Lorentzian manifold $(\mathcal{M}^{n+1}, \mathbf{g})$ with second fundamental form $k$, and let $\vec{n}$ be a unit normal vector to $M$. Show that along $M$, we have

$$\mathbf{G}(\vec{n}, \vec{n}) = \frac{1}{2}\left[R_g + (\operatorname{tr}_g k)^2 - |k|_g^2\right],$$
$$\mathbf{G}(\vec{n}, \cdot) = \operatorname{div}_g k - d(\operatorname{tr}_g k),$$

where $\mathbf{G}$ is the Einstein tensor of $\mathbf{g}$ and the $\cdot$ input is a vector tangent to $M$.

In light of the exercise above, it becomes clear why the assumptions on $g$ and $k$ in Theorems 7.13 and 7.14 are necessary.

Given an observer whose worldline in a spacetime $(\mathcal{M}, \mathbf{g})$ has causal tangent vector $v$, recall that the quantity $T_{\mu\nu}v^\mu$ may be regarded as the energy-momentum density of the gravitational sources, as seen by the observer. We say that $(\mathcal{M}, \mathbf{g})$ satisfies the *dominant energy condition (or DEC)* if for any future-pointing causal vector $v$, the covector $T_{\mu\nu}v^\mu$ is always future-pointing causal. This a natural physical assumption to impose on the gravitational sources that roughly corresponds to the assumption of nonnegative mass densities in Newtonian gravity. Note that the dominant energy condition is a general condition on a spacetime and has nothing to do with any particular matter field, but many physical models will naturally satisfy the dominant energy condition. This discussion motivates the following definition.

**Definition 7.16.** An *initial data set* $(M^n, g, k)$ is a Riemannian manifold $(M, g)$ equipped with a symmetric $(0,2)$-tensor $k$. We define

$$\mu := \frac{1}{2}\left[R_g + (\operatorname{tr}_g k)^2 - |k|_g^2\right],$$
$$J := (\operatorname{div}_g k)^\sharp - \nabla(\operatorname{tr}_g k).$$

The quantity $\mu$ is called the *energy density* while $J$ is called the *current density*. Here we have defined $J$ to be a vector quantity rather than a 1-form as some other texts do. (This is why we use the raising operator $\sharp$. Also note that for $n = 3$, our definition of energy density $\mu$ differs from the Newtonian mass density $\rho$ described earlier by a factor of $8\pi$.) Together, these equations are known as the *(Einstein) constraint equations*, and we will refer to $(\mu, J)$ as the *constraints* of $(g, k)$. They are called the constraints because if one is given a stress-energy tensor $T$, this determines the pair $(\mu, J)$, which constrains (but does not determine) the initial data $(g, k)$ according to the equations in Definition 7.16. In particular, in a vacuum spacetime, $T = 0$, and consequently $\mu$ and $J$ both vanish. More generally, any initial data set with vanishing $\mu$ and $J$ is said to satisfy the *vacuum (Einstein) constraint equations*, that is, the same conditions appearing in the hypotheses of Theorem 7.13. We say that $(M, g, k)$ satisfies the *dominant energy condition (or DEC)* whenever $\mu \geq |J|_g$ everywhere. We say that the *strict DEC* holds if $\mu > |J|_g$ everywhere.

We can emphasize the fact that $J$ is a divergence by writing

$$J = \operatorname{div}_g \pi,$$

where $\pi$ is the symmetric $(2,0)$ tensor defined by

$$\pi^{ij} := k^{ij} - (\operatorname{tr}_g k) g^{ij},$$

where the indices on $k$ have been raised using $g$. Note that $\pi$ contains the same information as $k$ since we can invert the relationship via

$$k_{ij} = \pi_{ij} - \frac{1}{n-1}(\operatorname{tr}_g \pi) g_{ij}.$$

We may sometimes (abusively) refer to the triple $(M, g, \pi)$ as an initial data set when the meaning is clear.

We say that $(M, g, k)$ *sits inside* a spacetime $(\mathcal{M}, \mathbf{g})$ if $M$ embeds into $\mathcal{M}$ in such a way that $\mathbf{g}$ induces $g$, and $k$ is the second fundamental form of the embedding.

In light of Theorems 7.13 and 7.14, we see that an initial data set $(M, g, k)$ solving the vacuum constraints is the appropriate Cauchy data for solving the vacuum Einstein equations. Building on the work of Theorems 7.13 and 7.14, similar theorems can be established for Einstein equations with various matter fields. For a thorough discussion of these various Cauchy problems, see the book [**CB09**].

The study of the Einstein equations using methods of hyperbolic PDE is an extensive field of current research, but it is not the focus of this book. However, we will mention a couple of the large, motivating problems in the field. In four spacetime dimensions, it is conjectured that the Kerr solutions are the only vacuum stationary spacetimes. (The analogous statement for the Myers-Perry solutions in higher dimensions is known to be false.) In fact, this conjecture is very close to being a known fact: S. Hawking [**HE73**] showed that any *analytic* vacuum stationary spacetime must be axisymmetric, and then work of Brandon Carter [**Car73**] and David C. Robinson [**Rob75**] shows that an axisymmetric vacuum stationary spacetime must either be static, or else it lies in the Kerr family. But if it is static, then one can show that it must correspond, via Exercise 7.10, to a vacuum static asymptotically flat manifold with minimal boundary. Theorem 6.25 then implies that it must be Schwarzschild, which is a special case of Kerr.

A mathematically rigorous version of this overall argument, drawing on work of various contributors, appears in a paper by P. Chruściel and J. Costa [**CC08**] (see also references cited therein). The analyticity assumption is nearly removed by work of S. Alexakis, A. Ionescu, and S. Klainerman [**AIK10**], which is strong enough to prove the uniqueness result for small perturbations of Kerr. The general topic of black hole uniqueness theorems has grown in a number of directions. See [**Rob09**] for a survey of developments.

The uniqueness of Kerr leads to a far more ambitious "final state conjecture" that all vacuum solutions of the Einstein equations should settle down to a Kerr solution in the long-time limit. A more tractable piece of this conjecture is simply the conjecture that the Kerr family is *stable* in the sense that initial data that is a small perturbation away from Kerr initial data will asymptotically settle down to a (possibly different) Kerr solution in the long-time limit. This is a highly active area of research that was set into motion by D. Christodoulou and S. Klainerman's pioneering proof of the stability of Minkowski space [**CK93**]. (H. Lindblad and I. Rodnianski later gave an alternative proof under stronger hypotheses [**LR10**].) Lydia Bieri extended the Christodoulou-Klainerman result by relaxing assumptions on both the regularity and decay [**Bie09**]. These results and questions have natural analogs for the Einstein-Maxwell equations. Specifically, it is also hoped that the Kerr-Newman family is stable and more generally that any electrovacuum solution of the Einstein equations must settle down to a

Kerr-Newman solution. Nina Zipser proved that Minkowski space is stable under evolution via the Einstein-Maxwell equations, thereby generalizing the Christodoulou-Klainerman result [**Zip09**]. In fact, a much broader version of the final state conjecture is that for various physical matter models, all matter except electromagnetic and purely gravitational energy should "radiate away" and leave us with a development that settles down to a Kerr-Newman solution.

**7.3.2. Asymptotically flat initial data sets.** In this book we are mainly interested in initial data sets. There is extensive literature on constructing initial data sets solving the Einstein constraint equations (chiefly the *conformal method*), but that is not our focus. We are primarily concerned with general properties of asymptotically flat initial data sets that satisfy the dominant energy condition.

**Definition 7.17.** Let $n \geq 3$. An initial data set $(M^n, g, k)$ is said to be *asymptotically flat* if there exists a bounded set $K$ such that $M \smallsetminus K$ is a finite union of *ends* $M_1, \ldots, M_\ell$ such that for each $M_k$, there exists a diffeomorphism

$$\Phi_k : M_k \longrightarrow \mathbb{R}^n \smallsetminus \bar{B}_1(0),$$

where $\bar{B}_1(0)$ is the standard closed unit ball, such that if we think of each $\Phi_k$ as a coordinate chart with coordinates $x^1, \ldots, x^n$, then in that coordinate chart (which we will often call the *asymptotically flat coordinate chart* or sometimes the *exterior coordinate chart*), we have

$$\begin{aligned} g_{ij}(x) &= \delta_{ij} + O_2(|x|^{-q}), \\ k_{ij}(x) &= O_1(|x|^{-q}) \end{aligned}$$

for some $q > \frac{n-2}{2}$. Moreover, we also require that $\mu$ and $J$ are integrable over $M$.

The case when $k$ is identically zero is called the *time-symmetric (or Riemannian)* case, and in this case the definition above reduces to the statement that $(M, g)$ is an asymptotically flat manifold. We reiterate that in the literature, the precise definition of asymptotic flatness can vary from paper to paper.

**Definition 7.18.** Let $(M^n, g, k)$ be a smooth, asymptotically flat initial data set. We define the *ADM energy-momentum* $(E, P)$ of an end of $M$ to be

$$E = \lim_{\rho\to\infty} \frac{1}{2(n-1)\omega_{n-1}} \int_{S_\rho} \sum_{i,j=1}^{n} (g_{ij,i} - g_{ii,j})\overline{\nu}^j \,\overline{d\mu}_{S_\rho},$$

$$P_i = \lim_{\rho\to\infty} \frac{1}{(n-1)\omega_{n-1}} \int_{S_\rho} \sum_{j=1}^{n} (k_{ij} - (\operatorname{tr}_g k) g_{ij})\overline{\nu}^j \,\overline{d\mu}_{S_\rho}$$

for $i = 1, \ldots, n$, where the right sides are calculated in the asymptotically flat coordinates of the chosen end, and barred quantities are calculated using the Euclidean metric in the end. If $(E, P)$ is future causal, then the *ADM mass* of that end is defined to be $m = \sqrt{E^2 - P^2}$. Observe that we can also write

$$P_i = \lim_{\rho\to\infty} \frac{1}{(n-1)\omega_{n-1}} \int_{S_\rho} \sum_{j=1}^{n} \pi_{ij}\overline{\nu}^j \,\overline{d\mu}_{S_\rho} \text{ for } i = 1, \ldots, n.$$

The definition above is due to Arnowitt, Deser, and Misner [**ADM60**, **ADM61**, **ADM62**]. Note that this definition of ADM energy is the exact same one we used for the ADM mass of asymptotically flat manifold. This apparent clash of nomenclature is fine because ADM mass and ADM energy may be regarded as the same thing for a time-symmetric asymptotically flat initial data set $(M, g, k = 0)$.

**Exercise 7.19.** Check that the ADM energy-momentum is well-defined for asymptotically flat initial data.

**Conjecture 7.20** ((Spacetime) positive mass conjecture)**.** *Let $(M^n, g, k)$ be a complete asymptotically flat initial data set satisfying the dominant energy condition. Then in each end the ADM energy-momentum vector $(E, P)$ is future causal. Or, in other words, $E \geq |P|$.*

**Conjecture 7.21** ((Spacetime) positive mass rigidity conjecture)**.** *Assume the hypotheses of the previous conjecture, and suppose that we also have $E = |P|$ in some end. Then $(M, g, k)$ sits inside Minkowski space.*

We will discuss the known cases of these conjectures in greater detail in the following chapter.

Noether's Theorem [**Wik**, Noether's_theorem] states that every symmetry of a physical system gives rise to a conserved quantity. In classical physics in flat space, spatial translation symmetry gives rise to conservation of total linear momentum, while time translation symmetry gives rise to conservation of total energy. Although an asymptotically flat initial data set need not have any symmetries, one can think of it as having asymptotic symmetries at infinity since the asymptotic flatness allows us to think of it as being "close" to a slice of Minkowski space near infinity. Since Minkowski space does have spacetime translation symmetries, it is possible to motivate the above definitions for the ADM energy-momentum using heuristic reasoning along these lines. The reader may recall that the Poincaré group of symmetries of Minkowski space also contains spatial rotations and boosts. The spatial rotations give rise to total angular momentum, while boosts give rise to a concept of center of mass. (In nonrelativistic physics, the Galilean transformations give rise to the usual concept of center of mass.) However, in order to define these quantities, we require stronger decay assumptions on $(g, k)$ than asymptotic flatness gives us.

