ETH *Nachdiplom* lectures:

The geometry and analysis of black hole spacetimes in general relativity

Mihalis Dafermos*

December 18, 2013

Abstract

These notes accompany my Nachdiplom lectures at ETH. The notes are still under construction–any comments welcome!

Contents

*University of Cambridge, Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB United Kingdom; Princeton University, Fine Hall, Washington Road, Princeton, NJ 08544 United States of America

1 Introduction

Black holes–so evocatively named by John Wheeler in 1967–are one of the central theoretical predictions of general relativity.

Despite the mystique that surrounds the term, the notion of black hole is in principle not as terribly complicated as its reputation may suggest. It is a notion with a clean mathematical definition, and the concept is already exhibited by the simplest non-trivial solution of the Einstein vacuum equations (the governing equations of general relativity), the celebrated Schwarzschild solution, discovered already–but barely understood!–in December 1915. These "black hole spacetimes" give rise to many natural mathematical problems in the analysis of partial differential equations (which are in turn closely tied to the physics of black holes). Study of these problems has witnessed a resurgance in recent years, promising to make this subject one of the central directions in geometric analysis in the next decade.

These notes, which accompany a set of *Nachdiplom* lectures given at ETH, aspire to provide a quick introduction to black hole geometries and problems in the analysis of partial differential equations which these give rise to. They are directed mainly towards mathematicians at the advanced undergraduate or beginning graduate stage and are meant to be elementary and reasonably self contained–see Section 1.2 for a discussion of prerequisites–starting from the basics of Lorentzian geometry (Section 2). They are not meant to be a full-blown introduction to general relativity–for references in the latter direction, see Section 1.3.

Before proceeding further, let us begin with an "impressionistic" first glance of our subject.

1.1 A first look at black holes

General relativity geometrises the notion of spacetime–i.e. the collection of all events–as a *Lorentzian manifold* $(\mathcal{M}, g)$, and this geometric structure allows one to discuss whether spacetime event p (the reader can think of this as the "moment" of explosion of a supernova in some "far-away" galaxy) can in principle communicate with spacetime event q (the reader can think of this as him or herself, sitting at the Hermann-Weyl-Zimmer at 10.23 am on September 23, 2013–by his or her (Swiss) watch).

The set of all events that can send signals to q is known as the *causal past* of q, and is denoted $J^-(q)$, thus "p can send a signal to q" is denoted $p \in J^-(q)$. The *black hole* region of spacetime is then the set of spacetime points characterised by the fact that they cannot communicate with (i.e. they are not in the causal past of) the totality of "far-away" observers–typically the latter means *us*, or the devices we create to detect electromagnetic and (soon, hopefully!) gravitational radiation.

Given a suitable idealisation of the latter notion of "far-away" observer–and this requires some sort of asymptotic "boundary" of spacetime, in this context known as *future null infinity* and often denoted $\mathcal{I}^+$–then the black hole region $\mathcal{B}$ is simply by definition the *complement* of the past of this ideal asymptotic boundary $\mathcal{I}^+$ in the spacetime $\mathcal{M}$, in symbols:

$$\mathcal{B} = \mathcal{M} \setminus J^-(\mathcal{I}^+).$$

As we shall see soon enough, the most basic example of a Lorentzian manifold, *Minkowski space*, denoted $\mathbb{R}^{3+1}$, does *not* contain a black hole region: All spacetime events can successfully communicate with the set of "far-away" observers: in symbols,

$$\mathbb{R}^{3+1} = J^-(\mathcal{I}^+), \qquad \text{i.e.} \quad \mathcal{B} = \mathbb{R}^{3+1} \setminus J^-(\mathcal{I}^+) = \emptyset.$$

This spacetime, the "trivial"[1] solution to Einstein's celebrated vacuum equations

$$\mathrm{Ric}(g) = 0 \tag{1}$$

defines *special relativity*. But already the next simplest Lorentzian manifold $(\mathcal{M}, g)$ whose metric g satisfies (1), the celebrated *Schwarzschild spacetime* (discovered in December 1915), does indeed contain a black hole region, and this simple, completely explicit, metric will form the basis for our own route into the topic.

Our primary interest in black holes will concern their influence on the behaviour of *waves*. In the simplest setting, this means we study the covariant wave equation

$$\Box_g \psi = 0 \tag{2}$$

on a Lorentzian manifold $(\mathcal{M}, g)$ containing a black hole region. Study of this can be thought of as the analogue of study of the Laplace-Beltrami operator on Riemannian manifolds. From the physical point of view, we are interested in waves for two reasons:

(i) By their very definition, black hole regions are not directly "seen" by far away observers. Behaviour of waves on black hole backgrounds could yield a signature of the presence of a black hole region. Thus, it is through waves that black holes can be in principle observed.

(ii) Even more fundamentally, it is waves which govern the stability of spacetimes as solutions to the Einstein vacuum equations. For the Einstein vacuum equations (2) can be viewed as a non-linear system of wave equations, which upon linearisation, yields a more complicated analogue of (2). In particular,

[1] The curvature tensor of Minkowski space vanishes identically, so in particular, so does the Ricci tensor, which is simply a suitable trace.

this means that, if solutions of equations such as (2) cannot be shown to be "well behaved", then we should not expect that black holes can occur in the first place!

We will not discuss (i) at all in these notes. In Section 6.4, however, we will discuss the mathematical formulation of the problem of black hole stability (ii), and the relation with the study of equation (2).

1.2 A note on pre-requisites

What do you need to know to read these notes?

Ideally, the reader should be familiar with the basics of differential manifolds (definition of smooth manifolds, tangent space, vector fields, as well as the formal apparatus of Riemannian geometry: connections, curvature, geodesics) and basic functional analysis (Sobolev spaces, etc.).

The more ambitious reader may wish to start with a mere "nodding aquaintance" of some of these fields, in which case, parallel reading of a basic text will be essential as things begin to get more involved.

1.3 Background reading

Classic texts in the subject are Hawking and Ellis's *The large scale structure of space-time* [21], Wald's *General Relativity* [37], and the encyclopedic *Gravitation*, by Misner, Thorne and Wheeler [24].

A concise introduction to general relativity is given in Christodoulou's *Mathematical problems in General Relativity I* [13].

These lectures will borrow material from a previous set of notes, *Lectures on black holes and linear waves*, written jointly with Igor Rodnianski, available on the arxiv and now published as [19].

Let me remark that, in comparison to [19], the present notes are less ambitious in scope as they do not intend to survey large tracts of the field. So, I like to think of them as "Little lecture notes", as opposed to [19], which I may refer to as "Big lecture notes". On the other hand, these notes will contain certain topics absent from [19] and will try–I may have to erase these words after I have finished writing!–to do things in more detail.

2 Basic Lorentzian geometry

Lorentzian manifolds are the basic objects of general relativity. At a formal level, they are very similar to their more familiar Riemannian cousins, thus, the reader familiar with the latter can quickly absorb the basic concepts. A word of warning, however! Lorentzian geometry is much richer[2] than Riemannian and, at the same time, further removed from our "intuition". So it takes time to develop any sort of feel for the true wealth that the concept encompasses.

[2]Riemannian geometry is contained, in the strict sense, in Lorentzian geometry, as all Riemannian manifolds arise as distringuished submanifolds of Lorentzian ones.

2.1 Lorentzian inner product spaces

Just as Riemannian geometry is based on the standard Euclidean inner product (i.e. a positive definite bilinear form) Lorentzian geometry is based again on a bilinear form, only the assumption of positive definitivity is replaced by a different one. As this type of bilinear form may already be unfamiliar, let us begin our discussion here.

Definition 2.1. *Let V be an $n+1$-dimensional vector space. A* Lorentzian inner product *m is a map $V \times V \to \mathbb{R}$ which is bilinear, symmetric ($m(v,w) = m(w,v)$), and non-degenerate ($m(v,w) = 0 \forall w \implies v = 0$) and such that the maximal dimension of any subspace $W \subset V$ such that $g|_W$ is positive definite is n.*

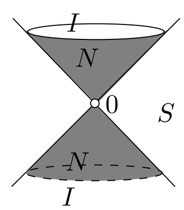
A corollary of the classification theory of bilinear forms then yields that

Proposition 2.1.1. *If V is a vector space and m a Lorentzian inner product, then there exists a basis e_0, e_1, $\ldots$, e_n such that $m(e_i, e_i) = 1$ for $i = 1, \ldots n$, $m(e_0, e_0) = -1$, and $m(e_i, e_j) = 0$ for $i \neq j$.*

We see thus that there are necessarily vectors of negative "length squared", e.g. e_0, but also (if $n \geq 1$!), non-vanishing vectors of 0 "length squared", e.g. $e_0 + e_1$. We define

Definition 2.2. *Let V, m be as above. A vector v is said to be* timelike *if $m(v,v) < 0$, null if $v \neq 0$ and $m(v,v) = 0$, and spacelike otherwise, i.e. if $m(v,v) > 0$ or $v = 0$.*[3] *Timelike and null vectors are collectively known as* causal.

One easily sees that the set of null vectors form a double cone N in V (minus the origin 0).



$N \cup \{0\}$ then partitions the tangent space into three connected components (two inside the cone, corresponding to the set of timelike vectors; let us call this set I) and one (unless $n = 1$, in which case two, or $n = 0$, in which case none) outside (corresponding to the set of non-zero spacelike vectors, let us call this set S).

The cone N is also known as the *lightcone* and null vectors *lightlike*.

More generally, we have

[3]Conventions tend to differ as to whether $v = 0$ should be considered spacelike.

Definition 2.3. *A subspace $W \subset V$ is said to be* spacelike *if the restriction $m|_W$ defines a Euclidean inner product,* timelike *if the restriction $m|_W$ defines again a Lorentzian inner product, and* null *if the restriction defines a degenerate inner product.*

One easily sees that a vector v is timelike, etc., iff the 1-dimensional subspace Span(v) is timelike, etc., in the above definition. (To check this, note what a 1-dimensional Lorentzian inner product space is by our definition.)

Besides the 1-dimensional case of vectors, the co-dimensional-1 case of hyperplanes will be the most important.

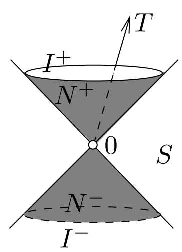
We may pick one of the two components of N, denote it N^+, and call it *the future null cone*. (Let us then denote the other component by N^- and call it the past null cone.)

This partitions then the set of timelike vectors I into two connected components $I^+ \cup I^-$ where I^+ is characterized as the component whose boundary is N^+ . Alternatively we can think that we have picked a timelike vector T, and N^+ is to be by definition the component containing N^+. The "topological" definition is equivalent to

$$N^+ = \{v \in N : g(v,T) < 0\},$$

$$I^+ = \{v \in I : g(v,T) < 0\}.$$

Then, N^- and I^- are similarly characterised by $N^- = \{v \in N : g(v,T) > 0\}$, $I^-\{v \in I : g(v,T) > 0\}$.



We shall call $v \in N^+ \cup I^+$ future directed, and $v \in N^- \cup I^-$ past directed.

2.2 Lorentzian manifolds

We may now pass to the notion of Lorentzian manifold. These stand in exact analogy to Lorentzian inner product spaces as Riemannian manifolds stand to usual (Euclidean) inner product spaces.

Definition 2.4. *Let $\mathcal{M}$ be a smooth manifold. A* Lorentzian metric *on $\mathcal{M}$ is an assignment to each $p \in \mathcal{M}$, of a Lorentzian inner product on $g_p : T_p\mathcal{M} \times T_p\mathcal{M} \to \mathbb{R}$, such that g_p depends smoothly on p.*

To understand abstractly smooth dependence, then one could have defined more formally a Lorentzian metric as a smooth section $g : \mathcal{M} \to T^*\mathcal{M} \otimes T^*\mathcal{M}$ such that for all $p \in \mathcal{M}$, its restriction $g|_p$ determines a Lorentzian inner product.

More pedestrianly, note that a Lorentzian metric (as any tensor) is completely determined by determing its values on a basis. Given local coordinates x^α in a chart $\mathcal{U} \subset \mathcal{M}$, then $\frac{\partial}{\partial x^\alpha}|_p$ form a basis for $T_p\mathcal{M}$ at each $p \in \mathcal{U}$. Defining

$$g_{\alpha\beta}(p) = g_p(\frac{\partial}{\partial x^\alpha}|_p, \frac{\partial}{\partial x^\beta}|_p)$$

then these $(n+2)(n+1)/2$ functions completely determine g.

In particular, since $\frac{\partial}{\partial x^\alpha}$ are smooth vector fields, then the smoothness of g is equivalent to the smoothness of the $g_{\alpha\beta}(p)$. In fact, more explicitly, we can define the induced basis $dx^\alpha \otimes dx^\beta|_p$ of $T^*\mathcal{M}$, and $dx^\alpha \otimes dx^\beta : p \mapsto dx^\alpha \otimes dx^\beta|_p$ are then smooth sections, i.e. elements of $\Gamma(T^*\mathcal{U} \otimes T^*\mathcal{U})$. We have then

$$g = \sum_{\alpha,\beta} g_{\alpha\beta} dx^\alpha \otimes dx^\beta.$$

Following Einstein, for repeated indices it is customary to omit the sum.

We will call the pair $(\mathcal{M}, g)$ a *Lorentzian manifold*. When g is understood, we will often refer simply to $\mathcal{M}$. Note that in contrast to Riemannian geometry[4], <u>not all manifolds admit a Lorentzian metric</u> (**Exercise!**).

The most simple example of a Lorentzian manifold is that naturally induced on the vector space V^{n+1} by the choice of a Lorentzian inner product. Without loss of generality, we can identify the underlying vector space with $\mathbb{R}^{n+1}$, with canonical coordinates $x^0, \ldots, x^n$. We may now define g_p by identifying the basis $\frac{\partial}{\partial x^0}, \ldots \frac{\partial}{\partial x^n}$ with the basis e_0, e_1, $\ldots e_n$, i.e.

$$g_p(\frac{\partial}{\partial x^0}|_p, \frac{\partial}{\partial x^0}|_p) = -1, \qquad \text{etc.}$$

Note that this trivially defines a smooth metric in the sense we defined it (**Exercise!**). The Lorentzian manifold $(\mathbb{R}^{n+1}, g)$ is $n+1$-dimensional *Minkowski space*. We will often refer to the metric in classical notation as

$$g = -(dx^0)^2 + (dx^1)^2 + \cdots + (dx^n)^2,$$

which can be interpreted as shorthand for the more modern notation

$$g = -dx^0 \otimes dx^0 + dx^1 \otimes dx^1 + \cdots + dx^n \otimes dx^n.$$

Given a Lorentzian manifold, we have by definition a notion of timelike, null, spacelike (and causal) for tangent vectors at any point p. These appelations are immediately inherited by vector fields and curves: We will call a vector field V *timelike*, *null*, *causal*, *spacelike at* p if $V(p)$ is *timelike*, *null*, *causal*, *spacelike*, and simply *timelike*, etc., if it is timelike, etc., for all p. Similarly for "vector fields

[4]Recall how Riemannian metrics can be constructed using a partition of unity and exloiting the convexity of the set of Euclidean inner products in the space of symmetric bilinear forms. The set of Lorentzian inner products does not enjoy this convexity property!

along curves γ".[5] We will call a curve γ itself *timelike*, *null*, *causal*, *spacelike* if its tangent vector field $\dot{\gamma}$ is *timelike*, *null*, *causal*, *spacelike* respectively.

Similarly, for smooth hypersurfaces $\mathcal{X} \subset \mathcal{M}$, we say $\mathcal{X}$ is *spacelike*, *timelike*, *null at p* if $T_p\mathcal{N} \subset T_p\mathcal{M}$ is spacelike, timelike, null respectively.

2.3 Time orientation and causal structure

Let T be a global smooth timelike vector field on $\mathcal{M}$.[6] We say that T defines a *time-orientation*.

At each point p, we can use the vector $T(p)$ to define N_p^+, the future light cone in the tangent space $T_p\mathcal{M}$ by

$$N_p^+ = \{v \in N_p : g(v, T(p)) < 0\}.$$

Here N_p denotes the set of all null vectors in $T_p\mathcal{M}$. Thus, a smooth timelike vector field defines a smooth choice of future light cones N_p^+. Denoting by I_p the set of all timelike vectors in $T_p\mathcal{M}$, we similiarly define

$$I_p^+ = \{v \in I_p : g(v, T(p)) < 0\}$$

Of course, any other $\tilde{T}$ such that $g(\tilde{T}, T) < 0$ will define the same notion of future and past, thus, more correctly, a time-orientation is the equivalence class $[T]$ under the obvious equivalence relation. Show (**Exercise**) that if $\mathcal{M}$ admits a time orientation[7] then there are exactly two equivalence classes defined by the above, while if $\mathcal{M}$ does not admit a time orientation, then there is double cover $\pi : \widetilde{\mathcal{M}} \to \mathcal{M}$ such that $(\widetilde{\mathcal{M}}, \pi^* g)$ does.

If T, or better, $[T]$ is as above, then the triple $(\mathcal{M}, g, [T])$ is known as a *time-oriented manifold*. For convenience let us define

Definition 2.5. *A spacetime is a time-oriented $n+1$ dimensional Lorentzian manifold $(\mathcal{M}, g, [T])$.*

We will essentially always deal with spacetimes, but always delete explicit mention of the time orientation, sometimes even dropping g from the notation when it is obvious. The choice of time-orientation will thus be implicit in referring to the sets N_p^+, etc. Also, we will typically consider the case $n = 3$, so "spacetime" will often mean "$3+1$-dimensional spacetime".

So, in what follows below, <u>let $\mathcal{M}$ be a spacetime</u> (initiating our convention of ellipsis!).

Following the terminology of Section 2, we say that causal vectors in $N_p^+ \cup I_p^+$ are *future directed*, while causal vectors in $N_p^- \cup I_p^-$ are *past directed*. As with "timelike", etc., these appellations are inherited now by vector fields, and curves.

[5] Recall that we often consider in differential geometry vector fields which are only defined along a curve or more generally along an immersed submanifold.

[6] Note that by definition such a vector field cannot vanish–thus, there are topological obstructions for the existence of such a T. So we are assuming in particular that $\mathcal{M}$ admits such a vector field! See other footnotes below.

[7] We can now give such Lorentzian manifolds a name: they are called *time-orientable*.

Future-directed causal curves are particularly important, because, in the physical interpretation, physical point-particles (in an approximation in which these make sense) are constrained to travel on future-directed causal curves in spacetime.

We define the *causal future* of p by

$$J^+(p) = \{p\} \cup \{q : \exists\gamma[0,1] \to \mathcal{M}, \dot{\gamma}(s) \in N^+_{\gamma(s)} \cup I^+_{\gamma(s)}, \forall s \in [0,1]\}.$$

That is to say, the causal future of p consists of all points q which lie on future directed causal curves emanating from p. (Note that there is no loss of generality in assuming $q = \gamma(1)$ (**Exercise**).)

We define similarly the *causal past* of p:

$$J^-(p) = \{p\} \cup \{q : \exists\gamma[0,1] \to \mathcal{M}, \dot{\gamma}(s) \in N^-_{\gamma(s)} \cup I^-_{\gamma(s)}, \forall s \in [0,1]\}.$$

Note that $q \in J^+(p) \Leftrightarrow p \in J^-(q)$ (**Exercise**).

More generally now, we may define the causal future of an arbitrary subset $\Sigma \subset \mathcal{M}$ by

$$J^+(\Sigma) = \cup_{p\in\Sigma} J^+(p)$$

similarly J^-.

Example (Causal structure of Minkowski space)**.** *We time orient $\mathbb{R}^{n+1}$ with $\frac{\partial}{\partial x^0}$. Defining the standard Lorentzian inner product on $\mathbb{R}^{n+1}$ (remembering that $\mathbb{R}^{n+1}$ is itself a vector space), then we can define the null cone $N \subset \mathbb{R}^{n+1}$, as well as N^+, I^+, etc. Show (**Exercise**) that $J^+(0) = N^+ \cup I^+$ and more generally, $J^+(p) = p + N^+ \cup I^+$.*

2.4 Length of curves, proper time, isometries, Killing fields

Like in Riemannian geometry, the Lorentzian metric allows one to discuss the length of curves. Typically, one restricts consideration to curves which are everywhere causal or everwhere spacelike.

