

Lectures on Lorentzian Geometry and hyperbolic pdes

Jacques Smulevici*

April 8, 2019

Contents

1 Preliminary remarks	3
2 Review of differential geometry	4
2.1 Differential manifolds	4
2.2 Differential mappings	5
2.3 Tangent vectors	5
2.4 The differential of a map	6
2.5 The inverse function theorem	6
2.6 Curves	6
2.7 Vector fields and tangent bundle	6
2.8 One forms	7
2.9 Einstein summation convention	8
2.10 Submanifolds	8
2.11 Immersions, submersions, embedding	9
2.12 Integral curves	10
2.13 Tensor fields	10
2.14 Tensors at a point	11
2.15 Contraction	12
2.16 Pull-back	13
2.17 Vector bundles	13
3 Semi-Riemannian manifolds	13
3.1 Semi-Riemannian metrics	13
3.2 Examples and basic constructions	17
3.3 The Levi-Civita connection	22
3.4 Isometries	26
3.5 The induced connection on a curve, parallel translations and geodesics	28
3.6 Physical interpretation of causal geodesics	32
3.7 The exponential map	33
4 Curvature	37
4.1 The Riemann curvature tensor and its symmetries	37
4.2 The Ricci tensor and the scalar curvature	40

*Laboratoire de Mathématiques, Université Paris-Sud 11, bât. 425, 91405 Orsay, France.

5	Divergence free vector and tensor fields	41
5.1	Some differential operators	41
5.2	Remarks on conservation laws and Lagrangian theory	43
5.3	The energy-momentum tensor	44
5.4	The Einstein tensor	46
6	The Einstein equations	47
6.1	A model problem: the Maxwell equations	48
6.2	Hyperbolicity I: Wave equations satisfied by the Riemann tensor	49
6.3	Hyperbolicity II: Wave coordinates	50
7	Semi-Riemannian submanifolds	52
7.1	The induced connection	53
7.2	Semi-Riemannian submanifolds and the Levi-Civita connection of the induced metric	53
7.3	The shape tensor	55
7.4	The Gauss equation	55
7.5	The normal connection and the Codazzi equation	56
7.6	The constraint equations	57
7.7	Normal frame field	57
7.8	The normal bundle	58
8	Local geometry	59
8.1	Two parameters maps	59
8.2	The Gauss lemma	60
8.3	Hyperquadrics	62
8.4	Convex neighborhoods	64
8.5	Arc lengths	66
8.6	The Riemannian case	67
8.7	The Lorentzian case	73
9	Deformation of curves	77
9.1	Variations of a curve	77
9.2	From causal to timelike curves	77
9.3	First variation	79
10	Causality and global Lorentzian geometry	81
10.1	Causality and topology	82
10.2	Various limits for sequences of curves	84
10.3	Causality conditions	88
10.4	Achronal sets	89
10.5	Cauchy hypersurfaces	92
10.6	Global hyperbolicity	94
10.7	Cauchy developments	95
10.8	Time and temporal functions	98
11	The domain of dependence in globally hyperbolic spacetime	106
11.1	Three extra geometric lemmas	107
11.2	A local version of the domain of dependence property	107
11.3	Proof of the domain of dependence	110

12 Symmetric hyperbolic systems	110
12.1 Solving PDEs: representation formula vs abstract methods	110
12.2 Symmetric hyperbolic systems on \mathbb{R}^{n+1}	111
12.3 A priori estimates for symmetric hyperbolic systems	111
12.4 A rough local uniqueness statement	113
12.5 Existence by duality	114
12.6 Applications to linear wave equations on \mathbb{R}^{n+1}	115
13 Existence theory for the wave equation on globally hyperbolic spacetimes	116
14 Local existence for the vacuum Einstein equations	117
A Global volume form and orientable manifolds	120
B The pull-back bundle	121
C Not so useful facts	121
D Taylor expansions of the metric tensor and volume form, following Alfred Gray [Gra73, Gra04]	121
E Additional exercises	124
E.1 Problems	124
E.2 Solutions	124

1 Preliminary remarks

These lectures have been designed for a course given at the Ecole Normale of Paris in 2017.

I used the following books to prepare these lectures

1. Semi-Riemannian Geometry by O’Neil [O’N83],
2. Global Lorentzian Geometry by Beem, Ehrlich and Easley [O’N83],
3. The Cauchy problem in General Relativity by H. Ringström [Rin09],
4. Mathematical problems of General Relativity I. by D. Christodoulou [Chr08].

Most of the beginning follows [O’N83]. The introduction to the Einstein equations is more personal. I used [O’N83] in Section 10. The construction of the time and temporal functions is taken from [Rin09], which itself followed the papers [BS05], [O’N83], [BS03]. We claim no originality in these notes, neither in their content, nor in the way it is written (but we do hope the students will find these notes useful!).

For a student who would be interested in an introduction to General Relativity from a Physics point of view, a good reference is D’Inverno, *Introducing Einstein’s Relativity* [d1992introducing].

The goals of the lectures are to

1. define the basic objects of Riemannian and Lorentzian geometry,
2. give some of the fundamental results in Riemannian and Lorentzian geometry.

- introduce the student to some analysis (pdes) problems and to the (by now) standard techniques to tackle them. More specifically, we will focus on hyperbolic pdes arising naturally in the context of Lorentzian geometry.

To set up a target, one of the main goal of these lectures will be to prove the following theorem¹

Theorem 1.1. *Let (M, g) be a smooth, oriented, time-oriented Lorentzian manifold. Assume that (M, g) is globally hyperbolic and that Σ is a smooth Cauchy hypersurface with future unit normal n . Let ψ_0, ψ_1 be smooth functions on Σ and F a smooth function on M . Then the Cauchy problem*

$$\begin{aligned}\square_g \psi &= F, \\ \psi_\Sigma &= \psi_0, \\ n(\psi_\Sigma) &= \psi_1\end{aligned}$$

admits a unique smooth solution ψ .

Obviously, we need to define everything above. In particular, \square_g will be a wave type operator generalizing the classical wave operator $-\partial_t^2 + \Delta$ and Σ will be replacement for the hypersurface $t = 0$ in \mathbb{R}^{n+1} . The pure pde aspect of the problem will actually be easy (say easier than solving a general hyperbolic pde in \mathbb{R}^{n+1} as seen in the PDE class). The hard part will be essentially to split the manifold on small pieces, on which we can solve the wave equation, and then glue everything together. To allow such a cut and paste argument we need the geometric assumption of "globally hyperbolicity" and a large part of the lectures will try to prepare us to define and use that assumption.

2 Review of differential geometry

I expect the students to already be acquainted with most of the content of this section. Most proofs are left to the reader.

The basic object we will need is that of a differential manifold.

2.1 Differential manifolds

Definition 2.1. *Let M be a topological space. A coordinate system on M is an homeomorphism ξ from an open set $U \subset M$ into an open set of \mathbb{R}^n .*

Two coordinate ξ_1 and ξ_2 systems overlap smoothly if $\xi_1 \circ \xi_2^{-1}$ and $\xi_2 \circ \xi_1^{-1}$ are smooth (exercise: on which domains of definition ?).

Definition 2.2. *An atlas on M is a collection of coordinate systems such that the union of all their domains of definition cover M and such that any two coordinate systems overlap smoothly.*

Exercise 2.1. *Recall the definition of the dimension of M . What if we replace smooth by continuous ? (a lot harder)*

¹This theorem could be attributed to Leray, following his 1952 lectures on hyperbolic pdes. The notion of global hyperbolicity used in his lectures is slightly different and in fact he proves much more, since his main interest was general hyperbolic pdes, not just the wave equation.

An atlas A is called complete if it contains each coordinate system that overlaps smoothly with every coordinate systems of A .

Lemma 2.1. *Every atlas is contained in a unique complete atlas.*

Definition 2.3. *A smooth manifold is a Hausdorff space endowed with a complete atlas.*

We will also assume in all of these lectures that all manifolds considered are second countable (i.e. admits a countable basis of open sets).

Exercise 2.2. 1. *what does Hausdorff mean again ?*

2. *what would be a good definition of an open submanifold of M ?*

3. *Same with product manifolds of M and N ?*

2.2 Differential mappings

Definition 2.4. *Let M and N be manifolds. A mapping $\phi : M \rightarrow N$ is smooth if for any coordinate systems (U_M, ξ_M) and (U_N, ξ_N) of M and N , we have $\xi_N \circ \phi \circ \xi_M^{-1}$ smooth.*

Notation : We will write $\mathcal{F}(M)$ for the set of smooth real valued functions of M .

2.3 Tangent vectors

Let $p \in M$ (M manifold.) A tangent vector X at p is a real-valued function

$$X : \mathcal{F}(M) \rightarrow \mathbb{R},$$

such that

1. X is \mathbb{R} -linear.
2. $X(fg) = f(p)X(g) + g(p)X(f)$.

(Think of X as a derivation at p).

Let $T_p M$ denote the set of all tangent vectors at p . $T_p M$ is a real vector space (for which operations ?) called the tangent space to M at p .

Definition 2.5. *Let $(U, \xi = (x^\alpha))$ be a coordinate system of M and $p \in U$. Let $f \in \mathcal{F}(M)$. Then,*

$$\partial_{x^i} f(p) := \frac{\partial f \circ \xi^{-1}}{\partial x^i}(\xi(p)).$$

Let $[\partial_{x^i}]|_p : f \rightarrow \partial_{x^i} f(p)$. One then easily check that $[\partial_{x^i}]|_p \in T_p M$. Often, we will drop the p index and simply write ∂_{x^i} .

Lemma 2.2. *Let $X \in T_p M, f, g \in \mathcal{F}(M)$. Then,*

1. *if $f = g$ on some neighborhood of p , then $X(f) = X(g)$.*
2. *if $f = \text{const}$ on some neighborhood of p , then $X(f) = 0$.*

This lemma implies that tangent vectors are purely local objects.

Theorem 2.1. *If $\xi = (x^\alpha)$ is a coordinate system of M , then (∂_{x^α}) form a basis of $T_p M$. In particular, $\dim T_p M = \dim M$.*

2.4 The differential of a map

Definition 2.6. Let M, N be smooth manifolds and $\phi : M \rightarrow N$ smooth. For any $X \in T_p M$, define $Y_X : \mathcal{F}(N) \rightarrow \mathbb{R}$ as $Y(g) = X(g \circ \phi)$. Then, $Y \in T_{\phi(p)} N$. The map sending each such X to Y_X is called the differential (or pushforward) of ϕ at p and denoted $d\phi_p$. ($d\phi_p : T_p(M) \rightarrow T_{\phi(p)} N$).

Exercise 2.3. Show that $d\phi_p$ is linear and compute it in local coordinates.

A specific case is that of $N = \mathbb{R}$ and $f \in \mathcal{F}(M)$. By definition,

$$df_p : T_p M \rightarrow T_{f(p)} \mathbb{R} \simeq \mathbb{R}$$

so it can be considered as an element of $(T_p M)^*$.

2.5 The inverse function theorem

Theorem 2.2. Let $\phi : M \rightarrow N$ be a smooth mapping. The differential map $d\phi_p$ at some $p \in T_p M$ is a linear isomorphism if and only if there is a neighborhood \mathcal{N} of p in M such that $\phi|_{\mathcal{N}}$ is a diffeomorphism from \mathcal{N} onto a neighborhood $\phi(\mathcal{N})$ of $\phi(p) \in N$.

Proof. This is a consequence of the usual inverse function theorem for subsets of \mathbb{R}^n . \square

2.6 Curves

A curve is smooth map $\gamma : I \rightarrow M$, where I is a non-empty interval. Given a curve γ , we write

$$\dot{\gamma}(s) = \frac{d\gamma}{ds}(s) = d\gamma_s \left[\frac{d}{ds} \right]_s \in T_{\gamma(s)} M.$$

A piecewise smooth curve is a continuous map $\gamma : I \rightarrow M$, where I is a non-empty interval such that if $I = [a, b]$, then there exists a finite sequence $a = t_0 < t_1 < \dots < t_k = b$ such that each curve segment $\gamma|_{[t_i, t_{i+1}]}$ is a smooth curve, while for an interval with open endpoints, say $I = (a, b)$, we require that for each non-empty closed interval $J \subset I$, the restriction $\gamma|_J$ is piecewise smooth (this implies that the break points have no cluster point in I).

2.7 Vector fields and tangent bundle

Let $TM = \bigsqcup_{p \in M} T_p M$ be the disjoint union of all the tangent planes of M . We call TM the tangent bundle of M . Associated to TM is the a canonical projection defined by

$$\begin{aligned} \pi : TM &\rightarrow M \\ (p, X) &\rightarrow p \end{aligned}$$

TM can be given a natural manifold structure as follows. Let (x^α) be a coordinate system of M . Let $p \in T_p M$. For any $v \in T_p M$, we have

$$v = v^\alpha \partial_{x^\alpha}$$

The (ν^α) then forms a global coordinate system on $T_p M$ (called conjugate to the (x^α)). Given $(p, X) \in TM$, let $\tilde{\xi}(p, X) = (\xi(p), X^\alpha)$, then $\tilde{\xi}$ defines a local coordinate system of TM and one easily verifies that any two such local coordinate systems overlap smoothly.

The tangent bundle is in fact the basic example of a vector bundle over M .
A vector field on M is by definition a section of TM , i.e. a map of the form

$$X : M \rightarrow TM$$

satisfying $\pi \circ X = Id_M$. We will write X_p to denote the second component of $X(p)$, i.e. $X_p \in T_p M, \forall p \in M$.

Given a vector field X and $f \in \mathcal{F}(M)$, the function $X(f)$ is by definition $X(f) : p \rightarrow X_p(f)$. The set $\Gamma(M)$ of vector fields on M can be naturally given the structure of a module over the ring $\mathcal{F}(M)$.

A vector field X is smooth if $X(f)$ is smooth for any f .

Given a local coordinate system (x^α) , we associate the vector fields ∂_{x^α} , defined so that $[\partial_{x^\alpha}]_p(f) = [\partial_{x^\alpha}]_p(f)$.

For any vector field X , we have

$$X = \sum_{\alpha} X(x^\alpha) \partial_{x^\alpha}.$$

Identification with derivation

A derivation on $\mathcal{F}(M)$ is a \mathbb{R} -linear function $\mathcal{D} : \mathcal{F}(M) \rightarrow \mathcal{F}(M)$ which satisfies the usual Leibniz property

$$\mathcal{D}(fg) = \mathcal{D}(f) \cdot g + f \cdot \mathcal{D}(g).$$

Exercise 2.4. Prove that every vector field defines a derivation and that every derivation can be realized as a vector field.

The Lie bracket

Definition 2.7. Let X, Y be vector fields. Then, we define the vector field $[X, Y]$ by

$$[X, Y]_p(f) = X_p(Y(f)) - Y_p(X(f)).$$

Push forward of a vector field by diffeomorphism

Definition 2.8. Let ϕ be a diffeomorphism $\phi : M \rightarrow N$ and X a vector field on M . Then, we can define a vector field on N by the formula

$$(d\phi X)g = X(g \circ \phi) \circ \phi^{-1}, \quad \forall g \in \mathcal{F}(N).$$

Note that this definition breaks down if ϕ is merely a smooth function.

2.8 One forms

Let $T_p M^* := (T_p M)^*$ be the vector space dual of $T_p M$. Its elements are called covectors. Let TM^* denote the cotangent bundle, obtained as above by the disjoint union of the $T_p M^*$. A one form is then a smooth section of TM^* (ex: give TM^* the structure of a smooth vector bundle). The set of 1-forms is denoted $\Lambda^1(M)$.

Definition 2.9. Let $f \in \mathcal{F}(M)$. Then, we define $df \in \Lambda^1(M)$ as $df(X) = X(f)$ for all tangent vector X to M .

Given a coordinate system (x^α) on M , for each α , the differential of the function x^α , dx^α is thus a 1-form.

Let ∂_{x^α} denotes the coordinate vector fields, then the (dx^α) provides at each point a dual basis to the (∂_{x^α}) . Moreover, we have

$$df = \sum_{\alpha} \partial_{x^\alpha}(f) dx^\alpha.$$

2.9 Einstein summation convention²

By convention, we will remove the \sum symbols everytime we have repeated indices. For instance, in the previous formula, we write

$$df = \partial_{x^\alpha}(f) dx^\alpha.$$

Again by convention, indices corresponding to the components of vectors are lowered, and indices corresponding to components of covectors are put in higher positions. With this convention, a repeating sum will always involve indices up and down (and no sum with indices only up should appear).

2.10 Submanifolds

Definition 2.10. A manifold P is a submanifold of M if

- $P \subset M$ and P has the induced topology of M .
- The inclusion map $i : P \rightarrow M$ is smooth and its differential is injective.

Exercise 2.5. Consider $M = \mathbb{R}^2$ and $P = \mathbb{R}(0, 1) + \mathbb{R}(1, 0)$ as a topological subspace of M .

1. We consider the map

$$\psi : P \rightarrow \mathbb{R}$$

defines by $\psi(y, 0) = y$, $\psi(0, x) = -x$. Prove that ψ is homeomorphism and thus that it defines a global chart on P . It follows that P is manifold.

2. Prove that the inclusion map $i_P : P \rightarrow M$ is not smooth.
3. Let now $Q = \{(t, 0) : t \in [0, 1]\} \cup \{(0, t) : t \in [0, 1]\}$. Again, we consider Q as a topological subspace of \mathbb{R}^2 . We us define the map

$$\phi : Q \rightarrow \mathbb{R}$$

by $\phi(e^{-1/t^2}, 0) = t$, $\phi(0, e^{-1/t^2}) = -t$, $\phi(0, 0) = 0$. We recall that $t \neq 0 \rightarrow e^{-1/t^2}$ can be extended smoothly by 0 at $t = 0$. Prove that ϕ is homeomorphism and thus defines a global chart on Q , making Q a manifold.

4. Prove that the inclusion map $i_Q : Q \rightarrow M$ is smooth but that its differential is not injective.

²According to the usual folklore knowledge, Einstein thought of this as his greatest invention. Let me know if you know an exact citation on this.

5. The restriction of ψ to Q defines another potential candidate for a coordinate system on Q . Prove that the two coordinate systems do not overlap smoothly.

Proposition 2.1. *If P^m is a submanifold of M^n , then for all $p \in P$, there exists a coordinate system $(x^1, \dots, x^m, x^{m+1}, \dots, x^n)$ defined on $U \subset M$ such that $p \in U$ and*

$$P = \{(x^{m+1} = 0, \dots, x^n = 0)\}.$$

Proof. Use the inverse function theorem. □

We call such a coordinate system adapted to P .

Proposition 2.2. *A subset $P \subset M$ is a submanifold of M if for each point p of P , there is a coordinate system of M adapted to P .*

Let P be a submanifold of M , $p \in P$ and denote the inclusion map by i . Then

$$di_p : T_p P \rightarrow T_p M.$$

Since di_p is injective, we will identify $T_p P$ with $di_p(T_p P)$ and therefore view $T_p P$ as a vector subspace of $T_p M$. A vector field X on M is then tangent to P , if for all $p \in P$, $X_p \in T_p P$.

2.11 Immersions, submersions, embedding

Definition 2.11. *An immersion is a smooth map $\phi : M \rightarrow N$ such that $d\phi_p$ is injective for all $p \in M$.*

Definition 2.12. *An embedding of P into M is a immersion $\phi : P \rightarrow M$ such that*

1. ϕ is injective.
2. the induced map $\phi^{\phi(P)} : \phi \rightarrow \phi(P)$ is an homeomorphism.

Clearly, the restriction of any immersion to any sufficiently small open set is an imbedding.

Exercise 2.6. 1. *If P is a submanifold of M then the inclusion map $i : P \rightarrow M$ is automatically an imbedding.*

2. *Conersely, let $\phi : P \rightarrow M$ be an imbedding. Give $\phi(P)$ a manifold structure such that $\bar{\phi} : P \rightarrow \phi(P)$ is a diffeomorphism. Then $\phi(P) \subset M$ and the inclusion is an immersion, so that $\phi(P)$ is a submanifold of M .*

Definition 2.13. *Let $\psi : M^m \rightarrow N^n$ be a smooth map and let $q \in N$. q is said to be regular provided that $d\psi_p$ is onto, for every $p \in \psi^{-1}(q)$.*

Lemma 2.3. *If $q \in \psi(M)$ is regular, then $\psi^{-1}(q)$ is a submanifold of M and $\dim M = \dim N + \dim \psi^{-1}(q)$.*

By definition, a *hypersurface* is a submanifold of codimension 1. They can be constructed from the previous lemma with $N = \mathbb{R}$.

Corollary 2.1. *Let $f : M \rightarrow \mathbb{R}$ and $c \in \text{Im}(f)$. If $df_p \neq 0$ for every $p \in f^{-1}(c)$, then $f^{-1}(c)$ is a hypersurface of M (called a level set of f).*

Definition 2.14. *A submersion $\psi : M \rightarrow B$ is a smooth mapping onto B such that $d\psi_p$ is onto for all $p \in M$.*

2.12 Integral curves

A curve α is an integral curve of $X \in \Gamma(M)$ if $\dot{\alpha} = X_\alpha$.

In coordinates, an integral curve is a solution to

$$\frac{d(x^i \circ \alpha)}{ds} = X^i \circ \alpha.$$

By Cauchy-Lipshitz, for any $X \in \Gamma(M)$ and any $p \in M$, there exists a unique maximal integral curve of X starting at p .

Exercise 2.7. Prove that for any vector field X such that $X_p \neq 0$, there exists a local coordinate system (x^α) at p such that $X = \partial_{x^1}$.

Definition 2.15. Let $p \in M$. A local flow of X is a neighborhood U of p , a neighborhood I of $0 \in \mathbb{R}$ and a map ϕ of the form

$$\begin{aligned} \phi: I \times U &\rightarrow M \\ (t, p) &\rightarrow \phi_t(p) \end{aligned}$$

where $\phi_t(p)$ is the value at time t of the solution of the integral curve equation for X with initial data $(0, p)$.

In fact, still by standard ode theory, there exists an open set $D \subset \mathbb{R} \times U$ containing $\{0\} \times U$, such that ϕ is well-defined on D .

A vector field is called *complete* if its flow is defined on $\mathbb{R} \times M$, or another words, if its integral curves are all defined for all times.

2.13 Tensor fields

Definition 2.16. A tensor field T of type (r, s) over a manifold is a map of the form

$$T: (\Lambda^1(M))^r \times (\Gamma(M))^s \rightarrow \mathcal{F}(M)$$

which is $\mathcal{F}(M)$ -multilinear.

We will denote by T_s^r the set of tensor fields of type (r, s) .

Definition 2.17 (Tensor product). Given T_1 a tensor field of type (r_1, s_1) and T_2 of type (r_2, s_2) , we define $T_1 \otimes T_2$ as

$$\begin{aligned} T_1 \otimes T_2: (\Lambda^1(M))^{r_1+r_2} \times (\Gamma(M))^{s_1+s_2} &\rightarrow \mathcal{F}(M) \\ (\theta_1, \dots, \theta_{r_1+r_2}, X_1, \dots, X_{s_1+s_2}) &\rightarrow T_1(\theta_1, \dots, \theta_{r_1}, X_1, \dots, X_{s_1}) T_2(\theta_{r_1+1}, \dots, \theta_{r_1+r_2}, X_{s_1+1}, \dots, X_{s_1+s_2}) \end{aligned}$$

Exercise 2.8. • Define a (r, s) tensor bundle and tensor fields as a section of that bundle.

- Recall that the Lie Bracket $[\cdot, \cdot]$ maps a pair of vector fields to a vector field. Thus, we can define a map ψ on $(\Lambda^1(M))^1 \times (\Gamma(M))^2$ by

$$\psi(\omega, X, Y) = \omega([X, Y]).$$

Prove that ψ does not define a tensor $(1, 2)$ tensor field.

- Show that $(0, 0)$ tensor fields, $(1, 0)$ tensor fields and $(0, 1)$ tensor fields can be identified with functions, vector fields and one-forms respectively.

2.14 Tensors at a point

First a classical lemma.

Lemma 2.4. *Given any neighborhood \mathcal{U} of a point $p \in M$, there is a function $f \in \mathcal{F}(M)$, called a bump function at p , such that*

1. $0 \leq f \leq 1$ on M ,
2. $f = 1$ on some neighborhood of p ,
3. The support of f is included in \mathcal{U}

$$\text{supp } f := \overline{\{p \in M : f(p) \neq 0\}} \subset U.$$

Proof. Exercise. □

Proposition 2.3. *Let $p \in M$ and $A \in T_s^r$. Let $(\theta^i)_{1 \leq i \leq s}$, $(v^i)_{1 \leq i \leq s}$ be $2s$ one forms such that $\theta_p^i = v_p^i$ for all i . Similarly, let $(X_i), (Y_i)$ for $1 \leq i \leq r$ be $2r$ vector fields on M such that $X_i(p) = Y_i(p)$.*

Then,

$$A(\theta^1, \theta^2, \dots, \theta^s, X^1, \dots, X^r)(p) = A(v^1, \dots, v^s, Y^1, \dots, Y^r)(p).$$

Proof. We do the proof in the case of tensor fields of type $(1, 1)$. Let θ be a one form and X be a vector field. Let (U, x^α) be a coordinate system around p and let f be a bump function at p with support in U . We can decompose X and θ ,

$$X = X^\alpha \partial_{x^\alpha}, \tag{1}$$

$$\theta = \theta_\alpha dx^\alpha, \tag{2}$$

Here any of the components X^α , θ_α are smooth functions on U . Moreover, fX^α , $f\theta_\alpha$, $f\partial_{x^\alpha}$, $f dx^\alpha$ are all smooth in U and can be extended trivially to the whole of M by 0. (We denote the resulting object by the same letters).

Thus, by $\mathcal{F}(M)$ -multilinearity

$$A(f^2\theta, f^2X) = f\theta_\alpha \cdot fX^\beta A(f dx^\alpha, f\partial_{x^\beta}).$$

Evaluated at p , this gives

$$f^4(p)A(\theta, X)(p) = A(f^2\theta, f^2X)(p) = f(p)\theta_\alpha(p)f(p)X^\beta(p)A(f dx^\alpha, f\partial_{x^\beta})(p),$$

so that

$$A(\theta, X)(p) = \theta_\alpha(p)X^\beta(p)A(f dx^\alpha, f\partial_{x^\beta})(p).$$

□

This implies that given a tensor field A of type (r, s) , we can define a map

$$A_p : (T_p^*M)^r \times (T_pM)^s \rightarrow \mathbb{R}$$

as follows. Given $\omega^1, \dots, \omega^r \in T_p^*M$ and $x_1, \dots, x_s \in T_pM$, let θ^i, X^i be respectively one forms and vector fields such that $\theta^i(p) = \omega^i$, $X_i(p) = x_i$. Then, define

$$A_p(\omega^1, \dots, \omega^r, x_1, \dots, x_s) = A(\theta, X)(p).$$

At each p , A_p is \mathbb{R} -multilinear (it is thus a tensor over the vector space T_pM).

Moreover, the above proposition implies that given any open set U of M , the restriction $A|_U$ of A to U is a well defined tensor field on U .

In particular,

Definition 2.18. Let $(U, \xi = (x^\alpha))$ be a coordinate system. Then, for any tensor A the components of A (relative to ξ) are the real-valued functions

$$A_{j_1 \dots j_s}^{i_1 \dots i_r} = A(dx^{i_1}, \dots, dx^{i_r}, \partial_{j_1}, \dots, \partial_{j_s}).$$

Lemma 2.5. With the above notation, the tensor A verifies on U

$$A = A_{j_1 \dots j_s}^{i_1 \dots i_r} \partial_{i_1} \otimes \dots \otimes \partial_{i_r} \otimes dx^{j_1} \dots \otimes dx^{j_s}.$$

Proof. Show that the left and right-hand side take the same value on each $(dx^{i_1}, \dots, dx^{i_r}, \partial_{j_1}, \dots, \partial_{j_s})$ and use the $\mathcal{F}(U)$ multi-linearity. \square

Tensor transformation rules

We give the transformation rule for tensor of type $(1, 1)$ and leave the general case as an exercise.

Lemma 2.6. Let A be a tensor field of type $(1, 1)$ and let $(x^\alpha), (y^i)$ be two coordinate systems defined on a common open set U . Denote the components of A with respect to (x) as a_β^α and with respect to (y) as b_j^i , then

$$a_\beta^\alpha = \frac{\partial x^\alpha}{\partial y^i} \frac{\partial y^j}{\partial x^\beta} b_j^i.$$

Remark 2.1. In the above formula, the interpretation of say $\frac{\partial x^\alpha}{\partial y^i}$ is that for each α , x^α is a function on U and for each i , ∂_{y^i} is a vector field. Thus, $\partial_{y^i}(x^\alpha) (= \frac{\partial x^\alpha}{\partial y^i})$ is well defined.

Proof. Just use the previous formula giving A in terms of its components and apply the change of coordinate rules to each of $\partial_{x^\alpha}, dx^\alpha$. \square

Remark 2.2. Let $(U, \xi = (x^\alpha))$ be a coordinate system and let $X^\alpha(x_1, \dots, x_n)$ be smooth functions defined on $\xi(U) \subset \mathbb{R}^n$. Then $X := X^\alpha \circ \xi \cdot \partial_{x^\alpha}$ defines a vector field on U . Now let $\xi' = (y^\alpha)$ be another coordinate system on U and let $(X')^\alpha$ be another set of functions defined on $\xi'(U)$. Then $X' := (X')^\alpha \circ \xi' \cdot \partial_{y^\alpha} = X$ if and only if the usual rule for changes of coordinates applies to X^α and X'^α . Thus, if for any coordinate system ξ on U , we can construct components X^α (defined on $\xi(U)$), we can then define a unique tensor field provided that the constructed components satisfy the usual rule for coordinate transformations.

Remark 2.3. The components of tensor need not be only with respect to coordinate induced basis. More precisely, a basis of (local) vector fields is a set of (local) vector fields X_i such that at each p , the $X_i(p)$ form a basis of $T_p M$. One can similarly introduce a basis of one-forms (which may or may not correspond to the dual basis of the basis of vector fields), and finally decompose all tensors with respect to these basis.

2.15 Contraction

The contraction is an operation on tensor fields taking (r, s) types to $(r-1, s-1)$ types. It can be viewed as a trace operation. More precisely,

Lemma 2.7. There is a unique $\mathcal{F}(M)$ -linear function C defined on tensor of type $(1, 1)$ such that $C(X \otimes \theta) = \theta(X), \forall (X, \theta) \in \Gamma(M) \times \Lambda(M)$.

Proof. In coordinates, given A_α^β the components of a tensor of type $(1, 1)$, check that $C(A) = A_\alpha^\alpha$ necessarily and that this defines C properly. \square

Exercise 2.9. Define a contraction C_i^j over the indices i and j for tensors of type (r, s) and $1 \leq i \leq r, 1 \leq j \leq s$.

2.16 Pull-back

Let $\phi: M \rightarrow N$ be a smooth map. Let $\omega \in \Lambda^1(N)$. Then, we can define a one-form on M by

$$\forall p \in M, v \in T_p M, \quad \phi^*(\omega)(v) = \omega_{\phi(p)}(d\phi_p v).$$

Similarly, if A is a $(0, s)$ tensor field, then define for each p on M

$$\phi^*(A)(v_1, \dots, v_s) = A_{\phi(p)}(d\phi_p v_1, \dots, d\phi_p v_s).$$

2.17 Vector bundles

The tangent bundle can be viewed as an example of a *vector bundle*, defined as follows.

Definition 2.19. A k -vector bundle (E, π) over a manifold M consists in a manifold E together with a smooth map

$$\pi: E \rightarrow M,$$

such that

1. for all $p \in M$, each fiber $\pi^{-1}(p)$ is a k -dimensional vector space.
2. for each $p \in M$, there is a neighborhood \mathcal{U} of p and a diffeomorphism

$$\phi: \mathcal{U} \times \mathbb{R}^k \rightarrow \pi^{-1}(\mathcal{U}),$$

such that for each $q \in \mathcal{U}$, the map $v \rightarrow \phi(q, v)$ is a linear isomorphism from \mathbb{R}^k to $\pi^{-1}(q)$. (In particular $\pi \circ \phi(q, v) = q$.)

If M is a manifold, $p \in M$ and $\xi = (\mathcal{U}, x^\alpha)$ is a local coordinate system near p , then for any point $(q, v_q) \in TM$, we have the decomposition $v_q = V^\alpha(q)[\partial_{x^\alpha}]_q$. We can then consider

$$\begin{aligned} \phi: \mathcal{U} \times \mathbb{R}^n &\rightarrow \pi^{-1}(\mathcal{U}) \\ (q, V) &\rightarrow (q, V^\alpha[\partial_{x^\alpha}]_q) \end{aligned}$$

which shows that TM is an example of a vector bundle.

A section of a vector bundle E is a map $X: M \rightarrow E$ such that $\pi \circ X = Id_M$.

3 Semi-Riemannian manifolds

3.1 Semi-Riemannian metrics

Definition 3.1. A scalar product g on a finite dimensional real vector space V is a non-degenerate symmetric bilinear form on V .

Non-degenerate means that, for each $v \neq 0 \in V$, $\exists x \in V$, such that $g(v, x) \neq 0$. Equivalently, any matrix representing g is invertible.

We say that v and w are orthogonal, denoted $v \perp w$, if $g(v, w) = 0$.

If W is a subspace of V , we define as usual W^\perp and we have

Lemma 3.1.

$$\dim W + \dim W^\perp = \dim V, \quad (3)$$

$$(W^\perp)^\perp = W. \quad (4)$$

Exercise 3.1. Prove the lemma. Does (4) still hold for infinite dimensional vector space V equipped with non-degenerate symmetric bilinear form?

Because $g(v, v)$ can be negative, the *norm* $|v|$ of a vector is defined to be

$$|v| = |g(v, v)|^{1/2},$$

(but it does not define a norm in the sense of normed vector space unless g is positive definite). A *unit* vector is a vector of norm 1, i.e. such that $g(v, v) = \pm 1$. As usual a set of mutually orthogonal unit vectors is said to be *orthonormal*.

It follows from the standard theory of quadratic forms that

Lemma 3.2. A scalar product on $V (\neq \emptyset)$ has an orthonormal basis.

If (e_i) is an orthonormal basis, the number of i such that $g(e_i, e_i) < 0$ is independent of the choice of orthonormal basis (cf Sylvester's law of inertia on the signature of quadratic forms).

The signature of g is typically written $- + \dots +$ or $+ \dots + -$, ... (each minus sign correspond to an i such that $g(e_i, e_i) < 0$ etc..)

Definition 3.2. Let g be a $(0, 2)$ tensor field. g is said to be symmetric if for any vector fields X and Y , $g(X, Y) = g(Y, X)$.

If $g_{\alpha\beta}$ are the components of g in some local basis of vector fields, then g is symmetric if and only if $g_{\alpha\beta} = g_{\beta\alpha}$.

If g is a $(0, 2)$ symmetric tensor field, then at each p , g_p is a symmetric bilinear form on $T_p M$. We say that g is non-degenerate if g_p is non-degenerate at each p . Moreover, at each p , we can compute the signature of g_p .

We can thus define a metric tensor as follows.

Definition 3.3. A metric tensor g on M is a symmetric non-degenerate $(0, 2)$ tensor field on M of constant signature. (M, g) is then called a semi-Riemannian manifold. If g is positive definite, then (M, g) is called a Riemannian manifold, if sign $g = - + \dots +$, then (M, g) is called a Lorentzian manifold.

The name pseudo-Riemannian manifold is also used instead of semi-Riemannian manifold (not to be confused with sub-Riemannian).

Let (M, g) be a semi-Riemannian manifold and (x^α) a local coordinate system. Since g is non-degenerate, its matrix components $g_{\alpha\beta}$ is invertible and will be denoted $g^{\alpha\beta}$.

Proposition 3.1. The components $g^{\alpha\beta}$ defines a $(2, 0)$ tensor field, called the inverse metric tensor.

Proof. Let (x^α) and (y^i) be two coordinate systems defined on a common open set. Let $g_{\alpha\beta}$ be the components of g in the first system, g'_{ij} be the components of g in the second. Similarly, we write $g^{\alpha\beta}$ and $(g')^{ij}$ for the components of the inverse.

We have, by definition of the inverse,

$$g^{\alpha\beta} g_{\beta\gamma} = \delta_\gamma^\alpha, \quad (g')^{ik} (g')_{kj} = \delta_j^i,$$

while by the transformation rule

$$(g')_{kj} = \frac{\partial x^\alpha}{\partial y^k} \frac{\partial x^\beta}{\partial y^j} g_{\alpha\beta}.$$

Thus, we have

$$(g')^{ik} \frac{\partial x^\alpha}{\partial y^k} \frac{\partial x^\beta}{\partial y^j} g_{\alpha\beta} = \delta_j^i$$

As in usual differential calculus, the Jacobian matrix $\left(\frac{\partial x^\beta}{\partial y^j}\right)$ has inverse $\left(\frac{\partial y^j}{\partial x^\beta}\right)$. Thus, we have

$$(g')^{ik} \frac{\partial x^\alpha}{\partial y^k} \frac{\partial x^\beta}{\partial y^j} g_{\alpha\beta} \frac{\partial y^j}{\partial x^\gamma} = \delta_j^i \frac{\partial y^j}{\partial x^\gamma}$$

i.e.

$$(g')^{ik} \frac{\partial x^\alpha}{\partial y^k} g_{\alpha\gamma} = \frac{\partial y^i}{\partial x^\gamma}$$

Multiply now by the inverse metric $(g^{\gamma\rho})$ to obtain

$$(g')^{ik} \frac{\partial x^\rho}{\partial y^k} = \frac{\partial y^i}{\partial x^\gamma} g^{\gamma\rho}$$

Multiplying again by the inverse Jacobian matrix, we obtain

$$(g')^{ik} = \frac{\partial y^i}{\partial x^\gamma} \frac{\partial y^k}{\partial x^\rho} g^{\gamma\rho},$$

i.e. $g^{\alpha\beta}$ transforms as a $(2, 0)$ tensor field. In view of Remark 2.2, it follows that there exists a unique $(2, 0)$ tensor field of components given by $g^{\alpha\beta}$. \square

If X is a vector field, then the map

$$Y \in \Gamma(M) \rightarrow g(X, Y)$$

is $\mathcal{F}(M)$ -linear and therefore defines a 1-form, denoted³ $\flat X$. In components, with $X = X^\alpha \partial_{x^\alpha}$, we will write

$$\flat X = X^\alpha g_{\alpha\beta} dx^\beta := X_\beta dx^\beta,$$

i.e. the components of the corresponding 1-form are denoted by the same letter but with the indices down.

Lemma 3.3. *The map $\flat : \Gamma(M) \rightarrow \Lambda^1(M)$ is an $\mathcal{F}(M)$ -linear isomorphism.*

Proof. 1. Injectivity: Note that if $g(V, Y) = g(W, Y)$, for all $Y \in \Gamma(M)$, then $V = W$ (g is non-degenerate).

³Traditionally called the musical notation.

2. Surjectivity:

Let $\theta \in \Lambda^1(M)$ and consider $(U, (x^\alpha))$ a coordinate system, $g_{\alpha\beta}$ the components of g and $g^{\alpha\beta}$ the components of the inverse matrix of $g_{\alpha\beta}$. The map $\sharp\theta : p \rightarrow g^{\alpha\beta}(p)\theta_\alpha(p)\partial_{x^\beta}$ then defines a vector field on U . Moreover, since $g_{\alpha\beta}$ are the components of a $(2,0)$ tensor field and θ_α the components of a one-form, one easily check that $g^{\alpha\beta}\theta_\alpha$ transforms like a vector field. Thus, in view of Remark 2.2, it follows that $\sharp\theta$ can actually be defined uniquely as a global vector field on M .

Now we have, for all $X \in \Gamma(X)$, $g(\sharp\theta, X) = \theta(X)$ by construction (evaluate everything in the (x^α) coordinate system.)

□

From the above proof, for a one-form ν , $\sharp\nu$ is given by the map

$$\sharp\nu : \theta \in \Lambda(M) \rightarrow g^{-1}(\theta, \nu)$$

and by convention

$$\nu^\alpha := \sharp\nu^\alpha = \nu_\beta g^{\alpha\beta}.$$

The raising and lowering of indices can then be generalized to arbitrary tensors in the natural way.

Definition 3.4. Let (M, g) be a semi-Riemannian manifold.

- Let $X \in T_p M$, then X is said to be spacelike if $g_p(X, X) > 0$ or $X = 0$, timelike if $g_p(X, X) < 0$ and null if $X \neq 0$ and $g_p(X, X) = 0$.
- The set of all null vectors in $T_p M$ is called the null cone at $p \in M$ and the union of all null and timelike vectors is the set of all causal vectors. If α is a smooth curve, then α is said to be timelike (respectively spacelike, null, causal) if all its tangent vectors are timelike (respectively spacelike, null, causal).

Remark 3.1. One of the postulates of General Relativity is that a particle of mass m moves along a timelike curve if $m > 0$ and along a null curve if $m = 0$. This is the geometric analogue of the postulate that massive bodies move with a speed strictly less than the speed of light and that massless bodies move with the speed of light.

Exercise 3.2. Extend the above definition from curves to submanifolds of M (What is a spacelike hypersurface?).

The line element

The metric g can be written locally as

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

By polarization, g can of course be reconstructed from the map $q(v) : v \in T_p M \rightarrow g(v, v)$. The map q is called the *line element* of the metric g and is typically denoted ds^2 with

$$ds^2 = g_{\alpha\beta} dx^\alpha dx^\beta.$$

See for instance the explanations p56 in [O’N83].

3.2 Examples and basic constructions

3.2.1 Riemannian case

The simplest example of a Riemannian manifold is (\mathbb{R}^n, δ) , where δ is the Euclidean metric, of components δ_{ij} in the canonical coordinates of \mathbb{R}^n .

Submanifolds of a Riemannian manifold

Lemma 3.4. *Let Σ be a submanifold of a Riemannian manifold (M, g) and $j : \Sigma \rightarrow M$ the inclusion map. Then, the pull-back of g , $j^*(g)$ is a Riemannian metric on Σ .*

Proof. It follows from the definition that $j^*(g)$ is a $(0,2)$ (symmetric) tensor. Moreover, given vectors X and Y in $T_p\Sigma$, $p \in \Sigma$, we have by definition

$$j^*(g)_p(X, Y) = g_p(dj_p(X), dj_p(Y)).$$

Since Σ is a submanifold, the differential dj_p is injective at each p , thus $dj_p(X) = 0$ if and only if $X = 0$. It then follows easily that $j^*(g)_p$ is positive definite. \square

Exercise 3.3. *Consider the sphere $\mathbb{S}^2 = \{x^2 + y^2 + z^2 = 1\}$ in \mathbb{R}^3 . Recall that \mathbb{S}^2 endowed with the induced topology is a submanifold of \mathbb{R}^3 . We consider on \mathbb{S}^2 a local chart ξ given by the coordinates (θ, ϕ) defined on $(0, \pi) \times (0, 2\pi)$ by the inverse of the map*

$$(\theta, \phi) \rightarrow (\cos\theta, \sin\theta \cos\phi, \sin\theta \sin\phi) \in \mathbb{S}^2.$$

We consider on \mathbb{R}^3 spherical coordinates $\chi = (r, \theta, \phi)$ defined on $(0, +\infty) \times (0, \pi) \times (0, 2\pi)$ by the inverse of the map (here by a small abuse of language, we are using again the label ϕ and θ)

$$(r, \theta, \phi) \rightarrow (r \cos\theta, r \sin\theta \cos\phi, r \sin\theta \sin\phi).$$

- Check that in the (r, θ, ϕ) coordinate system

$$\delta = dr \otimes dr + r^2 (\sin^2(\theta) d\phi \otimes d\phi + d\theta \otimes d\theta).$$

and that the map $\chi \circ j \circ \xi^{-1}$ is given by

$$\chi \circ j \circ \xi^{-1} : (\theta, \phi) \rightarrow (1, \theta, \phi).$$

- Check that the induced metric on \mathbb{S}^2 is given by

$$\sigma_{\mathbb{S}^2} = \sin^2(\theta) d\phi \otimes d\phi + d\theta \otimes d\theta.$$

3.2.2 The Lorentzian case

Minkowski space

Minkowski space is the Lorentzian manifold (\mathbb{R}^{n+1}, η) , with η the Lorentzian metric such that if $(x^\alpha) = (t = x^0, x^i)$ are canonical coordinates on \mathbb{R}^{n+1} , then the ∂_{x^α} are orthonormal with⁴ $g_{00} = -1$. Minkowski space plays the role of the Euclidean space

⁴Here are considering units (i.e. a specific choice of inertial frame) such that the speed of light in the vacuum is 1. In general, we have $g_{00} = -c^2$ in an inertial frame.

in Lorentzian geometry, in the sense that it is the trivial example of a Lorentzian manifold.

Contrary to the Riemannian case, if N is a submanifold of a Lorentzian manifold (M, g) (such as the Minkowski space) with inclusion map j , the pull-back tensor $j^*(g)$ does not need to define a Lorentzian metric, in fact, it does not need to define a metric at all.

Exercise 3.4. In (\mathbb{R}^{n+1}, η) with coordinates (t, x, y, z) , consider the hypersurfaces given by the equations $t = 0$, $x = 0$, $t^2 = x^2 + y^2 + z^2 + 1$ with $t \geq \sqrt{x^2 + y^2 + z^2 + 1}$, and $t = x$ and compute the pull back of the metric in each of these cases. In which cases do we have a Riemannian or Lorentzian metric?

By a small abuse of language, given P a submanifold of a semi-Riemannian manifold (M, g) with inclusion map p , the pull-back tensor $j^*(g)$ will sometimes be called induced metric, even when it does not define a metric.

3.2.3 Products

Let M and N be semi-Riemannian manifold with metric g_M and g_N respectively. Consider the product manifold $M \times N$ and let $\pi : M \times N$ and $\sigma : M \times N$ be the projection on the first and second components. Then,

$$g := \pi^*(g_M) + \sigma^*(g_N)$$

is a metric on g .

For instance, if (N, g_N) is a Riemannian manifold, then we can construct a Lorentzian manifold on the product $\mathbb{R}_t \times N$ by considering the metric

$$g = \pi^*(-dt \otimes dt) + \sigma^*(g_N).$$

(The index t on \mathbb{R}_t simply expresses that t is a global coordinate on \mathbb{R} .) Often, we'll drop the pull back of the projections in the above formula and simply write

$$g = -dt \otimes dt + g_N.$$

3.2.4 Uniformly Lorentzian metrics on \mathbb{R}^{n+1}

In this section, we consider \mathbb{R}^{n+1} with canonical coordinates (x^0, x^i) . Latin indices $1 \leq i, j \leq n$ will always refer to the spatial directions, so that for instance, if T is a tensor field, T_{ij} denotes all the components of T associated to the ∂_{x^i} , $1 \leq i \leq n$. Greek indices will be used to denote all the components (including the ones associated to ∂_{x^0}) of tensors.

In elliptic pde, a standard definition is that of a *uniformly elliptic symbol*. This would correspond geometrically⁵ to a Riemannian metric h and a global chart⁶ such that, for some $a, b > 0$, we have

$$aX_i X^i \leq h_p(X, X) \leq bX_i X^i,$$

for all tangent vectors X and all points p .

⁵This correspondance only make sense for second order elliptic pde.

⁶One can easily remove this constraint provided we have another reference metric.

The aim of this section is to introduce a (naive) analogous notion for Lorentzian manifolds. We will also provide some formulae to compute or estimate the components $g^{\alpha\beta}$ in terms of those of $g_{\alpha\beta}$.

We start with some linear algebra.

Lemma 3.5. *Let $g \in M_{n+1}(\mathbb{R})$ be a symmetric matrix such that $g_{00} < 0$ and $h = (g_{ij})$, the submatrix obtained by removing the first row and the first column, is positive definite. Then g has sign $- + \dots +$.*

Remark 3.2. *The point of the lemma is that we do not need to know the values of g on the whole first row and the whole first column to determine the signature of g in the case, only the sign of g_{00} .*

Proof. Let O be an orthogonal matrix diagonalizing h and let $k = M_O^t g M_O$, where

$$M_O = \begin{pmatrix} 1 & 0 \\ 0 & O \end{pmatrix}.$$

Consider the submatrix $l = (k_{ij})$ obtained similarly by removing the first row and the first column of k . Then, l is diagonal and

$$l = O^t h O = \text{diag}(\lambda_1, \dots, \lambda_n),$$

with $\lambda_i > 0$ for all i . Furthermore $k_{00} = g_{00} < 0$ and the eigenvalues of g and k coincides. We compute the determinant of $k - \lambda I$ to get

$$p(\lambda) = \det(k - \lambda I) = \left(k_{00} - \lambda - \frac{k_{01}^2}{\lambda_1 - \lambda} - \dots - \frac{k_{0n}^2}{\lambda_n - \lambda} \right) (\lambda_1 - \lambda) \dots (\lambda_n - \lambda).$$

Let us define

$$f(\lambda) = k_{00} - \lambda - \frac{k_{01}^2}{\lambda_1 - \lambda} - \dots - \frac{k_{0n}^2}{\lambda_n - \lambda}$$

and its derivative

$$f'(\lambda) = -1 - \frac{h_{01}^2}{(\lambda_1 - \lambda)^2} - \dots - \frac{h_{0n}^2}{(\lambda_n - \lambda)^2}.$$

Let λ_m be the smallest of the λ_i . On $(-\infty, \lambda_m)$, f is smooth and $f' < 0$. Moreover, $f(-\infty) = +\infty$ and $f(0) < 0$. Thus, there is a unique negative value of $\lambda \in (-\infty, 0)$, say λ_0 for which $f(\lambda) = 0$ which has to be an eigenvalue of g . Moreover, $f'(\lambda_0) \neq 0$, i.e. it is a root with multiplicity one. Since $f' < 0$ on $(-\infty, \lambda_m)$, there can be no other eigenvalue in this interval. It follows that the remaining eigenvalues must be strictly positive.

