MIT lectures on the geometry and analysis of black hole spacetimes in general relativity

Mihalis Dafermos*

November 8, 2007

Contents

*University of Cambridge, Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB United Kingdom

1 Lecture I: What is a black hole?

General relativity concerns *Lorentzian manifolds* satisfying the so-called *Einstein equations*. We will turn to a more detailed discussion of the analysis of the Einstein equations in the next lecture. The purpose of this lecture is to introduce one of the most central notions of general relativity, that of *black hole*. This notion is tied to global Lorentzian causality. In the present lecture, we shall give the basics of Lorentzian geometry and then take the shortest route to the notion of black holes, that provided by *Penrose diagrams*.

1.1 Basic Lorentzian geometry

1.1.1 Lorentzian metrics

Definition 1.1. *Let $\mathcal{M}^{n+1}$ be a C^1 oriented $n+1$-dimensional manifold. A* **Lorentzian metric** *$g_{\mu\nu}$ is a C^0 non-degenerate covariant symmetric 2-tensor of signature $(-,+,\ldots,+)$. We call the pair $(\mathcal{M}, g)$ a* **Lorentzian manifold**.

In general relativity, physical spacetime is represented by a $3+1$-dimensional Lorentzian manifold $(\mathcal{M}^{3+1}, g)$ with a piece of extra structure to be defined in Section 1.1.3. Although the most important case is $n = 3$, note that one often considers submanifolds, quotient manifolds, etc. of lower dimension, which again inherit Lorentzian metrics. The case $n > 3$ is central for current (c. 2007) speculation amongst high energy physicists.

Definition 1.2. *Let $v \in T_p\mathcal{M}$. We call v* **timelike** *if $g(v, v) < 0$,* **null**[1] *if $v \neq 0$, $g(v, v) = 0$, and* **spacelike** *if*[2] *$g(v, v) > 0$. If $v \neq 0$, $g(v, v) \geq 0$, we say that v is* **causal**.

These appellations are inherited by vector fields and curves:

Definition 1.3. *Let $S \subset \mathcal{M}$, and let $V \in \Gamma(T\mathcal{M}|_S)$. We will call V* **timelike**, *etc., if $V(p)$ is timelike for all $p \in S$.*

Definition 1.4. *Let $\gamma : I \to \mathcal{M}$ be a C^1 curve where $I \subset \mathbb{R}$. We call γ* **timelike**, *etc., if $\gamma'(t)$ is timelike, etc., for all t.*

For general C^1 immersed submanifolds $i : \mathcal{N}^d \to \mathcal{M}$, with $d < n + 1$ we make the following definitions

Definition 1.5. *We call $\mathcal{N}^d$* **spacelike** *if i^*g is Riemannian,* **timelike** *if i^*g is Lorentzian, and* **null** *if i^*g is degenerate.*

The above definition agrees with the previous for curves.

For $d = n$, note that any hyperplane P through the origin in the tangent bundle has a unique 1-dimensional orthogonal complement. We will call a generator of this orthogonal complement a normal for P. $\mathcal{N}^n$ is spacelike iff its

[1] in some conventions denoted *lightlike*

[2] in some conventions, the 0 vector is also counted as spacelike

normal is timelike, $\mathcal{N}^n$ is timelike iff its normal is spacelike, and $\mathcal{N}^n$ is null iff its normal is null. Note that in the latter case, the normal is contained in the tangent space.

1.1.2 The Levi-Civita connection, geodesics, curvature

First let us note that the inverse metric $g^{\mu\nu}$ can be defined as in Riemannian geometry by $g^{\mu\nu}g_{\nu\lambda} = \delta^\mu_\lambda$ where δ here denotes the Kronecker symbol. Here, following the Einstein summation convention repeated indices are to be summed. Indices will be raised with the inverse metric in the usual way. In addition, all notations ever invented to denote geometric objects will be used interchangably when convenient in what follows.

Definition 1.6. *Let $\mathcal{M}$ be a C^3 manifold with g a C^2 Lorentzian metric. We define* **the Levi-Civita connection** *to be the connection ∇ in $T\mathcal{M}$ defined by $(\nabla_\mu V)^\nu = \partial_\mu V^\nu + \Gamma^\nu_{\mu\lambda}V^\lambda$ where*

$$\Gamma^\nu_{\mu\lambda} \doteq \frac{1}{2}g^{\nu\alpha}(\partial_\lambda g_{\mu\alpha} + \partial_\mu g_{\lambda\alpha} - \partial_\alpha g_{\mu\lambda}).$$

Definition 1.7. *Let $\mathcal{M}$ be a C^3 manifold with g a C^2 Lorentzian metric. We define the* **Riemann curvature tensor** *$R^\mu_{\nu\lambda\rho}$ to be the $(1,3)$-tensor defined by*

$$R^\mu_{\nu\lambda\rho} \doteq \partial_\lambda\Gamma^\mu_{\rho\nu} - \partial_\rho\Gamma^\mu_{\lambda\nu} + \Gamma^\alpha_{\rho\nu}\Gamma^\mu_{\lambda\alpha} - \Gamma^\alpha_{\lambda\nu}\Gamma^\mu_{\rho\alpha},$$

the **Ricci curvature** *$R_{\mu\nu}$ to be the covariant 2-tensor defined by*

$$R_{\mu\nu} \doteq R^\alpha_{\mu\alpha\nu}$$

and the **scalar curvature** *R to be the scalar defined by*

$$R \doteq g^{\mu\nu}R_{\mu\nu}.$$

Note: Our physical spacetimes $(\mathcal{M}, g)$ will satisfy the **Einstein equations**

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{1}$$

where here Λ is a fixed real constant, called the *cosmological constant*, and where $T_{\mu\nu}$ is a symmetric tensor which is determined by matter fields which themselves must satisfy a system of equations to close the system (1). Equations (1) implies that $T_{\mu\nu}$ is divergence free. In the special case where $T_{\mu\nu} = 0$, the system (1) is equivalent to $R_{\mu\nu} = \Lambda g_{\mu\nu}$. These are known as the **vacuum Einstein equations**. In the study of astrophysical isolated self-gravitating systems, it is appropriate to set $\Lambda = 0$. Current models for cosmology depend on a small $\Lambda > 0$. We will return to the system (1) later on…

Definition 1.8. *Let $\mathcal{M}$ be a C^3 manifold with g a C^2 Lorentzian metric. A C^2 parametrized curve $\gamma : I \to \mathcal{M}$ is a* **geodesic** *if*

$$\nabla_{\gamma'}\gamma' = 0,$$

i.e., if, in local coordinates x^α

$$\frac{d^2x^\mu(\gamma(t))}{dt^2} + \Gamma^\mu_{\nu\lambda}\frac{dx^\nu(\gamma(t))}{dt}\frac{dx^\lambda(\gamma(t))}{dt} = 0.$$

Proposition 1.1.1. *If $\gamma : I \to \mathcal{M}$ is a geodesic and $\gamma'(0)$ is timelike, spacelike, null, causal, resp., then γ is a timelike, spacelike, null, causal, resp., curve.*

Proof.

$$\gamma'(g(\gamma', \gamma')) = 2(\nabla_{\gamma'}\gamma', \gamma') = 0.$$

□

Thus we may define

Definition 1.9. *A geodesic γ is said to be a* **timelike**, **spacelike**, **null**, **causal**, *resp.,* **geodesic** *if its tangent vector at a single point is timelike, spacelike, null, causal, resp.*

Of these geodesics, the timelike and null have special importance in the physical interpretation of general relativity. Timelike geodesics correspond to "freely falling observers" or "test particles". In principle, the metric g is "measured" by examining the trajectories of these. Null geodesics correspond to the trajectories of "light rays" in the geometric optics limit. They also correspond to the characteristics of the Einstein equations for the metric g.

Definition 1.10. *A parametrized C^1 curve $\gamma : I \to \mathcal{M}$ is* **inextendible** *if it is not the restriciton to I of a C^1 curve $\gamma : J \to \mathcal{M}$ where $J \supset I$ properly.*

Definition 1.11. *Let ($\mathcal{M}$ be a C^3 manifold with C^2 Lorentzian metric g. We will say that $(\mathcal{M}, g)$ is* **geodesically complete** *if all inextendible geodesics have domain $\mathbb{R}$. Otherwise we say that $(\mathcal{M}, g)$ is* **geodesically incomplete**. *We say that $(\mathcal{M}, g)$ is* **timelike**, **null**, **spacelike**, **causally**, *resp.,* **geodesically (in)complete** *if the above holds where only timelike, etc., geodesics are considered.*

1.1.3 Time orientation and causal structure

Let $\mathcal{T} \subset \Gamma^0(T\mathcal{M})$ denote the subset of continuous timelike vector fields.[3] Define an equivalence relation on $\mathcal{T}$ by $T_1 \sim T_2$ if $g(T_1(p), T_2(p)) < 0$, for all p. One easily sees that either $\mathcal{T} = \emptyset$, or $\mathcal{T}/\sim$ has two elements.

Definition 1.12. *If $\mathcal{T} \neq \emptyset$, we say that $(\mathcal{M}, g)$ is* **time-orientable**. *An element of $\tau \in \mathcal{T}/\sim$ is called a* **time orientation**. *A triple $(\mathcal{M}, g, \tau)$, where τ is a time-orientation is called a* **time oriented Lorentzian manifold**. *We call such a triple a* **spacetime**.

[3] Note that a timelike vector field in particular does not vanish!

Definition 1.13. *Let $(\mathcal{M}, g, \tau)$ be a time-oriented Lorentzian manifold, and let $v \in T_p\mathcal{M}$ be causal. Let T a representative for τ. We say that v is* **future pointing** *if $g(T(p), v) < 0$ and* **past pointing** *if $g(T(p), v) > 0$.*[4]

By the above definition, $T(p)$ is future pointing. The definition does not depend on the choice of representative T for τ.

The above appellations are inherited by vector fields and causal curves as in Definitions 1.2 and 1.3, i.e. we may speak of **future-directed timelike curves**, etc.

Starting in Section 1.2, **all Lorentzian manifolds will be assumed time-oriented** and we shall often suppress the notation τ for the choice of time orientation.

Definition 1.14. *Let $S \subset \mathcal{M}$. Define the sets $J^{\pm}(S)$ by*

$$J^{+}(S) \doteq \{q \in \mathcal{M} : \exists\gamma : [0,1] \to \mathcal{M} \text{ causal future-directed}, \gamma(0) \in S, \gamma(1) = q\},$$

$$J^{-}(S) \doteq \{q \in \mathcal{M} : \exists\gamma : [0,1] \to \mathcal{M} \text{ causal past-directed}, \gamma(0) \in S, \gamma(1) = q\}.$$

We call $J^{+}(S)$ the **causal future** *of S, and $J^{-}(S)$ the* **causal past** *of S.*

There is also the notion of *chronological future* and *past* where "causal" is replaced in the above by "timelike". Those sets are denoted $I^{\pm}(S)$, etc. We will not use this notion here however.

We can now add to Definition 1.11 terms like **future-causal geodesic completeness**, etc.

1.1.4 Global hyperbolicity and Cauchy surfaces

Definition 1.15. *Let $(\mathcal{M}, g)$ be a Lorentzian manifold, and let $\Sigma \subset \mathcal{M}$ be a codimension-1 C^1 embedded spacelike submanifold. We say that Σ is a* **Cauchy surface** *if every inextendible immersed causal curve in $\mathcal{M}$ intersects Σ exactly once. A Lorentzian manifold admitting a Cauchy surface Σ is said to be* **globally hyperbolic with Cauchy surface** Σ.