In the causal case (i.e. for γ with $\dot{\gamma}(s) \in N_{\gamma(s)} \cup I_{\gamma(s)}$), we define

$$L(\gamma) = \int \sqrt{-g(\dot{\gamma},\dot{\gamma})dt},$$

while in the spacelike case, we define

$$L(\gamma) = \int \sqrt{g(\dot{\gamma},\dot{\gamma})dt}.$$

By our conventions, length is always non-negative.

In the case of causal curves, length has an important physical interpretation in general relativity–it is the *proper time* felt by the point-particle traversing the curve γ in spacetime. In particular, if you are in a space-ship following the (future-directed) timelike curve $\gamma : [a,b] \to \mathcal{M}$, then $L(\gamma)$ is the time elapsed (according to your watch–and your biological clock) between events $\gamma(a)$ and $\gamma(b)$.

In contrast, the length of spacelike curves is not directly measurable by physical processes in general relativity.

Let us note explicitly that the length function does not make $\mathcal{M}$ into a metric space–see also the discussion of geodesics in Section 2.5 below.

Like in Riemannian geometry, we can define an isometry of Lorentzian manifold to be a diffeomorphism $\phi : (\mathcal{M}, g) \to (\widetilde{\mathcal{M}}, \widetilde{g})$ such that

$$\phi^* \widetilde{g} = g.$$

(If we don't require that ϕ be a diffeomorphism, but only a diffeomorphism into its image, we call this an *isometric embedding*.)

Isometric embeddings can be characterised by the property that they preserve the lengths of all curves and the angles between two curves. (How does one define the latter, and what is its physical interpretation in the case where the curves are timelike (**Exercise**)?)

As in Riemannian geometry, two Lorentzian manifolds are thought to be abstractly the same if they are isometric, so the actual object of interest is $\{(\mathcal{M}, g)\}/\sim$ where $\sim$ denotes the equivalence relation defined by the notion of isometry.

Two spacetimes $(\mathcal{M}, g, [T])$ and $(\widetilde{\mathcal{M}}, \widetilde{g}, [\widetilde{T}])$ will be equivalent if they are isometric and the isometry preserves the time orientation $[\phi_* T] = [\widetilde{T}]$. Show that this defines an equivalence relation (**Exercise**). Needless, to say, in practice we will always deal with representatives–i.e. spacetimes–and not abstract equivalence classes.

The set of isometries from a spacetime to itself $\phi : (\mathcal{M}, g) \to (\mathcal{M}, g)$ forms a Lie group. The associated Lie algebra can be realised as so-called *Killing fields*.

A Killing field is a vector field V on $\mathcal{M}$ such that $\mathcal{L}_L g = 0$; here $\mathcal{L}$ denotes Lie-differentiation. Given a Killing field V, then the flow ϕ_s generated by V defines a local one-parameter group of diffeomorphisms which are isometries on their domains. Conversely, a local one-parameter group of diffeomorphisms ψ_s which are isometries gives rise to a Killing field W.

Example (Isometries and Killing fields of Minkowski space $\mathbb{R}^{3+1}$)**.** *Let*

$$(x^0, x^1, \ldots, x^n)$$

denote standard coordinates on Minkowski space $\mathbb{R}^{3+1}$, with metric

$$-(dx^0)^2 + (dx^1)^2 + (dx^2)^2 + (dx^3)^2.$$

The generators of Euclidean translation $\frac{\partial}{\partial x^1}$, $\frac{\partial}{\partial x^2}$, $\frac{\partial}{\partial x^3}$ are manifestly Killing, as is the generator of time translation $\frac{\partial}{\partial x^0}$. The generators of Euclidean rotations are $x^i \frac{\partial}{\partial x^j} - x^j \frac{\partial}{\partial x^i}$, $i = 1, \ldots 3$,[8] *and are also Killing (**Exercise**).*

[8]It is a time-honoured tradition that Latin indices range from $1 \ldots n$ and Greek indices from $0 \ldots n$, unless your convention is backwards!

Additional nontrivial Killing fields are given by the celebrated Lorentz boosts*: Given x^i, then one can define a 1-parameter family of isometries:*

$$\phi_s : (x^0, x^1, \dots, x^i, \dots x^n) \to$$
$$((\cosh s)x^0 + (\sinh s)x^i, x^1, \dots, (\sinh s)t + (\cosh s)x^i, \dots x^n)$$

*An easy computation (**Exercise**) shows that ϕ_s are indeed isometries, and they correspond to the Killing fields $x^0\frac{\partial}{\partial x^i} + x^i\frac{\partial}{\partial x^0}$. What is the behaviour of ϕ_s as $s \to \infty$ (**Exercise**)?*

Introducing the notation $x_\mu = g_{\mu\nu}x^\nu$ we may collect our Killing fields

$$\frac{\partial}{\partial x^\mu}, \qquad x_\mu \frac{\partial}{\partial x^\nu} - x_\nu \frac{\partial}{\partial x^\mu}.$$

*These are now-closed under the Lie bracket, generating thus a 10-dimensional Lie algebra (**Exercise**).*

*The group generated by the above Killing fields is known as the Poincaré group. What other isometries are left (**Exercise**)? What if we require also that the time-orientation be preserved (**Exercise**)?*

2.5 Connections, geodesics, curvature

The whole miraculous formal apparatus of Riemannian geometry–allowing one to define a natural connection and relate it with curvature–transfers to the Lorentzian case. Let us review this here very briefly.

Let $(\mathcal{M}, g)$ be a spacetime.

As in Riemannian geometry, there is a unique linear connection ∇ (call it again the *Levi-Civita connection*) in the tangent bundle $T\mathcal{M}$ characterised by the requirements that it (a) preserves the metric

$$\nabla_X g(Y, Z) = g(\nabla_X Y, Z) + g(Y, \nabla_X Z) \tag{3}$$

and (b) is torsion-free

$$\nabla_X Y - \nabla_Y X = [X, Y].$$

This connection is completely determined in local charts by the Christoffel symbols $\Gamma^\alpha_{\beta\gamma}$:

$$\nabla_{\frac{\partial}{\partial x^\alpha}} \frac{\partial}{\partial x^\beta} = \Gamma^\gamma_{\alpha\beta} \frac{\partial}{\partial x^\gamma}$$

defined by the same formula as in Riemannian geometry

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2} g^{\alpha\delta}(\partial_\beta g_{\delta\gamma} + \partial_\gamma g_{\beta\delta} - \partial_\delta g_{\beta\gamma}).$$

Here we are using $g^{\mu\nu}$ to denote the components of the inverse metric, and are applying again the Einstein summation convention.

As in Riemannian geometry, the connection again defines the notion of *parallel transport*: A vector field X along a curve γ is *parallel* if

$$\nabla_{\dot\gamma} X = 0,$$

and to any vector $X(p)$ and curve γ with $\gamma(0) = p$, there is a unique such parallel vector field X along γ with given $X(p)$.

We are using here the fact, as in Riemannian geometry, we may define $\nabla_{\dot\gamma} X$ when X is merely a vector field along γ.

We say a parametrised curve $\gamma : I \to \mathcal{M}$ is a *geodesic* if

$$\nabla_{\dot\gamma}\dot\gamma = 0. \tag{4}$$

Again, given a point $p \in \mathcal{M}$ and a vector $v \in T_p\mathcal{M}$, there is a unique *maximal* geodesic $\gamma : (-T_-, T_+) \to \mathcal{M}$ such that $\gamma(0) = p$, $\dot\gamma|_{\gamma(0)} = v$, where $-\infty \le T_- < 0 < T_+ \le \infty$.

We have the following easy proposition

Proposition 2.5.1. *Let $\gamma(s_0)$ be a geodesic such that $\dot\gamma(s_0)$ is timelike for some $s_0 \in I$. Then $\dot\gamma(s)$ is timelike for all $s \in I$.*

Proof. By (3) and (4):

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

Thus, geodesics preserve their type. □

In view of the above, all geodesics are either timelike, null or spacelike.

General relativity gives (future-directed) timelike and null geodesics and parallel transport a concrete interpretation with operational meaning. In a suitable approximation, small massive bodies in free fall (a spaceship which has turned off its engines, or even the earth and other planets themselves in a suitable approximation) traverse timelike geodesics of spacetime, whereas small massless bodies (photons in the geodesic optics limit) traverse null geodesics. A small gyroscope, on the other hand, in free fall, again in a suitable approximation, will rotate around an axis which can be idealised as a parallelly propagated vector along a timelike geodesic.

Spacelike geodesics, like the length of spacelike curves, do not have direct operational meaning.

We can define the notion of geodesic completeness, just as in Riemannian geometry.

Definition 2.6. *A spacetime $(\mathcal{M}, g)$ is said to be* geodesically complete *if $T_\pm = \infty$ for all $p \in \mathcal{M}$, $v \in T_p\mathcal{M}$.*

Note that whereas in Riemannian geometry, geodesic completeness is related to "metric" completeness, in Lorentzian geometry, the length functional does not naturally induce on $\mathcal{M}$ the structure of a metric space. Indeed, even if the underlying manifold $\mathcal{M}$ is compact[9], the spacetime $(\mathcal{M}, g)$ may be geodesically incomplete.

Whereas in Riemannian geometry, it is often a natural assumption to restrict to manifolds which are geodesically complete, in Lorentzian geometry,

[9] Rarely do we ever assume such a thing in general relativity!

and specifically in applications to general relativity, many spacetimes of interest will not be geodesically complete. That this is true in quite some generality follows from Penrose's celebrated incompleteness theorem, but will in particular be the case for the black hole spacetimes we shall study.

As in Riemannian geometry, geodesics in Lorentzian geometry can also be characterized by their length extremising properties, i.e. as stationary points of the Langragian defined by length (which however is invariant to reparametrisation) or alternatively energy, which singles out the right parametrisation.

But here the situation is more complicated: Timelike geodesics locally *maximise* their length[10], whereas spacelike geodesics are locally saddle points (unless $n = 1$).[11] Note of course that all null curves have 0 length.

Returning to the basic formalism, given the Levi-Civita connection, we can define curvature. Again, all fromulae are as in the Riemannian case: The *Riemann curvature tensor* can be defined as a measure of the infinitessimal failure of parallel transport to commute

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

and in coordinates

$$R^{\delta}_{\alpha\beta\gamma} = \partial_{?} \Gamma^{?}_{??} - \partial\Gamma + \Gamma\Gamma - \Gamma\Gamma.$$

Put the right indices on the above formula (**Exercise**)! Moreover, show the Bianchi identities (**Exercise**):

$$R^{\delta}_{\alpha\beta\gamma} = R^{\cdot}_{\cdots} = R^{\cdot}_{\cdots}$$

$$\nabla R + \nabla R + \nabla R = 0,$$

putting the right indices on the above formulas.

The *Ricci curvature tensor* is defined by the contraction

$$\mathrm{Ric}_{\alpha\beta} = R^{\gamma}_{\alpha\beta\gamma}.$$

As in Riemannian geometry, the Ricci curvature $\mathrm{Ric}(n,n) = \mathrm{Ric}_{\alpha\beta} n^{\alpha} n^{\beta}$ has an interpretation in terms of the second derivative of the area element of a hypersurface normal to n when deformed in the direction of n.

The Ricci tensor plays an important role in general relativity in view of the Einstein equations, which in the vacuum case simply read

$$\mathrm{Ric}(g) = 0. \tag{5}$$

In Section 6.4, we shall turn to the study of (5) as a *p.d.e.*, whose character is essentially hyperbolic, and which can thus be studied from the point of view of the Cauchy problem.

Until then, we will simply study particular Lorentzian manifolds $(\mathcal{M}, g)$ for which one can verify (by explicit computation) that g satisfies (5).

[10] This is the essential origin of the "twin paradox".

[11] If $n = 1$, then one can think of spacelike as timelike with respect to $-g$. We will return to this fact later on!

2.6 Global hyperbolicity and Cauchy hypersurfaces

Lorentzian manifolds, even time oriented ones, can display a variety of "pathological" behaviour. For instance, there can exist *closed timelike curves*–the simplest example arises from the quotient space $\mathbb{R}^{3+1}/\{t \mapsto t+1\}$, with its induced Lorentzian metric.

A restricted class of spacetimes which are free from many of the worst pathologies (but not all!) are provided by the class of "globally hyperbolic" spacetimes. The motivation of the definition of this notion has to do with properties of the Cauchy problem for hyperbolic equations (of which the vacuum equations (5) themeselves are an example). See the discussion below.

Let us remark at the onset that "restricting" to this class of spacetimes cannot be justified on *a prori* grounds. Indeed, one very reasonably can take the point of view that one has to come to terms with all the pathologies that the theory in principle allows. However, restricting to this class would indeed become justified *a postiori* if some very difficult conjectures[12] in the subject turn out to be true. And these conjectures can moreover be studied by restricting to this class.

With this said, let us make the basic definitions.

Definition 2.7. *A smooth curve $\gamma : I \to \mathcal{M}$ is inextendible if for all smooth $\tilde{\gamma} : \tilde{I} \to \mathcal{M}$ such that $I \subset \tilde{I}$ and $\tilde{\gamma}|_I = \gamma$, then $\tilde{I} = I$.*

Definition 2.8. *Let $(\mathcal{M}, g)$ be a spacetime. A* Cauchy surface *is a spacelike hypersurface Σ such that all inextendible causal curves $\gamma : I \to \mathcal{M}$ intersect Σ once and only once.*

Definition 2.9. *A spacetime $\mathcal{M}$ is said to be* globally hyperbolic *if it admits a Cauchy hypersurface Σ.*

At this point it is easy to verify from the basic definitions that $\mathbb{R}^{3+1}$ is globally hyperbolic, with Cauchy surface $t = 0$ for instance (**Exercise**). A slightly more difficult example is $\mathcal{U} \subset \mathbb{R}^{3+1}$ defined by $\mathcal{U} = \mathrm{int}(J^-(1,0,0,0) \cap J^+(-1,0,0,0))$, again with Cauchy surface $\{t = 0\} \cap \mathcal{U}$.

Non-examples are easy to come by. Removing a point p from Minkowski space: $\mathbb{R}^{3+1} \setminus \{p\}$ trivially fails to be globally hyperbolic. Another non-example are the subsets $\{r \leq R\}$ or $\{r \geq R\}$ of Minkowski space for any $R > 0$, where $r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}$.

The notion of global hyperbolicity is originally due to Leray in a much more general context. Definition 2.9 is in fact simply an equivalent characterisation, the equivalence due to Geroch.

The relevance of global hyperbolicity is that it is precisely the property needed to discuss the global initial value problem for hyperbolic equations (like the covariant wave equation) naturally associated to a Lorentzian metric: In particular, the Cauchy problem for the linear hyperbolic equation

$$\Box_g \psi = 0 \tag{6}$$

[12] in particular, the so-called strong cosmic censorship conjecture [21]

is globally well posed on a globally hyperbolic $(\mathcal{M}, g)$ for suitable initial data imposed on a Cauchy hypersurface Σ. We will discuss this in Section **??**.[13]

In this context, one can understand the difference between the examples $\mathcal{R}^{3+1}$ and $\mathcal{U}$ and the non-examples $\{r \leq R\}$ and $\{r \geq R\}$. While in the former two cases, initial data for (6) on $\{t = 0\}$ (here g is of course the Minkowski metric) suffices to uniquely determine a solution whereas, in the latter two cases, Cauchy data on $\{t = 0\}$ must be supplemented with boundary conditions.

Let us note that one consequence of global hyperbolicity is that $\mathcal{M}$ is diffeomorphic to $\mathbb{R} \times \Sigma$.

3 Schwarzschild and Reissner–Nordström

With the fundamentals of Lorentzian geometry presented, we now turn to understanding the black hole concept through its most basic example–Schwarzschild.

The correct understanding of Schwarzschild geometry–simple as it may be in retrospect–took a long time to develop and even longer to be disseminated to the wider physics and astrophysics community: This is because so many new features coincided. Indeed, the black hole concept was hidden by two more obvious features of the metric–a "coordinate singularity" which involved how the metric was originally written down, and a "real singularity".

One should perhaps not fault the early pioneers for confusion. The abstract notion of differentiable manifold (let alone Lorentzian manifold) had not yet been formulated–had it been, it would have helped a lot to clarify things. Indeed, what is more remarkable is that the correct notions were in fact understood relatively early by some–in particular, already in 1932 by Lemaitre.

We shall not have to repeat all this confusion here–although these notes will pay some hommage to (a rational reconstruction of the) historical sequence in their first look at the Schwarzschild solution. Indeed, the goal of this chapter will be to introduce the Schwarzschild geometry as quickly as possible, and from this point of view, the "original" (r, t) coordinate system is the most intuitive to the uninitiated, and has the advantage that the metric is completely explicit and takes a simple form in which one can do easily computations. We will thus start with these, then study *Lemaitre's extension*, which is again easily defined by a simple explicit coordinate transformation. With this, we will give a first discussion of the black hole property.

To claim to be a true relativist, one must however become comfortable with *null coordinates*, so we will then understand Lemaitre's extension from this point of view. This will help us find the maximal extension, but more importantly, finally obtain a concrete representation of null infinity $\mathcal{I}^+$.

[13] Moreover, in Section 6.4, we will see that Cauchy data for the Einstein equations (5) give rise to a maximal globally hyperbolic development spacetime $(\mathcal{M}, g)$. Of course in this latter case, the spacetime is not given a prori, one of the major difficulties of the problem, so it is only a posteriori that the initial data manifold is a Cauchy hypersurface in $\mathcal{M}$.

3.1 The "original" Schwarzschild metric

In Newtonian theory the simplest interesting gravitational objects are point-masses, and more generally, spherically symmetric stars. The gravitational field (defined by the Newtonian potential) can be understood as a solution to $\Delta\psi = \delta$, while in the latter caser, the star has a mass density μ and the gravitational field is given by $\Delta\psi = \mu$. A result (basically of Newton, transported into the Poisson formulation) says that in the vacuum region outside the support of μ, then–in both cases–$\psi = -M/r$ where, modulo units, $M = \int \delta$ or $M = \int \mu$, respectively.

It was Schwarzschild who set out to understand the analogue of both point-masses and spherically symmetric stars, and the vacuum gravitational field outside of them.

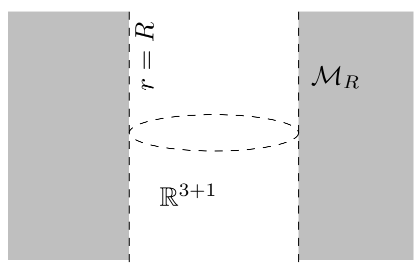
Let (t, x^1, x^2, x^3) denote standard coordinates on $\mathbb{R}^{3+1}$, and let

$$r = \sqrt{(x^1)^2 + (x^2)^2 + (x^3)^2}.$$

Let us fix $R > 0$ and consider the manifold

$$\mathcal{M}_R = \{r > R\} \subset \mathbb{R}^{3+1}$$

depicted here



The restriction of (t, x^1, x^2, x^3) to $\mathcal{M}_R$ defines a global coordinate chart, but in practice we can think of a spherical coordinate chart t, r, θ, ϕ where (θ, ϕ) are determined by

$$x^1 = r\cos\phi\sin\theta$$
$$x^2 = r\sin\phi\sin\theta$$
$$x^3 = r\cos\theta$$

and the coordinates range in

$$(-\infty, \infty) \times (R, \infty) \times [0, \pi] \times (0, 2\pi]$$

as global coordinates, even though they degenerate at $\theta = 0, \pi$.

