Black holes and geometrical methods in general relativity

Contents

Chapter 1

Introduction

The purpose of this course is to present in an intuitive but rigourous way the basic mathematical knowledge needed to understand the general theory of relativity with some emphasis on conformal methods, then to give a precise description of the fundamental examples of black hole space-times (Schwarzschild, Reissner-Nordstrøm, De Sitter-Schwarzschild, Kerr), and to present some applications of conformal techniques to the analysis of the behaviour of test fields on the spacetimes of general relativity : conformal scattering and also a complete treatment of an important question that until today has been open for more than 40 years, the peeling on the Schwarzschild metric.

General relativity is a geometrical theory of gravity. There are three essential principles that rule the theory :

1. material objects and fields cannot travel at speeds greater than that of light ;
2. the notion of simultaneity or even of an event being in the past of another depends on the observer (this is the reason for the name "relativity"), space and time are united in a single object called spacetime and no longer have an independent existence ;
3. gravity is itself a field that cannot travel faster than the speed of light, it is encoded in the theory as a geometrical quantity, a part of the curvature of the spacetime (it is a quantity defined locally from the knowledge of the local geometry of the universe), and it is determined by field equations that propagate the gravitational field, the Einstein equations.

With merely the first two principles, we have the theory of special relativity, which is not a theory of gravity.

The background of General Relativity in its original form is a four-dimensional spacetime, that is to say a manifold $\mathcal{M}$ of real dimension 4, describing the universe of space and time, or, if one prefers, although it is not fully in agreement with the spirit of general relativity, the evolution of a three dimensional universe throughout time. Since then, many other cases have been considered, universes with only space and no time or with a ludicrous amount of spatial dimensions for one time. We will not be concerned with these cases here. This course considers only the four-dimensional spacetimes that form the framework of what is now sometimes referred to as classical general relativity : Einstein's theory. To understand the notion of spacetime sufficiently well so as to work with it, we need to define properly a fair amount of geometrical concepts that are essential to the theory, among which metrics, connections, and the curvature tensor. Only a basic knowledge of differential geometry and tensor calculus will be assumed.

A bit of history.

1850 Mitchell, Laplace, dark star.

1905 Special Relativity.

1915 General Relativity.

1917 Schwarzschild metric (see chapter 5).

1917 Cosmological constant. Einstein introduces this extra term in his equations after realizing that the original form of his theory does not allow for a static universe (unless it is flat). The cosmological constant induces a repulsive force which he adjusted to counterbalance gravity, thus obtaining a static universe : the Einstein cylinder.

1919 Eddington's expedition confirms the deflection of light by the sun. Eddington travelled to the island of Principe off the West coast of Africa, to observe a total eclipse of the sun. Stars that could be seen near the sun at that time appeared to have shifted from their usual position with respect to the other stars, thus confirming the prediction of general relativity that the gravitational field of the sun should deflect light rays. The effect can only be seen during a total eclipse ; under normal conditions, the light of the sun is too bright and prevents "close" enough stars from being seen.

1924 Eddington discovers the Eddington-Finkelstein coordinates, re-discovered by Finkelstein in 1958 (see section 5.4.1).

1929 Hubble discovers the red-shift effect, thus proving the expansion of the universe. The cosmological constant is consequently abandoned as a reasonable model for physics, though it is still the object of mathematical studies. Einstein calls it his greatest mistake.

1939 Oppenheimer-Snyder model for the collapse of a star.

1960 Kruskal and Szekeres discover independently the Kruskal-Szekeres coordinates (see section 5.4.2).

1963 Kerr metric (see chapter 7).

1967 John Wheeler is credited with having coined the term "black hole" at a conference that year. He insists that someone else did. It seems to have appeared first in 1964 in a letter by Anne Ewing to the AAAS.

1967 Boyer-Lindquist coordinates.

1970 Hawking-Penrose singularity theorem. It establishes mathematically the existence of black holes as necessary in the framework of general relativity and by reversing time also entails the existence of a singularity at the origin of the universe : the big bang.

1998 The expansion of the universe is observed to be faster than expected. This will lead to the re-introduction of the cosmological constant in the Einstein equations as a reasonable model for the physics of the universe.

Chapter 2

Basic geometrical concepts

2.1 Submanifolds, manifolds, tensors

Definition 2.1. *A smooth submanifold of $\mathbb{R}^n$ of dimension $k \in \{1, ..., n\}$ is a subset S of $\mathbb{R}^n$ such that, for any point $p_0 \in S$, there exists V a neighbourhood of p_0 in $\mathbb{R}^n$, U a neighbourhood of* 0 *in $\mathbb{R}^n$ and $\phi : U \to V$ a $\mathcal{C}^\infty$ diffeomorphism such that $\phi(0) = p_0$ et*

$$S \cap V = \left\{ p = \phi(q)\,;\ q = (x^1, .., x^k, 0, ..., 0) \in U \right\} ,$$

i.e. it is a subset of $\mathbb{R}^n$ that can locally be straightened smoothly as a k-dimensional plane. This is called a local chart of S.

The topology of $\mathbb{R}^n$ induces naturally a topology on S by restriction to S of open sets of $\mathbb{R}^n$.

Such an object can be understood without reference to the ambient space $\mathbb{R}^n$, it is then referred to as a k-dimensional smooth manifold. The concept of submanifold of $\mathbb{R}^n$ is sufficient thanks to a theorem by H. Whitney in 1936 [26], stating that a manifold of dimension d can be embedded in $\mathbb{R}^{2d+1}$, i.e. realized as a submanifold of dimension d of $\mathbb{R}^{2d+1}$. With the definition above, we can easily introduce the notion of tangent and cotangent space to a given submanifold of $\mathbb{R}^n$.

The vectors

$$\frac{\partial \phi}{\partial x^1}(q)\,,\ \ldots,\ \frac{\partial \phi}{\partial x^k}(q)\,,$$

are tangent to S at p and generate a k-dimensional subspace of $\mathbb{R}^n$ since ϕ is a diffeomorphism. This subspace is denoted $T_p(S)$ and called the tangent space to S at p.

We denote by $T_p^*(S)$ and call cotangent space to S at p the dual of $T_p(S)$, i.e. the space of continuous linear forms acting on $T_p(S)$. The elements of $T_p^*(S)$ are called co-vectors at p.

Definition 2.2 (Tangent bundle, cotangent bundle)**.** *We denote by TS (resp. T^*S) and call tangent bundle (resp. cotangent bundle) the set of pairs (p, X) where $X \in T_pS$ (resp. $X \in T_p^*S$). Both are smooth manifolds of dimension $2k$ (it is very easy to realize them as smooth submanifolds of $\mathbb{R}^{2n}$ of dimension $2k$, using local charts of S, TS can be trivialized locally in p and globally in X as $\Omega \times \mathbb{R}^k$ where Ω is an open set of $\mathbb{R}^k$, it is $\phi^{-1}(S \cap V)$*

for the local chart described above). Explicitely, for a given local chart $\phi : U \to V$ *for* S*, denoting* $D\phi(q)$ *the differential of* ϕ *at* q*, the map*

$$\psi : U \times R^n \to X \times \mathbb{R}^n , \; \psi(q, X) = (\phi(q), D\phi(q)(X))$$

is a local chart for TS *and*

$$\begin{aligned} &\{(p, X) \in TS ; \; p \in U \cap S\} \\ = &\{\psi(q, Y) ; \; q \in U , \; Y \in \mathbb{R}^n , \; q = (q_1, q_2, ..., q_k, 0, ..., 0) , \; X = (X^1, X^2, ..., X^k, 0, ..., 0)\} ; \end{aligned}$$

the map

$$\chi : U \times R^n \to V \times \mathbb{R}^n , \; \psi(q, X) = \left(\phi(q), \left((D\phi(q))^*\right)^{-1} (V)\right)$$

is a local chart for T^*S *and*

$$\begin{aligned} &\{(p, X) \in T^*S ; \; p \in U \cap S\} \\ = &\{\chi(q, Y) ; \; q \in U , \; Y \in \mathbb{R}^n , \; q = (q_1, q_2, ..., q_k, 0, ..., 0) , \; X = (X_1, X_2, ..., X_k, 0, ..., 0)\} . \end{aligned}$$

Definition 2.3 (Vector fields, 1-forms)**.** *A vector field is a function that to each point* $p \in S$ *associates a vector at* p*, i.e. an element of* T_pS*. The graph of this function is a subset of* TS *that is referred to as a section of* TS*. The vector field and the section (the function and its graph) are usually identified. Using local charts, a vector field can be understood locally as a function from* S *(or even from an open set of* $\mathbb{R}^k$*) to* $\mathbb{R}^k$ *and it is therefore straightforward to talk about the regularity of such an object. Instead of the usual notation* $\Gamma(TS)$ *for the sections of* TS*, we shall use a notation that makes clear the regularity of the fields we consider, for example* $\mathcal{C}^\infty(S\,;\; TS)$ *will denote the set of smooth sections of* TS*,* $\mathcal{D}'(S\,;\; TS)$ *will denote the space of vector-valued distributions on* S*. We shall also consider sections with a regularity defined by Sobolev or Hölder spaces.*

A 1-form is defined in a similar way as a section on T^*S*, we shall use similar notations, such as for example* $H^k_{\mathrm{loc}}(S\,;\; T^*S)$*.*

Now we can consider vectors (resp. vector fields) as forms acting on co-vectors (resp. 1-forms). In other words, we can trivially identify TS and $T^{**}S$. The advantage of this is that the notion of tensor product is then very easily defined, and hence so are tensors and tensor fields.

Definition 2.4 (Tensor at a point)**.** *We define tensors at a point in an inductive manner as follows. Unfortunately there is no pleasant unambiguous notation for the space of tensors of a given valence at a point, but such notations will exist when talking about tensor fields.*

- *First we call vectors at a point* $p \in S$ *tensors at* p *of valence* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *and covectors at* p *will be referred to as tensors at* p *of valence* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$*. Tensors of valence* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *are then understood as linear forms on tensors of valence* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *and vice versa.*

- *Given two tensor at p of valence* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$*, say* X *and* Y*, we define their tensor product as the bilinear form on the space of tensors of tensors of valence* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *at* p

$$X \otimes Y \ : \ (\alpha, \ \beta) \mapsto \alpha(X) \times \beta(Y) \, .$$

The space of tensors of valence $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ *at* p *is then defined as the space of finite linear combinations of all such objects, it is the space of bilinear forms on* T_p^*S*.*

We can define in a similar way tensors at p *of valence* $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ *as finite linear combinations of the tensor products of two tensors of valence* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *(i.e. bilinear forms on* T_pS*), and tensors at* p *of valence* $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *as finite linear combinations of the tensor products of a tensor of valence* $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ *and another of valence* $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ *(i.e. bilinear forms on* $T_pS \times T_p^*S$*). There should be two notions of tensors of valence* $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *depending on whether we put a vector or a covector first in the tensor products, but the commutativity of the product in* $\mathbb{R}$ *gives a canonical identification between the two notions.*

- *Tensors of any given valence* $\begin{bmatrix} m \\ n \end{bmatrix}$*,* $m, n \in \mathbb{N}$ *at* p *can be defined analogously : we consider the tensor product of* m *vectors* $V_1, \, ..., \, V_m$ *and* n *covectors* $X_1, \, ..., \, X_n$ *at* p*, denoted* $V_1 \otimes ... \otimes V_m \otimes X_1 \otimes ... \otimes X_n$ *as the* $m+n$ *multi-linear forms acting on* m *covectors* $Y_1, \, ..., \, Y_m$ *and* n *vectors* $W_1, \, ..., \, W_n$ *as follows*

$$\begin{aligned} & V_1 \otimes ... \otimes V_m \otimes X_1 \otimes ... \otimes X_n \, (Y_1, ..., Y_m, W_1, ..., W_n) \\ & \quad = Y_1(V_1)...Y_m(V_m) X_1(W_1)...X_n(W_n) \, . \end{aligned}$$

The tensor bundle of valence $\begin{bmatrix} m \\ n \end{bmatrix}$ *is the space of finite linear combinations of such tensor products, it is the space of the* $m+n$ *multi-linear forms acting on* m *copies of* T_p^*S *and* n *copies of* T_pS*.*

Definition 2.5 (Tensor fields)**.** *Doing a similar construction with vector fields and* 1*-forms instead of vectors and co-vectors at a point, we obtain the notion of tensor fields of valence* $\begin{bmatrix} m \\ n \end{bmatrix}$*. Such tensor fields are sections of a vector bundle referred to as the tensor bundle of valence* $\begin{bmatrix} m \\ n \end{bmatrix}$ *: it is analogous to the tangent bundle but instead of the tangent space, we attach to each point the space of tensors of valence* $\begin{bmatrix} m \\ n \end{bmatrix}$ *at this point. The tensor bundle refers to the collection of all the tensor bundles of given valence. The abstract index* *formalism will provide us with convenient ways of denoting the tensor bundles of a given valence*[1].

For tensors of valence $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ or $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$, the notion of symmetry is straightforward. For a tensor α of valence $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$, it goes as follows : α is said to be symmetric is for any vectors V, W, we have $\alpha(V, W) = \alpha(W, V)$; α is said to be anti-symmetric is for any vector fields V, W, we have $\alpha(V, W) = -\alpha(W, V)$. The notion is analogous for a tensors of valence $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ and extends naturally to fields of such tensors. For a tensor of type $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ however, there is no notion of symmetry since the arguments cannot be exchanged.

For a tensor T of valence $\begin{bmatrix} m \\ n \end{bmatrix}$, we can consider a set of $m!n!$ tensors of the same valence obtained from T by symmetry operations, consisting of permutations of the set of vectors and permutations of the set of covectors to which it is applied. Permutations between vectors and co-vectors are not allowed for the reason seen above. If among these $m!n!$ tensors, some of them are equal, then T is said to have some symmetries. If they are all equal, T is said to be totally symmetric. If they are all equal to T multiplied by the product of the signatures of the permutations, T is said to be totally antisymmetric. Intermediate situations are numerous, such as tensors that are totally symmetric in their vector arguments and totally anti-symmetric in the co-vector arguments.

Tensor fields of valence $\begin{bmatrix} 0 \\ p \end{bmatrix}$ are called p-forms. A differential p-form will be a totally antisymmetric tensor fields of valence $\begin{bmatrix} 0 \\ p \end{bmatrix}$.

Definition 2.6 (Bases)**.** *A basis of TS, also referred to as a frame, is a family of k vector fields V_i, $i = 1, ..., k$, on S such that at each point $p \in S$, $\{V_1(p), V_2(p), ..., V_k(p)\}$ is a basis of $T_p(S)$. To such a basis is associated a dual basis of 1-forms $\{\alpha^i\}_{i=1,...,k}$ such that $\alpha^i(V_j) = \delta^i_j$, i.e. is equal to 1 if $i = j$ and to 0 otherwise. These then induce bases of all the bundles of tensors of a given valence by tensor product. For instance, $\{\alpha^i \otimes \alpha^j \otimes V_l\}_{i,j,l=1,...,k}$ is a basis of the tensor bundle of valence $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. A given tensor field can then be described by its components in the relevant basis, that will be referred to as components in the basis $\{V_i\}_{i=1,...,k}$, since all the bases of the tensor bundles stem from this original one. For example, for a tensor T of valence $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, we shall denote*

$$T = \sum_{i=1}^{k} \sum_{j=1}^{k} \sum_{l=1}^{k} T^{ij}_l V_i \otimes V_j \otimes \alpha^l \, . \tag{2.1}$$

Note that the frame $\{V_i\}_{i=1,...,k}$ also induces bases on tensor sub-bundles with symmetries. For instance, $\{\alpha^i \otimes \alpha^j \otimes V_l + \alpha^j \otimes \alpha^i \otimes V_l\}_{i,j,l=1,...,k}$ is a basis of the symmetric tensor

[1] The degree of convenience of such notations is relative to the observer but also to the use to which the notation is put, they have the advantage of labelling clearly the type of quantities one deals with and of allowing calculations as explicit as when using bases, but while remaining totally intrinsic.

bundle of valence $\left[\begin{array}{c}1\\2\end{array}\right]$.

Remark 2.1 (Einstein convention). *The Einstein notational convention says that if an index is repeated in an expression, appearing once up and once down, then it is "contracted", i.e. the sum is taken oven all the numerical values of the index. With this convention, the expression (2.1) becomes*

$$T = T_l^{ij} V_i \otimes V_j \otimes \alpha^l \,.$$

We shall systematically use this convention.

Sometimes, we may consider a frame that is only defined locally on an open set $\mathcal{U}$ of S. We shall then refer to it as a local frame, sometimes a local frame over $\mathcal{U}$.

An important example of local frame is that associated to a coordinate system. In such a case, the dual basis is naturally yielded by the coordinate system, then the frame is obtained by duality.

Definition 2.7 (Coordinate bases). *Consider a local coordinate system $x^1, x^2, ..., x^k$ on an open set $\mathcal{U}$ of S. The family of 1-forms $\{\mathrm{d}x^1, \mathrm{d}x^2, ..., \mathrm{d}x^k\}$ is by definition a local basis of T^*S over $\mathcal{U}$, i.e., at each point $p \in \mathcal{U}$, $\{\mathrm{d}x^1, \mathrm{d}x^2, ..., \mathrm{d}x^k\}$ is a basis of T_p^*S. Its dual basis is a local frame over $\mathcal{U}$ denoted $\{\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, ..., \frac{\partial}{\partial x^k}\}$.*

Such a local coordinate system is usually obtained via a local chart from the local coordinates $\{x^1, ..., x^k)$ on $\phi^{-1}(V \cap S)$. The 1-forms $\mathrm{d}x^1$, ..., $\mathrm{d}x^k$ then refer to the coordinate 1-forms on U that act on vector fields on $V \cap S$ pulled back to $\phi^{-1}(V \cap S)$ by the diffeomorphism ϕ. More precisely, the notation $\mathrm{d}x^i(W)(p)$ for a vector field W on $V \cap S$ and $p = \phi(q) \in V \cap S$ really means $\mathrm{d}x^1\left((D\phi(q))^{-1}(V) \right)$.

The notation of the coordinate basis vectors as differential operators reveals an important identification that is always made between vector fields and differential operators. The idea behind this identification is that vector fields define a flow and following geometrical objects along this flow, we can defferentiate them with respect to the parameter of the flow. This is the notion of Lie derivative that we shall encounter very soon.

Local frames are in fact more common than global ones. An important example is given by spherical coordinates on $\mathbb{R}^3$. The frame associated with the coordinate system is not globally defined. The coordinate system and the frame associated with it are singular on the North-South axis.

2.2 Abstract index formalism

Projecting tensors onto local bases is very useful for doing explicit calculations. The disadvantage of such calculations is that sometimes, they depend on the basis chosen. The intrinsic aspect of the result is therefore often a problem. However, in many cases, the advantage of a local basis is purely notational, in keeping track of the indices. This is what led Roger Penrose to developing the abstract index formalism. A complete axiomatic description of this set of notations is given in Spinors and space-time Vol.1 [24]. We simply intend to give a flavour of the essential idea here in order to be able to use this formalism for explicit calculations.

Abstract indices

Consider T a tensor field of valence $\begin{bmatrix} m \\ n \end{bmatrix}$. We shall denote T with indices, m up and n down, in order to be able to see the nature of this object purely from the way it is denoted. The indices used are always lower case lightface latin letters[2], possibly with indices themselves. Here for instance, we would do well to use a notation like

$$T^{a_1 a_2 \dots a_m}_{b_1 b_2 \dots b_n} .$$

For the moment, the respective position of an index that is up and another that is down is unimportant, for the reason already mentionned earlier that the product on $\mathbb{R}$ is commutative, therefore there is no reason a priori to distinguish between $\alpha \otimes V$ and $V \otimes \alpha$, where α is a 1-form and V a vector field. So we write the up and down indices above one another.

It is important to understand that the notation above does not refer to a collection of components in referrence to a basis. It is the intrinsic tensor field to which we have just put some stickers to see how many legs up and down it has[3]. The tensor T has m 1-form arguments and n vector field arguments. Suppose we wish to express

$$T(\alpha, \beta, ..., \gamma, U, V, ..., W) \tag{2.2}$$

with abstract indices, we shall denote the 1-forms with an index down, since they are tensor fields of valence $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and the vector fields with an index up since they are tensor fields of valence $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Then the notation for (2.2) will be

$$T^{a_1 a_2 \dots a_m}_{b_1 b_2 \dots b_n} \alpha_{a_1} \beta_{a_2} \dots \gamma_{a_m} U^{b_1} V^{b_2} \dots W^{b_n} .$$

There is no sum over indices of course since these are not indices that take numerical values, purely labels. The fact that an index is present once up and once down in the same expression means that a contraction has to take place, this denotes the action of one "leg" of the tensor on a vector or a 1-form. This is the abstract index version of the Einstein convention. The order of the factors in the above expression is irrelevent, the repeated indices simply tell us in what slot a vector or a 1-form should be contracted.

The tensor bundles of a given valence can be denoted with abstract indices too, for example $T_a S$ denotes T^*S and $T^{ab}_c S$ is the tensor bundle of valence $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

The link between the quantities with indices and without indices is formally realized by objects denoted $\mathrm{d}x^a$ and $\frac{\partial}{\partial x^a}$. For instance,

$$V = V^a \frac{\partial}{\partial x^a} , \ \alpha = \alpha_a \mathrm{d}x^a , \ T = T^a_{bc} \frac{\partial}{\partial x^a} \otimes \mathrm{d}x^b \otimes \mathrm{d}x^c .$$

[2]In Penrose's abstract index formalism, indices denoted by upper case latin letters are for spinors and indices denoted by greek letters are for twistors. As for boldface indices, they are concrete indices with referrence to a basis.

[3]Indeed, a more abstract set of notations has been developed by Penrose, consisting purely of diagrams with legs. See [24] for a description of the "legged diagram" formalism.

This looks like a decomposition on a basis, but the indices are all abstract, this is purely a formal link between indexed and non indexed quantities. This type of link is required for the coherence of some expressions. Typically, if we integrate a 1-form on a curve, we wish to obtain a scalar, hence without an index, so it is clear that the expression

$$I = \int_{\mathcal{C}} \alpha_a \,,$$

is inadequate. Instead, the following expression should be used

$$I = \int_{\mathcal{C}} \alpha_a \mathrm{d}x^a \,.$$

Another good reason for using these $\mathrm{d}x^a$ and $\frac{\partial}{\partial x^a}$ conventions is that most expressions should be the same with abstract indices or with concrete indices referring to a basis.

Symmetrizers and anti-symmetrizers

The symmetry operations on a tensor can now be expressed explicitely. If we swap two indices (they have to be both up or both down for this to be legitimate), this means that when applying the tensor to 1-forms and vectors, we shall swap the corresponding arguments. The symmetry operations known as symmetrizers are denoted by parentheses on each side of the group of indices it applies to, and anti-symmetrizers are denoted by square brackets. For example

$$\begin{aligned} T^a_{(bc)d} &= \frac{1}{2}\left(T^a_{bcd} + T^a_{cbd}\right) \,, \\ N^{a[bc]d}_{ef} &= \frac{1}{2}\left(N^{abcd}_{ef} - N^{acbd}_{ef}\right) \,, \\ K_{[abc]} &= \frac{1}{6}\left(K_{abc} + K_{bca} + K_{cab} - K_{bac} - K_{acb} - K_{cba}\right) \,. \end{aligned}$$

If we wish to exclude an index or a group of indices from a symmetry operation, we put them between vertical bars, such as

$$T^{ab}_{(c|de|f)g} = \frac{1}{2}\left(T^{ab}_{cdefg} + T^{ab}_{fdecg}\right) \,.$$

Concrete indices

Concrete indices refer to a given basis and label the components of tensors with respect to this basis, they take numerical values. They are denoted by boldface lower case latin letters. They also label the basis vectors and 1-forms. For instance, a frame $\{V_1, ..., V_k\}$ will be denoted $\{V_{\mathbf{a}}\}_{\mathbf{a}=1,...,k}$ and the dual basis of 1-forms $\{\alpha^{\mathbf{a}}\}_{\mathbf{a}=1,...,k}$. As indexed objects, the basis vectors are denoted $V^a_{\mathbf{a}}$ and the 1-forms $\alpha^{\mathbf{a}}_a$, i.e. we can write

$$V_{\mathbf{a}} = V^a_{\mathbf{a}} \frac{\partial}{\partial x^a} \,.$$

There is no contraction possible between a concrete index and an abstract index, they are objects of different natures.

The decomposition of a vector or a 1-form in the basis is written as

$$W^a = W^{\mathbf{a}} V_{\mathbf{a}}^a \text{ or } W = W^{\mathbf{a}} V_{\mathbf{a}} ,$$
$$\beta_a = \beta_{\mathbf{a}} \alpha_a^{\mathbf{a}} \text{ or } \beta = \beta_{\mathbf{a}} \alpha^{\mathbf{a}} .$$

For handwriting, boldface letters are not exactly natural, instead we shall underline the indices to signify that they are concrete indices.

Remark 2.2. *There is no perfect notation. The abstract index formalism has advantages for some aspects, and inevitable drawbacks. In some cases, it becomes too heavy and abusive notations are then sometimes used. When resorting to such, we shall endeavour to point it out.*

2.3 Metrics

Definition 2.8 (metric)**.** *Let $\mathcal{M}$ be a smooth manifold of dimension n.*

1. *A metric on $\mathcal{M}$ is a symmetric 2-form on $\mathcal{M}$ (equivalently, a symmetric tensor field of valence $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$). We shall always assume a metric on $\mathcal{M}$ to be at least continuous on $\mathcal{M}$.*

2. *Consider a metric g on $\mathcal{M}$. We say that a local frame $\{V_1, .., V_n\}$ on an open set $\mathcal{U}$ of $\mathcal{M}$ is orthonormal for g if*

$$g(V_{\mathbf{i}}, V_{\mathbf{j}}) = \begin{cases} 0 & \text{if } \mathbf{i} \neq \mathbf{j} , \\ \pm 1 & \text{if } \mathbf{i} = \mathbf{j} . \end{cases}$$

3. *We say that a metric g on $\mathcal{M}$ is non degenerate if, for any point p of $\mathcal{M}$ there exists a neighbourhood $\mathcal{U}$ of p and an orthonormal local frame $\{V_1, .., V_n\}$ on $\mathcal{U}$.*

4. *We say that a metric g on $\mathcal{M}$ has signature $(+ + ... + - ... -)$ with k "+" and $n - k$ "−" if, for any point p of $\mathcal{M}$ there exists a neighbourhood $\mathcal{U}$ of p and an orthonormal local frame $\{V_1, .., V_n\}$ on $\mathcal{U}$ such that for exactly k values of $\mathbf{i} \in \{1, ..., n\}$ we have $g(V_{\mathbf{i}}, V_{\mathbf{i}}) = 1$ and for exactly $n - k$ values of $\mathbf{i} \in \{1, ..., n\}$ we have $g(V_{\mathbf{i}}, V_{\mathbf{i}}) = -1$. Such metrics are of course non degenerate. Note also that for a non degenerate continuous metric on a connected manifold $\mathcal{M}$, the signature is unambiguously and globally defined on $\mathcal{M}$.*

Definition 2.9. *A metric with signature $(+...+)$ is called riemannian (the case of signature $(-...-)$ is rarely considered as such and usually identified with the riemannian case). When the signature contains "+" and "−" signs, the metric is said to be pseudo-riemannian or semi-riemannian. When there is only one "+" and $n - 1$ "−" signs, the metric is said to be Lorentzian (the case $(- + ...+)$ is also referred to as Lorentzian by many authors, the choice of convention $(+ - ...-)$ or $(- + ...+)$ is purely a matter of personal taste).*

Examples. 1. *Euclidian metric on $\mathbb{R}^3$. It is expressed in cartesian coordinates as*

$$g = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2$$

and acts on vectors at a point or vector fields on $\mathbb{R}^3$ *as follows*

$$V = V^1\frac{\partial}{\partial x} + V^2\frac{\partial}{\partial y} + V^3\frac{\partial}{\partial z}\,,\ W = W^1\frac{\partial}{\partial x} + W^2\frac{\partial}{\partial y} + W^3\frac{\partial}{\partial z}\,,$$
$$g(V,W) = V^1W^1 + V^2W^2 + V^3W^3\,.$$

This can be understood in terms of matrices as

$$g(V,W) = \begin{pmatrix} V^1 & V^2 & V^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W^1 \\ W^2 \\ W^3 \end{pmatrix}\,,$$

where the 3×3 *matrix above is the matrix of* g *in the coordinate basis* $\{x,y,z\}$*. The signature of* g *is* $(+++)$*.*

2. *On* $\mathbb{R}^2$*, we consider the metric expressed in cartesian coordinates as*

$$g = 2\mathrm{d}x\mathrm{d}y\,,$$

where $\mathrm{d}x\mathrm{d}y$ *denotes the symmetric product*

$$\mathrm{d}x\mathrm{d}y = \frac{1}{2}\left(\mathrm{d}x\otimes\mathrm{d}y + \mathrm{d}y\otimes\mathrm{d}x\right)\,.$$

Its action on vectors at a point or vector fields on $\mathbb{R}^2$ *is described as*

$$g(V,W) = \begin{pmatrix} V^1 & V^2 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} W^1 \\ W^2 \end{pmatrix} = V^1W^2 + V^2W^1\,.$$

This is a Lorentzian metric of signature $(+-)$*. Putting* $x = u+v$ *and* $y = u-v$*,* g *takes the new expression*

$$g = 2\mathrm{d}u^2 - 2\mathrm{d}v^2$$

which makes its Lorentzian signature more explicit.

Definition 2.10 (Spacetime)**.** *A spacetime is a 4-dimensional connected smooth manifold* $\mathcal{M}$ *endowed with a metric* g *of Lorentzian signature. By convention, when dealing with bases on a spacetime, the basis vectors will be numbered from 0 to 3, when one of them is timelike and the others spacelike, the timelike one will receive the label 0 (for the notion of timelike and spacelike vectors on a spacetime, see definition 4.1).*

Definition 2.11 (The metric as an index raising and lowering operator)**.** *Consider a spacetime* $(\mathcal{M},g)$*. To a vector* V^a *at a point we can associate a covector by contracting* V^a *into the metric at that point. We denote by* V_a *the covector thus obtained*

$$V_a = V^b g_{ab}\,.$$

Since the metric is a non degenerate symmetric 2-form, this operation is an isomorphism between vectors and covectors. We denote by g^{ab} *the inverse operator, i.e.*

$$V^a = g^{ab}V_b\,.$$

Then g^{ab} *is a symmetric tensor of valence* $\left[\begin{array}{c}2\\0\end{array}\right]$ *and by construction we have*

$$g_{ab}g^{bc} = \delta_a^c \,,$$

where δ_a^c *is merely an operator that replaces the index a by the index c, i.e. it transforms a vector field into the same vector field but with the index denoted by another letter. In terms of concrete indices,* $g_{\mathbf{ab}}$ *will be the matrix of the metric in the chosen basis,* $g^{\mathbf{ab}}$ *will be the inverse matrix and* $\delta_{\mathbf{a}}^{\mathbf{c}}$ *is the usual Kronecker symbol, that is* 1 *is* $\mathbf{a} = \mathbf{c}$ *and* 0 *otherwise.*

Remark 2.3. *As soon as we start raising and lowering indices using the metric, we realize that the respective position of up and down indices may have some importance after all. Typically we want to avoid the following absurdity*

$$g^{cf}T^{ab}_{cde} = T^{abf}_{de} \ , \ g_{fc}T^{abf}_{de} = T^{ab}_{dec} \text{ and therefore } T^{ab}_{cde} = T^{ab}_{dec} \,,$$

which looks like a symmetry property whereas it should just be $T = T$*. Hence, in some cases where we wish to keep track of indices though raising and lowering operations, we will order all indices, irrespective of their position up or down. We will have notations like*

$$g_{ai}T^{a}{}_{bc}{}^{de}{}_{f} = T_{ibc}{}^{de}{}_{f} \,.$$

Definition 2.12 (Dual bases and index raising)**.** *Consider a spacetime* $(\mathcal{M}, g)$ *and a (possibly local) frame* $\{V^a_{\mathbf{a}}\}_{\mathbf{a}=0,1,2,3}$*. We consider the four* 1*-forms* $V^{\mathbf{a}}_a$ *defined by*

$$V^{\mathbf{a}}_a = g^{\mathbf{ab}}g_{ab}V^b_{\mathbf{b}} \,.$$

This is at each point a basis of covectors satisfying

$$V^{\mathbf{a}}_a V^a_{\mathbf{b}} = \delta^{\mathbf{a}}_{\mathbf{b}} \,,$$

i.e. it is the (local) basis of 1*-forms dual to* $\{V^a_{\mathbf{a}}\}_{\mathbf{a}=0,1,2,3}$*.*

Remark 2.4. *In particular in a coordinate basis, if* $V^{\mathbf{a}}_a \mathrm{d}x^a = \mathrm{d}x^{\mathbf{a}}$*, then* $V^a_{\mathbf{a}}\frac{\partial}{\partial x^a} = \frac{\partial}{\partial x^{\mathbf{a}}}$*.*

Remark 2.5. *With these notations, the components of a tensor field in the basis* $\{V^a_{\mathbf{a}}\}_{\mathbf{a}}$ *are given by, for example,*

$$T^{\mathbf{a}}_{\mathbf{bc}} = V^{\mathbf{a}}_a V^b_{\mathbf{b}} V^c_{\mathbf{c}} T^a_{bc} \,.$$

Useful notations for basis vectors and covectors are $g^a_{\mathbf{a}}$ *and* $g^{\mathbf{a}}_a$*.*

2.4 Connections, tortion and curvature

Definition 2.13. *We define the gradient operator*

$$\partial_a \,:\, \mathcal{D}'(\mathcal{M}\,;\,\mathbb{R}) \to T_a\mathcal{M}\,;$$

by

$$\partial_a f \mathrm{d}x^a = \mathrm{d}f \,,$$

i.e. the gradient operator is just the exterior derivative acting on functions. This is an intrinsic object depending neither on the metric nor on a choice of coordinate system.

A connection is an extension of the gradient to tensor fields which satisfies two natural properties :

Definition 2.14. *A connection ∇_a is an extension of the gradient to arbitrary tensor fields, such that :*

1. *it is linear from any tensor bundle F of given valence to $T^*\mathcal{M} \otimes F$;*
2. *it satisfies the Leibnitz rule.*

Theorem 2.1. *There exists a unique connection ∇_a such that :*

1. *it is tortion-free, meaning that $[\nabla_a\,,\ \nabla_b]\, f = 0$ for any scalar field f, where $[\nabla_a\,,\ \nabla_b]$ is the commutator of ∇_a and ∇_b, $[\nabla_a\,,\ \nabla_b] = \nabla_a \nabla_b - \nabla_b \nabla_a$;*
2. *it commutes with the metric, i.e. $\nabla_a g_{bc} = 0$ and $\nabla_a g^{bc} = 0$.*

It is called the Levi-Civita connection.

Proof. We start with uniqueness, existence will be a trivial consequence. Working with a local coordinate basis, we denote by $V_{\mathbf{a}}$, or $\frac{\partial}{\partial x^{\mathbf{a}}}$, or $\partial_{\mathbf{a}}$ the basis vectors as well as partial derivation with respect to the coordinate $x^{\mathbf{a}}$. The notation $\nabla_{\mathbf{a}}$ with a concrete index refers to

$$\nabla_{\mathbf{a}} = V^a_{\mathbf{a}} \nabla_a\,,$$

i.e. covariant differentiation in the direction $V_{\mathbf{a}}$. Let us first consider the action of ∇_a on a 1-form ω_b. If we decompose the 1-form ω on the basis $\mathrm{d}x^{\mathbf{a}}$ (or $V^{\mathbf{a}}_a$), dual to $\partial_{\mathbf{a}}$ (or $V^a_{\mathbf{a}}$),

$$\omega = \omega_{\mathbf{b}} \mathrm{d}x^{\mathbf{b}} \text{ or } \omega_b = \omega_{\mathbf{b}} V^{\mathbf{b}}_b\,.$$

using the linearity of ∇_a, we must have

$$\nabla_{\mathbf{a}} \omega_b = \left(\nabla_{\mathbf{a}} \omega_{\mathbf{b}}\right) V^{\mathbf{b}}_b + \omega_{\mathbf{b}} \nabla_{\mathbf{a}} V^{\mathbf{b}}_b\,,$$

Since for each $\mathbf{b}$, $\omega_{\mathbf{b}}$ is a scalar function on $\mathcal{M}$, by the assumption that ∇_a is an extension of the gradient operator, the action of $\nabla_{\mathbf{a}}$ on $\omega_{\mathbf{b}}$ reduces to the partial derivation of $\omega_{\mathbf{b}}$ in the direction or $V^a_{\mathbf{a}}$, i.e.

$$\nabla_{\mathbf{a}} \omega_{\mathbf{b}} = V^a_{\mathbf{a}} \partial_a \omega_{\mathbf{b}} = \partial_{\mathbf{a}} \omega_{\mathbf{b}}\,.$$

Hence the action of $\nabla_{\mathbf{a}} - \partial_{\mathbf{a}}$ on a 1-form is described by a zero order linear operator defined as follows by the action of $\nabla_{\mathbf{a}}$ on the basis 1-forms :

$$\left(\nabla_{\mathbf{a}} - \partial_{\mathbf{a}}\right) \omega_b = \left(\nabla_{\mathbf{a}} V^{\mathbf{c}}_b\right) \omega_{\mathbf{c}}$$

and if we project the equality above on the basis 1-form $\mathrm{d}x^{\mathbf{b}}$, we obtain

$$V^b_{\mathbf{b}} \left(\nabla_{\mathbf{a}} - \partial_{\mathbf{a}}\right) \omega_b = \left(\nabla_{\mathbf{a}} V^{\mathbf{c}}_b\right) \omega_{\mathbf{c}} V^b_{\mathbf{b}}\,. \tag{2.3}$$

The coefficients $V^b_{\mathbf{b}} \nabla_{\mathbf{a}} V^{\mathbf{c}}_b$ are referred to as the Christoffel symbols and denoted $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$. We see that $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$ is the projection on $\mathrm{d}x^{\mathbf{b}}$ of the action of $\nabla_{\mathbf{a}}$ on $\mathrm{d}x^{\mathbf{c}}$.