Some authors define global hyperbolicity slightly differently and prove that the definition is equivalent to the above. Other authors strengthen the above definition by adding the assumption that Σ be complete.

Note that if $(\mathcal{M}, g)$ is globally hyperbolic with Cauchy surface Σ, then there exists a globally defined C^1 function $t : \mathcal{M} \to \mathbb{R}$ such that ∇t is past timelike. Such a function is called a *temporal function.* In particular, global hyperbolicity implies time-orientability. Moreover, it implies that $\mathcal{M}$ is diffeomorphic to $\Sigma \times \mathbb{R}$.

1.2 $\mathbb{R}^{1+1}$

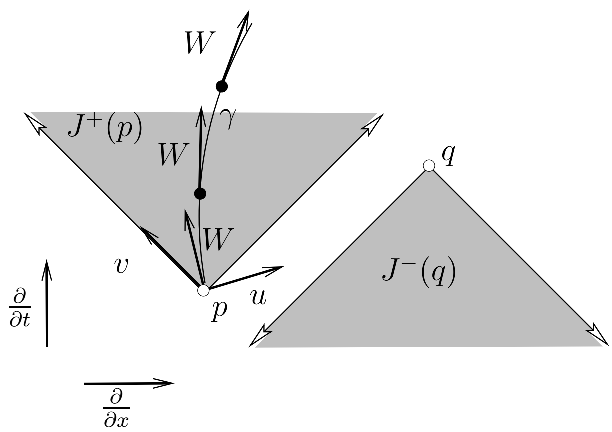
The above concepts are most easily illustrated in the case of **2-dimensional Minkowski space.** This is the Lorentzian manifold $(\mathbb{R}^2, -dt^2 + dx^2)$, where

[4] Note that one of these inequalities is necessarily satisfied for every causal vector.

(t, x) denote global coordinates on $\mathbb{R}^2$. Let us time-orient this manifold by $\frac{\partial}{\partial t}$. We shall denote the resulting spacetime by $\mathbb{R}^{1+1}$.

Let us agree to depict $\mathbb{R}^{1+1}$ on paper or the board so that the constant t-curves and constant x-curves are vertical and horizontal lines, equally calibrated with respect to the Euclidean geometry of the plane, and so that that $\frac{\partial}{\partial t}$ points upwards, and $\frac{\partial}{\partial x}$ to the right.

Let us also agree to superimpose tangent vectors on the manifold in the usual way. All concepts are illustrated in the diagram below:



It is to be understood that one can read off from the above the following information: The vector v is a future directed null vector. The vector u is spacelike. The curve γ is timelike with tangent vector field W. $\gamma \subset J^+(p)$, which is bounded by the rays emanating at 45 and 135 degrees from the horizontal at p, etc., etc.

1.3 $\mathbb{R}^{3+1}$

Let $\mathbb{R}^{3+1}$ denote the spacetime defined by the Lorentzian manifold $(\mathbb{R}^4, -dt^2 + dx_1^2 + dx_2^2 + dx_3^2)$, time-oriented by $\frac{\partial}{\partial t}$. This spacetime is called **Minkowski space**. Its geometry defines what is known as *special relativity*. We can relate the geometry of $\mathbb{R}^{3+1}$ to the geometry of $\mathbb{R}^{1+1}$ by exploiting the fact that $\mathbb{R}^{3+1}$ enjoys many symmetries with 2-dimensional group orbits and a 2-dimensional quotient. Since astrophysical systems are to a first approximation spherically symmetric, this symmetry plays by far the most important role in general relativity.

Definition 1.16. *We say that a 4-dimensional Lorentzian manifold $(\mathcal{M}, g)$ is* **spherically symmetric** *if $SO(3)$ acts by isometry.*

Consider the case of $\mathbb{R}^{3+1}$. Choose an $SO(3)$ action by isometry. Then the group quotient $\mathcal{Q} = \mathbb{R}^{3+1}/SO(3)$ inherits the structure of a 2-dimensional Lorentzian-manifold with boundary. In fact, $\mathcal{Q}$ is isometric to the right half plane $x \geq 0$ of $\mathbb{R}^{1+1}$ under the identification $x = r$, where r is the polar coordinate related to the $SO(3)$ action.

There is something more intelligent we can do. First a definition:

Definition 1.17. *Let $(\mathcal{M}, g, \tau)$, $(\widetilde{M}, \tilde{g}, \tilde{\tau})$ be time-oriented Lorentzian manifolds. We say that a C^1 map $\phi : \mathcal{M} \to \widetilde{\mathcal{M}}$ is* **(t.o.p.[5]) conformal** *if $\phi_*(v)$ is (future-pointing) null for all v (future-pointing) null.*

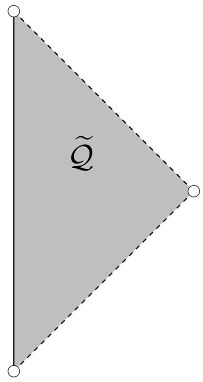
It follows from the definition that for all $p \in \mathcal{M}$, $\tilde{g}(\phi(v), \phi(w)) = \lambda(p) g(v, w)$ for a $\lambda > 0$. Thus, t.o.p. conformal maps correspond precisely to conformal maps in Riemannian geometry, modulo the additional stipulation that $\phi_*(T) \sim \tilde{T}|_{\phi(\mathcal{M})}$ for representatives T, $\tilde{T}$ of the time-orientations.

The causal structure of $(\mathcal{M}, g, \tau)$ is equivalent to the causal structure of $(\mathcal{M}, \phi^* \tilde{g}, \tau)$, in the sense that

$$J^+_g(S) = J^+_{\phi^* \tilde{g}}(S).$$

(In general, one must be careful to distinguish $J^+_{\tilde{g}|_{\phi(\mathcal{M})}}(\phi(S))$ and $J^+_{\tilde{g}}(\phi(S)) \cap \phi(\mathcal{M})$.)

Back to our example. It turns out that we can t.o.p. conformally map $\mathcal{Q}$ (realised as the subset $r \geq 0$) into $\mathbb{R}^{1+1}$ such that its image is the set:



for instance by

$$t = \frac{1}{2} \tan^{-1}(\bar{t} - r) + \frac{1}{2} \tan^{-1}(\bar{t} + r)$$
$$x = \frac{1}{2} \tan^{-1}(\bar{t} + r) - \frac{1}{2} \tan^{-1}(\bar{t} + r),$$

where we have temporarily denoted the t coordinate on $\mathbb{R}^{3+1}$ by $\bar{t}$ so as not to confuse it with the t coordinate of $\mathbb{R}^{1+1}$.

Let us denote the above submanifold-with-boundary of $\mathbb{R}^{1+1}$ as $\widetilde{\mathcal{Q}}$. Note that ϕ is a diffeomorphism onto its image $\widetilde{\mathcal{Q}}$, and, in view of the causal geometry of $\widetilde{\mathcal{Q}}$, it follows that for $S \subset \mathbb{R}^{1+1}$,

$$J^{\pm}_{\mathbb{R}^{1+1}}(S \cap \widetilde{\mathcal{Q}}) \cap \widetilde{\mathcal{Q}} = J^{\pm}_{\widetilde{\mathcal{Q}}}(S \cap \widetilde{\mathcal{Q}}).$$

[5] time-orientation preserving

The upshot of all this is that we can read off the causal structure of our original $\mathcal{Q}$, from the causal structure of $\mathbb{R}^{1+1}$.

But there's more. Since $\widetilde{\mathcal{Q}} \subset \mathbb{R}^{1+1}$ is bounded in the standard metric space topology of $\mathbb{R}^2$, we can consider its boundary

$$\partial_{\text{Penrose}} \widetilde{\mathcal{Q}} = (\text{clos}_{\mathbb{R}^2} \widetilde{\mathcal{Q}}) \setminus \widetilde{\mathcal{Q}}. \tag{2}$$

Recall that $\widetilde{\mathcal{Q}}$ was itself a manifold-with-boundary. We shall reserve the notation $\partial\widetilde{\mathcal{Q}}$ for the boundary of $\widetilde{\mathcal{Q}}$ in the sense of manifolds-with-boundary.

We can define two distinguished subsets, $\mathcal{I}^+$ and $\mathcal{I}^-$ of $\partial_{\text{Penrose}}\widetilde{\mathcal{Q}}$, as follows. First note the following more geometrical characterization of r:

Definition 1.18. *Let $\pi : \mathbb{R}^{3+1} \to \mathcal{Q}$ denote the standard projection, and let $r : \mathcal{Q} \to \mathbb{R}$ be defined by $r(p) = \sqrt{\text{Area}(\pi^{-1}(p))/4\pi}$. We call r the* **area-radius function**.

Note that r pulls back by ϕ^{-1} to a function on $\widetilde{\mathcal{Q}}$ which we shall unapologetically denote again by r.

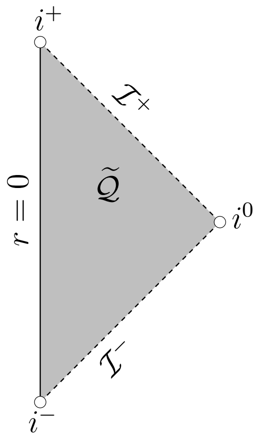
Definition 1.19. *Define $\mathcal{I}^\pm$ by*

$$\begin{aligned} \mathcal{I}^+ \quad &= \quad \{q \in \partial_{\text{Penrose}}\widetilde{\mathcal{Q}} : \exists \gamma : [0,1] \to \text{clos}(\widetilde{\mathcal{Q}}) \text{ null, future-directed,} \\ &\qquad \gamma([0,1)) \subset \widetilde{\mathcal{Q}}, \gamma(1) = q, \lim_{t\to 1} r(\gamma(t)) = \infty\}, \end{aligned} \tag{3}$$

$$\begin{aligned} \mathcal{I}^- \quad &= \quad \{q \in \partial_{\text{Penrose}}\widetilde{\mathcal{Q}} : \exists \gamma : [0,1] \to \text{clos}(\widetilde{\mathcal{Q}}) \text{ null, past-directed,} \\ &\qquad \gamma([0,1)) \subset \widetilde{\mathcal{Q}}, \gamma(1) = q, \lim_{t\to 1} r(\gamma(t)) = \infty\}. \end{aligned} \tag{4}$$

We call $\mathcal{I}^+$ **future null infinity**, *and $\mathcal{I}^-$* **past null infinity**.[6]

These sets are depicted below



The set $\partial_{\text{Penrose}}\widetilde{\mathcal{Q}}$ has three remaining points, call them i^0, i^+, i^-.

[6] Some people prefer to read out $\mathcal{I}^+$ as "scri plus".

Definition 1.20. *Let $\{i^0, i^+, i^-\} \subset \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$ be as depicted in the diagram. We call the point i^0* **spacelike infinity***, the point i^+* **future timelike infinity***, and the point i^-* **past timelike infinity***.*

We have the following fundamental (trivial) proposition

Proposition 1.3.1. $J^{\mp}(\mathcal{I}^{\pm})) \cap \widetilde{\mathcal{Q}} = \widetilde{\mathcal{Q}}$.