We can consider on $\mathcal{M}_R$ the family of metrics with parameter $M \in (-\infty, R/2]$, defined by the following expression:

$$-(1 - 2M/r)dt^2 + (1 - 2M/r)^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{7}$$

For each M, let us denote the metric by g_M. Check that these form a smooth one-parameter family of smooth metrics (**Exercise**)!). Moreover, these metrics are Ricci flat:

$$\mathrm{Ric}(g) = 0,$$

i.e. these are non-trivial solutions to the Einstein Vacuum Equations (**Exercise**)[14].

Let us time orient each $(\mathcal{M}_R, g_M)$ with the coordinate vector ∂_t.

We see that for all allowed M, the Killing fields of the Minkowski metric

$$\partial_t,$$

$$\Omega_1 = x^2 \frac{\partial}{\partial x^3} - x^3 \frac{\partial}{\partial x^2},\ \Omega_2 = x^3 \frac{\partial}{\partial x^1} - x^1 \frac{\partial}{\partial x^3},\ \Omega_3 = x^1 \frac{\partial}{\partial x^2} - x^2 \frac{\partial}{\partial x^1} = \frac{\partial}{\partial \phi}$$

remain Killing, but the Lorentz boosts and spatial translations are lost if $M \neq 0$. We refer to ∂_t as the static Killing field and Ω_i as the generators of spherical symmetry. (We say that $(\mathcal{M}_R, g_M)$ is "static" and spherically symmetric.) Note that ∂_t is timelike, while the Ω_i are spacelike.

As in Minkowski space, these Killing fields arise from global diffeomorphisms of $\mathcal{M}_R$. The group $\mathbb{R} \times SO(3)$ acts by isometry on $(\mathcal{M}_R, g_M)$ for all M in the allowed range, and the Lie algebra is spanned by the above Killing vector fields.

3.2 The extension of Lemaitre

Let us now fix $M > 0$ and take $R = 2M$. If we fix the ambient differential structure of $\mathbb{R}^{3+1}$, then we cannot extend the metric to $r = 2M$. It was essentially this fact which was originally mis-interpreted as meaning that $r = 2M$ is physically to be thought of as "singular".

On the other hand, let us institute a coordinate change in $r > 2M$. Define

$$t^* = t + 2M \log(r - 2M).$$

Then the metric can be written

$$-(1 - 2M/r)(dt^*)^2 + (4M/r)drdt^* + (1 + 2M/r)dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2).$$

In this expression, nothing goes wrong as $r \to 2M$.

If the underlying differential structure of $\mathbb{R}^{3+1}$ (which we used to identify $\mathcal{M}_R$) indeed had physical meaning, then the above computation would just be a cute remark of no significance. But a fundamental insight into the theory (which was not easy to understand before the manifold concept was formalised) is that the differential structure is not a priori prescribed. Thus, one can **use the metric to find the right underlying differential structure**.

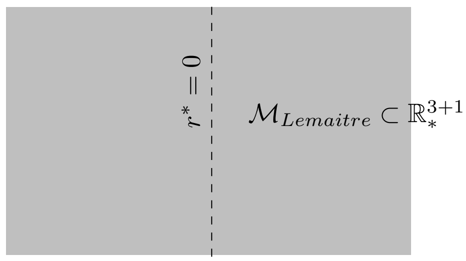
We implement the above below; again, though, as a crutch to aid our pictorial understanding, we will employ yet again an auxilliary ambient Minkowski space. We shall only dispense with this later.

[14] Actually, there is a "right" way to do this exercise–and we'll return to this when we study Birkhoff's theorem–but it is much more useful to do it brute force, if only just once.

To avoid confusion with the previous, let us call our new ambient Minkowski space $\mathbb{R}^{3+1}_*$, with polar coordinates $r^*, t^*, \theta^*, \phi^*$. Now define the open submanifold

$$\mathcal{M}_{Lemaitre} = \{r^* > 0\},$$

see below:



and, for each M, define on $\mathcal{M}_{Lemaitre}$ the metric

$$\hat{g}_M = -(1-2M/r)(dt^*)^2 + (4M/r)dr^*dt^* + (1+2M/r)dr^{*2} + r^{*2}(d\theta^{*2} + \sin^2\theta^* d\phi^{*2})$$

For any fixed M, $R \geq 2M$, our above computation, shows that we have an isometric embedding

$$(\mathcal{M}_R, g_M) \to (\mathcal{M}_{Lemaitre}, \hat{g}_M),$$

$$(r, t, \theta, \phi) \mapsto (r^* = r, t^* = t + 2M\log(r - 2M), \theta^* = \theta, \phi^* = \phi)$$

with range $r^* > R$. In particular, $\hat{g}_M$ again solves the Einstein equations in $r^* \geq 2M$.

Either by direct computation, or applying analytic continuation, it follows that $\hat{g}_M$ solves the Einstein equations in all of $\mathcal{M}_{Lemaitre}$ and moreover the Killing fields of g_M extend to Killing fields of $\hat{g}_M$. Explicitly, these are of course $\frac{\partial}{\partial t^*}$, Ω^*_i. Note that the Ω^*_i are always spacelike.

From now on, let $(\mathcal{M}_{Lemaitre}, \hat{g})$ be our *Schwarzschild solution.*

3.3 The basic causal structure and the black hole property

In what follows we fix M, fix the spacetime $(\mathcal{M}_{Lemaitre}, \hat{g}_M)$, time-oriented by a globally timelike vector field that coincides with $\frac{\partial}{\partial t^*}$ for $r^* > 2M + \epsilon$. (Show without explicit computation that one can find such a time orientation (**Exercise**), in particular that $(\mathcal{M}_{Lemaitre}, \hat{g}_M)$ is time-orientable; we will infer some of its properties below.)

Moreover, let us drop the $\hat{}$, the $*$'s off the coordinates, and the subscript M.

We turn to explore the causal structure of this spacetime.

Let us define the subsets

$$\mathcal{B} = \{r \leq 2M\} \subset \mathcal{M}_{Lemaitre}, \qquad \mathcal{R} = \{r > 2M\} \subset \mathcal{M}_{Lemaitre}$$

Thinking of the coordinate r as a smooth function, we can consider its gradient, which we can compute[15]

$$\nabla r = (1 - 2M/r)\frac{\partial}{\partial r} + (2M/r)\frac{\partial}{\partial t}.$$

In particular, ∇r is future-directed (according to our prescription of the time orientation (**Exercise**)) causal in $\mathcal{B}$, timelike in the interior and null on $\partial\mathcal{B} = \{r = 2M\}$.

In particular, $r = 2M$ is a null hypersurface in $(\mathcal{M}_{Lemaitre}, g_M)$.

Note that on $r = 2M$, $\nabla r = \frac{\partial}{\partial t}$ and ∂_t is thus a Killing field normal to $r = 2M$. This makes $r = 2M$ into something called a *Killing horizon*, on which more later.

Let us also already remark that $\frac{\partial}{\partial t}$ is spacelike in the interior of $\mathcal{B}$; thus, there are no Killing fields timelike there. We will return to the issue when we discuss the wave equation.

Let $p \in \mathcal{B}$ and consider a future directed timelike curve $\gamma(s)$, with $\gamma(0) = p$. We have

$$\dot\gamma r = dr(\dot\gamma) = g(\nabla r, \dot\gamma) < 0.$$

From this it follows that $\gamma(s) \in \mathrm{int}(\mathcal{B})$ for all $s > 0$.

Similarly, let $p \in \mathrm{int}\mathcal{B}$, and consider a future directed causal curve $\gamma(s)$ with $\gamma(0) = p$. We have

$$\dot\gamma r = dr(\dot\gamma) < 0$$

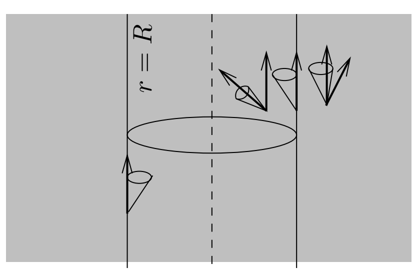
from which it follows that $\gamma(s) \in \mathrm{int}\mathcal{B}$.

It follows by continuity that for any $p \in \mathcal{B}$, and a future directed causal curve $\gamma(s)$ with $\gamma(0) = p$, then $\gamma(s) \in \mathcal{B}$ for $s > 0$.

Moreover, either $\gamma(s_0) \in \mathrm{int}\mathcal{B}$ for $s_0 > 0$ or $\dot\gamma(s)$ is in the direction of ∇r for $s \in [0, s_0]$.

Thus, we have shown that $J^+(\mathcal{B}) \subset \mathcal{B}$.

We draw the light cones:



On the other hand, let $p \in \mathcal{R}$. Consider the integral curve γ of $C\partial_t + \partial_r$ for sufficiently large C (such that $\gamma(0) = p$), i.e. the curve $\gamma(s) = (Cs + t(p), s + r(p), \theta(p), \phi(p))$.

[15] Beware of the fundamental confusion of calculus! Coordinate vector fields depend on the entirety of the coordinates. The gradient, on the other hand, is a Lorentzian-geometric concept.

Note that $\gamma(s) \in \mathcal{R}$, in fact $r(\gamma(s)) \geq r(p)$. For C large enough, we have that $C\partial_t + \partial_r$ is timelike in the region $r \geq r(p)$, thus, with C so chosen, γ is a timelike curve (for nonnegative s), existing for all positive s, and moreover $r(\gamma(s)) \to \infty$ as $s \to \infty$.

We can now understand the black hole property. As stated in the introduction, the black hole should be thought of as the region which cannot send signals to far away observers. We will see later that the latter should be thought of as an ideal boundary "null infinity".

In the absense of a good notion of null infinity so far, let us interpret the past of "far away observers" by the region

$$\bigcap_{R_\infty > 0} J^-(\{r \geq R_\infty\}).$$

We see then that this region is precisely $\mathcal{R}$, and thus

$$\mathcal{B} = \mathcal{M}_{Lemaitre} \setminus \bigcap_{R_\infty > 0} J^-(\{r \geq R_\infty\}).$$

Note that under this definition, Minkowski space has no black hole–a good sign.

Let us introduce the notation

$$\mathcal{H}^+ = \partial\mathcal{B} = \{r = 2M\}.$$

This is known as the *event horizon*, as it delimits the "events" which can communicate to "arbitrarily far away observers".

3.4 Further properties

At this point, one can already study a number of further properties.

Global hyperbolicity: Both $(\mathcal{M}_{2M}, g_M)$ and $(\mathcal{M}_{Lemaitre}, \hat{g}_M)$ are globally hyperbolic. The reader might try to see this already (find a Cauchy surface (**Exercise**)!), but it will be easier after the next section.

The curvature blows up as $r \to 0$: In Riemannian geometry, one could talk about the norm of the Riemannian curvature tensor. But in Lorentzian geometry, that norm is not positive definite. If one is lucky, a scalar invariant, like the Kretschmann scalar, blows up (**Exercise**). This indeed happens in Schwarzschild.[16]

Another easy thing to check (**Exercise**) is that all inextendible future-directed timelike geodesics $\gamma(s)$ which enter $\mathcal{M}$ are incomplete and satisfy $\sup_{s\in I} r(\gamma(s)) = 0$.

There exist some inextendible future-directed timelike geodesics which do not enter $\mathcal{B}$. Such geodesics are in fact future complete (**Exercise**).

Another thing that one can check at this point is that the spacetime is past causally geodesically *incomplete*. Where do these geodesics go (**Exercise**)?

[16] Since parallel transport has direct physical , so does . One can also compute curvature in a parallely propagated frame.

3.5 Null coordinates and Penrose–Carter diagrams

One can think of the Lie group quotient $\mathcal{M}_{\text{Lemaitre}}/SO(3)$ as forming a $1+1$-dimensional Lorentzian manifold

$$\mathcal{Q} = (-\infty, \infty) \times (0, \infty)$$

with metric

$$g_{\mathcal{Q}} = -(1-2M/r)(dt)^2 + (4M/r)drdt + (1+2M/r)dr^2. \tag{8}$$

Note that in Minkowski space the analogue of $\mathcal{Q}$ would be a manifold wtih boundary $(-\infty, \infty) \times [0, \infty)$.

So let us discuss briefly two dimensional Lorentzian metrics: A Lorentzian version of the uniformisation theorem says that locally around any point p, one can find a coordinate chart with range $\mathcal{U} \subset \mathbb{R}^2_{u,v}$ (where $\mathbb{R}^2_{u,v}$) denotes the plane with coordinates u and v), such that the metric takes the form

$$-\Omega^2(u,v)dudv.$$

The significance of this is that

$$-\Omega^2(u,v)dudv$$

is conformally equivalent to the Minkowsi metric

$$-dudv$$

on $\mathbb{R}^{1+1}$, where $u = x^0 - x^1$, $v = x^0 + x^1$. In what follows we will thus identify the $\mathbb{R}^2_{u,v}$ plane with $\mathbb{R}^{1+1}$, and draw $u = c$ and $v = c$ lines at 45 and 135 degrees from the horizontal. We will always choose these coordinates compatible with the usual time orientation on $\mathbb{R}^{1+1}$, i.e. ∂_u and ∂_v are future directed.

Let us note a general fact about conformally related metrics in Lorentzian geometry, defined on the same underlying manifold $\mathcal{M}$. If $\bar{g} = \Xi g$ for some function f, then $v \in T_p$ is timelike, null, spacelike with respect to $\bar{g}$ iff it is timelike, null, spacelike, respectively with respect to g. If we time orient both $\bar{g}$ and g by a timelike vector field T, then we see that

$$J^{\pm}_{g}(p) = J^{\pm}_{\bar{g}}(p)$$

for all $\mathcal{M}$. Thus the causal structure of the two Lorentzian manifolds coincide.

Returning to our example, in the case of $(\mathcal{M}_{Lemaitre}, g)$ and its quotient $\mathcal{Q}$, a little bit of extra work can convince one easily that in fact choose the (u, v) coordinates to be *global* and *bounded*.

(The latter is particularly easy to see, as $\tilde{u} = f(u)$, $\tilde{v} = f(v)$ define a new set of null coordinates.)

Thus, instead of realising the abstract manifold $\mathcal{M}_{Lemaitre}$ as the subset $\mathbb{R}^{3+1} \setminus \{r = 0\}$, then, denoting by $\mathbb{R}^2_{u,v}$ as the (u, v)-plane, we may realise $\mathcal{M}_{Lemaitre}$ as

$$\mathcal{M}_{Lemaitre} = \mathcal{Q} \times S^2$$

where $\mathcal{Q} \subset \mathbb{R}^{1+1}$ is a bounded subset with standard coordinate functions u, v, and the metric takes the form

$$-\Omega^2(u,v)dudv + r^2(u,v)(d\theta^2 + \sin^2\theta\, d\phi^2).$$

We will call $\mathcal{Q}$ a *Penrose–Carter diagramme* of $\mathcal{M}_{Lemaitre}$. We will let $\pi : \mathcal{M}_{Lemaitre} \to \mathcal{Q}$ denote the natural projection, though often we will suppress reference to the map π!

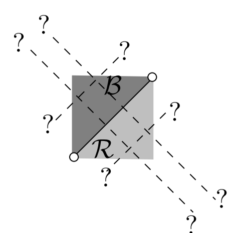
Our goal is to understand the causal structure of $\mathcal{Q}$ as a subset of $\mathbb{R}^{1+1}$, in particular, the causal structure of its boundary in $\mathbb{R}^{1+1}$.

3.6 The Penrose–Carter diagramme of Schwarzschild: first approach

Here, we shall infer the structure of the Lemaitre manifold's Penrose–Carter diagramme $\mathcal{Q}$ from what we already know, namely, from the explicit form of the metric (8).

Easy to see that $\mathcal{Q}$ is a union of "ingoing" null lines crossing $\mathcal{H}^+$, and a union of outgoing null lines. For this, we may use the flow of the Killing vector field ∂_t: Picking an "ingoing null curve"–(one has to convince oneself that there is one crossing the horizon)–then from the Lemaitre-coordinate description it follows the spacetime is covered by the ∂_t flow applied to this, which clearly takes such an ingoing null curve to another. On the other hand, one can show that picking an "outgoing null curve" emanating from a point in $\mathcal{R}$, then applying the flow generated by ∂_t covers $\mathcal{R}$, while, picking another outgoing null curve emanating from a point in $\mathrm{int}(\mathcal{B})$, then applying the flow generated by ∂_t covers $\mathrm{int}(\mathcal{B})$.

That is all one can obtain "for free". We thus have obtained at this point that $\mathcal{Q}$ looks something like this



but we don't know yet "how far" the null curves stretch out with respect to each other, i.e. what is the causal structure of the boundary.

To understand the structure of the boundary, we must write the spherically symmetric Einstein vacuum equations $\mathrm{Ric} = 0$ in u, v coordinates as

$$\partial_u\partial_v \log\Omega^2 = \frac{\Omega^2}{4r^2} + \frac{1}{r^2}\partial_v r \partial_u r \tag{9}$$

$$\partial_u\partial_v r = -\frac{\Omega^2}{4r} - \frac{1}{r}\partial_v r \partial_u r \tag{10}$$

$$\partial_u(\Omega^{-2}\partial_u r) = 0 \tag{11}$$

$$\partial_v(\Omega^{-2}\partial_v r) = 0 \tag{12}$$

Equations (11)–(12) are known as the *Raychaudhuri equations*; think about their geometric interpretation. We will return to this when we talk about Penrose's incompleteness theorem.

It turns out that we can in fact *rewrite* the equations (9)–(10) as follows:

Defining

$$m = \frac{r}{2}(1 - |\nabla r|^2) = \frac{r}{2}(1 + 4\Omega^{-2}\partial_u r \partial_v r) \tag{13}$$

we have in fact

$$\partial_u m = 0 \tag{14}$$

$$\partial_v m = 0 \tag{15}$$

The quantity m is known as the *Hawking mass*. Since on $\mathcal{H}^+$, we have $m = \frac{r}{2} = 2M/2 = M$, it follows that $m = M$ identically. We can now rewrite equations (10) as:

$$\partial_v(\partial_u r) = \frac{m}{8r^2}(-\Omega^{-2}\partial_u r)^{-1} \cdot (\partial_u r) \tag{16}$$

We have written it in this form because $\partial_u r < 0$ everwhere.

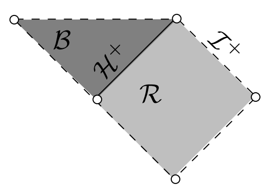
Since by Raychaudhuri (11), the quantity $(-\Omega^{-2}\partial_u r)$ is a known function of v, let us denote it by $F(v)$:

$$F(v) \doteq -\Omega^{-2}\partial_u r$$

Note that $\int F(v)dv$ is independent of the normalisation of the v coordinate.

Let $\mathcal{H}^+$ be $\{u_h\} \times (v_-, v_+)$ with $v_0 \in (v_-, v_+)$. We see immediately that $\int_{v_0}^{v_+} F(v) = \infty$.

We will see in what follows that the Penrose diagramme must look here:



First let us choose a point in $\mathcal{R}$ and an outgoing null geodesic γ. Integrating (16) in a diamond, it follows that r cannot become infinity "before" γ reaches the past null segment in the ambient $\mathbb{R}^{1+1}$ emanating from the endpoint of $\mathcal{H}^+$. From this it follows that the future boundary of $\mathcal{I}^+$ is as depicted.

Now let us turn to the future boundary of $\mathcal{B}$. Since $\partial_u\partial_v r < 0$ and on the other hand $r = 0$ in the limit, it follows that the boundary must be spacelike as depicted.

In particular, we can isolate the points of the boundary of $\mathcal{Q}$ in the ambient $\mathbb{R}^{1+1}$ which are endpoints of future directed null curves for which $r \to \infty$. (Interpret these upstairs).

We can define this as *(future) null infinity* $\mathcal{I}^+$.