Alternatively, from the assumptions, g restricted to $\mathbb{R}(1, \dots, 0)$ is definite negative while g restricted to the span of the $e_i = (0, \dots, 1, \dots, 0)$, $1 \leq i \leq n$ is definite positive. So it follows from basic linear algebra that g has one negative eigenvalue and n positive ones. \square

Definition 3.5. *Let g be a smooth symmetric $(0,2)$ tensor field on \mathbb{R}^{n+1} . We say that g is uniformly Lorentzian in the coordinate system (x^0, \dots, x^n) if there exists constants $a, b > 0$ such that $g_{00} < -a$, $(g_{ij}) > b$ and the components of g are uniformly bounded.*

Of course, it follows from the previous lemma that such a uniform Lorentzian metric is in particular Lorentzian.

Given a uniformly Lorentzian metric, it will be useful (for instance, when we want to prove estimates for solutions of wave equations) to have estimates controlling the inverse metric as well.

Let g be a matrix as in Lemma 3.5. Let h be the submatrix of g given in components by g_{ij} where we have removed the first row and the first column. From Lemma 3.5, we know in particular that g is invertible. Let $g^{\mu\nu}$ be the components of the inverse matrix of g and let H to denote the submatrix given in components by g^{ij} obtained by removing the first row and the first column (note that in general $H \neq h^{-1}$).

Finally, we will denote the vector of components g_{0i} by $s(g)$ (s is sometimes called the shift vector).

Lemma 3.6. *Let g be a matrix as in Lemma 3.5, i.e. $g_{00} < 0$ and $h > 0$, where h is the 3×3 submatrix defined by $h_{ij} = g_{ij}$, $1 \leq i, j \leq n$. Then,*

- $$g^{00} = \frac{1}{g_{00} - d^2},$$

where $d^2 = (h)^{-1}(s(g), s(g))$.

- $H = (g^{ij})$ is positive definite, with

$$\frac{g_{00}}{g_{00} - d^2} h^{-1} \leq H \leq h^{-1}.$$

- The new shift vector is given by

$$s(g^{-1}) = \frac{1}{d^2 - g_{00}} h^{-1} s(g).$$

Proof. Let A be the square root of h^{-1} , i.e. the positive definite symmetric matrix such that $A^2 = h^{-1}$. Then $A^t h A = Id$. Let $k = M_A^t g M_A$, where

$$M_A = \begin{pmatrix} 1 & 0 \\ 0 & A \end{pmatrix}.$$

Then, $k_{00} = g_{00}$, the submatrix (k_{ij}) obtained by erasing the first row and the first column is simply the identity matrix and $s(k)$, the new shift vector, is given by $A^t s(g) = A s(g)$. Let B be an orthogonal matrix such that $B^t A^t s(g) = |A^t s(g)| e_1$, where $|\cdot|$ denotes the Euclidean norm and $e_1 = (1, \dots, 0)$ is the first vector of the canonical basis of \mathbb{R}^n . Note that

$$\begin{aligned} |A^t s(g)|^2 &= [(A^t s(g))^t \cdot A^t s(g)]^{1/2} \\ &= [s(g) A A^t s(g)]^{1/2} = [s(g) A^2 s(g)]^{1/2} = [h^{-1}(s(g), s(g))]^{1/2} = d. \end{aligned}$$

Consider $\rho = M_B^t k M_B$, where M_B is defined similarly to M_A . Then, $\rho_{00} = g_{00}$, the submatrix (ρ_{ij}) obtained by removing the first column and the first row is simply the identity and $s(\rho) = d e_1$. Note that the inverse of the 2×2 submatrix with components $\rho_{\mu\nu}$, $0 \leq \mu, \nu \leq 1$ is given by

$$\frac{1}{g_{00} - d^2} \begin{pmatrix} 1 & -d \\ -d & g_{00} \end{pmatrix}.$$

Since $g^{-1} = M_A M_B \rho^{-1} M_B^t M_A^t$ and the matrices M_A and M_B preserves the 00 components of a matrix, we obtain the first claim.

Moreover, let $H_\rho = (\rho^{ij})$ be the submatrix obtained by removing the first column and the first row of ρ^{-1} . Then,

$$H = ABH_\rho B^t A^t.$$

We have, for any $w \neq 0$, denoting $\langle \cdot, \cdot \rangle$ the usual scalar product on \mathbb{R}^n ,

$$\begin{aligned} \frac{H(w, w)}{h^{-1}(w, w)} &= \frac{\langle ABH_\rho B^t A^t w, w \rangle}{\langle Aw, Aw \rangle} \\ &= \frac{\langle H_\rho B^t A^t w, B^t A^t w \rangle}{\langle B^t A^t w, B^t A^t w \rangle}, \end{aligned}$$

using that B is orthogonal and A symmetric. The second claim then follows since H_ρ is a diagonal matrix with all diagonal element 1 except for the first which is given by $\frac{g_{00}}{g_{00}-d^2} \leq 1$. We leave the last claim as an exercise. \square

3.2.5 A first look at the Schwarzschild spacetime

The Schwarzschild spacetime is one of the most important example of Lorentzian manifold. It was discovered in 1915 by Schwarzschild as an explicit solution to the Einstein equations. It is not just one Lorentzian manifold, but a one parameter family of Lorentzian manifolds, indexed by a positive⁷ number $m > 0$ (referred to as the mass). Let thus $m > 0$ and consider the product manifold

$$M_{ext} = \mathbb{R}_t \times (2m, +\infty) \times \mathbb{S}^2.$$

The projection on the second variable will be denoted by r . r is just a smooth function on our manifold taking values in $(2m, +\infty)$. On M_{ext} , we consider the following Lorentzian metric

$$g_{ext} = -\left(1 - \frac{2m}{r}\right) dt \otimes dt + \left(1 - \frac{2m}{r}\right)^{-1} dr \otimes dr + r^2 \sigma_{\mathbb{S}^2},$$

where $\sigma_{\mathbb{S}^2}$ denotes the pull back on M_{ext} of the usual round metric on \mathbb{S}^2 . On M_{ext} , we can consider local coordinates of the form (t, r, θ, ϕ) . Note that the above expression for the metric g becomes singular as $r \rightarrow 2m$.

We consider a change of coordinates of the form $(t, r, \theta, \phi) \rightarrow (t + f(r), r, \theta, \phi)$, where f is a function which will be determined. Our aim will be to choose f such that in the new coordinate system, the metric will be regular at $r = 2m$.

Let us thus define $v(t, r) = t + f(r)$. We have $dv = dt + f'(r)dr$ and thus $dt = dv - f' dr$. The metric in the (v, r, θ, ϕ) coordinates can therefore be written as

$$\begin{aligned} g_{ext} &= -\left(1 - \frac{2m}{r}\right) (dv - f' dr) \otimes (dv - f' dr) + \left(1 - \frac{2m}{r}\right)^{-1} dr \otimes dr + r^2 \sigma_{\mathbb{S}^2} \\ &= -\left(1 - \frac{2m}{r}\right) dv \otimes dv + \left(1 - \frac{2m}{r}\right) f' (dv \otimes dr + dr \otimes dv) \\ &\quad - \left(1 - \frac{2m}{r}\right) (f')^2 dr \otimes dr + \left(1 - \frac{2m}{r}\right)^{-1} dr \otimes dr + r^2 \sigma_{\mathbb{S}^2}. \end{aligned}$$

⁷Schwarzschild spacetimes with $m < 0$ can also be defined, but their physical interpretation is different (naked singularity vs. black hole).

In order to regularize the metric and cancel the terms which becomes singular at $r = 2m$, we see that we need

$$(f'(r))^2 = \left(1 - \frac{2m}{r}\right)^{-2},$$

i.e.

$$f'(r) = \pm \left(1 - \frac{2m}{r}\right)^{-1}.$$

This can be integrated to give us

$$f(r) = \pm r \pm 2m \ln(r - 2M) + C, \quad C \in \mathbb{R}.$$

Let us consider the + sign, fix $C = 0$ and thus define $v(t, r) := t + r + 2m \ln(r - 2m)$. The metric then takes the form

$$g_{ext} = -\left(1 - \frac{2m}{r}\right) dv \otimes dv + dv \otimes dr + dr \otimes dv + r^2 \sigma_{\mathbb{S}^2}. \quad (5)$$

Note that we have only done a change of coordinates. In particular, for $r > 2m$, our metric is still Lorentzian, since this is independant of our choice of coordinates, but in fact the above expression defines a Lorentzian metric for any $r > 0$, since for any $\alpha \in \mathbb{R}$, the matrix $\begin{pmatrix} \alpha & 1 \\ 1 & 0 \end{pmatrix}$ has signature $-+$.

We consider thus another manifold given by $M = \mathbb{R}_v \times (0, +\infty) \times \mathbb{S}^2$. We can endow this manifold with the Lorentzian metric given by the expression found in (5), which we denote by g_M . The map

$$\begin{aligned} \psi : M_{ext} &\rightarrow M \\ (t, r, \omega) &\rightarrow (v(t, r), r, \omega) \end{aligned}$$

then provides a embedding of M_{ext} into M which preserves the metric $\psi^*(g_M) = g_{M_{ext}}$. Note that $\psi(M_{ext}) = \mathbb{R}_v \times (2m, +\infty) \times \mathbb{S}^2$ is strictly smaller than M .

Remark 3.3. *M is still not the largest regular Lorentzian manifold than "contains" M_{ext} . First, we could have defined an extension of M_{ext} by taking the minus sign instead of the plus sign. The two differents extensions can in fact be glued together, corresponding to an exterior region, a black hole region (the one corresponding to the interior of $M \setminus \psi(M_{ext})$) and a white hole region (a similar region obtained by taking the minus sign above). This manifold can in fact still be extended !!*

3.3 The Levi-Civita connection

Given a semi-Riemannian manifold, we are looking for an operation replacing the usual notion of derivative for tensorial objects. We start by axiomatizing the key properties that we would like.

Definition 3.6. *A connection D on a smooth manifold M is map of the form*

$$\begin{aligned} D : \Gamma(M) \times \Gamma(M) &\rightarrow \Gamma(M) \\ (V, W) &\rightarrow D_V W \end{aligned}$$

such that $D_V W$ is $\mathcal{F}(M)$ -linear in V , \mathbb{R} -linear in W and the Leibniz rule is verified

$$D_V(fW) = V(f).W + fD_V W,$$

for all $f \in \mathcal{F}(M)$.

$D_V W$ is called the *covariant derivative* of W with respect to V (for the connection D).

Let D be a connection on M . Let $p \in M$ and consider $v \in T_p M$. Let V be any vector field such that $V_p = v$. Then, for any $W \in \Gamma(M)$, $(D_V W)(p)$ is independent of the choice of V (it depends only on v). Thus, for each p , we can define map (still denoted D)

$$\begin{aligned} D: T_p M \times \Gamma(M) &\rightarrow T_p M \\ (v, W) &\rightarrow D_v W. \end{aligned}$$

Given a semi-Riemannian manifold, there can be plenty of connections, but there is a unique one that satisfies the two extra conditions written below⁸

Theorem 3.1. *Let (M, g) be a semi-Riemannian manifold. Then, there exists a unique connection D such that*

1. *D is torsion free: $[V, W] = D_V W - D_W V$,*
2. *$X(g(V, W)) = g(D_X V, W) + g(V, D_X W)$.*

D is called the Levi-Civita connection of (M, g) and is characterized by the Koszul formula

$$2g(D_V W, X) = Vg(W, X) + Wg(X, V) - Xg(V, W) - g(V, [W, X]) + g(W, [X, V]) + g(X, [V, W]).$$

Remark 3.4. *The first condition can be interpreted as a twist or torsion free condition. Without it, geodesics (which we are going to define soon) would not completely determined the connection. The second condition means essentially that $Dg = 0$, i.e. the metric is constant with respect to this notion of derivative.*

Proof. Let D be a connection satisfying the above two conditions. Inserting the conditions on the RHS of the Koszul formula, it follows that it must hold. Then the non-degeneracy of the metric implies that $D_V W$ is uniquely determined. Reciprocally, let $F(V, W, X)$ be given by the RHS of the Koszul formula. Let $V, W \in \Gamma(M)$ and check that $X \rightarrow F(V, W, X)$ is $\mathcal{F}(M)$ -linear. Thus, the formula defines a one-form and we denote its image by \sharp as $2D_V W$ (i.e. $2g(D_V W, X) = F(V, W, X)$). Then the Koszul formula holds and from it we can check that all the conditions hold. For instance,

$$g(X, [V, W]) = 2g(D_V W, X) - Vg(W, X) - Wg(X, V) + Xg(V, W) + g(V, [W, X]) + g(W, [V, X])$$

and similarly (exchanging V and W)

$$g(X, [V, W]) = -2g(D_W V, X) + Vg(W, X) + Wg(X, V) - Xg(V, W) - g(V, [W, X]) - g(W, [V, X])$$

Adding the equations give the torsion free property.

⁸To motivate, naively and quickly, the definition below, one can imagine a submanifold P of \mathbb{R}^n with its induced metric. Consider a point $p \in P$ and an orthonormal frame in $T_p P$ at that point. Consider moving that frame along a curve on P passing through p such that the moving frame stays parallel to the original one. Then, we would like that our notion of derivative to be such that the derivative of the moving frame in the direction of the tangent vector of the curve be zero. From this constraint, one can reconstruct the definition given below.

Let us also check that with this definition, D is indeed $\mathcal{F}(M)$ -linear in the first variable. Let $f \in \mathcal{F}(M)$. Then,

$$\begin{aligned}
2g(D_{fV}W, X) &= fVg(W, X) + Wg(X, fV) - Xg(fV, W) - g(fV, [W, X]) + g(W, [X, fV]) + g(X, [fV, W]) \\
&= fVg(W, X) + W(fg(X, V)) - X(fg(V, W)) - fg(V, [W, X]) \\
&\quad + g(W, X(f)V + f[X, V]) + g(X, f[V, W] - W(f)V) \\
&= f(Vg(W, X) + Wg(X, V) - Xg(V, W) - g(V, [W, X]) + g(W, [X, V]) + g(X, [V, W])) \\
&= f2g(D_VW, X).
\end{aligned}$$

□

Exercise 3.5. Check the other conditions and finish the proof.

From the Koszul formula, it follows that, for any p , $D_VW(p)$ only depends on the value of W in a neighborhood of p (since it only depends on $W(p)$ and its derivative at p). Thus, given any open set $U \subset M$, we can consider D defined on $\Gamma(U) \times \Gamma(U)$.

Definition 3.7. Let (x^α) be a coordinate system. The Christoffel symbols are the real valued functions $\Gamma_{\beta\gamma}^\alpha$ such that

$$D_{\partial_{x^\beta}}\partial_{x^\gamma} = \Gamma_{\beta\gamma}^\alpha\partial_{x^\alpha}.$$

Lemma 3.7. We have $\Gamma_{\beta\gamma}^\alpha = \frac{1}{2}g^{\alpha\rho}(\partial_{x^\beta}g_{\gamma\rho} + \partial_{x^\gamma}g_{\beta\rho} - \partial_{x^\rho}g_{\beta\gamma})$.

In particular, $\Gamma_{\beta\gamma}^\alpha = \Gamma_{\gamma\beta}^\alpha$.

Proof. We have,

$$g(D_{\partial_{x^\beta}}\partial_{x^\gamma}, \partial_{x^\rho}) = \Gamma_{\beta\gamma}^\alpha g_{\rho\alpha}.$$

On the other hand, using the Koszul formula, all the commutators vanish and

$$2g(D_{\partial_{x^\beta}}\partial_{x^\gamma}, \partial_{x^\rho}) = \partial_{x^\beta}g_{\gamma\rho} + \partial_{x^\gamma}g_{\beta\rho} - \partial_{x^\rho}g_{\beta\gamma}.$$

Conclude by inverting $g^{\rho\alpha}$ in the first equation. □

Exercise 3.6. 1. Let $(\mathbb{S}^2, \sigma_{\mathbb{S}^2})$ be the usual 2-sphere endowed with the round metric. Compute its Christoffel symbols in the (θ, ϕ) coordinate system.

2. Let (\mathbb{R}^{3+1}, η) be the Minkowski space and consider spherical coordinates so that the line element takes the form

$$ds^2 = -dt^2 + dr^2 + r^2 d\sigma_{\mathbb{S}^2}.$$

Compute its Christoffel symbols.

Of course, for the Euclidean or the Minkowski space in Cartesian coordinates⁹, the covariant derivatives in the direction of ∂_{x^α} just reduces to partial derivatives and the Christoffel symbols vanish.

We now seek to extend our connection as a derivation for arbitrary tensors. First a definition.

⁹We call Cartesian coordinates, global coordinates on \mathbb{R}^n or \mathbb{R}^{n+1} such that the metric components are given by the matrices $diag(1, 1, \dots, 1)$ in the Euclidean case and $diag(-1, 1, \dots, 1)$ in the Minkowskian case.

Definition 3.8. A tensor derivation D is a set of maps $D(= D_s^r) : \mathcal{T}_s^r(M) \rightarrow \mathcal{T}_s^r(M)$, where $\mathcal{T}_s^r(M)$ is the set of tensor fields of type (r, s) , such that,

1. for a tensor field $A \otimes B$, $D(A \otimes B) = DA \otimes B + A \otimes DB$,
2. $D(CA) = C(DA)$ for any contraction C .

We have easily.

Proposition 3.2. Given a vector field V and an \mathbb{R} -linear map $\delta : \Gamma(M) \rightarrow \Gamma(M)$ such that

$$\delta(fX) = V(f)X + f\delta(X),$$

there exists a unique tensor derivation D on M such that $D_0^0 = V$ and $D_0^1 = \delta$.

Thus, the Levi-Civita connection extends naturally to arbitrary tensor fields.

Proof. D_0^0 (derivation of functions) and D_0^1 derivation of vector fields are given. Assume that such a general D exists. Let θ be a one-form and X a vector field. Then, $\theta \otimes X$ is a $(1, 1)$ tensor. From the contraction rule of a derivation, we must have

$$D(C(\theta \otimes X)) = C(D(\theta \otimes X)).$$

Using the rule for the derivation of a tensor product,

$$D(\theta \otimes X) = D\theta \otimes X + \theta \otimes DX.$$

Using the definition of the contraction, this gives

$$D(\theta(X)) = D\theta(X) + \theta(DX)$$

and thus

$$D\theta(X) = V(\theta(X)) - \theta(\delta X),$$

which determines uniquely $D\theta$. Moreover, one easily check that the above expression for $D\theta$ is $\mathcal{F}(M)$ -linear and thus defines a one-form, since for any function f

$$\begin{aligned} V(\theta(fX)) - \theta(\delta(fX)) &= V(f\theta(X)) - \theta(V(f)X + f\delta(X)) \\ &= V(f)\theta(X) + fV(\theta(X)) - V(f)\theta(X) - f\theta(\delta(X)) \\ &= f(V(\theta(X)) - \theta(\delta(X))). \end{aligned}$$

For an arbitrary tensor A of type (r, s) , we have similarly,

$$\begin{aligned} DA(\theta^1, \dots, \theta^r, X_1, \dots, X_s) &= V(A(\theta^1, \dots, \theta^r, X_1, \dots, X_s)) - \sum_{i=1}^r A(\theta^1, \dots, D\theta^i, \dots, \theta^r, X_1, \dots, X_s) \\ &\quad - \sum_{j=1}^s A(\theta^1, \dots, \theta^r, X_1, \dots, \delta X_j, \dots, X_s). \end{aligned}$$

Thus D is uniquely determined. One then needs to check that the above formula indeed defines a tensor derivation. The fact that D distributes on tensor products is easy from the above formula. For the contraction property, let us prove that D_1^1 commutes with C . We have trivially that D commutes with C for tensor product $\theta \otimes X$. The result then follows since any $(1, 1)$ tensor can be written as a sum of such terms. \square

Exercise 3.7. 1. Recall that given two vector fields $X, Y \in \Gamma(M)$, $[X, Y]$ is a vector field. Show that it defines a tensor derivation called the Lie derivative (notation $\mathcal{L}_X Y = [X, Y]$).

2. Let (x^α) be a coordinate system. Compute $\mathcal{L}_{\partial_\alpha} g_{\beta\gamma} := (\mathcal{L}_{\partial_\alpha} g)_{\beta\gamma}$.

3. Deduce that \mathcal{L} can never coincide with the tensor derivation associated with the Levi-Civita connection of a metric. (Of course, this also follows from the fact that $D_V W$ is $\mathcal{F}(M)$ -linear in V).

4. Let X and Y be two vector fields and $p \in M$. Let ϕ_t denotes a local flow of X near $p \in M$. Prove that

$$\lim_{t \rightarrow 0} \frac{1}{t} (d\phi_{-t}(Y_{\phi_t(p)}) - Y_p)_{t=0} = [X, Y]_p. \quad (6)$$

Definition 3.9. Let T be any tensor field of type (r, s) . Then we define DT , the covariant differential of T , as the $(r, s+1)$ tensor field

$$(\theta_1, \dots, \theta_r, X_1, \dots, X_s, V) \rightarrow (D_V T)(\theta_1, \dots, \theta_r, X_1, \dots, X_s).$$

Notations:

Let (x^α) be a coordinate system and T a tensor field of type (r, s) . Then the covariant derivative of T in the direction of α is often denoted D_α instead of D_{∂_α} .

Moreover, the component of the covariant derivative $D_\alpha T$ are sometimes written as

$$T_{j_1, \dots, j_s; \alpha}^{i_1, \dots, i_r}.$$

The ";" should be compared with the usual "," used for instance often in PDEs notations to denote partial differentiations.

Exercise 3.8. Prove that $Dg = 0$.

3.4 Isometries

Definition 3.10. Let (M, g) and (N, h) be semi-Riemannian manifolds. An isometry from M to N is a diffeomorphism $\phi: M \rightarrow N$ such that

$$\phi^*(h) = g.$$

This means that

$$h_{\phi(p)}(d\phi_p(v), d\phi_p(w)) = g_p(v, w),$$

for all $v, w \in T_p M$, $p \in M$.

Recall the definition of the flow ϕ of vector field $X \in \Gamma(M)$, cf Definition 2.15. If the flow is complete, then for every $s \in \mathbb{R}$, ϕ_s is a diffeomorphism from M to M . Similarly, if the flow is non-complete, its flow will define diffeomorphisms of the form, $\phi_s: U \rightarrow \phi_s(U)$.

A natural question is: which vector field X have a flow generating isometries ?

Proposition 3.3. Let $X \in \Gamma(M)$ and $A \in T_s^0(M)$ (a tensor field of type $(0, s)$). Consider an open set $D \subset M$ on which the flow of X , ψ_s , is well defined for s sufficiently small. Then, in D ,

$$\mathcal{L}_X A = \lim_{t \rightarrow 0} \frac{1}{t} (\psi_t^*(A) - A),$$

where $\mathcal{L}_X A$ denotes the Lie derivative of A with respect to the vector field X .

Proof. We do the proof in the case of a (0,2) tensor field A and we will use formula (6) from above.

Since \mathcal{L}_X is a tensor derivation, for vector fields V and W ,

$$(\mathcal{L}_X A)(V, W) = X(A(V, W)) - A([X, V], W) - A(V, [X, W]).$$

Let $p \in D$. Then,

$$\begin{aligned} (\psi_t^* A - A)(V_p, W_p) &= A_{\psi_t(p)}(d\psi_t(V_p), d\psi_t(W_p)) - A_p(V_p, W_p) \\ &= A_{\psi_t(p)}(d\psi_t(V_p), d\psi_t(W_p)) - A_{\psi_t(p)}(V_{\psi_t(p)}, W_{\psi_t(p)}) \\ &\quad + A_{\psi_t(p)}(V_{\psi_t(p)}, W_{\psi_t(p)}) - A_p(V_p, W_p) = I + II. \end{aligned}$$

Let α be the integral curve of X starting at p , i.e. $\alpha(t) = \psi_t(p)$, $\alpha(0) = \psi_0(p) = p$.

Then, by definition,

$$\lim_{t \rightarrow 0} \frac{1}{t} II = \frac{d}{dt} (A_\alpha(V_\alpha, W_\alpha))|_{t=0} = X_p(A(V, W)).$$

For I , we write first

$$I = A(d\psi_t(V_p) - V_{\psi_t(p)}, d\psi_t(W_p)) + A(V_{\psi_t(p)}, d\psi_t(W_p) - W_{\psi_t(p)}).$$

Since $\psi_t \circ \psi_{-t} = Id$,

$$d\psi_t(V_p) - V_{\psi_t(p)} = d\psi_t(V_p - d\psi_{-t}V_{\psi_t(p)}),$$

and similarly for the second term. The result then follows from (6) and the fact that $d\psi_t \rightarrow Id$ as $t \rightarrow 0$. \square

Definition 3.11. A Killing vector field on a semi-Riemannian manifold is a vector field X for which the Lie derivative of the metric tensor vanishes $\mathcal{L}_X g = 0$.

Proposition 3.4. Let X be a vector field and denote its local flow by ψ_t . X is Killing if and only if all its (local¹⁰) flows are isometries.

Proof. Let $p \in M$ and \mathcal{U} a neighborhood of p on which ψ_t is well defined for all t sufficiently small. Then, if all ψ_t are isometries $\psi_t^*(g) = g$ and thus, by the above $\mathcal{L}_X(g) = 0$. Reciprocally, let $\mathcal{L}_X g = 0$ and let ψ_t be a local flow of X . By standard properties of ψ_t , we have $\psi_s \psi_t = \psi_{s+t}$ for s, t small enough. The previous proposition then implies that, for v, w tangent vectors at some point in the domain of the flow and s sufficiently small

$$\lim_{t \rightarrow 0} \frac{1}{t} (g(d\psi_{s+t}(v), d\psi_{s+t}(v)) - g(d\psi_s(v), d\psi_s(w))) = 0$$

Thus, the real-valued functions $s \rightarrow g(d\psi_s(v), d\psi_s(w))$ has zero derivative and is thus constant, i.e.

$$g(d\psi_s(v), d\psi_s(w)) = g(v, w),$$

which implies that ψ_s is an isometry. \square

¹⁰A Killing vector field on a geodesically complete semi-Riemannian manifold is in fact necessarily complete, cf [O'N83], page 254.

Definition 3.12. For any vector field X , define its deformation tensor as

$${}^X\pi(V, W) = \frac{1}{2} (g(D_V X, W) + g(V, D_W X)).$$

From the definition, ${}^X\pi$ is indeed $\mathcal{F}(M)$ -linear in each of its two arguments and thus a $(0, 2)$ tensor field. In coordinates

$$\begin{aligned} {}^X\pi_{\alpha\beta} &= \frac{1}{2} (D_\alpha X_\beta + D_\beta X_\alpha) \\ &:= D_{(\alpha} X_{\beta)}, \end{aligned}$$

where for any tensor $A_{\alpha\beta}$, $A_{(\alpha\beta)}$ denotes its symmetrized version

$$A_{(\alpha\beta)} := \frac{1}{2} (A_{\alpha\beta} + A_{\beta\alpha}).$$

Lemma 3.8.

$$2{}^X\pi = \mathcal{L}_X g.$$

Exercise 3.9. 1. Prove Lemma 3.8.

2. Consider Minkowski space with Cartesian coordinates $(x^\alpha) = (t, x^i)$. Show that the translations ∂_{x^α} , the rotations $x^i \partial_{x^j} - x^j \partial_{x^i}$, and the hyperbolic rotations $t \partial_{x^i} + x^i \partial_t$ are all Killing fields.
3. Prove that if $\phi : M \rightarrow N$ is an isometry, then $d\phi(D_X Y) = D_{d\phi(X)}(d\phi Y)$, for all $X, Y \in \Gamma(M)$. (Recall first what $d\phi(X)$ is well defined for X a smooth vector field and ϕ a diffeomorphism).

3.5 The induced connection on a curve, parallel translations and geodesics

Definition 3.13. Let N and M be smooth manifolds and $\phi : N \rightarrow M$ a smooth map. A vector field on ϕ is a map $Z : N \rightarrow TM$ such that $\pi \circ Z = \phi$, where π is the canonical projection $\pi : TM \rightarrow M$.

In other words, for any $q \in N$, $Z(q) = (\phi(q), v)$, where $v \in T_{\phi(q)}$.

In terms of vector bundles, Z is thus a section of the pullback bundle $\phi^*(TM)$.

Consider now a curve $\alpha : I \rightarrow M$. A vector field on α is according to the above definition a map that assigns to each $s \in I$ a tangent vector at $\alpha(s)$. Let $\Gamma(\alpha)$ denote the set of all vector fields along α .

A particular example is given by the tangent vector α' to α itself. Another one is obtained from any vector field $V \in \Gamma(M)$, as $V_\alpha : t \rightarrow V_{\alpha(t)}$.

Proposition 3.5. Let $\alpha : I \rightarrow M$ be a curve. Then, there is a unique map

$$\begin{aligned} \frac{d}{ds} : \Gamma(\alpha) &\rightarrow \Gamma(\alpha) \\ Z &\rightarrow Z' = \frac{dZ}{ds} \end{aligned}$$

called the induced covariant derivative such that

1. $\frac{d}{ds}$ is \mathbb{R} -linear.
2. Leibniz rule: $(hZ)' = \frac{dh}{ds}Z + hZ'$, for any $h \in \mathcal{F}(I)$,
3. $(V_\alpha)'(t) = D_{\alpha'(t)}(V)$, $\forall V \in \Gamma(M)$.

Moreover, we then have, for any vector fields Z_1 and Z_2 on α ,

$$\frac{d}{dt}g(Z_1, Z_2) = g(Z_1', Z_2) + g(Z_1, Z_2').$$

Proof. 1. Uniqueness: This follows from the first three properties only. Let us first assume that α lies in a single coordinate patch $(U, (x^i))$.

For any $Z \in \Gamma(\alpha)$, we can write

$$Z(t) = Z^i(t)\partial_{x^i},$$

where $Z^i(t) = Z(t)(x^i)$.

By the first two properties,

$$\frac{dZ}{ds} = (Z^i)'\partial_{x^i} + Z^i [(\partial_{x^i})|_\alpha]'$$

Note that by the usual bump function argument, it follows that if $V \in \Gamma(U)$, then the third property still holds, i.e.

$$(V_\alpha)'(t) = D_{\alpha'(t)}(V), \quad \forall V \in \Gamma(U).$$

(Indeed, given $t_0 \in I$, consider a bump function f around $\alpha(t_0)$, then fV can be extended to a global vector field on M , use then property 3, expands using property 2 and evaluate everything at t_0 .)

Thus by the third property,

$$[(\partial_{x^i})|_\alpha]'(t) = D_{\alpha'(t)}\partial_{x^i}.$$

In particular, the connection completely determined $\frac{dZ}{ds}$.

In the general case (i.e. α not restricted to a single coordinate patch), let J be an non-empty open subinterval on which $\alpha(J)$ lies in a single coordinate patch (U, x^i) . Let ${}^I \frac{d}{ds}$ the operation associated to the original α and ${}^J \frac{d}{ds}$ the one associated to $\alpha|_J$. From the above, ${}^J \frac{d}{ds}$ is already uniquely determined. Let us show that given any $Z \in \Gamma(\alpha)$,

$$\left[{}^I \frac{d}{ds} Z \right]_{|J} = {}^J \frac{d}{ds} Z|_J.$$

Let $t_0 \in J$, f be a bump function around $\alpha(t_0)$ with support in U and h be a bump function (defined on \mathbb{R}) around t_0 with support in J . As before we can write

$$Z(t) = Z^i(t)\partial_{x^i}$$

for $t \in J$. We consider hZ^i and $f\partial_{x^i}$ and extends them smoothly by 0 outside of respectively J and U .

Then, on one hand,

$$(Z^i h f \partial_{x^i})' = (h \cdot f \circ \alpha \cdot Z)' = (h f \circ \alpha)' Z + h f \circ \alpha Z',$$

and thus $(Z^i h f \partial_{x^i})'(t_0) = Z'(t_0)$. On the other hand,

$$\begin{aligned} (Z^i h f \partial_{x^i})' &= (Z^i h)' f \partial_{x^i} + h Z^i (f \partial_{x^i})|_{\alpha}' \\ &= (Z^i h)' f \partial_{x^i} + h Z^i D_{\alpha'}(f \partial_{x^i}). \end{aligned}$$

We then evaluate this at t_0 and recognise on the right-hand side the expression previously found for $\left[J \frac{d}{ds} Z|_J \right]$.

2. Existence: on any subinterval J of I such that $\alpha(J)$ lies in a coordinate neighborhood, define Z' be the above formula. One can then check that all four properties hold (with I replaced by J). Let J_1 and J_2 be two such subintervals of I , such that $J_1 \cap J_2 \neq \emptyset$. The above formula gives two vector fields on $J_1 \cap J_2$, called them Z'_1 and Z'_2 . Since the first three properties hold, we can apply our uniqueness statement and thus we must have $Z'_1 = Z'_2$ on $J_1 \cap J_2 \neq \emptyset$. In other words, the above formula defines a single vector field on α . □

Exercise 3.10. Check indeed that all four properties hold as claimed in the proof. (For the last property, check that it first hold if α lies in a single coordinate neighborhood for Z_1 and Z_2 given by $[\partial_{x^i}]_{\alpha}$, $[\partial_{x^j}]_{\alpha}$.)

From the above proof, we obtain the coordinate expression

$$\begin{aligned} Z' &= \frac{dZ^{\beta}}{ds} (\partial_{x^{\beta}})|_{\alpha} + Z^{\beta} D_{\alpha'}(\partial_{\beta}) \\ &= \frac{dZ^{\beta}}{ds} (\partial_{x^{\beta}})|_{\alpha} + Z^{\beta} (x^{\rho}(\alpha))' \Gamma_{\beta\rho}^{\gamma} \circ \alpha (\partial_{x^{\gamma}})|_{\alpha}, \end{aligned}$$

where $Z^{\beta} := Z(x^{\beta}) : I \rightarrow \mathbb{R}$ are the components of Z along α .

Recall that a special case of a vector field along α is given by the tangent vector field to α , α' . Its derivative along α , α'' is sometimes called the *acceleration* of the curve α . (Note that the first $'$ does not depends on the geometry, but the second does.)

A vector field Z on α such that $Z' = 0$ is said to be *parallel*. The above computation leads to

Proposition 3.6. Let $\alpha : I \rightarrow M$ be a curve, $a \in I$ and $z \in T_{\alpha(a)}M$. Then, there exists a unique parallel vector field Z on α such that $Z(a) = z$.

Proof. Let J be a non-empty open subinterval of I containing a and such that $\alpha(J)$ is contained on a unique coordinate chart. From the above, we must have

$$0 = \frac{dZ^{\beta}}{ds} + (x^{\rho}(\alpha))' \Gamma_{\beta\rho}^{\gamma} Z^{\beta}.$$

This is a system of linear odes which therefore has a unique solution defined on the whole interval J for given initial data.

Now, if J_1 is another interval such that $\alpha(J_1)$ is contained in a unique coordinate chart and $J_1 \cap J \neq \emptyset$, let then $t_1 \in J_1 \cap J \neq \emptyset$ and $z(t_1) = Z(t_1)$, where $Z(t_1)$ has been

obtained uniquely by solving the equation on J . By the same argument, $Z|_{J_1}$ is then given uniquely by solving the equations in coordinates with the data given by $z(t_1)$. Since I can be written as a union of such subintervals, we are done. (More precisely, given any $t_1 \in I$, $\alpha([a, t_1])$ can be covered by a finite number of coordinate neighborhoods. Using the uniqueness statement and the finite number of neighborhoods, Z can be defined on the whole of $[a, t_1]$ uniquely. Since this holds for any t_1 , the statement of the proposition holds.) □

Definition 3.14. *With the above notation, given $a, b \in I$ the map sending $z \in T_{\alpha(a)}M$ to $Z(b) \in T_{\alpha(b)}M$ such that $Z' = 0$ and $Z(a) = z$ is called a parallel translation.*

Exercise 3.11. *Prove that parallel translation is a linear isometry from $(T_{\alpha(a)}M, g_{\alpha(a)})$ to $(T_{\alpha(b)}M, g_{\alpha(b)})$.*

We can now define the notion of geodesics.

Definition 3.15. *A geodesic is a curve α such that $\alpha'' = 0$.*

Lemma 3.9. *Let $\alpha(s)$ be given in local coordinates as $(x^\alpha(s))$. Then the geodesic equations read*

$$(x^\rho)'' + \Gamma_{\beta\gamma}^\rho (x^\beta)' (x^\gamma)' = 0.$$

The geodesic equation is therefore a system of second order odes. Since it is non-linear, (due to the product of x' in the second term), it does not need to be defined globally, but given initial data it always has a unique maximal solution.

More precisely, we have by standard application of the Cauchy-Lipschitz theorem,

Lemma 3.10. *Let $v \in T_pM$. Then there exists a unique maximal solution $\gamma : I \rightarrow M$ to the geodesic equation such that $\gamma'(0) = v$ where I is an open interval and $0 \in I$. (Note that since $v \in T_pM$, $\gamma(0) = p$.)*

Such a maximal geodesic (what does maximality refer to here ?) is called a maximal geodesic or inextendible geodesic.

Definition 3.16. *We say that (M, g) is geodesically complete if every inextendible geodesic is defined on the whole of \mathbb{R} (i.e the geodesic equation in (M, g) always admit global solutions).*

Lemma 3.11. *Let $a \neq 0$, $\tau : s \rightarrow as + b$ be a straight line and $J = \tau^{-1}(I)$. Then, if $\gamma : I \rightarrow M$ is a geodesic, so is $\gamma \circ \tau : J \rightarrow M$.*

Exercise 3.12. *what about other changes of parametrization?*

We will note (γ_v, I_v) the geodesic associated to the vector v as above.

Note that any constant curve is a geodesic.

Exercise 3.13. • *Show that the (non-constant) geodesics of the Euclidean or Minkowski space are the straight lines.*

- A curve α is called a pregeodesic if it has a reparametrization as a geodesic. Let now α be a regular curve ($\alpha' \neq 0$) such that α'' and α' are collinear, i.e. $\alpha''(s) = f(s)\alpha'(s)$ for some f .
 1. Show that $\beta = \alpha \circ h$ for β a geodesic if and only if $h'' + (f \circ h)h'^2 = 0$. Note that this is a non-linear ode and that it may not have global solutions on the interval of definition of α .
 2. Assume now that $g(\alpha', \alpha') \neq 0$ and consider a constant speed reparametrization, i.e. such that $g(\beta', \beta') = c$ for some $c \in \mathbb{R}$. Prove that any such constant speed reparametrization of α is a geodesic.
 3. Prove that either $g(\alpha', \alpha') = 0$ identically on α , or that it never vanishes.
 4. Deduce from the above that if $g(\alpha', \alpha')(s_0) \neq 0$ for some s_0 , then α is a pregeodesic.
 5. The above proof does not work if $g(\alpha', \alpha') = 0$ identically on α but the result still holds. To prove it, assume that $\beta = \alpha \circ h$ is a geodesic with h a diffeomorphism such that $h' > 0$ and write $\alpha = \beta \circ k$, with $k = h^{-1}$ and derive the equation $(\log k')' = f$. This can then be solved.)
- Consider \mathbb{R}^3 with its usual coordinates (x, y, z) .
 1. Prove that there exists a unique connection D verifying

$$D_{\partial_x}(\partial_y) = \partial_z, \quad D_{\partial_x}(\partial_z) = \partial_y \quad (7)$$

$$D_{\partial_x}(\partial_z) = -\partial_y, \quad D_{\partial_z}(\partial_x) = \partial_y \quad (8)$$

$$D_{\partial_y}(\partial_z) = \partial_x, \quad D_{\partial_z}(\partial_y) = -\partial_x \quad (9)$$
 2. One of the conditions in the definition of the Levi-Civita connection is the torsion free connexion $[V, W] = D_V W - D_W v$. Prove that (7)-(9) indeed defines a unique connection but that it is not torsion free.
 3. Prove that its geodesics are straight lines. This proves in particular that geodesics do not define uniquely a connection (They actually define uniquely a torsion free connection).
 4. What is the parallel transports of $(1, 0, 0)$ and $(0, 1, 0)$ along the z axis ?

3.6 Physical interpretation of causal geodesics

Recall Newton's equation from classical mechanics. In any Galilean frame, for a particule of mass m ,

$$m\ddot{x} = \sum F.$$

When the forces are only gravitational, then $\sum F = -m\nabla\phi$, where ϕ is the gravitational potential. The equation then reduces to

$$\ddot{x} = -\nabla\phi(x). \quad (10)$$

This is a second order ode, like the geodesic equations, whose solution depends only on the initial position and initial velocity of the particule¹¹

¹¹The fact that the initial mass m appearing on the left-hand side of (??) equal the gravitational mass appearing on the right-hand side is the "weak equivalence principle" of Einstein: *the acceleration imparted by a body by a gravitational field is independent of the nature of the body.*

In General Relativity, the replacement of Newton's equation (10) is the geodesic equation. More precisely, a particular of mass m , which is free falling (no external forces apart the effect of gravity), moves along a geodesic whose tangent vector is timelike if $m > 0$, null if $m = 0$. These last requirements are the general relativistic requirement that no particles can move faster than the speed of light.

Thus in General Relativity, there are no gravitational force, there is only geometry. To obtain a closed physical theory, it remains to have a prescription for this geometry. More specifically, in Newtonian mechanics, the gravitational potential ϕ is obtained through Poisson equation

$$\Delta\phi = 4\pi G\rho, \quad (11)$$

where ρ is the local mass density, G a constant (the gravitational constant) and $\Delta = \sum_i \frac{\partial^2}{\partial x^i \partial x^i}$ is the usual Laplacian.

Thus, although we have already a replacement for Newton's equation, we still need to find a replacement for Poisson equation. Since g appears in the geodesic equation, we would like this replacement to be an equation on g to obtain a closed system.

3.7 The exponential map

Definition 3.17. Let $p \in M$. Let U_p be the set of vectors in T_pM such that for $v \in U_p$, γ_v is defined on $[0, 1]$. The exponential map is then

$$\begin{aligned} \exp_p : U_p &\rightarrow M \\ v &\rightarrow \gamma_v(1). \end{aligned}$$

Remark 3.5. By smooth dependence with respect to the initial data (cf Cauchy-Lipschitz), the exponential map is a smooth function on U_p .

Remark 3.6. With the above notation, for any $t \in \mathbb{R}$, the map $s \rightarrow \gamma_v(ts)$ is a geodesic. Moreover, it has initial velocity tv and position p . Thus, $\gamma_v(ts) = \gamma_{tv}(s)$ by uniqueness. In particular, we have

$$\exp_p(tv) = \gamma_{tv}(1) = \gamma_v(t).$$

The curve $t \rightarrow \gamma_v(t)$ will be called a radial geodesic (emanating from p).

Exercise 3.14. Check that the Cauchy-Lipschitz theorem from standard ode theory guarantees that U_p contains a neighborhood of $0 \in T_pM$ and in fact that U_p is open.

Proposition 3.7. For each $p \in M$, there exists an open neighborhood \mathcal{V} of 0 in T_pM such that \exp_p is a diffeomorphism from \mathcal{V} to some open neighborhood \mathcal{U} of $p \in M$.

Proof. Let (x^α) be coordinates on M . Recall that the (x^α) induces a global coordinate system¹² (X^α) on T_pM given by

$$X = (X^\alpha \partial_{x^\alpha})|_p,$$

for any $X \in T_pM$. Consider the differential of the exponential map at $0 \in T_pM$. We have

¹²These coordinates are sometimes called *conjugate* to the (x^α) .

$$d \exp_p(0) : T_0 T_p M \rightarrow T_{\exp_p(0)} M = T_p M.$$

Let $V \in T_0 T_p M$. Since $T_p M$ is a vector space, we can identify $T_0 T_p M$ with $T_p M$. More precisely. We can write $V = V^\alpha [\partial_{x^\alpha}]_{|0 \in T_p M}$. Then, $\alpha : s \rightarrow s V^\alpha [\partial_{x^\alpha}]_{|p \in M}$ is a curve in $T_p M$ whose tangent vector is V . Denote by \tilde{V} the vector $V^\alpha [\partial_{x^\alpha}]_{|p \in M}$ (the $\tilde{\cdot}$ map is sometimes called the canonical isomorphism in this context).

Then by definition and the above remark,

$$\begin{aligned} d \exp_p(0)(V) &= \left. \frac{d}{ds} (\exp_p(\alpha(s))) \right|_{s=0} \\ &= \left. \frac{d}{ds} (\exp_p(s \tilde{V})) \right|_{s=0} \\ &= \left[\left. \frac{d}{ds} \gamma_{\tilde{V}}(s) \right]_{|s=0} \\ &= \tilde{V}. \end{aligned}$$

Thus, $d \exp_p(0)$ coincides with the $\tilde{\cdot}$ map. The result then follows by the inverse function theorem. \square

Exercise 3.15. Show that the $\tilde{\cdot}$ map is independent of the choice of coordinates.

Definition 3.18. If \mathcal{U} and \mathcal{V} are as in the preceding proposition and if \mathcal{U} is star-shaped around 0, then \mathcal{V} is called a normal neighborhood of p .

Recall that \mathcal{U} starshaped around 0 means that for any $x \in \mathcal{U}$, $sx \in \mathcal{U}$ for all $0 \leq s \leq 1$. (Note that $\mathcal{U} \in T_p M$ is a subset of a vector space, so that sx makes sense.)

Note that if \mathcal{U} is a neighborhood of 0, then it contains a starsphaped neighborhood of 0.

Proposition 3.8. If $\mathcal{V} = \exp_p(\mathcal{U})$ is a normal neighborhood of $p \in M$, then for each point $q \in \mathcal{V}$, there exists a unique geodesic $\sigma : [0, 1] \rightarrow \mathcal{V}$ from p to q . Furthermore $\sigma'(0) = \exp_p^{-1}(q) \in \mathcal{U}$.

Proof. Let $q \in \mathcal{V}$ and $v = \exp_p^{-1}(q) \in \mathcal{U}$. Since \mathcal{U} is starshaped, the ray $\rho : t \rightarrow tv$, $0 \leq t \leq 1$ is contained in \mathcal{U} . Let $\sigma = \exp_p \circ \rho$. Then σ is a geodesic from p to q .

For the uniqueness, consider $\tau : [0, 1] \rightarrow \mathcal{V}$ be any geodesic joining p to q with values in \mathcal{V} . Let $w = \tau'(0)$, so that $\tau = \gamma_w$. The difficulty is that we do not know a priori that $w \in \mathcal{U}$.

Recall that by definition $t \rightarrow \exp_p(tw)$ is defined for t such that γ_{tw} is defined on $[0, 1]$. Since $\gamma_{tw}(1) = \gamma_w(t)$ for all t sufficiently small, it follows by uniqueness that $t \rightarrow \gamma_{tw}(1) = \gamma_w$. In particular, they have the same interval of definition, so that $\gamma_{tw}(1)$ is defined for $t \in [0, 1]$. Thus, $t \rightarrow \exp_p(tw)$ is defined for $t \in [0, 1]$. Moreover, we have $\tau(t) = \exp_p(tw)$ for all $t \in [0, 1]$. However, we still do not know that $w \in \mathcal{U}$.

Consider the set of all $t \in [0, 1]$, such that $tw \in \mathcal{U}$. It is a non-empty subset of $[0, 1]$. Since \mathcal{U} is star-shaped around 0, E is an interval. Since \mathcal{U} is open, E is open in $[0, 1]$. Let t_0 be the supremum of E .

Recall that the map τ takes value in \mathcal{V} by assumption, so that in particular, $\tau(t_0) \in \mathcal{V}$ and thus $\exp_p^{-1} \tau(t_0) \in \mathcal{U}$. Now by continuity,

$$\exp_p^{-1} \tau(t_0) = \lim_{t \rightarrow t_0, t < t_0} \exp_p^{-1} \tau(t) = \lim_{t \rightarrow t_0, t < t_0} tw = t_0 w.$$

Thus, $t_0 w \in \mathcal{U}$, $t_0 = 1$, $w \in \mathcal{U}$ and $q = \exp_p(w)$. Since \exp_p is a diffeomorphism on \mathcal{U} , it follows that $w = v$ and hence that $\tau = \sigma$. □

Remark 3.7. Note that if we do not specify the domain of definition of σ , then the uniqueness does not hold anymore since we can make a change of parametrization.

Exercise 3.16. Let \mathbb{S}^2 be the unit sphere in \mathbb{R}^3 endowed with its round metric.

1. Show that the rotations of \mathbb{S}^2 are isometries.
2. Let γ be a geodesic and $\phi : (M, g_M) \rightarrow (N, g_N)$ an isometry. Show that $\phi(\gamma)$ is also a geodesic.
3. Show that on \mathbb{S}^2 , the great circles are geodesics.
4. Show that given any two distinct points on \mathbb{S}^2 , they can be joined by two distinct geodesics (where we consider two geodesics to be equal if they have same image in \mathbb{S}^2 to account for the parametrization freedom).

Definition 3.19. A broken geodesic is a (continuous) piecewise smooth curve with each piece being a smooth geodesic.

Recall that a piecewise smooth curve is a continuous map $\gamma : I \rightarrow M$, where I is an non-empty interval such that if $I = [a, b]$, then there exists a finite partition $a = t_0 < t_1 \dots < t_k = b$ such that each curve segment $\gamma|_{[t_i, t_{i+1}]}$ is a smooth curve, while for an interval with open endpoints, say $I = (a, b)$, we require that for each non-empty closed interval $J \subset I$, the restriction $\gamma|_J$ is piecewise smooth.