**Definition 7.22.** A smooth asymptotically flat initial data set $(M^n, g, k)$ is said to satisfy the *Regge-Teitelboim conditions* if we have

$$g_{ij}^{\text{odd}}(x) = O_2(|x|^{-q-1}),$$
$$k_{ij}^{\text{even}}(x) = O_2(|x|^{-q-2})$$

in the asymptotically flat coordinate chart, where $q > \frac{n-2}{2}$ and we define

$$g_{ij}^{\text{odd}}(x) := g_{ij}(x) - g_{ij}(-x),$$
$$k_{ij}^{\text{even}}(x) := k_{ij}(x) + k_{ij}(-x).$$

If $(M, g, k)$ satisfies the Regge-Teitelboim conditions, we define the *ADM angular momentum*

$$\mathbb{J}_{\ell m} = \lim_{\rho\to\infty} \frac{1}{(n-1)\omega_{n-1}} \int_{S_\rho} \sum_{i,j=1}^{n} (k_{ij} - (\text{tr}_g\, k)g_{ij}) Z_{\ell m}^i \overline{\nu}^j \, \overline{d\mu}_{S_\rho}$$

for $1 \le \ell < m \le n$, where $Z_{\ell m}$ is the vector field $x^\ell \partial_m - x^m \partial_\ell$ generating rotations around the plane perpendicular to the $x_\ell x_m$-plane. In the case $n = 3$, we instead simply write $\mathbb{J}_k$ for the angular momentum around the $x^k$-axis.

If we moreover know that the ADM energy $E$ is nonzero, we define the *ADM center of mass* to be

$$C^\ell = \lim_{\rho\to\infty} \frac{1}{2(n-1)\omega_{n-1}E} \int_{S_\rho} \sum_{i,j=1}^{n} [x^\ell (g_{ij,i} - g_{ii,j})\overline{\nu}^j - (g_{i\ell}\overline{\nu}^i - g_{ii}\overline{\nu}^\ell)] \, \overline{d\mu}_{S_\rho}$$

for $\ell = 1, \ldots, n$.

Lan-Hsuan Huang [**Hua09**] showed that the center of mass can also be written as

$$C^\ell = \lim_{\rho\to\infty} \frac{1}{2(n-1)(n-2)\omega_{n-1}E} \int_{S_\rho} G(Z^\ell, \overline{\nu}) \, \overline{d\mu}_{S_\rho},$$

where $G$ is the Einstein tensor of $g$ and $Z^\ell = \sum_{i=1}^n (|x|^2 \delta^{\ell i} - 2x^\ell x^i)\partial_i$ generates a conformal symmetry of Euclidean space. This can be proved along similar lines as in the proof of Theorem 3.14. See [**MT16**].

## 7.4. Black holes and Penrose incompleteness

In this section we introduce the important concept of black holes. Unfortunately, we will have to be a bit loose with this discussion to avoid getting bogged down in technical definitions. Because we are using nonrigorous definitions, we will be careful not to state "official" theorems depending on them. One reason why we take this vague approach is that the optimal theorems can sometimes be quite sensitive to the precise definitions used, and another is that the issue of what the "right" definitions are is itself often an interesting research question. In any case, the purpose of our discussion of black holes is only to provide some physical context behind the geometric ideas to be studied later on. In particular, we would like to motivate the physical relevance of *marginally outer trapped surfaces* and *apparent horizons*, the latter of which we already introduced in Chapter 4 in the time-symmetric case.

We are interested in spacetimes that are "asymptotically flat" in some sense. This means that they should "look like" Minkowski space "near infinity." One simplifying assumption that we will adopt (for ease of presentation) is to consider spacetimes that are conformally compactifiable, in the same sense that Minkowski space is itself conformally compactifiable. Going back to the discussion in Section 7.1.2, this compactified space is equivalent to the original one as far as causality questions are concerned. Thus, the compactification allows us to regard "infinity" as the boundary of the compactified space, made up of *future null infinity* $\mathcal{I}^+$ and *past null infinity* $\mathcal{I}^-$, which meet at *spacelike infinity* $i^0$ which is a single point for each end. For example, any complete future-pointing null geodesic "ends" at a point on $\mathcal{I}^+$, which is pronounced "scri plus." (Actually, having a meaningful notion of $\mathcal{I}^+$ is the main desirable property for the following discussion.) Using the conformal compactification, one obtains a nice "picture" of the global causal structure of the spacetime. When this picture is "sketched out" in two dimensions, it is referred to as a Penrose diagram.

The *black hole region* of a spacetime $(\mathcal{M}, \mathbf{g})$ is all of the points in $\mathcal{M}$ that can never reach $\mathcal{I}^+$ via future-pointing causal curves. The boundary of the black hole region is called the *black hole event horizon*. Similarly, the *white hole region* is all of the points that can never reach $\mathcal{I}^-$ via past pointing causal curves, and its boundary is the *white hole event horizon*. The points that lie outside both the black hole and white hole horizons are considered to be in the *domain of outer communication* (or *d.o.c.*). Thus, the black hole event horizon serves as a "point of no return" because if one passes from the domain of outer communication into the black hole region, then by definition it is not possible to return to the domain of outer communication.

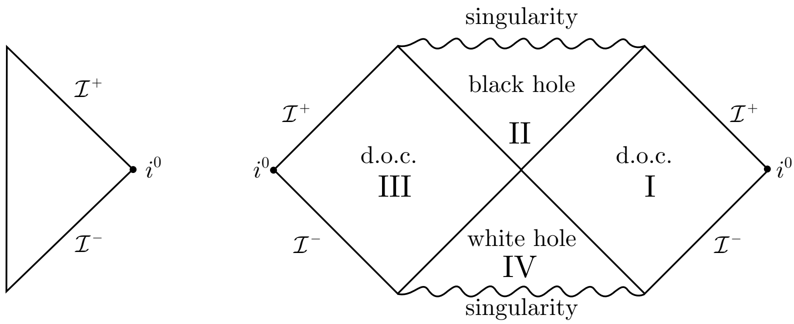


**Figure 7.3.** Penrose diagrams for the Minkowski spacetime (left) and the Schwarzschild spacetime (right).

We examine these concepts for the simple case of a Schwarzschild spacetime. The black hole region is where $U > 0$ and $V > 0$ in Kruskal-Szekeres coordinates, while the black hole event horizon is its boundary, made up of one piece $U = 0$, $V \geq 0$ bordering one end and another piece $U \geq 0$, $V = 0$ bordering the other. The white hole region is where $U < 0$ and $V < 0$. The domain of outer communication is the region where $UV < 0$, which has two components corresponding to the two infinite ends. Recall that each of these components separately corresponds to the $r > r_0$ region of (7.4). Although one does not approach a singularity as $r$ approaches $r_0$ from above, one does approach the event horizon (of the black hole and/or the white hole) as $r$ approaches $r_0$. Note that although our spacelike Schwarzschild constant $t$ slices (or more accurately, negative constant $U/V$ slices) do pass through the event horizon at $U = V = 0$, they do not actually intersect the interiors of the black hole or white hole regions.

**7.4.1. Geometry of null hypersurfaces.** In general, it is not clear whether an event horizon in $(\mathcal{M}, \mathbf{g})$ will be a smooth hypersurface of $\mathcal{M}$, but wherever it is smooth, it must be a null hypersurface. To see why, suppose $\mathcal{H}$ is a black hole event horizon which is a smooth hypersurface near $p \in \mathcal{H}$. If $\mathcal{H}$ is timelike at $p$, then there must exist past and future-pointing causal vectors at $p$ pointing toward both sides of $\mathcal{H}$. In particular, this means that both $J^+(p)$ and $J^-(p)$ intersect both the d.o.c. and the black hole region, which contradicts the definitions. If $\mathcal{H}$ is spacelike at $p$, then either the d.o.c. is on the past side of $\mathcal{H}$ while the black hole region is on the future side, which is a clear contradiction, or else we have the opposite. In the latter case $J^+(x)$ lies in the black hole region for all $x$ in a small enough neighborhood $U \subset \mathcal{H}$ of $p$. This means that for a point $q$ in the d.o.c. close enough to $p$, its entire causal future will have to pass through $U$ and consequently into the black hole region, which is a contradiction. Hence $\mathcal{H}$ is null at $p$. A similar argument works for white hole horizons.

For this reason and others, we will be interested in studying the geometry of null hypersurfaces. In general, it is important to consider null hypersurfaces of low regularity, but for our discussion here we will focus on the smooth case. Let $\mathcal{H}$ be a null hypersurface, and let $\ell$ be a nonvanishing future-pointing null normal vector field. This is equivalent to asking for $\ell$ to be a nonvanishing future-pointing null *tangent* vector field. Since there is no natural normalization for null vectors, $\ell$ is only determined up to multiplication by a positive function. That is, only its direction is naturally defined. Consequently, the integral curves of $\ell$ are defined independently of choice of $\ell$, and in fact, after reparameterization, these integral curves become null geodesics. To see why, let $p \in \mathcal{H}$ and $X \in T_p\mathcal{H}$, and then extend $X$ in the direction of $\ell$ so that $X$ remains tangent to $\mathcal{H}$ and $[\ell, X] = 0$. Then

$$\begin{aligned}\langle \nabla_\ell \ell, X\rangle &= \ell\langle \ell, X\rangle - \langle \ell, \nabla_\ell X\rangle \\ &= -\langle \ell, \nabla_X \ell\rangle \\ &= \frac{1}{2} X\langle \ell, \ell\rangle \\ &= 0.\end{aligned}$$

Thus $\nabla_\ell \ell$ is normal to $T_p\mathcal{H}$ and therefore points in the same direction as $\ell$. This means that the integral curves of $\ell$ are geodesics after reparameterization (or equivalently, after multiplying $\ell$ by an appropriate positive function). These null geodesics are called *(null) generators* of $\mathcal{H}$.

Because of the degeneracy of the induced metric on $\mathcal{H}$, we have to be careful about how to study the geometry of $\mathcal{H}$. In particular, defining the second fundamental form is not quite straightforward. To do it correctly, we must work modulo $\ell$. That is, we work with the quotient space $T_p\mathcal{H}/\langle \ell\rangle$ at each $p \in \mathcal{H}$, where $\langle \ell\rangle$ is the subspace spanned by $\ell$, which is independent of our choice of $\ell$. We will use bar notation to denote the quotient map from $T_p\mathcal{H}$ to $T_p\mathcal{H}/\langle \ell\rangle$. The degenerate induced metric on $T_p\mathcal{H}$ descends to an inner product on $T_p\mathcal{H}/\langle \ell\rangle$, which we will still abusively denote $\langle \cdot, \cdot\rangle$.

We define the *null second fundamental form* $A$ and *null shape operator* (or *null Weingarten map*) $S$ of $\mathcal{H}$ via

$$A(\bar{X}, \bar{Y}) := \langle S(\bar{X}), \bar{Y}\rangle := \langle \nabla_X \ell, Y\rangle$$

for any vector fields $X, Y$ tangent to $\mathcal{H}$, where $\nabla$ is the Levi-Civita connection of the ambient spacetime $(\mathcal{M}, \mathbf{g})$. Note that this definition is similar to equation (2.1). More precisely, $A$ is a symmetric bilinear form and $S$ is a symmetric operator on the space $T_p\mathcal{H}/\langle \ell\rangle$ at each $p \in \mathcal{H}$. The *null mean curvature* or *(null) expansion* of $\mathcal{H}$ is defined to be $\theta = \operatorname{tr} A = \operatorname{tr} S$, where the trace is taken over the space $T_p\mathcal{H}/\langle \ell\rangle$ at each $p \in \mathcal{H}$.

One can show that multiplying $\ell$ by a positive function has the effect of multiplying $A$, $S$, and $\theta$ by the same positive function. In particular, although the sizes of these quantities have no invariant geometric meaning, their *signs* do.

Next we would like to see how $A$, $S$, and $\theta$ evolve along a null generator $\gamma$. Given a vector field $X$ tangent to $\mathcal{H}$, one can see that $\nabla_{\gamma'}X$ must be tangent to $\mathcal{H}$. (Check this.) So we can define the covariant derivative of any $\bar{X} \in C^\infty(T\mathcal{H}/\langle\ell\rangle)$ along $\gamma$ by $\bar{X}' = \overline{\nabla_{\gamma'}X}$ for any $X \in C^\infty(T\mathcal{H})$ that projects to $\bar{X}$. This is well-defined in the sense that it is independent of the choice of $X$. With this definition, we can easily extend covariant differentiation along $\gamma$ to $A$ and $S$ in the standard way.