Remark 2.6. *In most geometry textbooks, the expression (2.3) is simply written*

$$\nabla_{\mathbf{a}}\omega_{\mathbf{b}} = \partial_{\mathbf{a}}\omega_{\mathbf{b}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\omega_{\mathbf{c}}\,. \tag{2.4}$$

This notation is dreadful and wrong! *Indeed, $\omega_{\mathbf{b}}$ is a scalar, it is the component of ω_b along the basis 1-form $V_{\mathbf{b}}$. Therefore, its covariant derivative is just its gradient $\partial_{\mathbf{a}}\omega_{\mathbf{b}}$, no additional term should be present! This terrible notation is completely standard in all geometry books, whether or not they use abstract indices. It is used in order to avoid a correct but much heavier expression. In the case of (2.4), what is meant by the left-hand side is that we look at the covariant derivative of ω_b in the direction $V_{\mathbf{a}}$, i.e. $\nabla_{\mathbf{a}}\omega_b$, and then we evaluate its component along the 1-form $V_{\mathbf{b}}$. The correct expression is therefore*

$$V^b_{\mathbf{b}}\nabla_{\mathbf{a}}\omega_b = \partial_{\mathbf{a}}\omega_{\mathbf{b}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\omega_{\mathbf{c}}\,.$$

However unsatisfactory the notation used in (2.4) may be, it is hard to avoid it, because the correct notation becomes totally unreadable when we differentiate tensors of arbitrary valence. But even though we shall with shame adopt such abusive notations, it is important to bear in mind what the correct meaning is for these incorrect expressions.

Now using the fact that ∇_a is tortion free, we have for any scalar field f (bearing in mind that the notation is again incorrect)

$$\begin{aligned} 0 = [\nabla_{\mathbf{a}}\,,\ \nabla_{\mathbf{b}}]\,f &= \nabla_{\mathbf{a}}\partial_{\mathbf{b}}f - \nabla_{\mathbf{b}}\partial_{\mathbf{a}}f \\ &= [\partial_{\mathbf{a}}\,,\ \partial_{\mathbf{b}}]\,f - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f + \Gamma_{\mathbf{ba}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f \\ &= -\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f + \Gamma_{\mathbf{ba}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f\,; \end{aligned}$$

whence $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}} = \Gamma_{\mathbf{ba}}{}^{\mathbf{c}}$, i.e. $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$ is symmetric in $(\mathbf{a},\mathbf{b})$.

We then use the Leibnitz rule to determine the action of ∇_a on a tensor field of any valence. First, we have

$$\begin{aligned} \nabla_{\mathbf{a}}\left(\omega_{\mathbf{b}}v^{\mathbf{b}}\right) &= \partial_{\mathbf{a}}\left(\omega_{\mathbf{b}}v^{\mathbf{b}}\right)\ \text{(since } \omega_{\mathbf{b}}v^{\mathbf{b}} \text{ is a scalar field)} \\ &= (\partial_{\mathbf{a}}\omega_{\mathbf{b}})v^{\mathbf{b}} + \omega_{\mathbf{b}}\partial_{\mathbf{a}}v^{\mathbf{b}}\,, \end{aligned}$$

and also

$$\nabla_{\mathbf{a}}\left(\omega_{\mathbf{b}}v^{\mathbf{b}}\right) = v^{\mathbf{b}}\left(\partial_{\mathbf{a}}\omega_{\mathbf{b}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\omega_{\mathbf{c}}\right) + \omega_{\mathbf{b}}\nabla_{\mathbf{a}}v^{\mathbf{b}}$$

and it follows that $\omega_{\mathbf{b}}\left(\nabla_{\mathbf{a}}v^{\mathbf{b}} - \partial_{\mathbf{a}}v^{\mathbf{b}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{b}}v^{\mathbf{c}}\right) = 0$ for any 1-form ω_a and any vector field v^a, hence

$$\nabla_{\mathbf{a}}v^{\mathbf{b}} = \partial_{\mathbf{a}}v^{\mathbf{b}} + \Gamma_{\mathbf{ac}}{}^{\mathbf{b}}v^{\mathbf{c}}\,.$$

Using again the Leibnitz rule and the fact that tensors fields are finite sums of tensor products of 1-forms and vector fields, we get for a tensor field of arbitrary valence :

$$\begin{aligned} \nabla_{\mathbf{a}}K^{\mathbf{i}_1\ldots\mathbf{i}_p}{}_{\mathbf{j}_1\ldots\mathbf{j}_q} = \partial_{\mathbf{a}}K^{\mathbf{i}_1\ldots\mathbf{i}_p}{}_{\mathbf{j}_1\ldots\mathbf{j}_q} &- \Gamma_{\mathbf{aj}_1}{}^{\mathbf{b}}K^{\mathbf{i}_1\ldots\mathbf{i}_p}{}_{\mathbf{b}\ldots\mathbf{j}_q} - \ldots - \Gamma_{\mathbf{aj}_q}{}^{\mathbf{b}}K^{\mathbf{i}_1\ldots\mathbf{i}_p}{}_{\mathbf{j}_1\ldots\mathbf{b}} \\ &+ \Gamma_{\mathbf{ab}}{}^{\mathbf{i}_1}K^{\mathbf{b}\ldots\mathbf{i}_p}{}_{\mathbf{j}_1\ldots\mathbf{j}_q} + \ldots + \Gamma_{\mathbf{ab}}{}^{\mathbf{i}_p}K^{\mathbf{i}_1\ldots\mathbf{b}}{}_{\mathbf{j}_1\ldots\mathbf{j}_q}\,. \end{aligned}$$

The fact that $\nabla_{\mathbf{a}}$ must commute with the metric will then give us the expression of $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$:

$$0 = \nabla_{\mathbf{a}}g_{\mathbf{bc}} = \partial_{\mathbf{a}}g_{\mathbf{bc}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{d}}g_{\mathbf{dc}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}g_{\mathbf{bd}}\,, \tag{2.5}$$

$$0 = \nabla_{\mathbf{b}}g_{\mathbf{ca}} = \partial_{\mathbf{b}}g_{\mathbf{ca}} - \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}g_{\mathbf{da}} - \Gamma_{\mathbf{ba}}{}^{\mathbf{d}}g_{\mathbf{cd}}\,, \tag{2.6}$$

$$0 = \nabla_{\mathbf{c}}g_{\mathbf{ab}} = \partial_{\mathbf{c}}g_{\mathbf{ab}} - \Gamma_{\mathbf{ca}}{}^{\mathbf{d}}g_{\mathbf{db}} - \Gamma_{\mathbf{cb}}{}^{\mathbf{d}}g_{\mathbf{ad}}\,. \tag{2.7}$$

Taking (2.5) + (2.6) -(2.7) and using the symmetry of $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$ and $g_{\mathbf{ab}}$, we obtain

$$2\Gamma_{\mathbf{ab}}{}^{\mathbf{d}} g_{\mathbf{cd}} = \partial_{\mathbf{a}} g_{\mathbf{bc}} + \partial_{\mathbf{b}} g_{\mathbf{ca}} - \partial_{\mathbf{c}} g_{\mathbf{ab}}$$

and multiplying by $g^{\mathbf{ec}}$,

$$\Gamma_{\mathbf{ab}}{}^{\mathbf{e}} = \frac{1}{2} g^{\mathbf{ec}} \left(\partial_{\mathbf{a}} g_{\mathbf{bc}} + \partial_{\mathbf{b}} g_{\mathbf{ac}} - \partial_{\mathbf{c}} g_{\mathbf{ab}} \right) .$$

Hence the uniqueness of the Levi-Civita connection. Existence is checked using the explicit formula above : all that needs to be verified is that ∇_a does not depend on the choice of coordinate system although the Christoffel symbols do ; it is a tedious but straightforward calculation. □

Corollary 2.1. *In a local coordinate basis, the action of the Levi-Civita connection on tensors of arbitrary valence is given by*

$$\begin{aligned} \nabla_{\mathbf{a}} K^{\mathbf{i}_1 \ldots \mathbf{i}_p}{}_{\mathbf{j}_1 \ldots \mathbf{j}_q} = \partial_{\mathbf{a}} K^{\mathbf{i}_1 \ldots \mathbf{i}_p}{}_{\mathbf{j}_1 \ldots \mathbf{j}_q} &- \Gamma_{\mathbf{a}\mathbf{j}_1}{}^{\mathbf{b}} K^{\mathbf{i}_1 \ldots \mathbf{i}_p}{}_{\mathbf{b} \ldots \mathbf{j}_q} - \ldots - \Gamma_{\mathbf{a}\mathbf{j}_q}{}^{\mathbf{b}} K^{\mathbf{i}_1 \ldots \mathbf{i}_p}{}_{\mathbf{j}_1 \ldots \mathbf{b}} \\ &+ \Gamma_{\mathbf{ab}}{}^{\mathbf{i}_1} K^{\mathbf{b} \ldots \mathbf{i}_p}{}_{\mathbf{j}_1 \ldots \mathbf{j}_q} + \ldots + \Gamma_{\mathbf{ab}}{}^{\mathbf{i}_p} K^{\mathbf{i}_1 \ldots \mathbf{b}}{}_{\mathbf{j}_1 \ldots \mathbf{j}_q} . \end{aligned} \quad (2.8)$$

where the Christoffel symbols $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$, are defined by

$$\Gamma_{\mathbf{ab}}{}^{\mathbf{c}} = \frac{1}{2} g^{\mathbf{cd}} \left(\partial_{\mathbf{a}} g_{\mathbf{bd}} + \partial_{\mathbf{b}} g_{\mathbf{ad}} - \partial_{\mathbf{d}} g_{\mathbf{ab}} \right) \quad (2.9)$$

and satisfy

$$\Gamma_{\mathbf{ab}}{}^{\mathbf{c}} = \Gamma_{(\mathbf{ab})}{}^{\mathbf{c}} .$$

Remark 2.7.

1. *We have established in the proof of theorem 2.1 that a connection ∇_a is characterized by Christoffel symbols $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$ whose expression depends on the choice of local coordinates and acts on tensor fields as described in equation (2.8).*
2. *The action of the connection is referred to as covariant differentiation.*
3. *The connection is said to be metric-compatible if it commutes with the metric, i.e. $\nabla_a g_{bc} = 0$ and $\nabla_a g^{bc} = 0$.*
4. *Given a connection ∇_a and a vector field V^a, the covariant directional derivative in the direction of V is defined as the contraction of V^a and ∇_a, i.e. $V^a \nabla_a$, and sometimes denoted ∇_V.*

Remark 2.8. *It is important to note that the Christoffel symbols $\Gamma_{\mathbf{ab}}{}^{\mathbf{c}}$ are not a tensor field : it is very easy to see that they depend on the choice of local coordinates (see exercice 3.2). However, the connection is an intrinsic object independent of the coordinate system. The transformation of Christoffel symbols under a change of coordinates is fixed by the independence of ∇_a of the coordinate system and the fact that ∇_a obeys the Leibnitz rule : more precisely, in two different coordinate systems, the action of ∇_a on a 1-form must be the same ; knowing the way in which the components of the 1-form change between the two bases and using the Leibnitz rule, we obtain the relation between the Christoffel symbols in the two coordinate systems.*

Proposition 2.1. *To a connection ∇_a is associated a tortion tensor defined by*

$$[\nabla_a\,,\ \nabla_b]\,f =: T_{ab}{}^c\nabla_c f = T_{ab}{}^c\partial_c f\,. \tag{2.10}$$

It is indeed a tensor field : since the connection is an intrinsic object, so is the tortion tensor. In a coordinate basis, the torsion tensor is expressed in terms of the Christoffel symbols as

$$T_{\mathbf{ab}}{}^{\mathbf{c}} = \Gamma_{\mathbf{ba}}{}^{\mathbf{c}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\,. \tag{2.11}$$

By definition, we have $T_{ab}{}^c = T_{[ab]}{}^c$. If the tortion tensor is zero, the connection is said to be tortion-free.

Proof. The action of the commutator of two covariant derivatives on a scalar field

$$[\nabla_a\,,\ \nabla_b]\,f = \nabla_a\nabla_b f - \nabla_b\nabla_a f$$

is linear on the gradient of f. Moreover, spelling out the formula explicitely in a coordinate basis,

$$[\nabla_{\mathbf{a}}\,,\ \nabla_{\mathbf{b}}]\,f = \partial_{\mathbf{a}}\partial_{\mathbf{b}}f - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f - \partial_{\mathbf{b}}\partial_{\mathbf{a}}f + \Gamma_{\mathbf{ba}}{}^{\mathbf{c}}\partial_{\mathbf{c}}f = \left(\Gamma_{\mathbf{ba}}{}^{\mathbf{c}} - \Gamma_{\mathbf{ab}}{}^{\mathbf{c}}\right)\partial_{\mathbf{c}}f\,, \tag{2.12}$$

we see that $[\nabla_a\,,\ \nabla_b]$ acts on $\nabla_a f$ as a differential operator of order zero, hence we have :

$$[\nabla_a\,,\ \nabla_b]\,f =: T_{ab}{}^c\nabla_c f \tag{2.13}$$

and the covariant derivative being an intrinsic object, $T_{ab}{}^c$ is a tensor field. The expression (2.11) of the tortion tensor in terms of Christoffel symbols follows from (2.12) and (2.13). □

Proposition 2.2. *When the commutator of two covariant derivatives acts on tensor fields of arbitrary valence, it involves another tensor field : the Riemann curvature tensor R_{abcd}. More precisely,*

$$\begin{aligned}
&([\nabla_a\,,\ \nabla_b] - T_{ab}{}^c\nabla_c)K^{i_1\ldots i_p}{}_{j_1\ldots j_q}\\
&\qquad\qquad = R_{abc}{}^{i_1}K^{c\ldots i_p}{}_{j_1\ldots j_q} + \ldots + R_{abc}{}^{i_p}K^{i_1\ldots c}{}_{j_1\ldots j_q}\\
&\qquad\qquad - R_{abj_1}{}^{d}K^{i_1\ldots i_p}{}_{d\ldots j_q} - \ldots - R_{abj_q}{}^{d}K^{i_1\ldots i_p}{}_{j_1\ldots d}\,.
\end{aligned} \tag{2.14}$$

In a local coordinate basis, its expression in terms of the Christoffel symbols is given by

$$R_{\mathbf{abc}}{}^{\mathbf{d}} = \partial_{\mathbf{b}}\left(\Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\right) - \partial_{\mathbf{a}}\left(\Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\right) + \Gamma_{\mathbf{bc}}{}^{\mathbf{e}}\Gamma_{\mathbf{ae}}{}^{\mathbf{d}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{e}}\Gamma_{\mathbf{be}}{}^{\mathbf{d}}\,. \tag{2.15}$$

Proof. We denote

$$\Delta_{ab} := [\nabla_a\,,\ \nabla_b] - T_{ab}{}^c\nabla_c\,.$$

First we check that Δ_{ab} acts on forms as a linear differential operator of order zero, and hence as a tensor field since all quantities involved are intrinsic. We have in a coordinate basis

$$\begin{aligned}
\Delta_{\mathbf{ab}}\omega_{\mathbf{c}} &= \partial_{\mathbf{a}}\left(\partial_{\mathbf{b}}\omega_{\mathbf{c}} - \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\omega_{\mathbf{d}}\right) - \Gamma_{\mathbf{ab}}{}^{\mathbf{d}}\nabla_{\mathbf{d}}\omega_{\mathbf{c}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\nabla_{\mathbf{b}}\omega_{\mathbf{d}} && (= \nabla_{\mathbf{a}}\nabla_{\mathbf{b}}\omega_{\mathbf{c}}) \\
&\quad - \partial_{\mathbf{b}}\left(\partial_{\mathbf{a}}\omega_{\mathbf{c}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\omega_{\mathbf{d}}\right) + \Gamma_{\mathbf{ba}}{}^{\mathbf{d}}\nabla_{\mathbf{d}}\omega_{\mathbf{c}} + \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\nabla_{\mathbf{a}}\omega_{\mathbf{d}} && (= -\nabla_{\mathbf{b}}\nabla_{\mathbf{a}}\omega_{\mathbf{c}}) \\
&\quad - \Gamma_{\mathbf{ba}}{}^{\mathbf{d}}\nabla_{\mathbf{d}}\omega_{\mathbf{c}} + \Gamma_{\mathbf{ab}}{}^{\mathbf{d}}\nabla_{\mathbf{d}}\omega_{\mathbf{c}} && \left(= -T_{\mathbf{ab}}{}^{\mathbf{d}}\nabla_{\mathbf{d}}\omega_{\mathbf{c}}\right) \\
&= -\partial_{\mathbf{a}}\left(\Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\right)\omega_{\mathbf{d}} - \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\partial_{\mathbf{a}}\omega_{\mathbf{d}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\partial_{\mathbf{b}}\omega_{\mathbf{d}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\Gamma_{\mathbf{bd}}{}^{\mathbf{e}}\omega_{\mathbf{e}} \\
&\quad + \partial_{\mathbf{b}}\left(\Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\right)\omega_{\mathbf{d}} + \Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\partial_{\mathbf{b}}\omega_{\mathbf{d}} + \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\partial_{\mathbf{a}}\omega_{\mathbf{d}} + \Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\Gamma_{\mathbf{ad}}{}^{\mathbf{e}}\omega_{\mathbf{e}} \\
&= \left(\partial_{\mathbf{b}}\left(\Gamma_{\mathbf{ac}}{}^{\mathbf{d}}\right) - \partial_{\mathbf{a}}\left(\Gamma_{\mathbf{bc}}{}^{\mathbf{d}}\right) + \Gamma_{\mathbf{bc}}{}^{\mathbf{e}}\Gamma_{\mathbf{ae}}{}^{\mathbf{d}} - \Gamma_{\mathbf{ac}}{}^{\mathbf{e}}\Gamma_{\mathbf{be}}{}^{\mathbf{d}}\right)\omega_{\mathbf{d}}\,.
\end{aligned}$$

This gives (2.14) in the case where Δ_{ab} acts on a 1-form and (2.15) .

Then, for a 1-form α_a and a vector field v^a, using the fact that $\alpha_a v^a$ is a scalar field,

$$\Delta_{ab}\alpha_e v^e = 0\,,$$

and also

$$\Delta_{ab}\alpha_e v^e = \alpha_e \Delta_{ab} v^e + v^e \Delta_{ab}\alpha_e$$

since the cross terms cancel one another. Hence, for any 1-form α_a and any vector field v^a,

$$\alpha_e \Delta_{ab} v^e = -v^e R_{abe}{}^d \alpha_d = -\alpha_e R_{abc}{}^e v^c\,,$$

which proves (2.14) in the case where Δ_{ab} acts on a vector field. As we have done in the special case of the contraction of a 1-form and a vector field, it is trivial to verify that Δ_{ab} satisfies the Leibnitz rule. Using this and the fact that tensor fields are finite sums of tensor products of 1-forms and vector fields, we obtain (2.14) in the general case. □

Corollary 2.2. *The commutator* $[\nabla_a\,,\ \nabla_b]$ *(and therefore also* $[\nabla_a\,,\ \nabla_b] - T_{ab}{}^c\nabla_c$*) satisfies the Leibnitz rule.*

Theorem 2.2. *The Riemann tensor has the following symmetries :*

1. $R_{(ab)cd} = 0$ *;*
2. $R_{ab(cd)} = 0$ *if the connection is metric-compatible ;*
3. $R_{[abc]}{}^d + \nabla_{[a}T_{bc]}{}^d + T_{[ab}{}^e T_{c]e}{}^d = 0$*, which, for a tortion-free connection, gives the first Bianchi identity* $R_{[abc]}{}^d = 0$ *;*
4. $\nabla_{[a}R_{bc]d}{}^e + T_{[ab}{}^l R_{c]ld}{}^e = 0$ *and if the connection is tortion-free, this gives the second Bianchi identity* $\nabla_{[a}R_{bc]d}{}^e = 0$*.*

Corollary 2.3. *Note that using* $R_{(ab)cd} = 0$*, the first Bianchi identity becomes*

$$R_{abc}{}^d + R_{bca}{}^d + R_{cab}{}^d = 0\,.$$

Proof of the theorem. The first property follows from the definition of the Riemann tensor.

Proof of 2. Consider the action of Δ_{ab} on the metric. Using the fact that the connection is metric compatible, we must have $\Delta_{ab} g_{cd} = 0$, hence

$$0 = \Delta_{ab} g_{cd} = -R_{abc}{}^{e} g_{ed} - R_{abd}{}^{e} g_{ce} = -R_{abcd} - R_{abdc}\,,$$

which gives the required symmetry.

Proof of 3. The properties of symmetry operations and of the tortion tensor give

$$\begin{aligned}
\Delta_{[ab}\nabla_{c]} &= 2\nabla_{[[a}\nabla_{b]}\nabla_{c]} - T_{[[ab]}{}^{e}\nabla_{|e|}\nabla_{c]} \\
&= 2\nabla_{[a}\nabla_{b}\nabla_{c]} - T_{[ab}{}^{e}\nabla_{|e|}\nabla_{c]} \\
&= 2\nabla_{[a}\nabla_{[b}\nabla_{c]]} + T_{[ab}{}^{e}\Delta_{c]e} - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} + T_{[ab}{}^{e}T_{c]e}{}^{d}\nabla_{d}\,.
\end{aligned} \tag{2.16}$$

Now for a given scalar field f, using (2.16) and the fact that Δ_{ab} vanishes on scalar fields,

$$\begin{aligned}
\Delta_{[ab}\nabla_{c]} f &= -R_{[abc]}{}^{d}\nabla_{d} f \\
&= 2\nabla_{[a}\nabla_{[b}\nabla_{c]]} f + T_{[ab}{}^{e}\Delta_{c]e} f - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} f + T_{[ab}{}^{e}T_{c]e}{}^{d}\nabla_{d} f \\
&= \nabla_{[a}T_{bc]}{}^{d}\nabla_{d} f - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} f + T_{[ab}{}^{e}T_{c]e}{}^{d}\nabla_{d} f \\
&= \nabla_{[a}T_{bc]}{}^{d}\nabla_{d} f - T_{[bc}{}^{d}\nabla_{a]}\nabla_{d} f + T_{[ab}{}^{e}T_{c]e}{}^{d}\nabla_{d} f \\
&= \left(\nabla_{[a}T_{bc]}{}^{d}\right)\nabla_{d} f + T_{[ab}{}^{e}T_{c]e}{}^{d}\nabla_{d} f\,,
\end{aligned}$$

which proves *3* since at any given point $\nabla_a f$ can be any covector.

Proof of 4. It is similar to the proof of *3* but we consider a vector field instead of a scalar field. First, using *3*, we have

$$\begin{aligned}
\Delta_{[ab}\nabla_{c]} v^{d} &= -R_{[abc]}{}^{e}\nabla_{e} v^{d} + R_{[ab|e|}{}^{d}\nabla_{c]} v^{e} \\
&= \left(\nabla_{[a}T_{bc]}{}^{e}\right)\nabla_{e} v^{d} + T_{[ab}{}^{i}T_{c]i}{}^{e}\nabla_{e} v^{d} + R_{[ab|e|}{}^{d}\nabla_{c]} v^{e}
\end{aligned} \tag{2.17}$$

and using (2.16), we also have

$$\begin{aligned}
\Delta_{[ab}\nabla_{c]} v^{d} &= 2\nabla_{[a}\nabla_{[b}\nabla_{c]]} v^{d} + T_{[ab}{}^{e}\Delta_{c]e} v^{d} - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} v^{d} + T_{[ab}{}^{e}T_{c]e}{}^{i}\nabla_{i} v^{d} \\
&= \nabla_{[a}\Delta_{bc]} v^{d} + \nabla_{[a}T_{bc]}{}^{e}\nabla_{e} v^{d} + T_{[ab}{}^{e}R_{c]ei}{}^{d} v^{i} \\
&\quad - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} v^{d} + T_{[ab}{}^{e}T_{c]e}{}^{i}\nabla_{i} v^{d} \\
&= \nabla_{[a}R_{bc]e}{}^{d} v^{e} + \nabla_{[a}T_{bc]}{}^{e}\nabla_{e} v^{d} + T_{[ab}{}^{e}R_{c]ei}{}^{d} v^{i} \\
&\quad - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} v^{d} + T_{[ab}{}^{e}T_{c]e}{}^{i}\nabla_{i} v^{d} \\
&= \left(\nabla_{[a}R_{bc]e}{}^{d}\right) v^{e} + R_{[bc|e|}{}^{d}\nabla_{a]} v^{e} + \nabla_{[a}T_{bc]}{}^{e}\nabla_{e} v^{d} \\
&\quad + T_{[ab}{}^{e}R_{c]ei}{}^{d} v^{i} - T_{[ab}{}^{e}\nabla_{c]}\nabla_{e} v^{d} + T_{[ab}{}^{e}T_{c]e}{}^{i}\nabla_{i} v^{d} \\
&= \left(\nabla_{[a}R_{bc]e}{}^{d}\right) v^{e} + R_{[ab|e|}{}^{d}\nabla_{c]} v^{e} + \left(\nabla_{[a}T_{bc]}{}^{e}\right)\nabla_{e} v^{d} \\
&\quad + T_{[ab}{}^{i}R_{c]ie}{}^{d} v^{e} + T_{[ab}{}^{i}T_{c]i}{}^{e}\nabla_{e} v^{d}\,.
\end{aligned} \tag{2.18}$$

Putting together (2.17) and (2.18), we obtain

$$\left(\nabla_{[a}R_{bc]e}{}^{d}\right) v^{e} + T_{[ab}{}^{i}R_{c]ie}{}^{d} v^{e} = 0$$

for any vector field v^e, which proves *4*. □

Remark 2.9. *The anti-symmetrized derivative of the Riemann tensor appearing in the fourth point or the previous theorem reads :*

$$\nabla_{[a}R_{bc]d}{}^e = \frac{1}{6}\left(\nabla_a R_{bcd}{}^e + \nabla_b R_{cad}{}^e + \nabla_c R_{abd}{}^e - \nabla_b R_{acd}{}^e - \nabla_c R_{bad}{}^e - \nabla_a R_{cbd}{}^e\right) ;$$

using the first symmetry of the Riemann tensor, this takes on a simpler form

$$\nabla_{[a}R_{bc]d}{}^e = \frac{1}{3}\left(\nabla_a R_{bcd}{}^e + \nabla_b R_{cad}{}^e + \nabla_c R_{abd}{}^e\right) .$$

Definition 2.15. *We define some important curvature quantities that are special parts of the full Riemann tensor :*

- *the Ricci tensor R_{ab} is the trace of the Riemann tensor in its second and fourth indices*

$$R_{ab} := R_{acb}{}^c = g^{cd}R_{acbd} ;$$

- *the scalar curvature R is the trace of the Ricci tensor*

$$R := R_a{}^a = g^{ab}R_{ab}$$

  and it is often denoted by Scal_g *;*

- *the Einstein tensor G_{ab} is defined as*

$$G_{ab} := R_{ab} - \frac{1}{2}Rg_{ab} ;$$

- *the Weyl tensor C_{abcd} is the trace-free part of the Riemann tensor*

$$C_{abcd} = R_{abcd} - \frac{1}{2}\left(g_{a[c}R_{d]b} - g_{b[c}R_{d]a}\right) + \frac{1}{3}Rg_{a[c}g_{d]b} .$$

Proposition 2.3. *For the Levi-Civita connection, we have the following properties :*

1. $R_{ab} = R_{(ab)}$ *(which implies immediately $G_{ab} = G_{(ab)}$) ;*
2. $\nabla^a G_{ab} = 0$.

Proof.

1. Using the fact that $R_{(ab)cd} = 0$,

$$R_{ab} - R_{ba} = \left(R_{acbd} - R_{bcad}\right)g^{cd} = -\left(R_{cabd} + R_{bcad}\right)g^{cd} .$$

   Then, by the first Bianchi identity (which requires the connection to be tortion-free),

$$R_{ab} - R_{ba} = R_{abcd}g^{cd} .$$

   Assuming in addition the connection to be metric compatible, we have $R_{ab(cd)} = 0$, i.e. R_{abcd} is antisymmetrical in the last two indices. Contracting with the metric (which is symmetric), we obtain 0.

2. We start from the second Bianchi identity in which we contract the indices a and e :

$$\begin{aligned} 0 &= \nabla_a R_{bcd}{}^a + \nabla_b R_{cad}{}^a + \nabla_c R_{abd}{}^a - \nabla_b R_{acd}{}^a - \nabla_c R_{bad}{}^a + \nabla_a R_{cbd}{}^a \\ &= 2\nabla_a R_{bcd}{}^a + 2\nabla_b R_{cad}{}^a - 2\nabla_c R_{bad}{}^a \text{ using } R_{(ab)cd} = 0\,, \\ &= 2\nabla_a R_{bcd}{}^a + 2\nabla_b R_{cd} - 2\nabla_c R_{bd} \\ &= 2\nabla^a R_{bcda} + 2\nabla_b R_{cad}{}^a - 2\nabla_c R_{bad}{}^a \\ &= -2\nabla^a R_{bcad} + 2\nabla_b R_{cad}{}^a - 2\nabla_c R_{bad}{}^a \end{aligned}$$

the last equality but one being obtained using the fact that the connection is metric compatible and the last one using the symmetry $R_{ab(cd)} = 0$ which also requires metric compatibility. Then, we contract the indices c and d. Using the metric compatibility again, we obtain :

$$\begin{aligned} \nabla^a R_{bca}{}^c &= \nabla_b R - \nabla^d R_{bad}{}^a \\ \text{i.e. } 0 &= 2\nabla^a R_{ba} - \nabla_b R \\ &= 2\nabla^a R_{ba} - \nabla^a (R g_{ab}) \\ &= 2\nabla^a G_{ba} = 2\nabla^a G_{ab}\,. \quad \square \end{aligned}$$

The **Einstein vacuum equations** that characterize the geometry of an empty universe are simply

$$G_{ab} = 0\,. \tag{2.19}$$

In the case of a universe containing energy or matter, the Einstein equation will become

$$G_{ab} = 8\pi T_{ab}$$

where T_{ab} is a tensor (referred to as the stress-energy tensor) describing the distribution of matter and energy in the universe.

Considered as an equation on the metric, Einstein's equation is a system of non linear second order partial differential equations. Taking the trace of G_{ab}, we obtain

$$G_a{}^a = R_a{}^a - \frac{1}{2} R g_a{}^a = R - 2R = -R\,,$$

whence (2.19) is equivalent to

$$R_{ab} = 0\,. \tag{2.20}$$

Einstein vacuum spacetimes are also referred to as Ricci-flat spacetimes.

There is a modified version of the Einstein equation, due to Einstein himself in 1917, involving a constant Λ called the "cosmological constant". It has the following form

$$G_{ab} + \Lambda g_{ab} = 8\pi T_{ab}\,. \tag{2.21}$$

Einstein introduced this modification because the original form of the theory did not allow for a static universe (unless it is also flat), it had to be expanding or contracting. The cosmological constant induces a repulsive force which Einstein adjusted so that it would counterbalance gravitation exactly. His new version of the theory thus allow for a static universe : the Einstein cylinder, see chapter 8. The reason for this is probably partly religious but also a static universe was the commonly accepted picture at that time. This unfortunately prevented him from discovering the expansion of the universe which Hubble proved in 1929. He subsequently declared that this was his greatest mistake. It is interesting to notice that observations made from 1993 to 2005 show that the expansion of the universe is now faster than we would expect. A well accepted explaination is that a repulsive force induced by a cosmological constant is responsible for it : in the early stages of the universe, the expansion from the big bang was slowed down by gravity, but as the universe expanded, the effects of gravity weakened and this repulsive force (referred to as dark energy) accelerated the expansion. The universe would appear to have a small but strictly positive cosmological constant. It is regrettable that Einstein never knew that his greatest mistake was just another brilliant idea.

Taking the trace of (2.21), we see that in the vacuum case, i.e. for $T_{ab} = 0$, the cosmological constant is a multiple of the scalar curvature :

$$\Lambda = \frac{1}{4} R \,.$$

2.5 Flow of a vector field, Lie derivative, Killing vectors

Beside the covariant derivative along a vector field, there is an important type of directional derivative called the Lie derivative. It is independent of a choice of connection and is a derivation along the flow of a vector field.

2.5.1 Flow of a vector field

Consider on a space-time $(\mathcal{M}, g)$ a $\mathcal{C}^1$ vector field V, i.e. a $\mathcal{C}^1$ section of $T\mathcal{M}$.

Definition 2.16 (Integral curve). *An integral curve of V is a curve in $\mathcal{M}$ that is a maximal solution to the equation*

$$\gamma'(s) = V(\gamma(s)) \,. \tag{2.22}$$

By the Cauchy-Lipschitz theorem (used in open sets of $\mathbb{R}^n$ through local charts), we have existence and uniqueness of maximal solutions of the Cauchy problem for (2.22). This allows us to define the propagator or flow of the vector field. A more detailed use of the machinery of the theory of ordinary differential equations shows that it is a local 1-parameter group of diffeomorphisms.

Definition 2.17 (Flow). *The flow of the vector field V is a family of mappings $\Phi_V(s)$ that to a point p in $\mathcal{M}$ associate $\gamma_p(s)$, where γ_p is the unique maximal solution to the Cauchy problem*

$$\gamma_p'(s) = V(\gamma_p(s)) \,,\ \gamma_p(0) = p \,.$$

Remark 2.10. *Since the maximal solution does not necessarily exist for all values of s, the mapping $\Phi_V(s)$ is not usually globally defined, except of course $\Phi_V(0)$ which is the identity. However, if $\Phi_V(t)$ is well defined at a point $p \in \mathcal{M}$, it is defined in a neighbourhood of p.*

Proposition 2.4. *The flow Φ_V of the vector field V is a local 1-parameter group of $\mathcal{C}^1$ diffeomorphisms, i.e. it has the following properties :*

1. *given $s \in \mathbb{R}$ and an open set $\mathcal{U}$ of $\mathcal{M}$ on which $\Phi_V(s)$ is well defined, $\Phi_V(s)$ is a $\mathcal{C}^1$ diffeomorphism from $\mathcal{U}$ onto $\mathcal{V} = \Phi_V(s)(\mathcal{U})$;*

2. *for any $s_1, s_2 \in \mathbb{R}$, we have $\Phi_V(s_1)\Phi_V(s_2) = \Phi_V(s_1 + s_2)$ wherever all quantities are defined.*

Moreover, if the vector field V is $\mathcal{C}^k$, then the flow Φ_V of the vector field V is a local 1-parameter group of $\mathcal{C}^k$ diffeomorphisms.

We omit the proof of this result and refer to the classic theory of ordinary differential equations for it. A good reference is the book by Zuily and Queffélec [27]. It is important to understand that the second property as well as the invertibility of $\Phi_V(t)$ are trivial consequences of the uniqueness of maximal solutions of the Cauchy problem. The delicate part of the proof is the regularity of Φ_V. This amounts to proving the regularity of the solution with respect to the intitial data.

2.5.2 Action of the flow on tensor fields

From here on, we shall use simplified notations for the flow of V : we denote Φ the flow Φ_V and Φ_t the local diffeomorphism $\Phi_V(t)$.

First, we observe that Φ_t acts on scalar functions on $\mathcal{M}$ via a mapping referred to as the pull-back and defined as follows :

Definition 2.18 (Action on scalar function). *Let $f : \mathcal{M} \to \mathbb{R}$ a continuous function on $\mathcal{M}$. We define the pulled back function $(\Phi_t)^* f$ as follows*

$$(\Phi_t)^* f = f \circ \Phi_t \, .$$

We see that if we evaluate, for a given differentiable scalar function f the quantity

$$\lim_{t \to 0} \frac{1}{t} \left((\Phi_t)^* f - f \right)(p) \, ,$$

we obtain simply

$$df(p) \left(\Phi'(0)(p) \right) = Vf(p) = V^a \partial_a f(p) \, . \tag{2.23}$$

We can define a similar action on vector fields. This action however is more naturally defined as a push forward, i.e. to a vector at the point p, we associate a vector at the point $\Phi_t(p)$. This is done as follows :

Definition 2.19 (Action on vector fields). *Let X be a continuous vector field on $\mathcal{M}$, i.e. $X \in \mathcal{C}(\mathcal{M}\, ;\, TS)$, we define the push-forward of X by Φ_t (denoted $(\Phi_t)_* X$) by its action on differentiable functions on $\mathcal{M}$;*

$$\left[\left((\Phi_t)_* X \right) f \right](p) = \left[X \left(f \circ \Phi_t \right) \right] \left(\Phi_{-t}(p) \right) = \left[X \left((\Phi_t)^* f \right) \right] \left(\Phi_{-t}(p) \right) \, .$$

We can also define a pulled back vector field by, instead of the push-forward mapping, applying, applying its inverse : we denote it $(\Phi_t)^* X$;

$$\left[\left((\Phi_t)^* X \right) f \right](p) = \left[\left(\left((\Phi_t)_* \right)^{-1} X \right) f \right](p) = X \left(f \circ \Phi_{-t} \right) \left(\Phi_t(p) \right) = \left[\left((\Phi_{-t})_* X \right) f \right](p) \, .$$

The pushed forward and pulled-back vector fields satisfy

$$\begin{aligned} ((\Phi_t)_*(X))(p) &= D(\Phi_t)(\Phi_{-t}(p))(X(\Phi_{-t}(p))) \,, \\ ((\Phi_t)^*(X))(p) &= D(\Phi_{-t})(\Phi_t(p))(X(\Phi_t(p))) \,. \end{aligned}$$

We can differentiate a vector field along the flow Φ_t just as we did for functions. We have for a differentiable vector field X on $\mathcal{M}$:

$$\lim_{t\to 0} \frac{1}{t}\left((\Phi_t)^* X - X\right) = [V, X] \,, \tag{2.24}$$

where $[V, X]$ is the Lie bracket of the two vector fields V and X, defined by

Definition 2.20. *The Lie bracket $[X, Y]$ of two differentiable vector fields on M is defined as $[X, Y] = XY - YX$, i.e. it is simply the commutator of the two vector fields as differential operators.*

We can then naturally extend the notion of pull-back to 1-forms on $\mathcal{M}$.