A classical result in differential geometry asserts that a manifold is connected if and only if it is *path connected*. Here is a slight variation in the case of semi-Riemannian manifolds.

Lemma 3.12. A semi-Riemannian manifold M is connected if and only if two points of M can be joined by a broken geodesic.

Proof. Let M be connected and $p \in M$. Let \mathcal{C} be the set of points that can be connected to p by broken geodesics.

\mathcal{C} is open: indeed if $q \in \mathcal{C}$, let \mathcal{V} be a normal neighborhood around q . Then from the above lemma, $\mathcal{V} \subset \mathcal{C}$.

$M \setminus \mathcal{C}$ is open: Let $q \in M \setminus \mathcal{C}$ and again \mathcal{V} be a normal neighborhood around q . Then $\mathcal{V} \subset M \setminus \mathcal{C}$. For, if $q' \in \mathcal{V} \cap \mathcal{C}$, then from the above, there exists a geodesic from q to q' and from p to q' and thus, a broken geodesic from q to p , which is a contradiction. Since $\mathcal{C} \neq \emptyset$, it follows from the connectedness of \mathcal{C} that $\mathcal{C} = M$. The converse is trivial. □

Let \mathcal{V} be a normal neighborhood of $p \in M$. Let (e_α) be an orthonormal basis for $T_p M$. For any $q \in \mathcal{V}$, let $(x^\alpha(q))$ be determined by

$$\exp_p^{-1}(q) = x^\alpha(q)e_\alpha.$$

In other words, given $q \in \mathcal{V}$, we first compute $v_q = \exp_p^{-1}(q)$. Then the components of v_q in the orthonormal basis e_α provides the x^α . The x^α are smooth functions since they can be written as a composition of smooth functions and it is also immediate that they form a local coordinate system near p .

Lemma 3.13. *The map $q \rightarrow x^\alpha(q)$ determines a local coordinate system on \mathcal{V} called normal coordinates.*

If ω^α is the dual basis to the e_α , then $x^\beta \circ \exp_p = \omega^\beta$ on \mathcal{U} . Moreover,

Proposition 3.9. *Let (x^α) be normal coordinates. Then, $g_{\alpha\beta}(p) = g(p)(e_\alpha, e_\beta)$ and $\Gamma_{\beta\gamma}^\alpha(p) = 0$.*

Remark 3.8. *Note that $g_{\alpha\beta}(p) = g(p)(e_\alpha, e_\beta)$ means that the matrix $g_{\alpha\beta}(p)$ coincides with the Minkowski (up to an reordering the e_α) or the identity matrix in the Lorentzian or Riemannian case. In physics, these coordinates are sometimes referred to as "local inertial frames" essentially because the laws of physics written in these coordinates will look similar (up to a first order term) to those written in Minkowski space or the Euclidean space, at least on a sufficiently small neighborhood of p . In other words, if your laboratory is sufficiently small¹³ and if you make measurements very quickly (we need a small region of spacetime, not just of space), then the measurements should converge to those obtained in say Minkowski space, i.e. we can neglect all gravitational effects (of course, if your lab becomes too small, then quantum effects might appear, but this is another story).*

Proof. Let $v = v^\alpha e_\alpha \in T_p M$, where e_α are orthonormal as above and v^α the components of v in the (e_α) basis. Since $\exp_p(tv) = \gamma_v(t)$ (for t small enough),

$$x^\alpha(\gamma_v(t)) = x^\alpha(\exp_p(tv)) = tv^\alpha.$$

Hence, $v = \gamma'_v(0) = v^\alpha [\partial_{x^\alpha}]_p$, i.e. the components of v in the $([\partial_{x^\alpha}]_p)$ basis are also given by v^α . It follows that $e_\alpha = [\partial_{x^\alpha}]_p$ and hence that $g_{\alpha\beta}(p) = g(p)(e_\alpha, e_\beta)$.

Moreover, from the above expression, we have $\frac{d^2}{dt^2}(x^\alpha(\gamma_v(t))) = 0$, and since γ_v is geodesic, this implies that

$$\Gamma_{\beta\gamma}^\alpha(\gamma_v(t)) \frac{d}{dt}(x^\beta(\gamma_v(t))) \frac{d}{dt}(x^\gamma(\gamma_v(t))) = 0.$$

Evaluated at $t = 0$, we get that

$$\Gamma_{\beta\gamma}^\alpha(p) v^\beta v^\gamma = 0,$$

for all v , which implies, by polarization, that $\Gamma_{\beta\gamma}^\alpha(p) = 0$. □

Remark 3.9. *With the above notations, the Jacobian matrix of the differential of the exponential map at p at any point $v \in \mathcal{U}$ with respect to the coordinates associated to the e_α on $T_p M$ and the normal coordinates on $\exp_p(\mathcal{U})$ is just the identity matrix.*

Exercise 3.17. *Prove that if (x^α) is a normal coordinate system and $X = X^\alpha \partial_{x^\alpha}$ with X^α constants, then the integral curve of X through p is a geodesic.*

¹³Smallness should always be measured with respect to something. Here, we want the Γ to stay small on some coordinate patch and since at p we have arranged that they vanish, we need the size of the lab to be small compared to the derivatives of the Γ . These derivatives are not very geometric, so if we want something more physical (i.e. less dependent on our choices of gauge), we need a tensorial quantity at the level of the first derivative of the Γ . This is the curvature, soon to be defined in these lectures.

4 Curvature

4.1 The Riemann curvature tensor and its symmetries

For a surface in \mathbb{R}^3 , a well known notion of curvature is that of the *Gaussian curvature*. This is an intrinsic invariant of the surface, in the sense that Gauss proved that the Gaussian curvature is independent of the fact that the surface happens to be embedded in \mathbb{R}^3 . The generalization of the the Gaussian curvature to arbitrary (semi-)Riemannian manifold, led Riemann to Riemannian geometry.

Lemma 4.1. *Let (M, g) be a semi-Riemannian manifold and D its Levi-Civita connection. The function*

$$\begin{aligned} R : \Gamma(M)^3 &\rightarrow \Gamma(M) \\ (X, Y, Z) &\rightarrow R_{XY}Z := D_X D_Y Z - D_Y D_X Z - D_{[X, Y]} Z \end{aligned}$$

defines a (1, 3) tensor field on M called the Riemann curvature tensor of (M, g) .

Remark 4.1. *Eventhough R only takes 3 vector fields as arguments, it is a (1, 3) tensor field in the sense that given any one-form ω , the map*

$$(X, Y, Z, \omega) \rightarrow \omega(R_{XY}Z)$$

is a $\mathcal{F}(M)$ multi-linear map from $\Gamma(M)^3 \times \Lambda^1(M)$ to \mathbb{R} .

Proof. For R to define a tensor field, one needs R to be $\mathcal{F}(M)$ linear in each of its arguments. Let $f \in \mathcal{M}$, then

$$\begin{aligned} R_{X, fY}Z &= D_X D_{fY}Z - D_{fY} D_X Z - D_{[X, fY]}Z \\ &= D_X (f D_Y Z) - f D_Y D_X Z - D_{X(f)Y + f[X, Y]}Z \\ &= X(f) D_Y Z + f D_X D_Y Z - f D_Y D_X Z - X(f) D_Y Z - f D_{[X, Y]}Z \\ &= f R_{X, Y}Z. \end{aligned}$$

The other cases are similar. □

Remark 4.2. *The sign of R is conventional, i.e. some authors, for instance O'Neil [O'N83]) define R as above but with an extra minus sign everywhere.*

Remark 4.3. *Even though, we have defined the Riemann curvature tensor from the Levi-Civita connection, every connection defines a curvature tensor.*

Since R is a tensor field, for every $p \in M$, it defines a map defined on $(T_p M)^3$, cf Section 2.14.

Given $x, y \in T_p M$, let R_{xy} be the map $T_p M \rightarrow T_p M$ sending z to $R_{xy}z$. We have

Proposition 4.1 (Symmetries of the Riemann tensor in the metric case).

$$R_{xy} = -R_{yx}, \tag{12}$$

$$g(R_{xy}v, w) = -g(R_{xy}w, v), \tag{13}$$

$$R_{xy}z + R_{yz}x + R_{zx}y = 0, \tag{14}$$

$$g(R_{xy}v, w) = g(R_{vw}x, y). \tag{15}$$

Remark 4.4. *The curvature can be defined for any linear connexion. Some symmetries always hold (such as the first), but others (such as the last) are specific to metric connexion.*

Remark 4.5. *Equation (14) is often called the first Bianchi identity.*

Proof. The first equation is obvious in view of the definition of R . For the second, we assume that X, Y, Z are local vector fields such that $X_p = x, Y_p = y, Z_p = z$ and such that all commutators $[X, Y] = [Y, Z] = [Z, X] = 0$. Since R is a tensor field, it does not depend on the choice of extensions. (ex: why do such extensions exists ?) Then,

$$\begin{aligned} g(R_{XY}V, V) &= g(D_X D_Y V, V) - g(D_Y D_X V, V) \\ &= Xg(D_Y V, V) - g(D_Y V, D_X V) - Yg(D_X V, V) + g(D_X V, D_Y V) \\ &= 1/2XYg(V, V) - 1/2YXg(V, V) \\ &= 0, \end{aligned}$$

since X and Y commutes. By polarization, we obtain the second equation.

We now consider equation (14). For a function $F := F(X, Y, Z)$, let $\mathcal{C}(F)$ denotes the sum over cyclic permutations of X, Y, Z i.e

$$\mathcal{C}(F)(X, Y, Z) = F(X, Y, Z) + F(Z, X, Y) + F(Y, Z, X).$$

Then, since any cyclic permutation leaves $\mathcal{C}(F)(X, Y, Z)$ unchanged, we have

$$\begin{aligned} \mathcal{C}R_{XY}Z &= \mathcal{C}D_X D_Y Z - \mathcal{C}D_Y D_X Z \\ &= \mathcal{C}D_Y D_Z X - \mathcal{C}D_Y D_X Z \\ &= \mathcal{C}D_Y [Z, X] = 0. \end{aligned}$$

Equation (15) follows from the first three and we leave the computations as an exercise detailed below. \square

Exercise 4.1. *With the above notation, consider the $E(Y, V, X, W) := g(\mathcal{C}R_{YV}X, W)$. Let $H(Y, V, X, W)$ be the sum over all cyclic permutations $\sigma(Y, V, X, W)$ of (Y, V, X, W) of $E(\sigma(Y, V, X, W))$. Compute H as a sum of 12 terms by expanding the terms of the form $\mathcal{C}R_{YV}X$. Deduce that equation (15) holds.*

Since R is a tensor field, we can consider its covariant differential $DR : V \in \Gamma(M) \rightarrow D_V R$. Recall that since D_V is a tensor derivation, it maps tensor fields into tensor fields, and since $D_V R$ is $\mathcal{F}(M)$ linear in V , one can thus view DR as a $(1, 4)$ tensor field which takes 4 vector fields X, Y, Z, V and return a vector field

$$DR(X, Y, Z, V) = D_V R_{XY}Z = (D_V R)(X, Y)V.$$

Again, at any $p \in M$, we get a map from $(T_p M)^4$ to $T_p M$ and given $(x, y, z) \in (T_p M)^3$, $D_z R(x, y)$ is an endomorphism of $T_p M$. We have

Proposition 4.2. *We have*

$$D_z R(x, y) + D_x R(y, z) + D_y R(z, x) = 0,$$

called the second Bianchi identity.

Proof. Choose a normal coordinate system (x^α) at p . Let $X_p, Y_p, Z_p \in T_p M$ and let X, Y, Z be local extensions of X_p, Y_p, Z_p by choosing X, Y, Z to have constant components in the ∂_{x^α} basis of vector fields. Note that X, Y, Z all commutes together and that their covariant derivatives vanish at p .

Then, from the tensor derivation laws¹⁴,

$$(D_Z R)(X, Y)V = D_Z(R(X, Y)V) - R(D_Z X, Y)V - R(X, D_Z Y)V - R(X, Y)(D_Z V).$$

At p , the two middle terms on the RHS vanishes, hence

$$\begin{aligned} (D_Z R)(X, Y)V &= D_Z[(D_X D_Y - D_Y D_X)V] - (D_X D_Y - D_Y D_X)D_Z V, \text{ at } p \quad (16) \\ &= [D_Z, [D_Y, D_X]]V, \text{ at } p, \end{aligned}$$

where $[D_Z, [D_Y, D_X]]V$ is a short-hand notation for the RHS of (16). The result then follows from the Jacobi identity

$$[D_Z, [D_Y, D_X]] + [D_X, [D_Z, D_Y]] + [D_Y, [D_X, D_Z]] = 0$$

(just write it out to see the cancellations). □

Let the components of the curvature tensor be defined as

$$R(X, Y)Z = R^\alpha_{\beta\gamma\delta} Z^\beta X^\gamma Y^\delta \partial_{x^\alpha},$$

i.e.

$$R^\alpha_{\beta\gamma\delta} \partial_{x^\alpha} = R(\partial_\gamma, \partial_\delta)(\partial_\beta).$$

Exercise 4.2. 1. Show that

$$R^\alpha_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha_{\delta\beta} - \partial_\delta \Gamma^\alpha_{\gamma\beta} + \Gamma^\sigma_{\delta\beta} \Gamma^\alpha_{\gamma\sigma} - \Gamma^\sigma_{\gamma\beta} \Gamma^\alpha_{\delta\sigma}. \quad (17)$$

2. Let

$$R_{\alpha\beta\delta\gamma} = g_{\alpha\rho} R^\rho_{\beta\gamma\delta}.$$

Show that $R_{\alpha\beta\gamma\delta} = R_{\gamma\delta\alpha\beta}$.

3. The Ricci identity: for $X \in \Gamma(M)$, define the second covariant of X , DDX as the covariant derivative $D(DX)$ of DX . It is a $(1,2)$ tensor field and we denote its components by $X^\alpha_{;\beta\gamma}$. Prove that

$$X^\alpha_{;\beta\gamma} - X^\alpha_{;\gamma\beta} = R^\alpha_{\delta\beta\gamma} X^\delta.$$

4. Let T be an arbitrary tensor field. Show that

$$[D_\alpha, D_\beta]T = (R \star T)_{\alpha\beta}.$$

where $R \star T$ denotes a tensor obtained from $R \otimes T$ and contractions.

5. There are plenty of nice geometric interpretations of the Riemann curvature tensor. Write something about it. (look for instance for the relation between curvature and holonomy).

¹⁴Note that strictly speaking, we should apply D_Z to the tensor $\tilde{R} : (\omega, X, Y, V) \rightarrow \omega(R(X, Y)V)$. One can view the formula given above for $D_Z R(X, Y)V$ as the correct definition such that $D_Z \tilde{R}(\omega, X, Y, V) = \omega(D_Z R(X, Y)V)$.

4.2 The Ricci tensor and the scalar curvature

Recall that a p -form on manifold is by definition a $(0, p)$ tensor field which is totally anti-symmetric. Given a semi-Riemannian manifold, we can define locally¹⁵ a n -form called *volume form* out of the metric, given by

$$\eta = \sqrt{|\det g|} dx^1 \wedge dx^2 \wedge \dots \wedge dx^n,$$

where $\det g = \det g_{\alpha\beta}$ is the determinant of the metric in the local coordinate system and $dx^1 \wedge dx^2 \wedge \dots \wedge dx^n$ is the unique n -form such that

$$[dx^1 \wedge dx^2 \wedge \dots \wedge dx^n](\partial_{x^1}, \partial_{x^2}, \dots, \partial_{x^n}) = 1.$$

Exercise 4.3. • Prove that if ω is a n -form and $V_i = A^{ij}W_j$, where V_i, W_j are vector fields and the A^{ij} are smooth functions, then

$$\omega(V_1, \dots, V_n) = (\det A) \omega(W_1, \dots, W_n).$$

- Let (x^α) and (y^i) be two local coordinate systems defined on a common open set. Let η be defined with respect to the (x^α) coordinate system. Let $\det g_y$ denote the determinant of g in the y coordinate system. Prove that

$$\eta = \pm \sqrt{|\det g_y|} dy^1 \wedge \dots \wedge dy^n.$$

- Prove that $D\eta = 0$. (Hint: use normal coordinates.)

As we will see later, the volume form appears naturally in many formulae involving geometric operators and of course, it is also the natural n -form to perform integrations on our manifold.

Choose normal coordinates (x^i) at some p and consider the function $\det g = \pm [\eta(\partial_{x^1}, \partial_{x^2}, \dots, \partial_{x^n})]^2$ in the (x^α) system of coordinates. Doing a Taylor expansion, we have

$$\det g = \det g(0) + \partial_\alpha(\det g)(0)x^\alpha + \frac{1}{2}\partial_\alpha\partial_\beta(\det g)(0)x^\alpha x^\beta + O(x^3),$$

where $\det g(0) = \pm 1$. The derivative terms on the right hand side can all be rewritten in terms of covariant derivatives. Using the Levi-Civita property of the metric ($Dg = 0$), the symmetrization obtained thanks to the $x^\alpha x^\beta$, and the total anti-symmetry of the determinant and the special properties of normal coordinates, all the second covariant derivatives can be rewritten using the Riemann curvature tensor. However, only specific combinations of the components of the curvature tensor actually appears and they can be rewritten as a tensor: the Ricci tensor, which is defined as follows.

Definition 4.1. The Ricci tensor, denoted $Ric(g)$, is the tensor obtained by contraction of the first and third positions of the Riemann curvature tensor, i.e. in coordinates

$$Ric(g)_{\beta\delta} = R^\alpha{}_{\beta\alpha\delta}.$$

¹⁵It can even be made global if (and only if) M is orientable. See Appendix A.

Remark 4.6. The Taylor expansion of the volume form in normal coordinates is given by the formula

$$\sqrt{|\det g|} = 1 - \frac{1}{6} Ric(g)_{\alpha\beta} x^\alpha x^\beta + O(x^3).$$

The Ricci tensor “measures” the change in volume in normal coordinates of an infinitesimally small volume element. We include a proof of this formula in Appendix D below.

Lemma 4.2. The Ricci tensor is symmetric: $Ric(g)_{\alpha\beta} = Ric(g)_{\beta\alpha}$.

Proof. This follows easily from the symmetries of the Riemann tensor. □

Similarly, the scalar curvature is defined as

Definition 4.2. The scalar curvature is then defined as the trace of $Ric(g)$ i.e., in local coordinates,

$$R(g) = Ric_{\alpha\beta} g^{\alpha\beta}.$$

Remark 4.7. The scalar curvature has again an interpretation in terms of Taylor expansion, this time, one needs to consider the surface of a geodesic ball centered at some point and compare it with the surface of a ball of flat space of the same radius.

5 Divergence free vector and tensor fields

5.1 Some differential operators

Definition 5.1. Let $f \in \mathcal{F}(M)$. Then, $grad f$ (the gradient of f) is by definition the unique vector field such that

$$g(grad f, V) = \langle df, V \rangle = V(f),$$

for any $V \in \Gamma(M)$.

In other words, $grad f = \sharp df$. In components,

$$(grad f)^\alpha = g^{\alpha\beta} \partial_\beta f.$$

One often write Df instead of $grad f$ for the gradient of f , which should then not be confused the map $Df: V \rightarrow D_V(f) = \langle df, V \rangle = df(V)$.

Definition 5.2. Let X be a vector field. We then define its divergence $div(X)$ as

$$div(X) := D_\alpha X^\alpha = C(DX)$$

where C is the contraction operator.

Exercise 5.1. 1. For a tensor field, we define similarly its divergence by computing first its covariant differential and then applying the contraction operator on the the first two indices. For instance, the divergence of a $(1, 1)$ tensor field T is a one form given in components by $D_\mu T^\mu_\nu$. Compute its components in terms of those of T^μ_ν .

2. (a) Let (a_{ik}) be a non-singular matrix, of inverse a^{ik} and determinant $\det a$. Let (A^{ik}) be the matrix of cofactors of (a_{ik}) . We regard A^{ik} as smooth functions of the a_{ik} . Prove that

$$\partial_{a_{ij}} \det a = A^{ik} = (\det a) a^{ki}.$$

- (b) Deduce from the first question that for a smooth metric g

$$\partial_{x^\alpha} \det g = (\det g) g^{\gamma\beta} \partial_{x^\alpha} g_{\gamma\beta}.$$

- (c) Show that $\text{div}(X) = \frac{1}{|\det g|^{1/2}} \partial_\alpha (|\det g|^{1/2} X^\alpha)$ in any system of coordinates (note the absolute value, which is necessary in the Lorentzian case).

3. Let $\eta = \sqrt{|\det g|} dx^1 \wedge dx^2 \wedge \dots \wedge dx^n$ be a local volume form associated to g and $X \in \Gamma(M)$. Show that

$$d(i_X \eta) = \text{div}(X) \eta = \mathcal{L}_X(\eta),$$

where $i_X \eta$ is an $(n-1)$ form, the interior product of η by X defined by

$$i_X \eta(X_1, \dots, X_{n-1}) = \eta(X, X_1, \dots, X_{n-1}), \quad X_1, \dots, X_{n-1} \in \Gamma(M).$$

(Hint: use a coordinate system such that $\partial_{x^1} = X$.)

We end this section by the definition of a natural second order differential operator associated to a Riemannian or Lorentzian metric.

Definition 5.3. • Assume that g is Riemannian, then the operator

$$\psi \in \mathcal{F}(M) \rightarrow \Delta_g(\psi) = g^{\alpha\beta} D_\alpha D_\beta \psi = \text{div}(\text{grad } \psi),$$

is called the Laplace-Beltrami operator associated to g .

- For g Lorentzian, we have similarly an operator \square_g

$$\psi \in \mathcal{F}(M) \rightarrow \square_g(\psi) = g^{\alpha\beta} D_\alpha D_\beta \psi = \text{div}(\text{grad } \psi),$$

called the wave (or D'Alembertian) operator associated to g .

Remark 5.1. One can define a similar operator for any semi-riemannian metric g but it is only for the Riemannian and Lorentzian case that there currently exists a good understanding (and, to the author knowledge, only these operators appears naturally in physical contexts).

Remark 5.2. For a tensor field T of arbitrary type, we can define similarly

$$\square_g T = g^{\alpha\beta} D_\alpha D_\beta T.$$

Remark 5.3. For a Riemannian metric, the principal symbol of Δ_g written in an arbitrary system of local coordinates is elliptic, while for a Lorentzian metric it is hyperbolic.

Exercise 5.2. • Show that the principal symbol of \square_g is $g^{\alpha\beta}(p) \xi_\alpha \xi_\beta$. It can thus be viewed as a real function defined on the cotangent bundle TM^* .

- Show that in any local system of coordinates such that $g_{00} < 0$, g_{ij} definite positive, if $\xi = (\tau, \xi_i)$ (i.e. $\xi_0 = \tau$), then, for each $(\xi_i) \neq 0$, the equation $P(\tau) := g^{\alpha\beta} \xi_\alpha \xi_\beta = 0$ has two distinct real roots.
- Show that in any local system of coordinates such that $g_{00} < 0$, g_{ij} definite positive, the wave equation can be recast as a symmetric hyperbolic system in the sense of Friedrichs.

5.2 Remarks on conservation laws and Lagrangian theory

Recall Stokes theorem. Let M be a smooth orientable manifold and $N \subset M$ a domain with smooth boundary ∂N , $j : \partial N \rightarrow M$ the inclusion map, then $\forall \omega \in \Lambda^{n-1}(M)$ such that $\text{supp}(\omega) \cap \partial N$ is compact, we have

$$\int_{\partial N} j^* \omega = \int_N d\omega.$$

Assume for simplicity that N is such that ∂N is compact and let X be a vectorfield. Assume also that M is orientable. In this case, the volume form η can actually be defined globally and agree with the given formula in any local chart¹⁶. Then, from Exercise 5.1.2c,

$$\int_{\partial N} j^*(i_X \eta) = \int_N \text{div}(X) \eta.$$

In particular, for every divergence free vector field X ,

$$\int_{\partial N} j^*(i_X \eta) = 0.$$

In physics, this is interpreted as a *conservation law*, in the sense that the total flux of X through ∂N vanishes, so that the incoming flux=the outgoing flux.

These conservation laws are important in PDEs (and of course in physics) applications because they will be the source of many of the estimates needed to control solutions.

The question becomes: how to construct divergence free vector fields ?

One answer: out of *divergence free, symmetric tensor fields and out of symmetries of the metric*.

To explain this, let T be a divergence free tensor field. Given T^μ_ν its components, the divergence free condition reads

$$D_\mu T^\mu_\nu = 0.$$

Let now X be a vector field and define the vector field ${}^X J$ by its components as

$${}^X J^\mu = T^\mu_\nu X^\nu.$$

Then,

$$\text{div}({}^X J^\mu) = T^\mu_\nu D_\mu X^\nu.$$

Thus, ${}^X J$ is divergence free and we get a conservation law provided that $T^\mu_\nu D_\mu X^\nu = 0$.

Assume now that T is a symmetric tensor field, so that $T^{\mu\nu} = T^{\nu\mu}$. In this case,

$$T^\mu_\nu D_\mu X^\nu = T^{\mu\nu} 1/2 (D_\mu X_\nu + D_\nu X_\mu) = T^{\mu\nu X} \pi_{\mu\nu},$$

where ${}^X \pi$ is the deformation tensor of X .

Recall from Definitions 3.11-3.12 and Lemma 3.8 that vector fields with vanishing deformation tensor, called Killing fields, arose from the symmetries of the space-time (their flow generates isometries).

Conclusion: Killing vector fields and symmetric divergence free tensors give conservation laws as explained above.

¹⁶See Appendix A

Remark 5.4. *Similar arguments based on Stokes theorem can be applied to symmetric tensor fields and vector fields even when they are not divergence free or Killing. In that case, we would get error terms from the divergence term, and the game becomes to control these errors. In particular, the control of solutions for non-linear problems often falls in that category.*

The question now becomes: how to construct or obtain divergence free symmetric tensors and Killing fields? It turns out that in many physical problems, both arise naturally.

5.3 The energy-momentum tensor

For all the standard equations of physics (say coming from electromagnetism, fluid dynamics, scalar fields etc.), it is not hard to write them down first in tensorial notation in special relativity (i.e. in the special case of a background geometry given by Minkowski space) and then in an arbitrary Lorentzian manifold by replacing partial derivatives in Cartesian coordinates by covariant derivatives in local coordinates. Now, in all these cases, the equation can be written simply under the form of a symmetric divergence free tensor field constructed out of the unknowns. These divergence free tensor fields themselves can be constructed from variational techniques and takes their roots in Lagrangian theory¹⁷. Thus, the physics is essentially going to tell us what are the divergence free symmetric tensor to consider. As for Killing fields, they will appear also naturally. For instance, if we imagine a physical system which is stationary (i.e. there is a sense in which the system is not evolving), then the physicist will model this by the existence of a Killing field T which has some timelike property ($g(T, T) < 0$ at least in some region of the manifold). There is another (similarly vague) argument for the existence of symmetry in many solutions of problems arising in physics. These solutions can often be constructed by variational arguments, for instance by constructing minimizers of certain energy functionals. Departure of symmetry will then often enhance the value of these functionals. A good example to have in mind is the isoperimetric problem in classical geometry.

We now provide two classical examples of equations and their associated divergence free tensor field.

Example 1 We consider a smooth function $\psi : M \rightarrow \mathbb{R}$ (a scalar field) on our manifold and assume that it solves the wave equation

$$\square_g \psi = F, \tag{18}$$

for some smooth source $F : M \rightarrow \mathbb{R}$.

To any such ψ , we can associate its so called *energy-momentum*¹⁸ tensor (field)

$$T[\psi] = d\psi \otimes d\psi - \frac{1}{2} g(g(D\psi, D\psi)).$$

Then,

Lemma 5.1.

$$\operatorname{div} T = \square_g \psi \cdot d\psi,$$

or in coordinates,

$$D_\mu T^\mu_\nu [\psi] = \square_g \psi \cdot D_\nu \psi.$$

¹⁷See for instance [Chr00] as well as the easier presentation given in [Chr08].

¹⁸Also refers to sometimes as stress energy tensor.

In particular, T is divergence free provided that ψ is a solution of the homogeneous wave equation, i.e. $F = 0$ in (18).

Example 2: Maxwell equations. In Minkowski space in usual coordinates, they can be written¹⁹ as

$$\partial_\mu F^{\mu\nu} = j^\nu, \quad (19)$$

where j^ν is a divergence free vector field (the source term, sometimes called a 4-current) and F (the Faraday tensor) is a closed two form on \mathbb{R}^{n+1} (thus, by the Poincaré Lemma in \mathbb{R}^n , $F = dA$ for some one-form A).

Exercise 5.3. Derive from (19) the usual form of the Maxwell equations (involving the divergence and rotational operators in flat space) in terms of the electric and magnetic fields $E_i = F(\partial_{x^0}, \partial_{x^i})$ and $B_i = -\frac{1}{2}\epsilon_{iab}\delta^{aj}\delta^{bk}F(\partial_{x^j}, \partial_{x^k})$, where ϵ is the totally antisymmetric symbol such that $\epsilon_{123} = 1$.

We then introduce the *energy-momentum tensor* of electromagnetism given in components by

$$T_{\mu\nu}[F] = F_{\mu\sigma}F^\sigma{}_\nu - \frac{1}{4}\eta_{\mu\nu}(F_{\alpha\beta}F^{\alpha\beta}).$$

Exercise 5.4. Check that the equation of motion (19) implies that

$$\partial^\mu T_{\mu\nu} = F^\mu{}_\nu j^\nu.$$

In particular, in the vacuum $j = 0$ and we again have a divergence free symmetric tensor field.

To obtain the equations in an arbitrary Lorentzian manifold (M, g) , just replace all ∂_μ by D_μ and the Minkowski metric η by g . Thus, the Maxwell equations in the vacuum (no source term) for a closed two form F is simply given by

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

We end this section with another useful property of the energy-momentum tensors.

Proposition 5.1. Let X_1, X_2 be two timelike vectors such that $g(X_1, X_2) \leq 0$. Let $T[\psi]$ and $T[F]$ be the energy-momentum tensor fields as defined above associated to a either a smooth function or a 2-form. Then,

$$T[\psi](X_1, X_2) \geq 0, \quad T[F](X_1, X_2) \geq 0.$$

In physics, these positivity properties are interpreted as positivity of local energy densities. For pde applications, these positivity properties allow to get good estimates from the conservation laws. The positivity property also extends to causal vectors.

Exercise 5.5. 1. Prove the proposition. What about the equality cases?

2. We consider again Stokes theorem in a domain N . We assume that T is the energy-momentum tensor of a scalar ψ of Maxwell field F and that the support of ψ or F is contained in a region where ∂N is spacelike and orientable.

- Prove that the normal to ∂N is a timelike vector field on ∂N .
- Deduce that on each connected component of ∂N , $T[\psi](X_1, X_1)$ has a sign.

¹⁹As for the rest of the equations written in these lectures, all physical constants have been set to 1 for simplicity.

5.4 The Einstein tensor

Let us now try to construct general relativity. In particular, we are looking for an equation replacing Poisson's equation

$$\Delta\phi = 4\pi\rho.$$

Let T be the energy-momentum tensor associated with the matter fields as introduced in the previous section. Recall that in the above examples, T is $(0,2)$ symmetric tensor field and T is divergence free. Moreover, from Proposition 5.1, the energy-momentum tensor has some positivity property, which is analogous to $\rho \geq 0$.

We interpret T as the correct tensorial replacement for ρ . On the other hand, the replacement for ϕ should be constructed out of the metric. Moreover, it should be second order (as is $\Delta\phi$) in terms of derivatives of g . Thus, we want an equation under the form

$$G(g) = T,$$

where G is a symmetric $(0,2)$ tensor (it has to be independent of the choice of coordinates), constructed out of the first and second derivatives of g and G is divergence free. A first candidate would have been the curvature tensor, since it is a tensor and depends on the second derivatives of g , but it is 4 tensor. The Ricci tensor then seems like a good candidate and indeed, in the first draft, Einstein (wrongly) chose it. The trouble is that *Ricci* is not divergence free in general.

The correct choice is given in the next lemma.

Lemma 5.2. *The differential $dS(g)$ of the scalar curvature $S(g)$ verifies*

$$dS(g) = 2 \operatorname{div} Ric.$$

Equivalently, the Einstein tensor defined as

$$G = Ric(g) - \frac{1}{2}gS(g)$$

is a divergence free (symmetric) tensor field.

Proof. Note that since $Ric(g)$ is a symmetric $(0,2)$ tensor field, by definition, $\operatorname{div} Ric$ is the 1-form of components

$$\begin{aligned} \operatorname{div} Ric_\gamma &:= g^{\alpha\beta} D_\beta Ric_{\alpha\gamma} \\ &= D^\alpha Ric_{\alpha\gamma} \\ &= D_\alpha Ric^\alpha_\gamma \\ &= Ric^\alpha_{\gamma;\alpha}. \end{aligned}$$

We first express the second Bianchi identity in coordinates

$$R^\alpha_{\beta\gamma\delta;\rho} + R^\alpha_{\beta\rho\gamma;\delta} + R^\alpha_{\beta\delta\rho;\gamma} = 0$$

Contracting in α and ρ gives, using also the symmetries of R and the definition of the Ricci tensor, ‡

$$\begin{aligned} 0 &= R^\alpha_{\beta\gamma\delta;\alpha} + R^\alpha_{\beta\alpha\gamma;\delta} + R^\alpha_{\beta\delta\alpha;\gamma} \\ &= R^\alpha_{\beta\gamma\delta;\alpha} + Ric_{\beta\gamma;\delta} - Ric_{\beta\delta;\gamma}. \end{aligned}$$

Contract again in $\beta\delta$ (and use again the symmetries of R) to conclude. \square

6 The Einstein equations

The Einstein equations for a Lorentzian²⁰ manifold (M, g) are given by

$$Ric(g) - \frac{1}{2}gS(g) + \Lambda g = T,$$

where $\Lambda \in \mathbb{R}$ is a constant, typically called in the physics literature the cosmological constant, T is a symmetric, divergence free, 2-tensor, the energy-momentum tensor (think of T as a source term depending on the choices of matter models). Note that the term Λg is also divergence free, by virtue of the fact that $Dg = 0$.

For simplicity, let us assume that $T = 0$ and that $\dim M = 4$. Then, taking the trace of the Einstein equations, we obtain that $S(g) = 4\Lambda$ and the Einstein equations reduce to the *vacuum Einstein equation*

$$Ric(g) = \Lambda g.$$

A solution to the Einstein vacuum equations is then a Lorentzian manifold (M, g) solving the above equations. Some remarks:

1. The manifold M is part of the unknown. We are not just solving for g . This is similar to ode theory, where a solution to an ode is a couple (f, I) , where I is an interval and f a solution to the ode defined on I .
2. A word about the real constant Λ . In the first formulation of the equations, Einstein did not add the Λ term. Then, trying to construct explicit static solutions (which have a time symmetry) in the presence of matter, he needed to balance the gravitational force that attracts matter together with another mechanism that stretch matter. Adding $\Lambda > 0$ allowed to construct such solutions. The solution that Einstein constructed is however highly unstable, thus not quite physical and moreover does not match the observations that Hubble made in 1929. He called the introduction of the Λ is "biggest mistake" and decided to remove it from the equations. However, even if the solution of Einstein did not match the physical observations, the new mechanism coming from $\Lambda > 0$ was legit. It is responsible for the so-called "accelerated expansion" of the universe which nowadays a standard of cosmology. In other words, yes the solution of Einstein was unphysical but that does not necessarily mean the equations were bad. There could be other solutions with better properties. The current models in astrophysics and cosmology use $\Lambda = 0$ for an event at the scale of a galaxy or a star and $\Lambda > 0$ (but very small²¹ compared to all other constants, such as G and c) for a global model of the observed universe. Finally, solutions with $\Lambda < 0$ are extremely popular in High Energy Physics (the most quoted paper in HEP is about such type of solutions).
3. As we shall see, the Einstein equations are dynamical (hyperbolic) equations. Thus, solutions are typically constructed by solving initial value problem (but other constructions are possible such as characteristic initial value problem, analytic extensions...) in which case we will need to explain what are the data and in which sense our solutions agree with the data.

²⁰The study of the Riemannian version of these equations is also an interesting subject. Here, we will focus on the Lorentzian case.

²¹According to the English wikipedia article the current proposed value is $1.19 \times 10^{-52} m^{-2}$.

4. In the non-vacuum case ($T \neq 0$), even though T is a source, the typical case is that T is not given a priori. Instead, T is also obtained by solving a dynamical equation, which itself depends on (M, g) . Thus, one obtained a coupled system.

Example: The Einstein-scalar field system. Consider a function $\psi \in \mathcal{F}(M)$ and recall that $T[\psi]$ is given by

$$T[\psi] = d\psi \otimes d\psi - \frac{1}{2}g(g(D\psi, D\psi)),$$

where $D\psi$ denotes the gradient of ψ . Recall also that

$$\operatorname{div} T = D\psi \cdot \square_g \psi.$$

Thus, we consider the coupled problem

$$\begin{aligned} Ric(g) - \frac{1}{2}g + \Lambda g &= T[\psi], \\ \square_g \psi &= 0. \end{aligned}$$

Typical other sources term: electromagnetism, Vlasov fields (kinetic theory), fluids etc..

Above, we claimed that the Einstein equations are actually hyperbolic, in fact wave, equations. There are several different notions of hyperbolicity in the PDE literature. Here, we are going to prove first that, from the Einstein equations, one can extract systems of wave equations, that is to say equations of the form

$$\square_g u = F(u, Du) + B,$$

for some tensor fields u related to the metric (These equations may only hold locally, say in some coordinate patch of the manifold).

6.1 A model problem: the Maxwell equations

The difficulties in identifying the hyperbolic character of the Einstein equations are actually already present in a much simpler, and better known systems of PDEs: the Maxwell equations.

Recall that they can be written (here in Minkowski space for simplicity) as $\partial_\mu F^{\mu\nu} = j^\nu$ for a closed two form F .

How do we know that these are actually wave equations ?

We will give two answers²².

First, in the (E, B) formulation, it is well known that from the above system we can write wave equations of the form

$$\square E = S_E(\partial j), \quad \square B = S_B(\partial j).$$

²²A third formulation of hyperbolicity for the Maxwell equations can be obtained directly from the (E, B) formulation. Discarding the two divergence equations, the two evolution equations for E and B actually form a 1st order symmetric hyperbolic system in the sense of Friedrich. Interestingly, the second Bianchi equations for the curvature tensor can also be separated into constraints and evolution equations and the evolution part forms again a symmetric hyperbolic system. That there are so many analogies between the maxwell equations $D_\mu F^{\mu\nu}$ and the Bianchi equations is actually not that surprising, since the Maxwell equations can also be viewed as the Bianchi equations for a connection 1-form on a principal bundle whose curvature is actually given by F .

Second, if A is such that $F = dA$, then the Maxwell equations written in terms of A takes the form

$$\square A_\nu - \partial_\nu \partial_\mu A^\mu = j^\mu.$$

Now this equation is not a wave equation because of the second term in the left-hand side. However, recall that given any scalar function χ , we have $F = d(A + d\chi)$. Claim: if χ solves $\square\chi = -\partial_\mu A^\mu$, then $A' = A + d\chi$ solves $\partial_\mu (A')^\mu = 0$. The new A' then solves a wave equation. Choosing χ is called making *gauge* choice (the one where $\partial_\mu A^\mu = 0$ is called the Lorentz gauge, another popular one is the Coulomb gauge, where only the spatial divergence of A is required to vanish $\partial_i A^i = 0$).

For the Einstein equations, we claim that similarly, we can write wave equations directly by essentially differentiating the equations (thus obtaining wave equations for R) or, that by making a gauge choice, we can write wave equations for g . The gauge freedom in this case is simply the choice of local coordinates on our manifold.

6.2 Hyperbolicity I: Wave equations satisfied by the Riemann tensor

Proposition 6.1. *Let (M, g) be a solution to the vacuum Einstein equations $\text{Ric}(g) = \Lambda g$. Then, the Riemann curvature tensor solves an equation of the form*

$$\square_g R = R \star R.$$

where $R \star R$ is obtained from $R \otimes R$ by contractions.

Proof. Taking the trace of the Bianchi equations, we obtain

$$D^\sigma R_{\alpha\beta\sigma\nu} + D_\beta R_{\alpha\nu} - D_\alpha R_{\beta\nu} = 0.$$

Thanks to the Einstein equations and the fact that $Dg = 0$, by definition of the Levi-Civita connection, the last two terms vanishes and we obtain that the Riemann curvature tensor is divergence free

$$D^\sigma R_{\alpha\beta\sigma\nu} = 0.$$

Recalling again the Bianchi equation

$$D_\sigma R_{\alpha\beta\mu\nu} + D_\beta R_{\sigma\alpha\mu\nu} + D_\alpha R_{\beta\sigma\mu\nu} = 0,$$

we can now compute

$$\begin{aligned} \square_g R_{\alpha\beta\mu\nu} &= g^{\rho\sigma} D_\rho D_\sigma R_{\alpha\beta\mu\nu}, \\ &= -g^{\rho\sigma} D_\rho D_\beta R_{\sigma\alpha\mu\nu} - g^{\rho\sigma} D_\rho D_\alpha R_{\beta\sigma\mu\nu} \\ &= -g^{\rho\sigma} [D_\rho, D_\beta] R_{\sigma\alpha\mu\nu} - g^{\rho\sigma} [D_\rho, D_\alpha] R_{\beta\sigma\mu\nu} - D_\beta g^{\rho\sigma} D_\rho R_{\sigma\alpha\mu\nu} - D_\alpha g^{\rho\sigma} D_\rho R_{\beta\sigma\mu\nu} \\ &= -g^{\rho\sigma} [D_\rho, D_\beta] R_{\sigma\alpha\mu\nu} - g^{\rho\sigma} [D_\rho, D_\alpha] R_{\beta\sigma\mu\nu} \\ &= R \star R, \end{aligned}$$

using Exercise 4.2. □

Remark 6.1. *Given (M, g) and appropriate data, it is not hard to construct local solutions to a non-linear system equations of the form*

$$\square_g \phi = \phi \star \phi.$$

However, this scheme does not really solves the Einstein equations in that R and g are of course not independent. We could imagine running a scheme of the form

1. Given appropriate data for g , we construct first a candidate for g^1 (for instance extending g locally to simply be a Lorentzian metric, whatever it is, as long as it agrees with the data).
2. Then, we compute data for R^1 and solve the associated wave equation.
3. We then try to construct g^2 , which would have R^1 as a curvature tensor (possibly modulo some error terms) and should still agree with the data.
4. We then run again our iterative scheme and hope to prove convergence.

There are many difficulties in doing so and in practice, the standard (and "fastest") proof of local existence of solutions to the Einstein equations does not use this formulation.

However, estimating the curvature first and returning to g has been a very successful strategy to address global problems for the Einstein equations, where one tries to understand the long time dynamics of the solutions. The important point which is exploited there is that the wave equation satisfied by R is independent of anything: it is gauge free.

6.3 Hyperbolicity II: Wave coordinates

Definition 6.1. Let (M, g) be a Lorentzian manifold. A system of wave coordinates²³ is by definition a coordinate system (x^α) such that for each α , the function x^α is itself a solution of the wave equation

$$\square_g x^\alpha = 0.$$

Proposition 6.2. Let (x^α) be wave coordinates. Then, for all α ,

$$g^{\rho\sigma} \Gamma_{\rho\sigma}^\alpha = 0.$$

Proof. We have for any function ψ

$$\begin{aligned} \square_g \psi &= g^{\rho\sigma} D_\rho \partial_\sigma(\psi) \\ &= g^{\rho\sigma} (\partial_\rho \partial_\sigma(\psi) - \Gamma_{\rho\sigma}^\gamma \partial_\gamma(\psi)). \end{aligned}$$

Thus,

$$\square_g x^\alpha = g^{\rho\sigma} (\Gamma_{\rho\sigma}^\gamma \delta_\gamma^\alpha) = g^{\rho\sigma} \Gamma_{\rho\sigma}^\alpha.$$

□

Proposition 6.3. In wave coordinates, the wave operator \square_g reduces to its principal part and the vacuum Einstein equations $\text{Ric}(g) = \Lambda g$ takes the form

$$\square_g g_{\alpha\beta} = Q_{\alpha\beta}(\partial g, \partial g) - 2\Lambda g,$$

where $Q_{\alpha\beta}(\partial g, \partial g)$ is quadratic in the first derivatives of g .

²³In the Riemannian case, similar coordinates can be constructed, replacing the wave operator by the Laplace-Beltrami operator. These are called harmonic coordinates and this name is sometimes also used in the Lorentzian case.

Proof. Recall from equation (17), that the components of the curvature tensor are given by

$$R^\alpha{}_{\beta\gamma\delta} = \partial_\gamma \Gamma^\alpha_{\delta\beta} - \partial_\delta \Gamma^\alpha_{\gamma\beta} + \Gamma^\sigma_{\delta\beta} \Gamma^\alpha_{\gamma\sigma} - \Gamma^\sigma_{\gamma\beta} \Gamma^\alpha_{\delta\sigma},$$

so that

$$Ric_{\beta\delta} = \partial_\alpha \Gamma^\alpha_{\delta\beta} - \partial_\delta \Gamma^\alpha_{\alpha\beta} + Q_{\alpha\beta}(\partial g, \partial g), \quad (20)$$

where, since we are interested only in the terms containing second order derivatives, we denote by a generic Q any terms quadratic in ∂g (and not dependent on $\partial^2 g$). Below, we continue to write any such terms by Q but it may change from line to line.

Define $\Gamma_{\alpha\beta\gamma}$ as

$$\Gamma_{\alpha\beta\gamma} := \frac{1}{2} (g_{\gamma\beta,\alpha} + g_{\alpha\beta,\gamma} - g_{\alpha\gamma,\beta}),$$

so that

$$\Gamma^\mu_{\alpha\beta} = g^{\mu\gamma} \Gamma_{\alpha\gamma\beta}.$$

With this notation,

$$\begin{aligned} \partial_\nu \Gamma^\nu_{\mu\rho} &= \partial_\nu (g^{\nu\beta}) \Gamma_{\mu\beta\rho} + g^{\nu\beta} \partial_\nu \Gamma_{\mu\beta\rho}, \\ \partial_\mu \Gamma^\nu_{\nu\rho} &= \partial_\mu (g^{\nu\beta}) \Gamma_{\nu\beta\rho} + g^{\nu\beta} \partial_\mu \Gamma_{\nu\beta\rho}. \end{aligned}$$

Thus, from (20), in order to compute $Ric_{\mu\rho}$, we need to compute

$$g^{\nu\beta} \partial_\nu \Gamma_{\mu\beta\rho} - g^{\nu\beta} \partial_\mu \Gamma_{\nu\beta\rho}.$$

Moreover,

$$\begin{aligned} 2\partial_\nu \Gamma_{\mu\beta\rho} &= g_{\beta\rho,\mu\nu} + g_{\mu\beta,\rho\nu} - g_{\mu\rho,\beta\nu}, \\ 2\partial_\mu \Gamma_{\nu\beta\rho} &= g_{\beta\rho,\nu\mu} + g_{\nu\beta,\rho\mu} - g_{\nu\rho,\beta\mu}. \end{aligned}$$

so that

$$2g^{\nu\beta} (\partial_\nu \Gamma_{\mu\beta\rho} - \partial_\mu \Gamma_{\nu\beta\rho}) = -g^{\nu\beta} g_{\mu\rho,\beta\nu} + g^{\nu\beta} (g_{\mu\beta,\rho\nu} - g_{\nu\beta,\rho\mu} + g_{\nu\rho,\beta\mu}).$$

The first term on the RHS is $-\square_g g_{\mu\rho}$, since from the wave coordinate conditions, \square_g reduce to its principal part. Thus, we would like the second term on the RHS to be a Q term. This means that the second order derivatives terms must cancel. For these to happen, we need to use again the wave coordinate condition. In terms of the $\Gamma_{\beta\mu\nu}$, the wave coordinate conditions reads

$$g^{\mu\beta} \Gamma_{\beta\mu\nu} = 0.$$

Thus, we try to rewrite the second terms in terms of the Γ .

For this, since $g^{\nu\beta}$ is symmetric in $(\nu\beta)$, we rewrite the second term on the RHS as

$$\begin{aligned} g^{\nu\beta} (g_{\mu\beta,\rho\nu} - g_{\nu\beta,\rho\mu} + g_{\nu\rho,\beta\mu}) &= \frac{1}{2} g^{\nu\beta} (g_{\mu\beta,\rho\nu} - g_{\nu\beta,\rho\mu} + g_{\nu\rho,\beta\mu} + g_{\mu\nu,\rho\beta} - g_{\beta\nu,\rho\mu} + g_{\beta\rho,\nu\mu}) \\ &= \frac{1}{2} g^{\nu\beta} (g_{\mu\beta,\nu\rho} + g_{\mu\nu,\rho\beta} - g_{\beta\nu,\rho\mu} + g_{\beta\rho,\nu\mu} + g_{\nu\rho,\beta\mu} - g_{\nu\beta,\rho\mu}) \\ &= \frac{1}{2} g^{\nu\beta} (g_{\mu\beta,\nu\rho} + g_{\mu\nu,\beta\rho} - g_{\beta\nu,\mu\rho} + g_{\beta\rho,\nu\mu} + g_{\nu\rho,\beta\mu} - g_{\nu\beta,\rho\mu}) \\ &= \frac{1}{2} g^{\nu\beta} (\partial_\rho \Gamma_{\beta\mu\nu} + \partial_\mu \Gamma_{\beta\rho\nu}). \end{aligned}$$

$$\begin{aligned} g^{\nu\beta} (g_{\mu\beta,\rho\nu} - g_{\nu\beta,\rho\mu} + g_{\nu\rho,\beta\mu}) &= \frac{1}{2} (\partial_\mu (g^{\nu\beta} \Gamma_{\beta\rho\nu}) + \partial_\rho (g^{\nu\beta} \Gamma_{\beta\mu\nu})) - \frac{1}{2} (\partial_\mu (g^{\nu\beta}) \Gamma_{\beta\rho\nu} + \partial_\rho (g^{\nu\beta}) \Gamma_{\beta\mu\nu}) \\ &= Q, \end{aligned}$$

using the wave coordinate condition.