With this definition, the black hole region $\mathcal{B}$ has the interpretation of

$$\mathcal{M} \setminus J^-(\mathcal{I}^+)$$

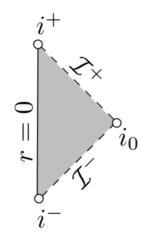
where the causal structure is interpreted in the ambient $\mathbb{R}^{1+1}$ and we really mean $\mathcal{M} \setminus \pi^{-1}(J^-(\mathcal{I}^+) \cap \mathcal{Q})$.

Note the symmetries $t \to -t$ *in the original coordinates* in both $\mathrm{int}(\mathcal{B})$ and $\mathcal{R}$. This tells us the structure of the remaining part of the boundary.

Note that for $\partial_u r$ to become 0 in the limit (it has to!) as one approaches the bottom ingoing boundary, then it follows that $\int_{v_-}^{v_0} F(v)$ is also ∞.

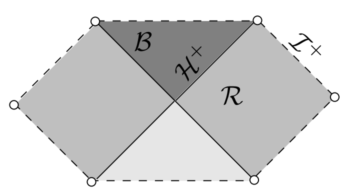
Now, defining $G(u) = (-\Omega^{-2}\partial_v r)^{-1}$ in $\mathcal{R}$, this shows that $\int_{u_-}^{u_0} G(u) = \infty$ and $\int_{u_0}^{u_+} G(u) = \infty$. Interpreted now at future null infinity $\mathcal{I}^+$, this says that future null infinity is *complete*.

At this point, it is useful to compare with Minkowski space. Again, we can proceed without actual computation: We have that $\mathcal{Q}$ is a *manifold with boundary*, where on the boundary, $r = 0$. it follows from (13) that $m = 0$ on this boundary, and thus m is identically 0. It follows that $\partial_v \partial_u r = 0$. From this, it is clear that the Penrose diagramme is as below.



3.7 Maximally extended Schwarzschild and Birkhoff's theorem

The Penrose diagramme of $\mathcal{M}_{Lemaitre}$, together with the symmetries $t \to -t$ *in the original coordinates* in both $\mathrm{int}(\mathcal{B})$ and $\mathcal{R}$ suggests that the spacetime is further extendible, and indeed it can be glued to a copy of itself.



The resulting spacetime, whose underlying abstract manifold we will denote $\mathcal{M}_{max}$, can be realized as

$$\mathcal{M}_{max} = \mathcal{Q}_{max} \times S^2$$

where $\mathcal{Q}_{max}$.

The reader can try to show this on his/her own at this point (**Exercise**).

Note the vector field T.

We can at this point prove a version of "Birkhoff's theorem".

Theorem 3.1. *Let $(\mathcal{M}, g)$ be a spherically symmetric spacetime. Then p is locally isometric to $(\mathcal{M}_{max}, g_M)$ with $M > 0$, or $(\mathbb{R}^4 \setminus \{r = 0\}, g_M)$ defined by (7) with $M \leq 0$.*

Proof. Let $p \in \mathcal{M}$. We can choose a neighbourhood of p with coordinates (u, v, θ, ϕ) ranging in $(-\epsilon, \epsilon) \times (\epsilon, \epsilon)$ so that the metric takes the form , and thus the Einstein equations . We can moreover have chosen the coordinates such that

$$\Omega^2 = 1 \text{ on } u = 0 \qquad \text{and} \qquad \Omega^2 = 1 \text{ on } v = 0. \tag{17}$$

It follows that m is constant $m = M$.

Consider the case $M > 0$. Denote $r(0,0)$ by R. Note that

$$M = \frac{r}{2}(1 + 4\partial_u r(0,0)\partial_v(r(0,)).$$

We have one more gauge freedom. If $\partial_u r(0,0) \neq 0$, we may rescale (u, v) preserving , such that, equal to ± 1. If $\partial_u r(0,0) = 0$, $\partial_v r(0,0) \neq 0$ we may set $\partial_v r \pm 1$. Otherwise, we shall not impose more gauge.

In all the above cases, we see that r is completely determined as a function of v on $u = 0$ and u on $v = 0$. Since Ω^2 is also similarly completely determined, then this completely determines everything by noting that (9)–(10) constitute a well-posed hyperbolic system.

Now we just have to find our data somewhere in the maximally extended Schwarzschild solution.

The case of $M \leq 0$ is left as an exercise. $\square$

3.8 Sketch of a revisionist approach

In the above, we have followed the "historical" approach to the Schwarzschild solution. In retrospect, however, it is clear that it is the above representation with respect to null coordinates which is perhaps the most useful. The "revisionist" approach is then to bypass completely the original coordinates and develop the theory from the beginning in null coordinates.

I.e., we consider from the beginning spacetimes defined on

$$\mathcal{M} = \mathcal{Q} \times S^2, \qquad \mathcal{Q} \subset \mathbb{R}^2$$

with a metric

$$-\Omega^2(u, v)dudv + r^2(u, v)d\sigma,$$

and derive directly the Schwarzschild metric in such coordinates.

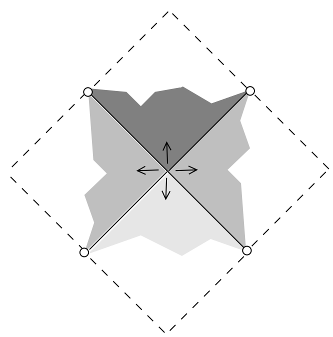
Using this approach one can easily obtain the qualitative features of the Penrose diagramme of maximally extended Schwarzschild without solving explicitly the equations.

Let us sketch this approach briefly.

We will construct a solution by solving an initial value problem for the evolution equations (9)–(9) with initial data on $(-\infty, \infty)\times\{0\} \to \{0\}\times(-\infty, \infty)$ such that the initial data satisfy also (11) and (12).

We shall take specifically initial data which in retrospect will corrspond to the horizons!

I.e, we will take data of the form $r = C$, $\Omega = C$ (Note that (11) and (12)) are trivially solved then). We will solve (9)–(10) in *each* of the four directions (note we are using the fact that the characteristic value problem for thhe $1+1$-dimensional system is well posed in any of the four directions) in the maximal region allowed (here again, we are anticipating the general notion of maximal development).



Lemma 3.8.1. *The equations* (11)–(12) *are propagated by* (9)–(10)*, i.e. the unique solution of* (9)–(10) *with characteristic data satisfying* (11)–(12) *satisfies* (11)–(12) *identically.*

Proof. **(Exercise)** □

It follows that we have constructed a non-trivial solution of the Einstein vacuum equations in a certain region.

What characterises the boundary of this region? This region should be a globally hyperbolic subset of each quadrant with respect to $dudv$ or $-dudv$, depending on the quadrant. It is easy to see from $1+1$-dimensional pde theory that if $p = (u, v)$ is a "first singularity" then

$$\sup_{0\le \mathrm{sgn}(u)\tilde{u}\le|u|, 0\le\mathrm{sgn}(v)\tilde{v}\le|v|} r, \Omega^2, \Omega^{-2}, r^{-1}(\tilde{u}, \tilde{v}) = \infty$$

Let us note that that the equation (10) determines the signs of $\partial_u r$ and $\partial_v r$ on the original axes. From (11)–(12) these are determined everywhere. Thus r has a lower bound $r > 2M$ in the mid-shaded regions and an upper bound $r < 2M$ in the darker and lighter regions.

From the form (16) it is clear that in the darker region, if (u, v) is a first singularity with $\liminf r(u, v) = r_0 > 0$ then one can bound $\log|\partial_u r|$ uniformly in the region $0 \le \tilde{u} \le u$, $0 \le \tilde{v} \le v$ by a constant depending on v and r_0. Then from (11) one infers upper and lower bounds on Ω^2. It follows that any first singularity must have $\liminf r(u,v) = 0$ (and thus, by monotonicity properties of r, $\lim r(u,v) = 0$. Moreover, revisiting again equation (16) gives us that the the whole boundary consists of first singularities and is in fact spacelike. Similar arguments hold for the lighter regions.

In the other two regions, we argue as we did in Section **??**, using again (16), that at a first singularity (u,v), $\log|\partial_u r|$ will be uniformly bounded by a constant in the region $0 \leq -\tilde{u} \leq -u$, $0 \leq \tilde{v} \leq v$, depending only on the values of u and v. It follows that r is bounded from above. On the other hand, it follows from (11) that Ω^2 is bounded from above and below, and now from (12) that $\partial_v r$ is bounded. *It follows that there are no first singularities.*

One sees easily now that $r \to \infty$ on future directed outgoing null rays. For suppose on such a $u = -u_0$ we had $r \leq R$. Then we would have by monotonicity $2M \leq r \leq R$ on $v \geq 0$, $-u_0 \leq u \leq 0$. Integrating (**??**) from $v = v_0$ for some $v_0 > 0$, then using that

$$\int \geq$$

we would obtain $-\partial_u r(u,v) \geq ce^{\int C/v dv}$ which upon inegration $\int_{-u_0}^{0}$ for sufficiently large v would contradict the upper bound on R.

Finally, notice that the existence of a Killing field T follows from the fact that the initial data admit a 1-parameter family of isometries. Write it down explicitly (**Exercise**).

3.9 The sub-extremal Reissner–Nordström family

As an example of the revisionist approach, we will derive the qualitative features of the so-called Reissner-Nordström family. This is a family of solutions to the Einstein–Maxwell system. This equations look of tremendous complexity:

$$\mathrm{Ric}_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi T_{\mu\nu} \tag{18}$$

$$\nabla^\mu F_{\mu\nu} = 0 \qquad dF = 0 \tag{19}$$

$$T_{\mu\nu} = \frac{1}{4\pi}(F_{\cdot\cdot}F^{\cdot}_{\cdot} - \frac{1}{4}g_{\mu\nu}F_{\cdot\cdot}F^{\cdot\cdot}) \tag{20}$$

However, under spherical symmetry, as we shall see, they reduce to something very similar to (9)–(12).

Let us take right away the point of view that spherical symmetry means that as a differentiable manifold

$$\mathcal{M} = \mathcal{Q} \times S^2$$

where $\mathcal{Q}$ is an open set of $\mathbb{R}^{1+1}$, and such that in the induced coordinates (u,v,θ,ϕ) the metric takes the form

$$g = -\Omega^2(u,v)dudv + r^2(u,v)(d\theta^2 + \sin^2\theta d\phi^2)$$

(Note that there is no loss in not allowing $\mathcal{Q}$ to be a manifold with boundary because this would imply finally that $F_{\mu\nu}$ must vanish identically.)

First we note that, simply fixing the metric g to be spherically symmetric in the above sense (i.e. not using the Einstein equations (18)), any solution of (19) must be of the form

$$F = Q_e r^2 \Omega^{-2} du \wedge dv + Q_m \sin\theta d\theta \wedge d\phi \tag{21}$$

where Q_e and Q_m are real constants.

Defining $Q = \sqrt{Q_e^2 + Q_m^2}$, the Einstein equations (19) then take the form

$$\partial_u \partial_v \log \Omega^2 = \frac{\Omega^2}{4r^2} + \frac{1}{r^2}\partial_v r \partial_u r - \frac{\Omega^2 Q^2}{2r^4} \tag{22}$$

$$\partial_u \partial_v r = -\frac{\Omega^2}{4r} - \frac{1}{r}\partial_v r \partial_u r + \frac{1}{4} r^{-3}\Omega^2 Q^2 \tag{23}$$

$$\partial_u(\Omega^{-2}\partial_u r) = 0 \tag{24}$$

$$\partial_v(\Omega^{-2}\partial_v r) = 0 \tag{25}$$

Note that (24) and (25) retain exactly the form of (11) and (12). Why? (**Exercise**).

We can rewrite the equations in terms of the re-normalised Hawking mass

$$\varpi = m + \frac{Q^2}{2r} = \frac{r}{2}(1 + 4\Omega^{-2}\partial_u r \partial_v r) + \frac{Q^2}{2r}$$

which satisfies

$$\partial_u \varpi = 0$$

$$\partial_v \varpi = 0.$$

Note that

$$1 - \frac{2\varpi}{r} + \frac{Q^2}{r^2} = -4\Omega^{-2}\partial_u r \partial_v r$$

When $\partial_u r \neq 0$, we can rewrite (23) as

$$\partial_v(\partial_u r) = \frac{1}{8r^2}(-\Omega^{-2}\partial_u r)^{-1}\left(\varpi - \frac{Q^2}{r}\right)(\partial_u r)$$

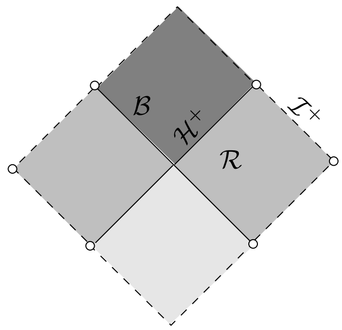
Pick now parameter a $M > Q$. We now define *subextremal Reissner-Nordström with parameters* (M, Q) as the evolution of data $\Omega^2 = 1$ and

$$r = M + \sqrt{M^2 - Q^2} \doteq r_+$$

defined on on $(-\infty, \infty) \times \{0\} \to \{0\} \times (-\infty, \infty)$. Note that we obtain that

$$\varpi = \frac{r_+}{2} + \frac{Q^2}{2r_+} = M$$

One can again obtain without explicitly solving these equations the form of the Penrose diagramme.



We proceed as in the Schwarzschild case:

First again, we can infer the signs of $\partial_u r$ and $\partial_v r$ on the axes. (For this, note that on the axes

$$\varpi - \frac{Q^2}{r} = M - \frac{Q^2}{M + \sqrt{M^2 - Q^2}} > 0$$

and thus $\partial_v \partial_u r < 0$.) By (11) and (12), this gives the signs everywhere. In particular, this tells us that as before r is bounded above in the darker shaded region.

But note that we have from the fact that $\partial_u r < 0$ and $\partial_v r < 0$ the relation

$$1 - \frac{2\varpi}{r} + \frac{Q^2}{r^2} < 0$$

whence the a priori bound

$$r > r_- \doteq M - \sqrt{M^2 - Q^2}.$$

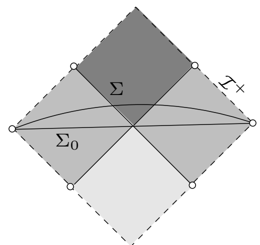
Continuing as before we obtain that there are no first singularities in this region, and thus the boundary must be as depicted. (With a little extra work (**Exercise**) one can show at this stage that r can be extended continuously to $r = r_- \doteq M - \sqrt{M^2 - Q^2}$ at the null dotted lines of the boundary. This will follow from what we will do below...)

We argue completely analogously to the Schwarzschild case that there are again no first singularities in the other regions, and that $r \to \infty$. Thus we may define future null infinity $\mathcal{I}^+$ and past null infinity $\mathcal{I}^-$ as before, and both are complete in the sense described in Section 3.4.

We can characterized the darker shaded region as $\mathcal{Q} \setminus J^-(\mathcal{I}^+)$; it is thus a black hole region.

Finally note that the symmetry of the initial data again implies again the existence of a Killing field T.

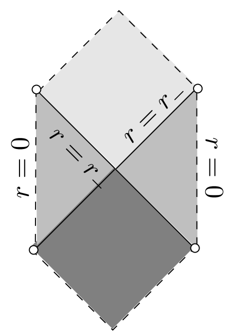
Let us note finally that the resulting spacetime, like maximally-extended Schwarzschild, is globally hyperbolic and admits an asymptotically flat (for those that no what this means) initial data surface Σ with two asymptotically flat ends.



In fact, one can chose a *time-symmetric* one Σ_0 whose second fundamental form vanishes[17].

As we shall see in the remainder of this section, the above spacetime is in fact extendible *beyond the future dotted-line boundary* of the darker shaded region.

For this, let us first define a new spacetime again by solving a characteristic initial value problem for (22)–(25): but this time choosing $r = r_- \doteq M - \sqrt{M^2 - Q^2}$ on two affine complete intersecting null rays. One again obtains that $\varpi = M$. One easily sees (**Exercise**) that the Penrose diagramme looks like:



Hint: first determine the signs of $\partial_u r$ and $\partial_v r$ on the initial data sets and mimick the arguments already used above in the Schwarzschild and Reissner–Nordström cases.

One easily sees (**Exercise**) that the curvature blows up as the boundary $r = 0$ is approached.

Again, a Killing field T extends everywhere.

As perhaps suggested by our shading conventions, the darker shaded region is isometric below is isometric to the the darker shaded region of the previous diagramme, and the two solutions can be pasted together by identifying these regions.

Moreover, the lighter shaded region is isometric to the lighter shaded region (which of course are isometric also to the darker ones) thus, the an infinite chain can be constructed.

[17]Infer the existence of one from the existence of discrete symmetry of our characteristic initial data on the axes (**Exercise**).

We will see this by proving something more general, namely Birkhoff's theorem for the Einstein–Maxwell equations

Theorem 3.2 (Birkhoff's theorem for Einstein–Maxwell, 1st version)**.** *Let*

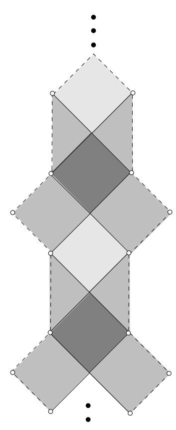
$$(\mathcal{M}, g, F),$$

be a spherically symmetric solution of (18)–(20) *with Maxwell part given by* (21) *with* $Q = Q_e^2 + Q_m^2$ *and with* $\varpi = M$ *for* $M > Q$*. Then for* $p \in \mathcal{M}$*, defining* $R \doteq r(p)$*, then near* p*,* $(\mathcal{M}, g)$ *can be isometrically embedded into a neighbourhood of any point* q *of physical or unphysical extremal Reissner–Nordström with parameters* (M, Q)*, so long as* $r(q) = R$*, and the signs of* $\partial_u r(q)$*,* $\partial_v r(q)$ *coincide with those of* $\partial_u r(p)$ *and* $\partial_v r(p)$*.*

One might wonder already at this point (**Exercise**) what solutions classify solutions of the for $\varpi \leq Q$. We may return to this later!

An immediate corollary of the above is

Corollary 3.1 (Pasted Reissner–Nordström and global version of Birkhoff's theorem)**.** *Physical subextremal Reissner–Nordström can be pasted to unphysical extremal Reissner–Nordström in series as infinite chain with Penrose diagramme:*



Any spherically symmetric solution as above with ambient manifold $\mathcal{M} = \mathcal{Q} \times S^2$ *where* $\mathcal{Q} \subset \mathbb{R}^{1+1}$ *and* u *and* v *are global null coordinates globally isometrically embeds into the above Penrose diagramme.*

Why have we isolated our original Reissner–Nordström as the "physical solution"? The reason is that it is globally hyperblic with an impeccable asymptotically flat Cauchy surface Σ (at least if you don't mind two asymptotically flat ends). In the language of the initial value problem of the Einstein–Maxwell equations, our "physical Reissner–Nordström" corresponds to the *maximal* globally hyperbolic Cauchy development of this Cauchy data. Any extension of this manifold to a larger one will fail to be globally hyperbolic with Cauchy hypersurface Σ–thus will fail to be uniquely determined from data[18].

One might think that the Reissner–Nordström behaviour should be considered preferable to that of Schwarzschild–at least if you are an observer entering the black hole. Because in the former case, all observers living for only finite time still "survive" in the globally hyperbolic domain, as opposed to being crushed by tidal forces, as is the case in Schwarzschild.

On the other hand, from the point of beaurocracy, Schwarzschild has the advantage that the fate of every observer is accounted for, whereas in Reissner–Nordström, the existence of observers who live for finite time but are not destroyed, puts forth the question–*What happens to them?*. In particular, as is demonstrated by pasted Reissner–Nordström, spacetime can be smoothly extended so that all finite-lifespan obeservers in physical Reissner-Nodrdström live for longer. As these extensions are non-unique (even as solutions to the Einstein–Maxwell equations), this gives a concrete realisation to the fact that the fate of such observers is indetermined.