1.4 Exercises

Exercise 1.1. *For $\mathcal{Q} = \mathbb{R}^{3+1}/SO(3)$, $\phi : \mathcal{Q} \to \widetilde{\mathcal{Q}} \subset \mathbb{R}^{1+1}$ t.o.p. conformal, bounded, show that the boundary $\partial_{\mathrm{Penrose}}(\widetilde{\mathcal{Q}})$ does not depend on the choice of ϕ, $\widetilde{\mathcal{Q}}$, in the sense that given another ϕ', $\widetilde{\mathcal{Q}}'$, then there exists a bijection $\psi : \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}} \to \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}'$ such that*

$$\phi'^{-1}(J^{\pm}(\psi(p)) \cap \widetilde{\mathcal{Q}}') = \phi^{-1}(J^{\pm}(p) \cap \widetilde{\mathcal{Q}}).$$

Exercise 1.2. *Let π denote the projection $\pi : \mathbb{R}^{3+1} \to \mathcal{Q}$. Show that if γ is an inextendible null geodesic in $\mathbb{R}^{3+1}$, then $\pi(\gamma)$ has exactly two limit points on $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$, one on $\mathcal{I}^+$ and one on $\mathcal{I}^-$. Show that if γ is an inextendible timelike geodesic in $\mathbb{R}^{3+1}$ then $\pi(\gamma)$ has exactly two limit points on $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$, namely the points $i^{\pm}$. Finally, show that if γ is an inextendible spacelike geodesic in $\mathbb{R}^{3+1}$, then γ has the unique limit point i^0 on $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$.*

Exercise 1.3. *Show that a C^1 spacelike hypersurface $\Sigma \subset \mathbb{R}^{3+1}$ is a Cauchy surface iff it is complete and $\pi(\Sigma)$ has limit point i^0.*

Exercise 1.4. *Let ϕ be a solution of the wave equation $\Box\phi = 0$ on $\mathbb{R}^{3+1}$. Let Σ be a spherically symmetric Cauchy surface, and suppose ϕ_Σ, $\nabla\phi|_\Sigma$ are spherically symmetric and compactly supported on Σ. Show that ϕ is spherically symmetric, and descends to a function on $\widetilde{\mathcal{Q}}$. Show that $r\phi$ extends continuously to $\mathcal{I}^+$ and that $r\phi$ is compactly supported on $\mathcal{I}^+$. Make sense of this for non-spherically symmetric ϕ.*

1.5 Penrose(-Carter) diagrams

We have dwelled on the previous construction because it is not limited to $\mathbb{R}^{3+1}$.

Let $(\mathcal{M}, g)$ be a spherically symmetric spacetime. Let us suppose that the $SO(3)$ action is such that $\mathcal{Q} = \mathcal{M}/SO(3)$ inherits the structure of a Lorentzian manifold (possibly with boundary), and that moreover, $\mathcal{Q}$ can be globally t.o.p. conformally mapped into a bounded connected subset of $\mathbb{R}^{1+1}$.[7]

We have all that is necessary to repeat the construction of Section 1.3.

Definition 1.21. *We call the image $\widetilde{\mathcal{Q}}$ of a bounded t.o.p. conformal map $\phi : \mathcal{Q} \to \widetilde{\mathcal{Q}} \subset \mathbb{R}^{1+1}$, together with $r : \widetilde{\mathcal{Q}} \to \mathbb{R}$ defined by Definition 1.18, a* **Penrose**[8]

[7] In applications to spherically symmetric dynamics, these assumptions do not turn out to be particularly restrictive.

[8] The spirit of these diagrams is due to Penrose, but they were first used as formal objects in the sense described here by B. Carter. Hence, they are often called Penrose-Carter or even Carter diagrams.

diagram *for $\mathcal{M}$. We define a subset $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}} \subset \mathbb{R}^{1+1}$ by* (2) *and subsets $\mathcal{I}^{\pm} \subset \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$ by Definition 1.19.*

At this level of generality, Exercise 1.1 may not hold, i.e. the causal structure of $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$ is not necessarily unique! Under suitable assumptions on the geometry of $\mathcal{Q}$, however, it is. As these assumptions will hold in our examples, we will often courageously drop the $\widetilde{\mathcal{Q}}$ notation in what follows.

In analogy with Definition 1.19, let us define

$$\begin{aligned}\mathcal{B}^{+} &= \{q \in \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}} : \exists\gamma : [0,1] \to \mathrm{clos}(\widetilde{\mathcal{Q}}) \text{ null, future-directed,} \\ &\qquad \gamma([0,1)) \subset \widetilde{\mathcal{Q}}, \gamma(1) = q, \lim_{t\to 1} r(\gamma(t)) = 0\}\end{aligned}$$

$$\begin{aligned}\mathcal{B}^{-} &= \{q \in \partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}} : \exists\gamma : [0,1] \to \mathrm{clos}(\widetilde{\mathcal{Q}}) \text{ null, past-directed,} \\ &\qquad \gamma([0,1)) \subset \widetilde{\mathcal{Q}}, \gamma(1) = q, \lim_{t\to 1} r(\gamma(t)) = 0\}.\end{aligned}$$

Be careful to distinguish these boundaries from $\partial\mathcal{Q}$ of $\mathbb{R}^{3+1}$, which is part of the spacetime. In all examples here, $\mathcal{B}^{\pm}$ will be singular.

1.6 Schwarzschild

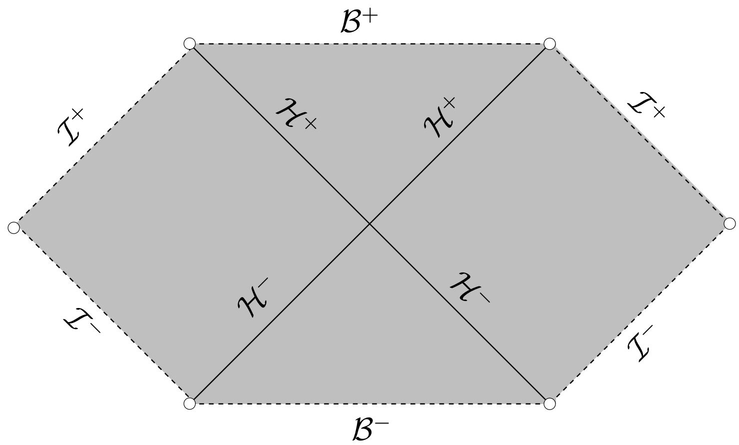
The Schwarzschild family $\mathbb{SCH}^{3+1}_M$ is a one parameter family of spherically symmetric solutions of the Einstein vacuum equations

$$R_{\mu\nu} = 0.$$

The parameter is traditionally called the *mass*, and is denoted M. The family contains Minkowski space: $\mathbb{SCH}^{3+1}_0 = \mathbb{R}^{3+1}$. For $M \neq 0$, the topology of $\mathbb{SCH}^{3+1}_M$ is $\mathbb{R} \times \mathbb{R} \times \mathbb{S}^2$. The quotient $\mathcal{Q} = \mathbb{SCH}^{3+1}_M/SO(3)$ is a 2-dimensional manifold without boundary, and with topology $\mathbb{R} \times \mathbb{R}$.

1.6.1 $M > 0$

For $M > 0$, $\mathbb{SCH}_M^{3+1}$ is globally hyperbolic with complete Cauchy surface Σ diffeomorphic to $\mathbb{S}^2 \times \mathbb{R}$. A Penrose diagram of $\mathbb{SCH}_M^{3+1}$ is given below:



At this point, let us make the following definitions:

Definition 1.22. *When $\mathcal{Q} \setminus J^-(\mathcal{I}^+) \neq \emptyset$, we call this set the* **black hole** *region of $\mathcal{M}$. We define the set $\mathcal{H}^+ = (\partial J^-(\mathcal{I}^+)) \cap \mathcal{Q}$ to be the* **event horizon** *of the black hole.*

We also have the dual notions

Definition 1.23. *When $\mathcal{Q} \setminus J^+(\mathcal{I}^-) \neq \emptyset$, we call this set the* **white hole** *region of $\mathcal{M}$. We define the set $\mathcal{H}^- = (\partial J^+(\mathcal{I}^-)) \cap \mathcal{Q}$ to be the* **event horizon** *of the white hole.*

From Proposition 1.3.1, it follows that $\mathbb{R}^{3+1}$ space does not have a black hole region. We see immediately from the Penrose diagram above that $\mathbb{SCH}_M^{3+1}$ for $M > 0$ contains both a black hole and a white hole region.

Let us note some additional geometric facts: Recall the function r defined by Definition 1.18. By definition of the sets $\mathcal{I}^{\pm}$, it follows that $r \to \infty$ along null rays terminating on these. Similarly, by definition of $\mathcal{B}^{\pm}$, $r \to 0$ along null rays terminating on these. Moreover, the Kretchmann scalar $R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta}$ blows up along such curves, thus $\mathcal{B}$ can be viewed as a singular boundary. The function r assumes the constant value $2M$ along $\mathcal{H}^+ \cup \mathcal{H}^-$. The spacetime $\mathcal{M}$ is both future and past timelike geodesically incomplete and null geodesically incomplete. Indeed, any causal geodesic which enters the black hole region reaches $\mathcal{B}^+$ in finite affine time.

Each point of $p \in \mathcal{Q}$ of the black hole region corresponds to a *future trapped surface* in $\mathcal{M}$. We shall define this notion precisely later on. In the spherical symmeric context, it just means that both future-directed null derivatives of r are negative at p. It turns out that the future geodesic incompleteness of $\mathcal{M}$

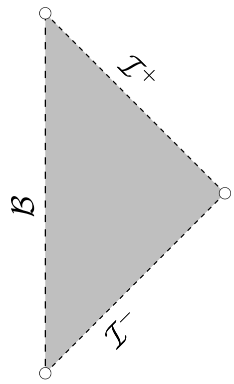
is guaranteed by the existence of a single such surface, in view of the global hyperbolicity of $\mathbb{SCH}_M^{3+1}$ and the non-compactness of Σ. This is known as the *Penrose incompleteness theorem*.[9]

All causal geodesics *not* entering the black hole region to the future exist for all positive affine time. Similarly, for causal geodesics not emerging from the white hole and negative affine time.

It follows easily from the above and from the properties described previously of geodesics entering into or emerging from the black/white holes, respectively, that the spacetime $\mathbb{SCH}_M^{3+1}$ is C^2-inextendible, i.e. there does not exist a 4-dimensional Lorentzian $(\widetilde{\mathcal{M}}, \tilde{g})$ with C^2 $\tilde{g}$ and an isometry $i : \mathbb{SCH}_M^{3+1} \to \widetilde{\mathcal{M}}$ such that $i(\mathbb{SCH}_M^{3+1}) \neq \mathcal{M}$.

1.6.2 $M < 0$

For $M > 0$, $\mathbb{SCH}_M^{3+1}$ is *not* globally hyperbolic and has Penrose diagram



The Kretchmann scalar blows up along any curve approaching $\mathcal{B}$.

One often describes the above situation as saying that $\mathcal{B}$ corresponds to a "naked singularity". Here "naked" means that it is "visible" to $\mathcal{I}^+$, i.e. such that

$$\mathcal{B} \subset J^-(\mathcal{I}^+). \tag{5}$$

In these lectures, I encourage a different point of view that characterizes naked singularities from the geometry of globally hyperbolic spacetimes with a suitable infinity. For such spacetimes, (5) cannot hold, so the notion of nakedness must be defined more carefully. More on this later.