Thus, in wave coordinates, we have

$$Ric(g)_{\alpha\beta} = -\frac{1}{2} \square_g g_{\alpha\beta} + Q_{\alpha\beta},$$

so that the vacuum Einstein equations imply that

$$\square_g g = -2\Lambda g + Q.$$

□

Exercise 6.1. *The lower order terms play no role as far as understanding the wave character of the Einstein equations but they are important for more precise statements, for instance, to describe the global behaviour of solutions.*

Prove that, in wave coordinates,

$$R_{\mu\rho} = -\frac{1}{2} g^{\alpha\beta} \partial_{\alpha\beta} g_{\mu\rho} + g^{\alpha\beta} g^{\gamma\delta} (\Gamma_{\alpha\gamma\mu} \Gamma_{\beta\delta\rho} + \Gamma_{\alpha\gamma\mu} \Gamma_{\beta\rho\delta} + \Gamma_{\alpha\gamma\rho} \Gamma_{\beta\mu\delta}).$$

Remark 6.2. *The existence of local wave coordinates will be relatively easy once we know how to solve the wave equation.*

Remark 6.3. *We have just proved that the vacuum Einstein equations take the form*

$$g^{\alpha\beta} \partial_{x^\alpha} \partial_{x^\beta} g = Q(\partial g, \partial g) + C_\Lambda g \quad (21)$$

in wave coordinates. The equations (21) form a standard system of quasilinear wave equations, which can be solved locally by standard methods, with initial data $g, \partial_{x^0} g$ given on some $x^0 = \text{const}$ slice provided for instance that $g^{00} < 0$ and $(g^{ij}) > 0$ is definite positive initially. See for instance [Sog95a, Section I.4]. On the other hand, to get a local solution to the full Einstein equations, one then needs to verify, after solving the equations, that the coordinates are then wave coordinates for the newly constructed metric. Moreover, one needs to explain where the data for g is coming from and explain what do we mean by a $x^0 = \text{const}$ slice while we have not yet constructed M ! Finally, in general, one needs several system of coordinates to cover our manifolds and therefore, one obtains several local solutions. One thus needs some way to patch all these solutions together to get a complete development of the data.

Remark 6.4. *For more on the gauge invariance of the Einstein or Maxwell equations and its relation to the hyperbolicity of the equations, one can consult [Chr08, Chapter 2.1.1].*

7 Semi-Riemannian submanifolds

Recall the definition of a vector field on a map, Definition 3.13 as well as the definition of a submanifold Definition 2.10. If P is a submanifold of M , then we call a vector field on the inclusion map $j : P \rightarrow M$ a P -vector field. We denote by $\Gamma(P, M)$ the set of all such vector fields. Note that $\Gamma(P)$, the set of vector fields of P can be viewed as a subset of $\Gamma(P, M)$.

7.1 The induced connection

Definition 7.1. Let P be a submanifold of M . Then the induced connection ∇ on P is a map

$$\nabla : \Gamma(P) \times \Gamma(P, M) \rightarrow \Gamma(P, M)$$

defined as follows. For any $V \in \Gamma(P)$ and $Z \in \Gamma(P, M)$, let V' and Z' be smooth local extensions to M of V and Z . Then we define $\nabla_V Z$ as the restriction of $D_{V'} Z'$ to P .

Exercise 7.1. 1. With the above notation, check that the restriction does not depend on the choice of extensions, and thus that the induced connection is well defined.

2. Redefine the induced connection using the pull-back bundle $j^*(TM)$.

We have immediately from the definition.

Lemma 7.1. Let ∇ be the induced connection on P . Then, if $V, W \in \Gamma(P)$ and $X, Y \in \Gamma(P, M)$, we have

1. $\nabla_V X$ is $\mathcal{F}(M)$ -linear in V and \mathbb{R} linear in X

2.

$$\nabla_V(fX) = V(f)X + f\nabla_V X,$$

3.

$$[V, W] = \nabla_V W - \nabla_W V,$$

4.

$$Vg(X, Y) = g(\nabla_V X, Y) + g(X, \nabla_V Y).$$

7.2 Semi-Riemannian submanifold and the Levi-Civita connection of the induced metric

The above definition make sense for any submanifold of M . Recall that in general, the pullback of the metric on a submanifold is not necessarily a semi-Riemannian manifold cf 3.4.

Definition 7.2. Let P be a submanifold of a semi-Riemannian manifold (M, g) and $j : P \rightarrow M$ denote the inclusion map. If $j^*(g)$ is a metric tensor on P , we say that P is semi-Riemannian submanifold of M .

$j^*(g)$ is called the induced metric or *first fundamental form*.

A submanifold of a Riemannian or Lorentzian manifold is called spacelike if the induced metric is Riemannian, timelike if the induced metric is Lorentzian, null if the induced metric²⁴ is degenerate. In the case of a submanifold of codimension 1, we say that we have a spacelike, null, or timelike hypersurface.

If P is a semi-Riemannian submanifold of M , then we have, for each $p \in M$ a direct sum decomposition

$$T_p M = T_p P + (T_p P)^\perp,$$

where

$$(T_p P)^\perp = \{v \in T_p M : g(v, x) = 0, \forall x \in P\}.$$

²⁴This is of course a slight abuse of language, since the induced metric in that case is not a metric.

Vectors in $(T_p P)^\perp$ are said to be normal to P while those of $T_p P$ are of course tangent to P . The resulting orthogonal projection will be denoted as

$$\tan : T_p M \rightarrow T_p P$$

and

$$\text{nor} : T_p M \rightarrow (T_p P)^\perp.$$

A vector field Z in $\Gamma(P, M)$ is normal (respectively tangent) to M provided that for each $p \in P$, Z_p is normal (respectively tangent) to P . We denote the set of all normal vector fields by $\Gamma(P)^\perp$, called the *normale bundle*.

For any Z in $\Gamma(P, M)$, we can apply \tan and nor at each point to construct the tangential and normal part of Z .

In general, if V, W are both tangent to P , it does not follow that $\nabla_V W$ is tangent to P . In the case of a semi-Riemannian submanifold, we have however

Lemma 7.2. *If $V, W \in \Gamma(P)$ and ${}^P D$ denotes the Levi-Civita connection associated to the induced metric on P , then*

$$\tan \nabla_V W = {}^P D_V W.$$

Remark 7.1. *In particular, we could reconstruct the properties of the Levi-Civita connection for surfaces (or any submanifolds of \mathbb{R}^n) in this way. Indeed, let S be a surface in \mathbb{R}^3 . Then, we know how to define the induced metric on S . Moreover, for $V, W \in \Gamma(P)$, we can consider extensions \tilde{V}, \tilde{W} to the whole of \mathbb{R}^n . Then, $\nabla_{\tilde{V}} \tilde{W}$ is given in Cartesian coordinates by $\tilde{V}^\alpha (\tilde{W}^\beta)$ and we can restrict to the surface as explained above. In other words, the formula just given together with basic calculus in \mathbb{R}^n can be seen as justification for the definition of the Levi-Civita connection.*

Proof. Recall the Koszul formula for the Levi-Civita connection

$$2g(D_V W, X) = Vg(W, X) + Wg(X, V) - Xg(V, W) - g(V, [W, X]) + g(W, [X, V]) + g(X, [V, W]).$$

Suppose that V, W, X are all vector fields on P and consider extensions $\tilde{V}, \tilde{W}, \tilde{X}$ to M . We have

$$2g(D_{\tilde{V}} \tilde{W}, \tilde{X}) = \tilde{V}g(\tilde{W}, \tilde{X}) + \tilde{W}g(\tilde{X}, \tilde{V}) - \tilde{X}g(\tilde{V}, \tilde{W}) - g(\tilde{V}, [\tilde{W}, \tilde{X}]) + g(\tilde{W}, [\tilde{X}, \tilde{V}]) + g(\tilde{X}, [\tilde{V}, \tilde{W}]).$$

Restricting to P , we obtain the equality

$$\begin{aligned} 2g(\nabla_V W, X) &= Vg(W, X) + Wg(X, V) - Xg(V, W) - g(V, [W, X]) + g(W, [X, V]) + g(X, [V, W]) \\ &= Vh(W, X) + Wh(X, V) - Xh(V, W) - h(V, [W, X]) + h(W, [X, V]) + h(X, [V, W]), \end{aligned}$$

where h is the induced metric on P . Since X is tangent to P , we have

$$2g(\nabla_V W, X) = 2g(\tan \nabla_V W, X).$$

It follows that the map

$$(V, W) \rightarrow \tan \nabla_V W$$

1. defines a connection, 2. satisfies the Koszul formula for the induced metric. By uniqueness, it must agree with the Levi-Civita connection of h . \square

In the rest of this section, P will denote a semi-Riemannian submanifold of M .

7.3 The shape tensor

For the normal part, we have

Lemma 7.3. *The function $\Pi : \Gamma(P) \times \Gamma(P) \rightarrow \Gamma(P)^\perp$ given by*

$$\Pi(V, W) = \text{nor} \nabla_V W$$

defines a symmetric tensor (in the sense of tensor bundles), i.e. it is symmetric and $\mathcal{F}(P)$ -bilinear.

Proof. The tensorial part follows easily from the definition. For the symmetric part, we have

$$\Pi(V, W) - \Pi(W, V) = \text{nor}(\nabla_V W - \nabla_W V) \quad (22)$$

$$= \text{nor}[V, W] = 0. \quad (23)$$

□

From the above, we have

$$\nabla_V W = {}^P D_V W + \Pi(V, W).$$

Note that in the case of (non-characteristic) hypersurfaces, the normal bundle is one dimensional²⁵ in the sense that every normal vector to P must be propotional to a unit normal. It follows that we can write $\Pi(V, W) = k(V, W)N$, where N is a unit normal and k is a $(0, 2)$ -tensor field. The tensor Π is called the shape tensor or second fundamental form, this last name being also used for the tensor k .

7.4 The Gauss equation

The relation between the curvature of P and that of M is given by the Gauss equation

Theorem 7.1. *Let ${}^P R$ denote the Riemann curvature tensor of P endowed the induced metric h . Then, for $V, W, X, Y \in \Gamma(P)$*

$$h({}^P R(V, W)X, Y) = g(R(V, W)X, Y) + g(\Pi(V, X), \Pi(W, Y)) - g(\Pi(V, Y), \Pi(W, X)),$$

where $R(V, W)X$ denotes (by a small abuse of language) $R(\bar{V}, \bar{W})\bar{X}$ restricted to P , where $\bar{V}, \bar{W}, \bar{X}$ are extensions to M of V, W, X .

Proof. Since the equation is tensorial, we can assume wlog that $[V, W] = 0$. With a small abuse of notations, we consider extensions of all vector fields to M and denote them by the same letter.

We can write schematically

$$g(R(V, W)X, Y) = (VW) - (WV),$$

where

$$VW = g(D_V D_W X, Y).$$

²⁵Strictly speaking it is a $(n - 1) + 1$ dimensional manifold.

Let $Z = D_W X$. By definition Z is a vector field on M and its restriction to P is a vector field over P . By definition, this vector field is in fact $\nabla_W X$. Moreover, the restriction of $D_V Z$ to P is also by definition $\nabla_V Z$. Thus, on P , we have

$$\begin{aligned}
VW &= g(\nabla_V \nabla_W X, Y) \\
&= g(\nabla_V (D_W^P X + \Pi(W, X)), Y) \\
&= g(\tan \nabla_V D_W^P X, Y) + g(\nabla_V \Pi(W, X), Y) \\
&= h(D_V^P D_W^P X, Y) + Vg(\Pi(W, X), Y) - g(\Pi(W, X), \nabla_V Y) \\
&= h(D_V^P D_W^P X, Y) - g(\Pi(W, X), \Pi(V, Y)).
\end{aligned}$$

Computing the difference $(VW) - (WV)$, we get the result. \square

7.5 The normal connection and the Codazzi equation

In the previous section, we considered the geometry of vectors tangent to P . Here, we will consider the geometry of vectors normal to P .

Definition 7.3. *The normal connection of P is the map*

$$\nabla^\perp : \Gamma(P) \times \Gamma(P)^\perp \rightarrow \Gamma(P)^\perp$$

defined by

$$\nabla_V^\perp Z = \text{nor } \nabla_V Z,$$

where $\nabla_V Z$ is the induced covariant derivative and $\Gamma(P)^\perp$ is considered as a subset of $\Gamma(P, M)$.

$\nabla_V^\perp Z$ is called the normal covariant derivative of Z with respect to V . The following properties are immediate consequence of the definition.

Lemma 7.4. *If $V \in \Gamma(P)$ and $Y, Z \in \Gamma(P)^\perp$, then*

1. $\nabla_V^\perp Z$ is $\mathcal{F}(M)$ -linear in V and \mathbb{R} -linear in Z
- 2.

$$\nabla_V^\perp (fZ) = V(f)Z + f\nabla_V^\perp Z,$$

- 3.

$$Vg(Y, Z) = g(\nabla_V^\perp Y, Z) + g(Y, \nabla_V^\perp Z).$$

Since the shape tensor is a tensor valued in $\Gamma(P)^\perp$, we cannot apply our standard definition of covariant derivative to it. Instead, we will use the following definition.

Definition 7.4. *Let Π be the shape tensor of P and V, X, Y three vector fields in $\Gamma(P)$. We define the covariant derivative of Π with respect to V by*

$${}^P D_V \Pi(X, Y) = \nabla_V^\perp (\Pi(X, Y)) - \Pi({}^P D_V X, Y) - \Pi(X, {}^P D_V Y).$$

As $\Pi, {}^P D_V \Pi(X, Y)$ is a symmetric tensor with values in $\Gamma(P)^\perp$.

Exercise 7.2. *Let P be a semi-Riemannian hypersurface with unit normal N (N could be defined locally only if P is non-orientable) i.e. $g(N, N) = \pm 1 = \epsilon$.*

1. *Prove that $\Pi(X, Y) = k(X, Y)N$, with $k(X, Y) = \epsilon g(N, \nabla_X Y)$.*

2. Prove that for any vector field $V \in \Gamma(P)$, $\nabla_V N \in \Gamma(P)$.
3. Prove that $\nabla^\perp(\Pi(X, Y)) = V(k(X, Y))N$.
4. Prove that $({}^P D_V \Pi)(X, Y) = ({}^P D_V k)(X, Y)N$.

The Gauss equation describes $\text{nor } R(V, W)X$, where V, W, X where are all tangent to P in terms of the intrinsic curvature of P and its shape tensor. The Codazzi equation completes this information by providing $\text{nor } R(V, W)X$.

Proposition 7.1. For $V, W, X \in \Gamma(P)$, we have

$$\text{nor } R(V, W)X = -({}^P D_V \Pi)(W, X) + ({}^P D_V \Pi)(V, X).$$

7.6 The constraint equations

From the Gauss-Codazzi equations, one obtain the following

Proposition 7.2. Let (M, g) be a Lorentzian manifold and P a spacelike hypersurface. We assume that P admit a unit normal N . Then, on P , we have

$$G(N, N) = \frac{1}{2} {}^P S - k_{ij} k^{ij} + (\text{tr}_h k)^2, \quad (24)$$

$$G(N, v) = {}^P D^j k_{ki} - {}^P D_i (\text{tr}_h k) v^i. \quad (25)$$

7.7 Normal frame field

Definition 7.5. Given a semi-Riemannian submanifold P of dimension r and $\mathcal{U} \subset P$ an open submanifold of P a (local) normal frame field in \mathcal{U} is an orthonormal family of $n - r$ vector fields in $\Gamma(U)^\perp$.

Example: Consider a submanifold P in M and an adapted local coordinate system (x^α) in which P is given locally by the equations $x^1 = 0, x^2 = 0$. The vector fields $\partial_{x^k}, k > 2$ then span (locally) the tangent bundle of P . A vector field X is normal to P provided $g(X, \partial_{x^k}) = 0$ for any $k > 2$. In coordinates, this reads as the conditions

$$X^\alpha g_{\alpha k} = 0.$$

In another words, the one form $X_\flat = X^\alpha g_{\alpha\beta} dx^\beta$ is of the form $X_\flat = a_1 dx^1 + a_2 dx^2$ and X is of the form $X = g^{\alpha 1} a_1 \partial_{x^\alpha} + g^{\alpha 2} a_2 \partial_{x^\alpha}$. Let now

$$\begin{aligned} X_1 &= g^{\alpha 1} a_1 \partial_{x^\alpha} + g^{\alpha 2} a_2 \partial_{x^\alpha} \\ X_2 &= g^{\alpha 1} b_1 \partial_{x^\alpha} + g^{\alpha 2} b_2 \partial_{x^\alpha}. \end{aligned}$$

(X_1, X_2) is orthonormal provided the equations

$$\begin{aligned} g^{11} a_1 b_1 + g^{22} a_2 b_2 + g^{21} a_1 b_2 + g^{12} a_2 b_1 &= 0 \\ g^{11} a_1^2 + g^{22} a_2^2 + 2g^{21} a_1 a_2 &= \pm \\ g^{11} b_1^2 + g^{22} b_2^2 + 2g^{21} b_1 b_2 &= \pm \end{aligned}$$

are satisfied. If for instance $g^{11} \neq 0$ and we are in the Riemannian case, then the signs on the RHS are all + and we can take $a_1 = (g^{11})^{-1/2}, b_2 = (g^{11})^{1/2} / \sqrt{\det g}, b_1 = -g^{12} / g^{11} b_2$ and obtain a normal frame field.

More generally, consider a submanifold P and an adapted local coordinate system (x^α) in which P is given locally by the equations $x^1 = x^2 = \dots = x^r = 0$. As above, the one-forms which are normal to P are of the form

$$\omega = \sum_{i=r+1}^n a_i dx^i.$$

Consider a non-zero one-form ω^1 of the above form. Since P is assumed to be non-characteristic, we can choose ω^1 to be a non-null one-form and then normalized it. Similar to the Gram-Schmidt, we can then consider an independent one-form ω^2 , such that ω^2 is also non-null and normalized. Repeating we then obtain an orthonormal family of one-forms and the corresponding family of vector fields then form a normal frame field. In particular, every semi-Riemannian submanifold admits local normal frame fields.

Exercise 7.3. Generalize the notion of normal frame field by considering null frames.

7.8 The normal bundle

Recall the definition of a vector bundle over M , definition 2.19. Let P be a semi-Riemannian manifold of M and let NP be the set of all points (p, v_p) where $p \in P$ and $v_p \in T_p P^\perp$. Let $\pi : NP \rightarrow P$ be the natural projection. For each p in P , there exists a normal frame field (E_1, \dots, E_k) defined in a neighborhood \mathcal{U} of p in P . NP can be given the structure of a manifold, for instance, we can use locally a coordinate system in an neighborhood $\mathcal{V} \subset \mathcal{U}$ using a local coordinate system (\mathcal{V}, x^α) of the base manifold P together with the coordinate systems induced by the E_i on each $T_p P^\perp$.

Define a map $\phi : \mathcal{U} \times \mathbb{R}^k \rightarrow NP$ by

$$\phi(q, a^1, \dots, a^k) = \sum_{i=1}^k a^i [E_i]_q.$$

This defines a smooth map which is a diffeomorphism from $\mathcal{U} \times \mathbb{R}^k$ to $\pi^{-1}(\mathcal{U})$ and makes NP a vector bundle over P .

The sections of NP are then the vector fields in $\Gamma(P)^\perp$.

Definition 7.6. For any p in P let \mathcal{D}_p^\perp be the set of all vectors v in $T_p P^\perp$ such that the geodesic starting at p with initial velocity v is well defined up to parameter 1. Let $\mathcal{D}^\perp = \bigcup_{p \in P} \mathcal{D}_p^\perp$. Then, we define the normal exponential map as

$$\begin{aligned} \exp^\perp : \mathcal{D}_p^\perp \subset NP &\rightarrow M \\ (p, v) &\rightarrow \gamma_v(1). \end{aligned}$$

As for the exponential map, \mathcal{D} is an open set containing all points $(p, 0_p)$ and \exp^\perp is a smooth map. Let Z be the set of all points $(p, 0_p)$.

Definition 7.7. A neighborhood \mathcal{U} of P is said to be normal if \mathcal{U}^\perp is the image by \exp^\perp of a neighborhood of Z in NP .

Lemma 7.5. If $p \in P$, then \exp^\perp carries some neighborhood of $(p, 0_p)$ in NP diffeomorphically onto a neighborhood of p in M .

Proof. Z is a submanifold of NP (of dimension that of P). Thus, \exp^\perp restricts to a smooth map from Z to P . This is in fact a diffeomorphism of inverse given by the zero vector field. Thus, $d\exp^\perp_{(p,0_p)}$ is one-to-one from $T_{(p,0_p)}Z$ to T_pP . Consider now $d\exp^\perp_{(p,0_p)}$ restricted to $T_{(p,0_p)}T_pP^\perp$. As in the case of the exponential from a point, this is the canonical isomorphism $T_{(p,0_p)}T_pP^\perp \simeq T_pP^\perp$. Since we have the direct sum $T_{(p,0_p)}\mathcal{D}_p^\perp = T_{(p,0_p)}Z \oplus T_{(p,0_p)}T_pP^\perp$, it follows that $d\exp^\perp_{(p,0_p)}$ is a linear isomorphism and we can apply the inverse function theorem. \square

We then prove

Proposition 7.3. *Every semi-Riemannian submanifold P has a normal neighborhood.*

Proof. For each $p \in P$, let \mathcal{N}_p be a neighborhood of $(p, 0_p)$ in NP on which \exp is a diffeomorphism. By shrinking \mathcal{N}_p , we can arrange that if $\exp(v) \in P$ for $v \in \mathcal{N}_p$, then $v = 0$. Moreover, we can assume that \mathcal{N}_p has compact closure. Let $\mathcal{N} = \bigcup \mathcal{N}_p$. One easily have that for any $K \subset P$ compact, $\pi^{-1} \cap \mathcal{N}$ is compact. Note that by construction, if $v \neq w$ and $\exp(v) = \exp(w)$, then $v = w = 0$.

Consider a decreasing sequence of neighborhood of Z of the form $\mathcal{N} = \mathcal{N}_0 \supset \mathcal{N}_1 \subset \dots$ such that $\bigcap_i \mathcal{N}_i = Z$. For any K , there is an i , such that \exp is one-to-one on $E_j = \pi^{-1}(K) \cap \mathcal{N}_i \cup Z$. For if not, there exists for all j , vectors v_j and w_j such that $\exp(v_j) = \exp(w_j)$. By the above, v_j and w_j are non zero, so they must lie in $\pi^{-1}(K) \cap \mathcal{N}_i \subset \pi^{-1}(K) \cap \mathcal{N}$. Taking subsequence and using the continuity of the exponential, we reach a contradiction. By second countability, P contains an increasing sequence of compact sets K_i such that $K_i \subset \text{int}K_{i+1}$ and $\bigcup K_i = P$. For K_1 , we construct E_i as above. The rest of the proof follows by an induction argument left as an exercise. \square

8 Local geometry

8.1 Two parameters maps

We start by some preliminaries on two parameter maps.

Definition 8.1. *Let $\mathcal{D} \subset \mathbb{R}^2$ be open and such that \mathcal{D} intersected with any vertical or horizontal line is an interval (possibly empty)²⁶. A two parameter map is a smooth map*

$$\begin{aligned} \kappa: \mathcal{D} &\rightarrow M \\ (u, v) &\rightarrow \kappa(u, v). \end{aligned}$$

κ is thus composed of two intertwined family of curves

- For any fixed v_0 such that $\{v = v_0\} \cap \mathcal{D} \neq \emptyset$, the u -parameter curve $v = v_0$ of κ is $u \mapsto \kappa(u, v_0)$.
- For any fixed u_0 such that $\{u = u_0\} \cap \mathcal{D} \neq \emptyset$, the v -parameter curve $u = u_0$ of κ is $v \mapsto \kappa(u_0, v)$.

²⁶A wedge (kind of boomerang) obtained from instance from the points $(10, 0)$, $(0, 0)$, $(0, 10)$ and $(1, 1)$ is an example of such \mathcal{D} which is non-convex.

At any $(u, v) \in \mathcal{D}$, we can compute $d\kappa_{(u,v)} : T_{(u,v)}\mathcal{D} \rightarrow T_{\kappa(u,v)}M$. The maps

$$(u, v) \rightarrow \kappa_u := d\kappa(\partial_u) \in T_{\kappa(u,v)}M$$

and

$$(u, v) \rightarrow \kappa_v := d\kappa(\partial_v) \in T_{\kappa(u,v)}M$$

are then vector fields on the map κ . Moreover, one has of course that $\kappa_u(u_0, v_0)$ coincides with the tangent vector of the $v = v_0$ curve at u_0 and $\kappa_v(u_0, v_0)$ is the tangent vector of the $u = u_0$ curve at v_0 .

If the image of κ lies in the domain of a coordinate system (x^α) then its coordinate functions $\kappa^\alpha = x^\alpha \circ \kappa$ are smooth real functions on \mathcal{D} . Moreover,

$$\kappa_u = \partial_u(\kappa^\alpha)\partial_{x^\alpha}, \quad \kappa_v = \partial_v(\kappa^\alpha)\partial_{x^\alpha}.$$

Let Z be a smooth vector field on the map κ , i.e.

$$Z : \mathcal{D} \rightarrow TM,$$

such that $\pi \circ Z = \kappa$, where π is the canonical projection $\pi : TM \rightarrow M$.

We can then define $Z_u = \frac{DZ}{\partial u}$ the covariant derivative of Z along the u -parameter curve, given in local coordinates by

$$Z_u = \partial_u Z^\alpha \partial_{x^\alpha} + Z^\alpha \Gamma_{\alpha\delta}^\beta \partial_u \kappa^\delta \partial_{x^\beta},$$

and similarly $Z_v = \frac{DZ}{\partial v}$ the covariant derivative of Z along the v -parameter curve.

As in the case of curves, in the particular case that $Z = \kappa_u$, Z_u is the acceleration of the u -parameter curve. Moreover,

Lemma 8.1. 1. If κ is a two parameter map,

$$\kappa_{uv} = \kappa_{vu}.$$

2. If Z is a vector field on κ , then

$$Z_{uv} - Z_{vu} = R(\kappa_v, \kappa_u)Z$$

where R is the Riemann curvature tensor.

Proof. Writing κ_{uv} in coordinates, it follows that it is symmetric in uv . We leave the proof of the second point as an exercise. \square

Remark 8.1. The induced covariant derivative on a curve, the induced covariant derivative on a submanifold and the two covariant derivatives for vector fields on a two-parameter map are in fact all special examples of the induced covariant derivative for sections of the pullback bundle. You can try to write a definition for it as an exercise.

8.2 The Gauss lemma

So far we have seen that given $p \in M$, \exp_p carries rays $t \rightarrow tv$ into radial geodesics $t \rightarrow \gamma_v(t)$. Moreover, we computed the differential of the exponential map at $0 \in T_pM$. The following result, called the *Gauss lemma* is concerned with the differential map of the exponential map at some radial vector. It implies in particular that orthogonality to radial directions is preserved by the exponential map.

Lemma 8.2. Let $p \in M$ and $0 \neq x \in T_p M$. If $v_x, w_x \in T_x T_p M$ with v_x radial, then

$$g_{\exp_p(x)}(d_x \exp_p(v_x), d_x \exp_p(w_x)) = g_p(v_x, w_x).$$

Remark 8.2. Here, on the RHS, we evaluate the metric g at p at the vectors v_x and w_x . Strictly speaking, we should replace them in this expression with their images by the canonical isomorphism that allows to identify $T_w T_p M$ and $T_p M$.

Remark 8.3. Recall that given $\Phi: V \rightarrow V$, for V a vector space endowed with a scalar product g , Φ is an isometry if $g(d\phi_x(v), d\phi_x(w)) = g(v, w)$, for all v, w . Thus, the Gauss lemma can be viewed as a statement about the exponential map being a partial isometry (it preserves certain directions). Note also that in the statement, the differential of the exponential map is taken at x (hence the notation d_x).

Proof. v_x radial means that $v = \lambda x$ (up to the canonical isomorphism) for some $\lambda \in \mathbb{R}$ and by linearity, we can assume that $\lambda = 1$, i.e. $v = x$ and we replace x by v below.

Consider the two parameter map $\tilde{x}(t, s) = t(v + sw)$ in $T_p M$ and its image by the exponential map in M

$$x(t, s) = \exp_p \tilde{x}(t, s).$$

We have $\tilde{x}(1, 0) = v$, $\tilde{x}_t(1, 0) = v_v$ and $\tilde{x}_s(1, 0) = w_v$, hence

$$x_t(1, 0) = d_v \exp_p(v_v), \quad x_s(1, 0) = d_v \exp_p(w_v).$$

The statement of the lemma is then

$$g(x_t(1, 0), x_s(1, 0)) = g(v, w).$$

By definition, the curve $t \rightarrow x(t, s)$ is a geodesic with initial velocity $v + sw$. Hence $x_{tt} = 0$ and $g(x_t, x_t) = g(v + sw, v + sw)$, for all t .

By Lemma 8.1, $x_{ts} = x_{st}$. Thus,

$$\partial_t g(x_t, x_s) = g(x_t, x_{st}) = g(x_t, x_{ts}) = \frac{1}{2} \partial_s g(x_t, x_t).$$

Since $g(x_t, x_t) = g(v, v) + s^2 g(w, w) + 2s g(w, v)$, we have

$$\frac{1}{2} \partial_s g(x_t, x_t)(t, 0) = g(w, v).$$

On the other hand, $x(0, s) = \exp_p(0) = p$, for all s , so $x_s(0, 0) = 0$ and $g(x_t, x_s)(0, 0) = 0$. Thus, we have

$$g(x_t, x_s)(t, 0) = t g(v, w),$$

which gives the lemma when evaluated at $t = 1$. \square

Exercise 8.1. Assume here that $\dim M = n \geq 2$ and that (M, g) is a Riemannian manifold. In Lemma 3.13, we introduced normal coordinates, by fixing an orthonormal basis of $T_p M$ and using the exponential map. Equivalently, we introduced a special coordinate system on $T_p M$, the one given by the orthonormal basis, and then used that \exp_p is a diffeomorphism to move these coordinates to M . Instead, fix again an orthonormal basis, and consider polar coordinate on $T_p M$ adapted to this basis, $r, \omega^1, \dots, \omega^{n-1}$, where $\omega := (\omega^1, \dots, \omega^{n-1})$ are standard coordinates on \mathbb{S}^{n-1} and $r(v) = \|v\| (= g(v, v)^{1/2})$. Denote such a coordinate system by ξ . For instance, if $n = 3$, if (x, y, z) is the coordinate system associated to the orthonormal basis, for any

$v \in T_p M$, $\xi(v) = (r, \theta, \phi)$, where (θ, ϕ) denotes the usual coordinates on \mathbb{S}^2 , so that $(z = r \cos \theta, y = r \sin \theta \sin \phi, x = r \sin \theta \cos \phi)$. Again, these coordinates can then be transferred to M via the exponential map. We denote them in this exercise by the same letters.

1. Check that $r \rightarrow \exp_p(\xi^{-1}(r, \omega))$ is a radial geodesic (this follows just by definition).
2. Check that the Gauss lemma translates into

$$g(\partial_r, \partial_r) = 1, \quad g(\partial_r, \partial_{\omega^i}) = 0.$$

In other words, the metric takes the form

$$g = dr \otimes dr + g_{ij} d\omega^i \otimes d\omega^j,$$

the points being that there are no cross terms, and that $g_{rr} = 1$.

3. What fails in the Lorentzian case?

8.3 Hyperquadrics

Denote by \tilde{q} the line element at p , i.e. the map

$$\tilde{q}: v \in T_p M \rightarrow g_p(v, v).$$

Let $q = \tilde{q} \circ \exp_p^{-1}$. (In the following, we will often use a upper $\tilde{}$ to denote objects defined on $T_p M$.)

Let \mathcal{U} be a normal neighborhood at p and $\tilde{\mathcal{U}} = \exp_p^{-1}(\mathcal{U})$.

At p , we can consider the level sets of \tilde{q} . These are called *hyperquadrics*. Since we can choose coordinates such that, at p , $g = \eta$ for g Lorentzian or $g = \delta$ for g Riemannian, we can identify the level sets of \tilde{q} with the level sets of the line element of the Minkowski space or the Euclidean space. Note that since the metric is non-degenerate, $d\tilde{q}_v \neq 0$ at any $v \neq 0$, so that, for any $c \in \mathbb{R}$ such that $\tilde{q}^{-1}(c) \neq \emptyset$, the hyperquadrics $\tilde{q}^{-1}(c)$ are honest hypersurfaces of $T_p M$ if $c \neq 0$. If g is Lorentzian, then $\tilde{q}^{-1}(0) \setminus \{0\}$ is a (non-connected) hypersurface of $T_p M \setminus \{0\}$ (we exclude $v = 0$ since it belongs to the level set $\tilde{q} = 0$ but $d\tilde{q}_0 = 0$), called the null cone at p .

Similarly, we can consider, for $c \neq 0$,

$$q^{-1}(c) = \{r \in \mathcal{U} : q(r) = c\},$$

called *local hyperquadrics*. Now, by definition,

$$q(r) = \tilde{q} \circ \exp_p^{-1}(r),$$

for all $r \in \mathcal{U}$, so if $v \in \tilde{q}^{-1}(c) \cap \tilde{\mathcal{U}}$, $q(\exp_p(v)) = c$ and $q^{-1}(c)$ is the image by the exponential map of $\tilde{q}^{-1}(c) \cap \tilde{\mathcal{U}}$. In the Lorentzian case, the *local null cone* at p , denoted $\Lambda(p)$, is by definition²⁷ $\Lambda(p) = q^{-1}(0) \setminus \{p\}$.

Let $c \neq 0$ such that $\tilde{q}^{-1}(c) \neq \emptyset$. Since $\exp_p: \tilde{q}^{-1}(c) \cap \tilde{\mathcal{U}} \rightarrow q^{-1}(c) \cap \mathcal{U}$ by construction, for any v in $\tilde{q}^{-1}(c) \cap \tilde{\mathcal{U}}$, we have

$$d_v \exp_p: T_v(\tilde{q}^{-1}(c) \cap \tilde{\mathcal{U}}) \rightarrow T_{\exp_p(v)}(q^{-1}(c) \cap \mathcal{U}).$$

²⁷The point p has to be removed so that these form regular submanifolds of M .

Excluding $0 \in T_p M$ and $p \in \mathcal{U}$, this is also true for $c = 0$, i.e. for the null cone and the local null cone.

Recall the definition of the position (also called scaling) vector field $\widetilde{\mathcal{P}}$. It is a vector field of $T_p M$, i.e. $\widetilde{\mathcal{P}} \in \Gamma(T_p M)$ and it is given in coordinates by

$$\widetilde{\mathcal{P}}(v) = v^\alpha \partial_{v^\alpha},$$

for $v \in T_p M$.

On the other hand, since \tilde{q} is a function on $T_p M$, its differential $d\tilde{q} \in \Lambda(T_p M)$ is a one-form of $T_p M$ and its gradient a vector field on $T_p M$.

For $v \in T_p M$, $\tilde{q}(v, v) = g_{\alpha\beta} v^\alpha v^\beta$ so that

$$d\tilde{q}_v = 2g_{\alpha\beta} v^\alpha dv^\beta,$$

and multiplying by g^{-1} , we obtain the following lemma.

Lemma 8.3. *We have*

$$\text{grad } \tilde{q} = 2\tilde{P}.$$

As a consequence, the position vector field is orthogonal to every hyperquadric and, in the Lorentzian case, is both tangent and orthogonal to the null cone $\tilde{q}^{-1}(0) \setminus \{0\}$. Finally, by construction, the position vector field is radial (trivial).

Using the exponential map, we can then define the *local position vector field* at p as follows. Let $q \in \mathcal{U}$ and $v = \exp_p^{-1}(q)$. Then define \mathcal{P} as

$$\mathcal{P}(q) = d_v \exp_p(\widetilde{\mathcal{P}}).$$

From the Gauss lemma, it follows that

Corollary 8.1. *The local position vector field P at p is orthogonal to every local hyperquadric of M at p . Furthermore, in the Lorentzian case, P is both orthogonal and tangent to the nullcone $q^{-1}(0) \setminus \{p\}$. Finally, P is radial, that is tangent to every radial geodesics emanating from p .*

Proof. Let $c \neq 0$ and consider a local hyperquadric $q^{-1}(c) \neq \emptyset$. Let X be tangent to $q^{-1}(c)$ at some point r . Let $v \in T_p M$ be such that $\exp_p(v) = r$. Now, since we have an isomorphism

$$d_v \exp_p : T_v \left(\tilde{q}^{-1}(c) \cap \widetilde{\mathcal{U}} \right) \rightarrow T_{\exp_p(v)} q^{-1}(c),$$

we can take $x \in T_v \left(\tilde{q}^{-1}(c) \cap \widetilde{\mathcal{U}} \right)$ be such that $d_v \exp_p(x) = X$. Then,

$$g(P, X)(r) = g(d_v \exp_p(\widetilde{\mathcal{P}}), d_v \exp_p(x)) = g(\mathcal{P}, x),$$

by the Gauss Lemma. Since x is tangent to $\tilde{q}^{-1}(c)$ and $\widetilde{\mathcal{P}}$ is orthogonal to $\tilde{q}^{-1}(c)$ by the previous lemma, we have $g(P, X)(r) = 0$. We leave the rest as an exercise. \square

Corollary 8.2. *Let Dq denotes the gradient vector field of q . Then, $Dq = 2P$.*

Proof. First, we have $D\tilde{q} = 2\widetilde{\mathcal{P}}$ from Lemma 8.3. Let $r \in \mathcal{U}$ and $v \in T_r \mathcal{U}$. Let w such that $\exp_p(w) = r$ and \tilde{v} be such that $d_w \exp_p \tilde{v} = v$. Then,

$$\begin{aligned} g_r(Dq, v) &= v(q) = \left(d_w \exp_p \tilde{v} \right) (q) = \tilde{v} \left(q \circ \exp_p \right) \\ &= \tilde{v}(\tilde{q}) = g(D\tilde{q}, \tilde{v}) = 2g(\widetilde{\mathcal{P}}, \tilde{v}) = 2g(P, v). \end{aligned}$$

where in the last step we used the Gauss lemma. \square

8.4 Convex neighborhoods

Recall the concept of convexity in Euclidean geometry: a set \mathcal{S} in an Euclidean space is convex if and only for all $0 \leq s \leq 1$ and all points $p, q \in \mathcal{S}$, $p(1-s) + sq \in \mathcal{S}$. Let us try to translate this in an arbitrary semi-Riemannian manifold. Given a set \mathcal{S} and $p, q \in \mathcal{S}$, the replacement for the curve $s \rightarrow ps + (1-s)q$ should clearly be a geodesic $s \rightarrow \gamma(s)$ from p to q . This motivates the following definition.

Definition 8.2. *An open set \mathcal{C} is convex provided it is a normal neighborhood of each of its points.*

In particular, for any two points p, q of \mathcal{C} , there is a unique geodesic segment $\sigma_{pq} : [0, 1] \rightarrow M$ from p to q that lies entirely in \mathcal{C} (but contrary to the Euclidean or Minkowskian case, there could well be others that leave \mathcal{C}).

The aim of this section will be to prove

Proposition 8.1. *Let M be a semi-Riemannian manifold. Then for any $p \in T_p M$, there exists a convex set \mathcal{C} containing p .*

Before we prove this proposition, let us introduce some extra material concerning the exponential map, viewed this time as a mapping defined on TM instead of the individual tangent plane. More precisely, let $\pi : TM \rightarrow M$ be the canonical projection. Let \mathcal{D} be the set of all vectors $v \in TM$, such that the curve γ_v defined as before as the unique maximal geodesic starting at $\pi(v)$ with initial tangent vector v is defined on $[0, 1]$ (i.e. its maximal time of existence is larger than 1). Define as well $\mathcal{D}_p = \mathcal{D} \cap T_p M$. (Here we identify $T_p M$ with a subset of TM , which if TM is defined as the disjoint union of the $T_p M$, is the subset $\{p\} \times T_p M$).

Lemma 8.4. *\mathcal{D} is open in TM and \mathcal{D}_p is an open set of $T_p M$ that is starshaped around 0.*

Proof. That \mathcal{D} is open follows from standard ode theory (in particular continuous dependance with respect to the initial data). That \mathcal{D}_p is starshaped around 0 is a consequence of the fact $\gamma_v(st) = \gamma_{sv}(t)$, for $0 \leq s, t, \leq 1$. \square

Let us now define E as

$$\begin{aligned} E : \mathcal{D} \subset TM &\rightarrow M \times M \\ v = (p, v_p) &\rightarrow (p(= \pi(v)), \exp_p(v_p)) \end{aligned}$$

Recall that a smooth map between two manifolds is called non-singular at some point if its differential at that point is injective²⁸ (one-to-one). We have

Lemma 8.5. *Let $p \in M$ and $w \in \mathcal{D}_p$. If $\exp_p : \mathcal{D}_p \rightarrow M$ is non-singular at w , then $E : \mathcal{D} \rightarrow M \times M$ is non-singular at w .*

Proof. Suppose $d_w E(v) = 0$ for $v \in T_w(TM)$. We must show that $v = 0$. Let $\pi : TM \rightarrow M$ denote the canonical projection and let π_1 be the projection of $M \times M$ on its first factor. Then $\pi_1 \circ E = \pi$, so that $d\pi_w(v) = d\pi_1(d_w E(v)) = 0$. It follows that v is vertical, that is tangent to $T_p M$. Since \mathcal{D}_p is open, we can find a curve in $\sigma : I \rightarrow \mathcal{D}_p$, such that $\sigma'(0) = v$, $\sigma(0) = w$. The result then follows since, for all $t \in I$, $E(\sigma(t)) = (p, \exp_p(\sigma(t)))$ and we have assumed that \exp_p is non-singular at w . \square

²⁸A regular point on the other hand is one where the differential is onto.

For any $p \in M$, it then follows from the inverse function theorem and Proposition 3.8 that E maps some open neighborhood of $(p, 0_p)$ in TM diffeomorphically onto an open neighborhood of $(p, p) \in M \times M$. We now proceed to the proof of Proposition 8.1.

Proof. Let $\xi = (x^\alpha)$ be a normal coordinate system on a neighborhood of $p \in M$. Let $N(\delta)$ be the set of points q such that

$$\sum_{\alpha} |x^\alpha(q)|^2 < \delta.$$

For δ small enough, this set defines a normal neighborhood of p which is diffeomorphic to an open ball. (Note that N_δ is clearly open and contains p . Thus, it is a neighborhood of p and clearly, $\exp_p^{-1}(N_\delta)$ is a neighborhood of $0 \in T_pM$. Using the definition of normal coordinates, it is not hard to check that this neighborhood is starshaped around 0.)

By the previous lemma, if δ is small enough, then E is a diffeomorphism from a neighborhood \mathcal{N} containing $(p, 0) \in T_pM$ onto $N(\delta) \times N(\delta)$. Moreover, for δ small enough, the tensor field whose components are given by $\delta_{\alpha\beta} - \Gamma_{\alpha\beta}^\gamma \delta_{\gamma\rho} x^\rho$ can be assumed to be positive definite. We claim that it then follows that $N(\delta)$ is a normal neighborhood of each of its points $q \in N(\delta)$.

Let $\mathcal{N}_q = \mathcal{N} \cap T_qM$. By construction, $E|_{\mathcal{N}_q}$ is a diffeomorphism onto $\{q\} \times N(\delta)$, hence $[\exp_q]_{|\mathcal{N}_q}$ is a diffeomorphism onto $N(\delta)$. It remains to show that \mathcal{N}_q is star-shaped about 0.

Let $r \in N(\delta)$, $r \neq q$ and let $v = E^{-1}(q, r)$. Then, $v \in \mathcal{N}_q$ and $\sigma = (\gamma_\nu)|_{[0,1]}$ is a geodesic from q to r . If σ lies in $N(\delta)$ then $t v \in \mathcal{N}_q$ for all $0 \leq t \leq 1$ and hence \mathcal{N}_q is starshaped around 0.

For any $k \in N(\delta)$, let $n(k) = \sum_{\alpha} |x^\alpha(k)|^2$. We consider the function $n \circ \sigma$ and differentiate twice, writing x^α for $x^\alpha(\sigma)$ and using the geodesic equation:

$$\begin{aligned} \frac{d^2 n \circ \sigma}{dt^2} &= 2 \left(\delta_{\beta\alpha} x^\beta \frac{d^2 x^\alpha}{dt^2} + \left[\frac{dx^\alpha}{dt} \right]^2 \right) \\ &= 2 \left(\delta_{\alpha\beta} - \delta_{\gamma\rho} x^\rho \Gamma_{\beta\alpha}^\gamma \right) \frac{dx^\alpha}{dt} \frac{dx^\beta}{dt} > 0, \end{aligned}$$

where we have used the geodesic equations to reach the last line. Thus, the function $n \circ \sigma$ is (strictly) convex and in particular, for any $t \in [0, 1]$, $n \circ \sigma(t) \leq (1-t)n \circ \sigma(0) + t.n \circ \sigma(1) < \delta$. \square

Remark 8.4. *In the above proof, we introduce a Riemannian metric (the Euclidean metric associated to the normal coordinates) on our neighborhood. This is a rare instance where one can draw conclusions in arbitrary semi-Riemannian manifolds by forcing a Riemannian structure.*

If p and q are points of a convex set \mathcal{C} and σ_{pq} is the geodesic in \mathcal{C} from p to q (unique up to parametrization), the *displacement vector* \vec{pq} is by definition

$$\vec{pq} := \sigma'_{pq}(0) \in T_pM.$$

Lemma 8.6. *Let \mathcal{C} be a convex (open) set. Then, the map $\Delta : \mathcal{C} \times \mathcal{C} \rightarrow TM$ sending (p, q) to \vec{pq} is smooth.*

Proof. Recall the map $E : \mathcal{D} \rightarrow M \times M$. We have that \mathcal{D} is open and that E is smooth. In particular, since $\mathcal{C} \times \mathcal{C}$ is open in $M \times M$, it follows that $E^{-1}(\mathcal{C} \times \mathcal{C})$ is open. One easily check that $E^{-1}(\mathcal{C} \times \mathcal{C}) = \Delta(\mathcal{C} \times \mathcal{C})$ and then that Δ is the inverse map of

$$E : \Delta(\mathcal{C} \times \mathcal{C}) \rightarrow \mathcal{C} \times \mathcal{C}.$$

□

Exercise 8.2. 1. Consider $\mathbb{S}^1 \subset \mathbb{R}^2$. Find two convex sets whose intersection is non connected, hence in particular non convex.

2. Prove that if U, V are two convex sets included into a convex set W , then $U \cap V$ is a convex set if non-empty.

Definition 8.3. A convex covering is a covering of M by convex open sets such that is U and V are elements of the covering, then their intersection $U \cap V$ is convex if not empty.

Let us end this section with the following lemma, whose proof is left as an exercise.

Lemma 8.7. Given any open covering C of M , there exists a convex covering R such that each element of R is contained in some element of C . In other words, the convex covering is a refinement of the original cover.

Proof. (sketch) Let C be an open covering of M . Recall that M is metrizable, so let d be a metric on M which is compatible with the topology of M . For any $x \in M$, let N_x be a convex neighborhood containing x . Shrinking N_x is necessary, we may assume that N_x is bounded (for d), has compact closure and is contained in some element of C . Then, let V_x be a convex neighborhood of x , such that $\text{diam}(V_x) < 1/3d(x, \partial N_x)$. The set of all such $V_x, x \in M$, is then a covering of M , such that each V_x is contained in some element of C and is convex. Moreover, if $V_x \cap V_y \neq \emptyset$, assume wlog that $\text{diam}(V_x) \geq \text{diam}(V_y)$ then by construction,

$$d(x, y) \leq 2\text{diam}(V_x) \leq 2/3d(x, \partial N_x).$$

Thus, $y \in N_x$ and then $V_y \in N_x$. Thus, $V_x \cap V_y \in N_x$ so that $V_x \cap V_y$ is convex from Exercise 8.2.2.

□

8.5 Arc lengths

Definition 8.4. Let α be a piecewise smooth curve defined on a closed interval $[a, b]$. We define its arc length as

$$L(\alpha) = \int_a^b |g(\alpha', \alpha')|^{1/2} ds.$$

A reparametrization function $h : [c, d] \rightarrow [a, b]$ is a piecewise smooth function such that either $h(c) = a, h(d) = b$ (h is orientation preserving), or $h(c) = b, h(d) = a$ (h is orientation-reversing). If its derivative does not change sign, h is monotone.

Lemma 8.8. The length of a piecewise smooth curve segment is invariant under a change of monotone reparametrization. Moreover, if $|g(\alpha', \alpha')| > 0$, there is a strictly increasing reparametrization function h such that $\beta = \alpha \circ h$ has $|g(\beta', \beta')| = 1$.

A curve such that $|g(\alpha', \alpha')| = 1$ is said to have *unit speed* or *arc length parametrization*.

Let \mathcal{U} be a normal neighborhood of $p \in M$. The function

$$r := k \in \mathcal{U} \rightarrow r(k) = \left| g \left(\exp_p^{-1}(k), \exp_p^{-1}(k) \right) \right|^{1/2}$$

is the *radius function* at p .