This situation for Reissner–Nordström was deemed philosphically to be so problematic that, in the context of the initial value problem for general relativity, it is conjectured to be *non-generic*:

Conjecture 3.1 (Strong cosmic censorship, Penrose 1972)**.** *For generic asymptotically flat initial data for the Einstein (or Einstein–Maxwell) equations, the maximal globally hyperbolic Cauchy development is inextendible as a suitably regular Lorentzian manifold.*

Further discussion of the correct formulation of inextendibility in the above, as well as the status of the above conjecture, is beyond the scope here.

3.10 Aside: Trapped surfaces and Penrose's incompleteness theorem

Already, it is worth it to put the above two solutions in a more general context.

In both Schwarzschild and Reissner–Nordström, points (u, v) in the darker shaded interior of the black hole region of the quotient $\mathcal{Q}$ satisfy $\partial_u r(u,v) < 0$, $\partial_v r(u,v) < 0$. Upstairs in $\mathcal{M}$, these correspond to 2-spheres S with the property that their two null expansions—i.e. their mean curvatures viewed as hypersurfaces in the two connected components of $\partial J^+(S)$ in a neighbourhood of S, are negative.

This defines the notion of a trapped surface

Definition 3.1. *Let* $(\mathcal{M}, g)$ *be a* 4*-dimensional spacetime. A* trapped surface *is a spacelike two-surface* $S \subset \mathcal{M}$ *such that, if* L *and* $\underline{L}$ *denote two future-directed null vector fields spanning the orthogonal complement to* T_pS*, then the two mean curvatures* $\mathrm{tr}\chi$ *and* $\mathrm{tr}\underline{\chi}$ *are* <u>*both negative*</u>*.*

[18]Wait a minute! Doesn't Birkhoff's theorem say that there is a unique spherically symmetric solution of the Einstein–Maxwell system? Yes, but here, we are not talking about spherically symmetric solutions but general solutions. Spherically symmetric data only imply spherical symmetry of the spacetime in a globally hyperbolic region.

One should compare this with the situation of standard spheres in Minkowski space, or the spheres in the mid-shaded region of Schwarzschild or Reissner–Nordström, for which $\partial_u r < 0$ and $\partial_v r > 0$.

The notion of trapped surface was first isolated by Penrose who proved the following celebrated theorem

Theorem 3.3 (Penrose, 1965)**.** *Let* $(\mathcal{M}, g)$ *be a globally hyperbolic* 4*-dimensional spacetime containing a closed*[19] *trapped surface, with non-compact Cauchy surface* Σ*. Then if* $\mathrm{Ric}(L, L) \geq 0$ *for all null vectors, it follows that* $\mathcal{M}$ *is future causally geodesically incomplete.*

Note that both Schwarzschild and "physical" subextremal Reissner–Nordström (recall that by the latter we mean the globally hyperbolic domain) satisfy the assumptions of the theorem–thus their geodesic incompleteness follows already from the above.

This theorem was very important because it showed that the property of geodesic incompleteness of Schwarzschild is a *stable* one to perturbation of Cauchy data. This is because, in the context of general solutions to the initial value problem for the Einstein vacuum equations (or Einstein–Maxwell equations), then by Cauchy stability, of $(\Sigma, \bar{g}, K)$ are Cauchy data leading to spacetime containing a closed trapped surface, then any spacetime arising from nearby initial data will similarly contain a closed trapped surface. In particular, by the above theorem, the incompleteness property of Schwarzschild and Reissner–Nordström is dynamically stable.

The above theorem is traditionally known as the "singularity theorem". Following Christodoulou, we will refer to it simply as the "incompleteness theorem". For the example of Reissner–Nordstrom shows that the "cause" of the incompleteness can have nothing to do with singularity in the sense of blow up[20]. Indeed, it is only a positive resolution of *strong cosmic censorship* (Conjecture 3.1) which would make Theorem 3.3–generically!–into a true "singularity theorem".

4 The wave equation on general Lorentzian manifolds

We now move to the *analysis* part of this course, namely study of the covariant wave equation (2) on Lorentzian manifolds.

The most basic example is the classical wave equation

$$-\partial_t^2 \psi + \triangle \psi = 0, \tag{26}$$

thought of now geometrically, as the covariant scalar wave equations (2) on Minkowski space $\mathbb{R}^{3+1}$.

[19] i.e. compact without boundary

[20] Of course, additional confusion is caused by the fact that in maximally analytic Reissner–Nordström there is a "singular boundary". But the above theorem has nothing to do with pasted RN!

Our goal in this section will be to prove local estimates and a domain of dependence theorem for (2) in the general setting.

We will see in Section 4.5 how our general theory applies to yield *global* estimates in the special case of (26). This will then serve as a model for the black hole case to be dicussed in Section 5.

4.1 Lagrangian structure and the energy momentum tensor

First some generalities: Let $(\mathcal{M}, g)$ be an $n+1$-dimensional Lorentzian manifold. For scalar functions $\psi : \mathcal{M} \to \mathbb{R}$, we define the covariant wave operator

$$\Box_g \psi \doteq \nabla^\alpha \nabla_\alpha \psi = g^{\mu\nu} \partial_\mu \partial_\nu \psi - g^{\mu\nu} \Gamma^\alpha_{\mu\nu} \partial_\alpha \psi = \frac{1}{\sqrt{-g}} (\partial_\mu (\sqrt{-g} g^{\mu\nu} \partial_\nu \psi))$$

The covariant wave equation

$$\Box_g \psi = 0 \tag{27}$$

on $(\mathcal{M}, g)$ arises as the Euler–Lagrange equations correponding to the Lagrangian

$$\mathcal{L}[\phi] = \int_{\mathcal{M}} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi \, d\mathrm{Vol}_g$$

On account of this, it is possible to define the so-called "energy-momentum tensor"

$$T_{\mu\nu}[\psi] = \partial_\mu \psi \partial_\nu \psi - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} \partial_\alpha \psi \partial_\beta \psi$$

which, for all solutions of (2) satisfies:

$$\nabla^\mu T_{\mu\nu}[\psi] = 0. \tag{28}$$

See [12] for a discussion of general Lagrangian theories.

To obtain an integral law, we must first contract $T_{\mu\nu}$ with a vector field X^μ. Given then an arbitrary X^μ, we obtain a one form

$$J^X_\nu[\psi] = T_{\mu\nu}[\psi] X^\mu$$

From (28) we derive the relation

$$\nabla^\mu J^X_\mu[\psi] = K^X[\psi] \tag{29}$$

where

$$K^X[\psi] \doteq {}^{(X)}\pi_{\mu\nu} T^{\mu\nu}[\psi]. \tag{30}$$

Here,

$${}^{(X)}\pi_{\mu\nu} = \frac{1}{2} X_{(\mu;\nu)} = \frac{1}{2} (\mathcal{L}_X g)_{\mu\nu} \tag{31}$$

is the so-called *deformation tensor* associated to the vector field X.

The relationship (29) between the one-form $J^X[\psi]$ and scalar $K^X[\psi]$ make these *compatible currents* in the terminology of [12], as both of these depend only the 1-jet of ψ, yet for solutions ψ of (27), the divergence identity (29) holds. It is this property, together with the positivity property to be shown in Proposition 4.2.1, on which the entire theory of the wave equation is in some sense based.

4.2 The energy identity

Let $(\mathcal{M}, g)$ now be time-oriented, i.e. a spacetime.

We may now integrate (29) in an arbitrary region $\mathcal{R} \subset \mathcal{M}$ with sufficiently regular boundary $\partial\mathcal{R}$ to obtain

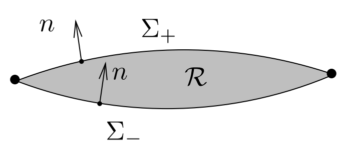
$$\int_{\partial\mathcal{R}} J^X_\mu[\psi] n^\mu_{\partial\Sigma} dVol_{g|_{\partial\mathcal{R}}} = \int_{\mathcal{R}} K^X_\mu[\psi] dVol_g \tag{32}$$

We will note the (perhaps slightly unfamiliar) sign conventions in a minute–as these are extremely important!

If X^μ is Killing, then the deformation tensor (31) vanishes, and thus $K^X[\psi] = 0$.

It follows that J^X_μ is a divergence free one-form and (32) constitutes a *conservation law*.

To examine the signs in identity (32), let us now restrict to the case where where $\mathcal{R}$ is bounded by two homologous compact spacelike hypersurfaces Σ_+ and Σ_- with common boundary.



Let n denote the *future* unit normal of $\Sigma_\pm$, and assume, as depicted, that n points into $\mathcal{R}$ on Σ_-, while $-n$ points into $\mathcal{R}$ on Σ_+.

We can rewrite (32)

$$\int_{\Sigma_+} J^X_\mu[\psi] n^\mu + \int_{\mathcal{R}} K^X_\mu[\psi] = \int_{\Sigma_-} J^X_\mu[\psi] n^\mu.$$

Let us note that we may rewrite the integrand as

$$T_{\mu\nu} X^\mu n^\nu,$$

the volume form always being understood.

We have the following fundamental proposition:

Proposition 4.2.1. *Let $Y, Z \in N^+$. Then*

$$T(Y, Z) \geq 0.$$

If $Y, Z \in I^+$, then, in any local coordinate system

$$T(Y,Z) \geq c \sum_{i=0}^{n} (\partial_i \psi)^2,$$

for a constant c depending on the coordinates and the choice of Y, Z.

One should think of the above positivity as the mathematical realisation of the wave equation's *hyperbolicity*.

An immediate corollary is the following

Corollary 4.1. *Let Σ_+ and Σ_- be homologous spacelike hypersurfaces with common boundary and let X be a timelike Killing field. Then*

$$\int_{\Sigma_+} J^X_\mu[\psi] n^\mu = \int_{\Sigma_-} J^X_\mu[\psi] n^\mu$$

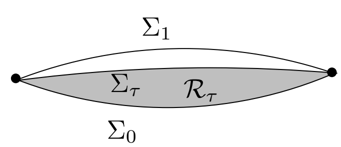
In particular, there exists a C depending only on the region $\mathcal{R}$ such that

$$\|\psi\|_{\mathring{H}^1(\Sigma_+)} + \|n\psi\|_{L^2(\Sigma_+)} \leq C \|\psi\|_{\mathring{H}^1(\Sigma_-)} + \|n\psi\|_{L^2(\Sigma_-)}.$$

The existence of a global timelike Killing field will be important for *global* in "time" properties of the wave equation, as we shall see in the Minkowski case in Section 4.5, and in the black hole case in Section 5. But for *local* properites, like well-posedness, and local-in-"time" estimates, one can make due with any timelike vector field.

To see this, let us specialise the above choice of hypersurfaces. Let Σ_τ, $\tau \in [0,1]$ be a foliation of a region $\mathcal{R}$ as above such that $\Sigma_0 = \Sigma_-$ and $\Sigma_1 = \Sigma_+$, and let X be a timelike vector field, *not necessarily Killing*. Let us moreover assume that the interior of Σ_τ are the level sets of a smooth function $\phi = \tau$ of the interior of $\mathcal{R}$, such that $-g(\nabla\phi, \nabla\phi) \geq c$ for some $c > 0$.[21]

Let us apply the identity (32) in the region $\mathcal{R}_\tau$:



bounded by Σ_0 and Σ_τ, for a general $0 < \tau \leq 1$, rewritten

$$\int_{\Sigma_\tau} J^X_\mu[\psi] n^\mu = \int_{\Sigma_-} J^X_\mu[\psi] n^\mu - \int_{\mathcal{R}_\tau} K^X[\psi]. \tag{33}$$

By Proposition 4.2.1, and compactness of $\mathcal{R}$, we have

$$|K^X[\psi]| \leq C J^X_\mu[\psi] n^\mu_{\Sigma_\tau}$$

[21]Note that an upper bound on $|g(\nabla\phi, \nabla\phi)|$ would be impossible.

for a constant independent of ψ and τ.

On the other hand, let us note that by the co-area formula

$$dVol_{\mathcal{R}} \leq c^{-1} d\tau \, dVol_{\Sigma_\tau},$$

again for a constant independent of τ, and $x \in \Sigma_\tau$.

It follows that we may rewrite

$$\int_{\Sigma_\tau} J^X_\mu[\psi] n^\mu \leq \int_{\Sigma_0} J^X_\mu[\psi] n^\mu + Cc^{-1} \int_0^\tau \int_{\Sigma_{\tau^*}} J^X_\mu[\psi] n^\mu.$$

Defining

$$f(\tau) = \int_{\Sigma_\tau} J^X_\mu[\psi] n^\mu \tag{34}$$

we see that $f(\tau)$ satisfies

$$f(\tau) \leq f(0) + C \int_0^{\tau *} f(\tau^*), \tag{35}$$

whence we obtain

$$f(\tau) \leq f(0) e^{C\tau}. \tag{36}$$

We have,

Theorem 4.1. *Under the above assumptions, there again exists a C depending only on the region $\mathcal{R}$ such that*

$$\|\psi\|_{\mathring{H}^1(\Sigma_+)} + \|n\psi\|_{L^2(\Sigma_+)} \leq C(\|\psi\|_{\mathring{H}^1(\Sigma_-)} + \|n\psi\|_{L^2(\Sigma_-)}). \tag{37}$$

As a corollary we may immediately obtain

Corollary 4.2. *Let ψ, $\tilde{\psi}$ be two sufficiently regular solutions of* (27) *such that $\psi = \tilde{\psi}$, $n^\mu \psi = n^\mu \tilde{\psi}$ on Σ_-. Then $\psi = \tilde{\psi}$ in $\mathcal{R}$.*

In the proof of the above, the reader paying attention will notice that we must use for instance a little Poincare type inequality which says that

$$\|\psi\|_{H^1(\Sigma_\tau)} \leq C(\|\psi\|_{\mathring{H}^1(\Sigma_\tau)} + \|\psi\|_{L^2(\partial\Sigma_-)})$$

Recall here that $\partial\Sigma_- = \partial\Sigma_\tau$ is a smooth compact submanifold of codimension 2.

Noting at the same time the trace theorem

$$\|\psi\|_{L^2(\partial\Sigma_-)} \leq C\|\psi\|_{H^1(\Sigma_-)}$$

We may substitute the homogeneous Sobolev norms in Theorem 4.1 with the inhomogeneous norms, i.e., we have

Theorem 4.2. *Under the above assumptions, there again exists a C depending only on the region $\mathcal{R}$ such that*

$$\|\psi\|_{H^1(\Sigma_+)} + \|n\psi\|_{L^2(\Sigma_+)} \leq C(\|\psi\|_{H^1(\Sigma_-)} + \|n\psi\|_{L^2(\Sigma_-)}).$$

Corollary 4.2 is a statement of local uniqueness. With a little Lorentzian geometry, we can "globalise" the statement as follows

Theorem 4.3 (Domain of dependence property)**.** *Let* $(\mathcal{M}, g)$ *be a spacetime and* Σ *an acausal*[22] *spacelike hypersurface (without boundary), with future directed normal* n. *Let* ψ, $\tilde{\psi}$ *be sufficiently regular solutions of* (27) *such that* $\psi|_\Sigma = \tilde{\psi}|_\Sigma$, $n\psi|_\Sigma = n\tilde{\psi}|_\Sigma$ *on* Σ. *Then* $\psi = \tilde{\psi}$ *on* $D(\Sigma)$.

Here $D(\Sigma) = \cup D^+(\Sigma) \cup D^-(\Sigma)$ denotes the so-called Cauchy development of Σ in $\mathcal{M}$. This is the set of all $p \in \mathcal{M}$ such that every inextendible causal curve through p intersects $\mathcal{M}$. It turns out that $D^\pm(\Sigma)$ is globally hyperbolic with Cauchy hypersurface Σ. If $\mathcal{M}$ was itself globally hyperbolic with Cauchy hypersurface Σ, then $D(\Sigma) = \mathcal{M}$.

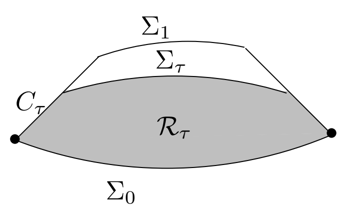
Proof. The proof requires quite a bit of Lorentzian causality theory that we won't get into here. Thus this is really only a sketch. It suffices of course to consider the case where $\tilde{\psi} = 0$ identically. One considers the set

$$\mathcal{S}^\pm = \{p \in D^\pm(\Sigma) : \exists \Sigma_p \text{ spacelike with } , p \in \Sigma_p, \psi|_{\Sigma_p} = 0, n\psi|_{\Sigma_p} = 0\}.$$

The set is non-empty as it contains the original Σ and is moreover open by Corollary 4.2. Let $\mathcal{S}_0^+$ denote the set of all $p \in \mathcal{S}^+$ such that $J^-(p) \subset \mathcal{S}^+$. This set also contains Σ and is open and connected.

Suppose $D^+(\Sigma) \setminus \mathcal{S}_0^+ \neq \emptyset$. Let $p \in D^+(\Sigma) \setminus \mathcal{S}_0^+ \neq \emptyset$ and consider the compact (by theory of global hyperbolicity) subset $J^-(p) \cap D^+(\Sigma)$. There exists a $q \in J^-(p) \cap D^+(\Sigma)$ such that $J^-(q) \cap D^+(\Sigma) \setminus q \subset \mathcal{S}_0^+$ (for instance, let q be a minimizer of the continuous function $x \to \mathrm{Vol} J^-(x)$ on the compact set $J^-(p) \cap D^+(\Sigma)$). Since $\mathcal{S}_0^+$ is open, it follows (from a little bit of Lorentzian geometry) that there is a spacelike Σ_q containing q such that $\Sigma_q \setminus \{q\} \subset \mathcal{S}_0^+$. But then it clearly follows that $q \in \mathcal{S}_0^+$, a contradiction. $\square$

Let us add finally that one often considers the wave equation in a region $\mathcal{R}$ which is foliated as below by Σ_τ, but which has also a smooth null boundary C_τ. Here $\Sigma\tau$ can be taken to be the level sets of a smooth function $\phi = \tau$, such that $0 < c < |\nabla\phi| \leq C$:



The energy estimate in each $\mathcal{R}_\tau$ gives

$$\int_{\Sigma_\tau} J^X_\mu[\psi]n^\mu + \int_{C_\tau} J^X_\mu[\psi]n^\mu_C = \int_{\Sigma_-} J^X_\mu[\psi]n^\mu - \int_{\mathcal{R}_\tau} K^X[\psi]. \tag{38}$$

[22] i.e. a hypersurface such that every timelike curve intersects it at most once

The second term on the left hand side of (38) requires some discussion. Null cones do not come with a natural volume form, neither a natural normal. But given a choice of normal, there is an induced volume form which makes Stokes theorem correct. Now it turns out that the integrand controls all derivatives of ψ normal to n_C, that is to say, all *tangential* derivatives to C_τ. Compare with the integrand of the other flux terms which control *all* derivatives.

In our discussion of black holes, we will see how to make use of this second term which naturally will come up along *event horizons* $\mathcal{H}^+$. For now, let us simply note that we are free to drop it, obtaining the inequality,

$$\int_{\Sigma_\tau} J^X_\mu[\psi]n^\mu \leq \int_{\Sigma_-} J^X_\mu[\psi]n^\mu - \int_{\mathcal{R}_\tau} K^X[\psi], \tag{39}$$

from which, defining f as before by (34), we may reprove (35), and thus (37).

How does one get the 0'th order term in this approach **(Exercise)**?