**Theorem 7.23** (Riccati equation for null generators)**.** *Let $\gamma$ be a null generator for a null hypersurface $\mathcal{H}$, and let $\ell = \gamma'$ be our choice of null normal along $\gamma$. Then along $\gamma$, the null shape operator satisfies*

$$\langle S'(\bar{X}), \bar{Y}\rangle = -\langle S^2(\bar{X})\bar{Y}\rangle - \mathbf{Riem}(X, \ell, Y, \ell),$$

*at any $p$ along $\gamma$, where $X, Y \in T_p\mathcal{H}$. The null expansion satisfies*

$$\theta' = -\frac{1}{n-1}\theta^2 - |\mathring{S}|^2 - \mathbf{Ric}(\ell, \ell),$$

*where $\mathring{S}$ is the trace-free part of $S$. The equation for $\theta'$ is usually referred to as the* Raychaudhuri equation *in the literature, and $|\mathring{S}|$ is called the* shear scalar*.*

**Proof.** First, we can extend $X, Y$ along $\gamma$ so that they remain tangent to $\mathcal{H}$ and $[\ell, X] = \nabla_\ell Y = 0$. Note that the curvature term becomes

$$\begin{aligned}\mathbf{Riem}(X, \ell, Y, \ell) &= \langle -\nabla_\ell\nabla_X\ell + \nabla_X\nabla_\ell\ell + \nabla_{[\ell,X]}\ell, Y\rangle \\ &= \langle -\nabla_\ell\nabla_\ell X, Y\rangle.\end{aligned}$$

Note that this is just the Jacobi equation for $X$, which should be expected when you observe that $\mathcal{H}$ is ruled by null geodesic generators, and that $X$ is simply varying through them. Thus

$$\begin{aligned}\langle S'(\bar{X}), \bar{Y}\rangle &= \nabla_\ell\langle S(\bar{X}), \bar{Y}\rangle - \langle S(\bar{X}'), \bar{Y}\rangle - \langle S(\bar{X}, \bar{Y}'\rangle \\ &= \nabla_\ell\langle\nabla_X\ell, Y\rangle - \langle S(\overline{\nabla_\ell X}), \bar{Y}\rangle \\ &= \nabla_\ell\langle\nabla_\ell X, Y\rangle - \langle S(\overline{\nabla_X\ell}), \bar{Y}\rangle \\ &= \langle\nabla_\ell\nabla_\ell X, Y\rangle - \langle S(S(\bar{X})), \bar{Y}\rangle \\ &= \mathbf{Riem}(X, \ell, Y, \ell) - \langle S^2(\bar{X}), \bar{Y}\rangle,\end{aligned}$$

where we used the Jacobi equation for $X$ in the last line.

To prove the Raychaudhuri equation, we simply take the trace of the Riccati equation for $S$ over the space $T_p\mathcal{H}/\langle\ell\rangle$ and use linear algebra. □

The formulas in Theorem 7.23 may look familiar from Riemannian geometry. Although we can use the same reasoning above to prove the Riemannian analog, it is interesting to note that it can also be seen as a special case.

**Corollary 7.24** (Riccati equation for parallel hypersurfaces)**.** *Let $\Sigma$ be a hypersurface of a Riemannian manifold $(M, g)$ with unit normal $\nu$. For each point $x \in \Sigma$, consider the geodesic $\exp_x t\nu$ emanating from $\Sigma$ with normal vector $\nu$. Let $\Sigma_t$ be the image obtained from $\Sigma$ by flowing along those geodesics for time $t$. (This $\Sigma_t$ is often called a parallel hypersurface.) Then along each geodesic, as long as $\Sigma_t$ is smooth near the geodesic, the shape operator $S_t$ of $\Sigma_t$ evolves according to*

$$\langle S'(X), Y\rangle = -\langle S^2(X), Y\rangle - \mathrm{Riem}(X, \nu, Y, \nu)$$

*along $\gamma$, where $X, Y \in T_pM$. The mean curvature satisfies*

$$H' = -\frac{1}{n-1}H^2 - |\mathring{S}|^2 - \mathrm{Ric}(\nu, \nu),$$

*where $\mathring{S}$ is the trace-free part of $S$.*

**Proof.** Given the hypotheses, we can construct a situation that satisfies the hypotheses of Theorem 7.23 as follows. Let $\mathcal{M} = \mathbb{R} \times M$ equipped with the Lorentzian metric $\mathbf{g} = -dt^2 + g$, and isometrically embed $(M, g)$ in $(\mathcal{M}, \mathbf{g})$ as its zero slice $\{0\} \times M$. Let $\ell = \partial_t + \nu$ along $\Sigma$ in $\mathcal{M}$, and then consider the null geodesics emanating from $\Sigma$ with initial tangent vector $\ell$. These null geodesics will generate the null hypersurface $\mathcal{H} = \{(t, x) \,|\, x \in \Sigma_t\}$. We can now apply Theorem 7.23 to the null hypersurface $\mathcal{H}$. All that remains is to check that the null shape operator $S$ of $\mathcal{H}$ at $(t, x) \in \mathcal{H}$ is essentially the same object as the shape operator $S$ of $\Sigma_t$ at $x \in \Sigma_t$, and that $\mathbf{Riem}(X, \ell, Y, \ell) = \mathrm{Riem}(X, \nu, Y, \nu)$, where $X, Y$ are tangent vectors to $\Sigma_t$ that can also be thought of as tangent to $\{t\} \times \Sigma_t \subset \mathcal{H}$. □

**Exercise 7.25.** Check those last two details in the proof above.

The Raychaudhuri equation is tremendously important in general relativity because we often have control over $\mathbf{Ric}(\ell, \ell)$. We say that a spacetime satisfies the *null energy condition* (or *NEC*) if $\mathbf{Ric}(v, v) \geq 0$ for all null vectors $v$. Physical spacetimes typically satisfy this condition. In particular, note that it is much weaker than the dominant energy condition.

**Exercise 7.26.** Let $(\mathcal{M}, \mathbf{g})$ be a spacetime satisfying the null energy condition, and let $\mathcal{H}$ be a smooth null hypersurface $\mathcal{H}$. If $\theta < 0$ at some point $p \in \mathcal{H}$, show that the null generator through $p$ cannot be future geodesically complete. Hint: Use the Raychaudhuri equation to show that if the geodesic exists for all parameter times $t > 0$, then $\theta$ has to blow up.

In light of this exercise, negative $\theta$ can be used to imply future geodesic incompleteness, as we shall see below.

**7.4.2. Trapped surfaces and the Penrose incompleteness theorem.** Let $(\mathcal{M}^{n+1}, \mathbf{g})$ be a spacetime, and let $\Sigma^{n-1}$ be a spacelike submanifold of $\mathcal{M}$, meaning that $\mathbf{g}$ induces a Riemannian metric on $\Sigma$. We will refer to $\Sigma^{n-1}$ as a "surface" since it is two-dimensional when $n = 3$. When $\mathbf{g}$ is restricted to the normal space $N\Sigma$, its signature will have one $-1$ and one 1, and therefore at each $p \in \Sigma$, $N_p\Sigma$ can be spanned by two future-pointing null vectors, which we denote $\ell_+$ and $\ell_-$. If $N\Sigma$ is a trivial bundle over $\Sigma$, then $\ell_+$ and $\ell_-$ can be defined globally over $\Sigma$. Once again, keep in mind that since these vectors are null, there is no notion of "normalizing" this basis in the way we can for orthogonal bases. We consider the components of the second fundamental form and mean curvature of $\Sigma$ in $\mathcal{M}$ with respect to these two null directions. For any $v, w \in T_p\Sigma$, we define

$$\chi_\pm(v, w) = \mathbf{g}(\nabla_v \ell_\pm, w),$$

and we define

$$\theta^\pm = \operatorname{tr}_\Sigma \chi_\pm = \operatorname{div}_\Sigma \ell_\pm.$$

We call the $\chi_\pm$ the *null second fundamental forms* and $\theta^\pm$ the *null mean curvatures*, or *null expansion scalars*. Once again, keep in mind that since $\ell_\pm$ cannot be normalized, these quantities depend on the choice of $\ell_\pm$. Specifically, multiplying $\ell_\pm$ by a positive function has the effect of multiplying $\theta^\pm$ by the same function, so that only the signs of $\chi_\pm$ and $\theta^\pm$ have geometric significance.

We can easily relate these quantities to the $A$ and $\theta$ that were defined on null hypersurfaces above. If we define $\mathcal{H}_\pm$ to be the null hypersurfaces generated by the geodesics leaving $\Sigma$ with tangent vectors $\ell_\pm$, respectively, it is easy to see from the definitions that

$$\chi_\pm(X, Y) = A_{\mathcal{H}_\pm}(\bar{X}, \bar{Y})$$

for any $X, Y$ tangent to $\Sigma$, and that $\theta^\pm$ for $\Sigma$ is equal to the null mean curvature $\theta$ for $\mathcal{H}_\pm$. Hence, our abusive choice to use the same notation for both is reasonable.

Generally, we like to think of $\ell_+$ as being outgoing and $\ell_-$ as being ingoing.

**Exercise 7.27.** Let $(M^n, g, k)$ be an initial data set sitting inside a spacetime $(\mathcal{M}^{n+1}, \mathbf{g})$. Let $\Sigma^{n-1}$ be a surface in $M$. Let $\nu$ be a unit normal of $\Sigma$ in $M$, and let $\vec{n}$ be the future-pointing unit normal to $M$ in $\mathcal{M}$. If we define $\ell_\pm = \vec{n} \pm \nu$, show that

$$\theta^\pm = \operatorname{tr}_\Sigma k \pm H,$$

where $H$ is the mean curvature scalar of $\Sigma$ in $M$ with respect to the normal $\nu$. Consequently, we can use this formula to *define* $\theta^{\pm}$ for any surface $\Sigma$ in an initial data set with a distinguished normal, even without specifying a spacetime $(\mathcal{M}^{n+1}, \mathbf{g})$.

If $\Sigma$ is a compact boundary surface, then we take $\nu$ to be outward pointing by convention.

**Definition 7.28.** Given a spacelike surface $\Sigma^{n-1}$, in either a spacetime $(\mathcal{M}, \mathbf{g})$ or an initial data set $(M^n, g, k)$, we say that $\Sigma$ is

- *outer trapped* if $\theta^+ < 0$,
- *weakly outer trapped* if $\theta^+ \leq 0$,
- *outer untrapped* if $\theta^+ > 0$,
- *weakly outer untrapped* if $\theta^+ \geq 0$,
- *marginally outer trapped* if $\theta^+ = 0$.

We often refer to a marginally outer trapped surface as a *MOTS* for convenience. We have similar definitions with "inner" in place of "outer" if we replace $\theta^+$ by $\theta^-$ on the right. A surface is called *trapped* if it is both outer trapped and inner trapped.

These definitions make sense as long as we have a distinguished choice of $\ell^+$ or $\nu$, and $\Sigma$ need not have an actual "outside" or "inside."

To get a sense for what the sign of $\theta^+$ means, recall from Proposition 2.10 that $\theta^{\pm}$ represents how the area form on $\Sigma$ is changing when we vary $\Sigma$ in the $\ell_{\pm}$ direction (which is why it is called a "null expansion"). Meanwhile, $\ell_+$ and $\ell_-$ represent the two most "extreme" directions that a light ray can travel away from $\Sigma$. We can think of $\ell_+$ as shooting light outward from $\Sigma$ and $\ell_-$ as shooting light inward. A trapped surface is one for which the following is true. If you flow $\Sigma$ in the direction of any smooth family of light rays emanating from $\Sigma$, this always has the effect of *decreasing* area. Intuitively, this is to be expected if you shoot light rays inward (corresponding to $\theta^-$), but it is not so expected when you shoot light rays outward (corresponding to $\theta^+$). For example, it is not hard to see that any large coordinate sphere in an asymptotically flat initial data set has $\theta^+ > 0$ and $\theta^- < 0$. The physical intuition is that only a "strong gravitational field" can cause light to be "trapped" in the sense that shooting light in any direction is area decreasing.