Exercise 8.3. Let (x^α) be normal coordinates at p . Give an expression for $r(q)$ in terms of the coordinates of q .

Lemma 8.9. Let r be the radius function on a normal neighborhood \mathcal{U} of $p \in M$. If σ is the radial geodesic from p to $q \in \mathcal{U}$, then $L(\sigma) = r(q)$.

Proof. If $v = \sigma'(0)$, then $v = \exp_p^{-1}(q)$. Since $g(\sigma', \sigma')$ is constant along σ ,

$$L(\sigma) = \int_0^1 |g(\sigma', \sigma')|^{1/2} = |g(v, v)|^{1/2} = r(q).$$

□

8.6 The Riemannian case

We assume here that g is a Riemannian metric. Then, $v \rightarrow \tilde{r}(v) = g(v, v)^{1/2}$ is a norm on each tangent plane and is smooth away from $v = 0$. so that, given a normal neighborhood \mathcal{U} of p , the radius function at p ,

$$r = \tilde{r} \circ \exp_p^{-1}$$

is smooth exact at p . Let $\tilde{q}(v) = g(v, v)$ and $q = \tilde{q} \circ \exp_p^{-1}$ and denote by P the local position vector field, i.e. $P = d \exp_p \tilde{P}$, with \tilde{P} given in local coordinates by

$$\tilde{P} = v^\alpha \partial_{v^\alpha}.$$

Then, for all $q = \exp_p(v)$,

$$|g(P, P)|^{1/2}(q) = g(d \exp_p \tilde{P}, d \exp_p \tilde{P})^{1/2}(q) = g(\tilde{P}, \tilde{P})^{1/2} = g(v, v)^{1/2} = r(q).$$

Hence, $U = P/r$ is the *outward unit radial vector field* (outward because r increase along the integral curves of U , radial because P is radial) and U is normal to all the hyperspheres at p , i.e. the hypersurfaces of constant r . Since $r = q^{1/2}$, it follows from Corollary 8.2 that

$$\text{grad } r = \frac{P}{r} = U$$

on $\mathcal{U} \setminus \{p\}$.

Proposition 8.2. Let (M, g) be a Riemannian manifold and \mathcal{U} be a normal neighborhood of p . If $q \in \mathcal{U}$, then the geodesic $\sigma : [0, 1] \rightarrow \mathcal{U}$ with $\sigma(0) = p$, $\sigma(1) = q$ is the unique (up to monotone reparametrization) shortest (piecewise smooth) curve in \mathcal{U} from p to q .

Remark 8.5. Note that here everything is constrained to stay in the neighborhood \mathcal{U} .

Proof. We must prove that if $\alpha : [0, b] \rightarrow \mathcal{U}$ is another curve joining p to q , then $L(\alpha) \geq L(\sigma)$ with equality if and only if $\alpha = \sigma \circ h$ for some monotone h . We will prove it in the case of smooth curves for simplicity, but modulo easy modifications the proof works for piecewise smooth curves.

First, restricting the vector field $U = P/r$ to α , we have

$$\alpha' = g(\alpha', U)U + N,$$

where N is a vector field on α orthogonal to U for all $t > 0$. At $t = 0$, the vector field U is a priori not well defined, but since $g(U, U) = 1$, its components are uniformly bounded in any local system of coordinates. (In fact, U can be extended continuously to $\frac{\alpha'(0)}{r(\alpha'(0))}$.)

Then,

$$g(\alpha', \alpha')^{1/2} = [g(\alpha', U)^2 + g(N, N)]^{1/2} \geq |g(\alpha', U)| \geq g(\alpha', U).$$

On the other hand, from $\text{grad } r = U$, we have $g(\alpha', U) = \frac{dr \circ \alpha}{dt}$. Hence,

$$L(\alpha) = \int_0^b g(\alpha', \alpha')^{1/2} ds \geq \int_0^b g(\alpha', U) ds = r(q) = L(\sigma).$$

Moreover, in the equality case, then we have $N = 0$, for all $t > 0$ as well as $|g(\alpha', U)| = g(\alpha', U)$, which implies that

$$\frac{dr \circ \alpha}{dt} \geq 0.$$

Thus, $\alpha' = \frac{dr \circ \alpha}{dt} U$, which implies that α is a monotone reparametrization of an integral curve of U . Since U is radial, the integral curves of U are geodesics passing through p , so it follows that α is a monotone reparametrization of a geodesic from p to q , which by uniqueness has to be σ . In fact, $\alpha(t) = \sigma(r \circ \alpha(t)/r(q))$. \square

8.6.1 The Riemannian distance

Definition 8.5. For any points p and q of a connected Riemannian manifold M , the Riemannian distance from p to q is the infimum of all the lengths of all piecewise smooth curves joining p to q .

Remark 8.6. Recall that any connected manifold is path connected, so that the above definition makes sense.

If $p \in M$ and $\epsilon > 0$, we define $B_\epsilon(p)$ to be the open ball of size ϵ for the Riemannian distance d , i.e.

$$B_\epsilon(p) = \{q \in M : d(p, q) < \epsilon\}.$$

Note that for the moment, we do not know that this is actually open for the topology of M .

Proposition 8.3. Let $p \in M$.

1. For any $\epsilon > 0$ sufficiently small, $B_\epsilon(p)$ is normal (in particular open).
2. For q in a normal neighborhood of the form $B_\epsilon(p)$, the radial geodesic σ from p to q is the unique shortest curve in M from p to q . In particular,

$$L(\sigma) = r(q) = d(p, q).$$

Remark 8.7. *The point of the second item above is that we are considering all possible geodesics, not only those constrained in the normal neighborhood.*

Proof. Let \mathcal{V} be a normal neighborhood of $p \in M$ and $\mathcal{U} = \exp_p^{-1}(\mathcal{V})$. For $\epsilon > 0$ sufficiently small, \mathcal{U} contains the starshaped open set

$$\tilde{B}_\epsilon(0) = \{v \in T_p M : g(v, v)^{1/2} < \epsilon\}.$$

Thus, $N := \exp_p(\tilde{B}_\epsilon(0))$ is also a normal neighborhood of p and by construction $N \subset B_\epsilon(p)$ (cf Lemma 8.9). Moreover, if $q \in N$, then by the previous lemma, the radial geodesic σ from p to q is the unique shortest curve in N from p to q and $L(\sigma) = r(q)$. Since $L(\sigma) < \epsilon$, it suffices to prove that if α is a curve in M starting at p and leaving N , then $L(\alpha) \geq \epsilon$. In particular, this shows that if $q \notin N$, then $d(p, q) \geq \epsilon$ and hence that $N = B_\epsilon(p)$. To prove the claim, note that since α leaves N , it meets every sphere $S(a)$ of constant r equal to a , for all $a < \epsilon$. If α_a is the shortest initial segment of α from p to $S(a)$, then α_a lies in N . Hence $L(\alpha_a) \geq a$ by the previous lemma. Since this holds for all $a < \epsilon$, the claim is proved. \square

Proposition 8.4. *For a connected Riemannian manifold M , the Riemannian distance function $d : M \times M \rightarrow \mathbb{R}$ defines a metric on M , i.e. it satisfies*

1. *for all $p, q \in M$, $d(p, q) = 0$ if and only if $p = q$.*
2. *for all $p, q \in M$, $d(p, q) = d(q, p)$,*
3. *for all $p, q, r \in M$, $d(p, q) \leq d(p, r) + d(r, q)$.*

Moreover, d is compatible with the topology of M .

Proof. The symmetry property is trivial. Assume that $d(p, q) = 0$, since M is Hausdorff, there is a normal neighborhood of p that does not contain q and from the (proof of the) previous proposition, there exists an open ball around p of radius $\epsilon > 0$ that does not contain q . Thus, $d(p, q) > \epsilon > 0$.

For the triangle inequality, let $\epsilon > 0$ and choose α a curve from p to q and β a curve from q to r , such that $L(\alpha) < d(p, q) + \epsilon$ and $L(\beta) < d(q, r) + \epsilon$. Joining α and β gives a curve of length at most $L(\alpha) + L(\beta) + 2\epsilon$. In other words, the triangle inequality follows from the definition of infimum of a set and the fact that given curves from p to q and q to r , we can get a curve from p to r whose length is the sum of the previous length.

Finally, let $p \in M$ and consider $q \in B_r(p)$, $r > 0$. We have $d(p, q) < r$ and so, by the triangle inequality, for any r' such that $0 < r' < r - d(p, q)$, the ball $B_{r'}(q)$ is contained in $B_r(p)$. Since for r' small enough, we know from the previous proposition that $B_{r'}(q)$ is open, it follows that $B_r(p)$ is open for the topology defined by M . Conversely, we have already proven that any open set for the topology of M contains a small enough ball around any point of that set. Thus, the two topologies are compatible. \square

A *minimizing segment* from p to q is by definition a curve from p to q that realizes the Riemannian distance function. Note that there could be several or none such segments. One has

Corollary 8.3. *A minimizing segment σ from p to q is a monotone reparametrization of a smooth geodesic segment from p to q .*

Proof. The domain I of α can be decomposed into subintervals I_i such that each $\alpha_i = \alpha|_{I_i}$ lies a convex open set. Since α is minimizing, each α_i is also minimizing (the restriction of a minimizing segment is a minimizing segment). Since each α_i lies in a convex set, it follows that it is a monotone reparametrization of a geodesic which we can assume to be of unit speed σ_i . Joining this gives a possibly broken geodesic σ from p to q and joining the reparametrization gives that α is a monotone reparametrization of σ . Moreover, we claim that in fact σ is unbroken, i.e. it is a smooth geodesic. For, consider two geodesic segments σ_i ending at r and a geodesic segment σ_{i+1} starting at r . Since α is minimizing, then σ is minimizing. Thus, the restriction of σ to any sufficiently small neighborhood of r is minimizing. But this implies that this restriction is a monotone reparametrization of a smooth geodesic. Since we had assume σ to be parametrized by arc length, it follows that in fact σ is a smooth geodesic, which concludes the proof. \square

8.6.2 The Hopf-Rinow Theorem

The aim of this section is to prove the following theorem.

Theorem 8.1 (Hopf-Rinow (1931)). *For a connected Riemannian manifold M the following are equivalent*

- (MC) *Metric completeness: (M, d) is a complete metric space, i.e. every Cauchy sequence is convergent.*
- (GC1) *Geodesic completeness from one point: There exists a point $p \in M$ such that every geodesic starting at p is defined on \mathbb{R} .*
- (GC) *Geodesic completeness: M is geodesically complete.*
- (HB) *Heine-Borel property: Every closed bounded subset of M is compact.*

With the notations of the theorem, we start by proving

Lemma 8.10. *If (GC₁) holds, then for any $q \in M$, there exists a minimizing geodesic segment from p to q .*

Proof. Consider a normal ball $B_\epsilon(p)$, where $\epsilon > 0$ small enough. If $q \in B_\epsilon(p)$, then the claim follows from Proposition 8.3, thus we assume that $q \notin B_\epsilon(p)$. Let r be the radius function at p . For $0 < \delta < \epsilon$, the level set $r = \delta$ is an $(n - 1)$ dimensional sphere in $B_\epsilon(p)$ denoted S_δ . In particular, S_δ is compact. The function $s \rightarrow d(s, q)$ is continuous on S_δ , hence reach its minimum at some point $m \in S_\delta$. We claim that

$$d(p, m) + d(m, q) = d(p, q).$$

To prove this, let $\alpha : [0, b] \rightarrow M$ be any curve from p to q . The function $\tau(s) : s \rightarrow d(\alpha(s), p)$ is continuous and satisfies $\tau(0) = 0$, $\tau(b) = d(p, q) > \epsilon$. By the intermediate value theorem, there exists an a , such that $d(\alpha(a), p) = \delta$. By definition, $\alpha(a) \in B_\epsilon(p)$ and in fact $\alpha(a) \in S_\delta$, cf Proposition 8.3. Moreover, $0 < a < b$. Let α_1 and α_2 be the restriction of α to $[0, a]$ and $[a, b]$ respectively. Then,

$$L(\alpha_1) \geq \delta = d(p, m)$$

by the minimizing property of radial geodesics and

$$L(\alpha_2) \geq d(m, q)$$

by the definition of m .

Thus,

$$L(\alpha) = L(\alpha_1) + L(\alpha_2) \geq d(p, m) + d(m, q).$$

Hence,

$$d(p, q) \geq d(p, m) + d(m, q).$$

The reverse inequality is just the triangle inequality, so the claim is proven.

Now, let $\gamma : [0, +\infty) \rightarrow M$ be the unit speed geodesic whose initial segment runs radially from p to m . (Note that γ is defined on $[0, +\infty)$ since we assumed (GC_1) .)

Let $d = d(p, q)$ and let T be the set of $t \in [0, d]$ such that

$$t + d(\gamma(t), q) = d.$$

It suffices to show that $d \in T$. For, it then follows that $d(\gamma(d), q) = 0$, which implies that $\gamma(d) = q$. Since γ has unit speed, $L(\gamma|_{[0, d]}) = d$, so that $\gamma|_{[0, d]}$ is indeed a minimizing geodesic from p to q .

Note first that $\gamma|_{[0, t]}$ is minimizing for any $t \in T$, since its length is t and

$$d \leq d(p, \gamma(t)) + d(\gamma(t), q) = d(p, \gamma(t)) + d - t,$$

which implies that $d(p, \gamma(t)) \leq t$ and thus that $d(p, \gamma(t)) = t$.

By continuity, T is closed. Moreover T is non-empty by the first part of the proof. Thus, it contains a largest number $t_0 \leq d$. Assuming $t_0 < d$, we deduce a contraction as follows. In a normal neighborhood \mathcal{W}' of $\gamma(t_0)$, repeating the first part of the proof produces a unit speed radial geodesic $\sigma : [0, \delta'] \rightarrow \mathcal{W}'$ from $\gamma(t_0)$ to $m' \in \mathcal{W}'$ such that

$$\delta' + d(m', q) = d(\gamma(t_0), q). \quad (26)$$

Moreover, because $t_0 \in T$, we have

$$t_0 + d(\gamma(t_0), q) = d,$$

so that from (26),

$$t_0 + \delta' + d(m', q) = d.$$

Since $d \leq d(p, m') + d(m', q)$ and since $t_0 + \delta'$ is the length of a curve from p to m' , we must have $t_0 + \delta' = d(p, m')$. Thus, the sum of $\gamma|_{[0, t_0]}$ and σ is minimizing, so from Corollary 8.3 and the fact that they have unit speed (so the parametrization is fixed), the sum is a smooth minimizing geodesic. This means that $m' = \sigma(\delta') = \gamma(t_0 + \delta')$, so that

$$t_0 + \delta' + d(\gamma(t_0 + \delta'), q) = d(p, m') + d(m', q) = d,$$

hence $t_0 + \delta' \in T$ a contradiction. \square

For the proof of the Hopf-Rinow theorem, we need the notion of an *extendible* curve.

Definition 8.6. A piecewise smooth curve $\alpha : [0, B) \rightarrow M$ is extendible provided

$$\lim_{s \rightarrow B, s < B} \alpha(s)$$

exists. We then say that α has a continuous extension $\tilde{\alpha} : [0, B] \rightarrow M$ and $q = \tilde{\alpha}(B)$ is called an endpoint of α .

Equivalently, there exists a $q \in M$ such that for every sequence $s_i \rightarrow B$, $\alpha(s_i) \rightarrow q$. An extendible curve need not have a piecewise smooth extension however, it does in the special case of integral curves, as a consequence of the Cauchy-Lipshitz theorem and the standard continuation criterion for odes. Note that this above definition makes sense for manifolds in general (it is not restricted to Riemannian manifolds).

- Exercise 8.4.** 1. Let (M, g) be semi-Riemannian manifold. With the above notation, prove that if α is a geodesic, then it is extendible if and only if it extendible as a smooth geodesic. (In the Riemannian case, this follows from Cauchy-Lipschitz and the standard continuation criterion for odes, for a solution in the general case, see for instance Lemma 8, p.130 in [O'N83].)
2. Let X be vector field on M that never vanishes (equivalently its flow has no fixed point). Prove that the maximal integral curves of M are inextendible. In particular, the maximal integral curves of a globally causal vector field are all inextendible.

Proof of the Hopf-Rinow Theorem 8.1.

(MC) \implies (GC). Let $\gamma : [0, b)$ be a unit speed geodesic. Assume that $b < +\infty$. If (t_i) is a sequence in $[0, b)$ converging to b then $(\gamma(t_i))$ is Cauchy, since

$$d(\gamma(t_i), \gamma(t_j)) \leq |t_i - t_j|.$$

Hence, $(\gamma(t_i))$ converges to some $q \in M$. Moreover, if (s_j) is another such sequence, then we still have $d(\gamma(t_i), \gamma(s_j)) \leq |t_i - s_j|$, so that the limit points is unique. It follows that γ is continuously extendible and it follows from the above exercise that in fact it is extendible as a smooth geodesic. Hence, every maximal geodesic is defined on $(-\infty, +\infty)$ i.e. M is geodesically complete.

(GC) \implies (GC1) Trivial.

(GC1) \implies (HB) Let p be as in (GC1). Let A be closed bounded subset of M . Note that by definition, A bounded means that d is bounded on $A \times A$ and by the triangle inequality, it follows that $q \in A \rightarrow d(q, p)$ is bounded on A .

Let $q \in A$. By Lemma 8.10, there exists a minimizing geodesic σ_q defined on $[0, 1]$ from p to q . Then $g(\sigma_q'(0), \sigma_q'(0))^{1/2} = L(\sigma_q) = d(p, q) < r$, for some $r > 0$ independent of q . In other words, $\sigma_q'(0)$ lies in the compact ball B_r of $T_p M$. Since $\exp_p(B_r)$ is compact and contains the closed set A , A is compact.

(HB) \implies (MC) Always true in metric spaces, so left as an exercise.

As a consequence of the Hopf-Rinow theorem (and the lemma), we have

Corollary 8.4. *If a connected Riemannian manifold is complete then any two of its points can be joined by a minimizing geodesic segment.*

Corollary 8.5. *Any compact Riemannian manifold is complete.*

Proof. Property (HB) holds trivially in this case. □

Finally, let us recall Whitney's embedding theorem²⁹ and obtain yet another corollary of the Hopf-Rinow theorem (combined with Whitney's embedding theorem).

²⁹Actually, Whitney proved that M can be embedded into \mathbb{R}^{2n} , but the $2n+1$ version is easier to prove and we do not care here about the dimension of the target space.

Theorem 8.2. *Let M be a smooth connected manifold of dimension n . Then, there exists an embedding of M into \mathbb{R}^{2n+1} such that the image under the embedding is a closed subset.*

For a proof, see for instance Theorem 10.8, p92 in [Bre97].

Corollary 8.6. *Let M be a smooth connected manifold. Then, there exists a complete Riemannian metric on M .*

Proof. Let $\phi : M \rightarrow \mathbb{R}^{2n+1}$ denotes an embedding as in Whitney's theorem and give M the pullback metric by the embedding. Let d_M be the corresponding Riemannian distance. Then, closed and bounded set of M with respect to d are easily seen to be compact (use for instance that $d_{\mathbb{R}^3}(\phi(p), \phi(q)) \leq d_M(p, q)$, $p, q \in M$), so that property (HB) is satisfied. \square

For an alternative, more constructive, proof, see Lemma 11.1, p111 in [Rin09].

8.7 The Lorentzian case

Many (most) of the previous properties proved in the case of Riemannian manifolds *do not* hold for Lorentzian manifolds. As a start, g does not defines a norm on each tangent space, so in particular, for any $\epsilon > 0$, the set of all v , such that $|g(v, v)| \leq \epsilon$ is non compact. The bottom line is that, yes there are analogies between the Riemannian and Lorentzian cases (for instance, we will soon see a local *maximizing* property of timelike geodesics analogous, including the proof, to the *minimizing* property of geodesics in Riemannian geometry), but overall Lorentzian manifolds have their own, more complicated, geometry.

We now assume for the rest of this section that (M, g) is a smooth Lorentzian manifold.

8.7.1 Timecone and time-orientability

Recall that $v \in T_p M \neq 0$ is called spacelike, null or timelike depending on whether $g_p(v, v) > 0, = 0$ or < 0 and that the set causal vectors is the union of the null and timelike vectors. Similarly, a submanifold is also called spacelike, null or timelike depending on whether corresponding the induced metric is Riemannian, degenerate or Lorentzian. In particular, this applies to smooth curves. Moreover, if $z \in T_p M$ is timelike, then z^\perp is spacelike (in particular, by convention, $0 \in T_p M$ is spacelike).

The proof of the following lemma is left as an exercise.

Lemma 8.11. *[Reverse Cauchy-Schwarz inequality] Let v, w be timelike vectors. Then,*

$$(g(v, w))^2 \geq |g(v, v)| \cdot |g(w, w)|.$$

Let \mathcal{T}_p denotes the set of all timelike vectors at some $p \in M$. On \mathcal{T}_p , we consider a binary relation defined as follows. We say for $v, w \in \mathcal{T}_p$ that $v C w$ if and only if $g(v, w) < 0$.

Lemma 8.12. *1. The relation C is an equivalence relation and there are exactly two equivalent classes, referred to as timecones.*

2. The two timecones are open and connected, so that \mathcal{T}_p has two connected components.

3. The two timecones are convex sets.

4. The map $v \in \mathcal{T}_p \rightarrow -v$ maps one timecone into the other.

Proof. Clearly, C is a symmetric binary relation since g is symmetric. That C is reflexive follows from the definition of a timelike vector. Moreover, if $g(v, w) < 0$ and $g(w, x) < 0$, then since w is timelike, it follows that $w \neq 0$, and we can find an orthonormal basis (e_i) such that $w = |g(w, w)|^{1/2} e_0$ where $g(e_0, e_0) = -1$. Then, $g(v, w) < 0$ implies that $v = v^0 e_0 + v^i e_i$, with $v^0 > 0$ and $v^0 = |v^0| > (v^i v^j \delta_{ij})^{1/2}$ since v is timelike and similarly for x . Then,

$$g(x, v) = -v^0 x^0 + v^i x^j \delta_{ij}$$

so that the result follows by the Cauchy-Schwarz inequality. We leave the rest as an exercise. \square

With the notion of timecones, we can extend the notation of timelike, null or causal to curves which are only piecewise smooth. For a piecewise smooth curve, we say that it is timelike if its tangent vector at every non-break point is timelike and if all its half tangents (left and right ones) are all timelike and points in the same timecone. Similarly for a null or causal curve. Note that by definition, a causal curve is a "regular" curve, in the sense that its tangent vector at every non-break point is nonzero, nor does any all its half-tangents.

A priori, there is no intrinsic way to distinguish one timecone from the other. To remediate this, we need an extra structure.

Definition 8.7. (M, g) is said to be time-orientable if there exists a global timelike vector field T on M . A choice of time-orientation is the prescription of such a global timelike vector field T . We then say that (M, g) is time-oriented by T .

Remark 8.8. The standard notion of orientation for manifolds and the notion of time-orientation for a Lorentzian manifold are unrelated, cf [O'N83], p145.

Remark 8.9. As for the standard orientation, every Lorentzian manifold admits a double cover³⁰ that is time-orientable, cf [O'N83, Chapter 7].

In the remainder all the Lorentzian manifolds that we will consider will be time-orientable. In fact, we will give a name to those.

Definition 8.8. A spacetime is a triple (M, g, T) where (M, g) is a connected Lorentzian manifold (of dimension $n \geq 2$) and T is a global timelike vector field on M .

For simplicity³¹, on top of being time-orientable, we will assume also that our spacetimes are orientable in the rest of the lectures. (Thus, for us, all spacetimes will be orientable by assumption, and we may omit to say it later). Moreover, we often write (M, g) instead of (M, g, T) .

A causal vector $v \in T_p M$ is called *future-oriented* if $g(v, T_p) < 0$, i.e. v and T_p points in the same timecone (if v is null then X does not belong to the timecone of T). Similarly for *past-oriented* and again curves inherit the labels if their tangent vectors at each point have it.

We have easily

³⁰i.e. a smooth map $k: \tilde{M} \rightarrow M$, where \tilde{M} is a smooth manifold, such that each $p \in M$ has a connected neighborhood \mathcal{U} such that $k^{-1}(\mathcal{U})$ has two connected components and k maps each of them diffeomorphically onto \mathcal{U} .

³¹While time-orientability is important to talk about causality, most constructions below and all the physical interpretations do not need the spacetime manifolds to be orientable.

Proposition 8.5. *Let (M, g) be a spacetime. Then, there exists a globally defined timelike vector field on M whose flow is defined on $M \times \mathbb{R}$, i.e. the vector field is complete.*

Proof. Since M is time-oriented, there exists a globally timelike vector field T_1 . Using Corollary 8.6, we can introduce a complete Riemannian metric h on M . Let then $T = \|T_1\|_h^{-1} T_1$. We claim that T is complete. Indeed, let γ an integral curve of T defined on (t_-, t_+) . If $t_+ < \infty$ then the integral curve restricted to $[a, t_+)$ ($a \in (t_-, t_+)$) is contained in a compact set, since it has bounded length for the Riemannian metric. Hence, the curve is extendible through t_+ a contradiction. \square

8.7.2 Local Lorentzian geometry

We have

Lemma 8.13. *Let $p \in M$ and $\beta: [0, b] \rightarrow T_p M$ a piecewise smooth curve at p (with values in $T_p M$ not in M !) such that*

1. $\beta(0) = 0$,
2. $\alpha = \exp_p \beta$ is well defined on $[0, b]$,
3. $\alpha = \exp_p \beta$ is timelike.

Then, for all $t \in]0, b]$, $\beta(t)$ is timelike and remains in a single timecone of $T_p M$.

Remark 8.10. *Recall that, for $t \in [0, b]$, $\exp_p \beta(t)$ is well defined if the geodesic $\gamma_{\beta(t)}$ starting at p and with initial tangent vector $\beta(t)$ is defined on $[0, 1]$.*

Proof. Note first that that if $\beta(t)$ is timelike for all $t > 0$, then β must stay in a single timecone (since there are only two timecones and there are disconnected).

In the following, we write that property P holds *initially* for “there exists $\epsilon > 0$, such that for all $t \in]0, \epsilon]$, Property P holds”. Assume first that β is smooth. Since $\beta'(0) (= \alpha'(0))$ is timelike and $\beta(t) = t\beta'(0) + O(t^2)$, initially β is in a single timecone, denoted \mathcal{C} . Note that the position vector field \tilde{P} on $T_p M$ is by definition radial and points in the same direction as β in particular, it is timelike in \mathcal{C} . Thus, initially $g(\beta', \tilde{P}) < 0$. Since $D\tilde{q} = 2\tilde{P}$, we have

$$\frac{d}{dt} \tilde{q} \circ \beta = 2g(\beta', \tilde{P}),$$

and initially $\frac{d}{dt} \tilde{q} \circ \beta$ is strictly negative.

By the Gauss lemma,

$$g(\alpha', P) = g(\beta', \tilde{P})$$

and so initially, $g(\alpha', P) < 0$.

Consider the set E all τ such that for all $0 < t \leq \tau$,

$$\frac{d}{dt} \tilde{q} \circ \beta(t) \leq 0.$$

By construction, E is a closed in $]0, b]$ and it is not-empty by the above. Let $a = \sup E$. By the above, we know that $a > 0$. Moreover, on $]0, a]$, $\tilde{q} \circ \beta$ is decreasing and thus $\tilde{q} \circ \beta < 0$, so β is timelike on $]0, a]$ and must stay in a single timecone. Thus, \tilde{P}_β (the position vector field at β) and P_α (by the Gauss lemma) are timelike on $]0, a]$ and must stay in a single timecone. Since α' and P_α are timelike, from Lemma 8.11,

$g(\alpha', P)$ cannot vanish and thus, it follows that $g(\alpha', P) < 0$ (since true initially) and then, as above, that $g(\beta', \tilde{P}) < 0$ and finally that $\frac{d}{dt} \tilde{q} \circ \beta < 0$.

Thus E is open and hence $E =]0, b]$. Hence β is timelike for all t and therefore stays in a single timecone.

Assume now that β is merely piecewise smooth. We know from step 1 that on its first smooth segment β stays in \mathcal{C} . Thus at the first break t_1 , $g(\alpha'(t_1^-), P_1) < 0$, where $P_1 = d \exp_p(P_\beta(t_1))$. Since α is timelike, $\alpha'(t_1^+)$ and $\alpha'(t_1^-)$ must point in the same direction, and thus $g(\alpha'(t_1^+), P_1) < 0$. It then follows that $\frac{d}{dt} (\tilde{q} \circ \beta)$ does not change sign at breaks and the rest of the argument follows similarly. \square

Exercise 8.5. Prove the analogue of the above lemma replacing timelike and timecone by causal and causal cone.

The following proposition is the Lorentzian analogue of Proposition 8.2.

Lemma 8.14. Let (M, g) be a Lorentzian manifold and \mathcal{U} be a normal neighborhood of p . If $k \in \mathcal{U}$ and there exists a timelike curve from p to k in \mathcal{U} , then the radial geodesic segment σ from p to k is the unique (up to reparametrization) longest timelike curve in \mathcal{U} from p to k .

Proof. Let α be a timelike curve in \mathcal{U} from p to k . Let r be the radius function of M at p which we recall is defined as

$$r(k) = \left| g(\exp_p^{-1}(k), \exp_p^{-1}(k)) \right|^{-1/2}.$$

In other words $r(k)$ gives the norm of the tangent vector to the geodesic σ going from p to k in an affine time of 1. In yet another words, $r(k)$ gives the norm of the position vector field. In yet yet another words, $r(k) = L(\sigma)$, where σ is the unique geodesic in \mathcal{U} from p to k .

Note that r is smooth apart at p and at the local null cone of p .

By the previous lemma $\beta = \exp_p^{-1} \circ \alpha$ lies in a single timecone C . Hence, apart at $t = 0$, α lies in a region on which $U = \frac{p}{r}$ is a unit timelike vector field. In particular, U is timelike at k , so σ is timelike at k and hence just timelike (since $g(\alpha', \alpha')$ is constant along α). Let now

$$\alpha' = -g(\alpha', U)U + N,$$

where N is a (spacelike) vector field on α orthogonal to U . Then,

$$|g(\alpha', \alpha')|^{1/2} = [(g(\alpha', U))^2 - (g(N, N))]^{1/2} \leq |g(\alpha', U)|.$$

Since $Dq = 2P$ and $r = (-q)^{1/2}$, we have $Dr = -\frac{p}{r} = -U$. By the previous lemma, $g(\beta, \tilde{P})$, hence $g(\alpha', U)$ is negative, hence

$$|g(\alpha', U)| = -g(\alpha', U) = \frac{d}{dt} g(r \circ \alpha).$$

Consequently,

$$L(\alpha) = \int |g(\alpha', \alpha')|^{1/2} dt \leq r(k) = L(\sigma).$$

If the length are equal, then $N = 0$ which implies that α is a monotone reparametrization of σ . \square

Again, this lemma holds for piecewise smooth curves (again assuming that, by definition, piecewise smooth timelike curves have their half-tangents at break points pointing in the same timecone).

9 Deformation of curves

9.1 Variations of a curve

Definition 9.1. A (smooth) variation of a curve segment $\alpha : [a, b] \rightarrow M$ is a two parameter (smooth) mapping

$$\begin{aligned} x : [a, b] \times (-\delta, \delta) &\rightarrow M \\ (u, v) &\rightarrow x(u, v) \end{aligned}$$

such that $x(u, 0) = \alpha(u)$ for $u \in [a, b]$.

The u parameter curves of a variation are called longitudinal and the v parameters curve transverse. The vector field V on α given by $V(u) = x_v(u, 0)$ is called the *variation vector field* of α .

If $x(a, s) = \alpha(a)$ and $x(b, s) = \alpha(b)$ for all $s \in (-\delta, \delta)$, we say that x is a variation with *fixed endpoints* (in which case V must vanish at the endpoints).

One easily have

Lemma 9.1. Given any curve $\alpha : [a, b] \rightarrow M$ and V a vector field on α , there exists a variation of α whose variation vector field is V . Moreover, if V vanishes at the endpoints, there exists a fixed endpoints variation of α whose variation vector field is V .

Proof. Just take $x(t, s) = \exp_{\alpha(t)}(sV)$. Since α is compact, there exists a $\delta > 0$ such that x is well-defined on $[a, b] \times (-\delta, \delta)$. \square

9.2 From causal to timelike curves

Lemma 9.2. Let α be a causal curve segment in a Lorentzian manifold and let x be a variation of α with variation vector field V . If $g(V', \alpha') < 0$, then for all sufficiently small $v > 0$, the longitudinal curve α_v of x is timelike.

Proof. Since α is causal,

$$g(x_u, x_u)(u, 0) = g(\alpha', \alpha') \leq 0.$$

But α is defined on a closed interval $[a, b]$ and

$$\partial_v g(x_u, x_u)|_{v=0} = 2g(x_{uv}, x_u)|_{v=0} = 2g(x_{vu}, x_u)|_{v=0} = 2g(V', \alpha') < 0.$$

Thus, if $v > 0$ is sufficiently small, then $g(x_u, x_u) < 0$ for all u . \square

By gluing two smooth curves, one often ends up with a piecewise smooth curve. Thus, it seems natural to consider piecewise smooth variations of piecewise smooth curves.

Definition 9.2. A variation x of a piecewise smooth curve $\alpha : [a, b] \rightarrow M$ is a continuous map of the form

$$\begin{aligned} x : [a, b] \times (-\delta, \delta) &\rightarrow M \\ (u, v) &\rightarrow x(u, v) \end{aligned}$$

with $x(u, 0) = \alpha(u)$ for $u \in [a, b]$ and such that for breaks

$$a = u_0 < u_1 < \dots < u_k < b = u_{k+1}$$

the restriction of x to each set $[u_{i-1}, u_i] \times (-\delta, \delta)$ is smooth.

A priori, x and α could have different breaks, but we can always add trivial breaks to x and α (i.e. add points to the partition of $[a, b]$ corresponding to α or x where actually there is no regularity issue).

Note that x_ν is then smooth on each $[u_{i-1}, u_i] \times (-\delta, \delta)$. In particular, the variation vector field $V : u \rightarrow x_\nu(u, 0)$ is continuous³² on each $[u_{i-1}, u_i]$ and hence on $[a, b]$. On the other hand, the velocity vector field α' will generally have a discontinuity at each break u_i . The discontinuity is measured by the tangent vector

$$\Delta\alpha'(u_i) = \alpha'(u_i^+) - \alpha'(u_i^-) \in T_{\alpha(u_i)}M.$$

- Exercise 9.1.**
1. Prove that Lemma 9.2 is valid for piecewise smooth curve (with the usual hypothesis on the half-tangents at each break points).
 2. Prove that the conclusion Lemma 9.2 still holds if $g(V', \alpha') \leq 0$ and strictly negative whenever α' is null.

This implies the following result.

Lemma 9.3. *In a Lorentzian manifold, if α is a causal curve that is not a null pregeodesic (cf Exercise 3.13) from p to q , then there exists a timelike curve from p to q arbitrarily close to α .*

Remark 9.1. *Here arbitrarily close means "constructed out of a fixed endpoints variation of α " so that the longitudinal curves are all timelike for $\nu > 0$. In particular, given any open set containing α , one can choose such a timelike curve lying in the open set.*

Proof. Without loss of generality, assume that α is defined on $[0, 1]$. Assume first that $\alpha'(1)$ is timelike. Let W be the vector field obtained by parallel translation of $\alpha'(1)$ along α . Then, W and α' are always in the same causal cone, and since W is timelike $g(W, \alpha') < 0$. Since $\alpha'(1)$ is timelike, by continuity, there exists $\delta, \epsilon > 0$, such that $g(\alpha', \alpha') < -\delta$ on $[1-\epsilon, 1]$. Let f be any smooth function on $[0, 1]$ that vanishes at the endpoints and such that $f' > 0$ on $[0, 1-\epsilon]$. Set $V = fW$, then $g(V', \alpha') = f'g(W, \alpha')$ is negative on $[0, 1-\epsilon]$. Let x be a fixed endpoint variation with variation vector field V . Then for $\nu > 0$ small enough, the longitudinal curve is timelike on $[0, 1-\epsilon]$ by the previous lemma, and on $[1-\epsilon, 1]$ simply by continuity. Now if α is timelike at 0, the above proof can be repeated and if α is timelike at a nonendpoint s , then we can apply the above result on $[0, s]$ and $[s, 1]$. (Note that if α is regular at s and the above construction yields a timelike curve that is regular (no break point) at s).

Thus, we are left with the case where α is a piecewise smooth null curve. Assume first that α is a smooth null curve. Differentiating $g(\alpha', \alpha') = 0$, we obtain that $g(\alpha'', \alpha') = 0$. Moreover α'' cannot be colinear everywhere on α to α' , otherwise (cf Exercise) α could be reparametrized to be a null geodesic. Thus the function $g(\alpha'', \alpha'')$ is not identically zero, since orthogonal null vectors are necessarily colinear. Let W be a parallel timelike vector field on α in the same causal cone as α' at each point, i.e. $g(W, \alpha') < 0$. Let $V = fW + h\alpha''$, where f and h are functions vanishing at the endpoints that will be determined below so that $g(V', \alpha') < 0$. From $g(\alpha'', \alpha') = 0$, we have

$$g(\alpha''', \alpha') = -g(\alpha'', \alpha'')$$

³²Note that here V is given one value at the break points. This is because since we require $x_{[u_{i-1}, u_i]}$ to be smooth, we have $V(u_i) = \partial_\nu x(u_i, \nu)$, i.e. we can take a partial derivative after evaluation at u_i .

and hence

$$\begin{aligned} g(V', \alpha') &= f'g(W, \alpha') + h'g(\alpha'', \alpha') + hg(\alpha''', \alpha') \\ &= f'g(W, \alpha') - hg(\alpha'', \alpha'') \\ &= g(W, \alpha')(f' - hk), \end{aligned}$$

where

$$k := \frac{g(\alpha'', \alpha'')}{g(W, \alpha')}.$$

Since k is not identically zero, so there exists a smooth function h vanishing at the endpoints such that

$$\int_0^1 khdu = -1$$

Let now $f(u) = \int_0^u (kh + 1)du$. Then, f vanishes at the endpoints, and $f' = kh + 1 > kh$. Consequently $g(V', \alpha') < 0$.

Now, if α is a piecewise smooth null curve, any non-pregeodesic segment can be deformed to a timelike curve, and then by the first argument we can deformed the whole curve to a timelike curve. Hence, we are left with the case of piecewise null curve for which each segment is a pregeodesic. By reparametrization, we might even assume that each segment is a null geodesic, and without loss of generality, we can consider only the case of two segments. Let us thus assume that α is a null geodesic on $[0, s]$ and $[s, 1]$ with a break at s . Let W be the vector field obtained by parallel transportation of $\Delta\alpha'(s) = \alpha'(s^+) - \alpha'(s^-)$. We have $g(W, \alpha'(s^-)) = g(\alpha'(s^+), \alpha'(s^-)) < 0$ because at breaks, we have assumed by definition that semi-tangents points in the same direction, and because if $= 0$, then there are no breaks (in the sense that null orthogonal vectors must be collinear and that a change a parametrization then allows to remove the break). On the other hand, $g(W, \alpha')$ is constant on $[0, s^-]$ by parallel transport and α being a geodesic on that segment. Thus, $g(W, \alpha')$ is negative on $[0, s^-]$. Similarly, it is positive on $[s^+, 1]$. Now choose a piecewise smooth function f that vanishes at the endpoints (i.e. at 0 and 1) and such that $f' > 0$ on $[0, s_-]$ and $f' < 0$ on $[s_+, 1]$. Then, for $V = fW$, we have $g(V', \alpha') < 0$. \square

Exercise 9.2. Give an example of 2 points in Minkowski space that can be joined by a null curve but cannot be joined by a timelike curve.

9.3 First variation

Let $x : [a, b] \times (-\delta, \delta) \rightarrow M$ be a variation of a curve segment α . For each $v \in (-\delta, \delta)$, let $L_x(v)$ be the length of the longitudinal curve $u \rightarrow x(u, v)$. Then $L_x(v)$ is a real-valued function with $L_x(0)$ the length of α . If $g(\alpha', \alpha') \neq 0$ on α , then for δ small enough, L_x is a smooth function (cf below).

Note that $g(\alpha', \alpha') > 0$ on α means that α is spacelike and $g(\alpha', \alpha') < 0$ means that it is timelike. Let ϵ be the sign of α , i.e. $\epsilon = \text{sign}(g(\alpha', \alpha')) = \{-1, 0, +1\}$.

Lemma 9.4. Let x be a variation of a curve segment $\alpha : [a, b] \rightarrow M$ with $g(\alpha', \alpha') \neq 0$. If $L_x(v)$ is the length of $u \rightarrow x(u, v)$ then,

$$L'_x(0) = \epsilon \int_a^b \frac{g(\alpha', V')}{|g(\alpha', \alpha')|^{1/2}} du,$$

where ϵ is the sign of α and V is the variation vector field of x .

Proof.

$$L_x(v) = \int_a^b |g(x_u(u, v), x_u(u, v))|^{1/2} du.$$

In particular, $|g(x_u(u, v), x_u(u, v))|^{1/2}$ is smooth for δ small enough since it does not vanish by continuity. Then,

$$L'_x(0) = \int_a^b \frac{d}{dv} |g(x_u(u, v), x_u(u, v))|^{1/2} du$$

where $|g(x_u(u, v), x_u(u, v))|^{1/2} = (\epsilon g(x_u(u, v), x_u(u, v)))^{1/2}$ and we compute

$$\begin{aligned} \frac{d}{dv} |g(x_u(u, v), x_u(u, v))|^{1/2} &= \frac{1}{2(\epsilon g(x_u(u, v), x_u(u, v)))^{1/2}} \epsilon \cdot 2g(x_u, x_{uv}) \\ &= \frac{\epsilon g(x_u, x_{vu})}{|g(x_u(u, v), x_u(u, v))|^{1/2}}, \end{aligned}$$

using $x_{uv} = x_{vu}$. Just evaluate at $v = 0$ to conclude. \square

Exercise 9.3. Check that the previous lemma holds for x a piecewise smooth variation of α a piecewise smooth curve.

Proposition 9.1 (First variation formula³³). Let $\alpha : [a, b] \rightarrow M$ be a piecewise smooth curve segment with constant speed $c = |g(\alpha', \alpha')|^{1/2} \neq 0$ and sign ϵ . If x is a piecewise smooth variation of α , with break points at $(u_i)_{1 \leq i \leq k}$ such that $u_0 = a < u_1 < \dots < u_i < \dots < u_k < u_{k+1} = b$, then

$$L'_x(0) = -\frac{\epsilon}{c} \int_a^b g(\alpha'', V) du - \frac{\epsilon}{c} \sum_{i=1}^k g(\Delta \alpha'(u_i), V(u_i)) + \left[\frac{\epsilon}{c} g(\alpha', V) \right]_a^b.$$

Proof. We simply split the previous integral at the break points and integrate by parts in u , using that

$$g(\alpha', V') = \frac{d}{du} g(\alpha', V) - g(\alpha'', V).$$

\square

Remark 9.2. For a fixed endpoint variation x , i.e. $x(a, v) = p$ and $x(b, v) = q$ for all $v \in (-\delta, \delta)$, the variation vector field V vanishes at a and b and thus the last term also vanishes.

Corollary 9.1. A piecewise smooth curve α of constant speed $c > 0$ is an (unbroken) geodesic if and only if the first variation of arc length is zero for every fixed endpoints variation of α .

Proof. If α is geodesic then $\alpha'' = 0$ and there are no breaks. Conversely, suppose $L'_x(0) = 0$ for every fixed endpoint variation x . First, we show that each segment $\alpha|_{[u_i, u_{i+1}]}$ is geodesic. It suffices to show that $\alpha'' = 0$ for $u_i < t < u_{i+1}$. Let y be any tangent vector to M at $\alpha(t)$ and let f be a bump function on $[a, b]$ with support

³³The first variation formula will help us to find critical points of the length functional, but do not determine whether these critical points are actual minimizers. For this, one needs to go beyond the first variation, for instance, by computing a second variation formula, at the level of two derivatives of L .

included in $[t-\delta, t+\delta] \subset [u_i, u_{i+1}]$. Let Y be the vector field on α obtained by parallel translation of y and set $V = fY$.

Since $V(a) = V(b) = 0$, the exponential formula $x(u, v) = \exp_{\alpha}(u)(vV(u))$ produces a fixed endpoint variation of α whose variation vector field is V .

Since $L'_x(0) = 0$, and given the support assumption on f , we have

$$0 = \int_{t-\delta}^{t+\delta} g(\alpha'', fY) du.$$

Since this holds for all $y, \delta > 0$ and bump function f , we have $g(\alpha'', y) = 0$ for all y , i.e. $\alpha'' = 0$. It remains to show that the breaks are trivial, so that α is indeed unbroken. As before, let y be an arbitrary tangent vector at $\alpha(u_i)$ and let f be a bump function at u_i with $\text{supp} f \subset [u_{i-1}, u_{i+1}]$. For a fixed variation with vector field fY , the first variation formula gives

$$0 = -\epsilon/cg(\Delta\alpha'(u_i), y)$$

for all y , since $\alpha'' = 0$ on each subsegment. Since g is non-degenerate, this implies that all the breaks are trivial. \square

10 Causality and global Lorentzian geometry

Recall that to distinguish between the future and the past, we say that M is *time oriented* if there exists a globally defined timelike vector field, denoted here T . Then, at every $p \in M$, for $v \in T_pM$ a causal vector, we say that v is future oriented if $g_p(v, T_p) < 0$, which amounts to v and T_p lying in the same time cone if v is timelike. Recall also that the notion of future-oriented or past-oriented descend to curves.

From now on, we will assume that all the Lorentz manifolds we consider are time-oriented. Moreover, we will assume that the manifold is oriented and connected (this ensures that we can define a global volume forme and the existence of a complete Riemannian metric).

Definition 10.1. *Let $p, q \in M$. We say that $p \ll q$ if there is a future pointing timelike curve in M from p to q and $p < q$ if there is a future pointing causal curve in M from p to q . Finally, $p \leq q$ means $p = q$ or $p < q$.*

Definition 10.2. *For $A \subset M$, we define*

$$\begin{aligned} I^+(A) &= \{p \in M : \exists q \in A / q \ll p\}, \\ J^+(A) &= \{p \in M : \exists q \in A / q < p\}, \\ I^-(A) &= \{p \in M : \exists q \in A / p \ll q\}, \\ J^-(A) &= \{p \in M : \exists q \in A / p < q\}. \end{aligned}$$

The set $I^+(A)$ is called the *chronological future* of A , $J^+(A)$ is the *causal future* and similarly for I^- and J^- , replacing future by past.

In the above definition, we require that the curves to be piecewise smooth. Note that in view of the previous section, if there is piecewise smooth causal curve from p to q that is not null pregeodesic, then it can be deformed to be a piecewise smooth timelike curve joining p to q . On the other hand, a piecewise smooth timelike curve from p to q can be smoothed out as follows.

Lemma 10.1. *If there exist smooth timelike, future-oriented, curves from p to q and from q to r , then there exists a smooth timelike future oriented curve from p to r .*

Proof. Let p, q, r be as in the statement of the lemma. Let α be a smooth timelike curve from p to q and β be a smooth timelike curve from q to r . The curve α followed by β is not necessarily smooth at q , but we claim that we can smooth it out and still maintain its timelike character. First, note that this is a local problem at q , and since the condition of being timelike is an open condition, it follows by continuity that it is sufficient to prove it in the case of the Minkowski metric. Let us thus assume that $\eta = g$. Without loss of generality, we can consider coordinates (t, r, ω) on \mathbb{R}^{n+1} such that $q = (0, \dots, 0)$ and we can also assume that $\alpha : [-1, 0] \rightarrow M$ is given in local coordinates by $(s, 0, \dots, 0) + O(s^2)$ and that $\beta : [0, 1] \rightarrow M$ with $\beta(0) = q$. Again, since this is a local problem near $s = 0$, we can neglect the $O(s^2)$ and we shall do so in the following.

From the timelike condition, we have $\dot{\beta}^t > |\dot{\beta}^i|$. In particular, $\tau := t(\beta(1)) > R := r(\beta(1))$. We consider a curve of the form $\gamma(s) = (s, Rf(s), \omega(q))$ defined on $[0, \tau]$ with $f \geq 0$. The lemma then follows since we can choose $f \geq 0$ such that $f(x) = 0$ for $x \leq 0$, $f(\tau) = 1$ and $|f'| < \frac{1}{R}$. For instance, rescaling in τ means we are looking for a function f defined on $[0, 1]$, such that $f(x) = 0$ for $x \leq 0$, $f(1) = 1$ and $|f'| < \frac{\tau}{R}$. Draw on a picture the straight line $t = x$, and then draw the straight line of slope $\frac{\tau}{R}$ passing by $(1, 1)$. Then any convex function lying between the two straightlines and vanishing for $x \leq 0$ will do the job. \square

We have easily from the definitions and Lemma 9.3.

Lemma 10.2. *The following relations hold*

$$I^+(A) = I^+(I^+(A)) = I^+(J^+(A)) = J^+(I^+(A) \subset J^+(J^+(A)) = J^+(A).$$

10.1 Causality and topology

For an open set \mathcal{U} , we denote by $I^+(A, \mathcal{U})$ the chronological future of $A \subset \mathcal{U}$ in the open submanifold $\mathcal{U} \subset M$. In other words, $q \in I^+(A, \mathcal{U})$ if and only if there exists $r \in A$ and a future oriented timelike curve with values in \mathcal{U} joining r to q . In particular, $q \in \mathcal{U}$, so that $I^+(A, \mathcal{U}) \subset I^+(A) \cap \mathcal{U}$.