Definition 2.21 (Action on 1-forms)**.** *Consider a continuous 1-form ω on $\mathcal{M}$, we define the pulled-back 1-form $(\Phi_t)^*\omega$ by its action on a differentiable vector field X :*

$$((\Phi_t)^*\omega)(X)(p) = \omega\left((\Phi_t)_* X\right)(\Phi_t(p)) \,.$$

The pulled-back 1-form satisfies

$$((\Phi_t)^*\omega)(p) = \left[D(\Phi_t)(p)\right]^* \left(\omega(\Phi_t(p))\right) \,.$$

The pull-back is then extended to arbitrary tensor fields by first defining it on tensor products of vector fields and 1-forms and then extending it by linearity to the tensor bundles of a given valence.

Definition 2.22 (Action of tensor fields)**.** *The pull-back of a tensor product of m vector fields and n 1-forms is simply defined as*

$$(\Phi_t)^*\left(U \otimes ... \otimes V \otimes \alpha \otimes ... \otimes \beta\right) = (\Phi_t)^* U \otimes ... \otimes (\Phi_t)^* V \otimes (\Phi_t)^*\alpha \otimes ... \otimes (\Phi_t)^*\beta \,.$$

Definition 2.23 (Lie derivative)**.** *Consider a differentiable tensor field T on $\mathcal{M}$, its Lie derivative along V is defined as*

$$\mathcal{L}_V T := \lim_{t\to 0} \frac{1}{t}\left((\Phi_t)^* T - T\right) \,.$$

Its action on scalar functions is given by (2.23) and its action on vector fields by (2.24). Moreover, due to the definition of the pull-back on tensor products, the Lie derivative satisfies the Leibnitz rule. This is easy to check : consider two differentiable tensor fields T ans S, we have for $t > 0$

$$\begin{aligned} \frac{1}{t}\left((\Phi_t)^*(T\otimes S) - T\otimes S\right) &= \frac{1}{t}\left((\Phi_t)^* T \otimes (\Phi_t)^* S - T \otimes S\right) \\ &= \frac{1}{t}\left((\Phi_t)^* T - T\right) \otimes (\Phi_t)^* S + T \otimes \frac{1}{t}\left((\Phi_t)^* S - S\right) \\ &\longrightarrow \mathcal{L}_V T \otimes S + T \otimes \mathcal{L}_V S \text{ as } t \to 0 \,. \end{aligned}$$

This suffices to characterize the action of the Lie derivative on any type of tensor field. In particular, the Lie derivative of a differentiable 1-form on $\mathcal{M}$ along V

$$\mathcal{L}_V \omega_a = \lim_{t \to 0} \frac{1}{t} \left((\Phi_t)^* \omega_a - \omega_a \right)(p) = V^b \nabla_b \omega_a + \omega_b \nabla_a V^b \tag{2.25}$$

can be obtained using the fact that for a differentiable vector field X^a, $\omega_a X^a$ is a differentiable scalar function and that we know the Lie derivatives of both vector fields and scalar functions. Of particular interest is the expression of the Lie derivative of the metric along a vector field.

Proposition 2.5. *The Lie derivative of the metric along a vector field V^a is given by*

$$\mathcal{L}_V g_{ab} = g_{cb} \nabla_a V^c + g_{ac} \nabla_b V^c = 2 \nabla_{(a} V_{b)} \,. \tag{2.26}$$

Proof. As a consequence of the Leibnitz rule, we have

$$\mathcal{L}_V g_{ab} = V^c \nabla_c g_{ab} + g_{cb} \nabla_a V^c + g_{ac} \nabla_b V^c$$

and (2.26) then follows from the metric-compatibility of the Levi-Civita connection. □

Proposition 2.6. *The Lie derivative is independent of the connection, i.e. it can be expressed using any connection, it will remain the same.*

Proof. This is clear for its action on vector fields and scalars. Now given a vector field X and a 1-form ω,

$$\mathcal{L}_V (\omega_a X^a) = \omega_a \mathcal{L}_V X^a + X^a \mathcal{L}_V \omega_a \,,$$

whence

$$X^a \mathcal{L}_V \omega_a = \mathcal{L}_V (\omega_a X^a) - \omega_a \mathcal{L}_V X^a$$

is the sum of two terms independent of the connection. This extends to all types of tensors by the Leibnitz rule. □

Definition 2.24 (Killing vector)**.** *A Killing vector field on a manifold $\mathcal{M}$ equipped with a metric g (assumed differentiable) is a differentiable vector field K^a on $\mathcal{M}$ such that its flow leaves the metric invariant, i.e. $\Phi_K(t)^* g_{ab} = g_{ab}$, or equivalently, $\mathcal{L}_K g_{ab} = 0$. As a consequence of proposition 2.5, a differentiable vector field K^a on $(\mathcal{M}, g)$ is Killing if and only if K^a satisfies the Killing equation*

$$\nabla_{(a} K_{b)} = 0 \,. \tag{2.27}$$

2.6 Geodesics

It is a classic notion that the most direct path between two points is the straight line. The notion of straight line however only has a meaning in affine spaces. We of course do not live in an affine space, so this classic image is in fact wrong and even meaningless. It is however true to a very good degree of accuracy provided the two points are not too far from each other (wich may mean arbitrarily close to each other if the cuvature is arbitrarily large), since a local diffeomorphism that straightens our spacetime to $\mathbb{R}^4$ in a small enough neighbourhood of these two points will be very close to the identity. In an affine space, a useful notion is that of a "freely falling object", i.e. an object that is not accelerated. The trajectories of such objects are of course exactly the straight lines. The advantage is that the notion of an object that is not accelerated can be extended to a general manifold, its trajectory is then a particular type of curve referred to as a geodesic. We have some freedom in the way we define the acceleration, i.e. on how we differentiate the speed vector along the curve. We choose a way of differentiating along the curve that transforms a tensor of a given valence into another tensor of the same valence, it is the so-called absolute derivative

$$\frac{D}{Ds} := \nabla_{\dot{\gamma}(s)}$$

i.e. the covariant derivative along the speed vector.

This provides us with the following definition of a geodesic, i.e. a curve with zero acceleration.

Definition 2.25 (Geodesics)**.** *A geodesic on a spacetime $(\mathcal{M}, g)$ is a $\mathcal{C}^2$ curve on $\mathcal{M}$ (i.e. the data of a pair (I, γ) where I is an interval and $\gamma\ : I \to \mathcal{M}$ is a $\mathcal{C}^2$ function such that $\dot{\gamma}(s)$ does not vanish on I) such that its acceleration, defined by $\frac{D}{Ds}\dot{\gamma}(s) = \nabla_{\dot{\gamma}(s)}\dot{\gamma}(s) = 0$. Expressing the covariant derivative in a coordinate basis, this immediately gives the equation of a geodesic*

$$\frac{\mathrm{d}^2\gamma^{\mathbf{a}}}{\mathrm{d}s^2} + \Gamma^{\mathbf{a}}_{\mathbf{bc}}\frac{\mathrm{d}\gamma^{\mathbf{b}}}{\mathrm{d}s}\frac{\mathrm{d}\gamma^{\mathbf{c}}}{\mathrm{d}s} = 0\,.$$

If we consider a differentiable vector field T^a that is propagated parallel along itself, i.e. such that $T^a\nabla_a T^b$ is colinear to T^a its integral curves are geodesics. Indeed, modulo re-parametrization, we can assume that $T^a\nabla_a T^b = 0$; the parameter of the integral curves that gives a tangent vector field satisfying this is called the affine parameter.

The geodesic equation is a differential equation whose coefficients are the Christoffel symbols, i.e. involve first order derivatives of the metric. Therefore, the metric needs to be such that its derivative is locally Lipschitz in order to ensure the existence and uniqueness of maximal solutions by the Cauchy-Lipschitz theorem. For a $\mathcal{C}^2$ metric, this is naturally guaranteed.

Remark 2.11. *In euclidian space in cartesian coordinates, the Christoffel symbols are zero and the geodesics are the straight lines. This property is shared by Minkowski spacetime which is the subject of chapter 3.*

Remark 2.12. *In Riemannian signature, a geodesic between two points can be understood as a length minimizing curve. There is no such property in Lorentzian signature (see figure 2.1).*

The definition of a geodesic entails the existence of a conserved quantity along such a curve. Moreover, any Killing vector field will give another conserved quantity along a geodesic.

Proposition 2.7. *Consider a spacetime $(\mathcal{M}, g)$ whose metric is $\mathcal{C}^2$ (or has locally Lipschitz first derivative), let γ be a geodesic. Then the quantity*

$$g(\dot{\gamma}(s), \dot{\gamma}(s)) = g_{ab}(\gamma(s))\dot{\gamma}^a(s)\dot{\gamma}^b(s)$$

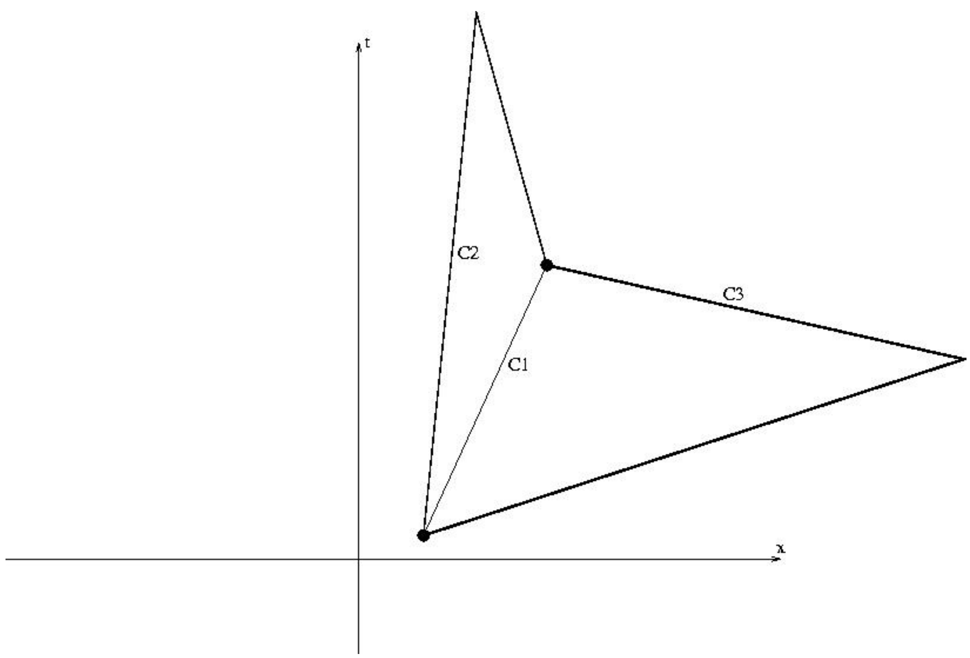

Figure 2.1: In Lorentzian signature, geodesics between two points are not extrema of the arc length. Here we consider the manifold $\mathbb{R}^2$ equipped with the Lorentzian metric $\mathrm{d}t^2 - \mathrm{d}x^2$. The curve $C1$ is a geodesic but $C2$ and $C3$ are not. The length of $C2$ is larger than that of $C1$ which is larger than that of $C3$. Also, $C2$ and $C3$ can be continuously deformed to $C1$ whilst retaining the same ordering of lengths.

is conserved along the curve. Moreover, if K is a Killing vector field on $(\mathcal{M}, g)$ or on an open neighbourhood of γ, then

$$g(\dot\gamma(s), K) = g_{ab}(\gamma(s))\dot\gamma^a(s)K^b(\gamma(s))$$

is conserved along γ.

Proof. For the first quantity, we have

$$\frac{\mathrm{d}}{\mathrm{d}s} g_{ab}\dot\gamma^a(s)\dot\gamma^b(s) = \left(\nabla_{\dot\gamma(s)} g_{ab}\right)\dot\gamma^a(s)\dot\gamma^b(s) + 2g_{ab}\dot\gamma^a(s)\nabla_{\dot\gamma(s)}\dot\gamma^b(s) = 0$$

since the connection is metric compatible and the curve γ is a geodesic.

Now for K^a a Killing vector field on $(\mathcal{M}, g)$,

$$\frac{\mathrm{d}}{\mathrm{d}s} g_{ab}K^a\dot\gamma^b(s) = \left(\nabla_{\dot\gamma(s)} g_{ab}\right)K^a\dot\gamma^b(s) + g_{ab}\left(\nabla_{\dot\gamma(s)}K^a\right)\dot\gamma^b(s) + g_{ab}K^a\nabla_{\dot\gamma(s)}\dot\gamma^b(s)\,.$$

The first term is zero since the connection is metric compatible and the third since γ is a geodesic. As for the second term, it can be written as

$$\begin{aligned}
g_{ab}\dot\gamma^b(s)\nabla_{\dot\gamma(s)}K^a &= g_{ab}g_{cd}\dot\gamma^b(s)\dot\gamma^c(s)\nabla^d K^a \\
&= g_{ab}g_{cd}\dot\gamma^b(s)\dot\gamma^c(s)\nabla^{[d}K^{a]} \text{ since } K^a \text{ is Killing} \\
&= -g_{db}g_{ca}\dot\gamma^b(s)\dot\gamma^c(s)\nabla^{[d}K^{a]} \\
&= -g_{ca}\dot\gamma^c(s)\nabla_{\dot\gamma(s)}K^a \\
&= -g_{ab}\dot\gamma^b(s)\nabla_{\dot\gamma(s)}K^a \text{ by symmetry of } g_{ab}\,.
\end{aligned}$$

This concludes the proof. □

2.7 Exercices

Exercice 2.1. *Prove property (2.24) using the definition of the flow $\Phi_V(t)$.*

Exercice 2.2. *Obtain the expression (2.25) of the Lie derivative of a 1-form.*

Exercice 2.3. *Prove proposition 2.27.*

Chapter 3

Minkowski spacetime

3.1 Definition and tangent structure

Minkowski space $\mathbb{M}$ is $\mathbb{R}^4$ endowed with the Minkowski metric, whose expression in cartesian coordinates is given by (the speed of light being equal to 1, as is common knowledge)

$$\eta = \mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \,. \tag{3.1}$$

Another useful expression of the metric η is in terms of spherical coordinates. It is particularly useful in relation to the Schwarzschild metric that we shall encounter in chapter 5. Is it a straightforward calculation to show that

$$\eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2\mathrm{d}\omega^2 \,,\ \mathrm{d}\omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2 \,, \tag{3.2}$$

where the spherical coordinates (r, θ, φ) are related to (x, y, z) by

$$x = r\sin\theta\cos\varphi \,,\ y = r\sin\theta\sin\varphi \,,\ z = r\sin\theta \,.$$

The metric $\mathrm{d}\omega^2$ defined in (3.2) is the euclidian metric on the 2-sphere.

The Minkowski metric acts on vectors at a point or vector fields on $\mathbb{M}$ as follows

$$\begin{gathered} V = V^0\frac{\partial}{\partial t} + V^1\frac{\partial}{\partial x} + V^2\frac{\partial}{\partial y} + V^3\frac{\partial}{\partial z} \,,\ W = W^0\frac{\partial}{\partial t} + W^1\frac{\partial}{\partial x} + W^2\frac{\partial}{\partial y} + W^3\frac{\partial}{\partial z} \,, \\ \eta(V, W) = \eta_{ab}V^aW^b = V^0W^0 - V^1W^1 - V^2W^2 - V^3W^3 \,, \\ \eta(V, V) = (V^0)^2 - (V^1)^2 - (V^2)^2 - (V^3)^2 \,. \end{gathered} \tag{3.3}$$

Remark 3.1. *Note that the tangent space to $\mathbb{M}$ at a given point p is $\mathbb{R}^4$ endowed with the Minkowski metric, but as a vector space. Minkowski space has the structure of an affine space. The tangent space at any given point will be referred to as Minkowski vector space. We shall see in the next chapter that it is the model for the tangent space to any spacetime.*

We see that for each point $p \in \mathbb{M}$, (3.3) distinguishes three disjoint classes of tangent vectors.

Definition 3.1. *Let $p \in \mathrm{M}$, a vector $V \in T_p\mathbb{M}$ is said to be*

- *spacelike if $\eta(V,V) < 0$ (the projection of V on the space directions is longer than its time component),*
- *null if $\eta(V,V) = 0$ (the time and space parts of the vector are of equal length),*
- *timelike if $\eta(V,V) > 0$ (the time part of the vector is longer than its space part),*
- *causal (or also non-spacelike) if $\eta(V,V) \geq 0$.*

A trajectory $\gamma : I \to \mathbb{M}$, where I is an interval of $\mathbb{R}$ and γ a differentiable function on I is said to be

- *timelike if its tangent vector $\dot\gamma(t)$ is timelike for each $t \in I$,*
- *spacelike if its tangent vector $\dot\gamma(t)$ is spacelike for each $t \in I$,*
- *null if its tangent vector $\dot\gamma(t)$ is null for each $t \in I$,*
- *causal (or non spacelike) if its tangent vector $\dot\gamma(t)$ is causal for each $t \in I$.*

Definition 3.2. *Given $p \in \mathbb{M}$, the set of null vectors in $T_p(\mathbb{M})$ is the cone*

$$C_p = \left\{ V = V^0 \frac{\partial}{\partial t} + V^1 \frac{\partial}{\partial x} + V^2 \frac{\partial}{\partial y} + V^3 \frac{\partial}{\partial z} \, ; \; (V^0)^2 = (V^1)^2 + (V^2)^2 + (V^3)^2 \right\} .$$

It is called the lightcone at p.

There are some useful orthogonality properties between vectors in the spacelike, timelike and lightlike cases. They are worth writing and proving in details since the orthogonality for an indefinite symmetric 2-form is less intuitive than for a positive definite one. First, let us introduce some notations that will be used extensively in the following proofs. Let $U \in T_p(\mathbb{M})$, we denote

$$U = U^0 \partial_t + U' ,$$

where U' is the projection of U on the spatial directions, i.e.

$$U' = U^1 \partial_x + U^2 \partial_y + U^3 \partial_z .$$

We shall also denote $|U'|$ the euclidian norm of U'

$$|U'|^2 = |U^1|^2 + |U^2|^2 + |U^3|^2 .$$

Let $U, V \in T_p(\mathbb{M})$, we denote by $\langle U', V' \rangle$ the euclidian inner product of U' and V' :

$$\langle U', V' \rangle = U^1 V^1 + U^2 V^2 + U^3 V^3 .$$

Proposition 3.1 (Orthogonal to a timelike vector)**.** *Let T a timelike vector at a point p and $V \in T_p\mathbb{M}$ such that $\eta(T,V) = 0$, then V is spacelike or zero.*

Proof. We assume that $V \neq 0$. We know that T is timelike, i.e.

$$|T^0| > |T'| .$$

Moreover,

$$\eta(T,V) = T^0V^0 - \langle T', V'\rangle = 0 .$$

This implies in particular that $V' \neq (0,0,0)$, otherwise the equality above would imply also that $V_0 = 0$ ad this would contradict $V \neq 0$. In addition, it follows that

$$|V^0| = \frac{\langle T', V'\rangle}{|T^0|} \leq \frac{|T'||V'|}{|T^0|} < |V'| .$$

This concludes the proof. □

Remark 3.2. *This means that the orthogonal in $T_p\mathbb{M}$ to a timelike vector at p for the metric η is a hyperplane in $T_p\mathbb{M}$ containing only spacelike vectors.*

The orthogonal to a spacelike vector is not necessarily timelike, a simple example is given by the vectors ∂_x and ∂_y, but if we restrict ourselves to a plane spanned by a timelike and a spacelike vector, then the result becomes true.

Proposition 3.2. *Consider at a point p in $\mathbb{M}$ a spacelike vector V and a timelike vector T. Let W a vector in the plane spanned by T and V and that is orthogonal to V, i.e. $\eta(W,V) = 0$, then W is timelike or zero.*

Proof. The restriction of η to the plane spanned by T and V is a quadratic form whose matrix in the basis $\{T, V\}$

$$A := \begin{pmatrix} \eta(T,T) & \eta(T,V) \\ \eta(T,V) & \eta(V,V) \end{pmatrix}$$

is real symmetric and has negative determinant

$$\det A = \eta(T,T)\eta(V,V) - \eta(T,V)^2 .$$

Hence A has one positive and one negative eigenvalue. In the basis $\{V\ ,\ W\}$ (assuming of course $W \neq 0$), the matrix of the quadratic form is diagonal since $\eta(V,W) = 0$. Since $\eta(V,V) < 0$ and the determinant of the matrix must still be strictly negative, it follows that $\eta(W,W) > 0$, i.e. W is timelike. □

Remark 3.3. *There is an alternative way of proving this. Since $\eta(V,V) \neq 0$, the vector W is of the form $W = \mu\,(T + \lambda_0 V)$ with $\mu \neq 0$ and we just need to show that $\tau = T + \lambda_0 V$ is timelike. The vector τ is orthogonal to V, hence*

$$\lambda_0 = -\frac{\eta(T,V)}{\eta(V,V)} .$$

Now

$$\eta(T + \lambda V, T + \lambda V) = \lambda^2\eta(V,V) + 2\lambda\eta(T,V) + \eta(T,T) .$$

This is a polynomial in λ with two real roots given by

$$\lambda_{\pm} = -\frac{\eta(T,V) \pm \sqrt{(\eta(T,V))^2 - \eta(T,T)\eta(V,V)}}{\eta(V,V)},$$

and it is positive between these two roots since $\eta(V,V) < 0$. Moreover we have

$$\lambda_0 = \frac{1}{2}(\lambda_+ + \lambda_-) ,$$

hence $\eta(\tau,\tau) > 0$. Note that the value λ_0 such that the vector $\tau = T + \lambda_0 V$ is orthogonal to V actually realizes the maximum of the quantity

$$\eta(T + \lambda V, T + \lambda V) .$$

When looking at the space of vectors orthogonal to a null vector field, the situation gets more unusual.

Proposition 3.3. *Let V be a non-zero null vector at a point p in $\mathbb{M}$. The subspace of $T_p\mathbb{M}$ of vectors orthogonal to V contains V ; except for the straight line generated by V, it is entirely composed of spacelike vectors ; it is the hyperplane tangent to the light-cone containing V.*

Proof. The fact that V is orthogonal to itself is trivial since V is assumed to be null. The vector V can be decomposed as follows

$$V = V^0 \partial_t + V' .$$

We can find two linearly independent vectors U and W in the hyperplane spanned by ∂_x, ∂_y, ∂_z which are orthogonal to V' for the euclidian inner product on $\mathbb{R}^3$. Then U, V, W are three linearly independent vectors orthogonal to V and which consequently span the hyperplane orthogonal to V. Moreover they are mutually orthogonal and since V is null and U and W are spacelike, it follows that any linear combination of the three is spacelike unless it is parallel to V. □

Definition 3.3. *Let S be a $\mathcal{C}^1$ hypersurface in $\mathbb{M}$. We say that S is :*

- *spacelike if its normal vector at each point is a timelike vector, this means that its tangent plane at each point is entirely composed of spacelike vectors ;*
- *null if its normal vector at each point is a null vector, this means that its tangent plane at each point is composed of spacelike vectors and one null direction given by the normal vector ;*
- *achronal or weakly spacelike if its normal vector at each point is a causal vector ;*
- *timelike if its normal vector at each point is a spacelike vector, this means that its tangent plane at each point is generated by one timelike and two spacelike vectors ;*

3.2 Causality

Let us consider on $\mathbb{M}$ the trajectory of a particle whose "experience" of time is described by the variable t. This is a curve $\gamma(t) = (t, x(t), y(t), z(t))$. Its tangent vector is

$$\dot{\gamma}(t) = \frac{\partial}{\partial t} + \dot{x}(t)\frac{\partial}{\partial x} + \dot{y}(t)\frac{\partial}{\partial y} + \dot{z}(t)\frac{\partial}{\partial z}$$

and

$$\eta(\dot{\gamma}(t), \dot{\gamma}(t)) = 1 - \dot{x}(t)^2 - \dot{y}(t)^2 - \dot{z}(t)^2 \,.$$

In the framework of classical mechanics, the vector

$$V(t) = \dot{x}(t)\frac{\partial}{\partial x} + \dot{y}(t)\frac{\partial}{\partial y} + \dot{z}(t)\frac{\partial}{\partial z}$$

is understood as describing the speed of the particle at time t. At a given time t, we know that the particle goes faster than, slower than, or at the speed of light, depending whether $|V(t)|^2 = \dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2 > 1$, $|V(t)|^2 < 1$ or $|V(t)|^2 = 1$. However there is nothing unique about the choice of time parameter t, it is relative to the observer. A change of time parameter t will change the value of the time component of $\dot{\gamma}$ and the length of the space part of the tangent vector will then need to be compared to some quantity other than 1 (in fact the length of the time part) to compare the speed of the particle with that of light. As a matter of fact, even the notion of time and space part is not well defined, many other choices are possible corresponding to different choices of coordinates.

In relativity, the notion that replaces that of speed vector is that of 4-velocity vector, it is $\dot{\gamma}(t)$, the tangent vector to the trajectory of the particle. This is still a non unique notion since its "length" changes with a change of parameter of the curve. Its direction however is an intrinsic notion. And this gives us an intrinsic way of comparing the speed of a particle with that of light : a particle at a given point moves faster than, slower than, or at the speed of light depending whether the tangent vector field to its trajectory at that point (measured for any choice of parameter that is not singular at that point) is spacelike, timelike or null.

A massive particle will move along a timelike curve, a massless particle will move along a null curve. If the particle is freely falling (free of any exterior influence), these curves will be geodesics.

The geodesics on Minkowski space are straight lines. This is obvious from the expression of the metric in cartesian coordinates since its coefficients are constants and therefore the connection coefficients (Christoffel symbols) are all zero. Just like general curves, they can be distinguinsed according to their timelike, spacelike, null, causal character, but of course, an important difference is that if a geodesic is timelike at a given point, it is timelike everywhere[1]. Causal geodesics can also be distinguished according to their time orientation, a notion that needs to be defined first.

Definition 3.4. *The vector field ∂_t defines a time orientation on Minkowski space. A causal vector V at a given point p is said to be future oriented (resp. past oriented) if $V^0 = \eta(V, \partial_t) > 0$ (resp. $\eta(V, \partial_t) < 0$).*

[1]This is in fact true in any spacetime as a consequence of proposition 2.7 (see proposition 4.1) but it is trivial in Minkowski space since all geodesics are straight lines.

Remark 3.4. *The previous definition is a classification of causal vectors, since for such vectors* $V^0 = \eta(V, \partial_t)$ *cannot be zero except if* $V = 0$.

Remark 3.5. *We have only defined the notion of future or past oriented for causal vectors although we could have extended it to all vectors such that* $\eta(V, \partial_t) \neq 0$. *We shall see that only causal vectors have an intrinsic time orientation. For spacelike vectors, the notion would depend on the vector field we choose as reference.*

In fact we can define a time orientation using other vector fields. The idea is that such a vector field must choose one out of the two components of the light-cone at each point, this will be labeled as the future component ; moreover it must do this is a manner that is consistent throughout the whole of Minkowski space. Such a vector field must therefore be continuous and nowhere vanishing so as to prevent it from "jumping" from one component to an incompatible one. These intuitive comments are of course far from rigorous but they seem to indicate that Minkowski spacetime will only have two time orientations, depending whether we choose to label as future the components of the light-cones containing ∂_t, or the others. Let us now give a proper definition of time orientation and prove this claim rigorously.

Definition 3.5 (Time orientation)**.** *A globally defined nowhere vanishing continuous time-like vector field* T^a *on* $\mathbb{M}$ *defines a time orientation on* $\mathbb{M}$. *For such a choice of vector field, a causal vector* V *at a given point* p *is said to be future oriented (resp. past oriented) if* $\eta(V, T(p)) > 0$ *(resp.* $\eta(V, T(p)) < 0$*).*

Proposition 3.4. *Consider a timelike vector* $T = T^0\partial_t + T^1\partial_x + T^2\partial_y + T^3\partial_z$ *and a non-zero causal vector* $V = V^0\partial_t + V^1\partial_x + V^2\partial_y + V^3\partial_z$ *at a point* p *in* $\mathbb{M}$. *Then the sign of* $\eta(T, V)$ *is that of* T^0V^0.

Proof. We have $T^0 \neq 0$ since T is timelike and $V^0 \neq 0$ since V is non-zero and causal. Then

$$\begin{aligned} \eta(p)(T, V) &= T^0V^0 - \langle T', V' \rangle \\ &= T^0V^0 \left(1 - \frac{\langle T', V' \rangle}{T^0V^0}\right). \end{aligned}$$

Now

$$\left|\frac{\langle T', V' \rangle}{T^0V^0}\right| \leq \frac{|T'|}{|T^0|}\frac{|V'|}{|V^0|} < 1$$

since T is timelike and V is non-zero an causal. Hence the result. □

This has the important following consequence.

Corollary 3.1. *Consider two time orientations of* $\mathbb{M}$ *determined respectively by two vector fields* T^a *and* τ^a. *Then one of the two following assertions is true :*

(ı) for any causal vector V *at a given point, the signs of* $\eta(V, T)$ *and* $\eta(V, \tau)$ *are the same ; the orientations are then said to be the same ; this corresponds to the case where* $\eta(T, \tau) > 0$ *;*

(ı) for any causal vector V *at a given point, the signs of* $\eta(V, T)$ *and* $\eta(V, \tau)$ *are opposite ; the orientations are then said to be opposite ; this corresponds to the case where* $\eta(T, \tau) < 0$.

Proof. Since the vector fields T^a and τ^a are timelike and continuous, then τ^0 and T^0 are nowhere vanishing and cannot change sign. The result then follows from proposition 3.4. □

This means that there are two time-orientations only on $\mathbb{M}$, the one given by ∂_t and the one given by $-\partial_t$. We choose the orientation given by ∂_t.

Proposition 3.5. *Consider a spacelike vector V at a point p. Then there exist two future oriented timelike vectors T and τ at p such that $\eta(V,T) > 0$ and $\eta(V,\tau) < 0$.*

Remark 3.6. *This shows that the time orientation of a spacelike vector V has no meaning. In fact there is even a timelike vector T at p such that $\eta(V,T) = 0$ as was clearly shown by proposition 3.2.*

Proof of proposition 3.5. We denote

$$V = V^0\partial_t + V^1\partial_x + V^2\partial_y + V^3\partial_z \,.$$

Consider for $\lambda \in \mathbb{R}$ the vector $T = T(\lambda) = \partial_t + \lambda V$. Now

$$\eta(T(\lambda), V) = V^0 + \lambda\eta(V,V)\,.$$

We know that $\eta(V,V) < 0$ since V is spacelike. For $\lambda_0 = -V^0/\eta(V,V)$, the vector $T(\lambda_0)$ is orthogonal to V and is in the plane spanned by V, a spacelike vector, and ∂_t, a timelike vector. Hence it is timelike by proposition 3.2. It is also future oriented, indeed we have

$$\eta(\partial_t, T(\lambda_0)) = 1 + \lambda_0 V^0 = 1 - \frac{(V^0)^2}{\eta(V,V)} > 0\,.$$

By continuity, $T(\lambda)$ is timelike and future oriented for λ close to λ_0. Moreover, since $\eta(T(\lambda),V)$ is affine in λ and vanishes for $\lambda = \lambda_0$, it changes sign around λ_0. We can therefore chose $\varepsilon > 0$ small enough such that $T(\lambda_0 \pm \varepsilon)$ are both timelike and future oriented and $\eta(T(\lambda_0 \pm \varepsilon), V)$ have opposite signs. □

Proposition 3.6. *Consider a causal geodesic $\gamma(s)$ on $\mathbb{M}$. Its time orientation is the same everywhere along the curve.*

Proof. There are at least two trivial ways of proving this result. First, the tangent vector to the geodesic is constant (always the same expression in the coordinate basis (t,x,y,z)), hence its time orientation is always the same. Second, the vector ∂_t is clearly a Killing vector field on $\mathbb{M}$, hence the quantity $\eta(\dot\gamma(s), \partial_t)$ is constant along the curve. □

3.3 Symmetries, Killing vectors

The symmetry group of Minkowski spacetime (preserving the metric, orientation and time-orientation) is the Poincaré group. It is the 10-dimensional group generated by the four cartesian coordinate translations, the three space rotations and the three boosts or hyperbolic rotations. The infinitesimal generators of these transformations provide the 10 independent Killing vector fields of Minkowski spacetime :

translations : ∂_t, ∂_x, ∂_y, ∂_z ;

rotations : $x\partial_y - y\partial_x$, $y\partial_z - z\partial_y$, $z\partial_x - x\partial_z$;

boosts : $x\partial_t + t\partial_x$, $y\partial_t + t\partial_y$, $z\partial_t + t\partial_z$, which are sometimes viewed as generating rotations in the planes (it, x), (it, y) and (it, z).

3.4 Exercices

Exercice 3.1. *Obtain the expression (3.2) of the Minkowski metric in spherical coordinates starting from its expression (3.1) in cartesian coordinates.*

Exercice 3.2. *Calculate the Christoffel symbols associated to the Minkowski metric for Cartesian coordinates and for spherical coordinates. Conclude that the Christoffel symbols are not a tensor field.*

Exercice 3.3. *Prove corollary 3.1 using proposition 3.4.*

Exercice 3.4. *Prove that the* 10 *vectors listed in the last section of this chapter are indeed Killing vector fields.*

Chapter 4

Curved spacetime

4.1 Tangent space, lightcones

As we have seen in the definition of Lorentzian metrics, if $(\mathcal{M}, g)$ is a spacetime, then we can find in the neighbourhood of each point an orthonormal basis. In such a basis, the metric g is described by the matrix

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The tangent space at each point is therefore a copy of Minkowski vector space. This gives us natural definitions of timelike, spacelike, null and causal vectors and a similar classification for curves and hypersurfaces.

Definition 4.1. *Let $p \in \mathrm{M}$, a vector $V \in T_p\mathcal{M}$ is said to be*

- *spacelike if $g(V, V) < 0$,*
- *null if $g(V, V) = 0$,*
- *timelike if $g(V, V) > 0$,*
- *causal (or also non-spacelike) if $g(V, V) \geq 0$.*

The definitions of timelike, spacelike, etc... for curves and hypersurfaces follow exactly as they do in Minkowski space. For geodesics, as a consequence of proposition 2.7, we have the following property :

Proposition 4.1. *Consider a geodesic $\gamma(s)$ in a spacetime $(\mathcal{M}, g)$, if γ is timelike (resp. spacelike, resp. null., resp. causal) at a given point, it is timelike (resp. spacelike, resp. null., resp. causal) everywhere.*

4.2 Causality

4.2.1 Time orientation

Definition 4.2. *A time orientation on a spacetime $(\mathcal{M}, g)$ is a globally defined nowhere vanishing continuous timelike vector field on $\mathcal{M}$. If a time orientation exists on $(\mathcal{M}, g)$, the spacetime is said to be time orientable.*

Definition 4.3. *Let $(\mathcal{M}, g)$ be a time orientable spacetime and T^a a time orientation. A causal vector V at a point is then said to be future oriented (resp. past oriented) if $g_{ab}V^aT^b > 0$ (resp. $g_{ab}V^aT^b < 0$).*

Proposition 4.2. *Let $(\mathcal{M}, g)$ be a time orientable spacetime on which we consider T^a and τ^a two time orientations. Then one of the two following assertions is true :*

- *($\imath$) for any causal vector V at a given point, the signs of $g_{ab}V^aT^b$ and $g_{ab}V^a\tau^b$ are the same ; the orientations are then said to be the same ;*
- *($\imath$) for any causal vector V at a given point, the signs of $g_{ab}V^aT^b$ and $g_{ab}V^a\tau^b$ are opposite ; the orientations are then said to be opposite.*

Proof. We consider an open covering $\{\Omega_i\}_{i\in I}$ of $\mathcal{M}$ such that on each Ω_i we can define globally an orthonormal frame

$$e_0 := \frac{T}{g(T,T)}\,,\ e_1\,,\ e_2\,,\ e_3\,.$$

Then in this frame the metric g is the Minkowski metric on Ω_i and the result follows on each Ω_i from the Minkowski case. This extends to the whole of $\mathcal{M}$ using the fact that $\{\Omega_i\}_{i\in I}$ is a covering. □

Proposition 4.3. *Let $(\mathcal{M}, g)$ be a time orientable spacetime. A spacelike vector has no time orientation. More precisely, given V a spacelike vector at a point p, there exist two choices T^a and τ^a of time orientation on $\mathcal{M}$ such that $g_{ab}V^aT^b > 0$ and $g_{ab}V^a\tau^b < 0$.*

Proof. We work in a small open neighbourhood of p in which whe can find a global orthnormal frame. The result follows from the same result for the Minkowski metric inside Ω. We then merely need to extend the two timelike vector fields to timelike vector fields defined on the whole of $\mathcal{M}$. Here is a completely explicit version of the proof.