1.6.3 Historical digression

Schwarzschild found the form of the metric of $\mathbb{SCH}_M^{3+1}$ in local coordinates in 1916. This metric was presented a few years later by Hilbert in what are now called Schwarzschild coordinates

$$-(1 - 2M/r)dt^2 + (1 - 2M/r)^{-1}dr^2 + r^2 d\sigma_{\mathbb{S}^2}.$$

[9] often called "singularity theorem"

It is important to remember that at that time, neither the notion of manifold, nor global Lorentzian geometry were understood.

In our language, we would say that the underlying manifold one then had in mind was $(-\infty, \infty) \times (2M, \infty) \times \mathbb{S}^2$. This corresponds precisely to the subset $J^-(\mathcal{I}^+_A) \cap J^+(\mathcal{I}^-_A)$ of our $\mathbb{SCH}^{3+1}_M$, where $\mathcal{I}^\pm_A$ denote say the right connected components of $\mathcal{I}^\pm$.

The above coordinate system cannot be extended to a larger manifold so as to be regular at $r = 2M$. The question of "what happens at $r = 2M$" was posed early on, but somehow often misunderstood, even by the pioneers of the subject Hilbert and Einstein.[10] The extendibility of the metric, at least beyond say $\mathcal{H}^+ \setminus \mathcal{H}^-$, was in fact known to Lemaitre and Eddington in the 1930's. The whole $\mathbb{SCH}^{3+1}_M$ can be identified as a subset of an even larger (obviously non-regular, in view of inextendibility) manifold first described by Synge around 1950 in an obscure Irish journal. The manifold $\mathbb{SCH}^{3+1}_M$ was more succinctly described by Kruskal around 1960.

Of course, irrespective of the existence of $\mathbb{SCH}^{3+1}_M$ and its inextendibility, one can still ask whether it is all of $\mathbb{SCH}^{3+1}_M$ or just some subset which is "physical". Again, this was a source of much confusion, and in some sense could only be understood when the problem of dynamics was correctly formulated (see next lecture) allowing one to attach a unique spacetime $\mathcal{M}$ to a notion of initial data for a closed system of equations. The question "physical" is then transferred to initial data, where it can be more rationally addressed.

With the benefit of this hindsight, the above question in some sense had already been answered in 1939 in pioneering work of Oppenheimer and Snyder where a class of matter spacetimes was constructed–arising from physically plausible initial data with trivial topology–which included isometerically the subregion $J^+(\gamma) \subset \mathbb{SCH}^{3+1}_M$, where γ is some inextendible null geodesic emerging from $\mathcal{I}^-$.[11] This region contains part of $\mathcal{H}^+$ and part of the black hole region. This then is the region which strictly speaking is "physical".

As for the spacetime $\mathbb{SCH}^{3+1}_M$ itself, its Cauchy surface with its $\mathbb{R} \times \mathbb{S}^2$ topology hardly seems physically plausible. On the other hand, by soft arguments (Cauchy stability, domain of dependence, etc.), it turns out that questions about the region $J^+(\gamma) \subset \mathbb{SCH}^{3+1}_M$ are most succinctly expressed as questions about $\mathbb{SCH}^{3+1}$, in particular, the question of stability that we will dwell on later. In addition, $\mathbb{SCH}^{3+1}$ is a convenient bookkeeping device, for instance, in stating and proving Birkhoff's theorem (See Exercise 1.8).

1.6.4 Note on the definition of black holes

The notion of black hole is appropriate any time there is a reasonable notion of "future infinity". In the spherically symmetric setting considered here, a "poor man's infinity" can be captured in the manner outlined here by considering the boundary $\partial_{\mathrm{Penrose}}\widetilde{\mathcal{Q}}$ and the limiting behaviour of the function r.

[10] At that time, $r = 2M$ was known as "the Schwarzschild singularity".

[11] Of course, to actually deduce this from the paper of Oppenheimer and Snyder needs considerable hindsight, as none of the notions employed here existed at that time...

An important issue in general relativity has been to capture the notion of black hole for spacetimes arising from *asymptotically flat* initial data. Although we have not yet given a definition of this term (see next lecture), suffice it to say here that $\mathbb{R}^{3+1}$ and $\mathbb{SCH}^{3+1}_M$ are indeed examples of such spacetimes, and the asymptotically flat ends correspond to the limits $r \to \infty$ along Cauchy surfaces. Whereas attempts at general definitions have been made, these all have the problem of making *a priori* assumptions on the spacetime, as opposed to just on initial data. Thus, the relevance of these definitions to the problem of dynamics is far from clear.

Definition 1.22, though restricted in scope *a priori* to spherical symmetry, has the advantage that it is applicable to the dynamical problem, as spherical symmetry is an *evolutionary hypothesis*, i.e. it carries from data to solution. More on this in the next lectures.

A final note: The tentative Definition 1.22 gives a purely causal characterization of the notion of black hole. One often wants to add some sort of "completeness" criterion on $\mathcal{I}^+$. (For instance, one might object that under Definition 1.22, the spherically symmetric subset $\pi^{-1}(J^-(\gamma))$ of Minkowski space $\mathbb{R}^{3+1}$, where γ is a spacelike curve in $\widetilde{\mathcal{Q}}$ intersecting transversely $\mathcal{I}^+$, has a black hole region.) We shall give such a criterion later on. It turns out, however, that in this spherically symmetric setting, the completeness criterion is in practice superfluous, provided the definitions be applied always to suitable "maximal developments of initial data".

1.7 Reissner-Nordström

The Reissner-Nordström family $\mathbb{RN}^{3+1}_{e^2,M}$ is a two parameter family of spherically symmetric spacetimes solving the Einstein-Maxwell equations with a corresponding 2-form $F_{\mu\nu}$:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi T_{\mu\nu}$$

$$T_{\mu\nu} = \frac{1}{4\pi}\left(F_{\mu\nu}F_{\lambda\rho}g^{\nu\rho} - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F_{\gamma\delta}g^{\alpha\gamma}g^{\beta\delta}\right)$$

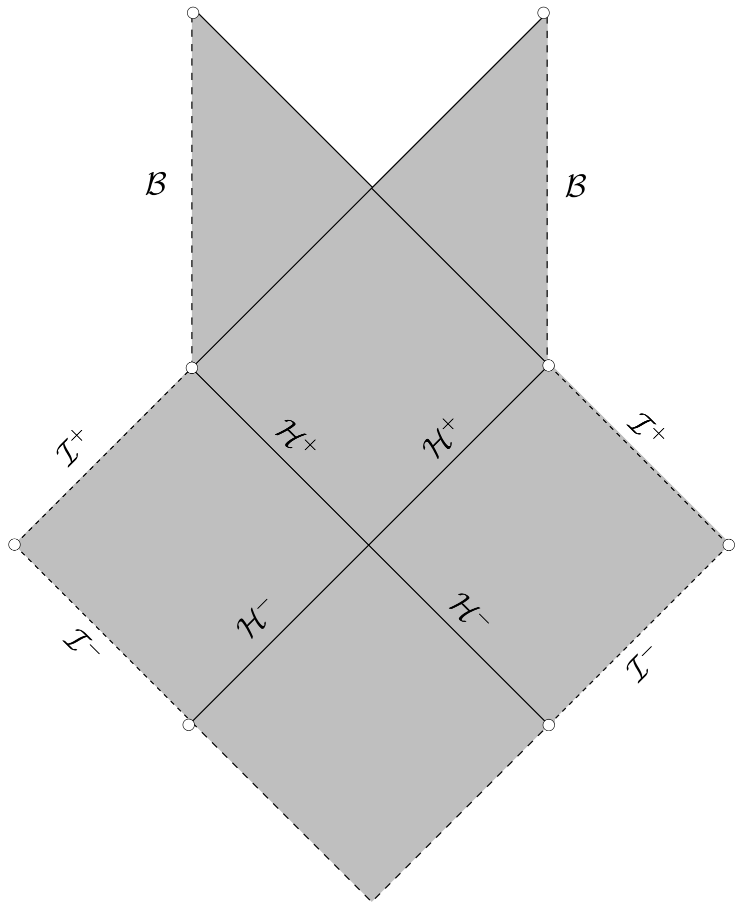
$$\nabla^\mu F_{\mu\nu} = 0$$

$$dF = 0.$$

The parameters range in $M \in (-\infty, \infty)$ and $e^2 \in [0, \infty)$. M is called the mass and e is called the charge. For $e^2 = 0$, the family corresponds to Schwarzschild, i.e. $\mathbb{RN}^{3+1}_{0,M} = \mathbb{SCH}^{3+1}_M$.

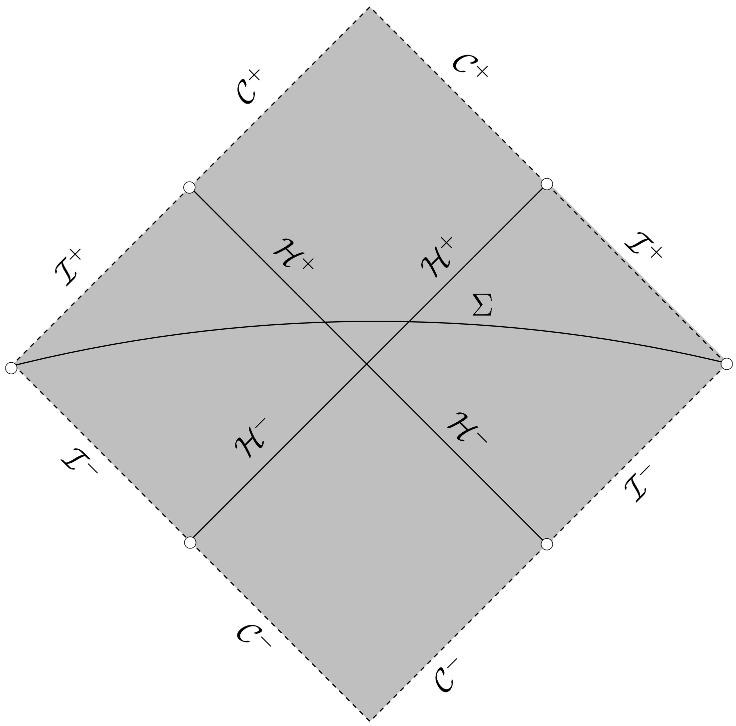
1.7.1 $M > 0$, $0 < e^2 < M^2$

The spacetime $\mathbb{RN}^{3+1}_{e^2,M}$ has a discrete global time periodicity. Below is the Penrose diagram of a fundamental domain



Note the existence of black hole and white hole regions for the fundamental domain in the sense of Definitions 1.22 and 1.23. The geometry of these regions is completely different, however, in particular, the sets $\mathcal{B} = \mathcal{B}^{\pm}$ are timelike, and this allows for the repeating structure to develop. As in Schwarzschild, the Kretchmann scalar blows up as $\mathcal{B}$ is approached along any curve. The spacetime $\mathbb{RN}^{3+1}_{e^2,M}$ is null geodesically incomplete but timelike geodesically complete. On $\mathcal{H}^{\pm}$, $r = M + \sqrt{M^2 - e^2}$ identically, whereas on the other two horizons, $r = M - \sqrt{M^2 - e^2}$.