Exercise 10.1. *Give an example for which the above inclusion is strict.*

Lemma 10.3 (Topological properties of causal sets in convex regions). *Let $\mathcal{C} \subset M$ be open and convex. Then,*

1. *For $p \neq q \in \mathcal{C}$, $q \in J^+(p, \mathcal{C})$ if and only if the unique geodesic from p to q is a future oriented causal curve (similar for I^+ and timelike).*
2. *$I^+(p, \mathcal{C})$ is open in \mathcal{C} (hence in M).*
3. *$J^+(p, \mathcal{C})$ is the closure in \mathcal{C} of $I^+(p, \mathcal{C})$.*
4. *The relation \leq is closed on \mathcal{C} , i.e if $p_n \rightarrow p$ and $q_n \rightarrow q$ with all points in \mathcal{C} then $q_n \in J^+(p_n, \mathcal{C})$ for all n implies that $q \in J^+(p, \mathcal{C})$.*

Proof. Let $p \neq q \in \mathcal{C}$.

Assume first that $q \in I^+(p, \mathcal{C})$. Let then α be a future-oriented timelike curve joining p to q . Since \mathcal{C} is a normal neighborhood of p , we have $\alpha = \exp_p \beta$ for some curve $\beta : [0, 1] \rightarrow T_p M$ with $\beta(0) = 0$. From Lemma 8.13 and the fact that α is timelike,

β stays in a single timecone on $]0, 1]$. This implies that $\overrightarrow{pq} = \beta(1)$ is timelike and future-oriented. Thus, the unique geodesic from p to q is a future oriented timelike curve.

If now $q \in J^+(p, \mathcal{C})$ and γ is a causal curve joining p to q , then either it can be deformed in \mathcal{C} to a timelike curve joining p to q and we run the first argument, or it is null (pre)geodesic, in which case we are also done. This proves the first point.

For the second point, let $q \in I^+(p, \mathcal{C})$. From the first point, the unique geodesic from p to q is timelike. Thus $g(\overrightarrow{pq}, \overrightarrow{pq}) < 0$. The result then follows since all the operations here are continuous, using Lemma 8.6.

Similarly, the third and fourth points follow from the first and the continuity of the map $(p, q) \rightarrow \overrightarrow{pq}$. □

Recall the definition of a continuously extendible curve, Definition 8.6.

We now prove

Lemma 10.4. *A causal curve contained in a compact subset $K \subset \mathcal{C}$ convex is continuously extendible.*

Proof. Let α be causal curve defined on $[0, B)$ whose image is included K . Since K is compact, there exists a sequence (s_i) such that $\alpha(s_i) \rightarrow q$ and $s_i \rightarrow B$ as $i \rightarrow +\infty$. We must show that every such sequence gives the same limit. Let t_i be another sequence, such that $\alpha(t_i) \rightarrow p$. By the previous lemma and the fact that α is causal it follows that $q \in J^+(p, \mathcal{C})$ and $p \in J^+(q, \mathcal{C})$. Now if $p \neq q$, then from the previous lemma, \overrightarrow{pq} is would be both future pointing and past pointing, a contradiction. □

We now prove

Lemma 10.5. *The relation $p \ll q$ is open, that is if $p \ll q$ there exists neighborhood \mathcal{U} of p , and \mathcal{V} of q such that for all $p' \in \mathcal{U}$, $q' \in \mathcal{V}$, $p' \ll q'$.*

Proof. Let σ be a timelike curve from p to q . If \mathcal{C} is a convex neighborhood of q , let q^- be a point of \mathcal{C} on σ before q . Dually, let p^+ be a point of σ between p and q^- and contained in a convex neighborhood \mathcal{C}' of p . By the lemma, $I^+(q^-, \mathcal{C})$ and $I^-(p^+, \mathcal{C}')$ are open in M and have the required property. □

Corollary 10.1. *For any subset $A \subset M$, $I^+(A)$ and $I^-(A)$ are open.*

Proof. Follows directly from the previous lemma. (Let $p \in A$ and $q \in I^+(A)$. Then, by the previous lemma, there exists a neighborhood \mathcal{V} of q including in $I^+(A)$.) □

In Minkowski space, for any set A , $J^+(A)$ is the closure of $I^+(A)$, but it is easy to construct spacetimes where this is no longer the case.

Example: In 1+1 Minkowski space, fix a point p and consider $J^+(p)$. The boundary of $J^+(p)$ consists of two null curves starting at p . Remove a point on these curves and consider the resulting spacetime to get an example where the closure of $I^+(p)$ is much larger than $J^+(p)$.

In general, we only have

Lemma 10.6. *For any subset A ,*

1. $\text{int}(J^+(A)) = I^+(A)$.

2. $J^+(A) \subset \overline{I^+(A)}$, with equality if and only if $J^+(A)$ is closed.

Proof. $I^+(A)$ is open and contained in $J^+(A)$, hence is contained in its interior. If $q \in \text{int}(J^+(A))$, then there exists a convex neighborhood of q , $\mathcal{C} \subset J^+(A)$. Take a point $r \in \mathcal{C}$, such that $r \in I^-(q, \mathcal{C})$. Since $r \in \mathcal{C}$, $r \in J^+(A)$. Hence, $q \in I^+(J^+(A)) \subset I^+(A)$.

Since $I^+(A) \subset J^+(A)$, we have $\overline{I^+(A)} \subset \overline{J^+(A)}$. Hence, if $J^+(A) = \overline{I^+(A)}$, then $J^+(A)$ is closed and if $J^+(A)$ is closed then, $\overline{I^+(A)} \subset \overline{J^+(A)} = J^+(A)$. Thus, the equality case is clear provided that the inclusion assertion in (2) holds.

It suffices to just prove (2) for one point, i.e. we will prove $J^+(p) \in \overline{I^+(p)}$. Let $q \in J^+(p)$. If $q = p$, then clearly $q \in \overline{I^+(p)}$. Otherwise, let α be a causal curve from p to q . Let \mathcal{C} be a convex neighborhood of q and let q^- be a point of $\alpha \in J^-(q, \mathcal{C})$. We have $q \in J^+(q^-, \mathcal{C})$ and by Lemma 10.6,

$$J^+(q^-, \mathcal{C}) = \overline{I^+(q^-, \mathcal{C})}.$$

The result then follows since $I^+(q^-, \mathcal{C}) \subset I^+(J^+(p)) \subset I^+(p)$. \square

10.2 Various limits for sequences of curves

We are interested in the construction of limits of a sequence of causal curves. Since we lack good compactness properties on such sets of curves, the limits can be rough. There are several approaches to this problem:

1. Quasi-limits: this is the notion used in [O'N83]. It has the advantage that it does not require any knowledge of weak compactness or the Arzela-Ascoli theorem, but the definition is quite cumbersome (see next section).
2. Limits from the C^0 topology on curves.
3. Limit curves defined using accumulation points.

We present the quasi-limits in the next section. We will use it later in order to follow O'Neill but we also present briefly limit curves for the interested reader as in [BEE96]. The connection between the C^0 topology on curves and the limit curves is explained in [BEE96].

10.2.1 Quasi-limits

Definition 10.3. Let (α_n) be a sequence of future oriented causal curves in M and let R be a convex covering (cf Definition 8.3) of M . A limit sequence for (α_n) (relative to R) is a finite or infinite sequence of points (p_i) , which are causally related $p_0 < p_1 < \dots$ in M and such that

L_{1a} . For each i , there exists are subsequences ${}^0\alpha_n := \alpha_{\phi^0(n)}$, ${}^1\alpha_n := {}^0\alpha_{\phi^1(n)}$, ${}^2\alpha_n := {}^1\alpha_{\phi^2(n)}$, ..., ${}^i\alpha_n := {}^{i-1}\alpha_{\phi^i(n)}$ such that there exists, for each n , a set of $i+1$ points along the curve ${}^i\alpha_n$, denoted ${}^i\alpha_n(s_{n,0})$, ${}^i\alpha_n(s_{n,1})$, ${}^i\alpha_n(s_{n,2})$, ..., ${}^i\alpha_n(s_{n,i})$, with $s_{n,0} < s_{n,1} < s_{n,2} < s_{n,3} \dots$ such that for all $j \leq i$, as $n \rightarrow +\infty$,

$${}^i\alpha_n(s_{n,j}) \rightarrow p_j.$$

L_{1b} . For each $j < i$, the points p_j , p_{j+1} and for all n , the segments of ${}^i\alpha_n$ restricted to $[s_{n,j}, s_{n,j+1}]$ are all contained in a single convex neighborhood \mathcal{C}_j of the covering R .

L₂. If the p_i are infinite, it is nonconvergent. If they are finite, say $p_1 < p_2 < \dots < p_k$, then they contains more than one point ($k > 1$) and there exists no p_{k+1} such that $p_1 < p_2 < \dots < p_k < p_{k+1}$ and a further subsequence of ${}^k\alpha_n$ satisfies L1 for the extended sequence.

Proposition 10.1. *Let (α_n) be a sequence of future oriented causal curves such that $\alpha_n(0) \rightarrow p \in M$ and (α_n) does not converge to just p (i.e. there exists a neighborhood of p that contains only finitely many of the α_n ³⁴). Then relative to any convex covering R , (α_n) has a limit sequence starting at p .*

Proof. i) Construction of the p_i .

Since M is paracompact³⁵ (since every topological manifold is paracompact cf Prop. 2.3 in Paulin's differential geometry course), it has a locally finite covering R' by open sets B such that each \overline{B} is compact and contained in some member of R . By the hypotheses on (α_n) , we can arrange for R' to be such that it contains B_0 such that infinitely many α_n start in B_0 and leave $\overline{B_0}$. Relabel these curves as ${}^1\alpha_n$ and for each ${}^1\alpha_n$, let s_n be such that ${}^1\alpha_n(s_n)$ is the first point in bdB_0 . Passing to a further subsequence, we can assume that $({}^1\alpha_n(s_n))$ converges to a point $p_1 \in bdB_1$. Now choose $B_1 \in R'$ containing p_1 . If infinitely many ${}^1\alpha_n$ s leave B_1 , we obtain as before a subsequence ${}^2\alpha_n$ whose first departure points from B_1 converges to a point p_2 in bdB_1 . Repeat this step as many times as possible with the additional rule on subsequent choices of B_i , if there is more than one candidate element of R' containing p_i , pick one that has been used fewest times before.

ii) Basic properties and L1.

Clearly, the first two conditions of the definition of limit sequences hold (with C_i any element of R that contains $\overline{B_i}$). Since the relation \leq is closed on \mathcal{C}_i , it follows that $p_{i+1} \geq p_i$. By construction, $p_{i+1} \neq p_i$, hence $p_{i+1} > p_i$.

iii) Checking L2.

If the sequence (p_i) is infinite, we must show that it is non-convergent. Assume that it converges to some q . Let $B \in R'$ such that $q \in B$, then for all sufficiently large i , $p_i \in B$. Since \overline{B} is compact and R' is locally finite, only finitely many number of R' meet B . Hence, some must have been chosen for B_i for infinitely many i . But this violates the additional rule, for B itself was a candidate infinitely many times, but was chosen only finitely many times (since it is open and contains all p_i but a finite number, only a finite number can be in bdB).

Finally, suppose that the sequence of the p_i is finite. Since the construction above cannot continue, there exists a k such that only a finite number of the ${}^k\alpha_n$ leave B_k . Let (α_m) be those trapped in B_k . By Lemma 10.4 (since the closure of B_k is compact) they are extendible. Assume thus that there are defined on closed interval $[0, b_m]$ (by a change of parametrization and continuous extension, we can always assume that). By compactness, we can take a further subsequence and assume that $\alpha_m(b_m)$ converges to some q . If $q = p_k$, then the sequence $p_0 < \dots < p_k$ cannot be extended to still satisfies the first two limit sequence conditions (hence it is a limit sequence). If $q \neq p_k$, then $q > p_k$ and posing $p_{k+1} = q$, we have a limit sequence. \square

³⁴The case where (α_n) just converges to p can happen by taking just one curve passing through p and shrinking the interval to 0

³⁵Recall that this means that M is Hausdorff (séparé in French) and every open cover admits a subcover that is locally finite.

Let now (p_i) be a limit sequence for (α_n) and let λ_i be the (future causal) geodesic from p_i to p_{i+1} in a convex set C_i as in L1. Assembling these segments for all i gives a broken geodesic λ called a *quasi-limit* of (α_n) with vertices p_i . λ is a future broken causal geodesic that starts at p . If the (p_i) are infinite, then by L2, λ is future-inextendible. In the finite case, the curve runs from p to some p_k . When all the curves (α_n) are future inextendible, we have

Lemma 10.7. *A quasi-limit λ of future inextendible curves (α_n) is future inextendible.*

Proof. Let p_i denotes the vertices of λ . If the p_i are infinite, then λ is inextendible, since they must be non-convergent. On the other hand, the limit sequence cannot be finite. Indeed, assume that the limit sequence is given by $p_0 < \dots < p_k$. Let C be a convex neighborhood containing p_k and let $B_k \subset C$ open and whose closure is compact. Let α_n be a subsequence satisfying L1. Now by hypothesis, all the curves are inextendible, thus by Lemma 10.4, they must leave B_k . But then we can pass to a subsequence to get an extra point in BdB_k , contradiction. \square

10.2.2 Limit curves

So far we have defined the notion of causal curves for smooth curves and piecewise smooth curves. The limit operation we will define below will in general not preserve these level of regularity so we need to introduce a broader class of causal curves.

Definition 10.4. *A continuous curve $\gamma : I \rightarrow M$ is said to be a future directed causal curve if for each $t_0 \in I$, there exists $\epsilon > 0$ such that $(t_0 - \epsilon, t_0 + \epsilon) \subset I$, and a convex neighborhood U of $\gamma_{(t_0 - \epsilon, t_0 + \epsilon)}$ such that given any t_1, t_2 satisfying $t_0 - \epsilon < t_1 < t_2 < t_0 + \epsilon$, there is a smooth future directed causal curve in U from $\gamma(t_1)$ to $\gamma(t_2)$.*

Exercise 10.2. 1. Let $M = \mathbb{S}^1 \times \mathbb{S}^1$ be the 2-torus. Let (θ_1, θ_2) be standard coordinates on M and consider the metric

$$g = -d\theta_1^2 + d\theta_2^2.$$

Prove that for any $p, q \in M$, $p \ll q$. (In the above definition, if we remove the reference to the convex neighborhood U , then every continuous curve would be causal and of course, we do not want that.)

2. *Prove that piecewise smooth causal curves are causal in the sense of Definition 10.4.*
3. *Prove that if $\gamma : I \rightarrow M$ is a continuous causal curve, then for any $s_0, s_1 \in I$, $s_0 < s_1$, there exists a piecewise smooth future directed causal curve λ joining $\gamma(s_0)$ to $\gamma(s_1)$.*

We can now define the notion of limit curves.

Definition 10.5. *A continuous curve $\gamma : I \rightarrow M$ is a limit curve of the sequence $\{\gamma_n\}$ if there is a subsequence $\{\gamma_m\}$ such that for all $p \in \gamma(I)$, each neighborhood of p intersects all but a finite number of curves of the subsequence $\{\gamma_m\}$. The subsequence $\{\gamma_m\}$ is said to distinguish the limit curve γ .*

Consider normal coordinates and a normal neighborhood \mathcal{U} around some point p . Let $t = x^0$ be the first coordinate, so that $g_{tt}(p) = -1$. For any $K > 0$, consider the auxiliary Lorentzian metric

$$g_0 = -K^2 dt^2 + \sum_{i=1}^n (dx^i)^2.$$

Shrinking the normal neighborhood \mathcal{U} is necessary and taking K large enough, we may assume that the null cone of g is small than the light cone of g_0 , i.e. causal vectors for g are always causal for g_0 . Let now γ be a continuous curve such that $\gamma(0) = p$. Let $s_0 < 0 < s_1$ be such that $\gamma_{[s_0, s_1]} \subset \mathcal{U}$. Then $t(\gamma(s_0)) < t(\gamma(s_1))$ since t is strictly increasing along future directed causal curves and since we can join $\gamma(s_0)$ to $\gamma(s_1)$ by a future directed piecewise smooth causal curve. It follows that t is strictly increasing along γ and in particular invertible (with continuous inverse!) and we can use it to change the parametrization of γ . Assume then that γ is parametrized by t . Since γ is causal for g , it must be causal for g_0 . The causal cones of g_0 are those of a flat spacetimes and this implies that for any t_1, t_2 ,

$$\|\gamma(t_1) - \gamma(t_2)\| \leq K'|t_1 - t_2|,$$

where $\|\gamma(t_1) - \gamma(t_2)\| = \|x^i \circ \gamma(t_1) - x^i \circ \gamma(t_2)\|$ only measures the spatial part of γ . (Assume for instance that $t_2 = 0$, then this simply says that $\gamma(t_1)$ lies in the future cone emanating from $\gamma(t_2)$.) It follows that in this coordinates and with this parametrization, γ is actually Lipschitz (in the sense that all its components are at least Lipschitz functions of t). Since Lipschitz functions are differentiable almost everywhere, we can in particular define almost everywhere its tangent vector.

We now introduce an (auxiliary) complete Riemannian metric h on M . From the above, for every causal curve γ defined on a compact interval, we can define its length $L_h(\gamma)$ by the usual formula, except that γ' is defined only almost everywhere along the curve. This length is actually finite, as we can use the above bounds on each time interval, after splitting the curve in sufficiently small pieces so that on each one we have access to bounds as derived above.

Finally, an inextendible causal curve must have unbounded length (cf Hopf-Rinow). Using Arzela's Theorem, we can then prove

Proposition 10.2. *Let $\{\gamma_n\}$ be a sequence of inextendible causal curves admitting some $p \in M$ as accumulation point. Then, there is a nonspacelike limit curve γ of the sequence γ_n such that $p \in \gamma$ and γ is inextendible.*

Proof. (sketch) Let h be an auxiliary complete Riemannian metric on M as above and parametrize each γ_n by arc length with respect to h and such that, after taking a subsequence if necessary, $\gamma_n(0) \rightarrow p$, $n \rightarrow +\infty$. Then each γ_n is defined on \mathbb{R} and

$$d_h(\gamma_m(t_1), \gamma_m(t_2)) \leq |t_1 - t_2|, \quad t_1, t_2 \in \mathbb{R}.$$

It follows that the sequence is equicontinuous. Moreover, for any fixed t_0 , the sequence restricted to $[-t_0, t_0]$ is uniformly bounded. By Arzela's theorem (exercise: state the appropriate version of this theorem), there exists a subsequence converging uniformly on compact set to a continuous curve γ . Moreover, we have

$$d_h(\gamma(t_1), \gamma(t_2)) \leq |t_1 - t_2|, \quad t_1, t_2 \in \mathbb{R}.$$

It remains to prove that γ is causal and inextendible.

To prove that γ is causal, fix t_1 and let U be convex with $\gamma(t_1) \in U$. Choose $\delta > 0$ so that the open ball for h , $B_h(\gamma(t_1), \delta)$ is in U . For $t_2 \in (t_1, t_1 + \delta)$, we have $\gamma_n[t_1, t_2] \subset U$ and using that the relation $p \leq q$ is closed in U (cf Lemma 10.6), we obtain that $\gamma(t_1)$ can be joined by a future directed causal curve to $\gamma(t_2)$.

Assume finally that γ is not future inextendible and thus that $\gamma(t) \rightarrow q$ as $t \rightarrow +\infty$. Take U a convex neighborhood of q with compact closure and such that in U ,

(x^α) are coordinates with x^0 strictly increasing along timelike curve (as in the above discussion). There exists a t_1 such that γ restricted to $[t_1, +\infty)$ lies U .

On the other hand, using the length function L_h and bounds as above, one can prove (exercise) that a causal curve with initial $x^0 = x^0(\gamma(t_1))$ and ending at a point with $x^0 = x^0(q)$ has length uniformly bounded by some L . On the other hand, for k sufficiently large the component x^0 of γ_k restricted to $[t_1 + 1, t_1 + L + 2]$ must lie within $[x^0(\gamma(t_1)), x^0(q)]$ by uniform convergence, continuity and the fact that x^0 is strictly increasing along γ . In view of the arc length parametrization, the length of γ_k restricted to $[t_1 + 1, t_1 + L + 2]$ is $L + 1$ which is a contradiction. \square

10.3 Causality conditons

First, an easy remark

Lemma 10.8. *A compact Lorentzian manifold admits a closed timelike curve.*

Proof. By compactness, the open covering $I^+(p)$ admits a finite subcover $I^+(p_1), \dots, I^+(p_k)$. If $p_1 \in I^+(p_i)$, for $i \neq 1$, then $I^+(p_1) \subset I^+(I^+(p_i)) = I^+(p_i)$. In that case (note that then $k > 1$), we can remove $I^+(p_1)$ and still obtain a open cover. Wlog, assume thus that p_1 is not included into any of the other $I^+(p_i)$. Then, $p_1 \in I^+(p_1)$, i.e. there exists a closed timelike curve at p . \square

Closed timelike curves are bad: from the physics point of view, and from evolution problems point of view. We want to exclude them. We want in fact a slightly stronger criterion.

Definition 10.6. *The strong causality condition is said to hold at $p \in M$ provided that given any neighborhood \mathcal{U} of p , there exists a neighborhood $V \subset \mathcal{U}$ of p such that every causal curve segment with endpoints in V lies entirely in U .*

Exercise 10.3. 1. *Prove that, in the above definition, one can use piecewise smooth causal curves or continuous causal curves equivalently.*

2. *Prove that if there exists a closed timelike curve passing through p , then M is not strongly causal at p .*

Lemma 10.9. *Suppose that the strong causality condition holds on a compact $K \subset M$. If α is a future-inextendible causal curve that starts in K , then α eventually leaves K never to return; that is there is an $s > 0$ such that $\alpha(t) \notin K$, for all $t \geq s$.*

Proof. Assume that α persistently returns to K . Assume that α is defined on $[0, B)$. Then, there exists a sequence $s_i \rightarrow B$, with $\alpha(s_i) \in K$. Taking a subsequence, we can assume that $\alpha(s_i)$ converges to some $p \in K$. Since α has no future endpoint, there must exist another sequence $t_j \rightarrow B$, such that $\alpha(t_j)$ does not converge to p . Taking yet another subsequence, we can assume that there exists a neighborhood U of p such that $\alpha(t_j) \notin U$ for all j . Since s_j and t_j both converges to B , they have subsequences that alternates $s_1 < t_1 < s_2 < t_2 \dots$. The curves $\alpha|_{[s_k, s_{k+1}]}$ are almost closed at p : Given any neighborhood $V \subset U$ of p , for k large enough the ends $\alpha(s_k)$ and $\alpha(s_{k+1})$ are contained in V and yet the curve escape U since $\alpha(t_k) \notin U$. \square

Lemma 10.10. *Suppose the strong causality condition holds on a compact subset K . Let (α_n) be a sequence of future oriented causal curve segments in K such that $\alpha_n(0) \rightarrow p$ and $\alpha_n(1) \rightarrow q \neq p$. Then, there exists a future oriented causal broken geodesic λ from p to q and a subsequence α_m whose length converges with the limit satisfying $\lim_{m \rightarrow +\infty} L(\alpha_m) \leq L(\lambda)$.*

Proof. From Proposition 10.1, the α_n admits a limit sequence p_i starting at p . If p_i is infinite, the corresponding quasi-limit λ is future inextendible and thus, by the above lemma, must leave K . In particular, some of the α_n must leave K , a contradiction. Thus the limit sequence is finite, starts at p and ends at q (again applying L2). The quasi-limit with these vertices is a broken geodesic from p to q . Thanks to L1b, we can consider a subsegment in some convex set C_i . On each convex set, the length separating two points $p < q$ is bounded above by the length of the corresponding geodesic. Thus,

$$L(\alpha_m|_{[s_{m,i}, s_{m,i+1}]}) \leq |p_{mi} \vec{p}_{m,i+1}|,$$

where $p_{mi} = \alpha(s_{mi})$. Moreover, we know that the length (in fact $\vec{p}q$ itself) between two points $p < q$ in a convex set depends continuously on p and q . Thus, $|p_{mi} \vec{p}_{m,i+1}| \rightarrow |p_i, \vec{p}_{i+1}|$. Summing on all the segments and taking the limit $m \rightarrow +\infty$ then gives the lemma. \square

10.4 Achronal sets

Definition 10.7. $A \subset M$ is said to be achronal if there exists no $p, q \in A$ such that $p \ll q$ and acausal if there exists no p, q such that $p < q$. Given A achronal, $D^+(A)$ is by definition, the set of all $p \in M$ such that every past inextendible causal curve through p meets A . Similarly, we define $D^-(A)$ as the set of all $p \in M$ such that every future inextendible causal curve through p meets A .

Definition 10.8. The edge (marge in French, not to be confused with the boundary in topological sense) of an achronal set A consists of all points p in the closure of A such that every neighborhood \mathcal{U} of p contains a timelike curve from $I^-(p, \mathcal{U})$ to $I^+(p, \mathcal{U})$ that does not meet A .

Exercise 10.4. Let A be the set $[0, 1] \times \{t = 0\}$ in 2-dimensional Minkowski space. What is the edge of A ?

Recall the definition of a topological hypersurface of M .

Definition 10.9. S is a topological hypersurface of M if for all $p \in S$, there exists a neighborhood U of p in M and a homeomorphism $\phi : \mathcal{U} \rightarrow \mathbb{R}^n$ such that $\phi(\mathcal{U} \cap S) = \phi(\mathcal{U}) \cap \Pi$ where Π is a hyperplane in \mathbb{R}^n .

Remark 10.1. The point is that topological hypersurface are rougher than differentiable submanifold. In particular, they do not necessarily admits tangent planes.

Lemma 10.11. Let $p \in M$ and \mathcal{U} a neighborhood of p . Then, there exists a coordinate chart (\mathcal{N}, ξ) of p such that

1. $\xi(\mathcal{N})$ has the form $(a - \epsilon, b + \epsilon) \times N \subset \mathbb{R}^1 \times \mathbb{R}^{n-1}$ with $a < b, \epsilon > 0$.
2. For any $(x^i) \in N$, the curve $s \rightarrow \xi^{-1}(s, x^i)$ is timelike.
3. The slice $x^0 = a$ of \mathcal{N} is in $I^-(p, \mathcal{U})$, the slice $x^0 = b$ of \mathcal{N} is in $I^+(p, \mathcal{U})$.

Proof. Consider a normal neighborhood U of p with $U \subset \mathcal{U}$. Let ξ be a normal coordinate system at p with $g(\partial_{x^0}, \partial_{x^0})(p) = -1$ and ∂_{x^0} future pointing. Consider the region \mathcal{N} such that, $|x^0| \leq \delta$ and $|(x^i)^2| \leq 1/8\delta^2$. Let V be defined by $V = \exp_p^{-1}(\mathcal{N})$ and let V_ρ be the subset of vectors which are timelike for the auxiliary Minkowski

metric $\eta_{1/4} := -1/4 dx^0 \otimes dx^0 + dx^i \otimes dx^i$. Now, let β be a curve in V_p which is timelike for $\eta_{1/4}$. Let $y^\alpha(t)$ be the coordinates of β

$$\beta(t) = y^\alpha(t) [\partial_{x^\alpha}]_p$$

and let $\gamma = \exp_p \circ \beta$. Then the normal coordinates of $\gamma(t)$ are also given by $y^\alpha(t)$ by construction.

Moreover, by construction, the coordinates of γ' are given by $(y^\alpha)'$ and

$$-1/4((y^0)')^2 + ((y^i)')^2 < 0.$$

Since at p , the metric in normal coordinates coincides with the Minkowski metric, by Taylor expansion, we have that,

$$g_{\alpha\beta}(x) = \eta_{\alpha\beta} + O(|x|)$$

and hence, for δ small enough,

$$|g_{\alpha\beta}(x) - \eta_{\alpha\beta}| \leq C\delta,$$

for some C .

This implies that

$$g(\alpha', \alpha') = \eta(\alpha', \alpha') + O(\delta)(|y'|^2),$$

where $|y'|^2 = ((y^i)')^2 + ((y^0)')^2 \leq 5/4((y^0)')^2$. Since $\eta(\alpha', \alpha') < -3/4((y^0)')^2$, we have that for δ small enough, $g(\alpha', \alpha') < 0$.

In particular, \exp_p maps each $x^i = \text{const}$ curve to a timelike curve in \mathcal{N} . We claim that the points of normal coordinates $(-\delta, x^i)$ and (δ, x^i) lies in respectively $I^-(p, \mathcal{U})$ and $I^+(p, \mathcal{U})$. Indeed, let $v = \delta[\partial_{x^0}]_p + x^i[\partial_{x^i}]_p$ and consider the curve $\beta: t \rightarrow tv$. Then, $\beta' = v$ and $\eta_{1/4}(v, v) < 0$. Thus, β is timelike for the auxiliary metric and $\gamma = \exp_p \beta$ is a future timelike (for g) geodesic from p to the point of normal coordinate (δ, x^i) . Finally, for $\epsilon > 0$ sufficiently small, we have again that the points of normal coordinates $(-\delta + \epsilon, x^i)$, $(\delta - \epsilon, x^i)$ lies in respectively $I^-(p, \mathcal{U})$ and $I^+(p, \mathcal{U})$. \square

Proposition 10.3. *An achronal set A is a topological hypersurface if and only if A contains no edge point (i.e. A and edge A are disjoint).*

Proof. First, let A be a topological hypersurface. Let $p \in A$ and let \mathcal{U} be a neighborhood of p as in Definition 10.9. Without loss of generality, we can assume that \mathcal{U} is connected and that $\mathcal{U} \setminus A$ has two connected components. (Consider $\phi(p)$ in \mathbb{R}^n . There exists a ball B centered at p such that $B \subset \phi(\mathcal{U})$. Then $\phi^{-1}(B)$ is a neighborhood of $p \in M$ that works). Since A is achronal, the open sets $I^-(p, \mathcal{U})$ and $I^+(p, \mathcal{U})$ are disjoint. For otherwise, there exists a closed (piecewise) smooth timelike curve through $p \in A$, which contradicts that A is achronal. For similar reasons, they do not meet A . Moreover, $I^-(p, \mathcal{U})$ and $I^+(p, \mathcal{U})$ are connected, since otherwise $\mathcal{U} \setminus A$ has at least three connected components. Any timelike curve through p meets both sets, hence they are contained in different components of $\mathcal{U} \setminus A$. Thus every timelike curve from $I^-(p, \mathcal{U})$ to $I^+(p, \mathcal{U})$ must go through some point of A . Thus, p is not in the edge of A .

Now suppose that A and $\text{edge } A$ are disjoint. Let $p \in A$. Since $p \notin \text{edge}(A)$, there exists a neighborhood \mathcal{U} of p , such that every timelike curve from $I^-(p, \mathcal{U})$ to $I^+(p, \mathcal{U})$ intersects A .

From the previous lemma, we consider a neighborhood \mathcal{N} of p such that

1. $\xi(\mathcal{N})$ has the form $(a - \epsilon, b + \epsilon) \times N \subset \mathbb{R}^1 \times \mathbb{R}^{n-1}$.
2. The slice $x^0 = a$ of \mathcal{N} is in $I^-(p, \mathcal{U})$, the slice $x^0 = b$ of \mathcal{N} is in $I^+(p, \mathcal{U})$.

If $y \in N \subset \mathbb{R}^{n-1}$, the x^0 coordinate curve

$$s \rightarrow \xi^{-1}(s, y)$$

is timelike by construction, and must then meet A . Since A is achronal, the meeting point is unique. Let $h(y)$ be its x^0 coordinate. It suffices now to show that the function

$$h : N \rightarrow (a, b)$$

is continuous, for then $\phi = (x^0 - h(x^1, \dots, x^{n-1}), x^1, \dots, x^{n-1})$ is a homeomorphism from \mathcal{N} onto its image (easy to check, just write explicitly ϕ^{-1}) that carries $A \cap \mathcal{N}$ to the slice $u^0 = 0$ of $\phi(N) \subset \mathbb{R}^n$.

Let y_n be a sequence that converges in N to y . Assume that $h(y_n)$ does not converge to $h(y)$. Then, there must exist (since the h values are bounded) a subsequence $y_{\psi(n)}$ converging to some $r \neq h(y)$. Now since $r \neq h(y)$, either $r > h(y)$, in which case one can reach $\xi^{-1}(y, r)$ from $\xi^{-1}(y, h(y))$ by following ∂_{x^0} , a timelike vector field forward, or $r < h(y)$ and then we follow ∂_{x^0} backwards. Thus, $\xi^{-1}(y, r)$ is in the set $I^-(q, \mathcal{N}) \cup I^+(q, \mathcal{N})$, where $q = \xi^{-1}(y, h(y))$. Since this set is open, it follows that there exists some n such that $\xi^{-1}(y_n, h(y_n)) \in I^-(q, \mathcal{N}) \cup I^+(q, \mathcal{N})$ for some n . But these belong to A so they cannot be in $I^-(q, \mathcal{N}) \cap I^+(q, \mathcal{N})$ since A is achronal. \square

Remark 10.2. One can in fact show that such an A is a Lipschitz submanifold, i.e. the function h constructed above satisfies a Lipschitz condition, cf Penrose [Pen72].

Corollary 10.2. An achronal set A is a closed topological hypersurface if and only if $\text{edge } A$ is empty.

Proof. Assume that A achronal is a topological hypersurface. Then A and $\text{edge } A$ are disjoint. But $\text{edge } A$ is included into the closure of A , so if A is closed then $\text{edge } A$ must be empty.

Reciprocally, suppose $\text{edge } A$ is empty. Then A is a topological hypersurface. Moreover, if A is achronal so is \bar{A} (because the relation \ll is open). Thus, if $q \in \bar{A} \setminus A$, then no timelike curve through q can ever meet A (otherwise, since $A \subset \bar{A}$, that would contradict that \bar{A} is achronal), which implies that q is in $\text{edge}(A)$ which we assumed was empty. \square

Definition 10.10. A subset F of M is a future set (respectively past set) provided $I^+(F) \subset F$ (respectively $I^-(F) \subset F$).

Example: For any set B , $J^+(B)$, $I^+(B)$ are future sets.

Exercise 10.5. If F is a future set, then $M \setminus F$ is a past set.

Corollary 10.3. *The topological boundary of a future set is a closed achronal topological hypersurface if non-empty.*

Proof. 1. ∂F is achronal: Let F be a future set and $p \in \partial F$. If $q \in I^+(p)$, then $I^-(q)$ is a neighborhood of p and hence contains a point of F . Thus, $q \in I^+(F) \subset F$, using that F is a future set. This proves that $I^+(p) \subset F$ and similarly, one has $I^-(p) \subset M \setminus F$. In particular, $I^+(\partial F)$ and $I^-(\partial F)$ are disjoint and hence ∂F is achronal.

2. The closed set ∂F has no edge points: Indeed, $I^+(p) \subset \text{int}F$ and $I^-(p) \subset \text{ext}F$ for $p \in \partial F$. (Let U be a neighborhood of p and γ be a timelike curve from $I^-(p, U)$ to $I^+(p, U)$, then since $I^-(p, U)$ and $I^+(p, U)$ are contained in $\text{int}F$ and $\text{ext}F$, γ must intersect the boundary of F). The conclusion then follows from the previous corollary. \square

Remark 10.3. *This last corollary has in fact plenty of applications in the study of possible singularities for non-linear wave equations. For instance, consider a wave equation of the form*

$$\square u = F(u, u')$$

where F is some non-linearity and \square is the usual wave operator. We imagine that we have initial data given at $t = 0$ and that the problem is well posed at locally, so that we can construct for any data a local solution. Now, using the domain of the dependence/finite speed of propagation, we can consider the largest domain on which the solution stays regular (see Alinhac's book [Ali09] for more on how to define this properly). That domain will be a future set, and hence its boundary (where something bad has to happen to our solution, otherwise we would continue) must be at least a closed achronal topological hypersurface.

10.5 Cauchy hypersurfaces

Definition 10.11. *A Cauchy hypersurface in M is a subset S that is met exactly once by every inextendible timelike curve in M .*

Exercise 10.6. 1. *Prove that any $t = \text{const}$ hypersurface in Minkowski space is a Cauchy hypersurface.*

2. *In 2d Minkowski space, for any $\rho > 0$, prove that the hyperboloid*

$$\{(t, x) : t^2 - x^2 = \rho, t > 0\}$$

is however not a Cauchy hypersurface.

We will need the following technical lemma.

Lemma 10.12. *Let α be a past-inextendible causal curve starting at p that does not meet a closed set C . Then, if $p_0 \in I^+(p, M \setminus C)$, there is a past inextendible timelike curve through p_0 that does not meet C .*

Remark 10.4. *Note that the original curve is only causal while the final curve is timelike. On the other hand, the point p_0 is in $I^+(p, M \setminus C)$ so it can be joined to p by a timelike curve.*

Proof. Since α is past inextendible, we may assume that α has domain $[1, +\infty)$ and that the sequence $\alpha(n)$ does not converge.

Since $p_0 \in I^+(p, M \setminus C)$, we can join p_0 to $\alpha(1) = p$ by a timelike curve that does not meet C . Then the combined curve, denoted γ , gives a causal curve starting at $\gamma(0) = p_0$ which is inextendible and has values in $M \setminus C$. Moreover, γ is not a null pregeodesic since it is initially timelike.

We claim that we can deform it, to be timelike everywhere and still avoiding C . We will denote β the deformed curve. Indeed, consider the points $p_1 = \alpha(1)$, $p_2 = \alpha(2)$, ..., $p_n = \alpha(n)$. First, we can deform γ on $[0, 2]$ to be a timelike curve keeping the ends p_0 and p_2 fixed and being arbitrary close to γ . Since $\gamma([0, 2])$ is compact and C is closed, if the deformation is small enough then the deformed curve β' does not meet C . (For instance, everything can be quantified using a metric d on M). We then define β on $[0, 1]$ to be the curve just obtained. To construct β on $[0, 2]$, we consider the curve β' constructed previously on $[1, 2]$ (and agreeing with β on $[0, 1]$ by construction) and join it to α from $[2, 3]$. This is a causal curve from $\beta(1)$ to p_3 which is initially timelike. Thus, it can be deformed so that the deformed curve β'' is timelike everywhere and avoids C . Extending β on $[0, 1]$ by β'' on $[1, 2]$ we obtain β on $[0, 2]$. Repeating the process, we construct β on $[0, \infty)$, which is timelike everywhere, avoids C and moreover $\beta(n)$ can be taking arbitrarily close to $\alpha(n)$ depending on n . For instance, we can arrange $d(\beta(n), \alpha(n)) < 1/n$, for an arbitrary metric d on M . It follows that β is inextendible. \square

We have

Proposition 10.4. *A Cauchy surface is a closed achronal topological hypersurface and is met by every inextendible causal curve.*

Remark 10.5. *Note that in Definition 10.11, a Cauchy surface was only assumed to be met by timelike curves.*

Proof. Let $p \in M$. Let γ be a maximally defined (thus inextendible) future oriented timelike geodesic γ defined in a neighborhood of 0 and such that $\gamma(0) = p$. Then, by definition of a Cauchy hypersurface, γ intersect S once and only once, i.e there exists a unique t_0 such that $\gamma(t_0) \in S$. Now, depending on whether $t_0 > 0$, $= 0$, or < 0 , we have that $p \in I^+(S)$, S , or $I^-(S)$. Thus, M is the disjoint union of $I^+(S)$, S and $I^-(S)$. Moreover, for any $q \in S$, any future oriented timelike curve γ such that $\gamma(0) = q$ is such that $\gamma(t) \in I^+(S)$ for $t > 0$ and $\gamma(t) \in I^-(S)$ for $t < 0$ and in particular, $S \subset \partial I^\pm(S)$ and in fact S is the common boundary of the sets $I^\pm(S)$ (note that $\partial I^+(S) \cap I^-(S) = \emptyset$ since $I^-(S)$ is open and disjoint from $I^+(S)$).

From Corollary 10.3, it follows that S is a closed achronal topological hypersurface. It remains to show that S is met, not just by every inextendible timelike curve, but by every inextendible causal curve α .

Assume that α is an inextendible causal curve that does not meet S . Wlog, assume that $\alpha(0) \in I^+(S)$. Then, by the previous lemma, there is a past inextendible timelike curve β starting in $I^+(S)$ that does not meet S . Since any future-pointing timelike curve starting at $\beta(0)$ must remain in $I^+(S)$, thus adjoining any such curve to β gives an inextendible timelike curve that avoids S , i.e. a contradiction. \square

Proposition 10.5. *Let S be a smooth spacelike Cauchy hypersurface. Then M is diffeomorphic to $\mathbb{R} \times S$. Furthermore, if S' is another smooth Cauchy hypersurface, then S' and S are diffeomorphic.*

Proof. Let T be a globally defined timelike vector field whose flow is defined on $\mathbb{R} \times M$, i.e. T is complete, cf. Proposition 8.5. Its restriction to $\mathbb{R} \times S$ gives a smooth map

$$f := \phi|_{\mathbb{R} \times S} : \mathbb{R} \times S \rightarrow M.$$

First, we claim that f is injective. Indeed, since if $\phi(t_1, x_1) = \phi(t_2, x_2)$, then by uniqueness and definition of the flow (or the semi-group property of $t \rightarrow \phi(t, \cdot)$), we must have $x_1 = \phi(t_2 - t_1, x_2)$. Since S is achronal and T timelike, we must then have $t_2 = t_1$ and $x_2 = x_1$.

Let now $(t_0, x_0) \in \mathbb{R} \times S$. We claim that $df_{(t_0, x_0)}$ is surjective. Indeed, let h be defined $h(p) = \phi(-t_0, p)$, then h is a diffeomorphism of M and $df_{(t_0, x_0)}$ is surjective is and only if $d(h \circ f)_{(t_0, x_0)}$ is surjective. Now by construction, for any $(t, p) \in \mathbb{R} \times S$,

$$h \circ f(t, p) = h(\phi(t, p)) = \phi(t - t_0, p).$$

Consider any curve $s \rightarrow (t, \gamma(s))$ where γ is a curve in S , as well as any curve of the form $s \rightarrow (s, p)$ for $p \in S$, it follows that the image of $d(h \circ f)_{(t_0, x_0)}$ contains both T_{x_0} and the tangent space of S . Since S is spacelike and T_{x_0} timelike, it follows that the image of $d(h \circ f)_{(t_0, x_0)}$ has dimension n and that $d(h \circ f)_{(t_0, x_0)}$ is surjective.

Since f is a local diffeomorphism and injective, it is in fact a global diffeomorphism onto its image. We claim that it is also surjective. Indeed, let $p \in M$ and let γ be the maximal integral curve of T starting at p . Then γ is inextendible and thus intersect S . It then follows that p is in the image of f .

Let now $\pi : \mathbb{R} \times S \rightarrow S$ be the projection on S . If S' is another smooth spacelike Cauchy hypersurface, then we get a map

$$\begin{aligned} r : S' &\rightarrow S \\ x &\rightarrow r(x) = \pi \circ f^{-1}(x). \end{aligned}$$

Then, r is smooth and r maps a point $x \in S'$ to the point of intersection between the integral curve of T starting at x and S . The inverse of r can be defined as the map that takes $y \in S$ to the point of intersection between the integral curve of T through y and S' . Thus, S and S' are diffeomorphic. \square

Recall that a Cauchy hypersurface needs not be differentiable, but is a topological hypersurface at minimum. In that latter case, we have

Proposition 10.6. *Let S be a Cauchy hypersurface in M and let X be a timelike vector field on M (which exists, since M is time-oriented). If $p \in M$, the maximal integral curve of X through p meets S at a unique point $\rho(p)$. Then $\rho : M \rightarrow S$ is a continuous open map onto S leaving S pointwise fixed. In particular S is connected (assuming M connected).*

Proof. See O'Neill [O'N83, p.417] (the proof requires invariance of domain, it would be nice to have another proof not based on it, though unlikely to be true since we are dealing with only topological manifolds here). \square

10.6 Global hyperbolicity

The notion of global hyperbolicity was introduced by Leray³⁶ [Ler55] in a more general context arising in the study of general hyperbolic pdes.

³⁶His definition was naturally slightly different, in particular outside from the realm of Lorentzian geometry. Instead of the compactness of $J^-(p) \cap J^+(q)$, he asked a C^0 compactness on the set causal curves from p to q .

Definition 10.12. A spacetime is said to be globally hyperbolic if it is strongly causal and if for any p, q , $J^-(p) \cap J^+(q)$ is compact.

We have the following results.

Proposition 10.7. If (M, g) is globally hyperbolic, then the causality relation \leq is closed, i.e. if $p_n \rightarrow p$ and $q_n \rightarrow q$, with $p_n \leq q_n$, then $p \leq q$.

Remark 10.6. Recall that in Lemma , we had proved that \leq was closed when restricted to a convex set.

Proof. The proof is trivial if $p = q$. Hence, we can assume that $p \neq q$. Let α_n be a causal curve from p_n to q_n . Take $p^- \ll p$ and $q \ll q^+$. Then, for n large enough, $p^- \ll p_n \leq q_n \ll q^+$. Thus, the curve α_n are all in the compact set $J^+(p^-) \cap J^-(q^+)$. From Lemma 10.10, there exists a causal curve from p to q . \square

Lemma 10.13. If (M, g) is globally hyperbolic and $p < q$, then there is a causal geodesic from p to q such that no causal curve connecting p to q has greater length.

Proof. Since $p < q$, the set $C(p, q)$ of causal curves from p to q is non-empty. Let $\tau(p, q) = \sup_{\gamma \in C(p, q)} L(\gamma)$. Note that a priori, $\tau(p, q) \in [0, +\infty]$. Let α_n be causal curves from p to q such that $L(\alpha_n) \rightarrow \tau(p, q)$ as $n \rightarrow +\infty$. These curves are all in $J^+(p) \cap J^-(q)$ which is compact and strongly causal by global hyperbolicity. Hence, by Lemma 10.10, there is a causal broken geodesic λ from p to q with $L(\lambda) = \tau(p, q)$. Moreover, if there is a break say at some r with $p \leq r \leq q$, then take a normal neighborhood U at r and consider points along the curve $p', q' \in U$, with $p' \leq r \leq q'$. Then, from Lemma 8.14, the geodesic segment from p' to q' is timelike and has bigger length than the corresponding segment of the original curve, a contradiction. \square

Lemma 10.14. Let (M, g) be a globally hyperbolic Lorentzian manifold and let $K_1, K_2 \subset M$ be compact. Then $J^-(K_1) \cup J^+(K_2)$ is compact.

Proof. Recall that since M is metrizable (since M is a manifold), a subset is compact iff it is sequentially compact. Let $q_n \in J^-(K_1) \cup J^+(K_2)$. By assumptions, there exists $p_{n,i} \in K_i$, such that $p_{n,2} \leq q_n \leq p_{n,1}$. Since the K_i are compact, up to subsequences, we can assume that $p_{n,i} \rightarrow p_i$. Let $r_i \in M$ be such that $r_2 \ll p_2$ and $p_1 \ll r_1$. Then, $r_1 \ll q_n \ll r_2$ for n large enough. Hence, $q_n \in J^-(r_1) \cup J^+(r_2)$ for n large enough and since this set is compact by global hyperbolicity, up to a subsequence q_n converges to some q . Since the relation \leq is closed on globally hyperbolic manifold, we conclude that $p_2 \leq q \leq p_1$. \square

10.7 Cauchy developments

Definition 10.13. If A is an achronal subset of M , the future Cauchy development of A is the set $D^+(A)$ of all points $p \in M$ such that every past-inextendible causal curve through p meets A . In particular $A \subset D^+(A)$. Similarly, $D^-(A)$ is the set of all points $p \in M$ such that every future-inextendible causal curve through p meets A . The Cauchy development of A is by definition $D^+(A) \cup D^-(A)$.

In particular, if S is a Cauchy hypersurface, then $D(S) = M$.

We have easily

Lemma 10.15. 1. $D^+(A) \subset A \cup I^+(A) \subset J^+(A)$,

2. $D^+(A)$ and $I^-(A)$ are disjoint,
3. $D^+(A) \cap D^-(A) = A$ and $D^+(A) \setminus A = D(A) \cap I^+(A)$.
4. If α is a past-directed causal curve starting in $D^+(A)$ and leaving $D^+(A)$, then it first leaves $D^+(A)$ through A .

Proof. We only prove the last claim, the first four follows from the definitions (and the fact that A is achronal!). Let $\alpha = [0, b) \rightarrow M$ be a past-directed causal curve such that $\alpha(0) \in D^+(A)$. Let $s > 0$ such that $\alpha(s) \notin D^+(A)$. By definition of D^+ , there must be a past-inextendible causal curve β starting at $\alpha(s)$ that does not meet A . But $\alpha|_{[0,s]} + \beta$ must meet A , thus $\alpha|_{[0,s]}$ must intersect A say at some t_0 . If there exists another $0 < s' < t_0$ such that $\alpha(s') \notin D^+(A)$, then we can produce another $t_1 < t_0$ such that $\alpha(t_1) \in A$, which contradicts the achronality of A . \square

We will also need.

Lemma 10.16. *If A is achronal and $p \in \text{int} D(A)$, then every inextendible causal curve through p meets both $I^-(A)$ and $I^+(A)$.*

Proof. Since $D(A) \subset A \cup I^+(A) \cup I^-(A)$, we can assume wlog that $p \in A \cup I^+(A)$. Let $\alpha : [0, +\infty) \rightarrow M$ be an inextendible causal curve starting at p . Consider the points $p_i = \alpha(i)$ along α . Let d be an auxiliary metric on M . Let $r_0 \in A$ be a point such that $r_0 \gg p_0 = p$. By induction, we can construct, a sequence of points (r_i) in A such that, for $i \geq 1$, $r_{i-1} \gg r_i \gg p_i$ and $d(r_i, p_i) < 1/i$. The curve β obtained by joining the r_i by timelike segment is then a past directed inextendible (the r_i cannot converge) timelike curve such that the past of any point of the curve contains a point of A . Since β is past inextendible, it must intersect A and thus $\alpha \in I^-(A)$. \square

(To get an idea why p must lie in the interior, just take A a closed horizontal interval in 2d Minkowski space and then take p on the boundary of $D(A)$ in the future of A .)