We consider τ^a a continuous nowhere vanishing timelike vector field on $\mathcal{M}$. Let $\mathcal{U}$ be a small enough neighbourhood of p such that there exists an orthonormal basis for g on $\mathcal{U}$. We can then choose another orthonormal basis $\{e_{\mathrm{a}}{}^a\}_{\mathrm{a}=0,1,2,3}$ such that

$$e_0{}^a = \frac{\tau^a}{\sqrt{g_{ab}\tau^a\tau^b}}\,.$$

We define on $\mathcal{U}$ the vector field V^a whose components in the basis $\{e_{\mathrm{a}}{}^a\}$ are those of the vector V in the basis $\{e_{\mathrm{a}}{}^a(p)\}$ of $T_p\mathcal{M}$. Let χ a smooth compactly supported scalar function on $\mathcal{M}$ with support in $\mathcal{U}$, such that $0 \leq \chi \leq 1$ and $\chi(p) = 1$; we define the vector field W on $\mathcal{M}$ by $W^a = \chi V^a$ in $\mathcal{U}$ and $W^a\partial_a = 0$ outside of $\mathcal{U}$. Finally, for $\lambda \in \mathbb{R}$, we consider the vector field $T^a(\lambda) = \tau^a + \lambda W^a$. This is a continuous vector field on $\mathcal{M}$.

Outside the support of χ, $T^a(\lambda) = \tau^a$ and is therefore timelike. Inside the support of χ, we have

$$g_{ab}T^a(\lambda)W^b = g_{ab}\tau^a W^b + \lambda g_{ab}W^a W^b = \chi g_{ab}\tau^a V^b + \lambda\chi^2 g_{ab}V^a V^b$$

and at the point p, this takes the form

$$g(p)(T(\lambda)(p), W(p)) = g(p)(\tau(p), V) + \lambda g(p)(V, V)\,.$$

This vanishes for

$$\lambda = \lambda_0 := -\frac{g(p)(\tau(p), V)}{g(p)(V, V)}$$

which is positive if $g(p)(\tau(p), V) > 0$ and negative if $g(p)(\tau(p), V) < 0$. For this value of λ, the vector $T(\lambda_0)$ at p is in the plane spanned by τ^a and V and is orthogonal to V, it is therefore timelike, i.e. $g(p)(T(\lambda_0)(p), T(\lambda_0)(p)) > 0$. Moreover,

$$\begin{aligned} g(p)(\tau(p), T(\lambda_0)(p)) &= g(p)(\tau(p), \tau(p)) + \lambda_0 g(p)(\tau(p), V) \\ &= g(p)(\tau(p), \tau(p)) - \frac{[g(p)(\tau(p), V)]^2}{g(p)(V, V)} > 0 \end{aligned}$$

since τ is timelike and V spacelike ; i.e. the vector $T(\lambda_0)$ at point p is future oriented with respect to τ. By continuity, $\tau^a + \lambda_0 V^a$ is timelike and future oriented with respect to τ^a in a small enough neighbourhood of p, and since T^a interpolates between τ^a and $\tau^a + \lambda_0 V^a$, it is also timelike and future oriented with respect to τ^a in a small enough neighbourhood of p. Then provided the support of χ is small enough, it follows that T^a is timelike and future oriented with respect to τ^a on $\mathcal{M}$. By continuity, for $\varepsilon > 0$ small enough, this is still true of $T^a(\lambda_0 + \varepsilon)$. Then $T^a(\lambda_0 + \varepsilon)$ and $T^a(\lambda_0 - \varepsilon)$ are two timelike vector fields that are future oriented with respect to τ^a such that $g(p)(T^a(\lambda_0 \pm \varepsilon)(p), V)$ have opposite signs. □

Proposition 4.4. *Let $(\mathcal{M}, g)$ be a time orientable spacetime with a time orientation given by a vector field T^a. Consider a causal geodesic $\gamma(s)$ on $\mathcal{M}$, then its time orientation is the same at all times.*

Proof. The function $s \mapsto g(\dot\gamma(s), T)$ is continuous on the interval I on which the parameter s varies. Besides it can never be zero since the vector $\dot\gamma(s)$ cannot vanish (unless it is always zero) or become spacelike (since the geodesic is causal by assumption). □

An important notion is that of the domain of dependence of a set :

Definition 4.4. *Let $(\mathcal{M}, g)$ be a time orientable spacetime on which a time orientation has been chosen. We consider a set A in $\mathcal{M}$. The future (resp. past) domain of dependence of A in $(\mathcal{M}, g)$ is the set of points of $\mathcal{M}$ that can be reached from a point of A along a future (resp. past) oriented causal curve. These are often merely referred to as the future or the past of A. The domain of dependence of A is the reunion of its future and past domains of dependence.*

4.2.2 Global hyperbolicity

The notion of global hyperbolicity is fundamentally related to that of the Cauchy problem. Of all the equivalent definitions that have been proposed for a globally hyperbolic spacetime, the first one due to Leray, the clearest is certainly that which R.P. Geroch put forward in 1970 [10]. The fundamental definition is that of a Cauchy hypersurface.

Definition 4.5 (Cauchy hypersurface)**.** *Let $(\mathcal{M}, g)$ be a time orientable spacetime. A Cauchy hypersurface on $(\mathcal{M}, g)$ is a hypersurface Σ satisfying :*

1. *Σ is spacelike ;*
2. *every inextendible timelike curve intersects Σ at exactly one point (which entails in particular that the domain of dependence of Σ is $\mathcal{M}$).*

We see that this is an adequate surface on which to impose initial data for covariant equations (a covariant equation on a Lorentzian space-time will necessarily be a generalization to the case of a curved spacetime of covariant equations on Minkowski space, which are hyperbolic equations), since they propagate the information at finite speed lower than or equal to the speed of light, the condition that the domain of dependence of Σ should be the whole spacetime is exactly what ensures that by specifying some data on Σ, we have enough information to propagate the solution to the whole spacetime. Moreover, the first and third conditions are here to guarantee that the information propagated along causal geodesics does not come back to a point where the solution is already determined, thus creating some possible incompatibility. A globally hyperbolic spacetime as defined by Geroch is simply a spacetime that admits a Cauchy hypersurface.

Definition 4.6. *A spacetime $(\mathcal{M}, g)$ is said to be globally hyperbolic if it admits a Cauchy hypersurface.*

So globally hyperbolic spacetimes are essentially the spacetimes for which the Cauchy problem makes sense. The spacetimes in which it is hardest to make any sense at all of the Cauchy problem are called totally vicious spacetimes, they are such that any point can be reached from any other point in the spacetime along a future oriented timelike curve. An example of a totally vicious part of a spacetime is the inner part of a Kerr black hole as we shall see in chapter 7.

In fact, global hyperbolicity has stronger consequences : the existence of a smooth time function t whose level hypersurfaces Σ_t are all Cauchy hypersurfaces and are diffeomorphic to a fixed 3-surface Σ. For a long time, the only available proof of this result was due to Geroch and his construction only guaranteed the existence of a continuous time function whose level hypersurfaces were homeomorphic to a fixed hypersurface. The work of Bernal and Sanchez [1, 2] proved that the time function can be chosen smooth when the metric is smooth. Their result in fact gives a $\mathcal{C}^k$ time function when the metric is $\mathcal{C}^k$. We will assume that the metric and the time function are $\mathcal{C}^\infty$ for simplicity, we will not consider here situations in which the precise regularity of the metric and the time function may be crucial.

So it turns out that global hyperbolicity (i.e. the existence of a single smooth Cauchy hypersurface) entails the existence of a complete foliation of spacetime by smooth Cauchy hypersurfaces which are the level hypersurfaces of a smooth time function. This can then be used to decompose the geometry into space and time parts. This is referred to as a $3+1$ decomposition.

4.3 3+1 decomposition, stationarity, staticity

We consider a smooth globally hyperbolic spacetime $(\mathcal{M}, g)$. Let t be a smooth time function on $\mathcal{M}$ inducing a foliation by its level-hypersurfaces $\{\Sigma_t\}_{t\in\mathbb{R}}$ such that the hypersurfaces Σ_t are all Cauchy hpersurfaces and are diffeomorphic to a given 3-surface Σ.

A time function is defined as follows :

Definition 4.7. *A time function on $\mathcal{M}$ is a continuous function t on $\mathcal{M}$ which is strictly increasing on any future-oriented causal curve.*

The function t being smooth, this defintion can be expressed in an equivalent way.

Proposition 4.5. *A smooth function t on $\mathcal{M}$ is a time function if and only if its gradient ∇t is a smooth timelike vector field on $\mathcal{M}$.*

Proof. First we see that t is a smooth time function if and only if for any causal non-zero vector V at any given point p, we have $\mathrm{d}t(V) > 0$. This is equivalent to saying that $g(\nabla t, V) > 0$. Consider at a given point p a vector W and $\mathcal{D}^+(p)$ the set of all future-oriented non-zero causal vectors at p. We have seen that

- if W is future-oriented and timelike, then $g(W, V) > 0$ for all $V \in \mathcal{D}^+(p)$;
- if W is past-oriented and timelike, then $g(W, V) < 0$ for all $V \in \mathcal{D}^+(p)$;
- if W is spacelike, then $g(W, V)$ changes sign when V spans $\mathcal{D}^+(p)$;
- if W is future-oriented and light-like, then $g(W, V) \geq 0$ for all $V \in \mathcal{D}^+(p)$ and vanishes for V colinear to W ;
- if W is past-oriented and light-like, then $g(W, V) \leq 0$ for all $V \in \mathcal{D}^+(p)$ and vanishes for V colinear to W.

So we see that the only case where we have $g(W, V) > 0$ for all $V \in \mathcal{D}^+(p)$ is when W future-oriented and timelike. □

We use the foliation to perform a 3+1 (or space/time) decomposition of the metric. Let T^a be the future-pointing timelike vector field normal to Σ_t, normalized to 1

$$T^a T_a = 1,$$

i.e.

$$T^a = \frac{1}{|\nabla t|}\nabla^a t\,,\ \text{where } |\nabla t| = \left(g_{ab}\nabla^a t \nabla^b t\right)^{1/2}. \tag{4.1}$$

At each point $p \in \mathcal{M}$, the metric g can be decomposed into its orthogonal parts along T^a and $(T^a)^\perp = \mathrm{T}_p\Sigma_t$:

$$g_{ab} = T_a T_b - h_{ab} \tag{4.2}$$

where $-h$ is the restriction of g to $\mathrm{T}_p\Sigma_t$, whence

$$T^a h_{ab} = 0\,, \tag{4.3}$$

and the 1-form T_a is given by

$$T_a \mathrm{d}x^a = \frac{1}{|\nabla t|}\nabla_a t\,\mathrm{d}x^a = \frac{1}{|\nabla t|}\mathrm{d}t\,. \tag{4.4}$$

We define the lapse function $N(p)$ by

$$T_a \mathrm{d}x^a = N\mathrm{d}t\,,\ \text{i.e. } N = \frac{1}{|\nabla t|} \tag{4.5}$$

and the decomposition of the metric g then takes the form

$$g = N^2 \mathrm{d}t^2 - h. \tag{4.6}$$

We now choose to define the product structure using the timelike vector field $\nabla^a t$ (or equivalently T^a), the vector field $\partial/\partial t$ is then defined independently of the choice of coordinates on Σ and is everywhere orthogonal to Σ_t. More explicitely, we have

$$\left(\frac{\partial}{\partial t}\right)^a = NT^a, \tag{4.7}$$

whence

$$h_{ab}\left(\frac{\partial}{\partial t}\right)^a = 0. \tag{4.8}$$

For this choice of product structure, let us consider a local coordinate system $x^0 = t\,, x^1, x^2, x^3$ on $\mathcal{M} \simeq \mathbb{R}\times\Sigma$. From (4.8), we infer that the expression of h in these coordinates is as follows

$$h_{ab}\mathrm{d}x^a\mathrm{d}x^b = \sum_{\mathbf{a},\mathbf{b}=1}^{3} h_{\mathbf{ab}}(t, x^1, x^2, x^3)\mathrm{d}x^{\mathbf{a}}\mathrm{d}x^{\mathbf{b}}.$$

Thus h is naturally interpreted as a time-dependent Riemannian metric on Σ.

We use the decomposition of the metric to project the connection ∇_a along T^a and along $(T^a)^\perp$. We obtain

$$\nabla_a = T_a T^b\,\nabla_b - h_a{}^b\,\nabla_b = T_a\nabla_T + \mathrm{D}_a, \tag{4.9}$$

where $\nabla_T = T^a\nabla_a$ is the covariant derivative along T^a and $\mathrm{D}_a = -h_a{}^b\,\nabla_b$ is the part of ∇_a orthogonal to T^a : $T^a\mathrm{D}_a = 0$. D_a is the four-dimensional covariant derivative restricted (by composition with the projection operator $-h_a{}^b$) to act tangent to Σ_t. It differs from the Levi-Civita connection on $(\Sigma_t, h(t))$ by a combination of the extrinsic curvature (or second fundamental form) of the leaves of the foliation. In particular $\mathrm{D}_a T_b = K_{ab} = K_{(ab)}$ is the extrinsic curvature. More precisely we have

$$K_{ab} = \mathrm{D}_a T_b = h_a{}^c h_b{}^d \nabla_c T_d = -\mathcal{L}_T h_{ab} \tag{4.10}$$

and obviously $T^a K_{ab} = 0$.

Let us prove (4.10). The second fundamental form K of Σ_t is defined as

$$K_{ab} v^a w^b = g(\nabla_v T, w) = g_{bc} v^a w^b \nabla_a T^c = v^a w^b \nabla_a T_b$$

for any two tangent vectors v and w to Σ_t, i.e. K_{ab} is the restriction tangent to Σ_t of $\nabla_a T_b$. This reads

$$\begin{aligned} K_{ab} = h_a{}^c h_b{}^d \nabla_c T_d = -h_b{}^d \mathrm{D}_a T_d &= \left(g_b{}^d - T_b T^d\right) \mathrm{D}_a T_d = \mathrm{D}_a T_b - T_b T^d \mathrm{D}_a T_d \\ &= \mathrm{D}_a T_b - \mathrm{D}_a \left(T_b T^d T_d\right) + T^d T_d \mathrm{D}_a T_b + T_b T_d \mathrm{D}_a T^d \\ &= \mathrm{D}_a T_b + T_b T_d \mathrm{D}_a T^d \,. \end{aligned}$$

Adding the first and third lines above and dividing by 2, we get $K_{ab} = \mathrm{D}_a T_b$.

4.4 Stationarity, staticity

Definition 4.8. *A spacetime is said to be stationary if it admits a globally defined timelike Killing vector field. It is said to be static if it admits a globally defined Killing vector field which is orthogonal to a family of spacelike hypersurfaces.*

Remark 4.1. *Staticity simply means that the distribution of hyperplaces orthogonal to the Killing vector field is integrable. The integral hypersurfaces are necessarily spacelike and foliate the spacetime. They are not necessarily Cauchy hypersurfaces however, see for example the block III of a subextremal Reissner-Nordstrøm black hole in chapter 6. If we perform a $3+1$ decomposition of the spacetime using this foliation, the lapse function N and the metric h will be independent of t and the extrinsinc curvature will vanish.*

Chapter 5

The Schwarzschild metric

When Karl Schwarzschild obtained his explicit solution of the Einstein vacuum equations, he got the expression (5.1) given below, in the coordinate system now referred to as Schwarzschild coordinates. What people found immediately worrying was the fact that the metric was singular not only at the origin but worse, on a sphere of positive radius. The solution was quickly dismissed as physically irrelevent because of this singularity. Eddington [7] was the first to realize that the sphere was not a singularity of the metric but merely a coordinate singularity. He found a coordinate system which allowed him to give the correct interpretation of the physical meaning of the sphere. Finkelstein [9] subsequently rediscovered this coordinate system in 1958, hence the name of Eddington-Finkelstein coordinates. After Oppenheimer and Snyder proposed a model for the collapse of a star where it appeared that the phenomenon could go well beyond white dwarfs and create a singularity, people suddenly remembered Schwarzschild's solution and the study of what John Wheeler would call black holes a few years later really started. Kruskal and Szekeres [14] completed the picture and built the maximal analytic extension on the Schwarzschild metric.

The Schwarzschild metric is expressed (in a coordinate system (t, r, ω) referred to as Schwarzschild coordinates), on $\mathbb{R}_t \times]0, +\infty[_r \times S^2_\omega$ as

$$g = F(r)\mathrm{d}t^2 - F(r)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\omega^2 \,,\ \mathrm{d}\omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}^2\varphi \,,\ F(r) = 1 - \frac{2M}{r} \,, \tag{5.1}$$

where m is the mass of the black hole and $\mathrm{d}\omega^2$ is the euclidian metric on the 2-sphere. Expressed in the form (5.1), this metric appears to have two singularities corresponding to $r = 2M$ and $r = 0$. The sphere $\{r = 2M\}$, referred to as the event horizon, is merely a coordinate singularity, the metric can be extended analytically through it, while the origin $\{r = 0\}$ which is a true curvature singularity. The horizon separates the space-time in two domains :

- the exterior of the black hole $\{r > 2M\}$ is a static domain where $\partial/\partial t$ is timelike and $\partial/\partial r$ spacelike ;
- the interior of the black hole $\{r < 2M\}$, is a dynamic region where $\partial/\partial t$ is spacelike, $\partial/\partial r$ timelike, so r should be thought of as a time variable inside the black hole, it is therefore oriented ; the usual understanding of a black hole says that things can fall into it but not come out of it ; this would correspond to the inertial frames in the

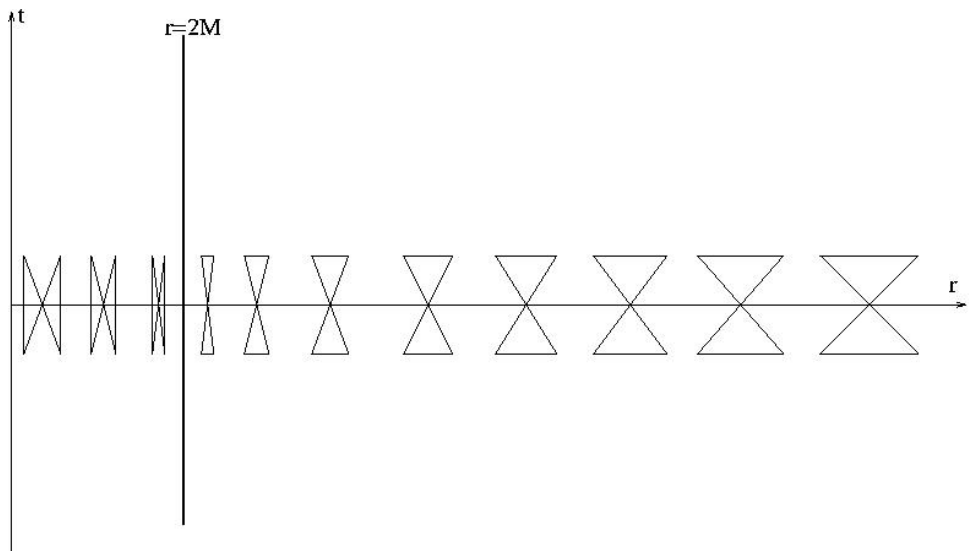

Figure 5.1: Profile of the light cones outside and inside the black-hole in the (t, r)-plane. The vectors $V^{\pm} = \partial_t \pm F\partial_r$ correspond to the upper parts of the cones.

interior being dragged towards the singularity at $\{r = 0\}$, i.e. $-\partial/\partial r$ being future oriented, but one may just as well consider the reverse time orientation which would correspond to a white hole ; nothing at this point indicates that one orientation is preferable to the other.

The two domains are globally hyperbolic. The surfaces

$$\{t\}\times]2M, +\infty[\times S^2_{\theta,\varphi}$$

are Cauchy hypersurfaces for the exterior and

$$\mathbb{R}_t \times \{r\} \times S^2_{\theta,\varphi}$$

are Cauchy hypersurfaces for the interior.

The shape of the lightcones outside and inside the black-hole is well described by the position of the null vectors

$$V^{\pm} := \frac{\partial}{\partial t} + F(r)\frac{\partial}{\partial r}\,.$$

The vectors V^+ and V^- get closer to each other as one approaches the horizon from the inside or the outside. The situation is however very different on either side of the horizon : outside the black hole, the light cones get narrower as one approaches the horizon, whereas inside they get wider (see figure 5.1).

5.1 Connection and curvature

In the Schwarzschild coordinates (t, r, θ, φ), the non zero Christoffel symbols of the Levi-Civita connection are

$$\Gamma^0_{01} = \frac{M}{r(r-2M)}\,,\ \Gamma^1_{00} = \frac{M(r-2M)}{r^3}\,,\ \Gamma^1_{11} = -\frac{M}{r(r-2M)}\,,$$
$$\Gamma^1_{22} = -(r-2M)\,,\ \Gamma^1_{33} = -(r-2M)\sin^2\theta\,,$$
$$\Gamma^2_{12} = \Gamma^3_{13} = \frac{1}{r}\,,\ \Gamma^2_{33} = -\sin\theta\cos\theta\,,\ \Gamma^3_{23} = \cot\theta\,,$$

and the non-zero components of the Riemann tensor

$$R_{0101} = -\frac{M(r-2M)}{r^2}\sin^2\theta\,,\ R_{0202} = \frac{2M}{r^3}\,,\ R_{0303} = -\frac{M(r-2M)}{r^2}\,,$$
$$R_{1212} = \frac{M}{r-2M}\sin^2\theta\,,\ R_{1313} = -2Mr\sin^2\theta\,,$$
$$R_{2323} = \frac{M}{r-2M}\,.$$

If, instead of the Schwarzschild coordinate basis, we evaluate the components of the Riemann tensor with respect to an orthonormal basis with vectors proportional to the coordinate basis vectors, namely (adopting Chandrasekhar's notations for frame indices between brackets)

$$e_{(0)}{}^a\partial_a = \frac{1}{\sqrt{F}}\frac{\partial}{\partial t}\,,\ e_{(1)}{}^a\partial_a = \sqrt{F}\frac{\partial}{\partial r}\,,\ e_{(2)}{}^a\partial_a = \frac{1}{r}\frac{\partial}{\partial\theta}\,,\ e_{(3)}{}^a\partial_a = \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}\,,$$

we find

$$R_{1010} = -R_{3232} = \frac{2M}{r^3}\,,\ R_{3131} = R_{1212} = R_{3030} = -R_{2020} = \frac{M}{r^3}\,,$$

and we see that the curvature, expressed in this frame, blows up at $\{r = 0\}$ but not at the horizon.

Remark 5.1. *Of course, if we express the components of the Riemann tensor with respect to the Schwarzschild coordinate basis, its components will be singular at $r = 0$ but also at $r = 2M$, as can readily be seen from the expression of the metric. This does not mean anything since the basis is not orthonormal. Orthonormality however is not enough to guarantee that the explosion of the coefficients of the curvature tensor corresponds to a real explosion of the curvature and not a singularity on the basis ; a basis could be singular by having an angular momentum that becomes infinite locally.*

A more usual way of seeing whether the curvature is singular is to calculate the curvature scalar which is an intrinsic quantity and is defined as follows :

$$g^{ae}g^{bf}g^{ci}g^{dj}R_{abcd}R_{efij} = R_{abcd}R^{abcd}\,.$$

It is easily calculated in the orthonormal frame above using the symmetries of the Riemann tensor :

$$R_{abcd}R^{abcd} = 32\frac{M^2}{r^6}\,.$$

5.2 Symmetries, Killing vectors, Birkhoff's theorem

Schwarzschild's spacetime has a four-dimensional space of global Killing vector fields, generated by

$$\partial_t\,,\ \ \sin\varphi\partial_\theta + \cot\theta\cos\varphi\partial_\varphi\,,\ \ \cos\varphi\partial_\theta - \cot\theta\sin\varphi\partial_\varphi\,,\ \ \partial_\varphi\,,$$

which are the timelike (outside the black hole) Killing vector field ∂_t already mentionned above and the three generators of the rotation group. In other words, the symmetry group of Schwarzschild's spacetime is $\mathbb{R}\times SO(3)$.

Schwarzschild's spacetime is the unique spherically symmetric and static solution of the Einstein vacuum equations. It is remarquable that if we remove the staticity assumption, Schwarzschild's geometry still remains the only solution. This is Birkhoff's theorem from 1923 [3] and which in fact says even a little more than this since it is a local result (or perhaps one should say semi-local since the assumption of spherical symmetry requires some sort of globality).

Theorem 5.1 (Birkhoff, 1923)**.** *If a given spacetime is spherically symmetric and satisfies the Einstein vacuum equations, then it is a part of Schwarzschild's spacetime.*

An important consequence of this theorem is that the spacetime outside a spherical uncharged star in an empty universe is Schwarzschild's spacetime.

5.3 The exterior of the black hole

We first consider the Schwarzschild geometry from the point of view of an observer static with respect to infinity. Such observers only see the exterior of the black hole and their perception of space-time is described by the time function t of the Schwarzschild coordinates outside the black hole. To their eyes, light rays falling into the black hole slow down infinitely as they approach the horizon and never cross it. One way of seeing this is to calculate the radial null geodesics.

Indeed, the fastest way of falling into the black hole, since the spacetime is spherically symmetric (i.e. in particular without rotation), is to go towards it radially and at the speed of light. Let us first evaluate the radial null directions. A radial vector at a given point (t,r,θ,φ) is of the form

$$V = \alpha\partial_t + \beta\partial_r\,.$$

For it to be null, α and β must satisfy

$$\frac{\beta}{\alpha} = F$$

since

$$g(V,V) = \alpha^2 F - \beta^2 F^{-1}\,.$$

So the two future oriented[1] radial null directions at a given point outside the black hole are those of the vectors

$$V^{\pm} = \partial_t \pm F\partial_r\,.$$

[1] Future-oriented provided we choose outside the black hole the time orientation given by ∂_t.

The apparent radial speed of these vectors for an observer static at infinity and measured using the variable r is $\pm F(r)$, it is ± 1 at infinity and slows down continuously to zero as one considers points closer and closer to the black hole horizon. Moreover, their integral curves are geodesics :

Proposition 5.1. *The radial null vectors $V^\pm$ satisfy*

$$\nabla_{V^+} V^+ = \frac{2M}{r^2} V^+ \, , \; \nabla_{V^-} V^- = -\frac{2M}{r^2} V^- \, .$$

Proof. Let us check this property for V^+. Dropping the "+" superscript for simplicity, using the values of the Christoffel symbols given above, we have

$$\begin{aligned}
\nabla_V V^{\mathbf{a}} \partial_{\mathbf{a}} &= V^{\mathbf{b}} \nabla_{\mathbf{b}} V^{\mathbf{a}} \partial_{\mathbf{a}} \\
&= V^0 \nabla_0 V^{\mathbf{a}} \partial_{\mathbf{a}} + V^1 \nabla_1 V^{\mathbf{a}} \partial_{\mathbf{a}} \\
&= \partial_t (V^{\mathbf{a}}) \partial_{\mathbf{a}} + \Gamma^{\mathbf{a}}_{0\mathbf{b}} V^{\mathbf{b}} \partial_{\mathbf{a}} + F \partial_r (V^{\mathbf{a}}) \partial_{\mathbf{a}} + F \Gamma^{\mathbf{a}}_{1\mathbf{b}} V^{\mathbf{b}} \partial_{\mathbf{a}} \\
&= 0 + \Gamma^0_{01} V^1 \partial_t + \Gamma^1_{00} V^0 \partial_r + F \partial_r (V^1) \partial_r + F \Gamma^0_{10} V^0 \partial_t + F \Gamma^1_{11} V^1 \partial_r \\
&= \frac{MF^{-1}}{r^2} F \partial_t + \frac{MF}{r^2} \partial_r + F \frac{2M}{r^2} \partial_r + F \frac{MF^{-1}}{r^2} \partial_t - F \frac{MF^{-1}}{r^2} F \partial_r \\
&= \frac{2M}{r^2} V \, .
\end{aligned}$$

The calculation is absolutely similar for V^- and left as an exercice. □

We note that the t, r-speed of radial light rays slows down as they approach the horizon. The question is whether this slowing down is strong enough to make t non-integrable along their worldlines. The answer is clearly yes since

$$\int_{2M}^{R} \frac{\mathrm{d}r}{F(r)} = \int_{2M}^{R} \frac{r \mathrm{d}r}{r - 2M} = +\infty \text{ for any } R > 2M \, .$$

This can be done in a more explicit way by introducing the Regge-Wheeler variable

$$r_* = r + 2M \mathrm{Log}(r - 2M) \tag{5.2}$$

which varies from $-\infty$ to $+\infty$ as r varies from $2M$ to $+\infty$. It satisfies

$$\frac{dr_*}{dr} = F^{-1}$$

and the metric g takes the form

$$g = F \left(\mathrm{d}t^2 - \mathrm{d}r_*^2 \right) - r^2 \mathrm{d}\omega^2 .$$

The radial null vectors take the expression

$$V^\pm = \partial_t \pm \partial_{r_*}$$

and their integral lines parametrized by r_* are the straight lines

$$\gamma^\pm_{C,\omega_0}(r_*) = \{ (t, r_*, \omega) \, ; \; \omega = \omega_0 \, , \; t = \pm r_* + C \} \; , \; C \in \mathbb{R} \, , \; \omega_0 \in S^2 \, .$$

The horizon $\{r = 2M\}$ (corresponding to $r_* \to -\infty$) is reached in infinite time t. A remarkable consequence of this property is that if we choose for a covariant field equation (Dirac, Maxwell, or the wave equation for instance) some initial data at time $t = 0$ whose support is contained in $\{r \geq 2M + \varepsilon\}$, $\varepsilon > 0$, then the support of the solution will only reach the horizon when t becomes infinite.

The intuitive description of a black hole tells us that the more we approach the horizon from the exterior, the harder it becomes to escape the attraction, until at the horizon, even a photon cannot escape anymore. But it is easier and easier to go towards the black hole. In terms of light-cones, this seems to indicate a picture where the lightcones are tilted towards the horizon and become tangent to the horizon as we reach it. When representing the lightcones in the Schwarzschild coordinates however, this does not appear to be correct after all. How do we solve this canondron? We will see that the intuitive picture has some degree of realism when we build the maximal analytic extension of the Schwarzschild spacetime, which gives the correct picture of the horizon.

An important consequence of this remark is that the interior of the black hole and the exterior should not be considered as co-existing simultaneously for the time t, in other words, a $t =$ constant slice for $r \in]0, +\infty[$ has no physical meaning whatsoever. Such hypersurfaces will be represented and put in their proper perspective once we have constructed the maximal extension of Schwarzschild's spacetime.

5.3.1 The spacelike geometry of the exterior of the black hole

The exterior of the black hole is globally hyperbolic. We consider the foliation by Cauchy hypersurfaces induced by the time function t, i.e. the slices are

$$\Sigma_t = \{t\} \times]2M, +\infty[_r \times S^2_\omega, \; t \in \mathbb{R},$$

with the induced Riemannian metric

$$h = F^{-1} \mathrm{d}r^2 + r^2 \mathrm{d}\omega^2. \tag{5.3}$$

The 3+1 decomposition of the geometry is given by (calling $\mathcal{M}$ the exterior of the black hole) :

$$\mathcal{M} = \mathbb{R}_t \times \Sigma \,, \; \Sigma =]2M, +\infty[_r \times S^2_\omega \,, \; g = F \mathrm{d}t^2 - h = \frac{N^2}{2} \mathrm{d}t^2 - h \tag{5.4}$$

with the lapse function $N = \sqrt{2} F^{1/2}$. The exterior of the black hole is static : $\frac{\partial}{\partial t}$ is a Killing vector field (since g does not depend on t), is timelike outside the black hole and is everywhere orthogonal to the Cauchy hypersurfaces Σ_t. The time orientation is chosen by deciding that $\frac{\partial}{\partial t}$ is future pointing and the normalized vector field T^a is then

$$T^a \partial_a = \sqrt{2}\, F^{-1/2} \frac{\partial}{\partial t} = \frac{2}{N} \frac{\partial}{\partial t} \,.$$

We consider a generic spacelike slice (Σ, h). The metric h appears singular at $r = 2M$. This is merely due to the choice of coordinates ; introducing as the new radial variable $u(r)$ the h-distance to the horizon, we show that (Σ, h) is a smooth manifold and that the horizon $H = \{2M\}_r \times S^2_{\theta,\varphi}$ is a smooth boundary.

Given $p = (r, \omega) \in \Sigma$, the h-distance from p to the horizon is given by

$$u(r) = \int_{[2M,r]} F^{-1/2}(s)\mathrm{d}s = \int_{[2M,r]} \frac{\sqrt{s}}{\sqrt{s-2M}}\mathrm{d}s\,. \tag{5.5}$$

This distance is finite and H thus appears as the boundary of (Σ, h). Since

$$\frac{\mathrm{d}u}{\mathrm{d}r} = F^{-1/2},$$

the metric h can be written as

$$h = \mathrm{d}u^2 + r^2\mathrm{d}\omega^2 \tag{5.6}$$

and

$$\Sigma =]0, +\infty[_u \times S^2_\omega\,.$$

The function $u(r)$ is continuous and strictly increasing from $[2M, +\infty[$ onto $[0, +\infty[$, it is $\mathcal{C}^\infty$ on $]2M, +\infty[$ but it is not differentiable at $2M$. However, the inverse function satisfies

Lemma 5.1. *The function $u \longmapsto r(u)$ is $\mathcal{C}^\infty$ on $[0, +\infty[$ and all its derivatives are uniformly bounded on $[0, +\infty[$. In particular, the first derivative $\frac{\mathrm{d}r}{\mathrm{d}u} = F^{1/2}$ (and therefore also the lapse function) is uniformly bounded as well as all its derivatives on $[0, +\infty[$.*

Proof of lemma 5.1 : the first and second derivatives $F^{1/2}$ and M/r^2 are continuous on $[0, +\infty[_u$ whence r is $\mathcal{C}^2$ on $[0, +\infty[_u$. If r is $\mathcal{C}^k$ on $[0, +\infty[_u$, then so is the second derivative and the lemma is thus proved by induction. □

This entails that h is smooth on $\bar\Sigma = [0, +\infty[_u \times S^2_\omega$; $(\bar\Sigma, h)$ is a smooth manifold with boundary. Moreover

Theorem 5.2. *The metric h is uniformly equivalent to the euclician metric on the exterior of the unit ball in $\mathbb{R}^3$*

$$\mathrm{d}u^2 + (1+u)^2\mathrm{d}\omega^2\,.$$

Proof. We see that

$$\frac{1+u}{r} \to \frac{1}{2M} \text{ as } r \to 2M\,,$$
$$\frac{1+u}{r} \to 1 \text{ as } r \to +\infty \text{ since } F(r) \to 1$$

and moreover $(1+u)/r$ is continuous on $[2M, +\infty[_r$, hence, there exists $C > 0$ such that

$$C < \frac{1+u}{r} < \frac{1}{C} \text{ for } 2M \leq r < +\infty\,.$$

This proves the theorem. □

5.3.2 Bending of light-rays : the photon sphere

We consider an extreme example of bending of light rays by gravity in the schwarzschild geometry : the photon sphere, which is a sphere of trapped geodesics around the black hole. Let us consider in the equatorial plane a null vector that is purely rotational, i.e. of the form $V = a\partial_t + b\partial_\varphi$, for example, we can take

$$V = r\partial_t + \sqrt{1 - \frac{2M}{r}}\partial_\varphi\,.$$

The integral curves of this vector field are circles in the equator (helices if we consider the time as well as space variables) whose tangent vectors are null. What is the acceleration of such curves? This is the following simple calculation :

$$\begin{aligned}
\nabla_V V = V^a \nabla_a V^b \partial_b &= \left(V^{\mathbf{a}}\nabla_{\mathbf{a}} V^{\mathbf{b}} + \Gamma^{\mathbf{b}}_{\mathbf{ac}} V^{\mathbf{c}}\right)\partial_{\mathbf{b}} \\
&= \left(V^0\partial_t V^{\mathbf{b}} + V^3\partial_\varphi V^{\mathbf{b}} + V^0\Gamma^{\mathbf{b}}_{0\mathbf{c}}V^{\mathbf{c}} + V^3\Gamma^{\mathbf{b}}_{3\mathbf{c}}V^{\mathbf{c}}\right)\partial_{\mathbf{b}} \\
&= \left(V^0\Gamma^{\mathbf{b}}_{0\mathbf{c}}V^{\mathbf{c}} + V^3\Gamma^{\mathbf{b}}_{3\mathbf{c}}V^{\mathbf{c}}\right)\partial_{\mathbf{b}} \\
&= r\left(\Gamma^0_{01}V^1\partial_t + \Gamma^1_{00}V^0\partial_r\right) \\
&\quad + \sqrt{1-\frac{2M}{r}}\left(\Gamma^1_{33}V^3\partial_r + \Gamma^3_{31}V^1\partial_\varphi + \Gamma^3_{32}V^2\partial_\varphi\right) \\
&= r\Gamma^1_{00}V^0\partial_r + \sqrt{1-\frac{2M}{r}}\Gamma^1_{33}V^3\partial_r \\
&= \left(r^2\frac{M}{r^3}(r-2M) + \left(1-\frac{2M}{r}\right)(-r)\left(1-\frac{2M}{r}\right)\right)\partial_r \\
&= \left(1-\frac{2M}{r}\right)(3M-r)\partial_r\,.
\end{aligned}$$

As could be expected, the acceleration is purely radial. It points towards the black hole if $r > 3M$, away from the black hole if $r < 3M$ and it is zero if $r = 3M$. This means that the integral curves of V for $r = 3M$ are geodesics : there are some "photon trajectories" orbiting the black hole at $r = 3M$. This is a very strong effect of light bending which requires a black hole or a very dense body of radius lower than three times its mass.