The famous Penrose incompleteness theorem [**Pen65**] states that in a spacetime satisfying the NEC, trapped surfaces force the formation of singularities in the spacetime (assuming there is a noncompact Cauchy hypersurface). We will instead state and prove a version appearing in [**Gal14**] better suited to our interests.

**Theorem 7.29** (Penrose incompleteness theorem for outer trapped surfaces)**.** *Let* $(\mathcal{M}^{n+1}, \mathbf{g})$ *be a spacetime containing a noncompact Cauchy hypersurface* $M^n$ *and satisfying the null energy condition. Suppose there exists a precompact open subset* $\Omega \subset M$ *such that* $\Sigma^{n-1} = \partial\Omega^n$ *is outer trapped. Then* $(\mathcal{M}, \mathbf{g})$ *is future null geodesically incomplete.*

According to Theorem 7.14, we saw that there always exists a maximal globally hyperbolic vacuum development of a vacuum initial data set, but it said nothing about future completeness. The Penrose incompleteness theorem is significant because it says that if the initial data contains an outer trapped surface, then its evolution *cannot* be future complete. The Penrose incompleteness theorem is often called a "singularity" theorem, though that might be slightly misleading since all it says is that there is a future null geodesic that cannot be extended forward with infinite parameter time. It does not mean that the curvature must blow up there, since there could be a smooth spacetime extension in which $M$ ceases to be a Cauchy hypersurface.

**Sketch of the proof.** Let $(\mathcal{M}^{n+1}, \mathbf{g})$ be a spacetime containing a Cauchy hypersurface $M^n$ and let $\Omega$ be an open subset of $M$ with $\Sigma^{n-1} := \partial\Omega^n$. Define $\partial^{\text{out}} J^+(\Sigma) := \partial J^+(\overline{\Omega}) \smallsetminus \Omega$. The notation on the left is chosen to abbreviate the right-hand side in such a way that reminds us of what it is, intuitively: it is supposed to represent the "outer boundary" of the causal future of $\Sigma$.

**Claim.** *Each* $q \in \partial^{\text{out}} J^+(\Sigma)$ *lies on a null geodesic leaving* $\Sigma$ *with tangent vector* $\ell_+ = \vec{n} + \nu$*, where* $\vec{n}$ *is the future unit timelike normal to* $M$ *and* $\nu$ *is the outward normal of* $\Sigma$ *in* $M$*. Moreover, this geodesic arrives at* $q$ *before it passes any conjugate points.*

The proof of Claim 1 only uses the global hyperbolicity and does not use any assumptions about compactness, trapping, or geodesic completeness. It should perhaps be thought of as a background lemma from causality theory. One can show that global hyperbolicity implies that $J^+(\overline{\Omega})$ is closed, that is, $\partial J^+(\overline{\Omega}) \subset J^+(\overline{\Omega})$. (Again, this is not so obvious from the way we defined global hyperbolicity.) Since being timelike is an open condition, one can see that $I^+(\overline{\Omega})$ is an open subset of $J^+(\overline{\Omega})$, and thus $\partial J^+(\overline{\Omega}) \subset J^+(\overline{\Omega}) \smallsetminus I^+(\overline{\Omega})$. By Proposition 7.4, it then follows that for any $q \in \partial^{\text{out}} J^+(\Sigma)$ there exists a null geodesic $\gamma$ starting at some $p \in \overline{\Omega}$ and ending at $q$.

We now argue that the starting point $p$ lies in $\Sigma$: if it started in $\Omega$, then we could move the starting point slightly to construct a causal curve that is timelike near its new starting point in $\Omega$. (Imagine the picture in Minkowski space to see why this is clear.) By Lemma 7.3, $q$ would lie in $I^+(\Omega)$, a contradiction.

Next we claim that $\gamma$ can be chosen so that $\gamma'(0) = \ell_+ = \vec{n}+\nu$. Without loss of generality, we can choose $\gamma$ so that $\gamma'(0) = \vec{n}+v$ for some unit vector $v \in T_pM$. If $v$ points into $\Omega$, then just as in the paragraph above, we can move the starting point in the $v$ direction to find a causal curve that is timelike near the new starting point in $\Omega$, which is a contradiction. If $\langle v, \nu\rangle \geq 0$, but $v \neq \nu$, we can instead move the starting point of $\gamma$ in the direction of $v - \langle v, \nu\rangle\nu \in T_p\Sigma$ in such a way that the new starting point remains in $\Sigma$, and such that we can construct a causal curve that is timelike near its new starting point. (Again, this should be intuitively clear in the local picture where $\Omega$ is a half plane in the $t = 0$ slice of Minkowski space.) Once again, Lemma 7.3 then implies that $q \in I^+(\overline{\Omega})$, a contradiction. Thus $v = \nu$.

Finally, we argue that $\gamma$ is free of conjugate points. If it did have conjugate points, we could use the corresponding Jacobi field to deform $\gamma$ to again obtain a broken null geodesic, which can then be deformed to a smooth causal curve that is timelike somewhere, and again use Lemma 7.3 to obtain a contradiction. This completes the proof of our Claim 1.

We now assume the full hypotheses of Theorem 7.29, and, in addition, we assume that $(\mathcal{M}, \mathbf{g})$ is actually future null geodesically complete and work toward a contradiction.

**Claim.** *$\partial J^+(\overline{\Omega})$ is compact.*

This part of the proof uses the Raychaudhuri equation, which lies at the heart of the Penrose incompleteness theorem. By the assumption of future null geodesic completeness, each null geodesic starting in $\Sigma$ with null vector $\ell_+$ can be extended for infinite parameter time. Let $\mathcal{H}_+$ be the union of all of these null geodesics. This $\mathcal{H}_+$ must be a smooth (possibly immersed) null hypersurface away from the conjugate points. Since $\Sigma$ is compact and trapped, there exists a constant $c > 0$ such that $\theta^+ < -c < 0$. By the Raychaudhuri equation, each generator must blow up before reaching parameter time $T = (n-1)/c$. (See Exercise 7.26.) In other words, it must reach a point of nonsmoothness of $\mathcal{H}_+$, which translates to saying that every generator must reach a conjugate point of the generator before parameter time $T$. By Claim 1, $\partial^{\text{out}}J^+(\Sigma)$ lies in the part of $\mathcal{H}_+$ with parameter times lying in the closed interval $[0, T]$. Since this latter space is clearly compact and $\partial^{\text{out}}J^+(\Sigma)$ is a closed subset of it, we conclude that $\partial^{\text{out}}J^+(\Sigma)$ is compact. Since $\overline{\Omega}$ is assumed to be compact, it follows that $\partial J^+(\overline{\Omega})$ is compact. See Figure 7.4 for an "illustration" of this (impossible) situation.

To complete the argument, we construct a map from $\partial J^+(\overline{\Omega})$ to $M$ as follows. By time-orientability, there exists a global timelike vector field on $\mathcal{M}$ generating timelike integral curves. For each point in $\partial J^+(\overline{\Omega})$, we

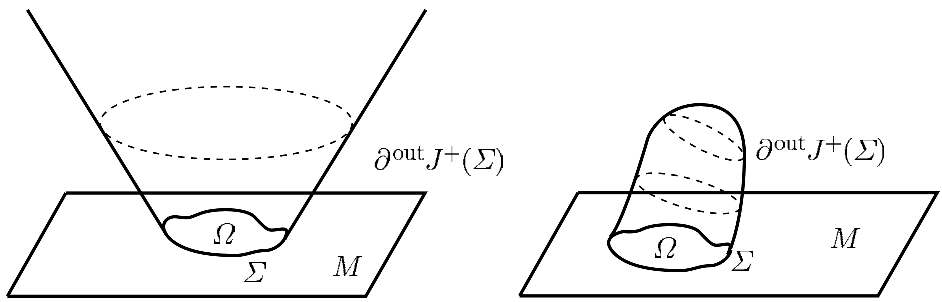


**Figure 7.4.** On the left is an example of $\partial^{\text{out}}J^+(\Sigma)$ in Minkowski space. On the right, if $\Sigma$ is trapped, then the NEC together with the Raychaudhuri equation and future null completeness forces $\partial^{\text{out}}J^+(\Sigma)$ to "close up," but this is intuitively impossible because the timelike future of $\Omega$ has nowhere to go.

follow one of these timelike integral curves to reach a point in $M$. By the definition of a Cauchy hypersurface, this map must be well-defined. Arguing as above (together with the fact that $\overline{\Omega}$ lies on a Cauchy hypersurface), one can see that no point in $\partial J^+(\overline{\Omega})$ can be in the chronological future of another point of $\partial J^+(\overline{\Omega})$, and thus the map from $\partial J^+(\overline{\Omega})$ to $M$ is actually injective and continuous. Finally, one can show that $\partial J^+(\overline{\Omega})$ is a Lipschitz hypersurface of $\mathcal{M}$ without boundary as a manifold. (The fact that it is a Lipschitz hypersurface is not obvious, but it is intuitively clear that $\partial J^+(\overline{\Omega})$ should have no manifold boundary since it is itself a boundary.) Putting it all together, this gives us an injective continuous map from a compact topological manifold to a noncompact manifold, which can be shown to be impossible for topological reasons. ☐

The hypotheses of Theorem 7.29 can almost be relaxed to *weakly* outer trapped surfaces. That is, it was shown in [**EGP13**] that, "generically," if $\Sigma$ is a MOTS, we obtain the same conclusion (where "generic" here means that certain curvatures do not vanish identically along the null generators).

Theorem 7.29 can be used to show that topology can force incompleteness.

**Theorem 7.30** (Gannon [**Gan75**], C. W. Lee [**Lee76**]). *Let $(\mathcal{M}^{n+1}, \mathbf{g})$ be a spacetime containing an asymptotically flat Cauchy hypersurface $M$ and satisfying the null energy condition. If $M$ is not simply connected, then $(\mathcal{M}, \mathbf{g})$ is future null geodesically incomplete.*

**Proof.** Assume that $(\mathcal{M}, \mathbf{g})$ is future null geodesically complete, and we work toward a contradiction. By Proposition 7.7, $\mathcal{M}$ is homeomorphic to $\mathbb{R} \times M$. We will use the nonsimply connected hypothesis in a manner similar to the way it was used in the proof of Theorem 4.11. Consider the universal cover $\tilde{M}$ of $M$, which sits inside the universal cover $\tilde{\mathcal{M}} \cong \mathbb{R} \times \tilde{M}$ of $\mathcal{M}$, which must also be future null geodesically complete. If $M$ is not simply connected, then $\tilde{M}$ has at least two ends. If we take $\Omega$ to be one of the infinite ends beyond a sphere of large enough radius, then it is easy to see that $\partial\Omega$ is outer trapped with respect to the unit normal pointing out of $\Omega$. We now run the same argument that was used in Theorem 7.29 with the difference being that $\overline{\Omega}$ is no longer compact, and therefore the space $\partial J^+(\overline{\Omega})$ is not compact. However, we still obtain an injective continuous map from $\partial J^+(\overline{\Omega})$ to $\tilde{M}$, where the former space is a topological manifold with one noncompact end, while the latter is a topological manifold with at least two noncompact ends. This is still impossible for topological reasons.

□

**7.4.3. Discussion.** At first glance, the Penrose incompleteness theorem might look like bad news. It means that if we start with any initial data set containing an outer trapped surface, then its maximal globally hyperbolic development under the Einstein equations (using any matter model satisfying the NEC, including vacuum) *must* come to some sort of abrupt end. (It is not too difficult to construct such initial data sets. See [**SY83**] for a theorem explaining how concentrating a lot of matter in a small place can force the existence of outer trapped surfaces.) In some sense, this suggests a failure of the Einstein equations as a physical theory. As a response to this problem, Penrose proposed the *weak cosmic censorship hypothesis* [**Pen02**]. Roughly, the weak cosmic censorship hypothesis is the conjecture that although singularities may form, they always stay inside the black hole region. Since the physically observable world exists in the domain of outer communication, weak cosmic censorship provides an elegant way for the theory to save face by only failing in a way that will never affect us. A more technical (but still vague) way to state the weak cosmic censorship conjecture is that "generic" initial data gives rise to a maximal globally hyperbolic development under the Einstein equations that admits a complete $\mathcal{I}^+$. Of course, the development depends on the matter model, but even the vacuum case is an important open problem. At the beginning of this section on black holes, we essentially started with the assumption of a spacetime with a complete $\mathcal{I}^+$ before we even defined the concept of a "black hole" (implicit in our vague assumption that our spacetime was asymptotic to Minkowski in some sense). This is one reason why we tried to state weak cosmic censorship without making mention of black holes, and it also illustrates the issue alluded to earlier about the trickiness involved in definitions.