Lemma 10.17. *Let A be achronal and $p \in \text{int} D(A) \setminus I^-(A)$. Then, $J^-(p) \cap D^+(A)$ is compact.*

Proof. Let x^n be a sequence in $J^-(p) \cap D^+(A)$. If x^n admits a subsequence converging to p we are done, so we assume that it is not the case. In particular, we may assume that $x_n \neq p$ for all n sufficiently large. Let α_n be past pointing causal curves from p to x_n . Using Proposition 10.1, since no subsequence of x_n converges to p , there is a past directed limit sequence $\{p_i\}$ for α_n starting at p . If $\{p_i\}$ is infinite, then if λ is a quasi-limit, λ is past inextendible causal curve starting at p . Thus, by Lemma 10.16, it must intersect $I^-(A)$, which is a contradiction. Thus, the p_i are finite and then some subsequence of the x_n (which are the endpoints of the α_n) must converges to the last p_i , denoted here x , with $x \in J^-(p)$. It remains to prove that $x \in D^+(A)$. Note first that $x \notin I^-(A)$, as this would imply that some $x_n \in I^-(A)$, which is impossible since $x_n \in D^+(A)$ and A is achronal.

Since $p \in \text{int} D(A) \setminus I^-(A)$, there exists $q \in D^+(A) \cap I^+(p)$. Let then σ be a past inextendible timelike curve σ from $q = \sigma(0)$ passing through $x = \sigma(t_0)$, at some $t_0 > 0$. (for instance σ can be constructed as in the previous lemma). Since $q \in D^+(A)$ and A is achronal, σ must intersect A once and only once, say at some t_1 . Note that it then follows that $\sigma|_{[0,t_1]} \subset D^+(A)$. Since $x \notin I^-(A)$, it follows that we must have $t_0 \leq t_1$ and then $x \in D^+(A)$. \square

Theorem 10.1. *If A is an achronal set, then $\text{int}D(A)$, if non-empty, is globally hyperbolic.*

Proof. • There exists no closed causal loops: If there is a closed causal loops, by passing through the loops infinitely many times, we can construct an inextendible causal loops. This must then intersect A infinitely many times and using 10.16, contradicts that A is achronal.

- Assume that the strong causality condition does not hold at some $p \in \text{int}D(A)$. Then, there exists a neighborhood $U \in \text{int}D(A)$ of p and future directed causal curve segments $\alpha_n : [0, 1] \rightarrow U$ such that $\alpha_n(0)$ and $\alpha_n(1)$ converges to p but none of the α_n are contained in U . Thus, α_n has a future directed limit sequence at p_i starting at p . If the p_i are finite, then it must end at p (if it ends at some $p' \neq p$, then we can extend the limit sequence, contradiction). It then follows that $p \in J^+(p)$, i.e. there is a closed future directed causal curve from p to p , which we have already excluded. Thus, the p_i are infinite and the corresponding quasi-limit λ is future inextendible. From Lemma 10.16, there exists some t_0 such that $\lambda(t_0) \in I^+(A)$ and from the achronality of A , $\lambda(t) \in I^+(A)$, for all $t \geq t_0$. In particular, some point p_i of the limit sequence must belong to $I^+(A)$. Thus, there is a subsequence α_m and some $s_m \in [0, 1]$ such that $\alpha_m(s_m) \in I^+(A) \rightarrow p_i$. Reparametrizing if necessary, we may assume that $s_m = s$ is independent of m . Since $p_i \neq p$, we can obtain another past-directed limit sequence q_i starting at p by considering the curves α_m restricted to $[s, 1]$. If the limit sequence is finite, it must end at p_i , and then $p_i < p$. Since, $p < p_i$, we again obtain a causal loop. Thus, the limit sequence is infinite and the corresponding quasi-limit σ is a past-directed inextendible causal curve starting at p . By Lemma 10.16 again, there exists some t_0 such that $\sigma(t_0) \in I^-(A)$ and in particular, some α_m restricted to $[s, 1]$ must meet $I^-(A)$. Since α_m is future pointing and has $\alpha_m(s) \in I^+(A)$, we again contradict the achronality of A .
- Let $p \neq q$, with $p, q \in \text{int}D(A)$. If $p = q$, then, since there are no causal loops, $J^-(q) \cap J^+(p) = \{p\}$ is compact. Suppose then that $p < q$. Let $\{x^n\}$ be a sequence in $J^-(q) \cap J^+(p)$. For any n , let α_n be a future directed causal curve from q to p passing by x_n . Let R be a covering of M by convex open sets C such that \bar{C} is compact and contained in a convex open set. Let p_k be a limit sequence relative to R starting at p .
 1. If the p_i are finite, they must end at some $p_k = q$. Let $(\alpha_m) = (\alpha_m)$ be a subsequence as in Definition 10.3, L1a. Recall that, for any m , x_m belongs to α_m . Thus, it must belong to one of segments $[s_{m,j}, s_{m,j+1}]$, where j potentially depends on m . On the other hand, since, for any m , they are only finitely many such segments (indexed by j not m !), there exists $i < k$ such that for infinitely many m , $x_m \in [s_{m,i}, s_{m,i+1}]$. Thus, there exists a subsequence, still denoted α_m , such that $x_m \in [s_{m,i}, s_{m,i+1}]$, for all m and hence the x_m all belongs to a single convex set C of compact closure. In particular, they admits a subsequence converging to some x , and since \leq is closed on convex sets, we have $p_i \leq x \leq p_{i+1}$ and hence $x \in J^-(q) \cap J^+(p)$.
 2. If the p_i are infinite, they the corresponding quasi-limit is a future inextendible causal curve starting at p . The proof is similar to the proof of the strong causality condition as above. First, there exists a subsequence α_m

and after a change of parametrization a fixed s such that $\alpha_m(s)$ converge to some point $p_i \in I^+(A)$. Since $p_i \neq q$, we can consider α_m restricted to $[s, 1]$ and construct a past directed limit sequence q_i starting at q . If the limit sequence is finite, then it must ends at $\lim \alpha(s) = p_i$ and it then follows that $p < p_1 < \dots < p_i < \dots < q_1 < q$ is a finite limit sequence for α_n , which we have already ruled out. Thus, q_i is infinite and the corresponding quasi-limit is a past inextendible causal curve μ starting at q . μ must intersect $I^-(A)$ hence some α_m restricted to $[s, 1]$ does and this contradicts achronality.

- This proves that $J^-(q) \cap J^+(p)$ is compact, but a priori this sets may differ from $J^-(q, \text{int}D(A)) \cap J^+(p, \text{int}D(A))$. We now prove that $J^-(q) \cap J^+(p) \subset \text{int}D(A)$, which implies that the two above sets are the same. As before, we may assume that $p < q$. It will be sufficient to consider only two cases $p, q \in I^+(A)$ or $p \in J^-(A), q \in J^+(A)$ (the other cases are similar exchanging $-$ and $+$. Note also that $I^\pm(A) \subset J^\pm(A)$).

 1. $p, q \in I^+(A)$. Let $q^+ \in I^+(q) \cap D(A) \subset D^+(A)$. Then, $U = I^+(A) \cap I^-(q^+)$ is an open set containing $J^-(q) \cap J^+(p)$. Let σ be a past directed timelike curve from q^+ to $y \in U$. Then, since A is achronal and $y \in I^+(A)$, σ does not meet A , thus we must have $y \in D^+(A)$.
 2. $p \in J^-(A), q \in J^+(A)$. Similarly, we can find points $p^- \in I^-(A) \cap D^-(A)$ and $q^+ \in I^+(q) \cap D^+(A)$ and consider the open set $U = I^+(p^-) \cap I^-(q^+)$. Let $x \in U$, σ be a past directed timelike curve from q^+ to x and τ be a past directed timelike curve from x to p^- . Since $A \subset D(A)$, we can also suppose that $x \neq A$. By achronality, at least one of τ and σ do not meet A , say σ . This then implies that $x \in D^+(A)$.

□

The last theorem implies in particular

Corollary 10.4. *A spacetime that admits a Cauchy hypersurface is globally hyperbolic.*

Proof. This follows from $M = D(S)$, if S is a Cauchy hypersurface and the previous theorem. □

10.8 Time and temporal functions

In the whole section, (M, g) is a globally hyperbolic Lorentzian manifold.

The aim of the section is to prove the existence of certain special, globally defined, functions on globally hyperbolic Lorentzian manifold. These functions will have the property that each of their level sets will be Cauchy hypersurfaces. Moreover, we will require more and more structure on them to eventually construct a smooth function, whose level sets are smooth spacelike Cauchy hypersurfaces and such that its gradient is timelike and past directed. This will be our replacement for say, the function t in Minkowski space.

First, the functions that we will construct will lack regularity. We will then modify it locally near any level set to eventually reach our goal.

10.8.1 Time function and Geroch's theorem

Let M be oriented, so that we can define globally a volume form

$$\eta = \sqrt{|\det g|} dx^1 \wedge \dots \wedge dx^n$$

Recall the definition of a smooth partition of unity.

Definition 10.14. *Given a open covering $U = \{U_i\}$ of M , a smooth partition of unity for the U_i is a collection of smooth maps ϕ_i such that*

- For any i , $\phi_i : M \rightarrow [0, 1]$.
- There is a locally finite open refinement $\{V_j\}$ of U such that the support of ϕ_i is included in V_j for any i ,
- $\sum_i \phi_i = 1$.

Let $(U_i, \xi_i)_{1 \leq i \leq +\infty}$ be an atlas of M compatible with its orientation (so that the Jacobian matrices of the change of coordinates have positive determinant on overlap) and let (ϕ_i) be a smooth partition of unity for the U_i . Finally, let m_i be the integral of the ϕ_i .

Then, defines the volume form

$$\omega = \sum_{i=1}^{+\infty} \frac{1}{m_i 2^i} \phi_i \eta.$$

Note that by construction,

$$\int_M \omega = 1 < +\infty, \text{ so that we have defined a finite measure on } M.$$

Let Λ be the linear functional that takes smooth function of compact support on M to $\Lambda(f) = \int_M f \omega$.

By Riesz representation theorem, there exists a positive measure μ and a σ algebra of subsets of M containing the Borel sets of M , such that μ is complete and such that for smooth functions of compact support,

$$\int f d\mu = \Lambda(f).$$

By construction, $\mu(M) = 1$ (this follows from the monotone convergence theorem), $\mu(U) > 0$ for any open non-empty set and for any measurable set V , $\mu(V) = 0$ if and only if for all i , $\xi_i(V \cap U_i)$ is of zero measure for the Lebesgue measure.

We will use this measure to prove the following theorem.

Theorem 10.2 (Geroch's theorem). *There exists a continuous, onto, function $\tau : M \rightarrow \mathbb{R}$ which is strictly increasing along any future oriented causal curve and such that if γ is an inextendibly causal curve defined on (t_-, t_+) , then $\tau(\gamma(t)) \rightarrow \pm\infty$, as $t \rightarrow t_{\pm}$. In particular, for any $a \in \mathbb{R}$ is an acausal Cauchy hypersurface.*

Proof. The fact that if τ is such a function, then its level sets are Cauchy hypersurface follows from the definition of a Cauchy hypersurface. Since the monotonicity of τ holds for any causal curve (not just timelike curve), then the level sets are acausal: they can not be met more than once by causal curves.

Let μ be the measure constructed above. Let $p \in M$ and let $C_p^+ = J^+(p) \setminus I^+(p)$. We claim that

$$\mu(C_p^+) = 0.$$

Indeed, let $q \in C_p^+$. Then, there is a causal curve from p to q . Moreover, since M is globally hyperbolic, from Lemma 10.13, there exists a causal geodesic connecting p to q of greater length and since $q \notin I^+(p)$, it must be a null geodesic. Thus, there exists a $v \in T_p M$ which is null and such that $\exp_p(v) = q$. Let N be the future directed null vectors in $T_p M$ that lies in the domain of \exp_p . This is an open submanifold of $T_p M$ of dimension $n-1$. Moreover, \exp_p is a smooth map from N to M and $C_p^+ \setminus \{p\} \subset \exp_p(N)$. An application of Sard's theorem then implies that $\mu(\exp_p(N)) = 0$ and hence C_p^+ has also zero measure.

Similarly, with $C_p^- = J^-(p) \setminus I^-(p)$, we have $\mu(C_p^-) = 0$.

Let us define now $f_-(p) = \mu[J^-(p)] (= \mu[I^-(p)])$ and $f_+(p) = \mu[J^+(p)] (= \mu[I^+(p)])$.

Since $I^\pm(p)$ are open and non-empty, $f_\pm(p) > 0$, for all $p \in M$.

Step 1: f_- and f_+ are continuous.

Let $p_j \rightarrow p$.

Assume first that $p_j \ll p$. Since $J^-(p_j) \subset J^-(p)$, we know that $f_-(p_j) \leq f_-(p)$. In view of the preliminary step, for all $q \in M$,

$$f_-(q) = \int_M \chi_{J^-(q)} = \int_M \chi_{I^-(q)} d\mu$$

Since the relation \ll is open, it follows that $\chi_{I^-(p_j)}$ converges pointwise to $\chi_{I^-(p)}$ and the result then follows by Lebesgue's dominated convergent theorem.

If now $p \ll p_j$, we have that $\chi_{J^-(p_j)}$ converges pointwise to $\chi_{J^-(p)}$ as a consequence of the fact that the relation \ll is closed on globally hyperbolic manifold and again $f_-(p_j) \rightarrow f_-(p)$ as $j \rightarrow +\infty$.

For a general sequence of p_j , let $\epsilon > 0$ and take points q_1 and q_2 such that $q_1 \ll p \ll q_2$ and

$$f_-(p) \leq f_-(q_2) \leq f_-(p) + \epsilon,$$

$$f_-(p) - \epsilon \leq f_-(q_1) \leq f_-(p).$$

for j large enough, $q_1 \ll p_j \ll q_2$, so that

$$f_-(p) - \epsilon \leq f_-(q_1) \leq f_-(p_j) \leq f_-(q_2) \leq f_-(p) + \epsilon.$$

This proves the continuity of f_- .

Step 2: f_\pm are strictly monotone along causal curves.

Let $p < q$. Then $q \notin J^-(p)$. (This follows from the strong causality condition). Since \leq is closed on globally hyperbolic spacetimes, $J^-(p)$ is closed and thus there exists an open neighborhood U of q which does not intersect $J^-(p)$. It follows that $J^-(q) \setminus J^-(p)$ contains a non-empty open set and thus $f_-(q) > f_-(p)$.

Step 3: For $\gamma : [0, a) \rightarrow M$ causal future inextendible, $f_+ \circ \gamma(t) \rightarrow 0$ as $t \rightarrow a$.

Let $K \subset M$ be compact. We claim that for t large enough $K_t = J^+[\gamma(t)] \cap J^-(K)$ is empty. Note that K_t is compact for every t , cf Lemma 10.14. Let $C = K_0 \cup \{\gamma(0)\}$. Then C is a compact set and γ start in C . From Lemma 10.9, there is a t_0 such that for $t > t_0$, $\gamma(t) \notin C$. Since $\gamma(t) \in J^+[\gamma(0)]$ for all $t \geq 0$, we conclude that $\gamma(t) \notin J^-(K)$ for $t > t_0$. Thus K_t is empty for $t > t_0$. For any i , let now C_i be the compact set defined by

$$C_i = \bigcup_{j=1}^i \text{supp } \phi_j.$$

Since $\mu(M - C_i) \leq 2^{-i}$, we conclude that for any i , we have $f_+(\gamma(t)) \leq 2^{-i}$, for t large enough.

Step 4: The function $\tau : M \rightarrow \mathbb{R}$ defined as

$$\tau(p) = \ln f_-(p) - \ln f_+(p)$$

then satisfies all the properties of the theorem. □

Definition 10.15. *A continuous function which is strictly increasing along future directed causal curves is called a time function.*

The time function we have just constructed still lacks regularity to be used in applications. It is only continuous. Moreover, eventhough it is strictly increasing along future directed causal curves, its gradient may not be timelike. We would like to construct a smooth function τ such that $D\tau$ is timelike and past-directed, what we will call a *temporal* function.

Definition 10.16. *A temporal function τ is smooth function such that $D\tau$ is timelike and past-directed.*

Exercise 10.7. • *Prove that temporal function are time functions.*

- *Give an example of a smooth time function which is not temporal.*

Our approach will be smooth out Geroch's time function. For this, we will use certain technical constructions presented in the next section.

10.8.2 Local constructions

First, we need the following lemma, whose classical proof is ommitted.

Lemma 10.18. *Let $V \subset U \subset \mathbb{R}^n$ be open sets. Let $f \in C^\infty(U)$ be such that $f(x) > 0$ for $x \in V$ and $f(x)=0$ for $x \in \partial V \cap U$. Let $g(x) = \exp[-1/f(x)]$, for $x \in V$ and 0 elsewhere. Then g is smooth.*

Let now $p \in M$ and consider a normal neighborhood U of p and $\tilde{U} = \exp_p^{-1}(U)$. Recall the functions $\tilde{q} : v \rightarrow g(v, v)$ and $q = \tilde{q} \circ \exp_p^{-1}$. Let \mathcal{F}_p^+ be the future timecone at p and let $V = \exp_p \mathcal{F}_p^+ \cap U$. Note that q vanishes on ∂V and $q < 0$ on V and thus we can define a function

$$f_p(x) = \exp[1/q(x)]$$

for $x \in V$ and $f_p(x) = 0$ elsewhere. f_p is smooth by the above lemma. Moreover, on V ,

$$Df_p = \text{grad } f_p = -\frac{f_p}{q^2} \text{grad } q = -2\frac{f_p}{q^2} P,$$

where P is the position vector field. Thus, Df_p is a past directed timelike vector field on V and is zero everywhere else.

Lemma 10.19. *Let S be an acausal Cauchy hypersurface. Let $p \in S$ and let U be a convex neighborhood of p . Then, there is a smooth non-negative function h_p such that*

1. $h_p(p) = 1$,
2. h_p has compact support contained in U ,
3. if $r \in J^-(S)$ and $h_p(r) \neq 0$, then $\text{grad } h_p(r)$ is past directed timelike.

Proof. Let $\gamma : (-\epsilon, \epsilon) \rightarrow M$ be a future directed timelike curve with $\gamma(0) = p$. We claim that for $t < 0$ close enough to 0, $K_t = J^+(\gamma(t)) \cup J^-(S)$ is contained in U . First, the K_t are compact by Lemma 10.17. Moreover, for $s < t$, $K_t \subset K_s$, i.e. the K_t are decreasing in t . Assume that for all $t < 0$, K_t is not included in U . Then, there exists an increasing sequence $0 > s_j \rightarrow 0$ such that for all j , $K_{s_j} \cap U^c \neq \emptyset$. Since $p_j \in K_{s_1}$ for all j , we can pass to a subsequence converging to r , and since U is open, U^c is closed, so that $r \in K_{s_j} \cap U^c$. Since $\gamma(s_j) \leq p_j$ and \leq is closed on globally hyperbolic manifold, we have $p \leq r$ and even $p < r$ (since $r \in U^c$ and $p \in U$). Since $r \in J^-(S)$ and $p \in S$, this contradicts the acausality of S . We conclude that for $t < 0$ small enough, $K_t \subset U$. Let $t_0 < 0$ such that $K_{t_0} \subset U$ and let $r_0 = \gamma(t_0)$. Since U is convex, we can construct the function f_{r_0} as introduced above, which has a past directed gradient whenever it is non-zero. Let $K \subset U$ be compact and containing K_{t_0} in its interior and let $\phi \in C_c^\infty(U)$ be a cutoff function associated to K , i.e. $\phi = 1$ on K and $\phi(U) \subset [0, 1]$. Let $H = \phi f$. It is a smooth function with compact support in U and can therefore be extended by 0 outside of U to a smooth function on M . Note that by definition of the function f_p (see the definition of the set V above), if $r \notin J^+(\gamma(t_0))$, then $f(r) = 0$. Thus, if $r \in J^-(S)$ and $H(r) \neq 0$, then $H = f$ on a neighborhood of r , so that $\text{grad } H$ is past pointing and timelike. Finally, note that $H(p) > 0$, since $p \gg r_0$ and that $\phi(p) = 1$. Thus, $h_p = H/H(p)$ has all the required property. \square

Proposition 10.8. *Let S be an causal Cauchy hypersurface. Given an open neighborhood W of S , there exists a smooth function $h_W : M \rightarrow [0, +\infty)$ such that*

1. The support of h_W is included in W ,
2. $h_W|_S > \frac{1}{2}$,
3. For $r \in J^-(S)$ and $h_W(r) \neq 0$, then $\text{grad } h_W(r)$ is past pointing timelike.

Remark 10.7. *Compared with the previous lemma, we are constructing a function h_W whose support and properties lies in a neighborhood of S instead of a neighborhood of some $p \in S$. Thus, we are gradually moving from local to more global properties.*

Proof. Let d be the distance associated to a complete Riemannian metric on M . Then, $\overline{B}_\rho(p)$, the closed ball of radius ρ centered at p , is compact due to the Hopf-Rinow theorem. Let $p \in M$, and define for $l \in \mathbb{N}$,

$$K_l = \overline{B}_l(p) - \overline{B}_{l-1}(p), \quad R_l = K_l \cap S.$$

For each $r \in S$, fix a convex set U_r of diameter strictly less than 1 for d and contained in W . Let h_r be the function constructed in the previous lemma for $(U, p) = (U_r, r)$. Let $V_r = h_r^{-1}[(1/2, +\infty)]$. Since the h_r are continuous with $h_r(r) = 1$, the V_r forms an open cover of S . Since R_l is compact and contained in S , for any l , there exists a finite number of points, denoted $r_{l,1}, \dots, r_{l,k_l} \in R_l$ such that the corresponding $V_{r_{l,1}}, \dots, V_{r_{l,k_l}}$ forms an open cover of R_l . Note that if $|l - m| \geq 3$, then $U_{r_{l,i}} \cap U_{r_{m,j}} = \emptyset$, in view of the diameter constraint and the fact that $d(r_{l,i}, r_{m,j}) \geq 2$. Thus, the sum

$$h = \sum_{l=1}^{+\infty} \sum_{i=1}^{k_l} h_{r_{l,i}}$$

defines a smooth function since each point has a neighborhood in which all but a finite number of the terms vanish. If $x \in S$, then $h(x) > 1/2$ since x must lie in some $V_{r_{l,i}}$. Further, the support of h is the union of the supports of $h_{r_{l,i}}$ (i.e. we do not need to take a further closure), since the covering $U_{r_{l,i}}$ is locally finite. Thus, the support of h is included into W . If $x \in J^-(S)$ and $h(x) \neq 0$, then $\text{grad } h_{r_{l,i}}(x)$ is timelike and past directed for every l, i such that $h_{r_{l,i}}(x) \neq 0$ and vanishes for every l, i such that $h_{r_{l,i}}(x) = 0$ (since $h_{r_{l,i}} \geq 0$, so that in that case x must be a minimum), so $\text{grad } h(x)$ is timelike and past directed. \square

10.8.3 Smooth temporal function

Let τ be the continuous time function guaranteed to exist by Geroch's theorem. We will denote its level set by $S_t := \tau^{-1}(t)$.

The idea will be to use the previous constructions to construct a replacement for τ near some S_t , and then glue the replacements to obtain a global temporal function.

Definition 10.17. Let $t_- < t_a < t < t_b < t_+$ and $S_\pm = S_{t_\pm}$. A function $\sigma : M \rightarrow \mathbb{R}$ is called a temporal step function around t (for (t_-, t_a, t, t_b, t_+)) if it satisfies

1. $\text{grad } \sigma$ is timelike and past pointing in $V = \{p \in M : \text{grad } \sigma(p) \neq 0\}$,
2. $\sigma(M) \subset [-1, 1]$,
3. $\sigma = \pm 1$ on $J^\pm(S_\pm)$,
4. $S_{t'} \subset V$ for all $t' \in (t_a, t_b)$.

Lemma 10.20. Let $t_- < t < t_+$. Then, there exists an open set U , such that

$$J^-(S_t) \subset U \subset I^-(S_{t_+})$$

and a function $h^+ : M \rightarrow [0, +\infty)$ whose support is contained in $I^+(S_{t_-})$ and such that

- If $p \in U$ and $h^+(p) > 0$, then $\text{grad } h^+(p)$ is timelike and past pointing,
- $h^+(p) > 1/2$ for $p \in J^+(S_t) \cap U$.

Proof. Take h^+ be the function h constructed in the previous Proposition with $S = S_t$ and $W = I^-(S_{t_+}) \cap I^+(S_{t_-})$. For $x \in S_t$, $h(x) > 1/2$ and $\text{grad } h(x)$ is past pointed and timelike. By continuity, there exists an open neighborhood of x , V_x , such that the same conditions on h and $\text{grad } h$ holds in V_x . Then define U as the union of $I^-(S_t)$ and the sets V_x for all $x \in S_t$. (that $J^-(S_t) \subset U$ is a consequence of the fact that S_t is a Cauchy hypersurface). \square

Lemma 10.21. *Let $t < t_+$ and $U \subset I^-(S_{t_+})$ be an open neighborhood of $J^-(S_t)$. Then, there exists a smooth function $h^- : M \rightarrow [-1, 0]$ such that*

- *the support of h^- is included in U ,*
- *if $\text{grad } h^-(p) \neq 0$, then it is timelike, past directed.*
- *$h^-(p) = -1$ for $p \in J^-(S_t)$.*

Proof. Reverse time orientation and construct a function h as in Proposition 10.8. Then we get a smooth function $h : M \rightarrow [0, +\infty)$ whose support is included in U and such that if $h(p) > 0$ for $p \in J^+(S_t)$, then $\text{grad } h(p)$ is timelike, future pointing. Moreover, $h(p) > 1/2$ for $p \in S_t$. Let $h_1 = -h$ and let $\phi : \mathbb{R} \rightarrow [-1, 0]$ be a smooth cutoff function such that $\phi = -1$ on $(-\infty, -1/2]$, $\phi' > 0$ on $(-1/2, 0)$, $\phi = 0$, on $[0, +\infty)$. Defines $h^- = \phi \circ h_1$ on $J^+(S_t)$ and $h^- = -1$ on $J^-(S_t)$, then h^- has all the desired property. \square

Proposition 10.9. *Let $t_- < t < t_+$. Then, there exists a smooth function $\sigma : M \rightarrow \mathbb{R}$ satisfying the first three properties of Definition 10.17 and such that $S_t \subset \{p \in M, \text{grad}(\sigma) \neq 0\}$.*

Proof. Let h^+ and U be as in Lemma 10.20 and given this U , let h^- be as in Lemma 10.21. Then, $h^+ - h^- > 1/2$ on U (since for $p \in U$, either $p \in J^+(S_t)$ or $p \in J^-(S_t)$). Thus, we can define σ as

$$\sigma = 2 \frac{h^+}{h^+ - h^-} - 1$$

on U and $\sigma = 1$ on the complement of U . We have

$$\text{grad } \sigma = 2 \frac{h^+ \text{grad } h^- - h^- \text{grad } h^+}{(h^+ - h^-)^2}$$

which at any given point is past pointing timelike or zero. \square

Corollary 10.5. *Let $t \in \mathbb{R}$ and $t_{\pm} = t \pm 1$. Let $t_- < t_a < t < t_b < t_+$ and let $K \subset \tau^{-1}([t_a, t_b])$ be compact. Then, there is a smooth function σ satisfying the first three property of Definition 10.17 and such that $K \subset \{p \in M, \text{grad}(\sigma) \neq 0\}$.*

Proof. For each $s \in [t_a, t_b]$, let σ_s be the function constructed in Proposition 10.9 with t_-, t, t_+ replaced by $t-1, s, t+1$ and let V_s be the set on which $\text{grad } \sigma_s$ is non-zero (i.e. the interior of the support of $\text{grad } \sigma$). Then, the set V_s constitute an open covering of K . Let s_1, \dots, s_k be such that the V_{s_i} is a finite open covering of K and define

$$\sigma = \frac{1}{k} \sum_{i=1}^k \sigma_{s_i}.$$

Then, σ has the desired properties. \square

In the proof of the next theorem we will need the following lemma, whose proof is left as an exercise. Note that the reason why this is not a triviality is that a limit of future-directed timelike vectors can be 0 or a null vector.

Lemma 10.22. *Let v_i be a sequence of future directed timelike vectors such that $\sum v_i$ is convergent. Then, $\sum_{i=1}^{\infty} v_i$ is timelike.*

Theorem 10.3 (Existence of temporal step function). *Let $t \in \mathbb{R}$ and let $t_{\pm} = t \pm 1$. Given $t_- < t_a < t < t_b < t_+$, there is a temporal step function around t as in definition 10.17.*

Proof. Let G_j be an increasing sequence of open sets of compact closure and such that $M = \cup_j G_j$. Let

$$K_j = \overline{G_j} \cap J^+(S_{t_a}) \cap J^- S_{t_b}.$$

K_j is compact and can be used to get a function σ_j as in Corollary 10.5. We will take a sum of the σ_j , but in order for the sum to converge we will need to divide each σ_j by certain weights. To define those, let $(V_i, {}^i\xi)_{1 \leq i}$ be an open covering by coordinate charts. (Note that i in ${}^i\xi$ labels the coordinate chart, not the individual coordinate functions). We may assume that the V_i are locally finite and that there exists another open cover by open sets U_i , such that $U_i \subset V_i$ for all i and the U_i have compact closure. (Exercise: construct such open covers).

For any j , let $A_j > 1$ be such that for all $1 \leq i \leq j$, $0 \leq m \leq j$ and for all multi-indices $|\beta| \leq m$,

$$\sup_{\overline{U_i}} \left| \frac{\partial^\beta \sigma_j}{\partial^m {}^i\xi} \right| \leq A_j,$$

i.e. the partial derivatives of order m of σ_j are uniformly bounded by A_j on $\overline{U_i}$ for the j first coordinate systems.

Let now σ be defined by

$$\sigma = \sum_{j=1}^{+\infty} \frac{1}{2^j A_j} \sigma_j.$$

Let $p \in M$. Then $p \in U_i$. To prove that σ is C^l , split the series into $j > i, l$ and $j \leq i, l$. The $j \leq i, l$ are finite, so poses no threat to regularity.

For $j > i, l$, $\frac{\sigma_j}{2^j A_j}$ and all its derivatives of order $\leq l$ with respect to the coordinates ${}^i\xi$ are bounded by $\frac{1}{2^j}$, so we have uniform converges. It follows that σ is smooth.

We still need to normalize σ to satisfy all properties of 10.17. For this, note that $\sigma = \sigma_-$ for some constant $\sigma_- < 0$ on $J^-(S_-)$ (since all the $\sigma_j = -1$) and similarly $\sigma = \sigma_+ > 0$ on $J^+(S_+)$. Let $\psi : \mathbb{R} \rightarrow [-1, 1]$ be such that, ψ is smooth, $\psi(t) = -1$ for $t \leq \sigma_-$, $\psi(t) = 1$ for $t \geq \sigma_+$ and $\psi' > 0$ for $t \in (\sigma_-, \sigma_+)$. Then $\psi \circ \sigma$ has all the desired properties. \square

Theorem 10.4. *There exists a smooth time function $\mathcal{T} : M \rightarrow \mathbb{R}$, whose gradient is timelike and past pointing and \mathcal{T} tends to $\pm +\infty$ along any future/past inextensible causal curve. In particular, the level sets of \mathcal{T} are smooth spacelike Cauchy hypersurfaces.*

Proof. Let $t_k = k/2$, $t_{k,\pm} = t_k \pm 1$, $t_{k,a} = t_k \pm 1/2$, $t_{k,b} = t_k \pm 1/2$. Let σ_k be the function just constructed, with $t = t_k$, $t_{\pm} = t_{k,\pm}$, $t_a = t_{k,a}$, $t_b = t_{k,b}$. Note that each

$p \in M$ is such that $\tau(p) \in (t_{k,a}, t_{k,b})$ for some k . Define

$$\mathcal{F} = \sigma_0 + \sum_{k=1}^{+\infty} (\sigma_{-k} + \sigma_k).$$

Note that if $k \geq 3$, then $\sigma_{-k}(p) + \sigma_k(p) = 0$ if $-k/2 + 1 \leq \tau(p) \leq k/2 - 1$ (because $\sigma_{-k}(p) = 1, \sigma_k(p) = -1$ on this interval). Thus, for any $p \in M$, there is a neighborhood of p such that only a finite number of the terms in the above sum are non-zero on that neighborhood. It follows from the definition and the fact that each p is such that $\tau(p) \in (t_{k,a}, t_{k,b})$ for some k that the gradient of \mathcal{F} is timelike and past directed. Let now $\gamma : (t_-, t_+) \rightarrow M$ be a future directed inextendible causal curve. From the gradient properties of \mathcal{F} , $\mathcal{F} \circ \gamma$ is strictly increasing along γ . Let $m \geq 1$. Since γ intersects each Cauchy hypersurfaces S_t , there exists $s_m \in (t_-, t_+)$ such that $\tau(\gamma(s_m)) = m$. Let $l = 2(m+1)$. We have $(\sigma_{-k} + \sigma_k)[\gamma(s_m)] = 0$ for $k \geq l$. Thus,

$$\mathcal{F}(\gamma(s_m)) = \left[\sigma_0 + \sum_{k=1}^l (\sigma_{-k} + \sigma_k) \right] (\gamma(s_m)).$$

Since $m \geq 1$, $\sigma_0[\gamma(s_m)] = 1$. Furthermore, $(\sigma_{-k} + \sigma_k)(\gamma(s_m)) \geq 0$ for all $k \geq 1$, since $\sigma_{-k}(\gamma(s_m)) = 1$ and $\sigma_k(\gamma(s_m)) \geq -1$. Finally, for $1 \leq k \leq 2(m-1)$, we have

$$(\sigma_{-k} + \sigma_k)(\gamma(s_m)) = 2$$

and thus

$$\mathcal{F}(\gamma(s_m)) \geq 4(m-1) + 1.$$

Combined with the monotonicity property of $\mathcal{F}(\gamma)$, we have obtained $\mathcal{F}(\gamma(s)) \rightarrow +\infty$, as $s \rightarrow t_+$. \square

The above theorem says in particular that given a globally hyperbolic spacetime, there exists a foliation by smooth spacelike Cauchy hypersurface. On the other hand, we know that a spacetime admitting a Cauchy hypersurface, and thus in particular, a smooth spacelike Cauchy hypersurface, is globally hyperbolic. The next theorem, whose proof is omitted, then states the existence of a foliation by smooth spacelike Cauchy hypersurface such that one of the leaf coincides with the original Cauchy hypersurface.

Theorem 10.5. *Let S be a smooth spacelike Cauchy hypersurface. Then, there exists a temporal function τ such that $\tau^{-1}(0) = S$ and τ goes to $\pm\infty$ along future/past directed inextendible causal curves.*

11 The domain of dependence in globally hyperbolic spacetime

The aim of this section is to prove the following statement.

Theorem 11.1 (Domain of dependence). *Let (M, g) be a globally hyperbolic spacetime and let t be a temporal function as in Theorem 10.4. Let $S = t^{-1}(0)$ and let $\Omega \subset S$ be open. Let u be a solution to $\square_g u = 0$ such that u and Du vanishes on Ω . Then $u = 0$ on $D^+(\Omega)$. In particular, taking $\Omega = S$, then we have proven uniqueness of the solution to the linear wave equation.*

Before proving the theorem, we will need a few more geometric statements, as well as a local version of it.

11.1 Three extra geometric lemmas

Lemma 11.1. *Let (M, g) be a smooth globally hyperbolic spacetime with a smooth spacelike Cauchy surface S . Then, for every $p \in S$, there is a neighborhood U_p of p such that for every $q \in U_p \cap J^+(S)$, there is a normal neighborhood V_q of q such that $J^-(q) \cap J^+(S)$ is compact and contained in V_q .*

We leave its proof as an exercise, but note that one cannot just take any convex neighborhood of p , for the condition $J^-(q) \cap J^+(S) \subset V_q$ is not always satisfied. (For a full proof, consult [Rin09], p135).

We will also need (again for proofs, see [Rin09, Chapter 12].)

Lemma 11.2 (The intersection of transverse manifolds is a manifold). *Let N_1 and N_2 be two n dimensional spacelike submanifolds. Assume for each $p \in N_1 \cap N_2$, the normals to N_1 and N_2 are linearly independent. Then, $N_1 \cap N_2$ is an $n-1$ dimensional spacelike submanifold.*

Lemma 11.3. *Let N_1 and N_2 be two n dimensional spacelike submanifolds of $(U \subset \mathbb{R}^{n+1}, g)$, for some Lorentzian metric g defined on U . Assume N_1 and N_2 are such that their normals are linearly independent on $N_1 \cap N_2$, so that from the previous lemma, $P = N_1 \cap N_2$ is a $n-1$ dimensional manifold. Assume moreover that P is compact. Fix n_1 and n_2 be the normals to N_1 and N_2 . Then, their restrictions to P give two vector fields on P such that at every $p \in P$, $n_1(p), n_2(p) \in T_p M^\perp$. Define*

$$f : P \otimes \mathbb{R}^2 \rightarrow \mathbb{R}^{n+1}$$

by

$$f(p, t, s) = p + tn_1(p) + sn_2(p).$$

Then f is smooth and there is an $\epsilon > 0$ such that f restricted to $P \times B_\epsilon(0)$ is a diffeomorphism onto a neighborhood of P .

11.2 A local version of the domain of dependence property

Lemma 11.4. *Let S be a Cauchy surface and $p \in J^+(S) \setminus S$. Assume that there is a normal neighborhood of p V_p such that $J^-(p) \cap J^+(S)$ is compact and contained in V_p . Then, if u is a smooth solution to the wave equation*

$$\square_g u = 0$$

and u and Du vanishes on $J^-(p) \cap S$, $u = 0$ on $J^-(p) \cap J^+(S)$.

Proof. Let $\tilde{V}_p = \exp_p^{-1}(V_p)$. We consider at p , the hyperquadric $\tilde{q}^{-1}(c) \in T_p M$, for $c < 0$. It has two connected components, and we consider the one corresponding to the past directed timelike vectors, denoted \tilde{Q}_c . Let Q_c be the image by the exponential map of $\tilde{Q}_c \cap \tilde{V}_p$. By construction $Q_c \subset I^-(p)$. Similarly, we denote by Q_0 the past directed local null cone at p . Let $D = J^-(p) \cap J^+(S)$ and $D_c = J^-(Q_c) \cap J^+(S)$. We leave as an exercise to verify that the following statement holds (Draw a picture !)

1. $D_c = \bigcup_{\gamma \leq c} Q_\gamma \cap J^+(S)$,
2. The interior of D_c is $I^-(Q_c) \cap I^+(S)$ and its boundary is $Q_c \cap J^+(S) \cup J^-(Q_c) \cap S$.
3. If $c_l \rightarrow 0$, $c_l < 0$ for all l , then $\text{int } D \subset \bigcup_l D_{c_l} \subset D$.

Let ρ be a Riemannian metric on V and let d be the associated topological metric. Let $\epsilon > 0$ and define

$$R_\epsilon = \{r \in S \cap D : d(r, Q_0) < \epsilon\}.$$

Then, R_ϵ is an open neighborhood of $Q_0 \cap S$ in $S \cap D$. Let $L_\epsilon = S \cap J^-(p) \setminus R_\epsilon$. Since D is compact and S is closed, we have $D \cap S$ compact and hence, L_ϵ is compact. Moreover, for every $r \in L_\epsilon$, $\exp_p^{-1}(r)$ is timelike (because r does not belong to some Q_0 but lies in $J^-(p)$). Consequently, $\exp_p^{-1}(L_\epsilon)$ is a compact subset of the interior of the past line cone in $T_p M$. Hence, for $c < 0$ small enough, L_ϵ does not intersect Q_c and Q_c and S must intersect in R_ϵ . Let T be a smooth unit normal to S . Then, T is timelike. Another vector field in R_ϵ is provided by the position vector field P , which we recall is normal to any of the Q_c . For $\epsilon > 0$ small enough, we claim that P and T are linearly independent in \bar{R}_ϵ . Since P and T are non-zero vector fields on \bar{R}_ϵ and \bar{R}_ϵ is compact, $\rho(T, T)$ and $\rho(P, P)$ are uniformly bounded above and below on \bar{R}_ϵ . On the other hand, $g(T, T) = -1$ and $g(P, P)$ tends to zero as $\epsilon \rightarrow 0$. This implies the linear independence.

Consequently, for $c < 0$, every point in $Q_c \cap S$ is such that the normal to Q_c and S at that point are linearly independent. Since Q_c and S are smooth spacelike n dimensional manifolds, $Q_c \cap S$ is a smooth $n-1$ dimensional submanifold. Moreover it is compact (exercise).

Let u be the solution assumed to exist in the statement of the lemma. Let T be its energy-momentum tensor, i.e.

$$T_{\alpha\beta} = D_\alpha u D_\beta u - \frac{1}{2} g_{\alpha\beta} (g(Du, Du)).$$

Let $f = -1/2g$ and $N = -P$, where P is the position vector field. Note that $\text{grad } f = N$.

Let $J := {}^N J_\beta = T_{\alpha\beta} N^\beta$ be the current associated to N and let

$$\eta = -e^{kf} |u|^2 N$$

and

$$\tilde{J} = e^{kf} J$$

be a weighted version of it. (The weights allows to absorb some error terms, otherwise, we would need an extra Gronwall argument to conclude). The constant k will be chosen later.

Note that in D_c , N is a future directed timelike vector field and that on $Q_c \cap I^+(S)$, it is the outward pointing normal relative to D_c . Recall that

$$D^\alpha T_{\alpha\beta} = \square_g u \cdot D_\beta u$$

and

$$\text{div } J = \square_g N(u) + T_{\alpha\beta} N^\alpha \pi_{\alpha\beta}.$$

Let $e = \frac{1}{2}|u|^2 + \sum_\alpha |\partial_\alpha u|^2$.

Then,

$$|\text{div } J| \leq Ce,$$

for some $C > 0$.

We also have

$$\text{div } \eta = -2e^{kf} u N(u) - e^{kf} |u|^2 \text{div } N - ke^{kf} |u|^2 g(N, N).$$

Since N is timelike on D_c , there exists a constant $c_0 > 0$ such that

$$\operatorname{div} \eta \geq e^{kf} (kc_0|u|^2 - Ce).$$

Moreover,

$$\operatorname{div} \tilde{J} = e^{kf} (\operatorname{div} J + kQ(N, N))$$

and again, since N is timelike on D_c (uniformly), there exists a constant c_1 such that

$$c_0|u|^2 + T(N, N) \geq c_1 e.$$

Hence,

$$\operatorname{div} \eta + \operatorname{div} J \geq e^{kf} (kc_1 - C)e.$$

For k large enough, the left hand side is therefore positive and controls $e^{kf} e$. Moreover,

$$\frac{g(N, \tilde{J})}{g(N, N)} = \frac{e^{kf} T(N, N)}{g(N, N)}, \quad \frac{g(N, \eta)}{g(N, N)} = e^{-kf} |u|^2.$$

Both of these quantities are non-positive. Recall that by Stokes theorem, if (M, g) is Lorentzian manifold with a smooth spacelike boundary ∂M and if N is the outward unit-normal, then for ξ is smooth vector field with compact support

$$\int_M \operatorname{div} \xi = \int_{\partial M} \frac{g(\xi, N)}{g(N, N)} \quad (27)$$

where both integrals are taken with respect to the natural (induced) volume forms on M and ∂M . If we could apply this directly in our setting we would be done, since both sides of the equation would have different signs.

The trouble is that our domain D_c has corners at the intersection of Q_c with S , so we are not just in the setting to apply the above.

We will use Lemma 11.3 above.

Using the lemma, we get a smooth map

$$h : Q_c \cap S \times B_\epsilon(0) \rightarrow V$$

which is a diffeomorphism onto its image and contains an open neighbourhood of $Q_c \cap S$, assuming that $\epsilon > 0$ is small enough. Let $\chi \in C_c^\infty(\mathbb{R}^2)$ be such that $\chi(x) = 1$ for $|x| \leq 1/2$ and $\chi(x) = 0$ for $|x| \geq 3/4$ and let $\chi(\delta/x) = \chi_\delta(x)$. For $\delta \leq \epsilon$, we can consider a function

$$\begin{aligned} \phi_\delta : Q_c \cap S \times B_\epsilon(0) &\rightarrow \mathbb{R}, \\ (p, x) &\rightarrow \psi(p, x) := \chi_\delta(x) \end{aligned}$$

Then, $\psi = \phi \circ h^{-1}$ is a smooth function defined on some neighborhood of $Q_c \cap S$ and we can extend it by 0 outside to get a smooth function on V . Moreover, the volume of the support of ψ_δ can be estimated by $C\delta^2$ for some constant C and

$$|\partial \psi_\delta| \leq C\delta^{-1}.$$

Now $D'_c = D_c \setminus S \cap Q_c$ is a smooth manifold with boundary and $(1 - \psi_\delta)X$ has compact support on this manifold. Applying the above divergence formula (27), we get

$$\int_{D'_c} \operatorname{div} ((1 - \psi_\delta)X) = - \int_{D'_c} X^\alpha \partial_\alpha \psi_\delta + \int_{D'_c} \operatorname{div} X - \int_{D'_c} \psi_\delta \operatorname{div} X.$$

The first and the last term converges to 0, so that the middle term converges to the divergence of X on D_c . The boundary integrals similarly converges to what they should. It follows that the above divergence formula (27) is valid even in the presence of "corners" (not that it was important that they were of co-dimension 2!) which concludes the proof. \square

11.3 Proof of the domain of dependence

We now move with the proof of Theorem 11.1.

Proof. Note first that $D^+(\Omega) \subset \overline{\text{int}(D^+(\Omega))}$. Let $p \in \text{int}(D^+(\Omega))$. We have $K = J^-(p) \cap D^+(\Omega)$ compact by Lemma 10.17. Recall that t is a smooth temporal function and that $S = t^{-1}(0)$. For any interval I and $t_0 \in \mathbb{R}$, we define the sets

$$R_I = t^{-1}(I) \cap K, \quad R_{t_0} = t^{-1}(t_0) \cap K, \quad S_{t_0} = t^{-1}(t_0).$$

Note that if I is compact then so is R_I and that R_{t_0} and S_{t_0} are compact. Let $I_\epsilon = [t_0 - \epsilon, t_0 + \epsilon]$.

Note that, for any t_0 such that R_{t_0} is not empty and any open set U containing R_{t_0} , by compactness of R_{t_0} , there exists $\epsilon > 0$ such that $R_{I_\epsilon} \subset U$.

Let now $0 \leq t_0 \leq T := t(p)$. For any $q \in R_{t_0}$, there is an open neighborhood U_q as in the Lemma 11.1, using S_{t_0} as a Cauchy surface. The U_q s form an open cover of R_{t_0} and by compactness of R_{t_0} , there is a finite subcovering U_{q_1}, \dots, U_{q_l} . Let U be their union and let ϵ be small enough so that $R_{I_\epsilon} \subset U$. Assume that u and Du vanishes on R_{t_0} . Let $q \in R_s$ for some $s \in [t_0, t_0 + \epsilon]$. There there is a normal neighborhood of q such that the conditions of Lemma 11.4 holds, which implies that u and Du vanishes on R_s . Thus, the set of $s \in [0, T)$ such that u and Du equal zero on $R_{[0,s]}$ is non-empty, open and closed in $[0, T)$, which proves the theorem. \square

12 Symmetric hyperbolic systems

12.1 Solving PDEs: representation formula vs abstract methods

There are many ways to solve linear wave equations. Here, we will go through a standard method which relies essentially on energy estimates and a duality argument. The advantage of this method is that it is an abstract method, that works in many different situations and that it sort of highlights the importance of a priori estimates. Another standard method is via the construction of *paramatrices* or representation of solutions. For instance, for the free wave equation in Minkowski space,

$$\square \phi := (-\partial_t^2 + \Delta) \phi = 0,$$

one can use the Fourier transform to get an explicit representation of ϕ given appropriate initial data. In a curved space, one cannot use the Fourier transform, but one can still obtain explicit various representation of the solutions³⁷ under certain conditions. See for instance [Fri75] in the case of spacetimes for which the exponential map is always a global diffeomorphism. The advantage is that from the representation formula one typically get more information than just existence. The drawback is that the more you ask about the solution, the less likely it is to apply in a general situation.

³⁷Often, the representation is only an approximate solution so that one typically have errors terms. One can then get rid of the error terms by iterating in a second step.

12.2 Symmetric hyperbolic systems on \mathbb{R}^{n+1}

In this section, we consider an equation of the form

$$L(u) := A^\mu \partial_\mu u + Bu = f, \quad (28)$$

$$u(0, \cdot) = u_0, \quad (29)$$

where u is an \mathbb{R}^N valued function on defined on \mathbb{R}^{n+1} (or a subset), the initial data u_0 is taken smooth for simplicity, f is smooth and decay fast at infinity in x (say $f(t, \cdot)$ is Schwarz for all t) and the A^μ are smooth $N \times N$ matrix valued functions defined on \mathbb{R}^{n+1} such that A^μ are symmetric for each μ and A^0 is *positive definite with a uniform lower bound*, i.e. $\exists c > 0$, such that

$$A^0(\xi, \xi) > c|\xi|^2, \quad \forall \xi \in \mathbb{R}^n.$$

The operator L is then said to be *symmetric hyperbolic*.