5.4 Maximal extension

After having adopted, in the previous section, the point of view of an observer static with respect to infinity, and thus limited our study to the exterior of the black hole foliated using Schwarzschild's time coordinate, we describe here briefly the global geometry of Schwarzschild's space-time. We define the Eddington-Finkelstein and the Kruskal-Szekeres coordinates inside and outside the black hole. These will allow us to show that the horizon is not a singularity of the metric. The maximal analytic extension of Schwarzschild's space-time will then appear naturally. Most of the material of this section is standard, it can be found under various forms in [5], [11] and [17] for example.

5.4.1 Eddington-Finkelstein coordinates

There are two types of Eddington-Finkelstein coordinates respectively referred to as advanced and retarded, or, more to the point, incoming and outgoing. They are based on the incoming (resp. outgoing) radial null geodesics.

The incoming Eddington-Finkelstein coordinates are

$$v = t + r_*, r, \theta, \varphi \,,$$

where $r_* = r + 2M \log(r - 2M)$ is the Regge-Wheeler coordinate. The Schwarzschild metric, in these coordinates, reads

$$g = \left(1 - \frac{2M}{r}\right) \mathrm{d}v^2 - 2\mathrm{d}v\mathrm{d}r - r^2\mathrm{d}\omega^2 \,. \tag{5.7}$$

This is fine outside the black hole but not inside where the expression of r_* is no longer valid. If we define r_* inside the black hole as

$$r_* = r + 2M \log(2M - r) \,, \tag{5.8}$$

r_* varies from $-\infty$ to $2M \log(2M)$ as r varies from $2M$ to 0. We keep the definition $v = t + r_*$ inside the black hole and we obtain the same expression (5.7) of the metric g. This is analytic on $\mathbb{R}_v \times]0, +\infty[_r \times S^2_\omega$ and does not degenerate anywhere (apart from the usual problem due to spherical coordinates) as we can see from the determinent of g :

$$\det g = -r^4 \sin^2 \theta \,.$$

The whole of Schwarzschild's spacetime is represented by the incoming Eddington-Finkelstein coordinates and we can wonder how to interpret the spacetime, and more particularly the horizon, physically.

A $v =$constant curve is a curve

$$(t = -r_* + v_0, r_*, \omega = \omega_0) \,,$$

with v_0 and ω_0 fixed ; i.e. this is an integral curve of the vector field $V^- = \partial_t - \partial_{r_*}$, in other words, a null geodesic. Outside the black hole, this is clearly the incoming radial null geodesic γ_{v_0,ω_0}. If we parametrize this curve by r, then it is an analytic curve in all positive values of r, in particular we see that the incoming null geodesic γ_{v_0,ω_0} outside the black hole extends analytically inside the black hole as the same $v = v_0$. As we follow the geodesic from infinity inwards, we move towards the future and r decreases (with r_* decreasing from $+\infty$ to $-\infty$ as r decreases from $+\infty$ to $2M$), the geodesic then crosses the horizon $\{r = 2M\}$ and keeps going towards the singularity at the origin (r_* increasing from $-\infty$ to $2M \log(2M)$ as r decreases from $2M$ to 0). The interior of the black hole is thus understood as lying in the future of the exterior. The correct time orientation of the interior of the black hole, consistent with that given by ∂_t outside the black hole, would appear to be given by $-\partial_r$.

The horizon is seen as the hypersurface $\mathbb{R}_v \times \{2M\}_r \times S^2_\omega$ and separates the exterior from the interior. Moreover, the horizon appears as a null hypersurface. Indeed, the metric does not degenerate there, but its restriction to the horizon is the 2-metric

$$-(2M)^2\mathrm{d}\omega^2 \,,$$

whereas the horizon is a 3-surface. This means that one of the tangent vectors to the horizon is null. At each point of the hypersurface $\{r = 2M\}$, the space of tangent vectors is spanned by ∂_v, ∂_θ and ∂_φ. The "squared norm" of ∂_v for the metric g is given by

$$g\left(\frac{\partial}{\partial v}, \frac{\partial}{\partial v}\right) = \left(1 - \frac{2M}{r}\right) .$$

So ∂_v is null for $r = 2M$. The correct picture of Schwarzschild's spacetime in incoming Eddington-Finkelstein coordinates is given by (FIGURE IncomEF) and we see that once inside the black hole, we cannot come back out of it.

We now perform a similar construction based on the outgoing Eddington-Finkelstein coordinates :

$$u = t - r_*, r, \theta, \varphi ,$$

and the Schwarzschild metric in these coordinates takes the expression

$$g = \left(1 - \frac{2M}{r}\right) \mathrm{d}u^2 + 2\mathrm{d}u\mathrm{d}r - r^2\mathrm{d}\omega^2 . \tag{5.9}$$

Similarly to the incoming case, this is analytic on $\mathbb{R}_u \times]0, +\infty[_r \times S^2_\omega$ and does not degenerate anywhere. The whole of Schwarzschild's spacetime is again represented, but the physical picture is different. Following an outgoing radial null geodesic (a u =constant line) towards the future, we emerge from the singularity at $r = 0$, cross the interior of the "black hole", the horizon, emerge from the "black hole" and go towards infinity. The black hole does not appear to be so black in this case since light rays emerge from it. The horizon is again a null hypersurface but this time it cannot be crossed from the exterior to the interior. This is a very different description of Schwarzschild's spacetime corresponding not to a black hole, but to a white hole (see figure OutgoEF). The time orientation of the interior consistent with the one given by ∂_t outside the black hole would now seem to correspond to ∂_r.

What we have constructed using the incoming and the outgoing Eddington-Finkelstein coordinates are similar objects but with the opposite time orientation. We shall see in the next section that the two descriptions are both present in the most complete picture of Schwarzschild's spacetime : the maximal analytic extension of it, also know as the Kruskal manifold.

5.4.2 Kruskal-Szekeres coordinates

Outside the black hole, Kruskal Szekeres coordinates (T, X, ω), ω denoting the angular variables of the Schwarzschild coordinate system, are defined by

$$T = \frac{1}{2} e^{\frac{r_*}{4M}} \left(e^{\frac{t}{4M}} - e^{-\frac{t}{4M}}\right) , \; X = \frac{1}{2} e^{\frac{r_*}{4M}} \left(e^{\frac{t}{4M}} + e^{-\frac{t}{4M}}\right) , \tag{5.10}$$

where r_* is the Regge-Wheeler variable outside the black hole given by (5.2)

$$r_* = r + 2M\mathrm{Log}(r - 2M) .$$

This coordinate system maps the exterior of the black hole $\mathbb{R}_t \times]2M, +\infty[_r \times S^2_\omega$ onto the quadrant $\{X > |T|\}$ of $\mathbb{R}_T \times \mathbb{R}_X \times S^2_\omega$. The horizon now appears as the hypersurface

$\{(T,X,\omega)\,;\ T=X>0\,,\ \omega\in S^2\}$. The outgoing (resp. incoming) radial null geodesics, represented in (t,r_*,ω) coordinates as the straight lines $\{(t,r_*=t+s,\omega)\,;\ t\in\mathbb{R}\}$ (resp. $\{(t,r_*=-t+s,\omega)\,;\ t\in\mathbb{R}\}$) for fixed $s\in\mathbb{R}$ and $\omega\in S^2$, are described in Kruskal-Szekeres coordinates as the straight lines $\{(T,X=T+S,\omega)\}$ (resp. $\{(T,X=-T+S,\omega)\}$) for fixed S and ω.

Inside the black hole, the definition is very similar. We consider the Regge-Wheeler coordinate adapted to this domain (given by (5.8))

$$r_*=r+2M\mathrm{Log}|r-2M|=r+2M\mathrm{Log}(2M-r)\,,$$

the expression of the variables T and X in terms of t and r_* is then given by

$$T=\frac{1}{2}e^{\frac{r_*}{4M}}\left(e^{-\frac{t}{4M}}+e^{\frac{t}{4M}}\right)\,,\ X=\frac{1}{2}e^{\frac{r_*}{4M}}\left(e^{-\frac{t}{4M}}-e^{\frac{t}{4M}}\right)\,. \tag{5.11}$$

The interior of the black hole $\mathbb{R}_t\times]0,2M[_r\times S^2_\omega$ is mapped onto the domain $\{(T,X,\omega)\in\mathbb{R}\times\mathbb{R}\times S^2\,;\ |X|<T<\sqrt{X^2+2M}\}$ and the singularity at $r=0$ is represented as the product of S^2_ω with the hyperbola in the (T,X)-plane : $\{(T,X)\,;\ T^2-X^2=2M\,,\ T>0\}$.

The expression of the metric in Kruskal-Szekeres coordinates is the same inside and outside the black hole

$$g=\frac{16M^2}{X^2-T^2}\left(1-\frac{2M}{r}\right)\left(\mathrm{d}T^2-\mathrm{d}X^2\right)-r^2\mathrm{d}\omega^2\,.$$

This can be simplified using the fact that

$$X^2-T^2=(r-2M)e^{\frac{r}{2M}} \tag{5.12}$$

and we obtain

$$g=\frac{16M^2}{r}e^{-\frac{r}{2M}}\left(\mathrm{d}T^2-\mathrm{d}X^2\right)-r^2\mathrm{d}\omega^2 \tag{5.13}$$

where r is determined implicitely in terms of T and X by (5.12). The function $(r-2M)e^{\frac{r}{2M}}$ is analytic in r and strictly increasing from $]0,+\infty[$ onto $]-2M,+\infty[$. It follows that r is an analytic function of X^2-T^2, and therefore of (T,X), on $-2M<X^2-T^2<+\infty$. An immediate consequence is the analyticity of the metric g on the whole Schwarzschild manifold, described in (T,X,ω) coordinates as $\{(T,X,\omega)\in\mathbb{R}\times\mathbb{R}\times S^2\,;\ T+X>0\,,\ T<\sqrt{X^2+2M}\}$ (the singularity at $r=0$ is not considered as a subset of the Schwarzschild manifold).

This construction is another way of showing that the metric g is not singular at the horizon of the black hole ; the expression (5.13) of g and the description of the horizon in (T,X,ω) coordinates reveal it to be a smooth null hypersurface of Schwarzschild's space-time. This can be seen as an alternative to the construction we performed earlier with the incoming Eddington-Finkelstein coordinates. This has an advantage over the previous construction however, it can now be extended into the "maximal Schwarzschild spacetime".

5.4.3 Maximal Schwarzschild space-time

As we have seen above, the metric (5.13) can be extended analytically on the region

$$\mathcal{M}^{\mathcal{K}}=\left\{(T,X,\omega)\in\mathbb{R}\times\mathbb{R}\times S^2_\omega\,;\ X^2-T^2>-2M\right\}\,.$$

We obtain a new space-time $(\mathcal{M}^{\mathcal{K}}, g)$ called the Kruskal extension, or maximal analytic extension, of Schwarzschild's space-time. It contains four blocks separated by a bifurcate horizon $\{|T| = |X|\}$ (see figure 5.2) :

$$\begin{aligned}
\mathrm{I} &:= \left\{(T,X,\omega)\,,\ X > |T|\,,\ \omega \in S^2\right\}\,,\\
\mathrm{II} &:= \left\{(T,X,\omega)\,,\ |X| < T < \sqrt{2M + X^2}\,,\ \omega \in S^2\right\}\,,\\
\mathrm{III} &:= \left\{(T,X,\omega)\,,\ X < -|T|\,,\ \omega \in S^2\right\}\,,\\
\mathrm{IV} &:= \left\{(T,X,\omega)\,,\ -|X| > T > -\sqrt{2M + X^2}\,,\ \omega \in S^2\right\}\,.
\end{aligned}$$

Blocks I and III are exteriors (corresponding to $r > 2M$) and the blocks II and IV are interiors (corresponding to $0 < r < 2M$). The realization of the Schwarzschild manifold that we constructed using the incoming (resp. outgoing) Eddington-Finkelstein coordinates is the union of blocks I and II (resp. I and IV) with the part of the horizon between them.

The union of blocks III and IV with the part of the horizon between them is also a realization of the Schwarzschild manifold ; it is isometric to the union of blocks I and II with the adequate part of the horizon with the time orientation reversed. More explicitely, blocks III and IV are the image of the Schwarzschild space-time, described in Schwarzschild coordinates, by the transformations (5.10) and (5.11) with the signs of T and X reversed.

The space-time $(\mathcal{M}^{\mathcal{K}}, g)$ is best pictured by a Penrose diagram, which can be constructed by defining the new coordinates (which are not smooth and only of practical use to get a picture of the general structure of $\mathcal{M}^{\mathcal{K}}$, not for any calculation) :

$$\alpha = \arctan\left(\frac{T+X}{\sqrt{2M}}\right) - \arctan\left(\frac{T-X}{\sqrt{2M}}\right)\,,$$

$$\beta = \arctan\left(\frac{T+X}{\sqrt{2M}}\right) + \arctan\left(\frac{T-X}{\sqrt{2M}}\right)\,.$$

This diagram will make more sense very soon after we have constructed the complete boundary (except for a few "points"). This will be done in chapter 8. Note that $(\mathcal{M}^{\mathcal{K}}, g)$ is globally hyperbolic, the hypersurface $\{\tau = 0\}$ is a Cauchy hypersurface.

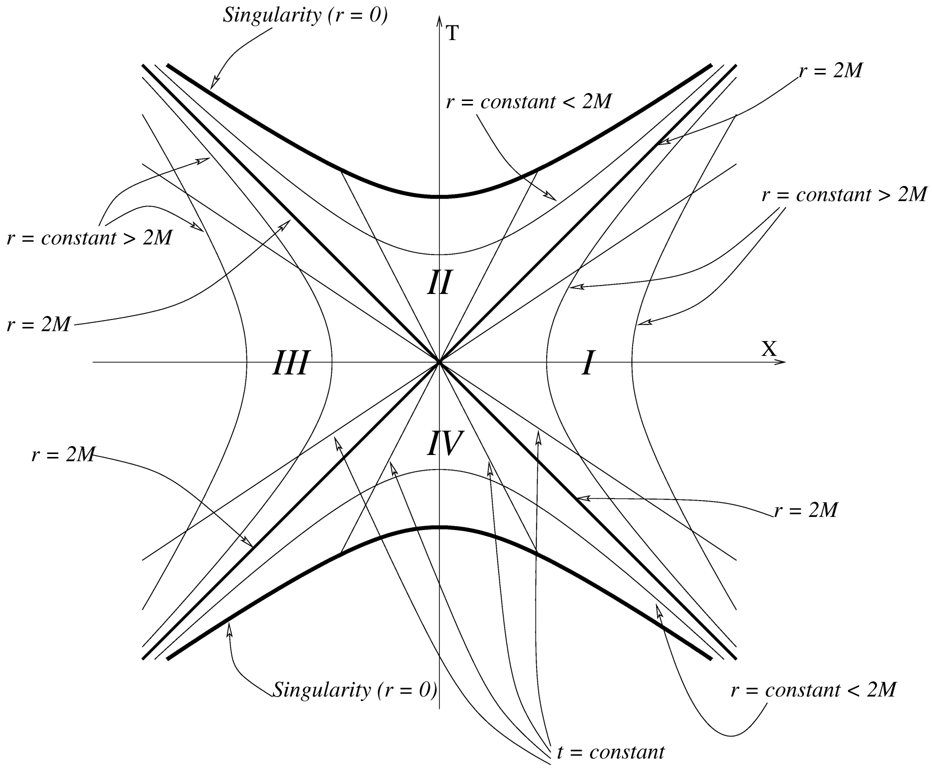

Figure 5.2: The maximal analytic extension of Schwarzschild's space-time in Kruskal-Szekeres coordinates : domains I and III correspond to $r > 2M$, domain II represents the interior of the black hole and domain IV the interior of the white hole.

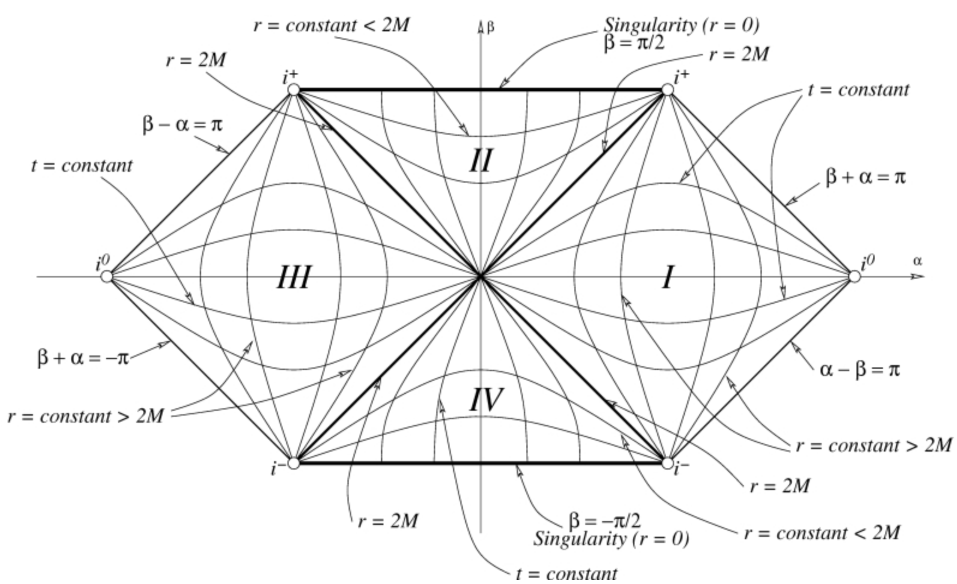

Figure 5.3: The Penrose diagram of maximal Schwarzschild space-time

5.5 exercices

Exercice 5.1. Domain of dependence

Let us consider for $2M < r_1 < r_2 < +\infty$, $-\infty < t_1 < t_2 < +\infty$, $0 < r_0 < 2M$, the domains defined in Schwarzschild coordinates by :

$$D_1 := \{(t, r, \theta, \varphi)\,,\ r_1 < r < r_2\,,\ t = 0\}\,,$$
$$D_2 := \{(t, r, \theta, \varphi)\,,\ r = r_0\,,\ t_1 < t < t_2\}\,.$$

1. *Determine the domain of dependence of D_1 in the exterior of the back hole.*
2. *Determine the domain of dependence of D_1 in the maximal extension of Schwarzschild's spacetime.*
3. *Determine the domain of dependence of D_2 in the interior of the black hole.*

Exercice 5.2. Global hyperbolicity

1. *Find a Cauchy hypersurface in the exterior of the black hole.*
2. *Find a Cauchy hypersurface in the interior of the black hole.*
3. *Find a Cauchy hypersurface in the maximal extension of Schwarzschild's spacetime.*

Exercice 5.3. Chute vers le trou noir

1. *Trouver toutes les géodésiques radiales à l'extérieur d'un trou noir de Schwarzschild.*
2. *Donner une illustration graphique du fait qu'un objet tombant directement vers le trou noir apparaît à un observateur lointain comme s'applatissant indéfiniment à l'horizon.*

Chapter 6

Other spherically symmetric black holes

The Schwarzschild metric has two extensions that still retain the feature of spherical symmetry :

- the Reissner-Nordstrøm metric describes a charged spherical static black hole ; it is no longer a solution of the Einstein vacuum equations but of the Einstein-Maxwell system, i.e. the Einstein equations with the stress-energy tensor of an electromagnetic field as a source, coupled to the Maxwell system ;
- the De Sitter-Schwarzschild metric describes a spherical eternal uncharged black hole in a universe with a positive cosmological constant.

In fact the two extensions are part of the De Sitter-Reissner-Nordstrøm family. This is the two-parameter family of metrics describing a spherical, charged, eternal black hole in a universe with a non negative cosmological constant. It is defined on $\mathbb{R}_t \times]0,+\infty[_r \times S^2_\omega$ by

$$g = F(r)\mathrm{d}t^2 - F(r)^{-1}\mathrm{d}r^2 - r^2\mathrm{d}\omega^2\,,\ F(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2} - \Lambda r^2\,, \tag{6.1}$$

where $M > 0$ is the mass of the black hole, Q its charge and $\Lambda > 0$ the cosmological constant. In the case where $\Lambda = 0$, g is the Reissner-Nordstrøm metric and when $Q = 0$, g is the De Sitter-Schwarzschild metric. When $M = Q = 0$ and $\Lambda > 0$, the geometry we obtain is known as De Sitter spacetime.

6.1 Reissner-Nordstrøm metrics

We consider the metric given by (6.1) with $\Lambda = 0$, i.e. with

$$F(r) = 1 - \frac{2M}{r} + \frac{Q^2}{r^2}\,.$$

Similarly to the case of the Schwarzschild metric, $\{r = 0\}$ is a curvature singularity and the zeros of the function F are the radii of the horizons, which are fictitious singularities that can be understood as smooth null hypersurfaces by means of Kruskal-Szekeres-type coordinates ; except now we may have two horizons. There are three types of Reissner-Nordstrøm metrics, depending on the respective importance of M and Q.

1. For $M > |Q|$, the function F has two roots
$$r_\pm := M \pm \sqrt{M^2 - Q^2}\,, \tag{6.2}$$
so the space-time has two horizons. The horizon $\{r = r_+\}$ will be called the outside horizon, or horizon of the black hole, while $\{r = r_-\}$ will be called the inner horizon.

2. For $M = |Q|$, $r_+ = r_- = M$ is the only root of F and there is only one horizon. The corresponding black hole is referred to as an extreme Reissner-Nordstrøm black hole.

3. For $M < |Q|$, the function F has no real root. There are no horizons in this case, the space-time contains no black hole and the singularity $\{r = 0\}$ is naked (i.e. not hidden beyond a horizon).

6.1.1 Sub-extremal case : $M > |Q|$

The two horizons decompose the Reissner-Nordstrøm manifold into three regions called blocks.

- *Block I* is the exterior of the black hole $\{r > r_+\}$. It is a static and globally hyperbolic region : the Killing vector ∂_t is timelike and orthogonal to the Cauchy hypersurfaces $\{t\}\times]r_+, +\infty[_r\times S^2_\omega$.
- *Block II* is the region between the two horizons $\{r_- < r < r_+\}$: it is a dynamic region where ∂_r is timelike and ∂_t is spacelike, as inside a Schwarzschild black hole. It is also globally hyperbolic.
- *Block III* is the region beyond the inner horizon $\{r < r_-\}$. It is another static region where ∂_t is Killing and timelike and orthogonal to the level hypersurfaces of t that are spacelike. But block III is not globally hyperbolic because of the singularity at $r = 0$. If we take any smooth connected spacelike hypersurface Σ in block III, there are inextendible timelike geodesics ending in the singularity and not meeting Σ. The singularity is timelike since the vector field ∂_t is timelike in block III.

The spacelike geometry of block I is similar to the Schwarzschild case in that the outer horizon is at finite spacelike distance from any point outside the black hole. This is a straightforward consequence of the fact that for any $r_0 > r_+$, the integral
$$\int_{r_+}^{r_0} \frac{1}{\sqrt{F(r)}}\mathrm{d}r = \int_{r_+}^{r_0} \frac{r}{\sqrt{(r-r_+)(r-r_-)}}\mathrm{d}r < \infty\,.$$
Similarly, in block III, the inner horizon is at finite spacelike distance from any point in block III. And so is the singularity from any point in block III.

We can define a Regge-Wheeler-type coordinate r_* in each of the three blocks easily. It needs to satisfy
$$\frac{\mathrm{d}r_*}{\mathrm{d}r} = \frac{1}{F(r)} = \frac{r^2}{(r-r_+)(r-r_-)} = 1 + \frac{r_+^2}{r_+ - r_-}\frac{1}{r-r_+} + \frac{r_-^2}{r_- - r_+}\frac{1}{r-r_-}\,,$$
i.e.
$$r_* = r + \frac{r_+^2}{r_+ - r_-}\log|r - r_+| + \frac{r_-^2}{r_- - r_+}\log|r - r_-| + R_0$$

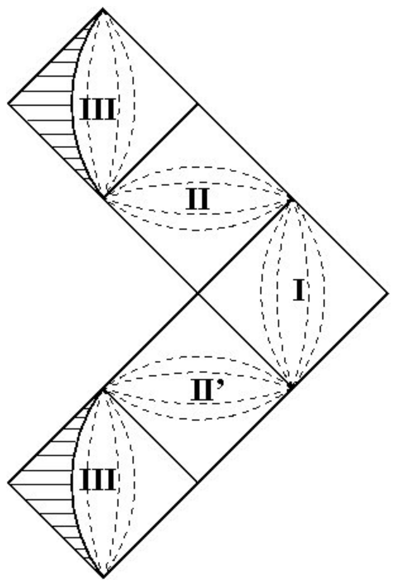

Figure 6.1: A first extension of the sub-extremal Reissner-Nordtrøm space-time drawn as a Carter-Penrose diagram. The dotted lines represent the r =cont. hypersurfaces. The thick continuous curved line is the singularity $r = 0$. The shaded regions are not part of the spacetime.

where R_0 is an arbitrary constant. In each block, the metric is expressed in terms of the variables (t, r_*, ω) as

$$g = F(r)\left(\mathrm{d}t^2 - \mathrm{d}r_*^2\right) - r^2\mathrm{d}\omega^2 .$$

This allows to construct Eddington-Finkelstein-type coordinates in each block $u = t - r_*$ and $v = t + r_*$. The corresponding expressions of the metric will be

$$\begin{aligned} g &= F(r)\mathrm{d}u^2 + 2\mathrm{d}u\mathrm{d}r - r^2\mathrm{d}\omega^2 \\ &= F(r)\mathrm{d}v^2 - 2\mathrm{d}v\mathrm{d}r - r^2\mathrm{d}\omega^2 . \end{aligned}$$

We can then glue blocks together as we did in the Schwarzschild case :

- using the incoming coordinates (v, r, ω), we glue block II to the future of block I via a future outer horizon and block III to the future of block II via a future inner horizon ;
- using the incoming coordinates (u, r, ω), we glue block II to the past of block I via a past outer horizon and block III to the past of block II via a past inner horizon.

We obtain a first extension of the Reissner-Nordstrøm manifold shown in figure 6.1. We notice that some radial null geodesics are incomplete in this picture and therefore the extension is not maximal. We can then construct the maximal analytic extention of the sub-extremal Reissner-Nordstrøm black hole by extending the incomplete radial null geodesics : the additional blocks to be glued are found by smoothness of the function r over the whole extension and by observations of time orientation. In total 6 types of blocks will be used in the construction of the maximal analytic extension : I, II, III (II here is understood as

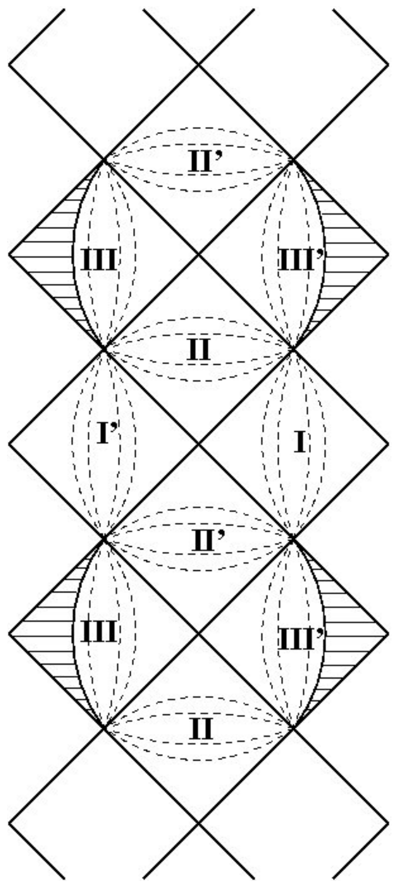

Figure 6.2: The Carter-Penrose diagram of the maximal analytic extension of the sub-extremal Reissner-Nordtrøm space-time.

having the time orientation given by $-\partial_r$) and the same blocks with their time orientation reversed I', II' and III'.

The Carter-Penrose diagram of the maximal analytic extension of the sub-extremal Reissner-Nordtrøm spacetime is shown in figure 6.2.

6.1.2 Extreme case : $M = |Q|$

Block II no longer exists in this case and only blocks I and III remain. The Regge-Wheeler coordinate r_* is now given as a primitive of the function

$$\frac{1}{F(r)} = \frac{r^2}{(r-M)^2} = 1 + \frac{2M}{r-M} + \frac{M^2}{(r-M)^2}\,,$$

i.e.

$$r_* = r + 2M \log|r-M| - \frac{M^2}{r-M}\,.$$

This has a very different behaviour at the horizon from the sub-extremal case.

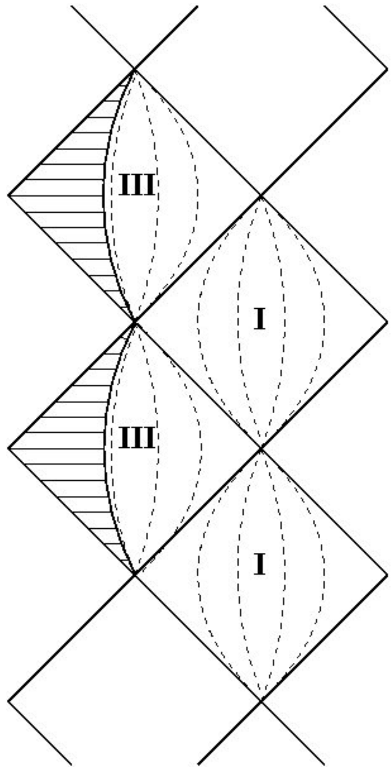

Figure 6.3: The Carter-Penrose diagram of the maximal analytic extension of the extreme Reissner-Nordtrøm space-time.

A similar difference appears in the spacelike geometry of extreme Reissner-Nordstrøm black holes. We note that

$$\int_M^{r_0} \frac{1}{\sqrt{F(r)}} \mathrm{d}r = \int_M^{r_0} \frac{r}{\sqrt{(r-M)^2}} \mathrm{d}r = \infty \,.$$

The horizon is at infinite spacelike distance from any point in block I and from any point in block III (the singularity remains of course at finite spacelike distance from points in block III).

This seems to indicate that the maximal analytic extension of an extreme Reissner-Nordstrøm black hole should be more puzzling than in the sub-extremal case. The Carter-Penrose diagram of the maximal analytic extension of the extremal Reissner-Nordtrøm spacetime is shown in figure 6.3. The structure is not so surprising, but the "spacelike horizon", i.e. $r = M$ for finite values of t is not a part of the spacetime.

6.1.3 Super-extremal case : $M < |Q|$

There is only one block with a singularity at $r = 0$. The spacetime is not extendible.

6.2 De Sitter-Schwarzschild metrics

We now study the metric (6.1) with $Q = 0$ and $\Lambda > 0$. The function $F(r)$ has the expression

$$F(r) = 1 - \frac{2M}{r} - \Lambda r^2 = -\frac{\Lambda r^3 - r + 2M}{r}.$$

The derivative $1 - 3\Lambda r^2$ of the numerator only vanishes on $]0, +\infty[$ for $r = 1/\sqrt{3\Lambda}$; it is positive in $]0, 1/\sqrt{3\Lambda}[$ and negative for $r > 1/\sqrt{3\Lambda}$. The value of the numerator at $r = 1/\sqrt{3\Lambda}$ is

$$\frac{1}{\sqrt{3\Lambda}} - \frac{1}{3\sqrt{3\Lambda}} - 2M = 2\left(\frac{1}{\sqrt{27\Lambda}} - M\right).$$

So there are three distinct situations.

- $27\Lambda M^2 < 1$: the function F has two zeros $0 < r_- < r_+ < +\infty$, there are two horizons and three blocks ; since F is positive between r_- and r_+, block II (the domain between the horizons) is static with ∂_t Killing, timelike and orthogonal to the Cauchy hypersurfaces of constant t. Block I ($r \in]0, r_-[$) is dynamic, it is the inside of the black hole or of the white hole. Block III ($r \in]r_+, +\infty[$) is also dynamic, either in expansion or in contraction. The horizon $r = r_+$ is called the cosmological horizon. The singularity at $r = 0$ is spacelike as in the Schwarzschild case.
- $27\Lambda M^2 = 1$: the function F has only one double zero at $r = 1/\sqrt{3\Lambda} = 3M$ and F is negative on either side. Block II vanishes and we only have blocks I and III. The singularity at $r = 0$ is spacelike.
- $27\Lambda M^2 > 1$: F has no zero and is negative everywhere, there is a naked singularity in a dynamic universe. The singularity at $r = 0$ is spacelike.

We give in figures 6.4, 6.5 and 6.6 the Carter-Penrose diagrams of the De Sitter-Schwarzschild space-time in the cases $27\Lambda M^2 < 1$, $27\Lambda M^2 = 1$ and $27\Lambda M^2 > 1$.

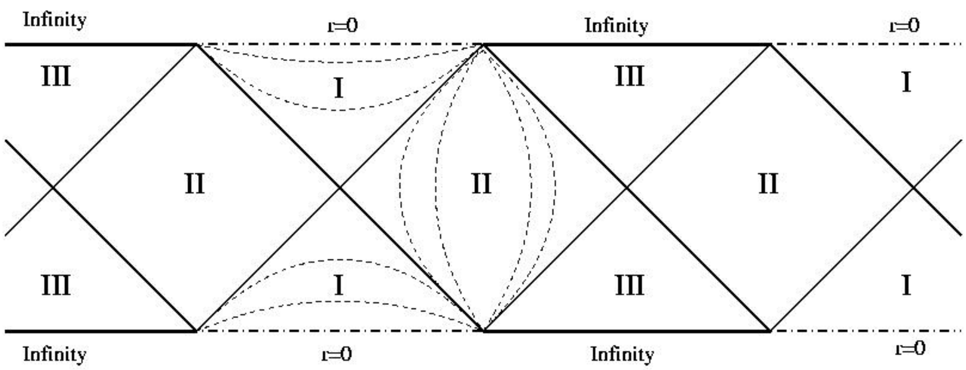

Figure 6.4: Carter-Penrose diagram of De Sitter-Schwarzschild spacetime in the case $27\Lambda M^2 < 1$.

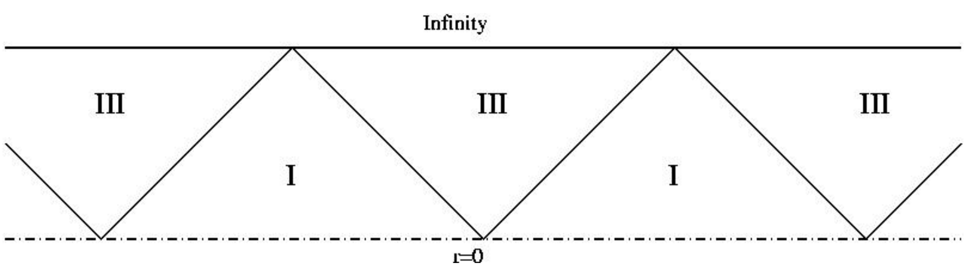

Figure 6.5: Carter-Penrose diagram of De Sitter-Schwarzschild spacetime in the case $27\Lambda M^2 = 1$.

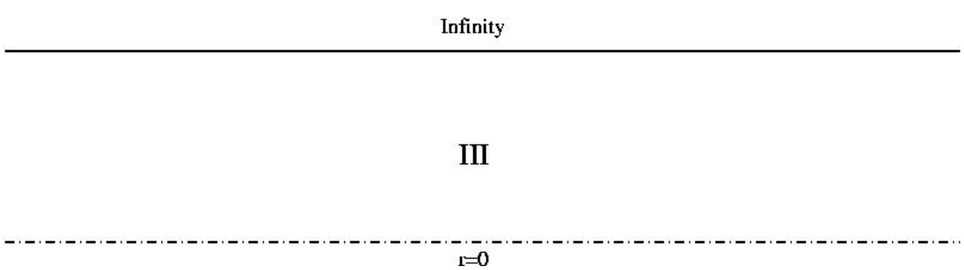

Figure 6.6: Carter-Penrose diagram of De Sitter-Schwarzschild spacetime in the case $27\Lambda M^2 > 1$.

Chapter 7

The Kerr metric

The Kerr metric is another extension of the Schwarzschild metric : it is no longer spherically symmetric. The additional parameter is the angular momentum per unit mass. It is a solution of the Einstein vacuum equations that describes a rotating uncharged black hole ; in Boyer-Lindquist coordinates on $\mathbb{R}_t \times \mathbb{R}_r \times S^2_\omega$, it takes the form

$$\begin{aligned} g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu = & \left(1 - \frac{2Mr}{\rho^2}\right)\mathrm{d}t^2 + \frac{2a\sin^2\theta\left(r^2+a^2-\Delta\right)}{\rho^2}\mathrm{d}t\mathrm{d}\varphi - \frac{\rho^2}{\Delta}\mathrm{d}r^2 \\ & -\rho^2\mathrm{d}\theta^2 - \left(\frac{\left(r^2+a^2\right)\rho^2 + 2Mra^2\sin^2\theta}{\rho^2}\right)\sin^2\theta\,\mathrm{d}\varphi^2, \end{aligned} \tag{7.1}$$

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

where a is the angular momentum per unit mass and $M > 0$ is the mass of the black hole. The black hole rotates around the axis going through its North and South poles. This results into a non-zero coefficient $g_{t\varphi}$ that couples the variables t and φ. The function Δ is the analogue of $r^2(1-2M/r)$ in Schwarzschild's space-time ; it defines the horizons as the sets of points where $\Delta = 0$. These horizons appear as singularities in the expression (7.1) above, but they are merely coordinate singularities, the metric can be extended smoothly through them. The only true curvature singularity of the metric is the equatorial ring defined by $\rho^2 = 0$, i.e. $r = 0$ and $\theta = \pi/2$. There are three types of Kerr space-times depending on the respective importance of the rotation and the mass :

- Slow Kerr space-time for $0 < |a| < M$ (the case $a = 0$ reduces to the Schwarzschild metric). Δ has two real roots r_- and r_+ :

  $$0 < r_- = M - \sqrt{M^2 - a^2} < M < r_+ = M + \sqrt{M^2 - a^2} < 2M, \tag{7.2}$$

  so there are two horizons on either side of the sphere $\{r = M\}$.