4.3 The inhomogeneous case and commutation

To consider higher order estimates, we must first enlarge our formalism to consider the inhomogeneous case of (27), namely,

$$\Box_g \psi = F. \tag{40}$$

We have

$$\nabla^\nu T_{\mu\nu}[\psi] = F\psi_\mu$$

whence, defining, for a vector field X as before J^X and K^X, and defining $\mathcal{E}^X = FX^\nu\psi_\nu$, we obtain

$$\nabla^\mu J^X_\mu[\psi] = K^X[\psi] + \mathcal{E}^X[\psi] \tag{41}$$

We will apply this to solutions of (27) commuted by a general vector field Z. Let Z then be a general vector field, and ψ a solution of (27). Then $Z\psi$ satisfies

$$\Box_g Z\psi = [\Box_g, Z]\psi$$

Now $[\Box_g, Z]$ is a second order operator, let us call it P_2. Note in particular that if Z is Killing, then $[\Box_g, Z] = 0$, i.e. $P_2 = 0$. In general, we have

$$\mathcal{E}^X[Z\psi] = P_2\psi X(Z\psi)$$

Now let X again be timelike in a region $\mathcal{R}$ as in before (33) or as in before (38). From (41), we obtain upon integration

$$\int_{\Sigma_\tau} J^X_\mu[Z\psi]n^\mu \leq \int_{\Sigma_0} J^X_\mu[Z\psi]n^\mu - \int_{\mathcal{R}_\tau} (K^X[Z\psi] + \mathcal{E}^X[Z\psi]), \tag{42}$$

Now let X be a timelike multiplier, and let Z_α denote a collection of vector fields which span $T_p\mathcal{M}$, for all p.

Note that with respect to any local coordinate system around p,

$$\sum_{\mu,\nu=0}^{n} |\partial^2_{\mu\nu}\psi|^2 \leq C(p) \sum_{\alpha=0}^{n} J^X_\mu[Z_\alpha\psi]n^\mu$$

In particular, by compactness we have

$$\sum_\alpha K^X[Z_\alpha\psi] + \mathcal{E}^X[Z_\alpha\psi] \leq C(\sum_\alpha J^X_\mu[Z_\alpha\psi]n^\mu + J^X_\mu[\psi]n^\mu)$$

in $\mathcal{R}$, where the presence of the second term on the right hand side is necessitated by the fact that P_2 has a first order part.

In view again of the co-area formula, then defining

$$f(x) = \int_{\Sigma_\tau} J^X_\mu[\psi]n^\mu + \sum_{\alpha=0}^{n} \int_{\Sigma_\tau} J^X_\mu[Z_\alpha\psi]n^\mu$$

We have that f again satisfies (35) whence again (36).

This gives H^2 bounds:

$$\|\psi\|_{H^2(\Sigma_\tau)} + \|n\psi\|_{H^1(\Sigma_\tau)} \leq C(\|\psi\|_{H^2(\Sigma_0)} + \|n\psi\|_{H^1(\Sigma_0)}) \tag{43}$$

(where, to include the 0'th order term we are using that we have already bounded it). Note that we have secretely thrown away the term

$$\|n^2\psi\|_{L^2(\Sigma_0)}$$

from the right hand side (which a priori is necessary) to relate the right hand side with $f(0)$, because the wave equation itself allows us to solve for $n^2\psi$. We shall see this again in the context of Minkowski space. (For the same reason, we have not bothered to include $|n^2\psi|_{L^2(\Sigma_\tau)}$ on the left hand side.)

We see at this point one disadvantage with the first region we considered (foliated by homologous spacelike hypersurfaces) as opposed to the region before (38). In the latter case, we may chose $n-1$ smooth vector fields Z_i on $\mathcal{M}$ such that Z_i are tangential to the leaves Σ_τ, and define f only with respect to these vector fields giving directly (43).[23]

Now show by an analogous commutation formula that one obtains the general H^k estimate **(Exercise)**.

Proposition 4.3.1. *For all $k \geq 1$, $\mathcal{R}$ as before, there exists a C_k such that*

$$\|\psi\|_{H^k(\Sigma_\tau)} + \|n\psi\|_{H^{k-1}(\Sigma_\tau)} \leq C_k(\|\psi\|_{H^k(\Sigma_0)} + \|n\psi\|_{H^{k-1}(\Sigma_0)}).$$

Note the two distinct uses of vector fields: in the role of *multipliers*, like X, and in the role of *commutation fields*, like Z.

[23] If one were to try to restrict to tangenetial vector fields in the former case, one would have to ensure that the blow up of their deformation tensor is canceled by that of the lapse $|\nabla\phi|$.

4.4 The Sobolev inequality and pointwise bounds

Let us recall the following local Sobolev inequality on the compact n-manifold with boundary Σ_τ: For all $s > n/2$, we have

$$\sup_{x\in\Sigma_\tau} |\psi| \leq C_s(\Sigma_\tau)\|\psi\|_{H^s(\Sigma_\tau)}$$

Thus, putting together the above with we obtain

$$\sup_{x\in\Sigma_\tau} |\psi| \leq C(\tau)\|\psi\|_{H^{[n/2]+1}(\Sigma_0)}$$

and more generally

$$\|\psi\|_{C^m(\Sigma_\tau)} \leq C(\tau)\|\psi\|_{H^{[n/2]+1+m}(\Sigma_0)}$$

4.5 The case of Minkowski space $\mathbb{R}^{3+1}$

In the case of Minkowski space, we can use the above methods to obtain *global* estimates for all time.

4.5.1 The conserved energy

Let $(t, x^1, \ldots, x^3)$ denote a standard Minkowski coordinate system. Recall from Section 2.4 the Lie algebra of Killing fields of Minkowski space.

The most convenient vector field to use as a multiplier vector field is the future-directed timelike vector field $T = \partial_t$.

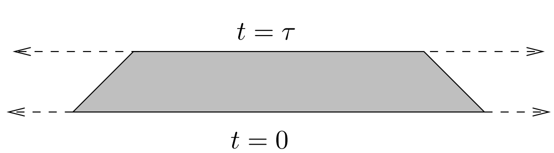
Choosing for instance the constant-t hypersurfaces, and noting that the normal $n = T$, we have that

$$J^T_\mu[\psi]n^\mu = (\partial_t\psi)^2 + \sum_i(\partial_{x_i}\psi)^2$$

and the estimate (32) gives

$$\int_{t=\tau}(\partial_t\psi)^2 + \sum_i(\partial_{x_i}\psi)^2 = \int_{t=0}(\partial_t\psi)^2 + \sum_i(\partial_{x_i}\psi)^2. \tag{44}$$

Wait a minute! How is this justified? Restrict to:



apply more correctly (39) and take the limit, then exchange the roles of $t = \tau$ and $t = 0$.

(Less explicity, to justify the above, one can alternatively argue as follows: We first note that smooth functions of compact support are dense in $H^s(\mathbb{R}^n)$

for all $s \geq 0$. On the other hand, by our general uniqueness result, we can show that if ψ, $\nabla\psi$ are of compact support on $t = 0$, then they are of compact support on $t = \tau$. Thus, it suffices to prove (44) on such functions for which one can manifestly drop the boundary terms if the boundary is sufficiently far.)

4.5.2 Commutation and pointwise estimates

As commutators, we may use $X_i = \partial_{x_i}$, where $x = 1, \ldots 3$. We obtain finally that

$$\|\psi(\tau,\cdot)\|_{\mathring{H}^1} + \|\psi(\tau,\cdot)\|_{\mathring{H}^2} \leq \|\psi(0,\cdot)\|_{\mathring{H}^1} + \|\psi(0,\cdot)\|_{\mathring{H}^2} + \|n\psi(0,\cdot)\|_{H^1} \tag{45}$$

Now let us note that we have the following global Sobolev inequality

Proposition 4.5.1. *For f of compact support,*

$$\sup_{\mathbb{R}^3} |f| \leq C\|f\|_{\mathring{H}^1(\mathbb{R}^3)} + \|f\|_{\mathring{H}^2(\mathbb{R}^3)}. \tag{46}$$

Proof. To prove (46), one can just revisit the proof of the usual Sobolev inequality by taking the Fourier transform.

Alternatively, one can note the following Hardy inequality

$$\int_{\mathbb{R}^3} |\mathbf{x} - \mathbf{x}_0|^{-2}|f|^2 \leq C|f|^2_{\mathring{H}^1(\mathbb{R}^3)}$$

on the closure of functions of compact support in H^1, for any $\mathbf{x}_0 \in \mathbb{R}^3$. To estimate $|\psi(\mathbf{x}_0)|$ we can then apply a local Sobolev inequality in the ball of radius say 1 at $\mathbf{x}_0$. □

It follows that we have

$$\sup_{\mathbb{R}^3} |\psi|(\tau,\cdot) \leq C(\|\psi(0,\cdot)\|_{\mathring{H}^1} + \|\psi(0,\cdot)\|_{\mathring{H}^2} + \|n\psi(0,\cdot)\|_{H^1}). \tag{47}$$

Applying further commutations one obtains

Theorem 4.4. *We have*

$$|\psi(\tau,\cdot)|_{C^{m+1}(\mathbb{R}^3)} \leq C(\|\psi(0,\cdot)\|_{\mathring{H}^{m+1}} + \|n\psi(0,\cdot)\|_{H^{m+2}} + \|\psi(0,\cdot)\|_{\mathring{H}^1})$$

We have secretly snuck in the t-derivatives on the left hand side, even though we have only used X_i as multipliers. Why can we do this?

It is the analogue of the above result which we aim to show for a suitable class of black hole exteriors, including Schwarzschild and subextremal Reissner–Nordström.

4.5.3 An alternative commutation argument

Let us note finally an alternative commutation which is in fact useful in the black hole case, if we only use T as a commutator instead of the X_i we obtain from keeping only the ∂_t derivative in the energy the estimate

$$\|\partial_t^2\psi(\tau,\cdot)\|_{L^2} \leq \|\partial_t\psi(0,\cdot)\|_{\mathring{H}^1} + \|\partial_t^2\psi(0,\cdot)\|_{L^2}. \tag{48}$$

Note that this is the estimate thrown away from (45). Now the wave equation itself can be rewritten

$$\partial_t^2\psi = \Delta\psi$$

whence we have

$$\|\Delta\psi(\tau,\cdot)\|^2_{L^2(\mathbb{R}^3)} = \|\partial_t^2\psi(\tau,\cdot)|^2_{L^2(\mathbb{R}^3)}$$

But now notice the elliptic estimate

$$\|f\|_{\mathring{H}^2(\mathbb{R}^3)} \leq \|\Delta f\|_{L^2(\mathbb{R}^3)}$$

valid for f smooth of compact support.

Thus we retrieve the estimate (**??**) just from (48) and (44). More generally, after further commutations just with T we can retrieve the full Theorem 4.4.

This latter approach is preferable in that it only requires T and not X. It is this approach that we shall use in the black hole case. As a warm-up, the reader may already try (**Exercise**) to generalise Theorem 4.4 to a class (defining this is part of the exercise) of asymptotically flat stationary spacetimes with a timelike Killing field. *Hint: one can get ideas from the more complicated black hole setting of Section 5.2.*

4.6 Aside: From estimates to well-posedness

The energy estimate of Propostion is also the key to proving well posedness of the Cauchy problem.

Let us state for reference

Theorem 4.5. *Let $(\mathcal{M}, g)$ be globally hyperbolic with Cauchy hypersurface Σ. Let $\uppsi \in H^s_{\mathrm{loc}}(\Sigma)$, $\uppsi' \in H^{s-1}_{\mathrm{loc}}(\Sigma)$. Then there exists a unique solution ψ of (2) in $\mathcal{M}$, such that for any spacelike S, $\psi|_S \in H^s_{\mathrm{loc}}$, $n_S\psi|_S \in H^{s-1}_{\mathrm{loc}}$, and $\psi|_\Sigma = \uppsi$, $n\psi|_\Sigma = \uppsi'$.*

Proof. First note that by a minor modification of the (only briefly sketched!) proof of Theorem 4.3, it suffices to show that given a region as in (or as in), in fact, assumed sufficiently small, one can solve this initial value problem in the class $\psi \in C^1(H^s_{\mathrm{loc}}(\Sigma_\tau))$, $n\psi \in C^0 \in H^{s-1}_{\mathrm{loc}}(\Sigma_\tau)$.

For this in turn there are various approaches. One can construct a solution in $L^2(\mathcal{R})$ via Hilbert space methods, and then infer a posteriori its regularity through the energy estimate (see [36]). One can approximate everything by analytic data, and an analytic metric, apply Cauchy-Kowalevsky, and use the energy estimates to take the limit in H^s. One could apply finite differences, or add artificial viscocity. Finally, one could also use much more elaborate tools like the existence of a parametrix (see the approach in [6]).

□

5 The wave equation on black hole backgrounds

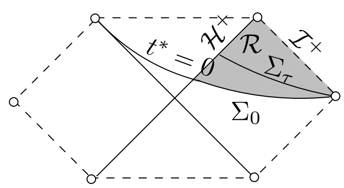
We now turn to the problem of proving definitive global boundedness-type results on a class of stationary black hole backgrounds. The class is motivated by the Schwarzschild (and sub-extremal Reissner–Nordström) setting, so let us first set up the problem in this case where everything is explict.

5.1 The Schwarzschild setting

We will consider a region, completely contained in the Lemaitre spacetime. Recall from Section 3.2 the Lemaitre coordinates $t*$ and r. (We will denote t^* and not t here just to emphasize that this is not the Schwarzschild coordinate t!)

Let us consider the region $\mathcal{R} = \{t^* \geq 0, r \geq 2M\}$. This region is foliated by the spacelike hypersurfaces $\Sigma_\tau = \{t^* = \tau\}$. Let us denote the Killing field $\partial_{t*} = T$, and the one parameter group of diffeomorphisms it generates by ϕ_τ. Note that $\Sigma_\tau = \phi_\tau(\Sigma_0)$.

Define $S_\tau = \{t^* = \tau\} \cap \{r = 2M\}$. Let us denote $\mathcal{H}^+ = \cup_{\tau \geq 0} S_\tau$. In our conventions Σ_τ is a manifold with boundary, with compact boundary S_τ. Again $S_\tau = \phi_\tau(\Sigma_0)$.



The Killing vector field T is null and tangent to $\mathcal{H}^+$. This makes $\mathcal{H}^+$ a *Killing horizon*. One easily shows (see below) that in this case $\nabla_T T = \kappa T$ for some function κ such that $T\kappa = 0$. In the Schwarzschild case, in view of this spherical symmetry, it is immediate that κ is constant on $\mathcal{H}^+$. But one can explicitly compute

$$\kappa = \frac{1}{4M}.$$

We defer this computation to Section 5.5 where the positivity of κ will be fundamental.

Finally, we note the surfaces Σ_0 are *asymptotically flat*. For us here, this will mean the following: There exists a compact set K_0, in this case, let us take $K = \{r \leq R\} \cap \Sigma_0$, for some large R, such that $\Sigma_0 \setminus K_0$ is covered by global Euclidean coordinates x^i as preceisely the region $|x| \geq R$, and the metric $g|_{\Sigma_0}$ of Σ_0 approaches the Euclidean metric

$$dx_i^2 + \ldots + dx_n^2$$

in a suitable sense. We will give the details of this later.

Let us also note that $\partial_t|_{\Sigma_0}$ approaches asymptotically n_{Σ_0}. We will need only the statement that

$$T|_{\Sigma_0} = \alpha n_{\Sigma_0} + \xi, \tag{49}$$

where $\xi \in T_p\Sigma_0$ and $g(\xi, \xi) \to 0$, $\alpha \to 1$, as $|x| \to \infty$.

5.2 A sufficiently general class of spacetimes

Let us abstract the above properties of Schwarzschild in the following definition.

Definition 5.1. *A* sub-extremal strictly stationary black-hole spacetime *will be a manifold* $(\mathcal{M}, g)$ *containing a region* $\mathcal{R}$ *with the following properties.*

$\mathcal{R}$ *is bounded to the future by a a future complete null hypersurface* $\mathcal{H}^+$*, not-necessarily connected, and to the past by an asymptotically flat spacelike hypersurface* Σ_0 *with compact boundary* $X_0 = \Sigma_0 \cap \mathcal{H}^+$*.*

There exists a smooth vector field T *on* $\mathcal{M}$ *which generates a* 1*-parameter global semigroup of diffeomorphisms* ϕ_τ *for* $\tau \geq 0$*. Defining* $\Sigma_\tau = \phi_\tau(\Sigma_0)$ *we suppose that* $\mathcal{R} = \cup_{\tau \geq 0}(\Sigma_0)$ *and* $\mathcal{H} = \cup_{\tau \geq 0}(X_0)$*.*

We assume moreover that T *is Killing on* $\mathcal{R}$ *and future-directed timelike in its interior, and that* $\mathcal{H}$ *is a null hypersurface, (making* T *null on* $\mathcal{H}^+$ *and* $\mathcal{H}^+$ *a so-called Killing horizon). Writing*

$$\nabla_T T = \kappa T$$

on $\mathcal{H}^+$*, we assume that* $\kappa \geq \kappa_0 > 0$*.*

Finally, we will assume as before that T *satsifies* (49) *with* $\xi \in T_p\Sigma_0$ *and* $g(\xi, \xi) \to 0$*,* $\alpha \to 1$*, as* $|x| \to \infty$*, where* $|x|$ *denotes Euclidean coordinates on the asymptotically flat end of* Σ_0 *given by the definition.*

The reader at this point can construct a region $\mathcal{R}$ of sub-extremal Reissner–Nordström satisfying the above assumptions. Note that in that case (**Exercise**) we have

$$\kappa = .$$

We stress that we are not assuming the spacetime to be spherically symmetric or the horizon to be connected.

We use the word *strictly stationary* to emphasize the assumption that T is assumed timelike everwhere in $\mathcal{R}$. **In particular, the Kerr spacetime, to be discussed later, does not satisfy the assumption of Definition 5.1.** See Section 5.5.

5.3 Preliminaries

For each $\epsilon > 0$, let B_0^ϵ be a non-empty neighbourhood of X_0 in Σ_0 such that $B_\epsilon \subset B_{\epsilon'}$ for $\epsilon < \epsilon'$, and $\cap B_\epsilon = \emptyset$. Define $\Sigma_0^\epsilon = \Sigma_0 \setminus B_\epsilon$, and define $\Sigma_\tau^\epsilon = \phi_\tau(\Sigma_0^\epsilon)$.

In the Schwarzschild case, we can take explictly $B_\epsilon = \Sigma_0 \cap \{r < 2M + \epsilon\}$. This makes $\Sigma_\tau^\epsilon = \{t^* = \tau\} \cap \{r \geq r + 2M + \epsilon\}$.

5.3.1 Positivity of the J^T current

Let n denote the normal of Σ_τ. This is a smooth vector field on $\mathcal{R}$. We have the following relations:

By compactness, our assumptions on T and our notion of asymptotic flatness, we have on Σ_ϵ, the estimate:

$$|J^T_\mu[\psi]n^\mu| \geq C_\epsilon(|\nabla|_{\Sigma_\tau}\psi|^2 + |n\psi|^2)$$

One can see this explicitly in the Schwarzschild case:

In fact, on Σ_τ we have

$$|J^T_\mu[\psi]n^\mu| \sim C((1-2M/r)(\partial_r\psi)^2 + (\partial_{t*}\psi)^2 + |\nabla\!\!\!/\psi|^2)$$

This makes clear the precise nature of the degeneration.

5.3.2 Reexpressing $\Box_g$

With the commutation method of Section 4.5.3 in mind, we make the following remarks relating $\Box_g$ and $\Delta_{|\Sigma_\tau}$.

Note that

$$\Box_g = \Delta|_{\Sigma_\tau} - n^2|_{\Sigma_\tau} + P$$

where P is a first order operator. If we rewrite $n = X + \xi$, where ξ is tangential to g_{Σ_τ}, then we have

$$\Box_g = -X^2 + \Delta|_{\Sigma_\tau} - \xi^2 - X\xi - \xi X + P$$

We have that if $X(p)$ is timelike, then at p

$$(\Delta_{\Sigma_\tau} - \xi^2)\text{``}\geq\text{''}\, c(p)\Delta_{\Sigma_\tau}$$

for some $c > 0$.

Our assumption of asymptotically flatness

5.3.3 Hardy and Sobolev inequalities

We have the following Hardy inequality on Σ_τ:

Proposition 5.3.1. *For smooth f of compact support*

$$\int_{\Sigma_\tau} \frac{1}{\nu^2}|f|^2 \leq C\|f\|^2_{\mathring{H}^1(\Sigma_\tau)}$$

where C is independent of τ and $\nu \geq c > 0$ is a ϕ_τ invariant smooth function such that $\nu \sim |x|$ near infinity.