Recall that $H^k = H^k(\mathbb{R}^n)$ is the space of tempered distributions u on \mathbb{R}^n such that its Fourier transform \hat{u} is a measurable function (we identify the function and its distribution) and such that

$$\|u\|_{H^k} := \|(1 + |\xi|^2)^{k/2} \hat{u}\|_{L^2} < C.$$

H^k with the norm $\|\cdot\|_{H^k}$ is then a Hilbert space. Recall also that for $k \geq 0$, we can identify H^k with the space of L^2 functions which have k weak derivatives in L^2 .

Recall also that $(H^k)^*$ can be identified with H^{-k} in the standard way.

We will need the following lemma, whose proof is left as an exercise (hint: use duality).

Lemma 12.1. *If u is Schwarz and ϕ is C^∞ with uniform bounds on ϕ and its derivatives, then*

$$\|\phi u\|_{H^k} \lesssim \|u\|_{H^k},$$

even for negative k .

12.3 A priori estimates for symmetric hyperbolic systems

Lemma 12.2 (Estimates for $k \geq 0$). *Let u be a solution of (28) which is smooth on $[0, T] \times \mathbb{R}^n$ and such that, for $k \geq 0$, $u(t, \cdot) \in H^k$, for all $t \in [0, T]$. Let E_k be defined by*

$$E_k[u] = \frac{1}{2} \sum_{|\alpha| \leq k} \int_{\mathbb{R}^n} A^0(\partial_\alpha u, \partial_\alpha u) dx.$$

Then, we have the following estimate

$$E_k^{1/2}(t) \lesssim E_k^{1/2}(0) + \int_0^t \|f(s, \cdot)\|_{H^k} ds.$$

Proof. We do the proof for $k = 0$, the other follows easily after commuting the equation by ∂_α and noting that the resulting equation can be put under the same form.

We have

$$\begin{aligned} \frac{d}{dt} \frac{1}{2} \int_x A^0(u, u) dx &= \frac{1}{2} \int_x \partial_t A^0(u, u) dx + \int_x A^0(\partial_t u, u) dx, \\ &= \frac{1}{2} \int_x \partial_t A^0(u, u) dx - \int_x u^t \left(A^i \partial_i u + Bu - f \right) dx, \\ &= \frac{1}{2} \int_x \partial_t A^0(u, u) dx + \frac{1}{2} \int_x \left(u^t \partial_i (A^i) u - (Bu, u) - (f, u) \right) dx, \end{aligned}$$

so that the lemma follows from the lower bounds on A^0 and the upper bounds on A and f and B . \square

Since we have uniform lower bounds on A^0 and uniform H^k bounds on f by assumptions, we obtain immediately

Corollary 12.1. *Under the same assumptions, we have uniform $\|\cdot\|_{H^k}$ bounds on u for $k \geq 0$.*

Because of the duality arguments we will use later, we also need uniform estimates on u in H^k for $k < 0$.

Lemma 12.3 (Estimates for $k < 0$). *Let u be a solution of (28) which is smooth on $[0, T] \times \mathbb{R}^n$, satisfying uniform Schwarz bounds, for all $(t, x) \in [0, T] \times \mathbb{R}^n$,*

$$(1 + |x|^i) |\partial^\beta u|(t, x) \leq C_{i, \beta}.$$

Let $k < 0$ be a negative integer. Then, we have the estimate

$$\|u(t, \cdot)\|_{H^k} \lesssim \|u(0, \cdot)\|_{H^k} + \int_0^t \|f(s, \cdot)\|_{H^k} ds.$$

Proof. Let $U(t, \cdot)$ be defined by

$$U(t, \cdot) = (1 - \Delta)^k u(t, \cdot),$$

where $(1 - \Delta)^k$ for k negative is to be understood as the Fourier multiplier $(1 + |\xi|^2)^k$. It follows from the assumptions on u that both U and LU satisfy Schwarz bounds, so that we can apply the results of the previous corollary. Thus

$$\begin{aligned} \|u(t, \cdot)\|_{H^k} &= \|U(t, \cdot)\|_{H^{-k}} \lesssim \|U(0, \cdot)\|_{H^{-k}} + \int_0^t \|LU(s, \cdot)\|_{H^{-k}} ds \\ &\lesssim \|u(0, \cdot)\|_{H^k} + \int_0^t \|LU(s, \cdot)\|_{H^{-k}} ds. \end{aligned}$$

It remains to estimate LU in H^{-k} . For this, we observe that

$$f = Lu = (1 - \Delta)^{-k} LU + [L, (1 - \Delta)^{-k}] U$$

yielding

$$\|LU\|_{H^{-k}} \leq \|f\|_{H^k} + \|[L, (1 - \Delta)^{-k}] U\|_{H^k}.$$

Note that

$$[L, (1 - \Delta)^{-k}] = \sum_{|\alpha| \leq 2|k|} a_\alpha \partial^\alpha$$

where α are multi-indices such that ∂^α contains at most one t derivative and a_α are smooth bounded coefficients depending only on A and B .

As a consequence of Lemma 12.1,

$$\|[L, (1 - \Delta)^{-k}] U\|_{H^k} \lesssim \|U\|_{H^{-k}} + \|\partial_t U\|_{H^{-k-1}}.$$

The last term contains t derivative of U , which we do not yet control. We will substitute the t derivative using the equation. More specifically, let

$$L_0 u = (A^0)^{-1} Lu.$$

Then,

$$\begin{aligned} (A^0)^{-1} f &= L_0 u = \partial_t u + (A^0)^{-1} A^i \partial_i u \\ &= (1 - \Delta)^{-k} \partial_t U + (1 - \Delta)^{-k} (A^0)^{-1} A^i \partial_i U + [(A^0)^{-1} A^i \partial_i, (1 - \Delta)^{-k}] U, \end{aligned}$$

yielding

$$\begin{aligned} \|\partial_t U\|_{H^{-k-1}} &= \|(A^0)^{-1} f\|_{H^{k-1}} + \|(A^0)^{-1} A^i \partial_i U\|_{H^{-k-1}} + \left\| \left[(A^0)^{-1} A^i \partial_i, (1 - \Delta)^{-k} \right] U \right\|_{H^{k-1}} \\ &\lesssim \|f\|_{H^{k-1}} + \|U\|_{H^{-k}} + \|U\|_{H^{-k-1}} \\ &\lesssim \|f\|_{H^{k-1}} + \|U\|_{H^{-k}}. \end{aligned}$$

Combining the above, we obtain

$$\|u(t, \cdot)\|_{H^k} \lesssim \|u(0, \cdot)\|_{H^k} + \int_0^t (\|u(s, \cdot)\|_{H^k} + \|f(s, \cdot)\|_{H^k}) ds.$$

The lemma then follows using Gronwall inequality. \square

The previous analysis already brings, in particular, a global uniqueness statement: if u and v agree at $t = 0$, they agree for all times (just apply Lemma 12.2 with $k = 0$). However, one can prove a stronger result, a local uniqueness statement.

12.4 A rough local uniqueness statement

Lemma 12.4 (local uniqueness: Domain of dependence). *Let L be as above with $A^\mu \in C^1$, A^0 positive definite and bounded from below, $B \in C^0$ and $f = 0$ on some slab $[T_1, T_2] \times \mathbb{R}^n$. Let u be a solution to (28) such that $u(0, \cdot) = 0$ on some ball $B_{x_0}(R)$ with $R > 0$. Then, there exists a $c > 0$ depending only on the bounds on A such that $u = 0$ on $I^-(p) \cap [T_1, T_2] \times \mathbb{R}^n$, where $p = (R/c, x_0) \in \mathbb{R}^{n+1}$ and $I^-(p)$ is the interior of a past null cone for the Minkowski metric with the speed of light c .*

Proof. We multiply the equation by $e^{-kt} u$ for some k large enough to get

$$\partial_\alpha [e^{-kt} A^\alpha(u, u)] = e^{-kt} (-k A_0(u, u) + \partial_\alpha A^\alpha(u, u) - 2B|u|^2)$$

and integrate this over a domain of the form $\mathcal{D} = I^-(p) \cap [0, T_2] \times \mathbb{R}^n$, where $p = (R/c, x_0)$ and c will be chosen large enough so that $R/c < T_2$. Moreover, here $I^-(p)$ refers to the interior of the past null cone of p associated with the Minkowski metric

$$\eta_c = -c^2 dt \otimes dt + \delta_{ij} dx^i \otimes dx^j.$$

We then use the usual Stokes' theorem in \mathbb{R}^{n+1} to relate the bulk terms to the boundary terms. The integral over $\partial\mathcal{D} \cap \{t = 0\}$ vanishes in view of the initial data hypothesis. The other boundary is the truncated cone given by the equation

$$c \left(t - \frac{R}{c} \right) = |x - x_0|, \quad t > 0,$$

and the vector $n = (n_\alpha) = \left(\begin{array}{c} c \\ -\frac{(x-x_0)^i}{|x-x_0|} \end{array} \right)$ is an outgoing normal. (ex: rewrite this using differential geometric language, then n is actually a one-form...)

Thus this boundary term give a contribution proportional to $n_\alpha A^\alpha(u, u)$ where n_0 can be made arbitrary large by choosing c large enough. Choosing c large enough, it follows that this boundary term is non-negative. Choosing k large, the bulk term is non-positive, which implies that all terms must vanish, i.e. $u = 0$. \square

Remark 12.1. *If we apply the above statement to the case of wave equation, this form of the domain of dependence is weaker than the one we proved in the previous section, because we are proving a domain of dependence using an auxiliary Minkowski metric rather than the true geometry of our spacetime. Essentially, we are saying that the true lightcone can always fit into a larger auxiliary flat lightcone, or yet in another words, we overestimated the speed of light.*

12.5 Existence by duality

We are now in a position to prove existence of solutions.

Theorem 12.1 (Existence). *Let $S_T = [0, T] \times \mathbb{R}^n$ and assume that L is above, with A^μ and $B \in C^\infty$ with bounded derivative A^0 positive definite with a uniform lower bound. Assume for simplicity that $u_0 \in C_c^\infty$, i.e it has compact support. Then, there is a unique solution u on S_T to (28). Moreover, for each t , $u(t, \cdot) = 0$ outside from a fix compact K_T .*

Proof. Let L be as above and define L^* be the formal adjoint of L

$$L^* u = -\partial_t(A^0 u) - \partial_j(A^j u) + B^t u.$$

Then, L^* is a symmetric hyperbolic operator, so in particular, we have, for any k integer and for every test function $\phi \in C_c^\infty((-\infty, T) \times \mathbb{R}^n)$,

$$\|\phi(t, \cdot)\|_{H^{-k}} \leq C \int_t^T \|L^* \phi(s, \cdot)\|_{H^{-k}} ds, \quad (30)$$

for $t \in [0, T]$, by applying the estimates of Lemma 12.3 with t replaced by $T - t$.

Let X be the Banach space

$$X = L^1([0, T], H^{-k}).$$

Note that $L^* \phi$ can be viewed as a member of X so we let M be the subspace of X composed of elements of the form $L^* \phi$ for ϕ as above, i.e.

$$M = L^*(C_c^\infty((-\infty, T) \times \mathbb{R}^n)).$$

Note that for any element ψ of M , there is a unique ϕ , so that $\psi = L^* \phi$, in view of the above a priori estimate.

For any $f \in L^1([0, T], H^k)$, we consider a functional $F(= F_f)$ on M , defined as follows. For $\psi = L^* \phi \in M$,

$$F(\psi) = \langle \phi, f \rangle := \int_0^T (\phi(t), f(t))_{L^2} dt.$$

Then F is a linear functional on the subspace M and it follows from (30) that it is a bounded functional on M in the norm of X . By the Hahn-Banach theorem, it admits a continuous extension to the whole of X satisfying the same bound. By duality³⁸, there exists a $u \in L^\infty([0, T], H^k)$ such that, for any $\psi \in X$,

$$F(\psi) = \langle \psi, u \rangle.$$

³⁸Recall that $(L^1)^*$ is isomorphic to L^∞ , while $(L^\infty)^*$ is not isomorphic in general to L^1 .

In particular, for all $\psi = L^* \phi \in M$,

$$F(\psi) = \int_0^T (\phi(t), f(t))_{L^2} dt = \int_0^T (L^* \phi(t), u(t))_{L^2} dt.$$

Thus u is a solution *in the sense of distributions* of the equation (since we have assumed that $u(0, \cdot) = 0$). From the equation, the weak time derivative $\partial_t u$ solves

$$\partial_t u = -(A^0)^{-1} A^i \partial_i u - Bu - f \in H^{k-1}.$$

Thus, we have $\partial_t u, u \in L^\infty((-\infty, T); H^{k-1})$, which implies that³⁹ $u \in C^0((-\infty, T); H^{k-1})$. Repeating the argument, we can obtain that $u \in C^1((-\infty, T); H^{k-2})$. In particular, if f is smooth, this implies (using k arbitrary large) that u is smooth. To solve the equation with non-vanishing data, assume first that $u(0, x) \in C_c^\infty$ and let $u_0(t, x) = u(0, x)$ so that u_0 is a spacetime function with the right-Cauchy data. Then, consider the modified equation $Lv = f - Lu_0$. \square

Corollary 12.2. *Let $u_0 \in C^\infty$ and $f \in C^\infty$. Assume that the operator L is as above, there exists a unique solution to (28)*

Proof. Use the previous existence result and the local uniqueness together. \square

12.6 Applications to linear wave equations on \mathbb{R}^{n+1}

We now consider a linear wave equation in \mathbb{R}^{n+1} with a curved Lorentzian metric g of components $g^{\mu\nu}$, given by

$$g^{\mu\nu} \partial_\mu \partial_\nu u + b^\alpha \partial_\alpha u + cu = f, \quad (31)$$

$$u(t=0, \cdot) = u_0, \quad (32)$$

$$\partial_t u(t=0, \cdot) = u_1, \quad (33)$$

where b^α, c, u_0 and u_1 are C^∞ functions. We assume that g is uniformly Lorentzian, as in Definition 3.5.

Without loss of generality, assume that $g^{00} = -1$ (otherwise, just divide). Let $v = (\partial_\alpha u, u)$ be a \mathbb{R}^{n+2} vector field over \mathbb{R}^{n+1} and define the symmetric $n+2 \times n+2$ matrices by $(A^0)^{ij} = g^{ij}$, $(A^0)^{n+1, n+1} = (A^0)^{n+2, n+2} = 1$, $(A^k)^{i, n+1} = (A^k)^{n+1, i} = -g^{ik}$, $(A^k)^{n+1, n+1} = -2g^{0k}$ and all other components vanishes. Let moreover $d^{n+1, i} = -b^i$, $d^{n+1, n+1} = -b^0$, $d^{n+1, n+2} = -c$, $d^{n+2, n+1} = -1$, $h_{n+1} = f$. Then, (31) is equivalent to the equation on v

$$A^\alpha \partial_\alpha v + dv = h. \quad (34)$$

Exercise 12.1. *Show reciprocally that if v solves the above equation and the data verifies $\partial_i v_{n+2} = v_i$, then $u = v_{n+2}$ solves the original wave equation.*

From our previous analysis, we therefore obtain

Theorem 12.2. *There exists a unique C^∞ solution u to (31) satisfying the initial condition.*

As well as a domain of dependence property

Proposition 12.1. *With the above notation, if $f = 0$ and if the initial data vanishes on some $B_R(x_0)$, then, there exists a $c > 0$, such that $u = 0$ on $\Gamma^-(p) \cup \{t \geq 0\}$, where the set $\Gamma^-(p)$ is constructed from the Minkowski metric with speed of light c and $p = (R/c, x^0)$.*

³⁹Recall that we denote by u both a distribution and a function representing u . Thus $uisC^0$, means there exists a C^0 function whose distribution defined by it equal u .

13 Existence theory for the wave equation on globally hyperbolic spacetimes

In Section 11, we proved the domain of dependence property for solutions of the wave equation on globally hyperbolic spacetimes but we still have not proven existence of actual solutions. In the previous section, we prove an existence result which holds for uniformly Lorentzian metrics.

We will use the following lemma to reduce the general case to the case of uniformly Lorentzian metrics. For the proof, see [Rin09], p141.

Lemma 13.1. *Let (M, g) be a globally hyperbolic spacetime and t a smooth temporal function whose level sets S_t are Cauchy hypersurfaces. If $p \in S_t$, then there exists $\epsilon > 0$ and open neighborhoods U and W of p such that*

1. $\overline{W} \subset U$ and is compact.
2. if $q \in W$ and $\tau \in [t_0 - \epsilon, t_0 + \epsilon]$, then the compact set $J^+(S_\tau) \cap J^-(q)$ is contained in U .
3. There is a coordinate system on U , $\phi = (x^\alpha)$ such that $x^0 = t$ and the metric components $g_{\alpha\beta}$ in this coordinate systems are uniformly Lorentzian on U .
4. For any compact $K \subset U$, there exists a uniformly Lorentzian metric h globally defined on \mathbb{R}^{n+1} such that h coincides with $g \circ \phi^{-1}$ on $\phi(K)$.

We then have

Theorem 13.1. *Let (M, g) be a globally hyperbolic spacetime and t a temporal function as above. Then, given smooth functions ϕ_0, ϕ_1 defined on $S := S_{t_0}$, there exists a smooth solution ϕ of*

$$\begin{aligned} \square\phi &= F, \\ \phi_{S_{t_0}} &= \phi_0, \quad N(\phi)_{S_{t_0}} = \phi_1, \end{aligned}$$

where N is one of the unit normals of S_{t_0} .

Proof. Let us assume that N is the future unit normal and construct the solution to the future of S . First, we assume that ϕ_0, ϕ_1 are compactly supported in $K_1 \subset S$ and that F is compactly supported in $K_2 \subset M$. Let $t_1 > 0$ and define R_{t_1} to be the closed set of all qs such that $0 \leq t(q) \leq t_1$, i.e. $R_{t_1} = t^{-1}([0, t_1])$. Let $K_3 = K_2 \cap R_{t_1}$. Note that K_3 is compact. The union of all $I^+(p)$ for $p \in I^-(S)$ is an open covering⁴⁰ of $K_1 \cup K_3$, so it admits a finite subcovering consisting of $I^+(p_l)$, $1 \leq l \leq k < +\infty$, $p_l \in I^-(S)$. Note that $K_4 = \bigcup_{l=1}^k J^+(p_l) \cap J^-(S_{t_1})$ is compact⁴¹. By the domain of dependence theorem 11.1, any solution in R_{t_1} must vanish in $R_{t_1} \setminus K_4$. Thus, we only need to define our solution ϕ in the compact $K_4 \cap R_{t_1}$. Let $K_4(t) = K_4 \cap S_t$. Note that if $K_4(t) \subset U$ for U open then there exists an $\epsilon > 0$ such that $K_4(s) \subset U$ for $s \in [t - \epsilon, t + \epsilon]$.

Let $0 \leq \tau \leq t_1$ and assume that we either have a solution on R_τ or on R_s for all $0 \leq s < \tau$. For all $p \in K_4(\tau)$, there are neighborhoods W_p, U_p and an $\epsilon_p > 0$ as in Lemma 13.1. By compactness, there is a finite number of points p_i , $1 \leq i \leq l$ such

⁴⁰Note that since $K_1 \subset S$, the union of all $I^+(p)$ for $p \in S$ does not cover S and hence does not cover K_1 .

⁴¹It would be more intuitive to consider the Cauchy development of the set where both the data and the source vanishes.

that the W_{p_i} forms an open covering of $K_4(\tau)$. Let $0 < \epsilon \leq \min_{1 \leq i \leq l}(\epsilon_{p_i})$ be such that the W_{p_i} form an open cover of $K_4(s)$ for all $s \in [\tau - s, \tau + s]$ and let $s_1 \in [\tau - \epsilon, \tau]$ such that there is a solution up to and including s_1 .

Let $s \in [s_1, \tau + \epsilon]$ and $p \in K_4(s)$. Then $K_p = J^-(p) \cap J^+(S_{s_1})$ is compact and, by Property 2 of the lemma, contained in one of the charts, say (U_{p_i}, ϕ) . Consider the wave equation on U_{p_i} with initial data given on S_{s_1} . We can truncate the data so that it coincides with induced data on $K_p \cap S_{s_1}$ and vanishes outside from some open set included with U_{p_i} . Similarly, we can truncate all lower order coefficients and source terms in the wave equations, keeping only the principal symbol $g^{\alpha\beta} \partial_\alpha \partial_\beta$. By the third property of the lemma, we can moreover, extend g outside of K_p to a globally defined uniform Lorentzian metric on \mathbb{R}^{n+1} . We then obtain a linear wave equation as in the previous section to which there exists a smooth global solution. This gives us a solution on K_p . For any $q \in I^-(p) \cap J^+(S_{s_1}) = V_p$, we defined ϕ to be this solution.

For $r \in V_p \cap V_q$, we have two candidate definitions for ϕ at r . Since $J^-(r) \cap J^+(S_{t_1}) \subset V_p \cap V_q$, it follows from our domain of dependence property that the two possible definitions coincide, so that our solution is well defined.

Let now

$$O_1 = \bigcup_{p \in K_4(s), s \in [s_1, \tau + \epsilon]} V_p.$$

Note that the interior of O_1 contains $K_4(s)$ for any $s \in (s_1, \tau + \epsilon)$. Let O_2 be the set q such that $s_1 \leq t(q) < \tau + \epsilon$ and $q \notin K_4$. Note that if $q \in O_2$ and $t(q) > s_1$ then q is in the interior of O_2 and similarly for O_1 . In O_1 , we have already define ϕ and in O_2 , we would like to defined it to be 0. If $q \in O_1 \cap O_2$ and $t(q) > s_1$, then ϕ and its normal derivatives vanish at $J^-(q) \cap S_{s_1}$ and F vanishes in $J^-(q) \cap J^+(S_{s_1})$. Moreover, there is an $r > 0$ such that $q \in V_r \subset O_1$. By uniqueness, the solution defined on O_1 must therefore vanish at q .

Thus, given a solution on R_τ or on R_s for any $s < \tau$, we get a solution on $R_{\tau+\epsilon}$, for some $\epsilon > 0$. Thus, the set of $s \in [0, +\infty)$ such that there is a solution on R_s is closed, open and non-empty, i.e. equal $[0, +\infty)$.

Finally, we need to remove the compact support assumption. Let $p \in J^+(S)$. Then, $K_p = J^-(p) \cap J^+(S)$ is compact. Truncate the data and F to be zero outside some compact set and such that the induced data coincides with the original data on K_p . We then get a smooth solution ϕ' . Defined ϕ to be ϕ' on $I^-(p) \cap J^+(S) = V_p$. Again, if $r \in V_p \cap V_q$, then by uniqueness, the two possible solutions must coincide at r .

□

14 Local existence for the vacuum Einstein equations

We have seen in Section 6.3 that, using wave coordinates, the Einstein equations reduce to a system of quasilinear wave equations. Given proper initial data, this kind of system can be solved locally in time, cf [Sog95b, Chapter I.4].

What remains to be done to turn this into a proper existence result for the Einstein equations is to explained how to set-up the initial data and how to propagate the wave gauge.

First a definition

Definition 14.1. *An initial data set for the vacuum Einstein equations is a triple (Σ, h, k) where (Σ, h) is a Riemannian manifold and k is a symmetric $(0, 2)$ tensor so*

that the following constraint equations are satisfied

$$\begin{aligned}\frac{1}{2}S(h) - |k|^2 + (tr_h k)^2 &= 0, \\ \operatorname{div} k - d(tr_h k) &= 0,\end{aligned}$$

where $S(h)$ denotes the scalar curvature of (Σ, h) and $|k|^2 = k_{ij}k^{ij}$.

Comparing the above equations with (24)-(25), we see that we want to think of k as the 2nd fundamental form of Σ . However, at the moment, Σ is not a submanifold of M , in fact, there is no manifold M yet.

Let us now define the concept of a solution, or *developement* associated to some initial data.

Definition 14.2. *Let (Σ, h, k) be an initial data set. Then, a developement of (Σ, h, k) is solution (M, g) to the vacuum Einstein equations together with an embedding of Σ as a smooth hypersurface of M ,*

$$\psi : \Sigma \rightarrow M,$$

such that (h, k) coincide after pullback, with the first and second fundamental form of the embedding.

Let us do a little bit of functions counting. Assume Σ is of dimension n . Then, M is of dimension $n + 1$. We are seeking for a Lorentzian metric g , which we can think as $(n + 1)(n + 2)/2$ functions in view of the symmetry of g . If we are thinking of solving the Einstein equations using the reduced equations in wave coordinates, we need data for g and $\partial_t g$, which is thus $(n + 1)(n + 2)$ functions. These functions are not independent because of the $n + 1$ wave gauge conditions. So we have only $(n + 1)^2$ free functions. (h, k) provide $n(n + 1)$ functions, but we should also take in account the $n + 1$ constraint equations, so that we really have

$$n(n + 1) - (n + 1) = (n - 1)(n + 1)$$

independent functions. Thus, we are missing $2(n + 1)$ initial data information. These are fixed by the choice of wave coordinates: in order to construct a function satisfying the wave equation, one needs data for it and its derivatives and there are $(n + 1)$ coordinate functions.

Consider now an initial data set (Σ, h, k) . Let $(U, (x^i))$ be a local coordinate system of Σ . We will solve the Einstein equations on an open subset M of $I \times U$, where I is an open interval containing 0. Moreover, M will contain the submanifold $P = \{t = 0\} \times U$. Note that such an M can be given a system of global coordinates (t, x^i) .

The metric g will be obtained by solving the reduced Einstein equations. Since this is a system of quasilinear wave equations, we need initial data for g and $\partial_t g$. The initial data will be imposed on the hypersurface $P = \{t = 0\} \times U$. We define them in local coordinates as

$$g_{ij} = h_{ij}, \quad g_{00} = -1, \quad g_{0i} = 0.$$

Note that with this choice, if we assume that we have already constructed a solution, the vector field ∂_t is a unit normal to P . The second fundamental form is then given

by

$$\begin{aligned}
k_{ij} &= g(\partial_{x^i}, D_j \partial_t) = g_{i\alpha} \Gamma_{jt}^\alpha \\
&= g_{i\alpha} \frac{1}{2} g^{\alpha\beta} (g_{j\beta,t} + g_{t\beta,j} - g_{jt,\beta}) \\
&= h_{ik} \frac{1}{2} h^{kl} g_{jl,t} \\
&= \frac{1}{2} g_{ji,t}.
\end{aligned}$$

Thus, we define $g_{ij,t} := 2k_{ij}$. We are still missing the data for $g_{00,t}$ and $g_{0i,t}$. Again, we imagine we already have a solution in wave coordinates. Then, the wave coordinate condition $g^{\alpha\beta} \Gamma_{\alpha\beta}^0 = 0$ give us on P ,

$$\begin{aligned}
g^{\alpha\beta} \Gamma_{\alpha\beta}^0 &= -\Gamma_{00}^0 + h^{ij} \Gamma_{ij}^0 \\
&= -\frac{1}{2} g^{0\alpha} (g_{\alpha 0,t} + g_{0\alpha,t} - g_{00,\alpha}) + h^{ij} \frac{1}{2} g^{0\beta} (g_{j\beta,i} + g_{i\beta,j} - g_{ij,\beta}) \\
&= \frac{1}{2} g_{00,t} + \frac{1}{2} h^{ij} g_{ij,t}
\end{aligned}$$

where $g_{ij,t} = 2k_{ij}$ in view of the above. Thus, we define

$$g_{00,t} := -2tr_h k.$$

From the wave coordinate conditions $g^{\alpha\beta} \Gamma_{\alpha\beta}^l = 0$, we would have on P

$$\begin{aligned}
g^{\alpha\beta} \Gamma_{\alpha\beta}^l &= -\Gamma_{00}^l + h^{ij} \Gamma_{ij}^l \\
&= -\frac{1}{2} g^{l\alpha} (g_{\alpha 0,0} + g_{0\alpha,0} - g_{00,\alpha}) + h^{ij} \frac{1}{2} g^{l\beta} (g_{j\beta,i} + g_{i\beta,j} - g_{ij,\beta}) \\
&= -\frac{1}{2} g^{lm} (g_{m 0,0} + g_{0m,0} - g_{00,m}) + h^{ij} \frac{1}{2} g^{lm} (g_{jm,i} + g_{im,j} - g_{ij,m}) \\
&= -h^{lm} g_{m 0,0} + h^{ij} \frac{1}{2} h^{lm} (g_{jm,i} + g_{im,j} - g_{ij,m}).
\end{aligned}$$

Thus, we need

$$g_{m 0,0} := h^{ij} \frac{1}{2} (g_{jm,i} + g_{im,j} - g_{ij,m}).$$

Since we now have enough data, we can apply a local existence theorem for quasilinear wave to obtain the existence of a metric g , solution of the reduced equations on some set M of the required form. However, so far the metric g only satisfies the reduced Einstein equations and we need to show that the wave coordinate conditions is indeed satisfied by such a g . It then would follows that g actually solves the Einstein equations.

Recall from the section on the wave coordinate conditions that the wave conditions can be equivalently formulated as

$$\Gamma_\alpha := g^{\mu\nu} \Gamma_{\mu\alpha\nu} = 0.$$

Going back to the computation of the Ricci tensor in wave coordinates, one has that, in general,

$$Ric(g)_{\mu\nu} = -\frac{1}{2} g^{\alpha\beta} \partial_{\alpha\beta} g_{\mu\nu} + Q_{\mu\nu}(\partial g, \partial g) + \frac{1}{2} (D_\mu \Gamma_\nu + D_\nu \Gamma_\mu).$$

From this, it follows that the Einstein tensor verifies

$$\begin{aligned} G_{\mu\nu} &= -\frac{1}{2}g^{\alpha\beta}\partial_{\alpha\beta}g_{\mu\nu} + Q_{\mu\nu}(\partial g, \partial g) + \frac{1}{2}(\partial_{\mu}\Gamma_{\nu} + \partial_{\nu}\Gamma_{\mu}) - \frac{1}{2}(\partial^{\alpha}\Gamma_{\alpha})g_{\mu\nu} \\ &= \frac{1}{2}(\partial_{\mu}\Gamma_{\nu} + \partial_{\nu}\Gamma_{\mu}) - \frac{1}{2}(\partial^{\alpha}\Gamma_{\alpha})g_{\mu\nu}, \end{aligned}$$

since g solves the reduced Einstein equations.

Recall now that the Einstein tensor is always divergence free and thus that

$$D^{\mu}G_{\mu\nu} = 0.$$

In view of the above, we obtain a system of wave equations of the form

$$g^{\alpha\beta}\partial_{\alpha\beta}\Gamma_{\nu} = F(\partial\Gamma_{\nu}).$$

In order for Γ_{ν} to vanish, it follows that it suffices to show that initially $\Gamma_{\nu} = 0$ and $\partial_t\Gamma_{\nu} = 0$.

That $\Gamma_{\nu} = 0$ on P is certainly true since we have set up our initial data for this to hold. For the derivative conditions, since P is a submanifold of M and g a Lorentzian metric, the Gauss-Codazzi equations (24)-(25) must hold for the metric g on P i.e.

$$\begin{aligned} G(N, N) &= \frac{1}{2}{}^P S - k_{ij}k^{ij} + (tr_h k)^2 \\ G(N, \nu) &= {}^P D^j k_{ki} - {}^P D_i (tr_h k) \nu^i \end{aligned}$$

On the other hand, the RHS vanishes by virtue of the constraint equations (35)-(35). Thus, we have $G(N, N) = 0$ and $G(N, \nu) = 0$, where $N = \partial_t$. The first condition gives $\partial_t\Gamma_0 = 0$. The second, gives with $\nu = \partial_{x^i}$,

$$\begin{aligned} G_{0i} &= \frac{1}{2}(\partial_t\Gamma_i + \partial_i\Gamma_0) \\ &= \frac{1}{2}\partial_t\Gamma_i. \end{aligned}$$

Thus, we have proven that the Γ_{ν} all have trivial data and in view of the above discussion, it follows that g solves the vacuum Einstein equations in set M of the required form.

A Global volume form and orientable manifolds

Lemma A.1. *A semi-Riemannian manifold M has a global volume form if and only if M is orientable.*

Proof. Recall that a manifold is orientable if it has a coordinate atlas all of whose transition functions have positive Jacobian determinants. For every such local coordinate system, we have a local n -form defined by $|\det(g)|dx^1 \wedge \dots \wedge dx^n$. Moreover, for every other coordinate in the atlas (thus positively oriented with the first) one can check that the local n -form in the new coordinate system agrees on overlap with the first. Thus, this defines a global volume form. \square

B The pull-back bundle

Recall first the definition of a smooth fiber bundle. Let E, B, F be smooth manifolds and $\pi : E \rightarrow B$ be a smooth surjective map such that for all $x \in E$, there is an open neighbourhood $U \subset B$ of $\pi(x)$ together with a diffeomorphism $\phi : \pi^{-1}(U) \rightarrow U \times F$ satisfying $\pi = \pi_U \circ \phi$, where $\pi_U : U \times F \rightarrow U$ is the projection on the first factor.

Let $\pi : E \rightarrow M$ be a fiber bundle over M and consider a smooth map

$$\phi : N \rightarrow M.$$

We define the pull-back bundle as

$$\phi^*(E) = \{(q, e) \in N \times E \mid \phi(q) = \pi(e)\} \subset N \times E,$$

equipped with the subspace topology and the projection map

$$\begin{aligned} \pi' : \phi^*E &\rightarrow N \\ (q, e) &\rightarrow q. \end{aligned}$$

C Not so useful facts

We state here too standard results concerning the existence of a Lorentzian metric.

Theorem C.1. *A compact manifold admits a Lorentzian metric if and only if its Euler characteristic is 0.*

Theorem C.2. *Every non-compact manifold admits a Lorentzian metric.*

The proof of this theorem is not difficult using the existence of a Riemannian metric on M (which is true for all manifolds) together with the following statement.

Theorem C.3. *Every non-compact manifold admits a real-valued function f whose differential is nowhere vanishing.*

Note that in General Relativity, Lorentzian manifold are constructed by solving the Einstein equations (in particular, the underlying manifold is never compact!), so that the above statements are seldom (never ?) used.

D Taylor expansions of the metric tensor and volume form, following Alfred Gray [Gra73, Gra04]

We consider a semi-Riemannian manifold (M, g) . Given vector fields X, X_1, \dots, X_q we write

$$D_{X_1 X_2 \dots X_q}^q X := D_{X_1} D_{X_2} \dots D_{X_q} X.$$

Let (x^α) denotes normal coordinates at some point $p \in M$. We consider normal coordinate vector fields, i.e. vector fields whose components in the ∂_{x^α} basis are constant. From Exercise 3.17, the integral curve of any such vector field through p is a geodesic.

We will need the following lemmas.

Lemma D.1. *Let X, Y be a normal coordinate vector fields and ξ the integral curve of X through p . Then,*

$$\begin{aligned} (D_{X..X..X}^q X)_{\xi(t)} &= 0, \\ (D_X Y)_p &= 0. \end{aligned}$$

Proof. We have $(D_X X)_{\xi(t)} = 0$, since $\xi(t)$ is a geodesic and $\dot{\xi} = X \circ \xi$. Moreover, by definition of the induced covariant derivative, we have

$$0 = \frac{d}{ds} [(D_X X)_{\xi(t)}] = D_{\dot{\xi}} D_X X = D_{X X}^2 X.$$

By induction, we obtain the first statement of the lemma. Now, since $(D_X X)_p = 0$, we have in particular,

$$\begin{aligned} 0 = (D_{X+Y}(X+Y))_p &= (D_X Y)_p + (D_Y X)_p + (D_Y Y)_p + (D_X X)_p \\ &= (D_X Y)_p + (D_Y X)_p. \end{aligned}$$

On the other hand, since X, Y have constant components with respect to a coordinate induced basis of vector fields, they commute, and the second statement of the lemma follows. \square

Lemma D.2. *Let X, Y be normal coordinate vector fields. Then,*

$$\begin{aligned} (D_{YX..X}^q X)_p &= \dots = (D_{X..XYX}^p X)_p, \\ (2D_{X..X}^q Y + (q-1)D_{X..X..XYX}^q X)_p &= 0, \\ (D_{X..X}^q Y)_p &= \frac{q-1}{q+1} D_{X..X}^{q-2} R(X, Y)X. \end{aligned}$$

Proof. Since $[X, Y] = 0$, we have $R(X, Y) = D_X D_Y - D_Y D_X$. Let $A_k := (D_{X..Y..X}^q X)_p$, where the Y is in k^{th} place. We have, for $k \geq 2$

$$A_k - A_{k-1} = D_{X..X}^{k-2} R(X, Y) D_{X..X}^{q-k} X.$$

Note that since R is a tensor, we have, if $q > k$

$$\left(R(X, Y) D_{X..X}^{q-k} X \right)_{\xi(t)} = 0,$$

since $\left(D_{X..X}^{q-k} X \right)_{\xi(t)} = 0$. It then follows by induction that $D_{X..X}^{k-2} R(X, Y) D_{X..X}^{q-k} X = 0$, for any $q > k$, which is the first statement of the lemma.

For the second statement, we have by the first lemma, that for any $t \in \mathbb{R}$,

$$(D_{X+tY}^q (X+tY))_p = 0.$$

The left hand side of the previous equation defined polynomial in t and we compute its linear coefficient.

$$\begin{aligned} \left(D_{(X+tY)..(X+tY)}^q (X+tY) \right)_p &= D_{(X+tY)..(X+tY)}^q X + t D_{(X+tY)..(X+tY)}^q Y \\ &= D_{(X+tY)..(X+tY)}^{q-1} D_X X + t D_{(X+tY)..(X+tY)}^{q-1} D_Y X + t D_{X..X}^p Y + D_{Y..Y}^q Y + O(t^2) \\ &= t(q-1) D_{YX..X}^{q-1} D_X X + t D_{X..X}^{q-1} D_Y X + t D_{X..X}^p Y + O(t^2) \\ &= t(q-1) D_{YX..X}^{q-1} D_X X + 2t D_{X..X}^p Y + O(t^2), \end{aligned}$$

where all terms should be evaluated at p and we have used that $D_{Y..Y}^q Y = 0$, the first statement of the lemma, and the fact that $D_X Y = D_Y X$. This proves the second statement of the lemma.

For the last statement, we start from the second, which we write as

$$(2D_{X..X}^q Y)_p = -(q-1) \left(D_{XXX}^{q-2} D_Y D_X X \right)_p.$$

We add $(q-1) \left(D_{XXX}^{q-2} D_X D_Y X \right)_p$ to both side of the equations, so to introduce the Riemann curvature tensor on the right-hand side

$$(2D_{X..X}^q Y)_p + (q-1) \left(D_{XXX}^{q-2} D_X D_Y X \right)_p = (q-1) \left(D_{XXX}^{q-2} R(X, Y) X \right)_p.$$

The formula then follows using $D_X Y = D_Y X$. \square

Lemma D.3. *Let X_1, X_2 and Y be normal coordinate vector fields. Then, we have*

$$D_{X_1} D_{X_2} Y + D_{X_1} D_{X_2} Y = \frac{1}{3} (R(X_1, Y) X_2 + R(X_2, Y) Y).$$

Proof. Apply the third statement of the previous lemma to $X_1 + tX_2$ and Y with $q = 2$ and then compute the coefficient linear in t . \square

Consider now the local volume form η

$$\eta = \sqrt{|\det g|} dx^1 \wedge \dots \wedge dx^n.$$

We recall that $D\eta = 0$. We consider a taylor expansion near $p = (x^\alpha = 0)$ for $\sqrt{g} = \eta(\partial_{x^1}, \dots, \partial_{x^n})$.

$$\begin{aligned} \eta(\partial_{x^1}, \dots, \partial_{x^n})(x^\alpha) &= \eta(\partial_{x^1}, \dots, \partial_{x^n})(0) + \partial_{x^\alpha} (\eta(\partial_{x^1}, \dots, \partial_{x^n}))(0) x^\alpha \\ &\quad + \frac{1}{2} \partial_{x^\alpha} \partial_{x^\beta} (\eta(\partial_{x^1}, \dots, \partial_{x^n}))(0) x^\alpha x^\beta + O(|x|^3). \end{aligned}$$

We have $\eta(\partial_{x^1}, \dots, \partial_{x^n})(0) = 1$ while

$$\partial_{x^\alpha} (\eta(\partial_{x^1}, \dots, \partial_{x^n}))(0) = ((D_\alpha \eta)(\partial_{x^1}, \dots, \partial_{x^n}))(0) + \eta(\partial_{x^1}, \dots, D_\alpha \partial_{x^\beta}, \dots, \partial_{x^n})(0) = 0$$

since $D\eta = 0$ and $D_\alpha \partial_{x^\beta}(0) = 0$. For the second order term, we compute similarly

$$\begin{aligned} \partial_{x^\alpha} \partial_{x^\beta} (\eta(\partial_{x^1}, \dots, \partial_{x^n})) &= \partial_{x^\alpha} ((D_\beta \eta)(\partial_{x^1}, \dots, \partial_{x^n}) + \eta(\partial_{x^1}, \dots, D_\beta \partial_{x^\gamma}, \dots, \partial_{x^n})) = 0 \\ &= D_\alpha D_\beta \eta(\partial_{x^1}, \dots, \partial_{x^n}) + D_\beta \eta(\partial_{x^1}, \dots, D_\alpha \partial_{x^\gamma}, \dots, \partial_{x^n}) \\ &\quad D_\alpha \eta(\partial_{x^1}, \dots, D_\beta \partial_{x^\gamma}, \dots, \partial_{x^n}) + \eta(\partial_{x^1}, \dots, D_\alpha D_\beta \partial_{x^\gamma}, \dots, \partial_{x^n}). \end{aligned}$$

After evaluation at p , the only term remaining is $\eta(\partial_{x^1}, \dots, D_\alpha D_\beta \partial_{x^\gamma}, \dots, \partial_{x^n})(0)$. Thus,

$$\begin{aligned} \frac{1}{2} \partial_{x^\alpha} \partial_{x^\beta} (\eta(\partial_{x^1}, \dots, \partial_{x^n}))(0) x^\alpha x^\beta &= \frac{1}{2} \eta(\partial_{x^1}, \dots, D_\alpha D_\beta \partial_{x^\gamma}, \dots, \partial_{x^n})(0) x^\alpha x^\beta \\ &= \frac{1}{2} \eta \left(\partial_{x^1}, \dots, \frac{D_\alpha D_\beta + D_\beta D_\alpha}{2} \partial_{x^\gamma}, \dots, \partial_{x^n} \right) (0) x^\alpha x^\beta \\ &= \frac{1}{12} \eta(\partial_{x^1}, \dots, R(\partial_{x^\alpha}, \partial_{x^\gamma}) \partial_{x^\beta} + R(\partial_{x^\beta}, \partial_{x^\gamma}) \partial_{x^\alpha}, \dots, \partial_{x^n})(0) x^\alpha x^\beta \end{aligned} \quad (35)$$

Recall now that we have, by definition of the components of R

$$R(\partial_{x^\alpha}, \partial_{x^\gamma})\partial_{x^\beta} = R_{\beta\alpha\gamma}^\rho \partial_{x^\rho}.$$

Moreover, by antisymmetry of η , if the curvature terms on the RHS of (35) are in γ th position, only the term proportional to ∂_{x^γ} will contribute. Thus, we have

$$\begin{aligned} \frac{1}{2}\partial_{x^\alpha}\partial_{x^\beta}(\eta(\partial_{x^1}, \dots, \partial_{x^n}))(0)x^\alpha x^\beta &= \frac{1}{12}(R_{\beta\alpha\gamma}^\gamma + R_{\alpha\beta\gamma}^\gamma)(0)x^\alpha x^\beta \\ &= -\frac{1}{6}Ric(g)_{\alpha\beta}(0)x^\alpha x^\beta. \end{aligned}$$

We have thus proven the following theorem.

Theorem D.1. *With respect to a normal coordinate system (x^α) at p , the volume form admits the Taylor expansion*

$$\eta(x) = \left(1 - \frac{1}{6}Ric(g)_{\alpha\beta}(0)x^\alpha x^\beta + O(|x|^3)\right) dx^1 \wedge \dots \wedge dx^n.$$

E Additional exercises

E.1 Problems

E.1.1 A commutation formula for the wave equation

Let \square_g denote the D'Alembertian acting on real valued functions $\psi : M \rightarrow \mathbb{R}$.

1. Let X be a Killing field. Prove that $[X, \square_g] = 0$, in the sense that for any smooth function ψ ,

$$X(\square_g \psi) = \square_g(X\psi).$$

2. Let now X be a vector field, not necessarily Killing and denote by π its deformation tensor, and by $tr(\pi)$ the trace of π , given in any local coordinate system by $tr(\pi) = g^{\alpha\beta}\pi_{\alpha\beta}$. Prove that for any smooth function ψ ;

$$\square_g(X\psi) = X(\square_g(\psi)) + \mathfrak{q}[X\psi] \quad \text{with}$$

$$\mathfrak{q}[X\psi] = 2\pi^{\alpha\beta}D_\alpha D_\beta \psi + [2D^\alpha \pi_{\alpha\mu} - D_\mu(tr\pi)]D^\mu \psi.$$

E.2 Solutions

E.2.1 Solution to E.1.1 (second question only)

This is a classical formula. The proof written below is taken from [HS13]), Lemma 6.2, p242.

Note that on functions f

$$\mathcal{L}_X D_\alpha f = D_\alpha \mathcal{L}_X f = X^\beta D_\beta D_\alpha f + (D_\alpha X^\beta) D_\beta f,$$

while on 1-forms

$$\mathcal{L}_X D_\alpha V_\beta - D_\alpha \mathcal{L}_X V_\beta = X^\gamma [D_\gamma, D_\alpha] V_\beta - V^\gamma [D_\gamma, D_\alpha] X_\beta + (D_\gamma D_\alpha X_\beta - 2D_\alpha \pi_{\gamma\beta}) V^\gamma.$$

Contracting with $g^{\alpha\beta}$, we obtain the formula

$$g^{\alpha\beta} \mathcal{L}_X D_\alpha D_\beta V_\beta = g^{\alpha\beta} D_\alpha \mathcal{L}_X V_\beta + V^\gamma D_\gamma (tr \pi) - 2V^\gamma D^\alpha \pi_{\gamma\alpha}.$$

With $V_\beta = D_\beta \psi$ (i.e. $V = d\psi$), we get

$$g^{\alpha\beta} \mathcal{L}_X D_\alpha D_\beta \psi = g^{\alpha\beta} D_\alpha D_\beta \mathcal{L}_X \psi + D_\gamma (tr \pi) D^\gamma \psi - 2D^\alpha \pi_{\gamma\alpha} D^\gamma \psi.$$

Finally, the desired formula follows from

$$X(\square_g \psi) = \mathcal{L}_X (g^{\alpha\beta} D_\alpha D_\beta \psi) = -2\pi^{\alpha\beta} D_\alpha D_\beta \psi + g^{\alpha\beta} \mathcal{L}_X D_\alpha D_\beta \psi.$$

References

- [Ali09] Serge Alinhac. *Hyperbolic partial differential equations*. Universitext. Springer, Dordrecht, 2009.
- [BEE96] John K. Beem, Paul E. Ehrlich, and Kevin L. Easley. *Global Lorentzian geometry*, volume 202 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, second edition, 1996.
- [Bre97] Glen E. Bredon. *Topology and geometry*, volume 139 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997. Corrected third printing of the 1993 original.
- [BS03] Antonio N. Bernal and Miguel Sánchez. On smooth Cauchy hypersurfaces and Geroch's splitting theorem. *Comm. Math. Phys.*, 243(3):461–470, 2003.
- [BS05] Antonio N. Bernal and Miguel Sánchez. Smoothness of time functions and the metric splitting of globally hyperbolic spacetimes. *Comm. Math. Phys.*, 257(1):43–50, 2005.
- [Chr00] Demetrios Christodoulou. *The action principle and partial differential equations*, volume 146 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2000.
- [Chr08] Demetrios Christodoulou. *Mathematical problems of general relativity. I*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2008.
- [Fri75] F. G. Friedlander. *The wave equation on a curved space-time*. Cambridge University Press, Cambridge-New York-Melbourne, 1975. Cambridge Monographs on Mathematical Physics, No. 2.
- [Gra73] Alfred Gray. The volume of a small geodesic ball of a Riemannian manifold. *Michigan Math. J.*, 20:329–344 (1974), 1973.
- [Gra04] Alfred Gray. *Tubes*, volume 221 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, second edition, 2004. With a preface by Vicente Miquel.
- [HS13] Gustav Holzegel and Jacques Smulevici. Stability of Schwarzschild-AdS for the spherically symmetric Einstein-Klein-Gordon system. *Comm. Math. Phys.*, 317(1):205–251, 2013.

- [Ler55] Jean Leray. *Hyperbolic differential equations*. The Institute for Advanced Study, Princeton, N. J., 1953 1955.
- [O’N83] Barrett O’Neill. *Semi-Riemannian geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983. With applications to relativity.
- [Pen72] Roger Penrose. *Techniques of differential topology in relativity*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1972. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 7.
- [Rin09] Hans Ringström. *The Cauchy problem in general relativity*. ESI Lectures in Mathematics and Physics. European Mathematical Society (EMS), Zürich, Zuerich, Switzerland, 2009.
- [Sog95a] Christopher D. Sogge. *Lectures on nonlinear wave equations*. Monographs in Analysis, II. International Press, Boston, MA, 1995.
- [Sog95b] Christopher D. Sogge. *Lectures on nonlinear wave equations*. Monographs in Analysis, II. International Press, Boston, MA, 1995.