The spacetime $\mathbb{RN}^{3+1}_{e^2,M}$ is not globally hyperbolic. Let us consider the complete spacelike hypersurface Σ and its *Cauchy development* (the largest globally hyperbolic subset containing it as a Cauchy surface) depicted below

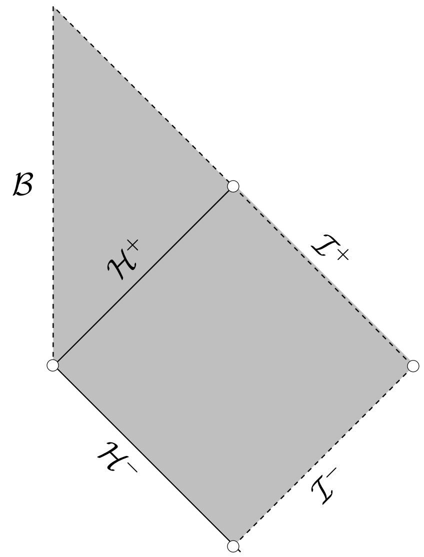


In the context of dynamics of the Einstein equations to be discussed in the next lecture, this is the unique spacetime arising from solving the Cauchy problem with initial data on Σ. In contrast to Schwarzschild, the black (white) hole does not have a singular future (past) boundary $\mathcal{B}^{\pm}$, but rather has a boundary beyond which the spacetime is everywhere smoothly extendible. Such a boundary in an extension is known as a *Cauchy horizon*. We have here denoted this by $\mathcal{C}^{\pm}$.

Thus, in principle, black holes are not necessarily related to singularity, if this term is to be interpreted as "a local geometric quantity blowing up". This is often a source of confusion, so it is important to keep this in mind. More on this when we discuss cosmic censorship in the next lectures.

1.7.2 $M > 0$, $e^2 = M^2$

For $M > 0$, $e^2 = M^2$, a Penrose diagram for a fundamental domain for $\mathbb{RN}^{3+1}_{e^2,M}$ is given below



This is the prototypical example of what is known as an *extremal black hole*. More on this later.

1.7.3 $e^2 > M^2$ or $M < 0$

For $e^2 > M^2$ or $M < 0$, a Penrose diagram for $\mathbb{RN}^{3+1}_{e^2,M}$ is as for $\mathbb{SCH}^{3+1}_M$ with $M < 0$.

1.8 Schwarzschild-de Sitter

The tentative definitions of black and white holes are made with asymptotically flat spacetimes in mind. As they refer only to the infinity defined by the function r, they can be used (and abused) in much more wide settings, in particular, in what are known as *cosmological* spacetimes. Because the term is reasonable in the setting of what is known as the Schwarzschild-de Sitter family, and we will discuss the analysis of the wave equation on these spacetimes, we might as well apply them here as well.

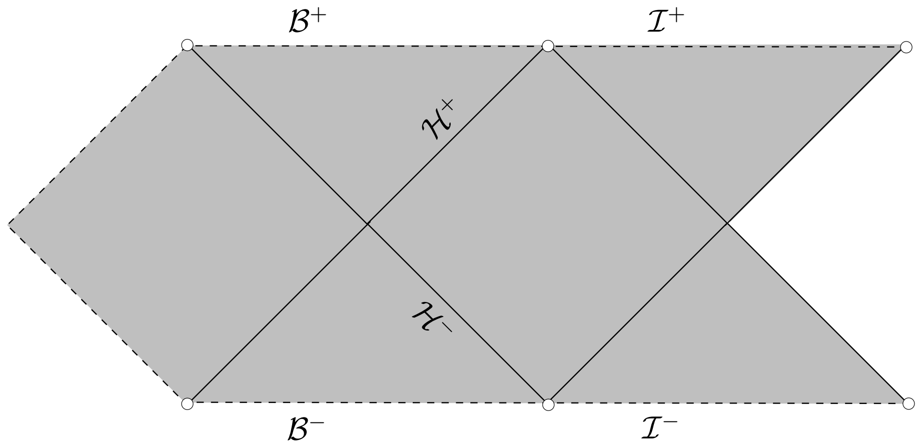
Fix $\Lambda > 0$. The Schwarzschild-de Sitter family is a one-parameter family of spherically symmetric spacetimes $\mathbb{SDS}^{3+1}_{M,\Lambda}$ solving

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

(The above system is equivalent to (1) with $T_{\mu\nu} = 0$.) The parameter M is again called the *mass*. Λ is called the *cosmological constant*. Let us only consider here the cases $0 < 3M\sqrt{\Lambda} \leq 1$.

1.8.1 $0 < 3M\sqrt{\Lambda} < 1$

The spacetime $\mathbb{SDS}^{3+1}_{M,\Lambda}$ admits a discrete isometry. A Penrose diagram for a fundamental domain of the quotient is given below:

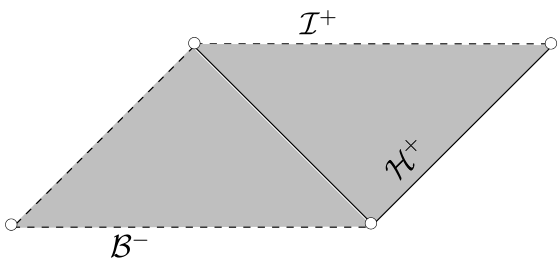


The spacetime $\mathbb{SDS}^{3+1}_{M,\Lambda}$ has a complete Cauchy surface with topology $\mathbb{R}\times\mathbb{S}^2$. It is thus globally hyperbolic with topology $\mathbb{R}\times\mathbb{R}\times\mathbb{S}^2$. By suitable identifications one can construct spatially compact globally hyperbolic quotients with topology $\mathbb{R}\times\mathbb{S}^2$.

On $\mathcal{H}^+$, $\mathcal{H}^-$, r takes on the constant smallest postive root of $1 - 2M/r - \frac{1}{3}\Lambda r^2 = 0$. On the "horizon" defined as the boundary of $J^+(\mathcal{B}^+)$, it takes on the larger positive root of the same expression. The latter "horizon" is called the "cosmological horizon".

1.8.2 $3M\sqrt{\Lambda} = 1$

The extremal case again admits a discrete isometry. The Penrose diagram of a fundamental domain is depicted as before.



The spacetime is again globally hyperbolic with topology $\mathbb{R}\times\mathbb{R}\times\mathbb{S}^2$. Globally hyperbolic spatially compact quotients can be constructed as before.

1.9 Exercises

Exercise 1.5. *Let us suppose that $(\mathcal{M}, g)$ is spherically symmetric and that $\mathcal{Q} = \mathcal{M}/SO(3)$ is a 2-dimensional manifold possibly with boundary such that the metric on $\mathcal{M}$ is a warped product of $\mathcal{Q}$ and $\mathbb{S}^2$. Note that $\mathcal{Q}$ can be realised as a totally geodesic submanifold of $\mathcal{M}$. Show that locally, null coordinates (u, v) can be defined on $\mathcal{Q}$ such that the metric of $\mathcal{M}$ takes the form*

$$-\Omega^2 du dv + r^2(u, v) d\sigma_{\mathbb{S}^2},$$

where r is defined as in Definition 1.18.

In these coordinates, show that the Einstein equations

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = 8\pi T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{6}$$

are equivalent to the following system on $\mathcal{Q}$:

$$\partial_u \partial_v r = -\frac{\Omega^2}{4r} - \frac{1}{r}\partial_u r \partial_v r + 4\pi r T_{uv} + \frac{1}{4} r \Omega^2 \Lambda \tag{7}$$

$$\partial_u \partial_v \log \Omega = -4\pi T_{uv} + \frac{\Omega^2}{4r^2} + \frac{1}{r^2}\partial_u r \partial_v r - \pi \Omega^2 g^{AB} T_{AB}, \tag{8}$$

$$\partial_u(\Omega^{-2}\partial_u r) = -4\pi r T_{uu} \Omega^{-2} \tag{9}$$

$$\partial_v(\Omega^{-2}\partial_v r) = -4\pi r T_{vv} \Omega^{-2}. \tag{10}$$

Give a geometric interpretation of the last two equations and derive them from a well known formula of differential geometry without computation.

Exercise 1.6. *Consider the quantity m defined by*

$$m = \frac{r}{2}(1 - g(\nabla r, \nabla r)).$$

The quantity m is called the Hawking mass*. Show that the Einstein equations give*

$$\partial_u m = r^2 \Omega^{-2}(8\pi T_{uv} \partial_u r - 8\pi T_{uu} \partial_v r) - \frac{1}{2} r^2 \Lambda,$$

$$\partial_v m = r^2 \Omega^{-2}(8\pi T_{uv} \partial_v r - 8\pi T_{vv} \partial_u r) - \frac{1}{2} r^2 \Lambda.$$

Use the above and a simple qualitative analysis of (7)–(10) to construct the Penrose diagrams for $\mathbb{SCH}^{3+1}_M$, and $\mathbb{SCH}^{3+1}_{M,\Lambda}$ and derive all properties claimed in this lecture. Don't bother trying to write explictly the metric functions Ω, r as functions of the coordinates.

Exercise 1.7. *Show that the vector field*

$$T = \Omega^{-2}\partial_v r \frac{\partial}{\partial u} - \Omega^{-2}\partial_u r \frac{\partial}{\partial v}$$

is Killing if $T_{\mu\nu} = 0$, i.e. show that $\mathcal{L}_g T = 0$. Show in general that $T_{\mu\nu}(\mathcal{L}_g T)^{\mu\nu} = 0$.

Exercise 1.8. *Prove Birkhoff's theorem: If* $(\mathcal{M}, g)$ *is spherically symmetric and satisfies* (6) *with* $T_{\mu\nu} = 0$, $\Lambda = 0$, *then it is locally isometric to* $\mathbb{SCH}^{3+1}_M$ *for some* M*. Give a suitable global version of this statement. Formulate a version which is true for* $\Lambda > 0$*. Hint: be careful.*

2 Lecture II: The analysis of the Einstein equations

In the previous lecture, we have gotten a taste for global properties of Lorentzian metrics, the mathematical structure that defines the "gravitational field". In this lecture we will be introduced to the equations which constrain the field, the so-called Einstein equations. In general, these equations form part of a larger system of closed equations–so called Einstein-matter systems–for the metric together with matter fields. In the special case of vacuum, the Einstein equations themselves close.

The equations can be naturally seen to be hyperbolic. Solutions are determined from a proper notion of initial data. Thus, just as in Newtonian mechanics, the natural framework is the dynamical framework, that is to say, the initial value problem.

Self-evident as the above may seem, in retrospect, the road to the realisation of the above facts was long and difficult. We will make some comments along the way about misconceptions that arose (and continue to arise...) when this point of view is not sufficiently well understood.