One bit of evidence in favor of weak cosmic censorship is that while outer trapped surfaces force future null incompleteness, they *also* indicate the existence of black holes, which makes one hope that they go hand in hand.

More precisely, given an asymptotically flat spacetime satisfying global hyperbolicity and the NEC, an outer trapped surface (with respect to a particular end) cannot intersect that end's domain of outer communication. We present the basic argument, due to S. Hawking [**HE73**]. Suppose that some part of a trapped surface $\Sigma = \partial\Omega$ lies outside the black hole region. This means that $J^+(\overline{\Omega})$ reaches all the way out to $\mathcal{I}^+$. Let $q$ be a point in the boundary of the intersection of $\mathcal{I}^+$ and $J^+(\overline{\Omega})$. Following similar reasoning as in Theorem 7.29, $q$ must lie at the end of a null geodesic leaving $\Sigma$ with tangent vector $\ell_+$, and moreover it should be free of conjugate points for all parameter times. But if $\theta^+ < 0$ at $\Sigma$, then the Raychadhuri equation (together with the NEC) forces the existence of a conjugate point in finite parameter time (Exercise 7.26), which is a contradiction. In fact, this reasoning can be extended to *weakly* outer trapped surfaces, because if the null generator starts with $\theta^+ = 0$ and has no conjugate points, then the Raychaudhuri equation implies that it must have $\theta^+ = 0$ for all parameter times. But since this null generator eventually reaches $\mathcal{I}^+$, it passes through a spacetime region that is close to Minkowski space, and there one can prove that such an "outgoing" null hypersurface with $\theta^+ = 0$ is impossible. However, in the argument above, note that we implicitly assumed that $\mathcal{I}^+$ was complete in the construction of the point $q$, so in some sense the argument rests upon the weak cosmic censorship hypothesis. Despite this, we tend to think of outer trapped surfaces as indicating the presence of a black hole.

There is another famous conjecture of Penrose called the *strong cosmic censorship*, which is logically independent from the weak cosmic censorship hypothesis. We will not discuss it here, except to note recent developments by M. Dafermos and Jonathan Luk [**DL17**], whose work implies that the "$C^0$-inextendibility" formulation of the conjecture requires revision.

For many years, it was an open question whether a vacuum initial data set free of outer trapped surfaces could develop outer trapped surfaces in its vacuum development. This represents a black hole forming from pure gravity, rather than from concentration of matter. Although it was generally believed to be possible, the first examples were constructed in a celebrated work of D. Christodoulou [**Chr09**].

Let us go back to considering a black event horizon $\mathcal{H}$. Assuming asymptotic flatness, global hyperbolicity, and the NEC, not only must weakly outer trapped surfaces lie on the inside side of $\mathcal{H}$, but a similar argument shows that $\mathcal{H}$ must have $\theta \geq 0$ wherever it is smooth [**HE73**]. (Again, we note that event horizons need not be smooth, so it is important to have proofs that work more generally. See [**CDGH01**].) We will summarize the argument. Suppose there is a point $p \in \mathcal{H}$ where $\mathcal{H}$ is smooth and $\theta < 0$. Consider a smooth spacelike surface $\Sigma$ in $\mathcal{H}$ passing through $p$. Then $\theta^+ < 0$, since the outgoing null normal $\ell_+$ of $\Sigma$ must be the null normal of $\mathcal{H}$. We can slightly push $\Sigma$ outward toward the d.o.c. near $p$ to obtain another surface $\Sigma'$ such that a small part of $\Sigma'$ leaks into the d.o.c. with $\theta^+ < 0$, while the rest of $\Sigma'$ is identically equal to $\Sigma$. As above, we can argue that there must exist a null geodesic from $\Sigma'$ out to $\mathcal{I}^+$ which is free of conjugate points. By definition of the d.o.c., this geodesic must originate from a point in the intersection of $\Sigma'$ with the d.o.c., where $\theta^+ < 0$. (And the geodesic must be outgoing, since an ingoing geodesic will surely enter the black hole region.) But this contradicts the Raychaudhuri equation via Exercise 7.26. This fact that the null expansion of a black hole horizon is nonnegative is usually called the *Hawking area theorem* or the *second law of black hole mechanics*. It is called the area theorem because if $\Sigma$ is a spacelike surface in $\mathcal{H}$, then $\theta^+_\Sigma \geq 0$ represents the rate of change of the area element of $\Sigma$ as it flows in the direction $\ell_+$. In particular, if $\Sigma_1$ and $\Sigma_2$ are smooth closed surfaces on a smooth part of $\mathcal{H}$ with $\Sigma_2$ in the causal future of $\Sigma_1$, then $|\Sigma_1| \leq |\Sigma_2|$.

A corollary of the Hawking area theorem is that under the same assumptions as above we have the following. If the spacetime is *stationary*, then a spacelike surface $\Sigma$ in $\mathcal{H}$ has $\theta_+ = 0$, or, in other words, it is a marginally outer trapped surface (or MOTS). Recall that stationarity means that there is a Killing vector field and hence a family of spacetime isometries. These isometries must preserve the event horizon and therefore they send $\Sigma$ to another cross-section $\Sigma'$ of $\mathcal{H}$ in its future, and $|\Sigma| = |\Sigma'|$. Because of the Hawking area theorem, this is only possible if $\theta^+ = 0$ along $\Sigma$.

## 7.5. Marginally outer trapped surfaces

Although the event horizon is an important concept, it can be difficult to study directly because it is determined by the global causal structure of the entire spacetime. The preceding discussion motivates the study of MOTS as a way of studying black holes using local geometry. Specifically, it allows us to use an initial data perspective.

**Definition 7.31.** Let $(M, g, k)$ be a complete initial data set with a distinguished noncompact end, let $\Sigma$ be a smooth enclosing boundary, and take $\nu$ to be the outward-pointing normal of $\Sigma$. We say that $\Sigma$ is an *outermost* MOTS if it cannot be enclosed by any other weakly outer trapped surfaces. An outermost MOTS will also be referred to as an *apparent horizon* for that end.

Be aware that, like many terms arising from physics, the phrase "apparent horizon" does not always have a consistent technical definition in the literature. Based on our earlier discussion, if $(M, g, k)$ lies in an asymptotically flat, globally hyperbolic spacetime satisfying the NEC, an apparent horizon will always lie on the inside of the black hole event horizon (including the horizon itself), and if the spacetime is stationary, the apparent horizon will lie on the black hole event horizon. It is the closest we can come to locating where the event horizon must intersect our initial data set. In particular, the concept has applications to numerical relativity.

**7.5.1. Stability of MOTS.** Observe that for a time-symmetric initial data set, a MOTS is just a minimal hypersurface. From a Riemannian geometry perspective, the MOTS equation $\theta^+ = 0$ can be thought of as a generalization of the minimal hypersurface equation $H = 0$. Although it shares many similarities with the minimal hypersurface equation, the most significant difference is that it does not arise from a variational principle. While volume is an essential tool for studying minimal hypersurfaces, there is no analogous quantity that can be used to study MOTS. Recall from Chapter 2 that the stability inequality (2.16) was a powerful tool for the study of scalar curvature. Even without a variational principle, we can still generalize the concept of stability from minimal hypersurfaces to the setting of MOTS.

Given a vector field $X$ defined along a hypersurface $\Sigma$ with unit normal $\nu$, we can decompose $X$ into its normal and tangential components

$$X = \varphi\nu + \hat{X}.$$

Given $X$, we adopt this convention for the notation $\varphi$ and $\hat{X}$.

We compute the linearization of the outward null expansion $\theta^+$.

**Proposition 7.32.** *Let $\Sigma^{n-1}$ be a hypersurface with unit normal $\nu$ in an initial data set $(M^n, g, k)$. The linearization of $\theta^+$ on $\Sigma$ in the direction of the vector field $X$ is given by*

$$D\theta^+|_\Sigma(X) = -\Delta_\Sigma\varphi + 2\langle W_\Sigma, \nabla\varphi\rangle + (\operatorname{div}_\Sigma W_\Sigma - |W_\Sigma|^2 + Q_\Sigma)\varphi + \frac{1}{2}\theta^+[\theta^- + 2k(\nu,\nu)]\varphi + \nabla_{\hat{X}}\theta^+_\Sigma,$$

*where*

$$Q_\Sigma := \frac{1}{2}R_\Sigma - \mu - \langle J, \nu\rangle - \frac{1}{2}|k_\Sigma + A_\Sigma|^2.$$

*Here $k_\Sigma$ denotes the restriction of $k$ to vectors tangent to $\Sigma$, and $W_\Sigma$ is the tangential vector field on $\Sigma$ that is dual to the 1-form $k(\nu,\cdot)$ along $\Sigma$.*

Note that the dominant energy condition implies that $Q_\Sigma \leq \frac{1}{2}R_\Sigma$.

**Proof.** As seen in Section 2.2, it is clear what the contribution from the tangential component $\hat{X}$ must be, so it suffices to consider the case of a normal variation $X = \varphi\nu$. Recall from (2.15) that we already know that the variation of $H$ is

$$DH|_\Sigma(\varphi\nu) = -\Delta_\Sigma\varphi + \frac{1}{2}(R_\Sigma - R_M - |A|^2 - H^2)\varphi.$$

Therefore the only thing we have to compute is the variation of $\operatorname{tr}_\Sigma k = \operatorname{tr}_g k - k(\nu,\nu)$. Since $g$ and $k$ are defined on the ambient space, the main quantity we need to understand is $\frac{\partial}{\partial t}\Phi_t^*\nu_t\big|_{t=0}$, where $\Phi_t$ is the family of diffeomorphisms generated by $X = \varphi\nu$ and $\nu_t$ is the unit normal of $\Sigma_t = \Phi_t(\Sigma)$. As is customary, we will use the abbreviated notation $\frac{\partial}{\partial t}\nu := \frac{\partial}{\partial t}\Phi_t^*\nu_t$.

**Exercise 7.33.** Show that for any hypersurface $\Sigma$ deformed in the $X = \varphi\nu$ direction, $\frac{\partial}{\partial t}\nu = -\nabla\varphi$.

So we can compute

$$\begin{aligned}\frac{\partial}{\partial t}(\operatorname{tr}_\Sigma k) &= \frac{\partial}{\partial t}(\operatorname{tr}_g k - k(\nu,\nu))\\ &= \frac{\partial}{\partial t}(\operatorname{tr}_g k) - \left(\frac{\partial}{\partial t}k\right)(\nu,\nu) - 2k\left(\frac{\partial}{\partial t}\nu,\nu\right)\\ &= \varphi\nabla_\nu(\operatorname{tr}_g k) - \varphi(\nabla_\nu k)(\nu,\nu) + 2k(\nabla\varphi,\nu)\\ &= [\nabla_\nu(\operatorname{tr}_g k) - (\nabla_\nu k)(\nu,\nu)]\varphi + 2\langle W,\nabla\varphi\rangle,\end{aligned}$$

where $W = W_\Sigma$ is as defined in the statement of the proposition. This is already enough to compute the variation of $\theta^+$, but we would like to put it in a nicer form. Specifically, we would like to see how the quantities $\mu$ and $J$ show up in the formula. Let $\tilde{W}$ be the vector field defined along $\Sigma$ that is dual to the $k(\nu,\cdot)$ so that $W$ is just the tangential part of $\tilde{W}$. We choose an orthonormal frame $e_1,\ldots,e_{n-1}$ for $\Sigma$ and compute (using Einstein notation)

$$\begin{aligned}(\nabla_\nu k)(\nu,\nu) &= (\operatorname{div}_g k - \operatorname{div}_\Sigma k)(\nu)\\ &= (\operatorname{div}_g k)(\nu) - (\nabla_{e_i}k)(\nu,e_i)\\ &= (\operatorname{div}_g k)(\nu) - [\nabla_{e_i}k(\nu)](e_i) + k(\nabla_{e_i}\nu,e_i)\\ &= (\operatorname{div}_g k)(\nu) - \operatorname{div}_\Sigma\tilde{W} + k(A_{ij}e_j,e_i)\\ &= (\operatorname{div}_g k)(\nu) - \operatorname{div}_\Sigma W + \langle \mathbf{H},\tilde{W}^\perp\rangle + A_{ij}k_{ij}\\ &= (\operatorname{div}_g k)(\nu) - \operatorname{div}_\Sigma W - Hk(\nu,\nu) + \langle A,k_\Sigma\rangle,\end{aligned}$$

where we used (2.6) to simplify $\operatorname{div}_\Sigma\tilde{W}$. Putting the three previous computations together, we have

$$\begin{aligned}D\theta^+|_\Sigma(\varphi\nu) = -\Delta_\Sigma\varphi + 2\langle W,\nabla\varphi\rangle + \frac{1}{2}[R_\Sigma - R_M - |A|^2 - H^2\\ + 2\nabla_\nu(\operatorname{tr}_g k) - 2(\operatorname{div}_g k)(\nu) + 2\operatorname{div}_\Sigma W\\ + 2Hk(\nu,\nu) - 2\langle A,k_\Sigma\rangle]\varphi.\end{aligned}$$

The rest of the computation is tedious but straightforward. □

**Exercise 7.34.** Complete the proof of Proposition 7.32. Also, what should the formula for the linearization of $\theta^-$ be?