In the Schwarzschild case, one can take for instance $\nu = r$.

Proof. Clearly suffices to prove on Σ_0.

We partition $\Sigma = K \cup \Sigma \setminus K$ where K is a compact manifold with boundary $X \cup S$ and $\Sigma \setminus K$ a manifold with boundary such that .

It suffices to $\square$

From this we have again the Sobolev inequality on Σ_τ,

Proposition 5.3.2. *For f smooth of compact support on Σ_τ,*

$$|f| \leq C(\|f\|_{\mathring{H}^2(\psi)} + \|f\|_{\mathring{H}^1(\psi)})$$

where C is independent of f and τ.

5.3.4 The wave equation on $\mathcal{R}$

By our assumptions, Σ_0 is a *past* Cauchy hypersurface for $\mathcal{R}$. We may impose initial data ψ, ψ' on Σ_0 for the wave equation on the hypersurface Σ_0 and solve the wave equation, i.e. we have the following Corollary of Theorem 4.5:

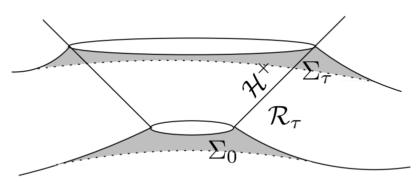
Corollary 5.1. *Given $k > 1$ and initial data $\psi \in H^k_{\mathrm{loc}}(\Sigma_0)$, $\psi' \in H^{k-1}_{\mathrm{loc}}(\Sigma_0)$, there exists a unique solution of (27) in $\mathcal{R}$ satisfying the regularity property of Theorem 4.5. In particular, if ψ, ψ' are C^∞ then ψ is C^∞. Moreover, if ψ, ψ' are of compact support then $\psi|_{\Sigma_\tau}, n\psi|_{\Sigma_\tau}$ will be of compact support for all τ.*

In what follows, for convenience, we will denote by ψ the solution of (27) arising from given data ψ, ψ' which are indeed assumed smooth and compactly supported.

The estimates will then hold more generally for the closure of this set in the space defined by the norms that appear on the right hand side of our inequalities.

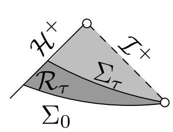
5.4 The conserved energy identity

Let us consider the region $\mathcal{R}_\tau = \cup_{t=0}^{\tau} \Sigma_t$.



This region is bounded to the future by $\Sigma_\tau \cup \mathcal{H}_\tau$, where $\mathcal{H}_\tau = \mathcal{H} \cap \mathcal{R}_\tau$ and to the past by Σ_0.

In the Schwarzschild case, we may alternatively depict this region by the formal two-dimensional Penrose diagramme:



For ψ, ψ′ of compact support, we may write the energy estimate

$$\int_{\mathcal{H}^+_\tau} J^T_\mu[\psi]n^\mu + \int_{\Sigma_\tau} J^T_\mu[\psi]n^\mu = \int_{\Sigma_0} J^T_\mu[\psi]n^\mu$$

where the absence of the boundary term is immediately justified by .

Since

$$\int_{\mathcal{H}^+_\tau} J^T_\mu[\psi]n^\mu \geq 0, \qquad \int_{\Sigma_\tau \setminus \Sigma^\epsilon_\tau} J^T_\mu[\psi]n^\mu \geq 0$$

we have immediately that for any $\epsilon > 0$,

$$c_\epsilon(\|\nabla\psi\|^2_{\mathring{H}^1(\Sigma^\epsilon_\tau)} + \|n\psi\|^2_{L^2(\Sigma^\epsilon_\tau)}) \leq \int_{\Sigma^\epsilon_\tau} J^T_\mu[\psi]n^\mu \leq \int_{\Sigma_0} J^T_\mu[\psi]n^\mu.$$

Thus, we have proven:

Proposition 5.4.1 (Degenerate energy bound)**.** *For all $\epsilon > 0$, there exists a C_ϵ such that for all τ and all solutions ψ, we have*

$$\|\nabla\psi\|^2_{\mathring{H}^1(\Sigma^\epsilon_\tau)} + \|n\psi\|^2_{L^2(\Sigma^\epsilon_\tau)} \leq C_\epsilon(\|\nabla\psi\|^2_{\mathring{H}^1(\Sigma_0)} + \|n\psi\|^2_{L^2(\Sigma_0)})$$

Using only this identity applied to ψ and to $T\psi$ (as in Section 4.5.3) one can already show (**Exercise**) that ψ in Σ_ϵ is pointwise bounded by

$$\sup_{\Sigma^\epsilon_\tau} |\psi| \leq C_\epsilon(\|\psi\|_{\mathring{H}^2(\Sigma_0)} + \|\psi\|_{\mathring{H}^1(\Sigma_0)} + \|n\psi\|_{H^1(\Sigma_0)}). \tag{50}$$

We see thus that the degeneracy of Proposition 5.4.1 is inherited by pointwise bounds.

The reader may want to work out in the Schwarzschild case the precise degeneration in the above as $\epsilon \to 0$ (**Exercise**).

5.5 The red-shift identity

In order to remove the above degeneracy, one must use the so-called *red-shift effect* at the horizon.

As stated, this is capture by the identity $\nabla_T T = \kappa T$ where $\kappa > c > 0$. Let us first see in Schwarzschild why this is the case. In fact, we can essentially argue without explicit computation of κ:

Proposition 5.5.1. *On $\mathcal{H}^+$ in Schwarzschild, $\nabla_T T = \kappa T$ for a constant κ.*

Proof. Let us extend T to a ϕ_τ-invariant null frame T, Y, E_1, E_2 on $\mathcal{H}^+$. Note the Killing equation

$$g(\nabla_X T, Z) + g(\nabla_Z T, X) = 0$$

It follows that t $g(\nabla_T T, T) = 0$. On the other hand

$$0 = g(\nabla_T T, E_i) + g(\nabla_{E_1} T, T) = g(\nabla_T T, E_i) + \frac{1}{2}E_i g(T,T) = g(\nabla_T T, E_i)$$

the last inequality since E_i is tangential to $\mathcal{H}^+$. It follows that $\nabla_T T = \kappa(x)T$ for some function $\kappa : \mathcal{H}^+ \to \mathbb{R}$.

Now note that since $\nabla_T T$ and T are ϕ_t-translation invariant then $T\kappa = 0$. On the other hand, since $\nabla_T T$ and T are also invariant to the spherically symmetric action, then $E_i\kappa = 0$. It follows that κ is constant.

Finally let us note that the equation $\nabla_T T = \kappa T$ tells us that $L = e^{-\kappa t^*}T$ satisfies

$$\nabla_L L = e^{-\kappa t^*}\nabla_T L = e^{-\kappa t^*}((-\kappa)e^{-\kappa t^*}T + \kappa e^{-\kappa t^* T}) = 0$$

Since T vanishes at the bifurcation sphere of the maximal Schwarzschild, it follows that $\kappa > 0$! $\square$

At the geometric optics level, the red-shift can be understood as follows: If $\underline{L}$ denotes a parallely propagated null vector field such that $g(\underline{L}, L) = -2$, then if n_{Σ_τ} denotes the normal to the foliation Σ_τ, then we can think of $\underline{L}$ as representing the displacement between to infinitessimally separated peaks of a wave emitted by an observer at $t^* = 0 \cap \mathcal{H}^+$ to $t^* = \tau$. The wave-length of the emitted wave will then be measured by the observer at time $t^* = 0$ represented by n_{Σ_0} as $g(n_{\Sigma_0}, \underline{L})$ while the wave length of the received signal by observer n_{Σ_τ} will be $g(n_{\Sigma_\tau}, \underline{L})$ But since ϕ_τ takes n_{Σ_0} to n_{Σ_τ}, and $\underline{L} = e^{+\kappa t^*}Y$ for Y as in the previous proposition, it follows that the wavelength *grows* exponentially, and thus the frequency is *damped* exponentially by the factor $e^{-\kappa t^*}$, i.e. the frequency is shifted *to the red*.

One can see this similarly from the point of view of the Gaussian beam approximation.

It turns out that this geometric optics effect manifests itself in the positivity properties near the event horizon $\mathcal{H}^+$ of the "bulk" (spacetime integral) term of an energy identity. (In some sense this energy identity is more general because geometric optics really should only concern the high-frequency limit.)

This is incorporated in the very general construction of a vector field given by the following

Proposition 5.5.2. *There exists a vector field N such that*

1. $(\phi_\tau) * N = N$
2. *N is future directed timelike*
3. *$N = T$ in $\Sigma_0 \setminus B$ where B is compact*
4. *There exists an $\epsilon > 0$ and a $c > 0$ such that on $\Sigma_\tau \setminus \Sigma_\tau^\epsilon$,*

$$K^N \geq cJ^N_\mu N^\mu.$$

Proof. First note that it suffices to construct a timelike vector field N_0 along $\Sigma_0 \setminus \Sigma_0^{2\epsilon}$ such that 4. holds.

For if χ is a cutoff function such that $\chi = 1$ in $\Sigma^\epsilon \setminus \Sigma^{2\epsilon}$, then by the convexity of I^+, we have that $\tilde{N}_0 = \chi N_0 + (1-\chi)T$ is future directed timelike, and 4. still holds in $\Sigma_0 \setminus \Sigma_0^\epsilon$, and $N = T$ in . But now define N globally by the push forward of $\tilde{N}_0$.

We now proceed to the construction of N_0 along $\Sigma_0 \setminus \Sigma_0^{2\epsilon}$. For this we consider the null vector T along S_0, and let Y be a null vector along S_0 with the property that $g(Y,T) = -2$.

We extend Y to $\Sigma_0 \setminus \Sigma_0^{2\epsilon}$ so that

$$\nabla_Y Y = -\sigma(Y+T)$$

along S_0, for σ a large constant to be determined. This is easily realised (**Exercise**). We claim that $N_0 = Y + T$ satisfies the required property.

For this we introduce a local orthonormal frame E_1, E_2 of the normal plane to (Y,T) at a point $p \in S_0$.

We compute at p:

$$\nabla_T Y = -\kappa Y + a^1 E_1 + a^2 E_2$$

$$\nabla_Y Y = -\sigma T - \sigma Y$$

$$\nabla_{E_1} Y = h_1^1 E_1 + h_1^2 E_2 - \frac{1}{2} a^1 Y$$

$$\nabla_{E_2} Y = h_2^1 E_1 + h_2^2 E_2 - \frac{1}{2} a^2 Y$$

(where $h_1^2 = h_2^1$).

We now compute

$$\begin{aligned} K^Y =& \frac{1}{2}(\mathbf{T}(Y,Y)\kappa + \mathbf{T}(T,Y)\sigma + \mathbf{T}(T,T)\sigma) \\ &- \frac{1}{2}(\mathbf{T}(E_1,Y)a^1 + \mathbf{T}(E_2,Y)a^2) \\ &+ \mathbf{T}(E_1,E_1)h_1^1 + \mathbf{T}(E_2,E_2)h_2^1 + \mathbf{T}(E_1,E_2)(h_1^2 + h_2^1) \end{aligned} \tag{51}$$

Now note that the second two lines can be bounded in absolute value by

$$c(\mathbf{T}(T, Y+T) + \sqrt{\mathbf{T}(T,Y+T)\mathbf{T}(Y,Y)}) \tag{52}$$

for large enough c independent of σ (since all constants appearing only depend on $Y|_{S_0}$). For this in turn we use the fact that the second two lines do not contain a $\mathbf{T}(Y,Y)$ term.

Now, by choosing σ large enough we clearly can absorb (52) by a little bit of the terms of the first line of (51), to obtain

$$K^Y \geq \frac{1}{4}(\mathbf{T}(Y,Y)\kappa + \mathbf{T}(T,Y)\sigma + \mathbf{T}(T,T)\sigma) \geq c J_\mu^Y n^\mu$$

□

5.6 Non-degenerate energy boundedness

We now proceed to prove a non-degenerate version of Proposition 5.4.1.

Proposition 5.6.1 (Non-degenerate energy bound)**.** *There exists a C such that for all τ and all solutions ψ, we have*

$$\|\nabla\psi\|^2_{\mathring{H}^1(\Sigma_\tau)} + \|n\psi\|^2_{L^2(\Sigma_\tau)} \leq C(\|\nabla\psi\|^2_{\mathring{H}^1(\Sigma_0)} + \|n\psi\|^2_{L^2(\Sigma_0)}).$$

Proof. Consider the energy identity for the vector field N of Proposition 5.5.2 in any region $\mathcal{R}_{(\tau_-,\tau_+)}$

$$\int_{\mathcal{H}^+_{\tau_-,\tau_+}} J^N_\mu[\psi]n^\mu + \int_{\Sigma_{\tau_+}} J^N_\mu[\psi]n^\mu + \int_{\mathcal{R}_{\tau_-,\tau_+}} K^N[\psi] = \int_{\Sigma_{\tau_-}} J^N_\mu[\psi]n^\mu$$

We may partition the integral

$$\int_{\mathcal{R}_{\tau_-,\tau_+}} K^N[\psi] = \int_{\mathcal{R}_{\tau_-,\tau_+}\setminus\cup\Sigma^\epsilon} K^N[\psi] + \int_{\cup_{\tau_-\leq\tau\leq\tau_+}\Sigma^\epsilon_\tau} K^N[\psi]$$

and thus rewrite the energy identity as

$$\begin{aligned}\int_{\mathcal{H}^+_{\tau_-,\tau_+}} J^N_\mu[\psi]n^\mu + \int_{\Sigma_{\tau_+}} J^N_\mu[\psi]n^\mu + \int_{\mathcal{R}_{\tau_-,\tau_+}\setminus\cup\Sigma^\epsilon} K^N[\psi]& \\ = \int_{\Sigma_{\tau_-}} J^N_\mu[\psi]n^\mu - \int_{\cup_{\tau_-\leq\tau\leq\tau_+}\Sigma^\epsilon_\tau} K^N[\psi].&\end{aligned}$$

We infer

$$\int_{\Sigma_{\tau_+}} J^N_\mu[\psi]n^\mu + \int_{\mathcal{R}_{\tau_-,\tau_+}\setminus\cup\Sigma^\epsilon} K^N[\psi] \leq \int_{\Sigma_{\tau_-}} J^N_\mu[\psi]n^\mu + \int_{\cup_{\tau_-\leq\tau\leq\tau_+}\Sigma^\epsilon_\tau} |K^N[\psi]|.$$

Note that by the coercive property 4. and the coarea formula, we have that

$$c\int_{\tau_-}^{\tau_+}\int_{\Sigma_{\tau*}\setminus\Sigma^\epsilon_{\tau*}} J^N_\mu[\psi]n^\mu \leq \int_{\mathcal{R}_{\tau_-,\tau_+}\setminus\cup\Sigma^\epsilon} K^N[\psi]$$

On the other hand, we have

$$\int_{\cup_{\tau_-\leq\tau\leq\tau_+}\Sigma^\epsilon_\tau} |K^N[\psi]| \leq C\int_{\tau_-}^{\tau_+}\left(\int_{\Sigma_\tau} J^T_\mu[\psi]n^\mu\right)d\tau$$

We now add a large constant times the following inequality:

$$\int_{\tau_-}^{\tau_+}\left(\int_{\Sigma_\tau} J^T_\mu[\psi]n^\mu\right)d\tau \leq (\tau_+-\tau_-)\int_{\Sigma_{\tau_-}} J^T_\mu[\psi]n^\mu$$

to obtain

$$\int_{\Sigma_{\tau_+}} J^N_\mu[\psi]n^\mu + c\int_{\tau_-}^{\tau_+}\int_{\Sigma_{\tau*}} J^N_\mu[\psi]n^\mu \leq \int_{\Sigma_{\tau_-}} +(\tau_+-\tau_-)C\int_{\Sigma_{\tau_-}} J^T_\mu[\psi]n^\mu$$

Of course

$$\int_{\Sigma_{\tau_-}} J^T_\mu[\psi]n^\mu \leq \int_{\Sigma_0} J^T_\mu[\psi]n^\mu = D$$

Defining

$$f(\tau) = \int_{\Sigma_\tau} J^N_\mu[\psi]n^\mu$$

we have that

$$f(t_+) + c\int f(t) \leq f(t_-) + CD(\tau_+ - \tau_-)$$

Lemma 5.6.1. *Let $f : [0, \infty) \to \mathbb{R}$ be a nonnegative function satisfying:*

$$f(t_+) + c\int f(t) \leq f(t_-) + \tilde{C}(\tau_+ - \tau_-)$$

Then

$$f(\tau) \leq B(f(0) + \tilde{C}).$$

Note that the statement of the lemma can be slightly refined (**Exercise**).

The statement of the proposition follows immediately in view of the relation of f with the geometric Sobolev norms. □

5.7 Commutation and pointwise estimates

5.7.1 The T-commutation

We may apply the energy identity to

$$J^N[T\psi]$$

to obtain the boundeness of

As we discussed before, the is is already sufficient to obtain. In fact $J^T[T\psi]$ is already sufficient.

5.7.2 The red-shift commutation

Let us denote $N = Y + T$ where Y is null at the event horizon.

We have the following remarkable commutation identity:

Proposition 5.7.1. *Under the assumptions of we have at $\mathcal{H}^+$*

$$\Box Y\psi = \kappa YY\psi + \sum_{|\mathbf{m}|\leq 2, m_i \leq 2, m_4 \leq 1} c_{\mathbf{m}} E_{m_1}E_{m_2}T_{m_3}Y_{m_4}\psi$$

Proof. Use the general formula

$$\Box Y\psi = 2^{(Y)}\pi^{\mu\nu}\nabla_\mu\nabla_\nu\psi + l.o.t$$

and the computation we already did! □

Let us quickly explain how this is used to obtain . Applying, the energy identity for

$$J^N_\mu[Y\psi]$$

the bulk terms which will appear are

$$K^N[Y\psi]$$

and

$$\mathcal{E}^N[Y\psi]$$

The latter is

$$\kappa NY\psi YY\psi$$

plus terms arising from the right hand side.

Using the re-expression of the Laplacian and elliptic estimates we obtain finally

$$\|\nabla\psi\|^2_{\mathring{H}^2(\Sigma_\tau)} + \|n\psi\|^2_{H^1(\Sigma_\tau)} \leq C_k(\|\nabla\psi\|^2_{\mathring{H}^2(\Sigma_0)} + \|n\psi\|^2_{H^1(\Sigma_0)}).$$

One can in fact iterate Proposition 5.7.1 to commutations of arbitrary high order. One easily shows then

Theorem 5.1 (Non-degenerate higher order energy bound)**.** *For all $k \geq 0$, there exists a C_k such that for all τ and all solutions ψ, we have*

$$\|\nabla\psi\|^2_{\mathring{H}^k(\Sigma_\tau)} + \|n\psi\|^2_{H^{k-1}(\Sigma_\tau)} \leq C_k(\|\nabla\psi\|^2_{\mathring{H}^k(\Sigma_0)} + \|n\psi\|^2_{H^{k-1}(\Sigma_0)}).$$

Using our Sobolev inequalities obtain finally pointwise bounds to all orders

Corollary 5.2. *For all $k \geq 0$, there exists a C^k such that for all solutions ψ we have*

$$|\psi|_{C^k(\mathcal{R})} \leq C_k(\|\nabla\psi\|^2_{\mathring{H}^{k+2}(\Sigma_0)} + \|n\psi\|^2_{H^{k+1}(\Sigma_0)}).$$

5.7.3 Aside: angular momentum operators

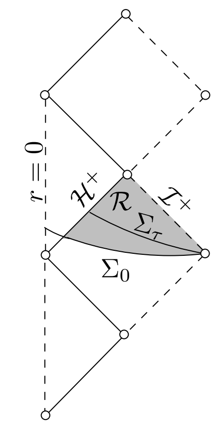
5.8 Aside: extremal Reissner–Nordström and the degeneracy of the red-shift

5.8.1 Extremal Reissner–Nordström

In discussion of the construction of the Reissner–Nordström family, we considered the case of initial data . What happens when we try to impose on both that .