2.1 Formulation of the equations

Definition 2.1. *Let* $(\mathcal{M}, g)$ *be a spacetime, let* $T_{\mu\nu}$ *be a symmetric covariant* 2*-tensor, and let* Λ *be a real constant. We say* $(\mathcal{M}, g)$ *satisfy the* **Einstein equations with energy momentum tensor** $T_{\mu\nu}$ **and cosmological constant** Λ *if*

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi T_{\mu\nu} - \Lambda g_{\mu\nu} \tag{11}$$

where $R_{\mu\nu}$ *denotes the Ricci curvature and* R *the scalar curvature.*

Proposition 2.1.1. *Let* $(\mathcal{M}, g)$ *satisfy the Einstein equations with energy momentum tensor* $T_{\mu\nu}$ *and cosmological constant* Λ*. Then*

$$\nabla^\mu T_{\mu\nu} = 0. \tag{12}$$

Proof. For a general metric, the Bianchi identities give that $R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R$ is divergence free, while $\Lambda g_{\mu\nu}$ is certainly also divergence free. □

The constraint of satisfying the above proposition, together with the Newtonian limit, led Einstein to the discovery of the equations bearing his name, once he had understood the role of what was then called "general covariance".[12] The

[12] See later remarks, however.

constant Λ was not present in the equations formulated in 1915. (The Newtonian limit requires $|\Lambda|$ to be very small compared with the masses and distances relevant in solar-system physics.) The Λ term was added by Einstein a few years later to allow for a static cosmological universe.[13]

Needless to say, any metric "satisfies the Einstein equations" if $T_{\mu\nu}$ is defined to be $8\pi^{-1}(R_{\mu\nu} - (1/2)g_{\mu\nu}R + \Lambda g_{\mu\nu})$. General relativity only acquires content when $T_{\mu\nu}$ is related to matter fields, and matter equations are prescribed so as to close the total system of equations. We shall give various examples in what follows.

2.1.1 The vacuum equations

The simplest example is when one postulates the absence of matter, i.e. vacuum.

Definition 2.2. *Let $(\mathcal{M}, g)$ be a spacetime. We say that $(\mathcal{M}, g)$ satisfies the* **Einstein vacuum equations with cosmological constant Λ** *if*

$$R_{\mu\nu} = \Lambda g_{\mu\nu}. \tag{13}$$

Alternatively, we say that $(\mathcal{M}, g)$ is a **vacuum spacetime with cosmological constant Λ**.

Note that (13) is equivalent to (11) with $T_{\mu\nu} = 0$.

The vacuum equations are important because (a) large regions of spacetimes can be considered vacuum to a good approximation (b) one can use them to understand "issues of principle" and (c) they are easier to analyse than more complicated Einstein-matter systems.

2.1.2 The Einstein-Maxwell system

We have already seen this in our discussion of Reissner-Nordström.

In the Einstein-Maxwell system, "matter" is described by an antisymmetric covariant 2-tensor $F_{\mu\nu}$, a.k.a. a 2-form. The equations for the matter are the celebrated Maxwell equations, given below.

Definition 2.3. *Let $(\mathcal{M}, g)$ be a spacetime and $F_{\mu\nu}$ an antisymmetric covariant 2-tensor. We say that the tensor $F_{\mu\nu}$* **satisfies the Maxwell equations for a source-free electromagnetic field on $(\mathcal{M}, g)$** *if*

$$dF = 0, \qquad \nabla^\mu F_{\mu\nu} = 0 \tag{14}$$

are satisfied. We call $F_{\mu\nu}$ a **source-free electromagnetic field**.

Note that, in contrast to the first, the second equation of (14) depends on the metric.

[13] From the dynamical point of view, such cosmologies are completely unstable.

Definition 2.4. *Let $F_{\mu\nu}$ denote a solution to* (14) *on* $(\mathcal{M}, g)$. *We define the* **energy momentum tensor $T_{\mu\nu}$ associated to $F_{\mu\nu}$** *by the expression*

$$T_{\mu\nu} \doteq \frac{1}{4\pi}\left(F_{\mu\nu}F_{\lambda\rho}g^{\nu\rho} - \frac{1}{4}g_{\mu\nu}F_{\alpha\beta}F_{\gamma\delta}g^{\alpha\gamma}g^{\beta\delta}\right). \tag{15}$$

Two important properties of the energy momentum tensor are given below

Proposition 2.1.2. *Let $F_{\mu\nu}$ denote a solution to* (14) *and let $T_{\mu\nu}$ denote its associated energy-momentum tensor. Then $T_{\mu\nu}$ is divergence free, i.e.*

$$\nabla^\mu T_{\mu\nu} = 0.$$

If v^μ is a causal vector then

$$T_{\mu\nu}v^\mu v^\nu \geq 0.$$

Definition 2.5. *Let $(\mathcal{M}, g)$ be a spacetime and $F_{\mu\nu}$ a covariant antisymmetric 2-tensor defined on $\mathcal{M}$. We say that the triple $(\mathcal{M}, g, F_{\mu\nu})$* **satisfies the Einstein-Maxwell system** *if equations* (11), (14), (15) *hold.*

Spacetimes $(\mathcal{M}, g)$ satisfying the above definition with an associated $F_{\mu\nu}$ are often called **electrovacuum spacetimes**.

The Einstein-Maxwell equations describe the interaction of gravitation and light.

2.1.3 The Einstein-scalar field system

Definition 2.6. *Let $(\mathcal{M}, g)$ be a spacetime and $\phi : \mathcal{M} \to \mathbb{R}$ a real-valued function on $\mathcal{M}$. We say that ϕ* **satisfies the wave equation on** $(\mathcal{M}, g)$ *if*

$$\Box_g \phi \doteq \nabla^\alpha \nabla_\alpha \phi = 0. \tag{16}$$

We call ϕ a **massless scalar field***.*

Definition 2.7. *Let ϕ denote a solution to* (16) *on* $(\mathcal{M}, g)$. *We define the* **energy momentum tensor $T_{\mu\nu}$ associated to ϕ** *by the expression*

$$T_{\mu\nu} \doteq \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi. \tag{17}$$

As in the Maxwell case, two important properties of the energy momentum tensor are given below

Proposition 2.1.3. *Let ϕ denote a solution to* (16) *and let $T_{\mu\nu}$ denote its associated energy-momentum tensor. Then $T_{\mu\nu}$ is divergence free, i.e.*

$$\nabla^\mu T_{\mu\nu} = 0.$$

If v^μ is a causal vector then

$$T_{\mu\nu}v^\mu v^\nu \geq 0.$$

Definition 2.8. *Let $(\mathcal{M}, g)$ be a spacetime and $\phi : \mathcal{M} \to \mathbb{R}$ a real valued function on $\mathcal{M}$. We say that the triple $(\mathcal{M}, g, \phi)$* **satisfies the Einstein-scalar field system** *if equations* (11), (14), (15) *hold.*

Scalar fields provide good model problems.

2.1.4 The Einstein-Vlasov system

This is the simplest matter model of so-called kinetic theory.

Definition 2.9. *Let $\mathcal{M}$ be a spacetime and let $P \subset T\mathcal{M}$ denote the set of future-directed timelike vectors v with $g(v,v) = -1$. We call P the* **mass shell** *of $\mathcal{M}$.*

The mass shell P of course inherits the structure of a manifold as regular as the metric, and the natural projection $\pi : T\mathcal{M} \to \mathcal{M}$ restricts to a projection $\pi : P \to \mathcal{M}$.

Given x^α local coordinates on $\mathcal{M}$, we can define coordinates (x^α, p^α) on $T\mathcal{M}$ by

$$x^\alpha(v) = x^\alpha(\pi(v)),$$
$$v = p^\alpha(v)\frac{\partial}{\partial x^\alpha}.$$

Let us define the vector field X on $T\mathcal{M}$ by

$$X = p^\mu \frac{\partial}{\partial x^\mu} - \Gamma^\mu_{\nu\lambda} p^\nu p^\lambda \frac{\partial}{\partial p^\mu}$$

X generates geodesic flow on $T\mathcal{M}$. X restricts to a vector field on P which lies in TP.

For the Einstein-Vlasov system, matter is described by a nonnegative function $f : P \to \mathbb{R}$. The equations for the matter are the Vlasov equation given by

Definition 2.10. *Let $\mathcal{M}$ be a spacetime and let $f : P \to \mathbb{R}$ be a nonnegative function. We say that f satisfies* **the Vlasov equation** *if $Xf = 0$, i.e.*

$$p^\mu \frac{\partial f}{\partial x^\mu} - \Gamma^\mu_{\nu\lambda} p^\nu p^\lambda \frac{\partial f}{\partial p^\mu} = 0. \tag{18}$$

Definition 2.11. *Let $\mathcal{M}$ be a spacetime and let $f : P \to \mathbb{R}$ satisfy the Vlasov equation. We define the* **energy momentum tensor $T_{\mu\nu}$ associated to** *f by*

$$T_{\mu\nu}(x) = \int_{P_x} f p_\mu p_\nu, \tag{19}$$

where the integration is with respect to the induced volume on the fibre $P_x = \pi^{-1}(x)$ of the mass shell P.

Proposition 2.1.4. *Let $\mathcal{M}$ be a spacetime, f satisfy the Vlasov equation, and $T_{\mu\nu}$ be its associated energy momentum tensor. Then $T_{\mu\nu}$ is divergence free, i.e.*

$$\nabla^\mu T_{\mu\nu} = 0.$$

If v^μ is a causal vector, then

$$T_{\mu\nu} v^\mu v^\nu \geq 0.$$

Finally we have

Definition 2.12. *Let $(\mathcal{M}, g)$ be a spacetime and $f : P \to \mathbb{R}$ a non-negative function. We say that the triple $(\mathcal{M}, g, f)$* **satisfies the Einstein-Vlasov system** *if equations* (11), (18), (19) *hold.*

Note that in contrast to the other examples, the Einstein-Vlasov system is a system of integro-differential equations, not a system of p.d.e.'s. Spacetimes $(\mathcal{M}, g)$ possessing an f such that the triple $(\mathcal{M}, g, f)$ satisfy the Einstein-Vlasov system are often called **collisionless matter spacetimes**.

The Einstein-Vlasov system is the simplest model of self-gravitating diffuse matter and can be used as a naive model for galactic dynamics. It is also often applied in a cosmological setting.

2.2 The initial value problem

2.2.1 The constraint equations

Let Σ be a spacelike hypersurface in $(\mathcal{M}, g)$. By definition, Σ inherits a Riemannian metric from g. On the other hand, we can define the so-called second fundamental form of Σ by

Definition 2.13. *Let $\mathcal{M}$ be a spacetime, Σ a spacelike hypersurface and N the unique future-directed timelike unit*[14] *normal. We define the* **second fundamental form** *of Σ to be the symmetric covariant 2-tensor in $T\Sigma$ defined by*

$$K(u, v) = -g(\nabla_u V, N)$$

where V denotes an arbitrary extension of v to a vector field along Σ, and ∇ here denotes the connection of g.

As in Riemannian geometry, one easily shows that the above indeed defines a tensor on $T\Sigma$, and that it is symmetric.

Proposition 2.2.1. *Let $(\mathcal{M}, g)$ satisfy the Einstein equations with energy momentum tensor $T_{\mu\nu}$ and cosmological constant Λ. Let Σ be a spacelike hypersurface in $\mathcal{M}$, and let $\bar{g}_{ab}$, $\bar{\nabla}$, K_{ab} denote the induced metric, connection, and second fundamental form, respectively, of Σ. Let barred quantities and Latin indices refer to tensors, curvature, etc., on Σ, and let $\Pi^\nu_a(p)$ denote the components of the pullback map $T^*\mathcal{M} \to T^*\Sigma$. It follows that*

$$\bar{R} + (K^a_a)^2 - K^a_b K^b_a = 16\pi\, T_{\mu\nu} n^\mu n^\nu + 2\Lambda \tag{20}$$

$$\nabla_b K^b_a - \nabla_a K^b_b = 16\pi\, \Pi^\nu_a T_{\mu\nu} n^\mu \tag{21}$$

Proof. Derive as in Riemannian geometry the Gauss and Codazzi equations, take traces, and apply (11). □

[14] i.e. $g(N, N) = -1$

2.3 Initial data

2.3.1 The vacuum case

Definition 2.14. *Let Σ be a 3-manifold, $\bar{g}$ a Riemannian metric on Σ, and K a symmetric covariant 2-tensor. We shall call $(\Sigma, \bar{g}, K)$ a* **vacuum initial data set with cosmological constant** Λ *if* (20)–(21) *are satisfied with $T_{\mu\nu} = 0$.*

Note that in this case, equations (20)–(21) refer only to Σ, $\bar{g}$, K.