**Definition 7.35.** Let $\Sigma$ be a hypersurface with unit normal $\nu$ in an initial data set $(M, g, k)$. By analogy with (2.15), we define the *(MOTS) stability operator* $\mathrm{L}_\Sigma$ on $\Sigma$ to be

$$\mathrm{L}_\Sigma \varphi := D\theta^+|_\Sigma(\varphi\nu) = \mathrm{L}^0_\Sigma \varphi + \frac{1}{2}\theta^+[\theta^- + 2k(\nu,\nu)]\varphi,$$

where

$$\mathrm{L}^0_\Sigma \varphi := -\Delta_\Sigma \varphi + 2\langle W_\Sigma, \nabla u\rangle + (\operatorname{div}_\Sigma W_\Sigma - |W_\Sigma|^2 + Q_\Sigma)\varphi$$

for any smooth function $\varphi$ on $\Sigma$, where $W_\Sigma$ and $Q_\Sigma$ are as defined in Proposition 7.32.

Since the MOTS stability operator generalizes the stability operator we already defined in Definition 2.16 in the time-symmetric case, we use the same notation. Again, the upright letter in $\mathrm{L}_\Sigma$ helps us distinguish the stability operator from the conformal Laplacian. Of course, if $\Sigma$ is a MOTS, then $\mathrm{L}_\Sigma$ and $\mathrm{L}^0_\Sigma$ are the same thing. Distinguishing these operators for non-MOTS hypersurfaces will be convenient later on.

Recall from Definition 2.16 that in the time-symmetric case, a compact minimal hypersurface (possibly with boundary) was defined to be *stable* if and only if the operator $\mathrm{L}_\Sigma$ is a nonnegative operator on smooth functions vanishing at the boundary. Since $\mathrm{L}_\Sigma$ is not self-adjoint in general, this definition of stability is not appropriate, but L. Andersson, M. Mars, and W. Simon proposed a concept of MOTS stability using the following observation [**AMS08**].

**Proposition 7.36.** *Let $\Sigma$ be a compact hypersurface (possibly with boundary) in an initial data set $(M, g, k)$. There exists a (Dirichlet) eigenvalue of the MOTS stability operator $\mathrm{L}_\Sigma$ with minimal real part, which is called the* principal (Dirichlet) eigenvalue*, denoted $\lambda_1(\mathrm{L}_\Sigma)$. Furthermore, this eigenvalue is real, and if $\Sigma$ is connected, the corresponding eigenspace is a one-dimensional space generated by a smooth* principal eigenfunction *that is positive on the interior of $\Sigma$. (The same is also true for $\mathrm{L}^0_\Sigma$.)*

*If $\Sigma$ is a MOTS and $\lambda_1(\mathrm{L}_\Sigma) \geq 0$, we say that $\Sigma$ is a* stable *MOTS.*

This proposition is a direct consequence of Theorem A.10, which in turn relies on the Krein-Rutman Theorem [**Wik**, Krein-Rutman_theorem].

We can "symmetrize" the stability condition to obtain the following result.

**Proposition 7.37** (Galloway-Schoen [**GS06**])**.** *Let $\Sigma$ be a compact hypersurface (possibly with boundary) in an initial data set $(M, g, k)$. Then*

$$\lambda_1(\mathrm{L}^0_\Sigma) \leq \lambda_1(-\Delta_\Sigma + Q_\Sigma),$$

*where the right side denotes the principal (Dirichlet) eigenvalue of the self-adjoint operator* $-\Delta_\Sigma + Q_\Sigma$*, where* $Q_\Sigma$ *was defined in Proposition* 7.32*. In particular, if* $\Sigma$ *is a stable MOTS, then for any smooth function* $u$ *compactly supported in the interior of* $\Sigma$*, we have*

$$\int_\Sigma |\nabla u|^2 + Q_\Sigma u^2 \geq 0. \tag{7.5}$$

**Proof.** We drop the $\Sigma$ subscripts in the following. For any function $\varphi$ that is positive in the interior of $\Sigma$,

$$\begin{aligned}
\mathrm{L}^0_\Sigma \varphi &= -\Delta\varphi + 2\langle W, \varphi^{-1}\nabla\varphi\rangle\varphi + (\operatorname{div} W - |W|^2 + Q)\varphi \\
&= -\Delta\varphi + (|\varphi^{-1}\nabla\varphi|^2 + |W|^2 - |W - \varphi^{-1}\nabla\varphi|^2)\varphi \\
&\quad + (\operatorname{div} W - |W|^2 + Q)\varphi \\
&= -\Delta\varphi + |\nabla\log\varphi|^2\varphi - |W - \nabla\log\varphi|^2\varphi + (\operatorname{div} W + Q)\varphi \\
&= -(\Delta\log\varphi)\varphi - |W - \nabla\log\varphi|^2\varphi + (\operatorname{div} W + Q)\varphi \\
&= [\operatorname{div}(W - \nabla\log\varphi)]\varphi - |W - \nabla\log\varphi|^2\varphi + Q\varphi.
\end{aligned}$$

Now let $u$ be any smooth function compactly supported in the interior of $\Sigma$. We multiply the above equation by $u^2\varphi^{-1}$ to obtain

$$\begin{aligned}
u^2\varphi^{-1}\mathrm{L}^0_\Sigma\varphi &= [\operatorname{div}(W - \nabla\log\varphi)]u^2 - |W - \nabla\log\varphi|^2u^2 + Qu^2 \\
&= \operatorname{div}(u^2(W - \nabla\log\varphi)) - \langle W - \nabla\log\varphi, 2u\nabla u\rangle \\
&\quad - |W - \nabla\log\varphi|^2u^2 + Qu^2 \\
&= \operatorname{div}(u^2(W - \nabla\log\varphi)) \\
&\quad + |(W - \nabla\log\varphi)u|^2 + |\nabla u|^2 - |(W - \nabla\log\varphi)u + \nabla u|^2 \\
&\quad - |W - \nabla\log\varphi|^2u^2 + Qu^2 \\
&= \operatorname{div}(u^2(W - \nabla\log\varphi)) + |\nabla u|^2 + Qu^2 \\
&\quad - |(W - \nabla\log\varphi)u + \nabla u|^2.
\end{aligned}$$

Integrating, we obtain

$$\int_\Sigma u^2\varphi^{-1}\mathrm{L}^0_\Sigma\varphi \leq \int_\Sigma |\nabla u|^2 + Qu^2. \tag{7.6}$$

By Proposition 7.36, we can choose $\varphi$ to be the principal eigenfunction of $\mathrm{L}^0_\Sigma$, so that the left side of the inequality becomes $\lambda_1(\mathrm{L}^0_\Sigma)\int_\Sigma u^2$. The result then follows from the Rayleigh quotient characterization of $\lambda_1(-\Delta_\Sigma + Q)$, as explained in the proof of Theorem A.10. $\square$

Moreover, using reasoning similar to that of Exercise 2.26, observe that if the DEC holds on $(M, g, k)$, then for any hypersurface $\Sigma$, we have

$$\lambda_1 (-\Delta_\Sigma + Q_\Sigma) \leq \frac{1}{2}\lambda_1(L_h), \tag{7.7}$$

where $L_h$ denotes the conformal Laplacian of the induced metric $h = g|_\Sigma$. Moreover, if the strict DEC holds, then the inequality is strict.

**7.5.2. Apparent horizons in initial data sets.** Here we discuss some theorems which are analogous to those presented in Section 4.1.1. One can prove that a version of the strong comparison principle for mean curvature (Corollary 4.2) also holds for $\theta^+$, and the proof is fairly similar. See [**AM09**, Proposition 2.4, **AG05**, Proposition 3.1] for details, as well as [**Gal00**] for a related maximum principle for null hypersurfaces.

**Proposition 7.38** (Strong maximum principle for $\theta^+$)**.** *Suppose we have open sets $\Omega_1 \subset \Omega_2$ in an initial data set $(M, g, k)$ and smooth hypersurfaces $\Sigma_1$ and $\Sigma_2$ (possibly with boundary) lie on $\partial\Omega_1$ and $\partial\Omega_2$, respectively, with $\theta^+_{\Sigma_1} \leq 0$ and $\theta^+_{\Sigma_2} \geq 0$, where these are computed using the outward-pointing unit normal. If $\Sigma_1$ touches $\Sigma_2$ anywhere in their interiors or are tangent to each other at a common boundary point, then they must be identically equal in a neighborhood of that point.*

Consequently, a closed MOTS can never "penetrate" a foliation by weakly outer untrapped surfaces (meaning $\theta^+ \geq 0$).

Next we would like to establish an existence result for apparent horizons, generalizing Theorem 4.7, but we cannot produce a MOTS via a minimization procedure, which is the standard technique used for producing minimal hypersurfaces. However, we do have the following existence theorem due to M. Eichmair [**Eic09**].

**Theorem 7.39** (Existence theorem for MOTS)**.** *Let $n < 8$. Let $(M^n, g, k)$ be an initial data set. Suppose $\Omega$ is an open subset of $M$ whose compact boundary $\partial\Omega$ can be divided into two smooth hypersurfaces with boundary pieces $\partial_1\Omega$ and $\partial_2\Omega$ that meet along a smooth $(n-2)$-dimensional submanifold $\Gamma$.*

*Assume that $\partial_1\Omega$ is weakly outer untrapped (meaning $\theta^+ \geq 0$) with respect to the outward-pointing normal, and that $\partial_2\Omega$ is weakly outer trapped (meaning $\theta^+ \leq 0$) with respect to the* inward*-pointing normal. Then there exists a smooth $\lambda$-minimizing stable MOTS $\Sigma$ such that $\partial\Sigma = \Gamma$ and $\Sigma$ is homologous to $\partial_1\Omega$. Moreover, if $(M, g, k)$ is asymptotically flat, then $\lambda > 0$ depends only on the geometry of $(M, g, k)$.*

The stability part of the conclusion was established in [**EM16**].

Briefly, a smooth boundary $\partial E$ in $\Omega$ is $\lambda$-*minimizing* if $|\partial E \cap \Omega| \leq |\partial^* F \cap \Omega| + \lambda |E \Delta F|$ for any open $F$ such that $E \Delta F \subset\subset \Omega$. (See [**Eic09**] for details.) The main point here is that much like minimizing boundaries, $\lambda$-minimizing boundaries have good regularity properties and come with useful volume bounds. When $n \geq 8$, one still obtains a solution $\Sigma$ that satisfies the MOTS equation in a weak sense, but it might have a singular set of Hausdorff dimension at most $n - 8$. Eichmair's proof of Theorem 7.39 followed a suggestion of Schoen to produce MOTS via limits of solutions of the regularized Jang equation—a phenomenon that was first observed by Schoen and Yau in their proof of the spacetime positive energy theorem [**SY81b**].

Notice that in the time-symmetric case $k = 0$, this theorem reduces to the previously known fact that one can always find a stable minimal hypersurface with given boundary $\Gamma$, as long as $\Gamma$ lies on the mean convex boundary hypersurface. (A result of this type was alluded to in our proof of Theorem 4.7.)