- Extreme Kerr space-time for $|a| = M$. M is then a double root for Δ and the sphere $\{r = M\}$ is the only horizon.

- Fast Kerr space-time for $|a| > M$. Δ has no real root and the space-time has no horizon. There is no black hole in this case, the ring singularity is a naked singularity.

We consider only the case of slow Kerr metrics. Horizons separate the space-time in connected regions called Boyer-Lindquist blocks :

Block I is the exterior of the black hole $\{r > r_+\}$. It is the simplest of all three blocks. In this region, the vectors $\partial/\partial r$, $\partial/\partial\theta$, $\partial/\partial\varphi$ are spacelike and, for $r >> 1$, $\partial/\partial t$ is timelike. However, block I contains a region called the ergosphere in which $g_{tt} < 0$ and thus $\partial/\partial t$ is spacelike. The ergosphere is the toroïdal domain around the outside horizon :

$$\mathcal{E} = \left\{(t,r,\theta,\varphi)\ ;\ r_+ < r < M + \sqrt{M^2 - a^2\cos^2\theta}\right\}.$$

Inside $\mathcal{E}$, the effects of the rotation are extreme and along every future-oriented non spacelike curve, the quantity $a\varphi$ is strictly increasing.

Block I, like any Boyer-Lindquist block, is not stationary, i.e. there is no timelike Killing vector field globally defined on it. However, the exterior of the ergosphere is stationary, and even absolutely stationary, since $\partial/\partial t$ is the unique (up to multiplication by a constant) timelike Killing vector field globally defined there. Also, every point in block I, even inside the ergosphere, has a stationary neighbourhood.

Block II is the region between the outer and inner horizons $\{r_- < r < r_+\}$; it only exists in the slow case. $\partial/\partial r$ is timelike there and $\partial/\partial t$, $\partial/\partial\theta$, $\partial/\partial\varphi$ are spacelike. It is a dynamic domain where the inertial frames are dragged towards the inner horizon (the time orientation implicit in this description is such that $\partial/\partial r$ is past pointing).

Block III lies beyond the inner horizon $\{-\infty < r < r_-\}$. It contains another ergosphere

$$\mathcal{E}' = \left\{(t,r,\theta,\varphi)\ ;\ M - \sqrt{M^2 - a^2\cos^2\theta} < r < r_-\right\},$$

the ring singularity and a time machine (being the only region where $\partial/\partial\varphi$ is timelike) which allows any two points in block III to be joined by a future-oriented timelike curve. Hence, not only is block III not stationary, it is not causal either.

For a detailed description of the geometry of Kerr black holes, see [19].

7.1 The exterior of the black hole

In this section, we study block I from the point of view of an observer who is static with respect to infinity. The perception of such observers is limited to block I and is described by the time function t of the Boyer-Lindquist coordinates. Just as in the Schwarzschild case, light rays in block I can only reach the horizon when t becomes infinite. To illustrate this property, we consider the principal null geodesics which play a role similar to the radial null geodesics on the Schwarzschild space-time. They are defined by

$$\dot r = \pm 1,\ \dot\theta = 0\,,\ \dot\varphi = \frac{a}{\Delta}\,,\ \dot t = \frac{r^2+a^2}{\Delta}\,.$$

Introducing a new coordinate r_* such that

$$\frac{\mathrm{d}r_*}{\mathrm{d}r} = \frac{r^2+a^2}{\Delta} > 0 \text{ on }]r_+, +\infty[$$

we get

$$\dot{r}_* = \pm \dot{t}$$

and therefore, along a principal null geodesic we must have

$$t = \pm r_* + C\,.$$

The horizon $r = r_+$ corresponds to $r_* \to -\infty$ and is consequently reached only when t becomes infinite.

We study the geometry of $\{t = \text{constant}\}$ slices ; their extrinsic geometry is non trivial and even singular at the horizon.

We denote by $\mathcal{M}$ the space-time outside the black hole and we choose the foliation of $\mathcal{M}$ by the level hypersurfaces of the time-function t :

$$\Sigma_t = \{t\} \times]r_+, +\infty[_r \times S^2_{\theta,\varphi}\,. \tag{7.3}$$

For each t, the hypersurface Σ_t is spacelike since at each point, its tangent plane is spanned by the three spacelike vectors $\frac{\partial}{\partial r}$, $\frac{\partial}{\partial \theta}$, $\frac{\partial}{\partial \varphi}$. This shows that t is indeed a time function, i.e. its gradient $\nabla^a t$ is a timelike vector field, in spite of the fact that in Boyer-Lindquist coordinates, $\frac{\partial}{\partial t}$ is not everywhere timelike in block I. The time orientation is fixed by deciding that $\nabla^a t$ is future pointing, which is equivalent to saying that $\frac{\partial}{\partial t}$ is future pointing in the region of block I where it is timelike.

7.1.1 The $3+1$ decomposition of the Kerr metric in block I

We perform the $3+1$ decomposition of the metric g relative to the foliation $\{\Sigma_t\}_{t\in\mathbb{R}}$. We calculate the expression of the vector

$$T^a = \frac{\nabla^a t}{|\nabla t|}$$

in Boyer-Lindquist coordinates. To do this, we look for a future pointing timelike vector field U^a orthogonal to Σ_t at each point and we normalize it to obtain T^a (we could also calculate the inverse metric and apply it to $\mathrm{d}t$). The time orientation yields that t increases along all timelike future pointing curves, hence we choose U^a of the form

$$U^a \partial_a = \frac{\partial}{\partial t} + A\frac{\partial}{\partial r} + B\frac{\partial}{\partial \theta} + C\frac{\partial}{\partial \varphi}$$

and imposing that U^a should be everywhere g-orthogonal to $\frac{\partial}{\partial r}$, $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \varphi}$, we obtain

$$U^a \partial_a = \frac{\partial}{\partial t} - \frac{g_{t\varphi}}{g_{\varphi\varphi}}\frac{\partial}{\partial \varphi} = \frac{\partial}{\partial t} + \frac{2aMr}{(r^2+a^2)\,\rho^2 + 2Mra^2\sin^2\theta}\frac{\partial}{\partial \varphi}\,. \tag{7.4}$$

We put

$$\alpha(r,\theta) = -\frac{g_{t\varphi}}{g_{\varphi\varphi}} = \frac{2aMr}{(r^2+a^2)\,\rho^2 + 2Mra^2\sin^2\theta}\,. \tag{7.5}$$

The norm of U^a is then given by

$$|U|^2 = U_a U^a = g_{tt} - \frac{(g_{t\varphi})^2}{g_{\varphi\varphi}} = \frac{-\Delta\sin^2\theta}{g_{\varphi\varphi}} = \frac{\Delta\rho^2}{(r^2+a^2)\,\rho^2 + 2Mra^2\sin^2\theta} > 0 \text{ in block I}\,,$$

and the vector T^a is

$$T^a = \frac{U^a}{|U|}\,.$$

If we introduce the vector fields r^a, θ^a, φ^a defined as

$$r^a\partial_a = |g_{rr}|^{-1/2}\frac{\partial}{\partial r}\,,\ \theta^a\partial_a = |g_{\theta\theta}|^{-1/2}\frac{\partial}{\partial \theta}\,,\ \varphi^a\partial_a = |g_{\varphi\varphi}|^{-1/2}\frac{\partial}{\partial \varphi}\,,$$

then $\{T^a, r^a, \theta^a, \varphi^a\}$ is a local orthonormal Lorentz frame in block I ; the metric can therefore be written as

$$g_{ab} = T_aT_b - h_{ab}\,,\ h_{ab} = r_ar_b + \theta_a\theta_b + \varphi_a\varphi_b$$

and the 1-forms T_a, r_a, θ_a and φ_a are given by

$$T_a\mathrm{d}x^a = |U|\,\mathrm{d}t = \sqrt{g_{tt} - \frac{(g_{t\varphi})^2}{g_{\varphi\varphi}}}\mathrm{d}t\,,\ r_a\mathrm{d}x^a = -|g_{rr}|^{1/2}\mathrm{d}r\,,\ \theta_a\mathrm{d}x^a = -|g_{\theta\theta}|^{1/2}\mathrm{d}\theta\,,$$

$$\varphi_a\mathrm{d}x^a = |g_{\varphi\varphi}|^{-1/2}\left(g_{t\varphi}\mathrm{d}t + g_{\varphi\varphi}\mathrm{d}\varphi\right) = -|g_{\varphi\varphi}|^{1/2}\left(\mathrm{d}\varphi - \alpha\,\mathrm{d}t\right)\,.$$

This gives the expression of the lapse function

$$N = |U| = \left(g_{tt} - \frac{(g_{t\varphi})^2}{g_{\varphi\varphi}}\right)^{1/2} = \left(\frac{\Delta\rho^2}{(r^2+a^2)\,\rho^2 + 2Mra^2\sin^2\theta}\right)^{1/2}\,.$$

In Boyer-Lindquist coordinates, the product structure is associated to the Killing vector field $\frac{\partial}{\partial t}$. If we wish our decomposition of the metric to be useful, we must interpret h_{ab} as a (time dependent) metric on

$$\Sigma\ :=]r_+, +\infty[_r \times S^2_{\theta,\varphi}\,.$$

This requires to choose the product structure associated with T^a. An explicit way of doing this is to define the new coordinates τ, R, Θ, Φ :

$$\tau = t\,,\ R = r\,,\ \Theta = \theta\,,\ \Phi = \varphi - (t-t_0)\alpha(r,\theta)\,(\mathrm{mod}\ 2\pi)$$

for a given $t_0 \in \mathbb{R}$. We obtain the following expression of g :

$$g(\tau) = N^2\mathrm{d}\tau^2 - h(\tau)$$

$$= \left(g_{tt} - \frac{(g_{t\varphi})^2}{g_{\varphi\varphi}}\right)\mathrm{d}\tau^2 + g_{rr}\mathrm{d}R^2 + g_{\theta\theta}\mathrm{d}\Theta^2 + g_{\varphi\varphi}\left(\mathrm{d}\Phi + (\tau-t_0)\frac{\partial\alpha}{\partial R}\mathrm{d}R + (\tau-t_0)\frac{\partial\alpha}{\partial\Theta}\mathrm{d}\Theta\right)^2$$

$$= \left(g_{tt} - \frac{(g_{t\varphi})^2}{g_{\varphi\varphi}}\right)\mathrm{d}\tau^2 + \left(g_{rr} + (\tau-t_0)^2\left(\frac{\partial\alpha}{\partial R}\right)^2 g_{\varphi\varphi}\right)\mathrm{d}R^2$$

$$+ \left(g_{\theta\theta} + (\tau-t_0)^2\left(\frac{\partial\alpha}{\partial\Theta}\right)^2 g_{\varphi\varphi}\right)\mathrm{d}\Theta^2 + g_{\varphi\varphi}\mathrm{d}\Phi^2$$

$$+2(\tau-t_0)^2\frac{\partial\alpha}{\partial R}\frac{\partial\alpha}{\partial\Theta}g_{\varphi\varphi}\mathrm{d}R\mathrm{d}\Theta + 2(\tau-t_0)\frac{\partial\alpha}{\partial R}g_{\varphi\varphi}\mathrm{d}R\mathrm{d}\Phi + 2(\tau-t_0)\frac{\partial\alpha}{\partial\Theta}g_{\varphi\varphi}\mathrm{d}\Theta\mathrm{d}\Phi\,. \quad (7.6)$$

Note that for these new variables, we have

$$\frac{\partial}{\partial\tau} = U^a\partial_a\,,\ \frac{\partial}{\partial R} = \frac{\partial}{\partial r}\,,\ \frac{\partial}{\partial\Theta} = \frac{\partial}{\partial\theta}\,,\ \frac{\partial}{\partial\Phi} = \frac{\partial}{\partial\varphi}\,,$$

$$T^a\partial_a = \frac{\sqrt{2}}{|U|}\frac{\partial}{\partial\tau} = \frac{2}{N}\frac{\partial}{\partial\tau}\,.$$

Remark 7.1. *The quantity α is the local rotation speed of the spacetime. We see that the function α has no singularity at r_+ (in fact it is only singular at the boundary of the time machine). And for $r = r_+$ the function α no longer depends on θ, indeed*

$$\begin{aligned}\alpha(r_+,\theta) &= \frac{2aMr_+}{(r_+^2+a^2)(r_+^2+a^2\cos^2\theta)+2Mr_+a^2\sin^2\theta}\\ &= \frac{2aMr_+}{(2Mr_+)(r_+^2+a^2\cos^2\theta+a^2\sin^2\theta)}\ \textit{since}\ r_+^2+a^2=2Mr_+\,,\\ &= \frac{a}{r_+^2+a^2}\,.\end{aligned}$$

The rotation speed of the outer horizon is the same everywhere on the horizon, it does not depend on the latitude. The same is true of the inner horizon with r_+ replaced by r_-.

7.1.2 The intrinsic and extrinsic geometry of the slices

All slices Σ_τ, $\tau \in \mathbb{R}$ have the same geometry (both intrinsic and extrinsic) since in Boyer-Lindquist coordinates, the metric g is independent of t ($\frac{\partial}{\partial t}$ is a Killing vector field). We consider a generic slice $(\Sigma, h(\tau_0))$ and we choose $t_0 = \tau_0$ in order to simplify the expression of $h(\tau_0)$:

$$h(\tau_0) = -g_{rr}\mathrm{d}R^2 - g_{\theta\theta}\mathrm{d}\Theta^2 - g_{\varphi\varphi}\mathrm{d}\Phi^2$$

$$= \frac{\rho^2}{\Delta}\mathrm{d}R^2 + \rho^2\mathrm{d}\Theta^2 + \left[\frac{(R^2+a^2)\rho^2 + 2MRa^2\sin^2\Theta}{\rho^2}\right]\sin^2\Theta\,\mathrm{d}\Phi^2,$$

$$\rho^2 = R^2 + a^2\cos^2\Theta\,,\ \Delta = R^2 - 2MR + a^2\,.$$

The coefficient ρ^2/Δ is singular at the horizon $H = \{r_+\}_R \times S^2_{\Theta,\Phi}$; we introduce a new radial coordinate to show that the metric $h(\tau_0)$ can be extended smoothly through H. Putting

$$F(R) := \frac{\Delta}{R^2} = 1 - \frac{2M}{R} + \frac{a^2}{R^2} = \frac{(R-r_+)(R-r_-)}{R^2}\,,$$

we define $u(R)$ for $R \in [r_+, +\infty[$ by

$$u(R) := \int_{r_+}^R F^{-1/2}(s)\mathrm{d}s\,.$$

(Note that for extreme Kerr space-time, we would have $r_+ = r_- = M$ and consequently, the integral defining $u(R)$ would diverge. Hence, the h-distance to the horizon would be everywhere infinite in block I.) The function u of R is continuous strictly increasing from $[r_+, +\infty[$ onto $[0, +\infty[$, it is $\mathcal{C}^\infty$ on $]r_+, +\infty[$ but is not differentiable at r_+. As in the Schwarzschild case, we easily show the following result ; the proof is identical to that of lemma 5.1 and we do not repeat it here :

Lemma 7.1. *The inverse function $u \longmapsto R(u)$ is smooth from $[0,+\infty[$ onto $[r_+,+\infty[$ and all its derivatives are uniformly bounded on $[0,+\infty[$.*

Lemma 7.1 will allow us to prove that each slice is a smooth manifold with boundary H and that the lapse function is smooth on $\bar{\Sigma}$. The following corollary expresses these properties as well as the fact that $h(\tau)$ depends regularly on τ :

Corollary 7.1. *The manifold*

$$\left(\bar{\Sigma} = [0,+\infty[_u \times S^2_{\Theta,\Phi}, h(\tau_0)\right)$$

is a smooth manifold with boundary. The lapse function N, which is independent of τ, is regular and uniformly bounded on $\bar{\Sigma}$ as well as all its derivatives. Moreover, the metric $h(\tau)$ is a smooth function of τ ; to be more explicit, we have

$$h_{ab} \in \mathcal{C}^\infty\left(\mathbb{R}_\tau\ ;\mathcal{C}_b^\infty\left(\bar{\Sigma}\ ;T_{ab}\mathcal{M}\right)\right)\ ,\ h^{ab} \in \mathcal{C}^\infty\left(\mathbb{R}_\tau\ ;\mathcal{C}_b^\infty\left(\bar{\Sigma}\ ;T^{ab}\mathcal{M}\right)\right)\ .$$

Remark 7.2. *The extrinsic curvature*

$$K_{ab} = -\mathcal{L}_T(h_{ab})$$

is singular at the horizon, however,

$$NK_{ab} \in \mathcal{C}^\infty\left(\mathbb{R}_\tau\ ;\mathcal{C}_b^\infty\left(\bar{\Sigma}\ ;T_{ab}\mathcal{M}\right)\right)\ .$$

Proof of corollary 7.1 : We write the metric $h(\tau_0)$ in the form

$$h(\tau_0) = \frac{\rho^2}{R^2}\mathrm{d}u^2 + \frac{\rho^2}{(1+u)^2}(1+u)^2\mathrm{d}\Theta^2 + \left[\frac{(R^2+a^2)\rho^2 + 2MRa^2\sin^2\Theta}{\rho^2(1+u)^2}\right](1+u)^2\sin^2\Theta\,\mathrm{d}\Phi^2.$$

The functions

$$\frac{\rho^2}{R^2}\ ,\ \frac{\rho^2}{(1+u)^2}\ ,\ \frac{(R^2+a^2)\rho^2 + 2MRa^2\sin^2\Theta}{\rho^2(1+u)^2}$$

are smooth on $\bar{\Sigma}$, positive, uniformly bounded as well as all their derivatives and uniformly bounded away from zero. Hence, $h(\tau_0)$ is a smooth, symmetric, positive definite 2-form on $\bar{\Sigma}$, uniformly controlled below and above by the euclidian metric on $\bar{\Sigma}$ considered as $\mathbb{R}^3 \setminus B(0,1)$:

$$\mathrm{d}u^2 + (1+u)^2\mathrm{d}\Theta^2 + (1+u)^2\sin^2\Theta\,\mathrm{d}\Phi^2.$$

This shows in particular that $(\bar{\Sigma}, h(\tau_0))$ is a smooth Riemannian manifold with boundary H. Given a regular coordinate system on $\bar{\Sigma}$, say the underlying euclidian coordinates on $\mathbb{R}^3 \setminus B(0,1)$, the 3×3 matrices h_{ij} and h^{ij}, representing the metric $h(\tau_0)$ and its inverse in this coordinate basis, are smooth and bounded on $\bar{\Sigma}$ as well as all their derivatives. This is expressed more intrinsically by

$$h_{ab}(\tau_0) \in \mathcal{C}_b^\infty\left(\bar{\Sigma}\ ;T_{ab}\mathcal{M}\right)\ ,\ h^{ab}(\tau_0) \in \mathcal{C}_b^\infty\left(\bar{\Sigma}\ ;T^{ab}\mathcal{M}\right)\ .$$

The lapse function N is given by

$$N(R,\Theta) = \left(\frac{2R^2\rho^2}{(R^2+a^2)\,\rho^2 + 2MRa^2\sin^2\Theta}\right)^{1/2} F^{1/2}.$$

It is the result of the multiplication of $F^{1/2}$ by a smooth function on $\bar{\Sigma}$, uniformly bounded as well as all its derivatives and uniformly bounded away from zero. Therefore, as a trivial consequence of lemma 7.1 and $\frac{\mathrm{d}R}{\mathrm{d}u} = F^{1/2}$, we have

$$N \in \mathcal{C}_b^\infty(\bar{\Sigma}) .$$

We now study the regularity of $h(\tau)$ with respect to τ. Let us consider the expressions of $h(\tau)$ and $h(\tau_0)$ in the coordinate system R, Θ, Φ with $t_0 = \tau_0$:

$$h(\tau) = -g_{rr}\mathrm{d}R^2 - g_{\theta\theta}\mathrm{d}\Theta^2 - g_{\varphi\varphi}\left(\mathrm{d}\Phi + (\tau-\tau_0)\frac{\partial\alpha}{\partial R}\mathrm{d}R + (\tau-\tau_0)\frac{\partial\alpha}{\partial\Theta}\mathrm{d}\Theta\right)^2 ,$$

$$h(\tau_0) = -g_{rr}\mathrm{d}R^2 - g_{\theta\theta}\mathrm{d}\Theta^2 - g_{\varphi\varphi}\mathrm{d}\Phi^2 .$$

Putting

$$\tilde{\Phi} = \Phi + (\tau-\tau_0)\alpha(R,\Theta) \pmod{2\pi} ,$$

we have

$$h(\tau) = -g_{rr}\mathrm{d}R^2 - g_{\theta\theta}\mathrm{d}\Theta^2 - g_{\varphi\varphi}\mathrm{d}\tilde{\Phi}^2 .$$

$h(\tau)$ is obtained from $h(\tau_0)$ by a rotation around the axis of the black hole whose angle (depending on τ, R and Θ) is

$$(\tau-\tau_0)\alpha(R,\Theta) = -(\tau-\tau_0)\frac{g_{t\varphi}(R,\Theta)}{g_{\varphi\varphi}(R,\Theta)} .$$

The function $\alpha(R,\Theta)$ is smooth on $\bar{\Sigma}$ and bounded as well as all its derivatives. Denoting by $G(\tau-\tau_0)$ the $\mathcal{C}^\infty$-diffeomorphism of $\bar{\Sigma}$

$$G(\tau-\tau_0) \ : \ (R,\Theta,\Phi) \longmapsto (R,\Theta,\Phi+(\tau-\tau_0)\alpha(R,\Theta)) ,$$

we have

$$h_{ab}(\tau) = h_{ab}(\tau_0)\circ G(\tau-\tau_0) , \ h^{ab}(\tau) = h^{ab}(\tau_0)\circ G(\tau-\tau_0) .$$

This entails

$$h_{ab} \in \mathcal{C}^\infty\left(\mathbb{R}_\tau \,; \mathcal{C}_b^\infty\left(\bar{\Sigma}\,; T_{ab}\mathcal{M}\right)\right) , \ h^{ab} \in \mathcal{C}^\infty\left(\mathbb{R}_\tau \,; \mathcal{C}_b^\infty\left(\bar{\Sigma}\,; T^{ab}\mathcal{M}\right)\right)$$

and concludes the proof of corollary 7.1. □

7.1.3 The Penrose process

If we consider a geodesic γ in block I, its energy as perceived by an observer static at infinity is

$$E_\gamma(s) := \langle \dot{\gamma}(s)\, \partial_t \rangle .$$

It is conserved since ∂_t is Killing, but since ∂_t is spacelike inside the ergosphere, for a timelike or null geodesic, the energy is allowed to be negative inside the ergosphere. This has led Roger Penrose to imagine a situation where a particle extracts some energy from the black hole : a particle is set towards the black hole ; its energy is of course positive ; once inside the ergosphere, it disintegrates into a particle with negative energy and another with positive energy ; the one with negative energy cannot leave the ergosphere and we assume that it falls into the black hole and that the other comes out of the ergosphere ; by conservation of the total energy, the energy of the particle that comes back out of the ergosphere is larger than the energy of the particle we sent towards the black hole. This is called the Penrose process.

The Penrose process has been given Subramanian Chandrasekhar the idea of a city built in orbit around a black hole and extracting its energy from the black hole. Shuttles are sent inside the ergosphere loaded with the city's litter. Inside the ergosphere, they eject the litter with an angular speed such that the energy of the litter bags is negative. The shuttles then come back lighter than they left but with more energy. A wheel slowing down the shuttles at their arrival back from the ergosphere then extracts the additional energy.

How can we choose a timelike direction τ^a at a given point inside the ergosphere such that $g(\tau, \partial_t) < 0$? Recall that for any spacelike vector at a given point, we can find two future oriented timelike vectors such that their inner products with the spacelike vector have opposite signs. In particular, this guarantees that there is a future timelike direction τ^a (and by continuity an open set of future timelike directions) such that $g(\tau, \partial_t) < 0$.

Want to be more explicit? An easy way is to look for τ of the form

$$\tau(\lambda) := \partial_t + \lambda \partial_\varphi \,.$$

Then

$$g(\tau, \tau) = g_{tt} + 2\lambda g_{t\varphi} + \lambda^2 g_{\varphi\varphi} \,.$$

We have

$$g_{t\varphi}^2 - g_{tt} g_{\varphi\varphi} = \Delta \sin^2 \theta \,.$$

This is positive in block I and we have two values of λ for which $g(\tau, \tau) = 0$:

$$\lambda_\pm = \frac{-g_{t\varphi} \pm \sqrt{g_{t\varphi}^2 - g_{tt} g_{\varphi\varphi}}}{g_{\varphi\varphi}} \,.$$

Now if we calculate $g(\tau(\lambda_\pm), \partial_t)$, we get TO BE CONTINUED...

7.1.4 Superradiance

The phenomenon of superradiance is the analogue of the Penrose process at the level of fields. TO BE CONTINUED...

7.2 Maximal extension of Kerr's space-time

The global geometry of Kerr's space-time (and in particular slow Kerr) is far more complex than that of Schwarzschild's space-time. An entire chapter of B. O'Neill's book [19] is devoted to the construction of the maximal extension. Our purpose in this section is to describe this construction schematically and to point out so-called Kruskal domains in maximal slow Kerr space-time.

7.2.1 Kerr-star and star-Kerr coordinates

Just as we did in the Schwarzschild case, we choose a coordinate system which will allow us to represent globally the whole of Kerr's space-time. This choice is guided by the following physical consideration : if a particle is to pass from block I to block II across the outer horizon and then from block II to block III across the inner horizon, its most direct course is to follow an incoming principal null geodesic. The whole idea of the Kerr-star coordinate system is to turn incoming principal null geodesics into coordinate lines. Such geodesics are defined on all three blocks in Boyer-Lindquist coordinates by

$$\dot{t} = \frac{r^2+a^2}{\Delta}\,,\ \dot{r} = -1\,,\ \dot{\theta} = 0\,,\ \dot{\varphi} = \frac{a}{\Delta}\,.$$

Keeping the coordinates r and θ, we introduce two new coordinates t^* and φ^* of the form

$$t^* = t + T(r)\,,\ \varphi^* = \varphi + A(r)$$

where the functions T and A are required to satisfy

$$\frac{\mathrm{d}T}{\mathrm{d}r} = \frac{r^2+a^2}{\Delta}\,,\ \frac{\mathrm{d}A}{\mathrm{d}r} = \frac{a}{\Delta}\,.$$

$(t^*, r, \theta, \varphi^*)$ defines a coordinate system in each Boyer-Lindquist block[1], called Kerr-star coordinates, in which the incoming principal null geodesics are described by

$$\dot{r} = -1\,,\ \dot{\theta} = 0\,,\ \dot{t^*} = \dot{t} + \frac{\mathrm{d}T}{\mathrm{d}r}\dot{r} = 0\,,\ \dot{\varphi^*} = \dot{\varphi} + \frac{\mathrm{d}A}{\mathrm{d}r}\dot{r} = 0\,,$$

i.e. they are the r coordinate curves parametrized by $s = -r$ (or $-r + C$). The expression of the Kerr metric in Kerr-star coordinates is given by

$$g = g_{tt}\mathrm{d}t^{*2} + 2g_{t\varphi}\mathrm{d}t^*\mathrm{d}\varphi^* + g_{\varphi\varphi}\mathrm{d}\varphi^{*2} - \rho^2\mathrm{d}\theta^2 - 2\mathrm{d}t^*\mathrm{d}r + 2a\sin^2\theta\,\mathrm{d}\varphi^*\mathrm{d}r\,, \tag{7.7}$$

where g_{tt}, $g_{t\varphi}$, $g_{\varphi\varphi}$ and $g_{\theta\theta} = -\rho^2$ are as defined in (7.1), i.e.

$$g_{tt} = \left(1 - \frac{2Mr}{\rho^2}\right)\,,\ g_{t\varphi} = \frac{a\sin^2\theta\left(r^2+a^2-\Delta\right)}{\rho^2}\,,$$

$$g_{\varphi\varphi} = -\left(\frac{\left(r^2+a^2\right)\rho^2 + 2Mra^2\sin^2\theta}{\rho^2}\right)\sin^2\theta\,,\ \rho^2 = r^2 + a^2\cos^2\theta\,.$$

We see from (7.7) that the metric g is smooth on all three blocks, with the exception of the ring singularity $\{\rho^2 = 0\} = \{r = 0 \text{ and } \theta = \pi/2\}$ in block III, and across both horizons (the component g_{rr} in Boyer-Lindquist coordinates was the only component of g to be singular at the horizons and it does not appear in (7.7)).

Kerr-star space-time is defined as the manifold

$$\mathcal{M}^* = \mathbb{R}_{t^*} \times \mathbb{R}_r \times S^2_{\theta,\varphi^*} \setminus \left\{(t^*, r, \theta, \varphi^*)\,;\ r = 0 \text{ and } \theta = \frac{\pi}{2}\right\}$$

[1] with the exception of the axis ($\theta = 0$ and $\theta = \pi$) ; this coordinate singularity can be dealt with simply (see [19] lemma 2.2.2), we shall systematically ignore it.

equipped with the smooth metric (7.7) and with the time orientation such that the null coordinate vector field $-\frac{\partial}{\partial r}$, defined and smooth on the whole of $\mathcal{M}^*$ and whose integral lines are the incoming principal null geodesics, be future oriented. This time orientation is consistent with the fact that, in Boyer-Lindquist coordinates, the Killing vector field $\frac{\partial}{\partial t}$ is future oriented outside the ergosphere in block I and also with the description of block II given at the beginning of the chapter, with $-\frac{\partial}{\partial r}$ (in Boyer-Lindquist coordinates) future pointing. This space-time contains all three blocks, glued smoothly at the horizons by the requirement that incoming principal null geodesics should cross horizons smoothly and that their orientation defines the time orientation. Block II is thus glued to block I in such a way that it lies in the future of block I and similarly, block III lies in the future of block II. The horizons $\{r = r_+\}$ and $\{r = r_-\}$ are smooth null hypersurfaces of $(\mathcal{M}^*, g)$. The fact that they are null is easily shown considering the metric induced by g on hypersurfaces of constant r

$$g_r = g_{tt}\mathrm{d}t^{*2} + 2g_{t\varphi}\mathrm{d}t^*\mathrm{d}\varphi^* + g_{\varphi\varphi}\mathrm{d}\varphi^{*2} - \rho^2\mathrm{d}\theta^2\,.$$

This induced metric has determinent

$$\det(g_r) = -\rho^2\left(g_{tt}g_{\varphi\varphi} - (g_{t\varphi})^2\right) = \rho^2\Delta\sin^2\theta$$

and thus degenerates for $\Delta = 0$, i.e. at the horizons. See figure 7.1 for a Penrose diagram of Kerr-star space-time.

This construction is similar to what we did in Schwarzschild's space-time, when we first used Kruskal-Szekeres coordinates to show that the metric could be extended smoothly across the horizon. In the Schwarzschild case, the maximal extension of the space-time followed naturally by extending the domain of definition of the Kruskal-Szekeres coordinate system. This we cannot do here since the domain of definition of Kerr-star coordinates is already maximal. We shall need to use other coordinate systems which will allow us to glue Boyer-Lindquist blocks in different manners.

Kerr-star coordinates were defined by modifying Boyer-Lindquist coordinates so that incoming principal null geodesics could become coordinate lines. Using outgoing principal null geodesics instead of the incoming ones, we obtain the star-Kerr coordinate system. These geodesics are defined on all three blocks in Boyer-Lindquist coordinates by

$$\dot{t} = \frac{r^2+a^2}{\Delta}\,,\ \dot{r} = 1\,,\ \dot{\theta} = 0\,,\ \dot{\varphi} = \frac{a}{\Delta}\,.$$

Keeping r and θ, we introduce the new coordinates

$${}^*t = t - T(r)\,,\ {}^*\varphi = \varphi - A(r)$$

where the functions T and A are the same used to define t^* and φ^*. In the star-Kerr coordinate system $({}^*t, r, \theta, {}^*\varphi)$, the outgoing principal null geodesics are the r coordinate lines parametrized by $s = r$ and the Kerr metric takes the form

$$\begin{aligned} g = {} & g_{tt}\mathrm{d}({}^*t)^2 + 2g_{t\varphi}\mathrm{d}({}^*t)\mathrm{d}({}^*\varphi) + g_{\varphi\varphi}\mathrm{d}({}^*\varphi)^2 - \rho^2\mathrm{d}\theta^2 \\ & +2\mathrm{d}({}^*t)\mathrm{d}r - 2a\sin^2\theta\,\mathrm{d}({}^*\varphi)\mathrm{d}r\,. \end{aligned} \tag{7.8}$$

This gives rise to star-Kerr space-time which is the manifold

$${}^*\mathcal{M} = \mathbb{R}_{{}^*t} \times \mathbb{R}_r \times S^2_{\theta,{}^*\varphi} \setminus \left\{({}^*t, r, \theta, {}^*\varphi)\,;\ r = 0 \text{ and } \theta = \frac{\pi}{2}\right\}$$

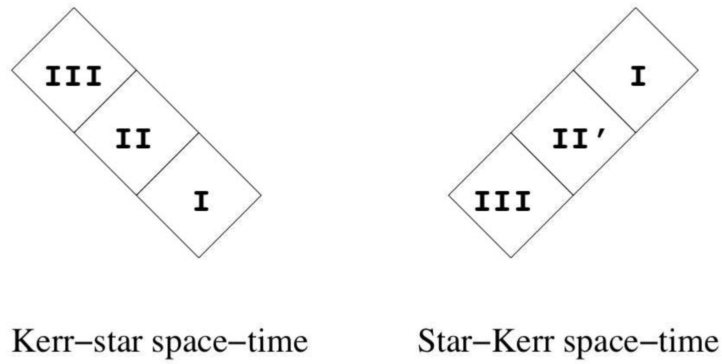

Figure 7.1: Penrose diagrams of Kerr-star and star-Kerr space-times

equipped with the smooth metric (7.8) and time orientation such that, in star-Kerr coordinates, the null coordinate vector field $\frac{\partial}{\partial r}$, which is defined and smooth all over ${}^*\mathcal{M}$ and whose integral lines are the outgoing principal null geodesics, is future pointing. This space-time contains all three blocks, glued together at the horizons which appear as regular null hypersurfaces. The gluing is done by requiring that the outgoing principal null geodesics should cross the horizons smoothly. The time orientation reflects this choice ; it is consistent with the fact that in Boyer-Lindquist coordinates $\frac{\partial}{\partial t}$ is future pointing outside the ergosphere in block I, but incompatible with $-\frac{\partial}{\partial r}$ future oriented in block II : in star-Kerr space-time, the inertial frames in bloc II are dragged outwards from the inner horizon to the outer horizon. There is a canonical isometry between star-Kerr and Kerr-star space-times. This isometry preserves the time orientation of blocks I and III but reverses that of block II. Star-Kerr space-time can be seen as a block I, to the past of which is glued a block II with its time orientation reversed, to the past of which is glued a block III : it describes a "slow Kerr white hole". See figure 7.1 for the Penrose diagram of star-Kerr space-time (II′ refers to a block II with reversed time orientation).

7.2.2 Maximal slow Kerr space-time

The maximal analytic extension of slow Kerr space-time is constructed using both Kerr-star and star-Kerr space-times. We start with Kerr-star space-time : all the incoming principal null geodesics are complete but the outgoing ones are not. The idea is to glue other blocks so as to make the outgoing principal null geodesics complete. The solution for blocks I and III is simple : we consider them as belonging to star-Kerr space-times, i.e. we glue to the future of block III a block II′ followed by a new block I and to the past of block I a block II′ preceded by a new block III. For block II, the situation is trickier ; we also wish to understand block II as part of a star-Kerr space-time, but this is incompatible with the time orientation of block II. The solution is to reverse the time orientation of the whole star-Kerr space-time. We are thus led to gluing to the future of block II a block III′ (block III with its time orientation reversed) and to its past a block I′ (block I with reversed time orientation). The resulting space-time is shown in figure 7.3. We keep on extending this new space-time wherever a family of principal null geodesics is incomplete. The extension is done step by step and is based on the same simple principle : if a family of principle null geodesics is incomplete, it means that the Kerr-star (in the incoming case) or star-Kerr (in the outgoing case) space-time which it generates lacks one or two blocks ; this is cured by gluing the lacking blocks, bearing in mind the consistency of the time orientation of the whole space-time. In this manner, we construct maximal slow Kerr space-time (see figure 7.2) as a reunion of four types of space-times : Kerr-star space-times, Kerr-star with their time orientation reversed, star-Kerr and star-Kerr with their time orientation reversed. Important objects in this maximal extension are the so-called Kruskal domains. They are “diamond shaped” reunions of four contiguous blocks. At their “centre” lies a 2-sphere, referred to as the crossing sphere, where the horizons intersect. Building this crossing sphere rigorously and extending the metric over it are important difficulties in the construction of maximal slow Kerr space-time. This is done by means of Kruskal-Boyer-Lindquist coordinates (see [19] for a fully detailed account). There are two types of Kruskal domains, as shown in figure 7.4. Type II-III contains two copies of block III ; it is not causal, therefore not globally hyperbolic, and contains two timelike singularities (the ring singularity of each block III). Because of the lack of causality, the notion of Cauchy problem is not even meaningful on type II-III domains. Type I-II domains are much more gentle. They are globally hyperbolic and contain no singularity. They can be treated in exactly the same manner as maximal Schwarzschild space-time.