This in fact leads to a *homogeneous* solution of the Einstein–Maxwell equations (**Exercise**) known as Bertotti–Robinson.

When we impose that r is constant on one and not on the other, we obtain so called extremal Reissner-Nordstrom.



This has a subregion with Penrose diagramme as follows. This region satisfies , except for the assumption $\kappa > 0$. We have in fact that $\nabla_T T = 0$ on $\mathcal{H}^+$.

All these properties can in fact be inferred from taking a limit of the subextremal case.

Let us note that the degenerate energy bound Proposition 5.4.1 still holds.

5.8.2 A conserved charge on the horizon

It turns out that one can prove that there does *not* exist a vector field N satisfying Proposition .

One can do the following remarkable elementary calculation: The wave equation along $\mathcal{H}^+$ takes the form

5.8.3 The Aretakis instability

5.9 Aside: the wave equation on the black hole interior

6 Looking ahead

The remainder of these notes will look ahead to more advanced topics concerning waves on black holes.

In Section 6.1, we shall discuss the issues involved in showing *decay properties*, as opposed to just boundedness, for solutions of (27), properties that turn out to be much more sensitive to the geometry and will require us to restrict to Schwarzschild or Reissner–Nordström. We will then turn in Section 6.2 to the remarkable Kerr family of black holes, an object of amazing complexity and beauty where all the difficulties are further . Finally, after a brief discussion of the Cauchy problem for the Einstein equations in Section 6.3, we will end in Section 6.4 with the formulation of the non-linear stability conjecture of the Schwarzschild and Kerr black hole spacetimes themselves.

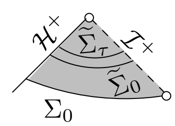
Whereas in the previous sections, we have attempted to give a self-contained treatment, the discussion here gives only the briefest of sketches. It is meant to wet the readers appetite so as to consult the literature. We give only the most basic references here; further references to these are provided in the lecture notes [19], and the survey article [20]. This is a fast moving field, and understanding of these problems is still developing and improving!

6.1 The problem of decay

With boundedness understood, at least for the simplest black hole spacetimes, the next problem to understand is that of decay.

It turns out that whereas the assumptions of Section 5.2 were sufficient for boundedness, good decay results will require more structure. Thus, in this section, we will restrict to Schwarzschild (or Reissner–Nordström). We will discuss in Section 6.1.3 what more general spacetimes one could expect the result to hold for.

Decay can of course be measured in various . To understand in what sense the energy decays, it is useful to have a foliation $\widetilde{\Sigma}_\tau$ or a subset of the future of the original asymptotically flat Cauchy hypersurface Σ_0, the leaves of which which now terminate at in null infinity $\mathcal{I}^+$, as depicted here:



In the Schwarzschild case, an explicit choice of such a foliation is provided by

$$\widetilde{\Sigma}_\tau =$$

A proto-type theorem is as follows:

Theorem 6.1 (Decay for Schwarzschild)**.** *We have*

$$\int_{\widetilde{\Sigma}_\tau} J^N_\mu[\psi]n^\mu \lesssim \tau^{-2} \int_{\Sigma_0} r^2 (J^N_\mu[TT\psi] + J^N_\mu[T\psi] + J^N_\mu[\psi]),$$

and for all $\delta > 0$

$$\int_{\widetilde{\Sigma}_\tau} J^N_\mu[N\psi]n^\mu \lesssim \tau^{-4+\delta} \sum_{\mathcal{C}_i} \int_{\Sigma_0} r^4 J^N_\mu[\mathcal{C}_i\psi]n^\mu,$$

for an appropriate collection of commutation vector fields $\mathcal{C}_i$ *up to order .*

Pointwise bounds follow easily from the above and Sobolev inequalities

Corollary 6.1. *We have the pointwise decay bounds*

$$\sup_{\widetilde{\Sigma}_\tau} r|\psi| \lesssim \sqrt{E}\tau^{-1/2}$$

$$\sup_{\widetilde{\Sigma}_\tau\cap\{r\leq R\}} |\psi| \lesssim \sqrt{E}\tau^{-3/2+\delta}$$

$$\sup_{\widetilde{\Sigma}_\tau\cap\{r\leq R\}} \sqrt{J^N_\mu[\psi]N^\mu} \lesssim E\tau^{-2+\delta}$$

where E denotes an appropriate higher order energy arising from further commutations of the the right hand side of the inequalities of Theorem 6.1.

Similar results are true in subextremal Reissner–Nordström. Note, however, that the relevant constants implicit in $\lesssim$ would degenerate as one approaches extremality.

Theorem 6.1 is a precise analogue of what is true in Minkowski space, except for the "loss of derivatives", i.e. the fact that the right hand side of the inequality refers to more derivatives than the left hand side. (There is of course always differentiability loss in the pointwise bounds of Corollary 6.1 due to Sobolev inequality.) The loss of differentiability for the global bounds even at the level of energy, for which there is no such loss locally, is a fundamental aspect of the problem, and has to with a feature of black hole geometry we have not previously discussed–namely the presence of *trapped null geodesics*.

6.1.1 The photon sphere and trapped null geodesics

The origin of the loss of derivatives in the estimate of Theorem 6.1 is the presence of so-called *trapped null geodesics*.

Trapped null geodesics are the analogue of trapped rays in the obstacle problem. The former are defined as rays which remain in a compact subset of space.

In general relativity, one has to think from the point of view of spacetime. Given a foliation $\widetilde{\Sigma}_\tau$ with normal $\widetilde{n}$, we say that a geodesic γ is trapped with respect to $\widetilde{\Sigma}_\tau$ if γ crosses $\widetilde{\Sigma}_\tau$ for all τ, and $|g(\dot\gamma(s),\widetilde{n})|$ remains uniformly bounded away from 0 for all parameters $s > 0$. A theorem of Sbierski [31] (proven in the context of a general class of Lorentzian manifolds) says that given the above, one can construct a sequence of solutions of the wave equation whose energy flux through $\widetilde{\Sigma}_\tau$ is bounded below by a uniform constant (independent of τ) times its energy flux at Σ_0. Thus, the existence of a single trapped null geodesic would imply that the first statement of Theorem 6.1 cannot hold if all T commutations on the right hand side were removed.

An example of such a geodesic is any null geodesic γ which is intially tangent to the sphere $r = 3M$. One can show (**Exercise**) that such geodesics remain tangent to $r = 3M$. Thus, clearly γ intersects every $\widetilde{\Sigma}_\tau$. On the other hand, since $T = N$ at $r = 3M$, and $g(\dot\gamma, T)$ is constant since T is Killing, the second property is also true.

It follows immediately that Theorem 6.1 cannot be true for a quantity .

Let us note a non-example of the notion of "trapped null geodesics" as we have formulated it: the null geodesic generators γ of the horizon. For although we have that γ reaches every $\widetilde{\Sigma}_\tau$, the quantity $g(\dot{\gamma}(s), N)$ goes to 0 as $s \to \infty$. The latter is precisely the statement that these geodesics are red-shifted.

Note that this theorem also can be invoked to understand the necessity of the r weight.

6.1.2 Integrated local energy decay

It turns out that the key to proving decay statements is something known as "integrated local energy decay".

To state this, let us define a nonnegative function χ_1 satisfying $T\chi_1 = 0$ which vanishes to second order at $r = 3M$, is strictly positive everwhere else, and which decays like $r^{-1-\delta}$ for some $\delta > 0$, and a second nonnegative function χ_0 which is everywhere positive and vanishes like r^{-3}.

We have

Theorem 6.2 (Integrated local energy decay with degeneration)**.** *With χ_1 and χ_0 as above, we have*

$$\int_{\mathcal{R}\cap J^+(\widetilde{\Sigma}_\tau)} \chi_1 |J^N_\mu[\psi]N^\mu| + \chi_0|\psi|^2 \leq \int_{\widetilde{\Sigma}_\tau} J^N_\mu[\psi]n^\mu$$

We note here again that the theorem of Sbierski [31] discussed in Section 6.1.1 above also shows that the statement of Theorem 6.2 cannot be correct if $\chi_1(3M) > 0$. On the other hand, one can refine slightly the statement: In particular, at $r = 3M$, one can control the Schwarzschild r-derivative without degeneration. One can also weaken the order of vanishing of the function χ_1.

A Corollary is

Corollary 6.2 (Integrated local energy decay with derviative loss)**.** *We have*

$$\int_{\mathcal{R}\cap J^+(\widetilde{\Sigma}_\tau)} \chi J^\mu_\nu[\psi]N^\mu \leq \int_{\widetilde{\Sigma}_\tau} J^\mu_\nu[T\psi]n^\mu$$

The proof of the above proceeds by a vector field identity. Whereas in our previous uses of such identities, the goal was to

6.1.3 A black-box approach to decay

It turns out that given versions of Theorems 5.1 and 6.2, the statement of Theorem 6.1 follows, using only the asymptotic flatness of the background. This is a useful "black-box" result which can be used to . In particular, this does not require stationarity (though remember that proving do require stationarity or almost stationarity).

6.2 The Kerr solution

We have restricted our attention in these notes to the simplest examples of black holes, Schwarzschild and Reissner–Nordström. For the beginner, these are certainly difficult enough! But nature always surprises us more.

The Schwarzschild family turns out to be one parameter subfamily of a larger family of vacuum spacetimes, the so-called *Kerr family*. These are much more subtle–in particular, they were quite subtle to find, as the family was discovered only in 1963.

6.2.1 The Kerr metric

To give an explicit expression for the metric, we first take the point of view of Section ,.

We fix parameters $M > 0$, $|a| < M$, we define

$$\Delta = r^2 - 2Mr + a^2,$$

and we let r_+ denote the larger root of $\Delta = 0$, though of as a quadratic polynomial in r.

Thinking of r as the $r = \sqrt{|x_1|^2 + |x_2|^2 + |x_3|^2}$ of an ambient $\mathbb{R}^{3+1}$, and θ, ϕ as the associated spherical coordinates, we will define the Kerr metric on the manifold $\{r > r_+\} \subset \mathbb{R}^{3+1}$ by the expression

$$g_{M,a} = -\frac{\Delta}{\rho^2}\left(dt - a\sin^2\theta d\phi\right)^2 + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + \frac{\sin^2\theta}{\rho^2}\left(a\,dt - (r^2 + a^2)d\phi\right)^2$$

where Δ is defined as above and

$$\rho^2 = r^2 + a^2\cos^2\theta$$

Let us think of the above abstractly as a manifold $(\mathcal{M}_0, g)$. Then, $\mathcal{M}_0$ extends so as for .

We see immediately that the above expression is no longer spherically symmetric. Indeed, for $|a| \neq 0$, the dimension of the Lie algebras is precisely two, and these are generated by ∂_t and ∂_ϕ.

On the other hand, there is an additional higher-order Killing tensor, given by the expression:

6.2.2 The ergoregion and superradiance

The most obvious difficulty of the Kerr metric is that $T = \partial_t$ is not everywhere timelike in the black hole exterior.

From the metric, we read immediately that $T = \partial_t$ is spacelike in the region

$$\begin{aligned}\mathcal{E} &= \{\Delta \leq a^2\sin^2\theta\}\\ &= \{r \leq M + \sqrt{M^2 - a^2\cos^2\theta}\}\\ &= \{\rho^2 \leq 2Mr\}\end{aligned}$$

The set $\mathcal{E}$ is known as the *ergoregion* and its boundary the *ergosphere*.

In view of the presence of an ergo-region, Kerr does not satisfy the assumptions of Defintion . Indeed, the problem of boundedness of solutions to the scalar wave equation (27) on Kerr is much more difficult than in the .

We see immediately why . Though we may apply the energy identity of the T vector field, we no-longer have .

In particular the flux on $\mathcal{I}^+$ can be greater than the initial data (or, in the scattering of). Hence the term *superradiance*.

Indeed, it is not a priori clear under what conditions this flux is even finite.

6.2.3 Kerr's trapped null geodesics

As far as decay is concerned, one has a similar problem with Schwarzschild. But whereas in Schwarzschild, all trapped null geodesics asymptote to $r = 3M$, in the Kerr case, trapped null geodesics occur in a region…

6.2.4 The global geometry of Kerr and the black hole interior

Just as in the case of Schwarzschild and Reissner–Nordström, one can ask what is the correct largest underlying manifold that the Kerr metric can be defined on.

It turns out that the situation is analogous to Reissner–Nordström.

6.3 The Cauchy problem in general relativity

As we explained already in the introduction, a motivation–perhaps the pricipal one–for studying the wave equation on black hole backgrounds is as this serves as a toy model problem for the non-linear stability properties of the spacetimes themselves as solutions to the Einstein vacuum equations. To discuss this in more detail, however, we first have to introduce the setup of the Cauchy problem for the vacuum equations, and the fundamental well-posedness result.

6.3.1 The initial value problem

6.3.2 The maximal globally hyperbolic Cauchy development

See [30, 11].

6.4 The black hole stability conjecture

We may now give a rough formulation of the black hole stability conjecture, the celebrated open problem whose final resolution would be the culmination of the ideas outlined here.

The conjecture in words: Spacetimes arising from data near Kerr data should remain close to *and dynamically approach* the Kerr family in the exterior-to-the-black-hole region.

Slightly more precisely, we have

Conjecture (Stability of Kerr). *Let $(\Sigma, \bar{g}, K)$ be a vacuum initial data set sufficiently close to the initial data on a Cauchy hypersurface in the Kerr solution $(\mathcal{M}, g_{M_i,a_i})$ for some subextremal parameters $0 \leq |a_i| < M_i$. Then the maximal Cauchy development $(\mathcal{M}, g)$ of the data under evolution by the vacuum equations*

$$\mathrm{Ric}(g) = 0$$

possesses a complete null infinity $\mathcal{I}^+$ such that the metric restricted to $J^-(\mathcal{I}^+)$ remains close to for all time and asymptotically approaches a Kerr solution $(\mathcal{M}, g_{M_f,a_f})$ in a uniform way with quantitative decay rates, where $|a_f| < M_f$ are near a_i, M_i respectively.

Note: $a_i = 0$ will not imply that $a_f = 0$!

References

[1] L. Andersson and P. Blue *Hidden symmetries and decay for the wave equation on the Kerr spacetime*, arXiv:0908.2265

[2] S. Aretakis *Stability and instability of extreme Reissner-Nordstrom black hole spacetimes for linear scalar perturbations I*, Comm. Math. Phys. **307** (2011), 17–63

[3] S. Aretakis *Stability and instability of extreme Reissner-Nordstrom black hole spacetimes for linear scalar perturbations II*, Ann. Henri Poincare **8** (2011), 1491–1538

[4] S. Aretakis *Horizon instability of extremal black holes*, arXiv:1206.6598

[5] S. Aretakis *The characteristic gluing problem and conservation laws for the wave equation on null hypersurfaces*, arXiv:1310.1365

[6] C. Bär, N. Ginoux and F. Pfäffle *Wave equations on lorentzian manifolds and quantization*, arXiv:0806.1036

[7] P. Blue and J. Sterbenz *Uniform decay of local energy and the semi-linear wave equation on Schwarzschild space* Comm. Math. Phys. **268** (2006), 481–504

[8] B. Carter *Black hole equilibrium states*, in Black Holes (Les Houches Lectures), edited B. S. DeWitt and C. DeWitt (Gordon and Breach, New York, 1972).

[9] Y. Choquet-Bruhat *Théorèmes d'existence et d'unicité pour les équations de la gravitation einsteinienne dans le cas non analytique* C. R. Acad. Sci. Paris **230** (1950) 618–620

[10] Y. Choquet-Bruhat *Théoreme d'existence pour certains systèmes d'équations aux dérivées partielles non linéaires* Acta Math. **88** (1952), 141–225

[11] Y. Choquet-Bruhat and R. Geroch *Global aspects of the Cauchy problem in general relativity* Comm. Math. Phys. **14** (1969), 329–335

[12] D. Christodoulou *The action principle and partial differential equations*, Ann. Math. Studies No. 146, 1999

[13] D. Christodoulou *Mathematical Problems of General Relativity Theory I*, ETH-Zürich lecture notes, 2003

[14] D. Christodoulou and S. Klainerman *The global nonlinear stability of the Minkowski space* Princeton University Press, 1993

[15] M. Dafermos *The interior of charged black holes and the problem of uniqueness in general relativity* Comm. Pure Appl. Math. **58** (2005), 445–504

[16] M. Dafermos and I. Rodnianski *A proof of Price's law for the collapse of a self-gravitating scalar field*, Invent. Math. **162** (2005), 381–457

[17] M. Dafermos and I. Rodnianski *The redshift effect and radiation decay on black hole spacetimes* Comm. Pure Appl. Math. 62 (2009), no. 7, 859–919

[18] M. Dafermos and I. Rodnianski *A new physical-space approach to decay for the wave equation with applications to black hole spacetimes*, in XVIth International Congress on Mathematical Physics, P. Exner (ed.), World Scientific, London, 2009, pp. 421433, arXiv:0910.4957

[19] M. Dafermos and I. Rodnianski *Lectures on black holes and linear waves*, in Evolution equations, Clay Mathematics Proceedings, Vol. 17. Amer. Math. Soc., Providence, RI, 2013, pp. 97–205, arXiv:0811.0354

[20] M. Dafermos and I. Rodnianski *The black hole stability problem for linear scalar perturbations*, in Proceedings of the Twelfth Marcel Grossmann Meeting on General Relativity, T. Damour et al (ed.), World Scientific, Singapore, 2011, pp. 132–189, arXiv:1010.5137

[21] S. W. Hawking and G. F. R. Ellis *The large scale structure of space-time* Cambridge Monographs on Mathematical Physics, No. 1. Cambridge University Press, London-New York, 1973

[22] S. Klainerman *Uniform decay estimates and the Lorentz invariance of the classical wave equation* Comm. Pure Appl. Math. **38** (1985), no. 3, 321–332

[23] M. Kruskal *Maximal extension of Schwarzschild metric* Phys. Rev. **119** (1960), 1743–1745

[24] C. Misner, K. Thorne and J. Wheeler *Gravitation*.

[25] C. S. Morawetz *The limiting amplitude principle* Comm. Pure Appl. Math. **15** (1962) 349–361

[26] E. Noether *Invariante Varlationsprobleme* Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse (1918) 235–257

[27] R. Penrose *Gravitational collapse and space-time singularities* Phys. Rev. Lett. **14**, 57–59

[28] E. Poisson and W. Israel *Inner-horizon instability and mass inflation in black holes* Phys. Rev. Lett. **63** (1989), no. 16, 1663–1666

[29] J. Ralston *Solutions of the wave equation with localized energy* Comm. Pure Appl. Math. **22** (1969), 807–823

[30] J. Sbierski *On the Existence of a Maximal Cauchy Development for the Einstein Equations - a Dezornification*, arXiv:1309.7591v2

[31] J. Sbierski *Characterisation of the Energy of Gaussian Beams on Lorentzian Manifolds - with Applications to Black Hole Spacetimes*, arXiv:1311.2477v1

[32] J. Schauder *Das Anfangswertproblem einer quasilinearen hyperbolischen Differential- gleichung zweiter Ordnung in beliebiger Anzahl yon urtabhängigen Veränderliehen* Fundam. Math. **24** (1935)

[33] K. Schwarzschild *Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie* Sitzungsber. d. Preuss. Akad. d. Wissenschaften **1** (1916), 189–196

[34] C. Sogge *Lectures on nonlinear wave equations* International Press, Boston, 1995

[35] J. L. Synge *The gravitational field of a particle* Proc. Roy. Irish Acad. **53** (1950), 83–114

[36] M. Taylor *Partial Differential Equations: Basic Theory*, Springer

[37] R. Wald *General relativity* University of Chicago Press, Chicago, 1984

[38] H. Weyl *Raum, Zeit, Materie* Springer, Berlin, 1919