In Theorem 7.39, $\Gamma$ could be empty, in which case $\partial_1\Omega$ and $\partial_2\Omega$ are made up of components of $\partial\Omega$ and the theorem produces a closed MOTS. This case (for $n < 7$) was also proved by Lars Andersson and Jan Metzger [**AM09**] who implemented Schoen's suggestion using different techniques. Using this, we obtain the following generalization of Theorem 4.7, due to Andersson, Eichmair, and Metzger [**AM09**, **Eic09**, **AEM11**].

**Theorem 7.40** (Existence and uniqueness of apparent horizons)**.** *Let $n < 8$, and let $(M^n, g, k)$ be a complete asymptotically flat initial data set (possibly with boundary).*

(1) *If $M$ has nonempty weakly outer trapped boundary and only one end, then there exists a smooth apparent horizon.*
(2) *If an end of $M$ has an apparent horizon, then it is unique, and moreover both the horizon and the region outside the horizon are orientable.*

**Sketch of the proof.** The proof is similar to that of Theorem 4.7. We start with the first statement and construct a stable MOTS homologous to a large coordinate sphere. By asymptotic flatness, the mean curvature of the coordinate sphere of radius $\rho$ is approximately $\frac{n-1}{\rho}$, while $k$ decays faster. Therefore the end is foliated by outer untrapped hypersurfaces. Let $\Omega$ be the region of $\operatorname{Int} M$ whose boundary is a large coordinate sphere $S_\rho$, so that $\partial_1\Omega$ can be taken to be $S_\rho$ and $\partial_2\Omega$ can be taken to be $\partial M$ (with normal vector pointing into $\Omega$), while $\Gamma = \emptyset$. We can now see that $\Omega$ satisfies the hypotheses of Theorem 7.39, and therefore we obtain closed stable MOTS homologous to $\partial_1\Omega = S_\rho$.

From here, we can argue as we did in Theorem 4.7, though there are some added technical issues to address. See [**AEM11**] for details. □

We immediately obtain a generalization of Corollary 4.9.

**Corollary 7.41.** *If $(M^n, g, k)$ is a complete initial data set with* more than one *end, then there is an apparent horizon corresponding to each end. The result still holds if $M$ has a boundary, as long as that boundary is weakly outer trapped.*

Recall from Proposition 2.18 that every stable, two-sided closed minimal surface in an orientable 3-manifold with positive scalar curvature must be a sphere. The exact same reasoning leads to the following.

**Exercise 7.42.** Let $(M^3, g, k)$ be an orientable initial data set satisfying the *strict* dominant energy condition, meaning that $\mu > |J|_g$ everywhere. Prove that every stable, two-sided closed MOTS is a sphere.

Using Proposition 7.37, we can control the topology of apparent horizons as we did in the Riemannian case in Corollary 4.10.

**Theorem 7.43** (Galloway-Schoen [**GS06**], Galloway [**Gal18**])**.** *If $\Sigma$ is an apparent horizon in an initial data set $(M^n, g, k)$ satisfying the dominant energy condition, then $\Sigma$ is orientable and Yamabe positive. In particular, if $n = 3$, then $\Sigma$ is a union of spheres.*

**Proof.** One critical observation is that an apparent horizon must be a stable two-sided orientable MOTS. This follows from the construction and uniqueness statement in Theorem 7.40, but it is instructive to see how the stability follows directly from the outermost property. Deform $\Sigma$ in the direction $\varphi\nu$, where $\varphi > 0$ is the principal eigenfunction of the MOTS stability operator (Definition 7.35) with eigenvalue $\lambda$. If $\Sigma$ were *not* stable, then we would have $\frac{\partial}{\partial t}\theta^+ = \mathrm{L}_\Sigma\varphi = \lambda\varphi < 0$, which would mean that these small deformations would have $\theta^+ < 0$ and therefore be outer trapped. By Theorem 7.39, we could then produce a MOTS strictly enclosing $\Sigma$, which would violate the outermost property of $\Sigma$.

As we saw in Proposition 7.37 and inequality (7.7), stability and the DEC imply that

$$0 \leq \lambda_1(\mathrm{L}_\Sigma) \leq \lambda_1(-\Delta_\Sigma + Q_\Sigma) \leq \frac{1}{2}\lambda_1(L_h),$$

where $L_h$ is the conformal Laplacian of $h = g|_\Sigma$. If any of these inequalities is strict, then it follows that $\Sigma$ is Yamabe positive, by Corollary 2.27.

Suppose that all of the equalities above hold and that $\Sigma$ is not Yamabe positive. We will argue that this leads to a contradiction by constructing a splitting of $M$ near $\Sigma$, which will contradict the outermost MOTS property of $\Sigma$. This part of the proof comes from [**Gal18**]. Note that we have already seen much of this argument in our proof of Theorem 2.41.

We follow the steps of the proof of Theorem 2.38 in order to construct a foliation with constant $\theta^+$. For each smooth function $u$ on $\Sigma$, we consider the image hypersurface $\Sigma[u]$ of $\Sigma$ under the map $F_u(x) = \exp_x(u(x)\nu)$. All hypersurfaces that are close to $\Sigma = \Sigma[0]$ in the smooth sense can be parameterized by functions $u$ that are close to zero. For $\alpha \in (0,1)$, consider the map $\Psi$ from a small ball in $C^{2,\alpha}(\Sigma) \times \mathbb{R}$ to $C^{0,\alpha}(\Sigma) \times \mathbb{R}$, defined by

$$\Psi(u,s) = \left( F_u^* \theta^+_{\Sigma[u]} - s, \frac{1}{|\Sigma|} \int_\Sigma u \, d\mu_\Sigma \right),$$

where $F_u^*\theta^+_{\Sigma[u]}$ is the mean curvature scalar of the image surface $\Sigma[u]$, pulled back to the original surface $\Sigma$. Then

$$D\Psi|_{(0,0)}(u,s) = \left( \mathrm{L}_\Sigma u - s, \frac{1}{|\Sigma|} \int_\Sigma u \, d\mu_\Sigma \right).$$

Since $\lambda_1(\mathrm{L}_\Sigma) = 0$, the kernel of $\mathrm{L}_\Sigma$ is spanned by a positive principal eigenfunction, and then it is an exercise to conclude that the kernel of $\mathrm{L}^*_\Sigma$ is also spanned by a positive function. From this one can see that $D\Psi|_{(0,0)}$ is an isomorphism. By the inverse function theorem (Theorem A.43), there exists $\epsilon > 0$ and a smooth map $(v, \theta^+) : (-\epsilon, \epsilon) \longrightarrow C^{2,\alpha}(\Sigma) \times \mathbb{R}$ such that $\Psi(v(t), \theta^+(t)) = (0,t)$ for all $t \in (-\epsilon, \epsilon)$. Just as in the proof of Theorem 2.38, this means that in a neighborhood of $\Sigma$, we can write the metric $g$ as

$$g = h_t + \varphi_t^2 dt^2,$$

where $h_t$ is the induced metric on the hypersurface level sets $\Sigma_t := \Sigma \times \{t\}$, and each level set has constant $\theta^+_{\Sigma_t} = \theta^+(t)$, and $t > 0$ is on the outside of $\Sigma$.

According to Proposition 7.32 and Definition 7.35, we have

$$(\theta^+)'(t) = \mathrm{L}^0_{\Sigma_t}\varphi_t + \frac{1}{2}\theta^+(t)\left[\theta^-_{\Sigma_t} + 2k(\nu_t,\nu_t)\right]\varphi_t.$$

Choose a constant $C$ large enough so that $\frac{1}{2}\left[\theta^-_{\Sigma_t} + 2k(\nu_t,\nu_t)\right]\varphi_t \le C$ for all $t \in [0,\epsilon)$ and all points in $\Sigma_t$. We claim that $\theta^+(t) > 0$ for all $t \in (0,\epsilon)$. If not, then we could use the existence theorem for MOTS (Theorem 7.39) to construct a MOTS outside of $\Sigma$, but this contradicts the outermost property of $\Sigma$, proving the claim. So we have

$$(\theta^+)'(t) \le \mathrm{L}^0_{\Sigma_t}\varphi_t + C\theta^+(t),$$

and thus

$$\frac{d}{dt}\left[e^{-Ct}\theta^+(t)\right] \le e^{-Ct}\mathrm{L}^0_{\Sigma_t}\varphi_t.$$

Since $e^{-Ct}\theta^+(t)$ is positive for all $t \in (0, \epsilon)$ but zero at $t = 0$, it follows that there exists a time $t = \tau$ such that the left side of the above inequality is positive. Hence $\mathrm{L}^0_{\Sigma_\tau}\varphi_\tau > 0$, and thus $\varphi_\tau^{-1}\mathrm{L}^0_{\Sigma_\tau}\varphi_\tau \geq c > 0$ for some constant $c$. Using this choice of $\Sigma_\tau$ and $\varphi_\tau$ in (7.6) and invoking (7.7), it follows that

$$0 < c \leq \lambda_1\left(-\Delta_{\Sigma_\tau} + Q_{\Sigma_\tau}\right) \leq \frac{1}{2}\lambda_1(L_{h_\tau}).$$

Thus $\Sigma$ is Yamabe positive, which is a contradiction. □

The $n = 3$ case of Theorem 7.43 was first observed by Hawking in the case of stationary spacetimes [**Haw72**]. (Interestingly, this predates Schoen and Yau's work on scalar curvature described in Chapter 1.) When the spacetime dimension is greater than 4, there do exist examples of apparent horizons with nonspherical topology. Most famously, there are the "black ring" stationary vacuum Einstein solutions of R. Emparan and H. Reall in 4+1 dimensions [**ER02**, **ER06**], whose apparent horizons have $S^2 \times S^1$ topology. See [**Chr15**, Chapter 2] for a mathematical exposition of the Emparan-Reall black rings. Other examples of apparent horizons with nonspherical topology have been constructed by Fernando Schwartz [**Sch08**], Kunduri and Lucietti [**KL14**], and Mattias Dahl and Eric Larsson [**DL16**]. See also recent work of M. Khuri, Y. Matsumoto, G. Weinstein, and S. Yamada on the topology of $(4 + 1)$-dimensional stationary bi-axisymmetric black holes [**KMWY18**] .

When $n = 3$, we have the following generalization of Theorem 4.11.

**Theorem 7.44** (Eichmair-Galloway-Pollack [**EGP13**]). *Let $(M^3, g, k)$ be an asymptotically flat initial data set whose boundary is either empty or a union of MOTS. Assume that $M$ contains no immersed MOTS in its interior. Then $M$ is diffeomorphic to the $\mathbb{R}^3$ minus a finite number of open balls.*

We already gave the proof as our proof of Theorem 4.11. The only difference is that now we have to use Corollary 7.41 in place of Corollary 4.9.

## 7.6. The Penrose inequality

We consider the following generalization of Conjecture 4.12.

**Conjecture 7.45** ((Spacetime) Penrose inequality conjecture). *Let $(M^n, g, k)$ be a complete asymptotically flat initial data set satisfying the dominant energy condition, and suppose it contans an apparent horizon $\Sigma$ with respect to one of the ends. Then if $m$ is ADM mass of that end, and $\Sigma'$ is the strictly minimizing hull of $\Sigma$, then*

$$m \geq \frac{1}{2}\left(\frac{|\Sigma'|}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}},$$

*and if equality holds, then the part of* $(M, g, k)$ *outside* $\Sigma$ *sits inside the Schwarzschild spacetime of mass* $m$.

Clearly, in the time-symmetric case $k = 0$, the inequality reduces to the one in Conjecture 4.12. We will now discuss Penrose's original physical motivation for this conjecture, though it is important to note that the original conjecture was only for $n = 3$ since that is the only case in which the physical motivation is valid. However, given that the Riemannian Penrose inequality is known to be true for $n < 8$, it seems reasonable to conjecture that the spacetime version also holds in higher dimensions.