2.3.2 The case of matter

Let us here provide only the case for the Einstein-scalar field case.

First note that were Σ a spacelike hypersurface in a spacetime $(\mathcal{M}, g)$ satisfying the Einstein-scalar field system with massless scalar field ϕ, and n^μ were the future-directed normal, then setting $\varphi' = n^\mu \partial_\mu \phi$, $\varphi = \phi|_\Sigma$ we have

$$T_{\mu\nu} n^\mu n^\nu = \frac{1}{2}((\varphi')^2 - \bar{\nabla}^a \varphi \bar{\nabla}_a \varphi),$$

$$\Pi^\nu_a T_{\mu\nu} n^\mu = \varphi' \bar{\nabla}_a \varphi,$$

where latin indices and barred quantities refer to Σ and its induced metric and connection.

This motivates the following

Definition 2.15. *Let Σ be a 3-manifold, $\bar{g}$ a Riemannian metric on Σ, K a symmetric covariant 2-tensor, and $\varphi : \Sigma \to \mathbb{R}$, $\varphi' : \Sigma \to \mathbb{R}$ functions. We shall call $(\Sigma, \bar{g}, K)$ an* **Einstein-scalar field initial data set with cosmological constant** Λ *if* (20)–(21) *are satisfied replacing $T_{\mu\nu} n^\mu n^\nu$ with $\frac{1}{2}((\varphi')^2 - \bar{\nabla}^a \varphi \bar{\nabla}_a \varphi)$, and replacing $\Pi^\nu_a T_{\mu\nu} n^\mu$ with $\varphi' \bar{\nabla}_a \varphi$.*

Note again that with the above replacements the equations (20)–(21) do not refer to an ambient spacetime $\mathcal{M}$.

2.3.3 Cosmology

There are two regimes where general relativity is typically applied: the astrophysical and the cosmological.

In cosmology, the object of interest is the "whole universe", that is to say one deals with a spacetime $(\mathcal{M}, g)$ which is meant to represent the history of the entire universe, together with matter fields defined on $\mathcal{M}$ which are meant to represent all matter. Needless to say, entertaining this subject is epistemologically fundamentally different from usual physics. Nonetheless, cosmological considerations were crucial to the development of general relativity from the very beginning.

The dynamical point of view is not unnatural in cosmology when trying to predict the future. But the main goal of cosmology is to explain the past. This is much hairier business: It is not at all clear, for instance, that our past should be stable to perturbation of our present.

In the mathematical relativity literature, it is very common for the study of cosmology to be defined as the dynamics of general compact initial data. Cosmologists on the other hand typically study homogeneous solutions and their formal perturbations. Their solutions are very rarely spatially compact. From this point of view, cosmology can be considered the study of the dynamics of initial data close to homogeneous.

2.3.4 Isolated self-gravitating systems: asymptotic flatness

In astrophysics, the object of interest is the "isolated self-gravitating system". The cosmological constant Λ can be set to 0. Isolated means that one can understand the dynamics of the system without taking into account the "rest of the universe". One can realise this mathematically by replacing the part of the Cauchy surface representing the initial state of the rest of the universe with an asymptotically flat end. One must be careful, however. The asymptotics of this end carry important information about how the system was formed.

Let us refer in this section to a triple $(\Sigma, \bar{g}, K)$ where Σ is a 3-manifold, $\bar{g}$ a Riemannian metric, and K a symmetric two-tensor on Σ as an **initial data set**, even though we have not specified a particular closed system of equations.

Definition 2.16. *An initial data set* $(\Sigma, \bar{g}, K)$ *is* **strongly asymptotically flat with one end** *if there exists a compact set* $\mathcal{K} \subset \Sigma$ *and a coordinate chart on* $\Sigma \setminus \mathcal{K}$ *which is a diffeomorphism to the complement of a ball in* $\mathbb{R}^3$*, and for which*

$$g_{ab} = \left(1 + \frac{2M}{r}\right)\delta_{ab} + o_2(r^{-1}), \qquad k_{ab} = o_1(r^{-2}),$$

where δ_{ab} *denotes the Euclidean metric.*

Definition 2.17. *The quantity* $E = 4\pi M$ *is known as the* **total energy**.

This quantity E was first defined in Weyl's book Raum-Zeit-Materie from a Noetherian point of view. Later relativists learned about it from a paper published 40 years later by authors whose initials ADM have given rise to the name "ADM mass". In the special case of strongly asymptotically flat initial data considered here, a quantity known as linear momentum vanishes. Thus, mass and energy are equivalent. This is not the case for more general asymptotically flat spacetimes.

A celebrated theorem of Schoen-Yau states

Theorem 2.1. *Let* $(\Sigma, \bar{g}, K)$ *be strongly asymptotically flat with one end and satisfy* (20), (21) *with* $\Lambda = 0$*, and where* $T_{\mu\nu}n^\mu n^\nu$, $\Pi_a^\nu T_{\mu\nu}n^\mu$ *are replaced by the scalar* μ *and the tensor* J_a*, respectively, defined on* Σ*, such that moreover* $\mu \geq \sqrt{J^a J_a}$*. Suppose moreover the asymptotics are strengthened by replacing* $o_2(r^{-1})$ *by* $O_4(r^{-2})$ *and* $o_1(r^{-2})$ *by* $O_3(r^{-3})$*. Then* $M \geq 0$ *and* $M = 0$ *iff* Σ *embeds isometrically into* $\mathbb{R}^{3+1}$ *with induced metric* $\bar{g}$ *and second fundamental form* K*.*

One can define the notion of **strongly asymptotically flat with k ends** by assuming that there exists a compact $\mathcal{K}$ such that $\Sigma \setminus \mathcal{K}$ is a disjoint union of k regions possessing a chart as in the above definiition. The Cauchy surface Σ of $\mathbb{SCH}^{3+1}_M$ for $M > 0$ and the partial Cauchy surface Σ depicted for Reissner-Nordström $\mathbb{RN}^{3+1}_{M,e^2}$ for $0 < M^2 < e^2$, $M > 0$, can be chosen to be strongly asymptotically flat with 2-ends. The mass coincides with the parameter M of the solution.

The above theorem applies to this case as well for the parameter M associated to any end. If $M = 0$ for one end, then it follows by the rigidity statement that there is only one end. Note why $\mathbb{SCH}^{3+1}_M$ for $M < 0$ does not provide a counterexample.

The association of "naked singularities" with negative mass Schwarzschild gave the impression that the positive energy theorem protects against naked singularities. This has proven utterly false.

2.4 The maximal development

Theorem 2.2. *Let $(\Sigma, \bar{g}, K)$ denote a smooth vacuum initial data set with cosmological constant Λ. Then there exists a unique spacetime $(\mathcal{M}, g)$ with the following properties.*

1. *$(\mathcal{M}, g)$ satisfies the Einstein vacuum equations with cosmological constant Λ.*
2. *There exists a smooth embedding $i : \Sigma \to \mathcal{M}$ such that $(\mathcal{M}, g)$ is globally hyperbolic with Cauchy surface $i(\Sigma)$, and $\bar{g}$, K are the induced metric and second fundamental forms, respectively.*
3. *If $(\widetilde{\mathcal{M}}, \widetilde{g})$ satisfies (1), (2) with embedding $\tilde{i}$, then there exists an isometric embedding $j : \widetilde{\mathcal{M}} \to \mathcal{M}$ such that j commutes with $\tilde{i}$.*

The spacetime $(\mathcal{M}, g)$ is known as the **maximal development of** $(\Sigma, \bar{g}, K)$. *The spacetime $\mathcal{M} \cap J^+(\Sigma)$ is known as the* **maximal future development** *and $\mathcal{M} \cap J^-(\Sigma)$ the* **maximal past development***.*

More generally, we may define the notion of development of initial data as follows

Definition 2.18. *Let $(\Sigma, \bar{g}, K)$ be as in the statement of the above theorem. We say that a smooth spacetime $(\mathcal{M}, g)$ is a smooth* **development of initial data** *if 1, 2 are satsfied.*

The original local existence and uniqueness theorems were proven in 1952 by Choquet-Bruhat. They can be formulated as follows

Theorem 2.3. *Let $(\Sigma, \bar{g}, K)$ be as in the statement of the above theorem. Then there exists a smooth development $(\mathcal{M}, g)$ of initial data.*

Theorem 2.4. *Let $\mathcal{M}$, $\widetilde{\mathcal{M}}$ be two smooth developments of initial data. Then there exists a third development $\mathcal{M}'$ and isometric embeddings $i : \mathcal{M}' \to \mathcal{M}$, $\tilde{i} : \mathcal{M}' \to \widetilde{\mathcal{M}}$ commuting with j, $\tilde{j}$.*

The proof of Theorem 2.2 is an easy application of Zorn's lemma[15], the above two theorems and simple facts about Lorentzian causality.

Theorem 2.2 is the statement that general relativity is a well-defined predictive theory. It should be considered as central to the formulation of the theory as the Einstein equations themselves. Indeed, much early confusion in the subject arose from the lack of conceptual clarity provided by the above statement.

The avant-guardness of the notion of "sameness" for two solutions of the Einstein equations as exemplified by the uniquness statement in Theorems 2.2 or 2.4 is well illustrated by the following episode in the history of the development of general relativity: Einstein and Marcel Grossman had arrived already by 1912 at the conclusion that the gravitation field was described by an object we now call a Lorentzian metric. But then they made the following observation: For a "generally covariant equation" relating g and $T_{\mu\nu}$, like (11), then if $g_{\mu\nu} : \mathbb{R}^4 \to \mathbb{R}$ is a solution, and if $\phi : \mathbb{R}^4 \to \mathbb{R}^4$ is a diffeomorphism such that $\phi = id$ outside $B_{1/2}(0)$, and if $T_{\mu\nu} = 0$ in $B_{1/2}(0)$, then $\tilde{g}_{\mu\nu} \doteq (\phi^* g)_{\mu\nu} : \mathbb{R}^4 \to \mathbb{R}$ is again a solution. In particular, there did not appear to be a sense in which the solution was "determined".[16] The above argument is known as the "hole argument" in view of the role of the "hole" $B_{1/2}(0)$.

This led Einstein and Marcel Grossman to abandon the notion of general covariance altogether, only for Einstein to return to it a few years later when he realised that solutions which differed by a diffeomorphism should be considered the "same".

2.5 Hyperbolicity and the proof of Theorem 2.2

The Einstein equations can be identified as hyperbolic in a natural sense once the diffeomorphism invariance is accounted for. At the time of their formulation, however, there was no notion of hyperbolicity for non-linear equations in general, or a corpus of corresponding results, like well-posedness, domain of dependence, etc.