For a type I-II Kruskal domain, we consider a foliation $\{S_\tau\}_{\tau\in\mathbb{R}}$ (see figure 7.5) by Cauchy hypersurfaces such that, outside the domain of dependence of a neighbourhood of the crossing sphere, for each $\tau \in \mathbb{R}$ the hypersurface S_τ coincides in block I with the level hypersurface $\Sigma_\tau = \{t = \tau\}$ of the time coordinate t of Boyer-Lindquist coordinates and in block I′ with $\Sigma_{-\tau}$ (suffice it to say that the Boyer-Lindquist coordinates in blocks I, II, I′ and II′ are defined unambiguously from the Kruskal-Boyer-Lindquist coordinates defined on the whole domain).

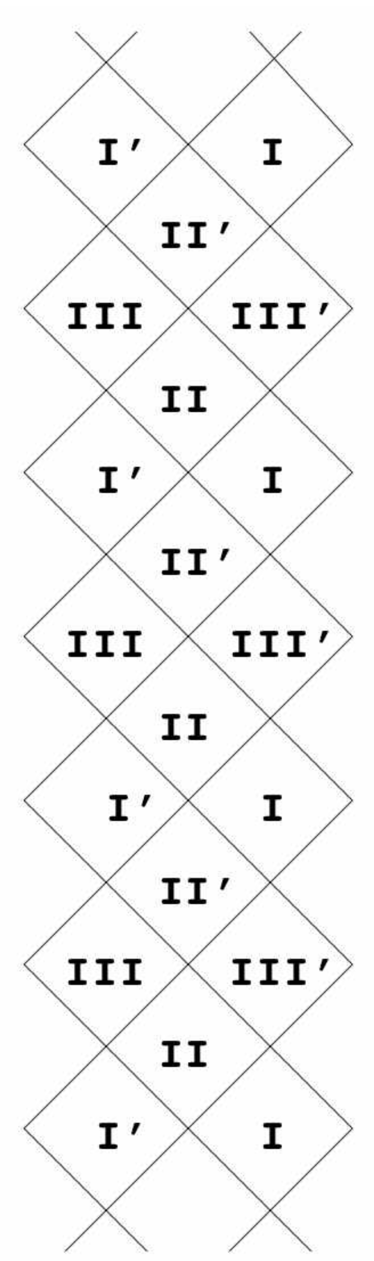

Figure 7.2: Maximal slow Kerr space-time

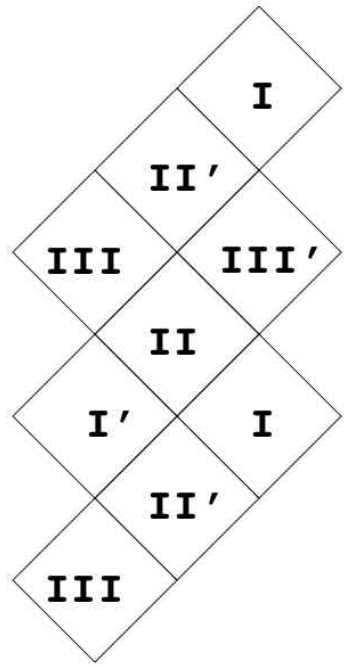

Figure 7.3: First step in the construction of maximal slow Kerr space-time

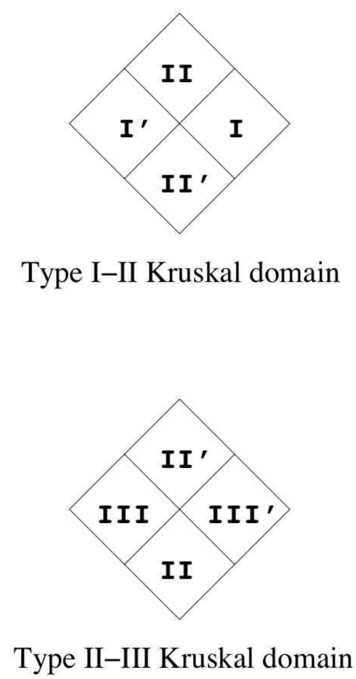

Figure 7.4: The two different types of Kruskal domains

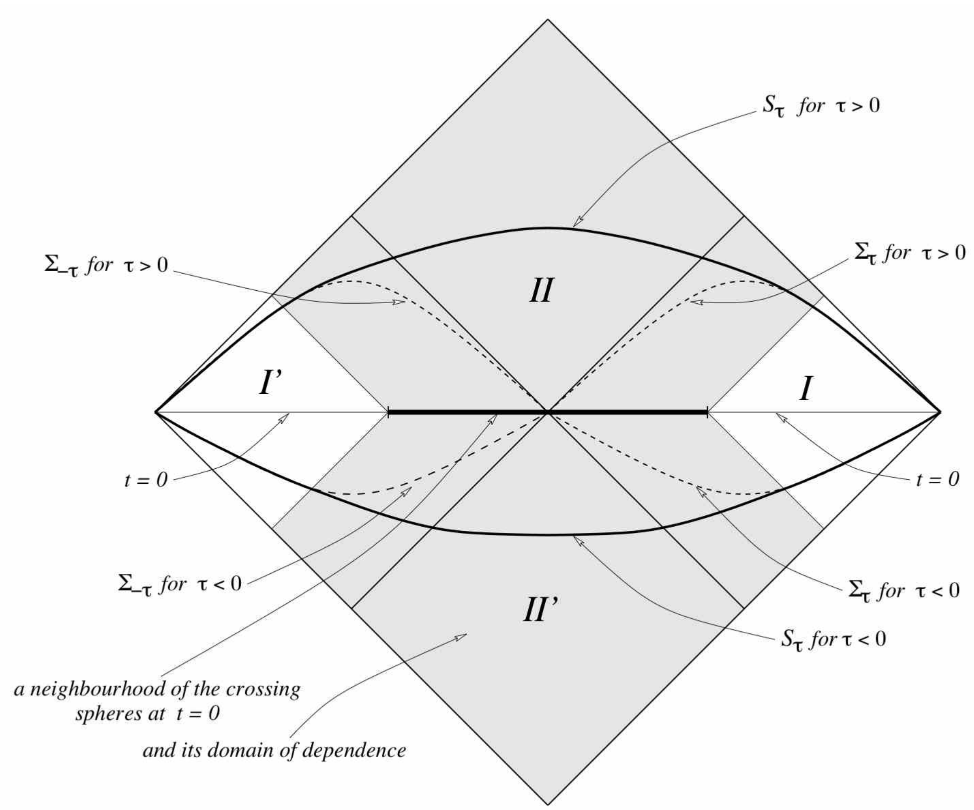

Figure 7.5: Foliation of a type I-II Kruskal domain

Chapter 8

Conformal compactifications

The notion of conformal compactification was introduced by Roger Penrose in the 1960's (voir Penrose 1963 [21], 1964 [22] and 1965 [23]). This chapter reviews classic material that can be found in details in Penrose & Rindler Vol. 2 [24].

Consider a spacetime $(\mathcal{M}, g)$, which may possess a boundary, but more importantly has some spacelike asymptotic end, i.e. some spacelike geodesics have infinite length. The idea of conformal compactification is to multiply the spacetime metric g by a positive function Ω^2 ($\Omega > 0$ being called the conformal factor), that tends to zero at infinity, in such a way that for the new metric $\hat{g} := \Omega^2 g$, the boundary of the spacetime $(\mathcal{M}, \hat{g})$ is larger than that of $(\mathcal{M}, g)$. We denote by $\left(\overline{\mathcal{M}}, \hat{g}\right)$ the closure of $(\mathcal{M}, \hat{g})$. The spacetime $\left(\overline{\mathcal{M}}, \hat{g}\right)$ is called a conformal compactification of $(\mathcal{M}, g)$. It is not required to be compact, in fact, the constructions in which the conformally compactified spacetime is compact are quite exceptional and usually do not correspond to physically relevent spacetimes.

8.1 Conformal rescalings, conformal invariance

Definition 8.1 (Conformal class)**.** *Consider a spacetime $(\mathcal{M}, g)$. We say that a metric $\hat{g}$ on $\mathcal{M}$ is conformally equivalent to g if there exists a positive nowhere vanishing smooth function Ω on $\mathcal{M}$ such that $\hat{g} = \Omega^2 g$. The conformal class $[g]$ of g is the set of all metrics on $\mathcal{M}$ that are conformally equivalent to g.*

Definition 8.2 (Conformal Killing vector)**.** *A conformal Killing vector field on a spacetime $(\mathcal{M}, g)$ is a vector field K^a on $\mathcal{M}$ such that there exists a metric $\hat{g} \in [g]$ for which K^a is a Killing vector.*

The conformal Killing vectors satisfy an equation similar to Killing vectors.

Proposition 8.1. *A vector field K^a on a spacetime $(\mathcal{M}, g)$ is a conformal Killing vector field if and only if its Killing form $\nabla_{(a}K_{b)}$ is proportional to the metric g_{ab}.*

Proof. (TO BE CONTINUED...) (should be established in the definition of Killing fields)

Provided we always work with the Levi-Civita connection associated with each metric, we can calculate the transformation of connection coefficients and of the curvature tensor under a conformal rescaling. A particularly useful formula is for the scalar curvature. For the sake of clarity we shall denote Scal_g the scalar curvature associated to a metric g and its Levi-Civita connection.

First, we calculate the change in the connection coefficients. It is expressed by the following formula : given a 1-form ω_a :

$$\hat{\nabla}_a \omega_b = \nabla_a \omega_b - C^c_{ab}\omega_c \,,$$

where $\hat{\nabla}$ and ∇ are respectively the Levi-Civita connections associated with the metrics $\hat{g}$ and g and $C^a_{bc} = C^a_{(bc)}$ denotes the change in the connection. The expression of the coefficients $\{C^c_{ab}\}$ can be obtained using the same method as for calculating the Christoffel symbols in a coordinate basis. We have $\hat{\nabla}_c \hat{g}_{ab} = 0$ and also

$$\hat{\nabla}_c \hat{g}_{ab} = \nabla_c \hat{g}_{ab} - C^d_{ca}\hat{g}_{db} - C^d_{cb}\hat{g}_{ad} \,.$$

This gives

$$C^d_{ca}\hat{g}_{db} + C^d_{cb}\hat{g}_{ad} = \nabla_c \hat{g}_{ab}$$

and performing a permutation of indices

$$C^d_{ab}\hat{g}_{dc} + C^d_{ac}\hat{g}_{bd} = \nabla_a \hat{g}_{bc} \,,$$
$$C^d_{bc}\hat{g}_{da} + C^d_{ba}\hat{g}_{cd} = \nabla_b \hat{g}_{ca} \,.$$

Using the symmetry of C and $\hat{g}$, taking the sum of the second and third equations minus the first, we get

$$2C^d_{ab}\hat{g}_{dc} = \nabla_a \hat{g}_{bc} + \nabla_b \hat{g}_{ca} - \nabla_c \hat{g}_{ab} \,,$$

i.e.

$$C^d_{ab} = \frac{1}{2}\hat{g}^{dc}\left(\nabla_a \hat{g}_{bc} + \nabla_b \hat{g}_{ca} - \nabla_c \hat{g}_{ab}\right) . \tag{8.1}$$

Now, using the relation between $\hat{g}$ and g and the fact that $\nabla g = 0$, we can make (8.1) more explicit :

$$\begin{aligned} C^d_{ab} &= \frac{1}{2}\hat{g}^{dc}\left(\nabla_a \hat{g}_{bc} + \nabla_b \hat{g}_{ca} - \nabla_c \hat{g}_{ab}\right) \\ &= \frac{1}{2}\hat{g}^{dc}\left(g_{bc}\nabla_a \Omega^2 + g_{ca}\nabla_b \Omega^2 - g_{ab}\nabla_c \Omega^2\right) \\ &= \frac{1}{2\Omega^2}g^{dc}2\Omega\left(g_{bc}\nabla_a \Omega + g_{ca}\nabla_b \Omega - g_{ab}\nabla_c \Omega\right) \\ &= 2g^d_{(b}\nabla_{a)}\log\Omega - g_{ab}\nabla^d \log\Omega \,. \end{aligned} \tag{8.2}$$

From this, the relation between the curvature tensors can be established as we established the expression of the Riemann tensor in terms of Christoffel symbols is chapter 2. We are merely interested in the transformation law for the scalar curvature here. To express it, we need to introduce the d'Alembertian on $(\mathcal{M}, g)$. It is a covariant differential operator on $(\mathcal{M}, g)$ denoted $\Box_g$ and defined by

$$\Box_g = \nabla_a \nabla^a \tag{8.3}$$

where ∇ is the Levi-Civita connection associated with g. When it acts on scalar fields, its expression in a local coordinate system $\{x^{\mathbf{a}}\}_{\mathbf{a}=0,1,2,3}$, is given by

$$\Box_g = \frac{1}{\sqrt{|\det g|}}\frac{\partial}{\partial x^{\mathbf{a}}}\sqrt{|\det g|}\, g^{\mathbf{ab}}\frac{\partial}{\partial x^{\mathbf{b}}} . \tag{8.4}$$

The transformation of the scalar curvature under a conformal rescaling can be expressed using the d'Alembertian of the initial or the resulting spacetime. For example, we have the following result :

Theorem 8.1. *Consider a spacetime $(\mathcal{M}, g)$ and a metric $\hat{g}$ in the conformal class of g with conformal factor Ω, i.e. $\hat{g} = \Omega^2 g$, then*

$$\mathrm{Scal}_{\hat{g}} = 6\Omega^{-3}\Box_g\Omega + \Omega^{-2}\mathrm{Scal}_g .$$

See [25] appendic D for a detailed proof.

8.2 Compactification of flat spacetime

8.2.1 The full compactification

The Minkowski metric in spherical coordinates is expressed as

$$\eta = \mathrm{d}t^2 - \mathrm{d}r^2 - r^2\mathrm{d}\omega^2 ,\ \mathrm{d}\omega^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\varphi^2 .$$

We choose the advanced and retarded coordinates

$$u = t - r ,\ v = t + r . \tag{8.5}$$

The metric (3.1) in terms of these new coordinates takes the form

$$\eta = \mathrm{d}u\mathrm{d}v - \frac{(v-u)^2}{4}\mathrm{d}\omega^2 .$$

We now introduce new null coordinates that allow us to describe the whole of Minkowski space as a bounded domain :

$$p = \arctan u ,\ q = \arctan v . \tag{8.6}$$

We obtain

$$\eta = (1+u^2)(1+v^2)\mathrm{d}p\,\mathrm{d}q - \frac{(v-u)^2}{4}\mathrm{d}\omega^2 .$$

Finally coming back to time and space coordinates as follows,

$$\begin{aligned} \tau &= p + q = \arctan(t-r) + \arctan(t+r) , \\ \zeta &= q - p = \arctan(t+r) - \arctan(t-r) , \end{aligned} \tag{8.7}$$

we get

$$\eta = \frac{(1+u^2)(1+v^2)}{4}\left(\mathrm{d}\tau^2 - \mathrm{d}\zeta^2\right) - \frac{(v-u)^2}{4}\mathrm{d}\omega^2 .$$

Choosing the conformal factor

$$\Omega^2 = \frac{4}{(1+u^2)(1+v^2)} = \frac{4}{(1+\tan^2 p)(1+\tan^2 q)} = (2\cos p \cos q)^2 \,, \tag{8.8}$$

we obtain the rescaled metric

$$\begin{aligned} \mathfrak{e} &:= \Omega^2 \eta = \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - \frac{(v-u)^2}{(1+u^2)(1+v^2)} \mathrm{d}\omega^2 \\ &= \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - ((\tan q - \tan p)\cos p \cos q)^2 \, \mathrm{d}\omega^2 \\ &= \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - (\sin q \cos p - \sin p \cos q)^2 \, \mathrm{d}\omega^2 \\ &= \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - (\sin(q-p))^2 \, \mathrm{d}\omega^2 \\ &= \mathrm{d}\tau^2 - \mathrm{d}\zeta^2 - (\sin\zeta)^2 \, \mathrm{d}\omega^2 \\ &= \mathrm{d}\tau^2 - \sigma^2_{S^3} \,, \end{aligned}$$

where $\sigma^2_{S^3}$ is the euclidian metric on the 3-sphere. Minkowski space is now described as the diamond

$$\mathbb{M} = \{ |\tau| + \zeta \leq \pi \,,\ \zeta \geq 0 \,,\ \omega \in S^2 \} \,.$$

The metric $\mathfrak{e}$ is the Einstein metric, it extends analytically to the whole Einstein cylinder $\mathfrak{E} = \mathbb{R}_\tau \times S^3_{\zeta,\theta,\varphi}$. The full conformal boundary of Minkowski space can be defined in this framework. It is described as

$$\partial\mathbb{M} = \{ |\tau| + \zeta = \pi \,,\ \zeta \geq 0 \,, \omega \in S^2 \} \,.$$

Several parts can be distinguished.

- Future and past null infinities :

$$\begin{aligned} \mathscr{I}^+ &= \left\{ (\tau \,,\ \zeta \,,\ \omega) \,;\ \tau + \zeta = \pi \,,\ \zeta \in]0,\pi[\,,\ \omega \in S^2 \right\} \,, \\ \mathscr{I}^- &= \left\{ (\tau \,,\ \zeta \,,\ \omega) \,;\ \zeta - \tau = \pi \,,\ \zeta \in]0,\pi[\,,\ \omega \in S^2 \right\} \,. \end{aligned}$$

Proposition 8.2. *The hypersurfaces $\mathscr{I}^\pm$ are smooth null hypersurfaces for $\mathfrak{e}$ (hence the terminology "null infinities"). Their null generators are respectively the vector fields*

$$\partial_\tau - \partial_\zeta \text{ for } \mathscr{I}^+ \text{ and } \partial_\tau + \partial_\zeta \text{ for } \mathscr{I}^- \,.$$

Proof. They are clearly smooth hypersurfaces since $\mathfrak{e}$ is analytic up to $\mathscr{I}^\pm$ and does not degenerate there : its determinent

$$\det(\mathfrak{e}) = -\sin^4\zeta \sin^2\theta$$

does not vanish on $\mathscr{I}^\pm$ (except for the usual coordinate singularity unavoidable when working with spherical coordinates). Now the vector fields $\partial_\tau - \partial_\zeta$ and $\partial_\tau + \partial_\zeta$ are null and tangent respectively to $\mathscr{I}^+$ and $\mathscr{I}^-$. They are orthogonal to the two other generators of $\mathscr{I}^\pm$: ∂_θ and ∂_φ. They are therefore normal to $\mathscr{I}^+$ and $\mathscr{I}^-$ respectively. This proves the proposition. □

- Future and past timelike infinities :

$$i^{\pm} = \left\{ (\tau = \pm\pi\,,\ \zeta = 0\,,\ \omega)\,;\ \omega \in S^2 \right\}\,.$$

They are smooth points for $\mathfrak{e}$ (2-spheres whose area is zero because they correspond to $\zeta = 0$).

- Spacelike infinity :

$$i^0 = \left\{ (\tau = 0\,,\ \zeta = \pi\,,\ \omega)\,;\ \omega \in S^2 \right\}\,.$$

It is also a smooth point for $\mathfrak{e}$.

The scalar curvature of $\mathfrak{e}$ can be calculated easily using theorem 8.1 for the conformal factor Ω and the fact that the curvature tensor vanishes for Minkowski space :

$$\frac{1}{6}\mathrm{Scal}_{\mathfrak{e}} = \Omega^{-3}\Box_\eta \Omega = 1\,. \tag{8.9}$$

8.2.2 A partial compactification

There is another way we can compactify Minkowski space ; this is very convenient when one is only interested in constructing null infinity and not the rest of the conformal boudary. It is a construction that in fact merely allows to construct $\mathscr{I}^+$ or $\mathscr{I}^-$ respectively but not both at the same time. It goes as follows.

We consider the retarded time variable $u = t - r$ and we perform an inversion on the radial variable $R = 1/r$. We express the Minkowski metric in terms of the coordinates u, R, θ, φ. We have

$$\begin{aligned} \mathrm{d}t^2 - \mathrm{d}r^2 &= (\mathrm{d}u + \mathrm{d}r)^2 - \mathrm{d}r^2 \\ &= \mathrm{d}u^2 + 2\mathrm{d}u\mathrm{d}r \\ &= \mathrm{d}u^2 - \frac{2}{R^2}\mathrm{d}u\mathrm{d}R\,, \end{aligned}$$

whence

$$\eta = \mathrm{d}u^2 - \frac{2}{R^2}\mathrm{d}u\mathrm{d}R - \frac{1}{R^2}\mathrm{d}\omega^2\,. \tag{8.10}$$

If now we multiply η by R^2, we obtain

$$\hat{\eta} = R^2\mathrm{d}u^2 - 2\mathrm{d}u\mathrm{d}R - \mathrm{d}\omega^2\,; \tag{8.11}$$

which extends as an analytic metric on the domain $\mathbb{R}_u \times [0, +\infty[_R \times S^2_{\theta,\varphi}$. Hence we can add to Minkowski spacetime the boundary $\mathbb{R}_u \times \{R = 0\} \times S^2_{\theta,\varphi}$. A point on this boundary, $(u = u_0, R = 0, \theta = \theta_0, \varphi = \varphi_0)$ is reached along a radial null geodesic

$$\gamma_{u_0,\theta_0,\varphi_0}(r) = (t = r + u_0, r, \theta = \theta_0, \varphi = \varphi_0)$$

as $r \to +\infty$ and there is a one-to-one correspondence between the outgoing radial null geodesics and the points on the boundary. The boundary therefore represents future null infinity $\mathscr{I}^+$.

A similar construction using an advance time variable $v = t + r$ instead of u allows to construct $\mathscr{I}^-$ instead of $\mathscr{I}^+$.

The scalar curvature of the rescaled metric $\hat{\eta}$ is zero.

8.2.3 Conformal Killing vectors

In the full compactification of Minkowski spacetime, we obtain a conformal Killing vector that is a Killing vector for $\mathfrak{e}$: the time translation along the Einstein cylinger ∂_τ. In the partial compactification there is another straightforward conformal Killing vector : ∂_u. However ∂_u is exactly ∂_t in the Schwarzschild coordinates and is therefore not a new vector field.

There is another Killing vector field for $\hat{\eta}$ that is not obvious in the expression of the metric. It is usually referred to as the Morawetz vector field and is obtained from ∂_t by a light-cone inversion or Kelvin transform. Its expression as it was found by Cathleen Morawetz in 1962 [18] is in spherical coordinates

$$(r^2+t^2)\frac{\partial}{\partial t}+2tr\frac{\partial}{\partial r}\,.$$

Its simplest expression is in terms of the coordinate system involving the advanced and retarded time variables $u=t-r$, $v=t+r$ and θ and φ :

$$K^a\partial_a=u^2\frac{\partial}{\partial u}+v^2\frac{\partial}{\partial v}\,.$$

In the coordinate system u,R,θ,φ, it has the expression

$$K^a\partial_a=u^2\frac{\partial}{\partial u}-2(1+uR)\frac{\partial}{\partial R}\,. \tag{8.12}$$

Theorem 8.2. *The Morawetz vector field (8.12) is a Killing vector for $\hat{\eta}$.*

The proof is left as an exercice. A more general result will be proved in the framework of the Schwarzschild metric.

The Morawetz vector field and the translation along the Einstein cylinder are in fact very close to one another. If we express ∂_τ in the coordinate system u,R,θ,φ, we obtain

$$\partial_\tau=\frac{1}{2}\left[\partial_u+u^2\partial_u-2(1+uR)\partial_R\right]\,.$$

We see that

$$2\partial_\tau=\partial_u+K^a\partial_a\,,$$

or in terms of variables t,r,θ,φ,

$$2\partial_\tau=\partial_t+K^a\partial_a\,.$$

8.3 Compactification of Schwarzschild's spacetime

Schwarzschild's spacetime contains mass. This is apparent in the asymptotic behaviour of the metric : some terms are proportional to the mass M of the black hole and fall off in $1/r$ at infinity. These terms prevent the construction of a complete regular compactification similar to what can be done with Minkowski spacetime. A partial compactification however is possible and just as easy as in the flat case. Instead of the variable $t-r$ (resp. $t+r$), it is based on the variable $t - r_*$ (resp. $t + r_*$) that is a parameter of the family of outgoing radial null geodesics.

In terms of variables $u = t - r_*$, $R = 1/r$, θ and φ, referred to as the Eddington-Finkelstein coordinates, the Schwarzschild metric g takes the form

$$g = (1 - 2MR)\mathrm{d}u^2 - \frac{2}{R^2}\mathrm{d}u\mathrm{d}R - \frac{1}{R^2}\mathrm{d}\omega^2 .$$

Rescaling the metric with the conformal factor $\Omega = R = 1/r$, we obtain

$$\hat{g} = R^2 g = R^2(1 - 2MR)\mathrm{d}u^2 - 2\mathrm{d}u\mathrm{d}R - \mathrm{d}\omega^2 ,$$

which extends as an analytic metric on the domain $\mathbb{R}_u \times \left[0, \frac{1}{2M}\right[_R \times S^2_{\theta,\varphi}$. Similarly to the Minkowski case, we can add a boundary to the exterior of the black hole : the hypersurface $\mathbb{R}_u \times \{0\}_R \times S^2_{\theta,\varphi}$. A point $(u_0, 0, \theta_0, \varphi_0)$ on the boundary is reached along the outgoing radial null geodesic

$$\begin{aligned}\gamma_{u_0,\theta_0,\varphi_0}(r) &= (t = r + 2M\mathrm{Log}(r - 2M) + u_0 \, , \; r \, , \; \theta = \theta_0 \, , \; \varphi = \varphi_0) \\ &= \left(u = u_0 \, , \; R = \frac{1}{r} \, , \; \theta = \theta_0 \, , \; \varphi = \varphi_0\right)\end{aligned}$$

as $r \to +\infty$ and there is a one to one correspondence between the points on the boundary and the outgoing radial null geodesics. The hypersurface therefore represents future null infinity, $\mathscr{I}^+$, for the Schwarzschild metric.

Using theorem 8.1 and the fact that the scalar curvature of the Schwarzschild metric is zero, we can calculate the scalar curvature of the rescaled metric $\hat{g} = R^2 g$ and we find

$$\frac{1}{6}\mathrm{Scal}_{\hat{g}} = 2MR .$$

8.4 Conformal compactification of Kerr's spacetime

This is analogous to the compactification of Schwarzschild's spacetime. It is done for block I only, using star-Kerr and Kerr-star coordinates and an inversion of the variable r. TO BE CONTINUED...

8.5 Asymptotically simple spacetimes

8.6 exercices

Exercice 8.1. *Prove theorem 8.2.*

Exercice 8.2. *Prove that the action on scalar fields of the d'Alembertian (8.3) is given by (8.4).*

Chapter 9

Conformal invariance and asymptotic behaviour

The zero rest-mass field equations satisfy a property referred to as conformal invariance. In essence, it means that when the metric is rescaled via a conformal transformation, the equation transforms to the corresponding covariant equation for the rescaled metric, provided the field is itself rescaled by a power of the conformal factor Ω. The appropriate power of Ω for a given field is called its conformal weight. Zero rest-mass fields have conformal weight -1. Formally, if we denote a zero rest mass field equation by

$$\mathcal{L}_g(\phi) = 0\,, \tag{9.1}$$

then ϕ satisfies (9.1) if and only if

$$\mathcal{L}_{\Omega^2 g}\left(\Omega^{-1}\phi\right) = 0\,.$$

9.1 The scalar wave equation

The scalar wave equation (or simply wave equation) is a covariant equation that models the evolution of scalar fields on a spacetime. On a given spacetime $(\mathcal{M}, g)$, it has the form

$$\Box_g \phi = 0\,, \tag{9.2}$$

where ϕ is a scalar function on $\mathcal{M}$ and $\Box_g$ is the d'Alembertian on $(\mathcal{M}, g)$ defined in a local coordinate basis by (8.4).

Definition 9.1 (Conformal invariance)**.** *The conformal invariance of a covariant equation means that there exists $s \in \mathbb{R}$ such that a field ϕ satisfies the equation for the metric g if and only if $\Omega^s\phi$ satisfies the equation for $\hat{g}$.*

The wave equation (9.2) is not conformally invariant. However, a slight modification of this equation involving the scalar curvature is conformally invariant. We shall refer to it as the conformal wave equation :

$$\Box_g \phi + \frac{1}{6}\mathrm{Scal}_g \phi = 0\,, \tag{9.3}$$

where Scal_g denotes the scalar curvature associated to the metric g and its Levi-Civita connection. More precisely, we have the following fundamental result that is a straightforward consequence of theorem 8.1 :

Theorem 9.1. *We consider a spacetime $(\mathcal{M}, g)$ and a metric $\hat{g}$ in the conformal class of g with conformal factor Ω, i.e. $\hat{g} = \Omega^2 g$. The we have the equality of operators acting on scalar fields on $\mathcal{M}$:*

$$\Box_g + \frac{1}{6}\mathrm{Scal}_g = \Omega^3 \left(\Box_{\hat{g}} + \frac{1}{6}\mathrm{Scal}_{\hat{g}} \right) \Omega^{-1} .$$

Proof. For a scalar field ϕ : $\hat{\nabla}_a \phi = \nabla_a \phi$ and given $s \in \mathbb{R}$,

$$\begin{aligned} \Box_{\hat{g}}(\Omega^s \phi) = \hat{g}^{ab} \hat{\nabla}_a \hat{\nabla}_b (\Omega^s \phi) &= \Omega^2 g^{ab} \left(\nabla_a \nabla_b (\Omega^s \phi) - C^c_{ab} \nabla_c (\Omega^s \phi) \right) \\ &= \Omega^{s-2} \Box_g \phi + (2s + n - 2) \Omega^{s-3} g^{ab} \nabla_a \Omega \nabla_b \phi \\ &\quad + (s \Omega^{s-3} \Box_g \Omega) \phi + (s(n+s-3)\Omega^{s-4} g^{ab} \nabla_a \Omega \nabla_b \Omega) \phi \end{aligned}$$

Putting $s = (2-n)/2$ gets rid of the second term in the right-hand side. The remainder is a multiple of the scalar curvature. TO BE CONTINUED... (needs a more detailed proof) $\Box$

This has the immediate consequence :

Corollary 9.1. *Let $\phi \in \mathcal{D}'(\mathcal{M})$, the following conditions are equivalent :*

1. *ϕ satisfies (9.3) in the sense of distributions on $\mathcal{M}$;*

2. *$\hat{\phi} := \Omega^{-1} \phi$ satisfies*

$$\Box_{\hat{g}} \hat{\phi} + \frac{1}{6}\mathrm{Scal}_{\hat{g}} \hat{\phi} = 0$$

in the sense of distributions on $\mathcal{M}$.

9.2 Pointwise decay

Conformal invariance combined with a conformal compactification allows to infer asymptotic properties of solutions to the wave equation on the "physical" spacetime from local properties of the rescaled solution at the boundary of the compactified spacetime. Such constructions however will typically be valid for certain classes of data for the rescaled equation ; these data should satisfy regularity assumptions at the boundary that can be translated as asymptotic properties at infinity for the corresponding physical data. The most explicit example of this type of construction is in the framework of Minkowski space. The situation in Schwarzschild's spacetime is more delicate as we shall see in chapter 10.

9.2.1 Pointwise decay in flat spacetime

Consider the conformal wave equation on the Einstein cylinder $\mathfrak{E} = \mathbb{R} \times S^3$. By classic results on hyperbolic equations due to Leray [15],

Proposition 9.1. *For any initial data*

$$\hat{\phi}_0\,,\ \hat{\phi}_1 \in \mathcal{C}^\infty(S^3)\,,$$

the Cauchy problem

$$\Box_{\mathfrak{e}}\hat{\phi} + \frac{1}{6}\mathrm{Scal}_{\mathfrak{e}}\hat{\phi} = 0\,,\ \hat{\phi}_{|\tau=0} = \hat{\phi}_0\,,\ \partial_\tau\hat{\phi}_{|\tau=0} = \hat{\phi}_1$$

has a unique solution in $\mathcal{C}^1(\mathbb{R}_\tau\,;\ \mathcal{D}'(S^3))$ and it is smooth on $\mathfrak{E}$.

The corresponding physical solution therefore admits the following pointwise decays :

Proposition 9.2. *There exist smooth functions $\hat{\phi}^\pm_\infty \in \mathcal{C}^\infty(\mathbb{R}\times S^2)$ such that*

$$\lim_{r\to+\infty} r\phi(t=r+u,r,\omega) = \frac{1}{\sqrt{1+u^2}}\,\hat{\phi}^+_\infty(u,\omega)\,,$$
$$\lim_{r\to+\infty} r\phi(t=-r+v,r,\omega) = \frac{1}{\sqrt{1+v^2}}\,\hat{\phi}^-_\infty(v,\omega)\,.$$

The functions $\hat{\phi}^\pm_\infty$ are simply the traces of $\hat{\phi}$ on $\mathscr{I}^\pm$; the two functions in the right hand side of the limits above are referred to as the future and past asymptotic profiles of ϕ.

There exists smooth functions $f^\pm \in \mathcal{C}^\infty(S^2)$ such that

$$\lim_{t\to+\infty} t^2\phi(t,r,\omega) = 2f^+(\omega)\,,$$
$$\lim_{t\to-\infty} t^2\phi(t,r,\omega) = 2f^-(\omega)\,.$$

These functions are the traces of $\hat{\phi}$ at i^+ and i^- respectively.

In other words, the physical solution ϕ decays like $1/r$ along radial null geodesics and like $1/t^2$ along the integral lines of ∂_t.

Proof. The proof is trivial ; it is based of the explicit expression of Ω and on the smoothness of $\hat{\phi}$ on $\mathfrak{E}$. □

It is important to note that the result above has been established for solutions ϕ of the wave equation on Minkowski spacetime such that $\hat{\phi} = \Omega^{-1}\phi$ extends as a smooth function on $\mathfrak{E}$. This entails some fall-off assumptions on ϕ :

Proposition 9.3. *The smoothness of $\hat{\phi}_0$ and $\hat{\phi}_1$ on S^3 entails that there exist two smooth functions $g_0, g_1 \in \mathcal{C}^\infty(S^2)$ such that*

$$\lim_{r\to+\infty} r^2\phi(0,r,\omega) = 2g_0(\omega)\,,$$
$$\lim_{r\to+\infty} r^4\partial_t\phi(0,r,\omega) = 4g_1(\omega)\,.$$

The functions g_0 and g_1 are the traces of $\hat{\phi}_0$ and $\hat{\phi}_1$ at i^0.

Proof. The first limit is a straightforward consequence of the regularity of $\hat{\phi}$ on $\mathfrak{E}$ and its relation to ϕ. As for the second limit, we have

$$\partial_t\phi = (\partial_t\Omega)\,\hat{\phi} + \Omega\frac{\partial\tau}{\partial t}\partial_\tau\hat{\phi} + \Omega\frac{\partial\zeta}{\partial t}\partial_\zeta\hat{\phi}$$

which for $\tau = 0$ (i.e. $t = 0$) gives

$$\partial_t\phi_{|t=0} = \frac{4}{(1+r^2)^2}\partial_\tau\hat{\phi}_{|\tau=0} = \frac{4}{(1+r^2)^2}\hat{\phi}_1\,.$$

This proves the proposition. □

9.2.2 Pointwise decay in Schwarzschild's spacetime

Since we only have a partial compactification for Schwarzschild's spacetime, using the same method based on a conformal compactification and straightforward observations, we cannot get a description of pointwise decay that is as complete as in the case of Minkowski space. First of all, the t =constant slices are not compactified, i.e. we have no natural notion af data that are regular up to the boundary of the $t = 0$ slice. Hence, in order to guarantee that the rescaled solution extends as a smooth function on the boundary of the compactified spacetime, we need to restrict ourselves to smooth compactly supported data. Otherwise, we would need to have a clear understanding of the fall-off assumptions on initial data that ensure smoothness at $\mathscr{I}$ of the rescaled solution ; this is a much more difficult question that is dealt with in details in chapter 10. We have the following result whose proof is straightforward.

Proposition 9.4. *Let ϕ_0, $\phi_1 \in \mathcal{C}_0^\infty(]2M,+\infty[_r \times S^2_\omega)$. Then the Cauchy problem*

$$\Box_g \phi = 0\,,\ \phi_{|_{t=0}} = \phi_0\,,\ \partial_t \phi_{|_{t=0}} = \phi_1$$

has a unique solution

$$\phi \in \mathcal{C}^1\left(\mathbb{R}_t\,;\ \mathcal{D}'(]2M,+\infty[_r \times S^2_\omega)\right)$$

and it is smooth on $\mathbb{R}_t \times]2M,+\infty[_r \times S^2_\omega$. Moreover the rescaled solution $\hat{\phi} = r\phi$ has a smooth trace on $\mathscr{I}^+$ and on $\mathscr{I}^-$, i.e. there exist two smooth functions $\hat{\phi}^\pm_\infty \in \mathcal{C}^\infty(\mathbb{R} \times S^2)$ such that

$$\lim_{t\to+\infty} r\phi(t, r_* = t-u, \omega) = \hat{\phi}^+_\infty(u,\omega)\,, \qquad (9.4)$$

$$\lim_{t\to+\infty} r\phi(t, r_* = -t+v, \omega) = \hat{\phi}^-_\infty(v,\omega)\,. \qquad (9.5)$$

In particular, this shows that ϕ falls off like $1/r$ along radial null geodesics.