Assuming weak cosmic censorship and the more general version of the final state conjecture, the initial data in Conjecture 7.45 will evolve under some appropriate matter model to a spacetime that is approaching a Kerr-Newman solution. By Hawking's argument described in Section 7.4.3, the apparent horizon $\Sigma$ must lie inside some black hole event horizon $\mathcal{H}$. By the Hawking area theorem, the cross-sections of $\mathcal{H}$ must have monotone nondecreasing area as we move forward in time. Meanwhile, since energy can only radiate away to infinity, the mass must be monotone nonincreasing. More precisely, it is the *Trautman-Bondi mass* [**BvdBM62**, **Tra58**, **Sac62**], which we have not discussed, that is nonincreasing, and we know that the Trautman-Bondi mass approaches the ADM mass in many situations (and hope or expect that it does so generally). Since the inequality

$$m_0 \geq \sqrt{\frac{|\mathcal{H} \cap M_0|}{16\pi}}$$

is known to be true for a standard slice $M_0$ of a Kerr-Newman spacetime of mass $m_0$, it should then follow that the inequality also holds for the original initial data set $M$. The final step is to replace $|\mathcal{H} \cap M|$ by $|\Sigma'|$, which holds because $\Sigma$ is enclosed by $\mathcal{H}$. Note that since an apparent horizon need not be outward-minimizing, we should not expect to be able to replace $|\mathcal{H} \cap M|$ by $|\Sigma|$.

The cleverness of Penrose's conjecture is that he took a complicated conjectural picture about the future development of an initial data set under Einstein's equations and used it to produce a highly nontrivial, nonobvious conclusion about quantities that are well-defined in the initial data set itself. The main purpose of this was to test the plausibility of weak cosmic censorship. For a much longer and better discussion of the motivation behind the Penrose inequality and its consequences, see the survey [**Mar09**].

As we have seen in Chapter 4, this conjecture has been proved in the time-symmetric case. The general conjecture is essentially wide-open, though we do have a result for the spherically symmetric case. For this purpose we consider a boundary $\partial M$ which is an "outermost MOTS/MITS," meaning that $\partial M$ is a MOTS or a MITS (that is, a marginally inner trapped surface) and there are no other MOTS nor MITS enclosing it.

**Theorem 7.46** (Spherically symmetric Penrose inequality [**MḾ94**, **Hay96**]). *Let $(M, g, k)$ be a complete asymptotically flat initial data set diffeomorphic to $[0, \infty) \times S^{n-1}$ which is spherically symmetric in the sense that the metric can be expressed as*

$$g = ds^2 + r^2 d\Omega^2$$

*for some smooth positive function $r(s)$, where $d\Omega^2$ is the standard metric on the sphere, while $k$ can be written as*

$$k = k_{\nu\nu} ds^2 + \frac{1}{n-1} \kappa r^2 d\Omega^2,$$

*where $k_{\nu\nu}$ and $\kappa$ are smooth functions of $s$. Note that $k_{\nu\nu} = k(\nu, \nu)$, where $\nu$ is the outward normal of a symmetric sphere at $s$, while $\kappa$ is the trace of $k$ over the symmetric sphere at $s$.*

*If $(g, k)$ satisfies the dominant energy condition and $\partial M$ is an outermost MOTS/MITS, then*

$$m \geq \frac{1}{2} \left( \frac{|\partial M|}{\omega_{n-1}} \right)^{\frac{n-2}{n-1}},$$

*where $m$ is the ADM mass of $(M, g)$. Moreover, if equality holds, then $(M, g, k)$ lies inside the Schwarzschild spacetime of mass $m$.*

**Proof.** This exposition is based on the proof appearing in [**Mar09**, Section 4]. We claim that $\partial M$ is outward-minimizing, which is equivalent to saying that $\frac{dr}{ds} > 0$ in the interior of $M$. Since $\frac{dr}{ds} = \frac{1}{n-1} H r$, this is also equivalent to saying that there are no minimal spheres strictly enclosing $\partial M$. If there were a minimal sphere strictly enclosing $\partial M$, then it would have $H = 0$ there, and since $H > |\kappa|$ for large $r$, it follows from continuity that there must be some sphere with $H = |\kappa|$ that strictly encloses $\partial M$. But that would be a MOTS or a MITS, violating the outermost property of $\partial M$. Note that this argument also shows that the statement in Conjecture 7.45 follows from Theorem 7.46 for spherically symmetric spaces, since in this case the minimizing hull $\Sigma'$ of any MOTS $\Sigma$ is either $\Sigma$ itself or else $\Sigma'$ is minimal. In either case, $|\Sigma'|$ must be less than or equal to the volume of the outermost MOTS/MITS enclosing it.

Since $\frac{dr}{ds} > 0$ on the interior of $M$, we can perform a change of variables, just as in the proof of Proposition 4.20, so that

$$g = V^{-1} dr^2 + r^2 d\Omega^2$$

on $[r_0,\infty)\times S^{n-1}$, where $\sqrt{V}=\frac{dr}{ds}$ and $r_0$ satisfies $|\partial M|=\omega_{n-1}r_0^{n-1}$. We consider the function

$$\begin{aligned} m(r) &:= \frac{1}{2}r^{n-2}\left(1+\frac{1}{(n-1)^2}(\kappa^2-H^2)r^2\right)\\ &= \frac{1}{2}r^{n-2}(1-V)+\frac{1}{2(n-1)^2}\kappa^2 r^n, \end{aligned}$$

where the second expression can be easily compared to the one used in the proof of Proposition 3.20. In the $n=3$ case, this can be written more geometrically as

$$m(r)=\sqrt{\frac{|S_r|}{16\pi}}\left(1+\frac{1}{16\pi}\int_{S_r}\theta^+\theta^-\,d\mu_{S_r}\right),$$

where the right side is the appropriate generalization of the Hawking mass of the sphere $S_r$ to the initial data setting. In the spherically symmetric setting, it is called the *Misner-Sharp energy.*

**Claim.**

$$m'(r)=\frac{1}{n-1}r^{n-1}\left(\mu-\frac{\kappa}{H}\langle J,\nu\rangle\right).$$

In the proof of Proposition 3.20, we already saw that

$$(n-1)^2\frac{d}{dr}\left[\frac{1}{2}r^{n-2}(1-V)\right]=\frac{n-1}{2}r^{n-1}R. \tag{7.8}$$

Meanwhile, looking at the $\kappa$ term in $m(r)$, we compute

$$(n-1)^2\frac{d}{dr}\left[\frac{1}{2(n-1)^2}\kappa^2 r^n\right]=r^{n-1}\left(r\kappa\frac{d\kappa}{dr}+\frac{n}{2}\kappa^2\right). \tag{7.9}$$

**Exercise 7.47.** Show that in the spherically symmetric case, the constraint equations (Definition 7.16) reduce to

$$\begin{aligned} \mu &= \frac{1}{2}\left(R+2\kappa k_{\nu\nu}+\frac{n-2}{n-1}\kappa^2\right),\\ \langle J,\nu\rangle &= H\left(-\frac{1}{n-1}r\frac{d\kappa}{dr}+\kappa_{\nu\nu}-\frac{1}{n-1}\kappa\right). \end{aligned}$$

Rearranging the result of the exercise above for inclusion in equations (7.8) and (7.9),

$$\begin{aligned} \frac{n-1}{2}r^{n-1}R &= r^{n-1}\left[(n-1)\mu-(n-1)\kappa k_{\nu\nu}-\frac{n-2}{2}\kappa^2\right],\\ r^n\kappa\frac{d\kappa}{dr} &= r^{n-1}\left(-(n-1)\frac{\kappa}{H}\langle J,\nu\rangle+(n-1)\kappa k_{\nu\nu}-\kappa^2\right). \end{aligned}$$

Feeding these into (7.8) and (7.9), we obtain the Claim. Another way to derive the formula for $m'(r)$ is to start with the general variation formula for

$\theta^+$ in Proposition 7.32 as well as the analogous formula for $\theta^-$ (Exercise 7.34) and specialize to the case of spherical symmetry.

Again, observe that since $H > |\kappa|$ for large $r$, any sphere with $H < |\kappa|$ will lead to the existence of a MOTS or MITS enclosing it. Therefore the outermost property of $\partial M$ implies that we must have $H > |\kappa|$ in the interior of $M$. From this, the dominant energy implies that $\mu \geq \frac{\kappa}{H}\langle J, \nu\rangle$, and hence $m'(r) \geq 0$.

Finally, it is easy to check that since $\kappa^2 = H^2 = (n-1)^2 V(r_0) r_0^{-2}$ at $\partial M$, we have

$$\frac{1}{2}\left(\frac{|\partial M|}{\omega_{n-1}}\right)^{\frac{n-2}{n-1}} = \frac{1}{2} r_0^{n-2} = m(r_0) \leq \lim_{r\to\infty} m(r) = E,$$

where $E$ is the ADM energy. The last equality follows from Exercise 3.21, since the $\kappa^2$ term in $m(r)$ decays too fast to contribute to the limit. Finally, one can easily see that spherical symmetry implies that the ADM momentum is zero (as you would expect), so that the ADM energy $E$ is the same as the ADM mass $m$.

Next we consider the case of equality, which is unfortunately a bit messy. We will show that if equality holds, then $M$ sits inside the Schwarzschild spacetime of mass $m$ as a graph over the $t = 0$ slice. To do this, let us first make some calculations on graphical hypersurfaces. Consider the spacetime Schwarzschild metric of mass $m$,

$$\mathbf{g}_m = -V_m dt^2 + V_m^{-1} dr^2 + r^2 d\Omega^2,$$

where $V_m(r) = 1 - 2mr^{2-n}$. Next consider the graph of a radial function $f(r)$ over the $t = 0$ slice, that is, we look at the hypersurface defined by $t = f(r)$. A quick calculation shows that if we write the induced metric $g_f$ on the graph as

$$g^f = V_f^{-1} dr^2 + r^2 d\Omega^2,$$

then

$$V_f^{-1} = V_m^{-1} - V_m (f')^2.$$

We can also consider the second fundamental form $k^f$ of the graph in $\mathbf{g}_m$ and look at its normal-normal component $k^f_{\nu\nu}$ and its trace over the sphere $\kappa^f$ (both computed with respect to $g^f$). Some routine but somewhat involved computations (try it) show that

$$\kappa^f = \frac{n-1}{r}\sqrt{V_f - V_m}, \tag{7.10}$$

$$k^f_{\nu\nu} = \frac{d}{dr}\sqrt{V_f - V_m}. \tag{7.11}$$

By spherical symmetry, these components determine all of $k^f$. Therefore, in order to prove that some given spherically symmetric $(M, g, k)$ sits inside the Schwarzschild spacetime of mass $m$, it is sufficient to find some $f$ such that $V_f = V$, $\kappa^f = \kappa$, and $k^f_{\nu\nu} = k_{\nu\nu}$.

Now assume that $(M, g, k)$ satisfies equality in the Penrose inequality. By our proof of that inequality, we have $m(r) = m$ for all $r \geq r_0$. This translates into the statement that

$$V - V_m = \left(\frac{\kappa r}{n-1}\right)^2. \tag{7.12}$$

In particular, this means that it is possible to find some $f$ such that $V_f = V$. We now just need to show that with this choice of $f$, $\kappa^f = \kappa$ and $k^f_{\nu\nu} = k_{\nu\nu}$. The equation for $\kappa^f$ follows directly from (7.10) and (7.12). To prove the equation for $k^f_{\nu\nu}$, observe that since $m'(r) = 0$ for all $r \geq r_0$, our Claim above says that $\mu = \frac{\kappa}{H}\langle J, \nu\rangle$. Recall that $\frac{|\kappa|}{H} < 1$ outside $\partial M$ by the outermost MOTS/MITS assumption, so together with the DEC, this tells us that $\mu = \langle J, \nu\rangle = 0$. By Exercise 7.47, it follows that

$$k_{\nu\nu} = \frac{d}{dr}\left(\frac{\kappa r}{n-1}\right).$$

Combining this with (7.11) and 7.12, we see that $k^f_{\nu\nu} = k_{\nu\nu}$, completing the proof. □

Getting back to the general Penrose conjecture (Conjecture 7.45), H. Bray and M. Khuri have suggested an approach that involves using a generalized version of the Jang equation used in [**SY81b**] and "coupling" it with either inverse mean curvature flow or the Bray flow. This approach reduces the problem to one of solving a system of coupled PDEs. See [**BK10**].

In fact, even a positive mass theorem for initial data sets with MOTS boundary, which would be the natural spacetime analog of Theorem 4.16, does not easily follow from the version without boundary, since there is no doubling trick for MOTS. However, we do have a result in the spin case. See Theorem 8.29.