2.5.1 Wave coordinates and Einstein's linearisation

There is a manifestation of hyperbolicity easily accessible to the traditional methodology of physics: Namely, suppose one imposed the coordinate condition

$$\Box_g x^\alpha = 0. \tag{22}$$

[15] It is a pity that a theorem has fundamental as Theorem 2.2 requires Zorn's lemma.

[16] Of course, this illustrates a second misunderstanding of Einstein of the theory, namely that $T_{\mu\nu}$ (together with implicit boundary conditions) should determine the metric. To decouple the two misunderstandings, let us substitute the word "determined" by the phrase "determined by suitable data on say $\{-1\} \times \mathbb{R}^3$" where g and $\tilde{g}$ clearly coincide.

Then the Einstein vacuum equations can be written as a system of the form

$$\Box_g g_{\mu\nu} = Q^{\alpha\beta\lambda\rho\sigma\tau}_{\mu\nu}(g)\partial_\alpha g_{\lambda\rho}\partial_\beta g_{\sigma\tau} \tag{23}$$

which upon linearisation around the flat solution $\mathbb{R}^{3+1}$ yield

$$\Box g_{\mu\nu} = 0,$$

which is the linear wave equation on Minkowski space.

Einstein himself was the first to make the above deductions, predicting thus gravitational waves. He later backtracked, believing the deduction to be gauge dependent. In any case, his entirely non-dynamical view point and his continued insistance on fundamentally misconceived Machian points of view ($T_{\mu\nu}$ determines $g_{\mu\nu}$) indicate that he never fully appreciated the mathematical content of the true hyperbolicity of his equations.

2.5.2 Local existence for quasilinear wave equations

The proof of Theorem 2.2 can essentially be derived from general local well posedness results concerning systems of the form (23). Let us discuss this here.

The structure that allows for all this stems from the following remark. Consider the equation

$$\Box_g \psi = F. \tag{24}$$

Let N be a timelike vector field with respect to g, and consider the current

$$J^N_\mu(\psi) = T_{\mu\nu}(\psi)N^\nu$$

where $T_{\mu\nu}$ is defined by (17). If ψ is a solution of (24), then J^N_μ is a **compatible current**, i.e. both J^N_μ and its divergence $K^N = \nabla^\mu J_\mu$ depend only on the 1-jet of ψ, specifically

$$K^N = FN^\nu\partial_\nu\phi + T_{\mu\nu}\nabla^\mu N^\nu.$$

The above comments are meant to be understood for fixed g and F. But one sees immediately that J_μ, K, remain compatible currents if $g = g(\psi, \partial\psi)$, $F = F(\psi, \partial\psi)$.

The integral of $J^N_\mu n^\mu$ over a spacelike hypersurface with normal n^μ controls $\nabla\psi$ in L^2. If the region is small then the integral of K^N can be controlled by ϵ times the supremum of its integral over spacelike leaves. The integrand only depends on ψ and $\nabla\psi$.

The general structure (24) is preserved after commuting the equation with ∂_x derivatives arbitrarily many times. Moreover, after commuting with k partial derivatives, F becomes more and more linear in the highest few derivatives, in the sense that they appear multiplied by derivatives of much lower order. It follows by the Sobolev inequality that the integral of K^N can be controlled ϵ times the integrand of $J^N_\mu n^\mu$.

This yields H^s estimates for the solution for any sufficiently high Sobolev space. Applying the scheme to differences yields local existence for $H^s \times H^{s-1}$ data.

2.5.3 From quasilinear wave equations to the Einstein vacuum equations

Of course, the above applied to the Einstein equations says that given harmonic coordinates on a spacetime, then the Einstein vacuum equations reduce to (23). To apply this to construct a solution, one must at the same time construct harmonic coordinates.

This can be done as follows:

2.6 Penrose's incompleteness theorem

First, a geometric definition

Definition 2.19. *Let $(\mathcal{M}, g)$ be a spacetime and let $S \subset \mathcal{M}$ be a codimension-2 spacelike submanifold. Let L and $\bar{L}$ be future-directed null vector fields along S, such that $g(L, \bar{L}) \neq 0$, $L \perp T_pS$. For $x, y \in T\Sigma$ define*

$$K(x,y) = -g(\nabla_x L, Y), \qquad \bar{K}(x,y) = -g(\nabla_x \bar{L}, Y)$$

where Y is an arbitrary extension of y to a vector field. K and $\bar{K}$ are a covariant 2-tensor on $T\Sigma$ called the **null second fundamental forms with respect to** *L, $\bar{L}$, respectively.*

Definition 2.20. *Let S, L, $\bar{L}$, K, $\bar{K}$ be as above. We say that S is a* **trapped surface** *if*

$$trK < 0, \qquad tr\bar{K} < 0$$

for all $p \in S$.

Example 2.1. *Let $p \in \mathcal{Q}$ where $\mathcal{Q}$ is the Penrose diagram depicted for $\mathbb{SCH}^{3+1}_M$ or for the globally hyperbolic subset depicted of $\mathbb{RN}^{3+1}_{M,e^2}$ with $M > 0$ and $0 \leq e^2 < M^2$, if applicable. Suppose p is in the interior of the black hole region, i.e. $p \in \mathcal{Q} \cap \overline{J^-(\mathcal{I}^+)}$. Then $\pi^{-1}(p)$ is a closed trapped surface.*

To see this, note that one can choose L $\bar{L}$ to be the null coordinate vectors $\frac{\partial}{\partial v}$ and $\frac{\partial}{\partial u}$ corresponding for instance to the null coordinate system induced by the Penrose diagram. The statement follows from the fact that $\frac{\partial r}{\partial v} < 0$, $\frac{\partial}{\partial v} < 0$ in the black hole region.

Theorem 2.5. *Let $(\mathcal{M}, g)$ be a globally hyperbolic spacetime with Cauchy surface Σ, and suppose S is a closed trapped surface. Suppose further that Σ is non-compact and that $R_{\mu\nu}v^\mu v^\nu \geq 0$ for all null vectors v. Then $(\mathcal{M}, g)$ is future-causally geodesically incomplete.*

The assumption $R_{\mu\nu}v^\mu v^\nu \geq 0$ is known as the **null convergence theorem**. If $(\mathcal{M}, g)$ satisfies the Einstein equations with energy momentum tensor $T_{\mu\nu}$ and cosmological constant Λ, then this assumption holds provided that $T_{\mu\nu}v^\mu v^\nu \geq 0$ for all null vectors v. In view of Propositions 2.1.2, 2.1.3, and 2.1.4 we have

Corollary 2.1. *Let $(\mathcal{M}, g)$ be the underlying spacetime of the maximal development of initial data with non-compact Σ for the Einstein-vacuum, Einstein-Maxwell, Einstein-scalar field, or Einstein-Vlasov system. Suppose that $\mathcal{M}$ contains a closed trapped surface. Then $(\mathcal{M}, g)$ is future-causally geodesically incomplete.*

Theorem 2.5 and subsequent generalisations are among the most abused theorems of mathematical physics. These are usually referred to as "singularity" theorems on account of the expectation that at their heart was a local breakdown of the geometry, for instance, curvature blowup. Indeed, in $\mathbb{SCH}^{3+1}_M$ we see that this is the case. On the other hand, considering the globally hyperbolic region of $\mathbb{RN}^{3+1}_{M,e^2}$ we see that it is geodesically incomplete on account of the above theorem, but there exits a smooth extension $\widetilde{\mathcal{M}}$ of $\mathbb{RN}^{3+1}$ such that *every* incomplete causal geodesic in $\mathbb{RN}^{3+1}$ enters $\widetilde{\mathcal{M}}$.

2.7 The cosmic censorship conjectures

The cosmic censorship conjectures are the conjectures which allow us to "live with" the incompleteness deduced in Theorem 2.5.

2.7.1 Strong cosmic censorship

Fix your favourite Einstein-matter system, fix a notion of "generic" initial data, and fix a notion of inextendibility.

The strong cosmic censorship conjecture states that

Conjecture 2.1. *Let $(\mathcal{M}, g)$ denote the maximal development of generic asymptotically flat or compact data. Then $(\mathcal{M}, g)$ is inextendible.*

For instance, one can reasonably conjecture some version of the above for any of the Einstein-matter systems described in these notes, including of course the Einstein vacuum equations.

One possible class of notions of inextendibility are given by the following

Definition 2.21. *Let $(\mathcal{M}, g)$ be an $n+1$-dimensional Lorentzian manifold with C^k metric, for an integer $n \geq 0$. We say that $(\mathcal{M}, g)$ is* C^k**-extendible** *if there exists an $n+1$-dimensional Lorentzian manifold $(\widetilde{\mathcal{M}}, \tilde{g})$, with C^k metric, and an isometric embedding $i : \mathcal{M} \to \widetilde{\mathcal{M}}$ such that $i(\mathcal{M}) \neq \widetilde{\mathcal{M}}$.*

In view of our previous remarks on Schwarzschild, it follows that it is C^k inextendible for any $k \geq 0$. On the other hand, the maximal development of the asymptotically flat Cauchy surface Σ depicted in for Reissner-Nordström is C^k extendible for all $k \geq 0$. Strong cosmic censorship conjectures in particular that this feature of the solution be unstable to generic perturbation of initial data.

2.7.2 Weak cosmic censorship

2.8 Exercises

Exercise 2.1. *Prove Propositions 2.1.2, 2.1.3, and 2.1.4.*

Exercise 2.2. *Let* $(\mathcal{M}, g)$ *be a* 4*-dimensional spacetime and let* $F_{\mu\nu}$ *be a sourceless electromagnetic field on* $\mathcal{M}$*. Suppose* $(\widetilde{\mathcal{M}}, \tilde{g})$ *is a second* 4*-dimensional spacetime, and* $\Psi : \widetilde{\mathcal{M}} \to \mathcal{M}$ *is conformal. Show that* $(\Psi^* F)_{ab}$ *is a sourceless electromagnetic field on* $\widetilde{\mathcal{M}}$*.*

Now let $(\mathcal{M}, g)$*,* $(\widetilde{\mathcal{M}}, \tilde{g})$*,* Ψ*, be as above, but where these are assumed to be* 2*-dimensional Lorentzian manifolds. Let* $\phi : \mathcal{M} \to \mathbb{R}$ *satisfy* $\Box_g \phi = 0$*. Show that* $\Psi^*\phi$ *satisfies* $\Box_{\tilde{g}}(\Psi^*\phi) = 0$ *on* $\widetilde{\mathcal{M}}$*.*

Exercise 2.3. *Write a clean proof of Theorem 2.2 for the vacuum.*

3 Lecture III: The linear wave equation on black hole backgrounds

3.1 Decay on Minkowski space

3.1.1 The Morawetz current J_μ^K

3.1.2 Angular momentum operators Ω_i

3.1.3 Digression: Penrose's conformal compactification

3.2 Schwarzschild: A closer look

3.2.1 Trapping

3.2.2 The redshift effect

3.3 The currents J_μ^K, J_μ^Y, J_μ^X

3.4 The wave equation on Schwarzschild-de Sitter

3.5 The geometry of the Kerr solution

3.6 Open problems, further reading

4 Lecture IV: Non-linear spherically symmetric problems

4.1 The spherically symmetric Einstein equations with matter

4.1.1 The Hawking mass, the trapped set and the regular region

4.1.2 A general extension principle

4.1.3 Spherically symmetric Penrose inequalities

4.2 Spherically symmetric "no hair" theorems

4.3 Decay rates for a self-gravitating scalar field

4.4 Open problems, further reading