The same result is valid for the Kerr geometry along principal null geodesics.

Chapter 10

Peeling

We have in the previous chapter used conformal compactifications to infer the asymptotic behaviour of solutions to the wave equation on Minkowski or Schwarzschild spacetimes such that the rescaled solution extend as smooth functions on the compactified spacetime. These constructions however only used the continuity of the rescaled solution at the boundary (except for the fall-off of physical initial data where we actually used the fact that the rescaled solution was $\mathcal{C}^1$).

The peeling studies the precise regularity of the rescaled solution at $\mathscr{I}$ and its characterization in terms of the regularity and fall-off at infinity of the physical initial data. The peeling was initially proposed by Penrose in 1963 [21] as a model for the behaviour of solutions to conformally invariant equations on generic asymptotically flat spacetimes. It is based on the observation that in Minkowski spacetime, using the complete compactification, we can establish a precise relation between the regularity at $\mathscr{I}$ of the rescaled solutions and the regularity and asymptotic behaviour of the initial data for the associated physical solutions. Penrose's conjecture was stating that on a generic asymptotically flat spacetime (including Schwarzschild, Kerr, and physically realistic models of black holes in an asymptotically flat universe) exactly the same assumptions on the physical data as in Minkowski space should guarantee precisely the same regularity at $\mathscr{I}^+$ for the rescaled solution.

Peeling in flat spacetime has been well understood for linear equations since Penrose's work. Concerning non-linear equations and more particularly the Einstein equations (i.e. concerning non-linear perturbations of flat space-time), peeling properties have only been recently established by Klainerman and Nicoló [13] (and in fact only for the first orders of regularity). On the Schwarzschild metric, which is the simplest non trivial asymptotically flat solution to the Einstein equations, the conjecture has remained totally open until now and has been an extremely controversial issue.

10.1 Flat spacetime

Recall that the conformal invariance of the wave equation entails the equivalence of the two properties :

(i) $\tilde{\psi} \in \mathcal{D}'(\mathbb{M})$ satisfies the conformal wave equation on $(\mathbb{M}, \eta)$;

(ii) $\psi := \Omega^{-1}\tilde{\psi}$ satisfies the conformal wave equation on the Einstein cylinder $(\mathfrak{E}, \mathfrak{e})$.

The conformal wave equation on flat space-time is simply the wave equation, which in spherical coordinates has the expression

$$\Box_\eta \tilde{\psi} = 0\,,\ \Box_\eta = \partial_t^2 - \frac{1}{r^2}\partial_r r^2 \partial_r - \frac{1}{r^2}\Delta_{S^2}\,; \tag{10.1}$$

while on the Einstein cylinder, the conformal wave equation (using the fact that the scalar curvature on the Einstein cylinder is equal to 6) is

$$\Box_{\mathfrak{e}}\psi + \psi = 0\,,\ \Box_{\mathfrak{e}} = \partial_\tau^2 - \Delta_{S^3}\,; \tag{10.2}$$

where Ω is defined by (8.8), i.e.

$$\Omega^2 = \frac{4}{(1+(t+r)^2)(1+(t-r)^2)}\,.$$

10.1.1 The usual description of peeling

The observation of the peeling in Minkowski space is usually derived from the property that the Cauchy problem on the Einstein cylinder is well-posed in the space of $\mathcal{C}^\infty$ functions (see proposition 9.1). This provides a natural definition of solutions that peel at all orders :

Definition 10.1. *A solution $\tilde{\psi}$ of (10.1) is said to peel at all orders if $\psi = \Omega^{-1}\tilde{\psi}$ extends as a $\mathcal{C}^\infty$ function on the whole Einstein cylinder. The latter property is satisfied by solutions ψ of (10.2) arising from initial data $\psi|_{\tau=0} \in \mathcal{C}^\infty(S^3)$ and $\partial_\tau\psi|_{\tau=0} \in \mathcal{C}^\infty(S^3)$. Going back to Minkowski space and to the physical field $\tilde{\psi}$, this gives us a corresponding class of data for (10.1), giving rise to solutions that peel at all orders.*

10.1.2 Description by means of vector field methods

Although it is not commonly used, we can give a description of the peeling in Minkowski space in terms of Sobolev spaces. This has the advantage of allowing a precise description at all orders of regularity, which is trickier in $\mathcal{C}^k$ spaces that are less well controlled by the evolution. To do so, we write equation (10.2) in its hamiltonian form

$$\frac{\partial}{\partial\tau}\begin{pmatrix}\psi\\ \partial_\tau\psi\end{pmatrix} = iH\begin{pmatrix}\psi\\ \partial_\tau\psi\end{pmatrix},\ H = -i\begin{pmatrix}0 & 1\\ \Delta_{S^3}-1 & 0\end{pmatrix}$$

and work on the Hilbert space

$$\mathcal{H} = H^1(S^3)\times L^2(S^3)$$

with the usual inner product

$$\left\langle \begin{pmatrix}f_1\\ f_2\end{pmatrix}, \begin{pmatrix}g_1\\ g_2\end{pmatrix}\right\rangle_{\mathcal{H}} = \int_{S^3}(\nabla_{S^3}f_1.\nabla_{S^3}\bar{g}_1 + f_1\bar{g}_1 + f_2\bar{g}_2)\,\mathrm{d}\mu_{S^3}\,,$$

where ∇_{S^3} is the Levi-Civitta connection and μ_{S^3} the measure induced by the Euclidian metric on S^3. We have :

Proposition 10.1. *The operator H with its natural domain $D(H)=H^2(S^3)\times H^1(S^3)$ is self-adjoint on $\mathcal{H}$ and its successive domains are*

$$D(H^k)=H^{k+1}(S^3)\times H^k(S^3)\,.$$

Let $k\in\mathbb{N}$. For any initial data $\psi_0\in H^{k+1}(S^3)$, $\psi_1\in H^k(S^3)$, there exists a unique solution

$$\psi\in\bigcap_{l=0}^{k+1}\mathcal{C}^l\left(\mathbb{R}_\tau\,;\;H^{k+1-l}(S^3)\right)$$

of (10.2) such that $\psi(0)=\psi_0$ and $\partial_\tau\psi(0)=\psi_1$. In particular, $\psi\in H^{k+1}_{\mathrm{loc}}(\mathfrak{E})$. Moreover, for any $0\leq l\leq k$, $\|\psi(\tau)\|^2_{H^{l+1}(S^3)}+\|\partial_\tau\psi(\tau)\|^2_{H^l(S^3)}$ is constant throughout time.

This gives us a definition of solutions that peel at a given order $k\in\mathbb{N}$ in terms of Sobolev spaces :

Definition 10.2. *A solution $\tilde\psi$ of (10.1) is said to peel at order $k\in\mathbb{N}$ if $\psi=\Omega^{-1}\tilde\psi$ extends as a function that is in H^{k+1}_{loc} on the whole Einstein cylinder. The latter property is satisfied by solutions ψ of (10.2) arising from initial data $\psi|_{\tau=0}\in H^{k+1}(S^3)$ and $\partial_\tau\psi|_{\tau=0}\in H^k(S^3)$. Going back to Minkowski space and to the physical field $\tilde\psi$, this gives us a corresponding class of data for (10.1), giving rise to solutions that peel at order k.*

What is essentially unsatisfactory in definition 10.2 is that it merely provides an inclusion : we know a class of data that gives rise to peeling at order k, but we do not know whether it is the largest possible class. An alternative approach consists in using vector field methods (energy estimates). Such techniques allow to prove easily the last property in proposition 10.1 but are much more flexible than a purely spectral result : we can just as naturally obtain estimates between the initial data surface and $\mathscr{I}^+$. This will provide us with a third description of the peeling on flat space-time. It will be more precise than the first two in that the optimal set of suitable data for a peeling at order k will be completely characterized.

We consider the stress energy tensor for equation (10.2)

$$T_{ab}=T_{(ab)}=\partial_a\psi\partial_b\psi-\frac{1}{2}\mathfrak{e}_{ab}\mathfrak{e}^{cd}\partial_c\psi\partial_d\psi+\frac{1}{2}\psi^2\mathfrak{e}_{ab} \tag{10.3}$$

and contract it with the Killing vector field ∂_τ. This yields the conservation law

$$\nabla^a\left(K^bT_{ab}\right)=0\,. \tag{10.4}$$

The energy 3-form $K^aT_{ab}\mathrm{d}^3x^b=K^aT_a^b\partial_b\lrcorner\,\mathrm{dVol}^4$ has the expression

$$K^aT_{ab}\mathrm{d}^3x^b=\psi_\tau\nabla\psi\lrcorner\,\mathrm{dVol}^4+\frac{1}{2}\left(-\psi_\tau^2+|\nabla_{S^3}\psi|^2+\psi^2\right)\partial_\tau\lrcorner\,\mathrm{dVol}^4\,. \tag{10.5}$$

Integrating (10.5) on an oriented hypersurface S defines the energy flux across this surface, denoted $\mathcal{E}_S(\psi)$. For instance, denoting $X_\tau=\{\tau\}\times S^3$ the level hypersurfaces of the function τ

$$\mathcal{E}_{X_\tau}(\psi)=\frac{1}{2}\int_{X_\tau}\left(\psi_\tau^2+|\nabla_{S^3}\psi|^2+\psi^2\right)\mathrm{d}\mu_{S^3}\,,$$

and parametrizing $\mathscr{I}^+$ as $\tau = \pi - \zeta$,

$$\begin{aligned}\mathcal{E}_{\mathscr{I}^+}(\psi) &= \frac{1}{\sqrt{2}}\int_{\mathscr{I}^+}\left(-2\psi_\tau\psi_\zeta + \psi_\tau^2 + |\nabla_{S^3}\psi|^2 + \psi^2\right)\mathrm{d}\mu_{S^3}\\ &= \frac{1}{\sqrt{2}}\int_{\mathscr{I}^+}\left(|\psi_\tau - \psi_\zeta|^2 + \frac{1}{\sin^2\zeta}|\nabla_{S^2}\psi|^2 + \psi^2\right)\mathrm{d}\mu_{S^3}\,.\end{aligned}$$

This is a natural H^1 norm of ψ on $\mathscr{I}^+$, involving only the tangential derivatives of ψ along $\mathscr{I}^+$.

Now consider a smooth solution ψ of (10.2). The conservation law (10.4) tells us that (10.5) is closed, hence, integrating it on the closed hypersurface made of the union of X_0 and $\mathscr{I}^+$, we obtain

$$\mathcal{E}_{\mathscr{I}^+}(\psi) = \mathcal{E}_{X_0}(\psi)$$

and since ∂_τ is a Killing vector, for any $k \in \mathbb{N}$, $\partial_\tau^k\psi$ satisfies equation (10.2), whence

$$\mathcal{E}_{\mathscr{I}^+}(\partial_\tau^k\psi) = \mathcal{E}_{X_0}(\partial_\tau^k\psi)\,.$$

Using equation (10.2), for $k = 2p$, $p \in \mathbb{N}$, we have

$$\begin{aligned}\mathcal{E}_{X_0}(\partial_\tau^k\psi) &= \|\partial_\tau^{2p}\psi\|^2_{H^1(X_0)} + \|\partial_\tau^{2p+1}\psi\|^2_{L^2(X_0)}\\ &= \|(1-\Delta_{S^3})^p\psi\|^2_{H^1(X_0)} + \|(1-\Delta_{S^3})^p\partial_\tau\psi\|^2_{L^2(X_0)}\\ &\simeq \|\psi\|^2_{H^{2p+1}(X_0)} + \|\partial_\tau\psi\|^2_{H^{2p}(X_0)}\,,\end{aligned} \tag{10.6}$$

and for $k = 2p+1$, $p \in \mathbb{N}$,

$$\begin{aligned}\mathcal{E}_{X_0}(\partial_\tau^k\psi) &= \|\partial_\tau^{2p+1}\psi\|^2_{H^1(X_0)} + \|\partial_\tau^{2p+2}\psi\|^2_{L^2(X_0)}\\ &= \|(1-\Delta_{S^3})^p\partial_\tau\psi\|^2_{H^1(X_0)} + \|(1-\Delta_{S^3})^{p+1}\psi\|^2_{L^2(X_0)}\\ &\simeq \|\psi\|^2_{H^{2p+2}(X_0)} + \|\partial_\tau\psi\|^2_{H^{2p+1}(X_0)}\,.\end{aligned} \tag{10.7}$$

Hence, we have for each $k \in \mathbb{N}$:

$$\|\psi\|^2_{H^{k+1}(X_0)} + \|\partial_\tau\psi\|^2_{H^k(X_0)} \simeq \mathcal{E}_{X_0}(\partial_\tau^k\psi) = \mathcal{E}_{\mathscr{I}^+}(\partial_\tau^k\psi) \simeq \|\partial_\tau^k\psi\|^2_{H^1(\mathscr{I}^+)}$$

and using the fact that the H^k norm controls all the lower Sobolev norms, this gives us the apparently stronger equivalence

$$\|\psi\|^2_{H^{k+1}(X_0)} + \|\partial_\tau\psi\|^2_{H^k(X_0)} \simeq \sum_{p=0}^{k}\|\partial_\tau^p\psi\|^2_{H^1(\mathscr{I}^+)}\,. \tag{10.8}$$

Remark 10.1. *This equivalence should not in principle be understood as providing a solution to a Goursat problem on $\mathscr{I}^+$. Indeed, in Lars Hörmander's paper on the Goursat problem for the wave equation [12], it is made very clear that such an equivalence only provides us with a trace operator on $\mathscr{I}^+$ that is a partial isometry, it is then necessary to prove the surjectivity of this operator in order to solve the Goursat problem. However, we know from the same paper that the Goursat problem for equation (10.2) with data $\psi_{|_{\mathscr{I}^+}} \in H^1(\mathscr{I}^+)$ is well posed and gives rise to solutions $\psi \in \mathcal{C}^0(\mathbb{R}_\tau\,;\; H^1(S^3)) \cap \mathcal{C}^1(\mathbb{R}_\tau\,;\; L^2(S^3))$. Hence* *equivalence (10.8) indeed provides us with a regularity result for the Goursat problem : data on $\mathscr{I}^+$ for which the norm on the right-hand side is finite give rise to solutions that are in $\mathcal{C}^l(\mathbb{R}_\tau ;\ H^{k+1-l}(S^3))$ for all $0 \leq l \leq k+1$.*

This is however stronger than the information we are interested in. Equivalence (10.8) simply says that for smooth solutions, the control of the transverse regularity on $\mathscr{I}^+$ described by $\mathcal{E}_{\mathscr{I}^+}(\partial_\tau^p \psi)$, $0 \leq p \leq k$, is equivalent to that of the H^{k+1} norm of the restriction of ψ to X_0 and the H^k norm of the restriction of $\partial_\tau \psi$ to X_0. By a standard density argument, this shows that if we wish to guarantee, by means of some control on the initial data, that the restriction to $\mathscr{I}^+$ of $\partial_\tau^p \psi$, $0 \leq p \leq k$, is in $H^1(\mathscr{I}^+)$, the optimal condition to impose is that $\psi|_{\tau=0} \in H^{k+1}(X_0)$ and $\partial_\tau \psi|_{\tau=0} \in H^k(X_0)$.

This provides us with our third definition of peeling at order k and a characterization by a function space of initial data.

Definition 10.3. *A solution $\tilde{\psi}$ of (10.1) is said to peel at order $k \in \mathbb{N}$ if the traces on $\mathscr{I}^+$ of $\partial_\tau^p \psi$, $0 \leq p \leq k$, are in $H^1(\mathscr{I}^+)$. The optimal function space of initial data giving rise to such property is defined by $\psi|_{\tau=0} \in H^{k+1}(S^3)$ and $\partial_\tau \psi|_{\tau=0} \in H^k(S^3)$.*

Going back to Minkowski space and to the physical field $\tilde{\psi}$, this gives us the exact function space of data for (10.1), giving rise to solutions that peel at order k.

Remark 10.2. *The description given in definition 10.2 corresponds to the slightly weaker approach, via the equality*

$$\mathcal{E}_{X_\tau}(\partial_\tau^k \psi) = \mathcal{E}_{X_0}(\partial_\tau^k \psi) \ \forall\ \tau \in \mathbb{R}\,,$$

which entails

$$\|\psi\|_{H^{k+1}(\Omega^+)} \lesssim \|\psi\|^2_{H^{k+1}(X_0)} + \|\partial_\tau \psi\|^2_{H^k(X_0)}\,, \tag{10.9}$$

where Ω^+ is the 4-volume in the future of X_0 and the past of $\mathscr{I}^+$. It is slightly weaker in the way we understand the transverse regularity at $\mathscr{I}^+$ (implicitly in terms of trace theorems for Sobolev spaces), hence the fact that we have merely inequalities instead of equivalences. But the spaces of initial data for which regularity at a given order is guaranteed near $\mathscr{I}^+$ are the same.

10.2 Peeling in the Schwarzschild case

The equation that we study is the conformally invariant wave equation

$$\left(\Box_{\hat{g}} + \frac{1}{6}\mathrm{Scal}_{\hat{g}}\right)\phi = \left(\Box_{\hat{g}} + 2mR\right)\phi = 0\,, \tag{10.10}$$
$$\Box_{\hat{g}} = -2\partial_u\partial_R - \partial_R R^2\left(1-2mR\right)\partial_R - \Delta_{S^2}\,.$$

A distribution ϕ on the exterior of Schwarzschild's black hole satisfies (10.10) if and only if $\tilde{\phi} := R\phi$ satisfies

$$\Box_g \tilde{\phi} = 0\,.$$

10.2.1 General strategy

We study peeling properties in the neighbourhood of i^0 where the real problem is localized. If one wishes, results can be extended to the whole of $\mathscr{I}^+$ and the whole initial data surface in a straightforward manner. Our work deals with what happens near i^0.

We work with vector field methods, that is to say essentially energy estimates. We start by obtaining some basic energy estimates between $\mathscr{I}^+$ and the initial data surface $\{t=0\}$. Then, by applying some well chosen vector fields (differential operators) to the equation, we get some higher order estimates.

For the basic energy estimates, we need to find some vector field that is close to being a Killing vector field for $\hat{g}$ and that is transverse to $\mathscr{I}$. We adapt the classic "Morawetz vector field" to Schwarzschild's space-time just as Dafermos and Rodnianski did in [6].

10.2.2 The Morawetz vector field

For $m=0$, the metric g is the Minkowski metric η and we have $u=t-r$. The Morawetz vector field is defined as the image of ∂_t, that is a Killing vector for η, by a light-cone inversion. Its simplest expression is in terms of variables $u=t-r$ and $v=t+r$:

$$T^a\partial_a = u^2\partial_u + v^2\partial_v \,,$$

which, in terms of variables u, R gives

$$T^a\partial_a := u^2\partial_u - 2(1+uR)\partial_R \,. \tag{10.11}$$

This vector field is a Killing vector for $\hat{g}$ for $m=0$, i.e. if we perform the time asymmetrical compactification for Minkowski spacetime, we obtain a rescaled metric for which T is a Killing vector. It is interesting to note that the Killing vector ∂_τ corresponding to the time translation on the Einstein cylinder is a simple combination of T and ∂_u in this coordinate system :

$$2\partial_\tau = \partial_u + T^a\partial_a \,,$$

it is therefore also a Killing vector for $\hat{g}$ for $m=0$.

We keep the expression (10.11) in terms of variables u, R in the Schwarzschild case to define our approximate Killing vector field. We still refer to it as the Morawetz vector field and denote it T.

T^a is uniformly timelike in a neighbourhood of i^0 and can therefore be used for obtaining energy estimates with positive definite energies on spacelike hypersurfaces.

10.2.3 Stress energy tensor and energy density

We choose the stress-energy tensor for the free wave equation $\Box_{\hat{g}}\phi = 0$

$$T_{ab} = T_{(ab)} = \partial_a\phi\partial_b\phi - \frac{1}{2}\hat{g}_{ab}\hat{g}^{cd}\partial_c\phi\partial_d\phi \,,$$

which, for ϕ solution of (10.10), satisfies $\nabla^a T_{ab} = \Box\phi\partial_b\phi = -2mR\phi\partial_b\phi$. Contracting T_{ab} with T^a, we get the conservation law

$$\nabla^a\left(T^b T_{ab}\right) = T_{ab}\nabla^{(a}T^{b)} - 2mR\phi T^b\partial_b\phi \,. \tag{10.12}$$

The hope is then that the error terms (on the right hand side), have a sufficiently nice behaviour to allow a control via Gronwall-type arguments.

The energy density 3-form $E(\phi)$ associated with T^a is given by :

$$\begin{aligned} E(\phi) &:= T^a T_{ab} \mathrm{d}^3 x^b = T^a T^b_a \partial_b \lrcorner \,\mathrm{dvol}^4 \\ &= \left[u^2 \phi_u^2 + R^2(1-2mR)\left(u^2\phi_u\phi_R - (1+uR)\phi_R^2\right)\right. \\ &\quad \left. +(1+uR)|\nabla_{S^2}\phi|^2\right] \mathrm{d}u \wedge \mathrm{d}^2\omega \\ &\quad +\frac{1}{2}\left[\left((2+uR)^2 - 2mu^2R^3\right)\phi_R^2 + u^2|\nabla_{S^2}\phi|^2\right] \mathrm{d}R \wedge \mathrm{d}^2\omega \\ &\quad + \text{angular terms} \ . \end{aligned} \tag{10.13}$$

For a hypersurface $\mathcal{S}$, we denote

$$\mathcal{E}_{\mathcal{S}}(\phi) := \int_{\mathcal{S}} E(\phi) \, .$$

For instance,

$$\mathcal{E}_{\mathscr{I}^+}(\phi) = \int_{\mathscr{I}^+} \left[u^2\phi_u^2 + |\nabla_{S^2}\phi|^2\right] \mathrm{d}u \wedge \mathrm{d}^2\omega \, .$$

We foliate the domain $\{u < u_0 << -1\}$ by the spacelike (except for $s = 0$) hypersurfaces

$$\mathcal{H}_s := \{u = -s r_*\} \, , \ 0 \le s \le 1 \, ;$$

$\mathcal{H}_1$ is the $\{t = 0\}$ hypersurface and $\mathcal{H}_0$ corresponds to $\mathscr{I}^+$.

The energy on the surface $\mathcal{H}_s$ is given uniformly equivalent to

$$\mathcal{E}_{\mathcal{H}_s}(\phi) \ \simeq \ \int_{\mathcal{H}_s} \left(u^2\phi_u^2 + \frac{R}{|u|}\phi_R^2 + |\nabla_{S^2}\phi|^2\right) \mathrm{d}u \wedge \mathrm{d}^2\omega \, . \tag{10.14}$$

10.2.4 The fundamental energy estimates

For $m = 0$, T^a is a Killing vector field for our rescaled metric $\hat{g}$ (it is a conformal Killing vector field for g and in fact a Killing vector field for $\hat{g}$). For $m \neq 0$, the Killing form for T^a is given by

$$\nabla_{(a}T_{b)} = 4mR^2(3+uR)\mathrm{d}u^2 \, .$$

This gives

$$T_{ab}\nabla^a T^b = 4mR^2(3+uR)\phi_R^2$$

which exhibits a nice fall-off near $\mathscr{I}^+$ and i_0.

We need to choose a vector field V that will identify the different hypersurfaces $\mathcal{H}_s$, and the conservation law (10.12) integrated over the domain $\{u < u_0\}$, $u_0 << -1$, gives (draw a picture of the domain $\{u < u_0\}$ and define the three hypersurfaces $\mathscr{I}^+_{u_0}$, $\mathcal{S}_{u_0}$ and $\mathcal{H}_{1,u_0}$)

$$\begin{aligned} \int_{\{u<u_0\}} \nabla^a\left(T^b T_{ab}\right) \mathrm{d}^4\mathrm{Vol} &= \mathcal{E}_{\mathscr{I}^+_{u_0}} + \mathcal{E}_{\mathcal{S}_{u_0}} - \mathcal{E}_{\mathcal{H}_{1,u_0}} \\ &= \int_{\{u<u_0\}} \left(T_{ab}\nabla^{(a}T^{b)} - 2mR\phi T^b\partial_b\phi\right) \mathrm{d}^4\mathrm{Vol} \\ &= \int_0^1 \left(\int_{\mathcal{H}_{s,u_0}} \left(T_{ab}\nabla^{(a}T^{b)} - 2mR\phi T^b\partial_b\phi\right) V \lrcorner \,\mathrm{d}^4\mathrm{Vol}\right) \mathrm{d}s \, . \end{aligned}$$

Another way of understanding this kind of energy estimates that is more familiar to PDE analysts is the following : we multiply equation (10.10) by $T\phi$ and integrate the result between $s = 0$ and $s = 1$. This is not quite precise enough however : it is implicitely assumed that we have a well defined product structure on the domain $\{u < u_0\}$, s being a "time" variable and the space variables being chosen on $\mathscr{I}^+$ or on $\mathcal{H}_1$ or any other $\mathcal{H}_s$. It is in order to define this product structure that we need an identifying vector field V. It will perform a splitting of the 4-volume measure into a part along the integral curves of V that is simply $\mathrm{d}s$ and a 3-volume measure on each $\mathcal{H}_s$ defined as $V \lrcorner \, \mathrm{d}^4\mathrm{Vol}$.

In this case, a natural identifying vector field ν [1] is given by

$$\nu = r_*^2 R^2 (1 - 2mR)|u|^{-1}\partial_R \, . \tag{10.15}$$

The corresponding 3-volume measure on each $\mathcal{H}_s$ is then

$$\nu \lrcorner \, \mathrm{dVol}^4|_{\mathcal{H}_s} = r_*^2 R^2 (1 - 2mR)|u|^{-1} \mathrm{d}u \mathrm{d}^2\omega|_{\mathcal{H}_s} \, .$$

Such a choice of identifying vector field ν^a, which is really associated with the choice of parameter s for the foliation, will lead to error terms that cannot be controlled by the energy density[2] and therefore to the impossibility of performing a priori estimates.

To be more precise, the 4-volume error terms will be multiplied by

$$T^a \partial_a \phi r_*^2 R^2 (1 - 2mR)|u|^{-1}$$

and then considered as terms on the 3-surface $\mathcal{H}_s$ that we shall try to control by the energy density. Supposing that the (squares of the) 4-volume error-terms are barely controlled by the energy density, we have a problem, indeed :

$$T^a \partial_a \phi \nu \lrcorner \, \mathrm{dVol}^4 \simeq u\phi_u - 2(1 + uR)\frac{1}{u}\phi_R \, .$$

The first term is naturally controlled by the energy density on the $\mathcal{H}_s$ slices, but not the second, since u^{-1} is infinitely larger than Ru^{-1} near $\mathscr{I}^+$.

All we need to do in order to solve this problem is the following change of parameter :

$$\tau := -2\left(\sqrt{s} - 1\right) \, , \text{ so that } \frac{\partial \tau}{\partial s} = -\frac{1}{\sqrt{s}}. \tag{10.16}$$

The change of sign and the -1 term are there purely for aesthetic reasons, the important part is $2\sqrt{s}$. This new parameter varies from 0 to 2 as s varies from 1 to 0. We denote

$$\Sigma_{\tau(s)} = \mathcal{H}_s \, . \tag{10.17}$$

[1] It needs to satisfy

$$\frac{\partial s}{\partial \nu} = -\frac{1}{r_*}\frac{\partial u}{\partial \nu} + \frac{u}{r_*^2}\frac{\partial r_*}{\partial \nu} = 1 \, ,$$

whence the expression (10.15) of ν obtained by imposing that it is colinear to ∂_R, i.e. choosing ν parallel to the level hypersurfaces of u ; this is natural given the shape of our domain $\{u < u_0\}$.

[2] To be more precise, this does not occur for the fundamental estimates, because the scalar curvature $2mR$ gives us some extra fall-off at $\mathscr{I}$. For higher order estimates, commuting ∂_R into the equation will give error terms without any fall-off. So the problem will occur as soon as we try to gain one extra degree of regularity from the fundamental estimates.

The natural new identifying vector field is

$$V = -\sqrt{s}\,\nu = -\sqrt{\frac{|u|}{r_*}r_*^2 R^2(1-2mR)|u|^{-1}\partial_R} = -\left(r_* R\right)^{3/2}(1-2mR)\sqrt{\frac{R}{|u|}}\,\partial_R\,. \quad (10.18)$$

We get error terms on each Σ_τ that are equivalent to :

$$\begin{aligned}
(\nabla^a T^b)T_{ab}V \lrcorner \,\mathrm{dvol}^4 &\simeq R^2\sqrt{\frac{R}{|u|}}\phi_R^2 \mathrm{d}u\mathrm{d}^2\omega\,,\\
\left(-2mR\phi T^b\partial_b\phi\right) V \lrcorner \,\mathrm{dVol}^4 &\simeq R\phi\sqrt{\frac{R}{|u|}}\left(u^2\partial_u\phi - 2(1+uR)\partial_R\phi\right)\mathrm{d}u\mathrm{d}^2\omega\\
&\lesssim \left(R^2\phi^2 + u^2\phi_u^2 + \frac{R}{|u|}\phi_R^2\right)\mathrm{d}u\mathrm{d}^2\omega\,.
\end{aligned}$$

The only difficulty comes from the zero order term. We solve it by proving (simply by integration by parts) the following estimate :

Lemma 10.1. *Given $u_0 < 0$, there exists a constant $C > 0$ such that for any $f \in \mathcal{C}_0^\infty(\mathbb{R})$, we have*

$$\int_{-\infty}^{u_0} (\phi(u))^2\,\mathrm{d}u \leq C\int_{-\infty}^{u_0} u^2\left(\phi'(u)\right)^2\mathrm{d}u\,.$$

This entails that the energy density controls the L^2 norm on the $\mathcal{H}_s$ slices.

We see that this is much more than what we need, but for higher order estimates, we will use all that this lemma gives us.

This allows us to obtain, via simple Gronwall estimates, the following estimates

Theorem 10.1. *For $u_0 < 0$, $|u_0|$ large enough, there exists a constant $C > 0$ such that,*

$$\mathcal{E}_{\mathscr{I}^+_{u_0}}(\phi) \leq C\mathcal{E}_{\mathcal{H}_{1,u_0}}(\phi)\,,$$

$$\mathcal{E}_{\mathcal{H}_{1,u_0}}(\phi) \leq C\left(\mathcal{E}_{\mathscr{I}^+_{u_0}}(\phi) + \mathcal{E}_{\mathcal{S}_{u_0,s_0}}(\phi)\right)\,.$$

10.2.5 Higher order estimates

Contrary to what happens for $m = 0$, where T is a Killing vector field, we cannot use it here to raise the regularity, since

$$[T,[T,2mR]] = 4mu(2+uR)\,,$$

so differentiating the equation several times using T will introduce potentials which blow up near i^0. Although this may appear as unfortunate, it is a blessing in disguise.

Instead, we use the vector field ∂_R. We obtain

$$(\Box + 2mR)\,\phi_R = 2(1-3m)R\partial_R\phi_R - 2(1-6mR)\phi_R - 2m\phi\,. \quad (10.19)$$

All the terms in the right hand-side can be controlled by the energy density for ϕ_R or ϕ using when necessary the lemma above. It is clear that further differentiations using ∂_R will not raise any difficulty. We get the estimates

Theorem 10.2. *For each $k \in \mathbb{N}$, there exists a constant $C_k > 0$ such that, for any solution ϕ of (10.10) associated to smooth compactly supported initial data, we have for all $0 \leq s \leq 1$,*

$$\begin{aligned} \mathcal{E}_{\mathscr{I}_{u_0}}(\partial_R^k \phi) &\leq C \sum_{p=0}^{k} \mathcal{E}_{\mathcal{H}_{1,u_0}}(\partial_R^p \phi)\,, \\ \mathcal{E}_{\mathcal{H}_{1,u_0}}(\partial_R^k \phi) &\leq C \sum_{p=0}^{k} \left(\mathcal{E}_{\mathscr{I}^+_{u_0}}(\partial_R^p \phi) + \mathcal{E}_{\mathcal{S}_{u_0}}(\partial_R^p \phi) \right) . \end{aligned}$$

10.3 Conclusion

This gives us a simple characterization of the peeling

Definition 10.4. *We say that a solution ϕ of (10.10) peels at order $k \in \mathbb{N}$ if for all polynomial P in ∂_R and ∇_S^2 of order lower than or equal to k, we have $\mathcal{E}_{\mathscr{I}^+_{u_0}}(P\phi) < +\infty$. This means than for all $p \in \{0, 1, ..., k\}$ we have for all $q \in \{0, 1, ..., p\}$, $\mathcal{E}_{\mathscr{I}^+_{u_0}}(\partial_R^q \nabla_{S^2}^{p-q} \phi) < +\infty$.*

This is of course valid for Minkowski space as well. So in flat spacetime, we have two definitions of the peeling :

- a first one obtained using the embedding in the Einstein cylinder and the time translation of the Einstein cylinder (quite close to the Morawetz cector field) ;
- a second one using the time asymmetric compactification and the vector field ∂_R.

It turns out these two definitions are different. The class of data given by the second definition is larger than that given by the first. This is due to the fact that the vector field we use for higher order estimates is characteristic (or null if one prefers this terminology), which gives less stringent conditions than a timelike vector field that controls all spacelike derivatives via the equation (the energies on $\mathcal{H}_1$ given by both constructions are equivalent).

So this is a complete verification of the peeling model at all orders for the wave equation on the Schwarzschild metric, as well a definition that is different and possibly more relevent definition of the sets of solutions admitting peeling at a given order. Why more relevent? Because we show that we merely need to control the null derivative in the direction of $\mathscr{I}^+$ instead of a timelike derivative. It is in a way a propagation estimate : the other null derivative does not give a contribution on $\mathscr{I}^+$ so even if it lies in very weakly regular spaces, it will not interfere with the regularity of the solution at $\mathscr{I}^+$.

Bibliography

[1] A.N. Bernal, M. Sánchez, *On smooth Cauchy hypersurfaces and Geroch's splitting theorem*, Commun. Math. Phys. **243** (2003), p. 461–470.

[2] A.N. Bernal, M. Sánchez, *Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes*, Commun. Math. Phys. **257** (2005), p. 43–50.

[3] G.D. Birkhoff, *Relativity and modern physics*, with the cooperation of R. E. Langer, Cambridge: Harvard University Press, 1923.

[4] R.H. Boyer & R.W. Lindquist, *Maximal Analytic Extension of the Kerr Metric*, J. Math. Phys. **8** (1967), p. 265–281.

[5] S. Chandrasekhar, *The mathematical theory of black holes*, Oxford University Press 1983.

[6] M. Dafermos & I. Rodnianski, *The redshift effect and radiation decay on black hole spacetimes*, arXiv gr-qc:/0512119, 2006.

[7] Eddington

[8] G.F.R. Ellis, S.W. Hawking, *The large scale structure of space-time*, Cambridge Monographs in Mathematical Physics, 1973.

[9] Finkelstein

[10] R.P. Geroch, *The domain of dependence*, J. Math. Phys. **11** (1970), 437–439.

[11] S.W. Hawking & G.F.R Ellis, *The large scale structure of space-time*, Cambridge monographs in mathematical physics, Cambridge University Press 1973.

[12] L. Hörmander, *A remark on the characteristic Cauchy problem*, J. Funct. Anal. **93** (1990), 270-277.

[13] S. Klainerman & F. Nicolò, *Peeling properties of asymptotically flat solutions to the Einstein vacuum equations*, Class. Quantum Grav. **20** (2003), p. 3215–3257.

[14] Kruskal, Szekeres

[15] J. Leray, *Hyperbolic differential equations*, Princeton Institute for Advanced Studies 1952.

[16] L. Mason, J.-P. Nicolas, *Conformal scattering and the Goursat problem*, J. Hyperbolic Differ. Equ. **1** (2004), 2, 197–233.

[17] C.W. Misner, K. Thorne, J.A. Wheeler, *Gravitation*, Freeman, San Francisco, 1973.

[18] C.S. Morawetz, *The decay of solutions of the exterior initial-boundary value problem for the wave equation*, Comm. Pure Appl. Math. **14** (1961), p. 561–568.

[19] B. O'Neill, *The geometry of Kerr black holes*, A.K. Peters, Wellesley, 1995.

[20] J.R. Oppenheimer & H. Snyder, *On continued gravitational contraction*, Phys. Rev. **56** (1939), p. 455–459.

[21] R. Penrose, *Asymptotic properties of fields and spacetime*, Phys. Rev. Lett. **10** (1963), 66–68.

[22] R. Penrose, *Conformal approach to infinity*, in Relativity, groups and topology, Les Houches 1963, ed. B.S. De Witt and C.M. De Witt, Gordon and Breach, New-York, 1964.

[23] R. Penrose, *Zero rest-mass fields including gravitation : asymptotic behaviour*, Proc. Roy. Soc. London **A284** (1965), 159–203.

[24] R. Penrose, W. Rindler, *Spinors and space-time*, Vol. I & II, Cambridge monographs on mathematical physics, Cambridge University Press 1984 & 1986.

[25] R. Wald, *General Relativity*, University of Chicago Press, 1984.

[26] H. Whitney, *Differentiable functions defined in arbitrary subsets of Euclidean space*, Trans. Amer. Math. Soc. **40** (1936), no. 2, p. 309–317.

[27] C. Zuily & H. Queffélec *Éléments d'analyse pour l'agrégation*, Masson, 1995.